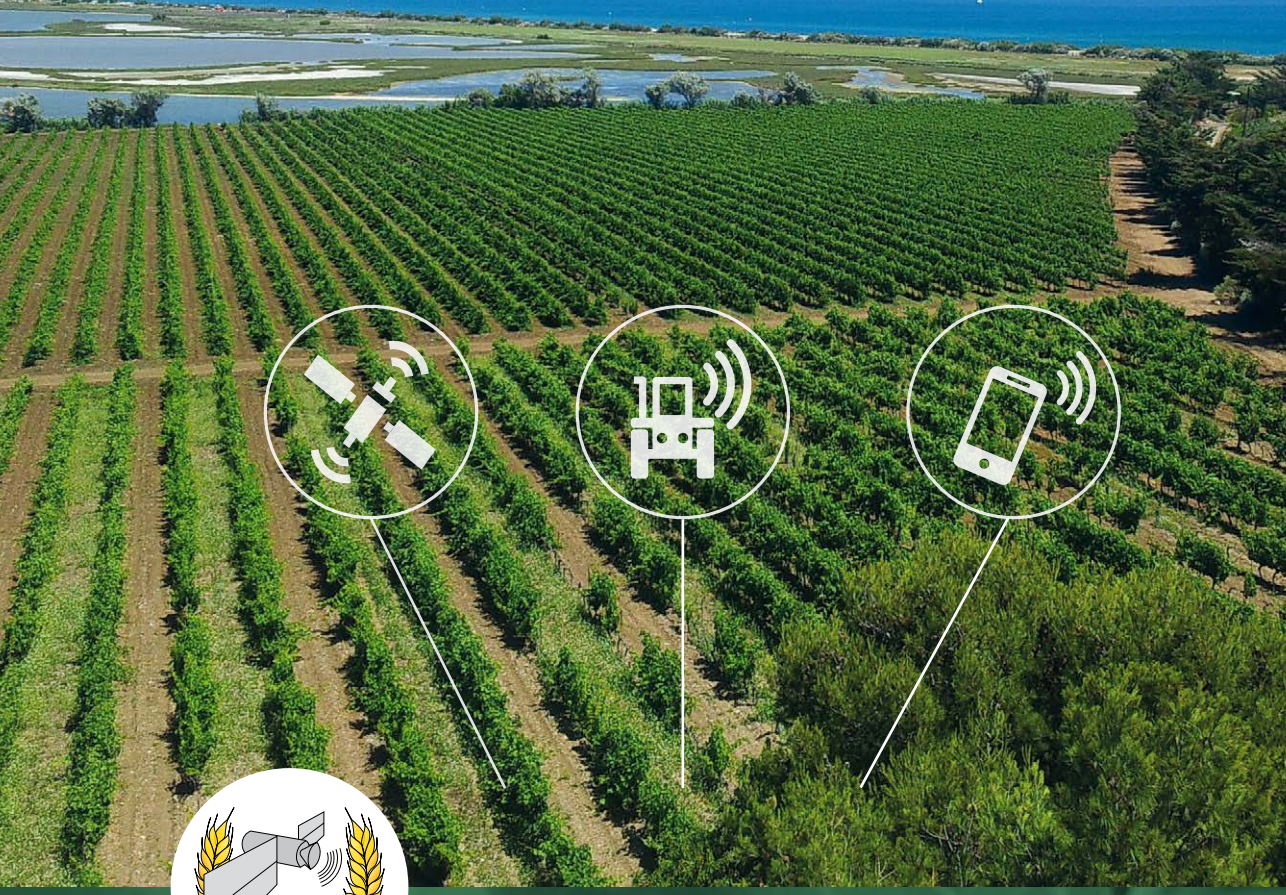


Precision agriculture '19

Wageningen Academic Publishers

edited by:
John V. Stafford

Precision agriculture '19

Precision agriculture '19

edited by:
John V. Stafford

Papers presented at the 12th European Conference on Precision Agriculture
Montpellier, France
8-11 July 2019



Wageningen Academic
P u b l i s h e r s

EAN: 9789086863372
e-EAN: 9789086868889
ISBN: 978-90-8686-337-2
e-ISBN: 978-90-8686-888-9
DOI: 10.3920/978-90-8686-888-9

Photo cover: View of the vineyards of Maguelone Island - Hérault, Languedoc © Photothèque Hérault Tourisme - C. Gauthier

First published, 2019
© Wageningen Academic Publishers The Netherlands, 2019

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned. Nothing from this publication may be translated, reproduced, stored in a computerised system or published in any form or in any manner, including electronic, mechanical, reprographic or photographic, without prior written permission from the publisher:
Wageningen Academic Publishers
P.O. Box 220
6700 AE Wageningen
The Netherlands
www.WageningenAcademic.com
copyright@WageningenAcademic.com
The individual contributions in this publication and any liabilities arising from them remain the responsibility of the authors.
The publisher is not responsible for possible damages, which could be a result of content derived from this publication.

Foreword

Prof. Bruno Tisseyre, Conference Chair
April 16, Montpellier SupAgro, Montpellier, France

Dear Reader,
In 2001, the city of Montpellier hosted the 3rd European Conference on Precision Agriculture (ECPA). Now, 18 years later, we are very pleased to welcome the Precision Agriculture scientific community back to Montpellier for 12ECPA.
We sincerely hope this 12th European Conference on Precision Agriculture will result in a profitable meeting for everyone and will provide some solutions to the challenges that modern agriculture is facing.
We are grateful to Montpellier SupAgro, Irstea and the International Society of Precision Agriculture (ISPA) for supporting us in the organisation of this event. We would especially like to thank and to congratulate John V. Stafford, the editor of these proceedings, for the enormous and accurate job he has done in producing this tome. In addition, we are also grateful to all the members of the conference Scientific Committee for their invaluable contribution in supporting John and assuring the scientific quality of the communications presented at this conference. There has been excellent interaction with Wageningen Academic Publishers in producing these proceedings which will be accessed for many years hence – they have been so helpful.
We appreciate the financial contribution of all the sponsors of 12ECPA, which includes private companies as well as public institutions and consortia. We would also like to express our gratitude to all the authors and attendees. The conference is nothing without your support and engagement. We will have more than 120 oral communications (compiled in this Conference Proceedings) and nearly 80 posters presented at the conference. This is a strong indication of the trust placed in this conference as a source of knowledge related to Precision Agriculture.
Finally, as the Chair, I am indebted to the support and hard work of the conference Organising Committee over the past 2 years in bidding for and delivering this conference. It is simply not possible to do this without a fantastic team behind the scenes. The Organising Committee is composed of academics and engineers from a joint team of Montpellier SupAgro (Institute for Higher Education in Agriculture) and Irstea (National Research Institute of Science and Technology for the Environment and Agriculture); The UMR ITAP team (joined research team in information-technologies-environmental analysis-agricultural processes). We are all both honoured and delighted to have worked for you and to have helped advance methods and techniques in Precision Agriculture.

Bon congrès à tous!

Bruno

Organizing committee

Chair

Prof Bruno Tisseyre (Montpellier SupAgro)

Vice Chair

Dr James Taylor (Irstea)

Members

Dr Guilhem Brunel (Montpellier SupAgro)
Eng. Thomas Crestey (Montpellier SupAgro)
Dr Nicolas Devaux (Montpellier SupAgro)
Dr Arnaud Ducanchez (Montpellier SupAgro)
Dr Serge Guillaume (Irstea)
Mrs Laure Haon (Montpellier SupAgro)
Dr Hazael Jones (Montpellier SupAgro)
Eng. Nina Lachia (Montpellier SupAgro)
Eng. Simon Moinard (Montpellier SupAgro)
Dr Olivier Naud (Irstea)
Eng. Pierre Péré (Irstea)
Eng. Léo Pichon (Montpellier SupAgro)
Dr Gilles Rabatel (Irstea)
Eng. Yoann Valloo (Montpellier SupAgro)
Prof Philippe Vismara (Montpellier SupAgro)

Scientific committee

The mission of the Scientific Committee has been to guarantee the academic quality of the proceedings of the 12th European Conference on Precision Agriculture by reviewing and assessing each of the 163 full papers submitted. The committee was chaired by Dr John V. Stafford, in his capacity as the editor of the 12th ECPA proceedings. The Scientific Committee of the 12th ECPA was made up of 135 internationally renowned researchers working collective across all facets of Precision Agriculture.

Poster and Oral communications have followed different parallel processes. In order to guarantee the academic quality of the proceedings, all abstracts submitted for consideration (295 total submissions) were assessed by the Editor to decide which were to be invited as full papers submissions. The full papers submitted for oral presentation then followed a peer review process carefully managed by the Editor. Each paper was assessed by two members of the Scientific Committee and, where recommended, revisions made before final acceptance by the Editor. The result is the Proceedings (124 papers) that you are currently reading. Poster abstracts were shorter (two pages) and assessed by a local committee chaired by Dr James Taylor. The accepted poster abstracts were edited and published as an electronic book 'Book of Poster abstracts', with a corresponding ISBN.

Copies of the proceedings of the 12th ECPA have been sent to Elsevier, Thomson Reuters and others considered of interest so that the peer-reviewed full papers can be considered for indexing in the Scopus and the Web of Knowledge academic databases. After the conference, authors of oral presentations will be invited to expand their manuscripts with additional data/analyses and submit them to a special ECPA '19 issue of the journal Precision Agriculture (https://link.springer.com/journal/11119).

Chair

Dr John V. Stafford

Members (alphabetical order)

Adamchuk, Viacheslav (Canada)
Agüera, Juan (Spain)
Alchanatis, Victor (Israel)
Amaral, Lucas (Brazil)
Andrade, Pedro (USA)
Antille, Diogenes (Australia)
Arno, Jaume (Spain)
Balsari, Paolo (Italy)
Bar-Hillel, Aharon (Israel)
Basso, Bruno (USA)
Behmann, Jan (Germany)
Berge, Therese (Norway)
Bishop, Tom (Australia)
Blasco, Jose (Spain)
Bongiovanni, Rodolfo (Argentina)
Bramley, Rob (Australia)
Bronson, Kevin (USA)
Brus, Dick (The Netherlands)
Cammarano, Davide (UK)
Carter, Paul (USA)
Casa, Raphael (Italy)
Castrignano, Anna Maria (Italy)
Castro, Ana de (Spain)
Chamen, Tim (UK)
Christensen, Svend (Denmark)
Clay, David (USA)
Clevers, Jan (The Netherlands)
Cohen, Yafit (Israel)
Colaco, Andre (Australia)
Cook, Simon (Australia)
Cugnasca, Carlos (Brazil)
Dille, Anita (USA)
Dillon, Carl (USA)
Ducati, Jorge (Brazil)
Escolà, Àlex (Spain)
Fitzgerald, Glen (Australia)

Fountas, Spyros (Greece)
Franzen, David (USA)
Fulton, John (USA)
Gaulton, Rachel (UK)
Geipel, Jakob (Norway)
Gemtos, Fanis (Greece)
Gerhards, Roland (Germany)
Germain, Christian (France)
Gil, Emilio (USA)
Gilbertsson, Mikael (Sweden)
Godwin, Richard (UK)
Granados, Francisca (Spain)
Griepentrog, Hans (Germany)
Griffin, Terry (USA)
Guillaume, Serge (France)
Hamilton, Alistair (UK)
Hedley, Carolyn (New Zealand)
Heuvelink, Gerard (The Netherlands)
Holland, Jonathan (UK)
Huang, Yanbo (USA)
Kerry, Ruth (USA)
Kersebaum, Christian (Germany)
Khosla, Raj (USA)
Khot, Lav (USA)
Kitchen, Newell (USA)
Koundouras, Stefanos (Greece)
Kross, Angela (Canada)
Lebeau, Frederic (Belgium)
Lee, Matthew (USA)
Leroux, Corentin (France)
Linker, Raphael (Israel)
Long, Dan (USA)
Longchamps, Louis (Canada)
Lowenberg-Deboer, Jess (USA)
Mahlein, Anne-Katrin (Germany)
Martínez-Casasnovas, José (Spain)
Matese, Alessandro (Italy)
Melakeberhan, Haddish (USA)
Meng, Jihua (China)
Meron, Moshe (Israel)
Miao, Yuxin (China)
Miller, Paul (UK)
Molin, Jose (Brazil)
Moltó, Enrique (Spain)
Moral, Francisco (Spain)
Morari, Francesco (Italy)
Moshou, Dimitrios (Greece)
Mulla, David (USA)
Nordmeyer, Henning (Germany)
Oberti, Roberto (Italy)
Oliver, Margaret (UK)
O'Shaughnessy, Susan (USA)
Pajares, Gonzalo (Spain)
Panneton, Bernard (Canada)
Paoli, Jean-Noel (France)
Pedersen, Soren (Denmark)
Pena-Yewtukhiw, Eugenia (USA)
Perez-Ruiz, Manuel (Spain)
Perry, Eileen (Australia)
Piikki, Kristin (Sweden)
Plant, Richard (USA)
Rabatel, Giles (France)
Rasmussen, Jesper (Denmark)
Ribeiro, Ángela (Spain)
Rodrigues, Francelino (Mexico)
Rosell-Polo, Joan R. (Spain)
Roudier, Pierre (New Zealand)
Rovira-Más, Francisco (Spain)
Rueda-Ayala, Victor (Norway)
Sadler, John (USA)
Santesteban, Gonzaga (Spain)
Sassenrath, Gretchen (USA)
Schepers, Jim (USA)
Schueller, John (USA)
Söderström, Mats (Sweden)
Solanelles, Francesc (Spain)
Stamatiadis, Stamatis (Greece)
Stombaugh, Tim (USA)
Sudduth, Ken (USA)
Tabb, Amy (USA)
Tardaguila, Javier (Spain)
Taylor, James (France)
Thomasson, Alex (USA)
Torres-Sanchez, Jorge (Spain)
Van Bergeijk, Jacob (USA)
Van Henten, Eldert (The Netherlands)
Vellidis, George (USA)
Vidović, Ivan (Croatia)
Vincini, Massimo (Italy)
Vougioukas, Stavros (USA)
Waine, Toby (UK)
Walsh, Kerry (USA)
Weber, Christian (Argentina)
Wendroth, Ole (USA)
Whelan, Brett (Australia)
Yang, Chenhai (USA)
Yule, Ian (New Zealand)
Zarco-Tejada, Pablo (Spain)
Zude, Manuela (Germany)

Table of contents

Foreword 5
Prof. Bruno Tisseyre, Conference Chair

Section 1 – Precision agriculture 21

Optimizing durum wheat cultivation in Northern Italy: assessing proximal and remote sensing derived from different platforms for variable-rate application of nitrogen 23
M.A. Bruce, J. Moretto, R. Polese and F. Morari

Integrating geospatial tools and a crop simulation model to understand spatial and temporal variability of cereals in Scotland 29
D. Cammarano, J. Holland, B. Basso, F. Fontana, T. Murgia, C. Lange, J. Taylor and D. Ronga

Comparing satellite and high-resolution visible and thermal aerial imaging of field crops for precision irrigation management and plant biomass forecast 37
A. Chen, V. Orlov-Levin, O. Elharar and M. Meron

Identifying yield stability and drivers of yield variability in cotton using multi-layered, whole-farm datasets 45
P. Filippi, T.F.A. Bishop and B.M. Whelan

Precision spraying by combining a variable rate application map with an on/off map 53
R. Hørfarter, M.D. Thorsted, K. Stougård and H.V. Poulsen

Understanding intra-field variation in N requirement for oilseed rape 61
S.L. Kendall, K. Storer, R. Wade and P.M. Berry

Evaluation of a functional Bayesian method to analyse time series data in precision viticulture 67
C. Laurent, M. Baragatti, J. Taylor, T. Scholasch, A. Metay and B. Tisseyre

Production gap analysis – an operational approach to yield gap analysis using historical high-resolution yield data sets 75
C. Leroux, J. Taylor and B. Tisseyre

Agriculture and digital sustainability: a Digitization Footprint 83
F. Marinello, R.G.V. Bramley, Y. Cohen, S. Fountas, H. Guo, M. Karkee, J.A. Martínez-Casasnovas, D.S. Paraforos, L. Sartori, C.G. Sørensen, B. Stenberg, K. Sudduth, B. Tisseyre, G. Vellidis and S.G. Vougioukas

Using the WOFOST crop growth model to assess within-field yield variability 91
A.C. Tagarakis, G. Mimić, H.M. Vaessen, F. Rodriguez-Moreno, F.K. van Evert and V. Ćirić

Empowering farmers by resolving the trust and legal issues emerging from precision farming 99
L. Wiseman and J. Sanderson

Section 2 – Precision horticulture **107**

3D point clouds from UAV imagery for precise plant protection in fruit orchards 109
M. Hobart, M. Schirrmann and M. Pflanz

Calculating the water deficit spatially using LiDAR laser scanner in an apple orchard 115
N. Tsoulias, D.S. Paraforos, S. Fountas and M. Zude-Sasse

Section 3 – Precision viticulture **123**

Comparison of water potential and yield parameters under uniform and variable rate drip irrigation in a cabernet sauvignon vineyard 125
I. Bahat, Y. Netzer, A. Ben-Gal, J.M. Grünzweig, A. Peeters and Y. Cohen

Use of an integrated model of water consumption as a decision support system for scheduling regulated deficit irrigation in a vineyard 133
J. Bellvert, M. Mata, X. Vallverdú, C. Paris and J. Marsal

Deep learning for in-field image-based grapevine downy mildew identification 141
J. Boulent, M. Beaulieu, P.-L. St-Charles, J. Théau and S. Foucher

Monitoring site-specific spraying in vineyards from a prescription map obtained with a UAV 149
J. Campos, M. Gallart, J. Llop and E. Gil

Investigation on LiDAR-based indicators for predicting agrochemical deposition within a vine field 157
A. Cheraiet, M. Carra, A. Lienard, S. Codis, A. Vergès, X. Delpuech and O. Naud

Missing plant detection and biomass estimation from 3D models generated from UAV in a vineyard 165
A. Matese, P. Cinat, Y. Romboli, A. Berton and S.F. Di Gennaro

Combining target sampling with route-optimization to optimise yield estimation in viticulture 173
B. Oger, P. Vismara and B. Tisseyre

Assessment of grapevine yield and quality using a canopy spectral index in white grape variety 181
M. Sozzi, A. Kayad, D. Tomasi, L. Lovat, F. Marinello and L. Sartori

Assessment of vineyard trimming and leaf removal using UAV photogrammetry 187
J. Torres-Sánchez, D. Marín, A.I. De Castro, I. Oria, F.M. Jiménez-Brenes, C. Miranda, L.G. Santesteban and F. López-Granados

Section 4 – Precision crop protection **193**

On-board colour imaging for the detection of downy mildew 195
F. Abdelghafour, F. Rançon, B. Keresztes, C. Germain and J.-P. Da Costa

Image-based assessment of hyperspectral reflectance characteristics of plants in the field: Lessons Learned 203
J. Behmann, D. Bohnenkamp and A.-K. Mahlein

Implications of dose-response relationships of herbicide droplet applications for leaf-specific weed control in leeks 209
N. Koukiasas, J. Martinez-Perez, R.A. Pilgrim, S. Sanford and A.J. Murdoch

Exploring the factors affecting spatio-temporal variation in grapevine powdery mildew 217
R. Melyon-Delage, B. Bois, S. Zito, M. Rega, G. Garin and A. Caffarra

Identifying the *Fusarium* spp. infestation in winter wheat based on RGB imaginary 225
G.G. Peteinatos, M. Sökefeld, J. Machleb, K. Cambel and R. Gerhards

Marrying futuristic weed mapping with current herbicide sprayer capacities 231
G.J. Somerville, R.N. Jørgensen, O.M. Bojer, P. Rydahl, M. Dyrmann, P. Andersen, N.-P. Jensen and O. Green

Use of simulations to study herbicide site-specific spraying 239
S. Villette, T. Maillot, J.P. Douzals, G. Jones, J.N. Paoli and J.P. Guillemin

Section 5 – Proximal and remote sensing of soil and crop 247

Design of a portable sensor suite for real-time monitoring of crop water stress index in maize breeding plots 249
O.E. Apolo-Apolo, M. Pérez-Ruiz, J. Martínez-Guanter and G. Egea

Characterisation of fungal diseases on winter wheat crop using proximal and remote multispectral imaging 255
R. Bebronne, A. Michez, V. Leemans, P. Vermeulen, B. Dumont and B. Mercatoris

Assessing infield temporal and spatial variability of leaf water potential using satellite imagery and meteorological data 263
O. Beeri, R. Pelta, T. Shilo, J. Raz and S. Mey-tal

Characterisation of cereal morphology through proximal stereo vision under contrasting nitrogen inputs 271
S. Dandrifosse, A. Bouvry, V. Leemans, B. Dumont and B. Mercatoris

Early detection of corn and sunflower stress induced by chemical spraying 279
S. Gad, Y. Edan, T. Sandovsky, I. Harary, T. Nacson, E. Kosover, A. Levi Bar Shalom and V. Alchanatis

Hyperspectral imaging application under field conditions: assessment of the spatio-temporal variability of grape composition within a vineyard 287
S. Gutierrez, M.P. Diago, J. Fernandez-Novales and J. Tardaguila

Potential of optical sensors for nitrogen management in spring barley 293
R. Hackett

Disentangling the sources of chlorophyll-content variability in banana fields 299
J. Lamour, C. Leroux, G. Le Moguédec, O. Naud, M. Léchaudel and B. Tisseyre

Sentinel-2 vegetation indices and apparent electrical conductivity to predict barley (*Hordeum vulgare* L.) yield 307
J.A. Martínez-Casasnovas, A. Uribeetxebarría, A. Escolà and J. Arnó

Accessing the plant architecture in 3D for plant phenotyping – recent approaches and requirements 315
S. Paulus

Vegetation indices from remote sensing imagery as proxies for yield and grain N in wheat 323
M. Quemada, J.L. Pancorbo, M. Alonso-Ayuso, J.L. Gabriel, J. López-Herrera and E. Pérez-Martín

Using terrestrial photogrammetry for leaf area estimation in maize under different plant growth stages 331
D. Reiser, A. Kamman, M. Vázquez Arellano and H.W. Griepentrog

Early detection of *Fusarium* infection in corn using spectral analysis 339
T. Sandovsky, Y. Edan, S. Gad, A. Etzioni, T. Nacson and V. Alchanatis

Proximal versus remote sensing to monitor pasture quality in a Mediterranean Montado ecosystem 347
J. Serrano, S. Shahidian, F. Moral and J. Marques da Silva

Section 6 – Applications of unmanned aerial systems 355

Investigation of spraying efficiency of an aerial spraying system in a super-high density olive grove in Greece 357
F. Aru, A. Gertsis, G. Vellidis and F. Morari

UAV-based hyperspectral imaging for weed discrimination in maize 365
R. Casa, S. Pascucci, S. Pignatti, A. Palombo, U. Nanni, A. Harfouche, L. Laura, M. Di Rocco and P. Fantozzi

A precision viticulture UAV-based approach for early yield prediction in vineyard 373
S.F. Di Gennaro, P. Toscano, P. Cinat, A. Berton and A. Matese

Estimating melon yield for breeding processes by machine-vision processing of UAV images 381
A. Kalantar, A. Dashuta, Y. Edan, A. Dafna, A. Gur and I. Klapp

Estimation of the leaf area index in maize based on UAV imagery using deep learning techniques 389
J. Martínez-Guanter, G. Egea, M. Pérez-Ruiz and O.E. Apolo-Apolo

A UAV-based system for monitoring crop growth in wheat, barley and triticale phenotyping field trials 397
J.M. Peña, F.J. Ostos-Garrido, J. Torres-Sánchez, F. Pistón and A.I. de Castro

The application of structure from motion techniques in late-season corn damage 405
J.M. Prince Czarnecki, S. Samiappan and L.A. Hathcock

A multispectral processing chain for chlorophyll content assessment in banana fields by UAV imagery 413
G. Rabatel, J. Lamour, D. Moura and O. Naud

Developing crop canopy model for irrigation of high-density olive groves by using UAV imagery 421
E. Zancanaro, A. Gertsis, G. Vellidis, F. Marinello and F. Morari

Section 7 – Satellite-based applications for precision agriculture 429

Assessing spatial and temporal variability in evapotranspiration for olive orchards in Tunisia using satellite remote sensing 431
A. Bchir, D.J. Mulla, A. Ben Dhiab, F. Ben Meriem, W. Bousetta and M. Braham

Accuracy of crop coefficient estimation methods based on satellite imagery 437
O. Beeri, R. Pelta, T. Shilo, S. Mey-tal and J. Tanny

Can time series of multispectral satellite images be used to estimate stem water potential in vineyards? 445
Y. Cohen, P. Gogumalla, I. Bahat, Y. Netzer, A. Ben-Gal, I. Lenski, Y. Michael and D. Helman

Interaction between soil variability and maize nitrogen status assessment from Sentinel-2 453
A. Crema, G. Vandini, M. Boschetti, F. Nutini, D. Cillis and R. Casa

Adopting precision agriculture to improve the cultivation of old wheat varieties in Tuscany (Italy) 461
C. Fabbri, M. Napoli, M. Mancini, G. Brandani, R. Vivoli and S. Orlandini

A cotton yield estimation model based on agrometeorological and high resolution remote sensing data 469
A. Falagas and K. Karantzalos

Can satellite-derived vigour maps be used to delineate homogeneous zones in hedgerow olive orchards? 477
F.J. Moral, F.J. Rebollo, S. Millán, H. Prieto, J.M. Pérez and C. Campillo

Estimating cotton water requirements using Sentinel-2: model development and validation 485
O. Rozenstein, N. Haymann, G. Kaplan, and J. Tanny

Satellite-based detection of sowing dates in field crops 493
Y. Sadeh, X. Zhu, K. Chenu and D. Dunkerley

Cost-effectiveness and performance of optical satellites constellation for Precision Agriculture 501
M. Sozzi, A. Kayad, D. Giora, L. Sartori and F. Marinello

Section 8 – Site-specific nutrient, lime and seed management **509**

Crop imaging and soil adjusted variable rate nitrogen application in winter wheat 511
F. Argento, T. Anken, F. Liebisch and A. Walter

Using yield monitor data to guide precision nitrogen application 519
B.W. Brorsen, E. Park and X. Li

Integrated approach for site-specific nitrogen management in North Dakota, USA 525
D. Franzen

Dividing the risk – theoretical exploration of increasing N management temporal granularity in maize 531
A.G. Hunt, J.M. Sharp, P.R. Johnstone and B.P. Searle

Creating soil texture maps for precision liming using electrical resistivity and gamma ray mapping 539
S. Meyer, C. Kling, S. Vogel, I. Schröter, A. Nagel, E. Kramer, R. Gebbers, G. Philipp, K. Lück, F. Gerlach, D. Scheibe and J. Ruehlmann

Effect of variable rate phosphorus and nitrogen fertilizing on winter wheat *(Triticum aestivum* L.*)* in Mezőföld, Hungary 547
G. Milics, A. Vér, L. Szekeres and J. Kauser

Precision nitrogen and water management for maize production in the western great plains of the US 555
E. Phillippi, R. Khosla, A. Andales and L. Longchamps

Statistical model to overcome rice variety effect in precision fertilisation 563
D. Sacco, E. Cordero, B. Moretti, E.F. Miniotti, D. Tenni, G. Beltarre, C. Grignani and M. Romani

Assessing sampling strategies and soil sensors performance in the detection of field scale variability of plant-available nitrogen 571
E. Wallor, M. Bourouah, K.C. Kersebaum and R. Gebbers

Section 9 – Site-specific canopy and harvest quality **579**

Satellite-based modelling of protein content in winter wheat and malting barley 581
T. Börjesson, S. Wolters and M. Söderström

Yield sensing technologies for horticultural crops: a short review 589
L. Longchamps, B. Panneton, S. Fountas and R. Khosla

Preliminary study for weed biomass prediction combining visible images with a plant-growth model 597
J. Merienne, A. Larmure and C. Gée

Using remote sensing to map in-field variability of peanut maturity 605
A.F. Santos, L.N. Lacerda, S. Gobbo, A. Tofannin, R.P. Silva and G. Vellidis

Section 10 – Precision pasture management **613**

Delineation of management zones in an agrosylvopastoral ecosystem based on the Rasch model 615
F.J. Moral, F.J. Rebollo and J.M. Serrano

Accurate segmentation of grass-sward structures in RGB images using machine learning algorithms 623
C. Rueda-Ayala, V. Rueda-Ayala, M. Höglind and D. Andujar

Long-term evaluation of the Grassmaster II probe to estimate productivity of pastures 631
J. Serrano, S. Shahidian, F. Moral and J. Marques da Silva

Species distribution mapping of grass clover leys using images for targeted nitrogen fertilization 639
S. Skovsen, M.S. Laursen, R. Gislum, J. Eriksen, M. Dyrmann, A.K. Mortensen, S. Farkhani, H. Karstoft, N.P. Jensen and R.N. Jørgensen

Section 11 – Drainage optimization and variable rate irrigation **647**

Evaluating and improving soil sensor-based variable irrigation scheduling on farmers' fields in Alabama 649
L. Bondesan, B.V. Ortiz, G.T. Morata, D. Damianidis, A.F. Jimenez, G. Vellidis and F. Morari

Artificial neural networks for irrigation management: a case study from southern Alabama, USA 657
A. Jimenez, B.V. Ortiz, L. Bondesan, G. Morata and D. Damianidis

Using computational optics for agricultural monitoring with an emphasis on irrigation management zones 665
I. Klapp, O. Brand, P. Yafin, S. Papini, N. Oz, I. Bahat, Y. Cohen, V. Alchanatis and N. Sochen

A variable rate irrigation decision support system for corn in the US eastern coastal plain 673
K. Stone, P. Bauer, S. O'Shaughnessy, M. Andrade and S. Evett

Use of soil electrical conductivity mapping in variable rate irrigation 681
R. Sui

Modeling spatio-temporal variations in crop water stress for variable-rate irrigation 687
J.D. Svedin, N.C. Hansen, R. Kerry and B.G. Hopkins

Comparison of precision and conventional irrigation management of cotton 695
E. Vories, S. O'Shaughnessy and M. Andrade

Section 12 – Geostatistics, mapping and spatial data analysis **703**

A geophysical and spectrometric sensor data fusion approach for homogeneous within-field zone delineation 705
G. Buttafuoco, R. Quarto, F. Quarto, M. Conforti, A. Venezia, C. Vitti and A. Castrignanò

Wheat yield forecast using contextual spatial information 713
M. Fajardo, B. Whelan, P. Filippi and T. Bishop

Spatial analysis of mycotoxins in stored grain to develop more precise management strategies 721
R. Kerry, B. Ingram, E. Garcia-Cela and N. Magan

An iterative region growing algorithm to generate fuzzy management zones within fields 729
C. Leroux, H. Jones and B. Tisseyre

Investigating the harmonization of highly noisy heterogeneous datasets hand-collected over the same study domain 735
L. Pichon, C. Leroux, V. Geraudie, J. Taylor and B. Tisseyre

Can mapping of within-vineyard variability be facilitated using data from multiple vineyards? 743
B. Sams, R. Bramley, V. Pagay, L. Sanchez, C. Ford and N. Dokoozlian

Assessing the spatial variability of winter wheat yield in large-scale paddy fields of Japan using structural equation modelling 751
T.S.T. Tanaka, Y. Kono and T. Matsui

Section 13 – Robotics, guidance and automation 759

Overfitting a convolutional neural network to support annotations of weeds 761
M. Dyrmann, R.N. Jørgensen, M.S. Laursen and S. Skovsen

Hybrid topological location and mapping for autonomous agricultural robots 767
L. Emmi, J. Dufour, V. Cadenat and M. Devy

Determination of a field boundary using ISO 11783 and D-GNSS data acquired during a ploughing operation 775
A. Heiß, D.S. Paraforos and H.W. Griepentrog

Proof-of-concept modular robot platform for cauliflower harvesting 783
F.B. Klein, A. Wilmot, V.F. de Tejada, B.L. Rodriguez, I. Requena, S. Busch, A. Rondepierre, T. Auzeeri, T. Sauerwald, W.F.P. Andrews, H. Rihan, M.P. Fuller and M.F. Stoelen

Deep learning-based image segmentation for grape bunch detection 791
R. Marani, A. Milella, A. Petitti and G. Reina

Soft manipulator robot for selective tomato harvesting 799
A. Mohamed, J. Shaw-Sutton, B.M. Green, W. Andrews, E.J. Rolley-Parnell, Y. Zhou, P. Zhou, X. Mao, M.P. Fuller and M.F. Stoelen

Analysis of historical auto-steering position data for optimising headland turning time 807
D.S. Paraforos, R. Hübner and H.W. Griepentrog

Section 14 – Economics of precision agriculture 815

The economic value of on-farm precision experimentation 817
D.S. Bullock and T. Mieno

The emerging structure of the U.S. precision agriculture industry 825
J.R. McFadden and D.E. Schimmelpfennig

Financial and environmental performance of integrated precision farming systems 833
S.M. Pedersen, M. Medici, T. Anken, G. Tohidloo, M.F. Pedersen, G. Carli, M. Canavari, Z. Tsiropoulos and S. Fountas

Section 15 – Adoption of precision agriculture 841

How quickly do farmers adopt technology? A duration analysis 843
T.W. Griffin and E.A. Yeager

A collective framework to assess the adoption of precision agriculture in France: description and preliminary results after two years 851
N. Lachia, L. Pichon and B. Tisseyre

How does European adoption of precision agriculture compare to worldwide trends? 859
J. Lowenberg-DeBoer and B. Erickson

Collection of agricultural data and evaluation of their usefulness for farm management in Australian cropping and red meat industries 867
A. Zhang, J. Trindall and E. Hobman

Section 16 – Small farms/small holders and precision agriculture 875

Evaluating the potential benefits of field-specific nitrogen management of spring maize in northeast China 877
X. Wang, Y. Miao, R. Dong, Y. Guan, D.J. Mulla

Combining crop modelling and remote sensing to create yield maps for management zone delineation in small scale farming systems 883
H. Zha, D. Cammarano, L. Wilson, Y. Li, W.D. Batchelor and Y. Miao

Section 17 – On farm experimentation with site-specific technologies 891

Operating Farmer Innovation Groups (FIGs) for testing yield enhancing ideas using on-farm experimentation 893
C. Clarke, D. Brightman, S. Roques, S. Kendall, D. Kindred and R. Sylvester-Bradley

Improving yield mapping accuracy using remote sensing 901
R. Gonçalves Trevisan, L.S. Shiratsuchi, D.S. Bullock and N.F. Martin

Model validation of the dynamics of a no-till seeding assembly with a magneto-rheological damper system 909
G.M. Sharipov, D.S. Paraforos and H.W. Griepentrog

Variation across scales indicates that best progress in crop yields should come from farmer-centric research 917
R. Sylvester-Bradley, S. Clarke, D. Kindred, S. Roques, P. Berry and S. Welham

Site-specific treatment responses in on-farm precision experimentation 925
R.G. Trevisan, D.S. Bullock and N.F. Martin

Section 18 – Software and mobile apps for precision agriculture 933

Easy water stress detection system for vineyard irrigation management 935
G. Brunel, L. Pichon, J. Taylor and B. Tisseyre

A model for precision irrigation scheduling of soybeans for the South-eastern U.S. 943
V. Liakos, W. Porter, J. Kichler, A. Sawyer, D. Pavlou, A. Orfanou and G. Vellidis

Automated irrigation scheduling for drip-irrigated plum trees. 951
S. Millán, C. Campillo, J. Casadesús, M.J. Moñino, A. Vivas and M.H. Prieto

Drone dataflow – a MATLAB toolbox for extracting plots from images captured by a UAV 959
A.K. Mortensen, M.S. Laursen, R.N. Jørgensen and R. Gislum

Section 19 – Decision support for precision agriculture 967

A multi-disciplinary approach for the precision management of lodging risk 969
P. Berry, A. Blackburn, M. Sterling, Y. Miao, D. Hatley, D. Gullick, G. Joseph, D. Whyatt, D. Soper, J. Murray and C. Baker

Extending the CERES-Beet model to simulate leaf disease in sugar beet 977
E. Memic, S. Graeff and W.D. Batchelor

User needs for decision support functionalities in future crop protection software 985
J. Nikander, S. Hara, L. Pesonen, M. Jalli, P. Erlund and A.I. Nissinen

Automated mixed-scale data fusion for mapping of within-field variation in a national decision support system – the example of pH correction 993
K. Piikki, M. Söderström, H. Stadig and J. Martinsson

Near-real time winter wheat N uptake from a combination of proximal and remote optical measurements: how to refine Sentinel-2 satellite images for use in a precision agriculture decision support system 1001
S. Wolters, M. Söderström, K. Piikki and M. Stenberg

Section 20 – Data mining for precision agriculture 1009

Maize yield prediction based on artificial intelligence using spatio-temporal data 1011
A. Nyéki, C. Kerepesi, B. Daróczy, A. Benczúr, G. Milics, A.J. Kovács and M. Neményi

Keyword index 1019

Author index 1025

Section 1
Precision agriculture

Optimizing durum wheat cultivation in Northern Italy: assessing proximal and remote sensing derived from different platforms for variable-rate application of nitrogen

M.A. Bruce, J. Moretto*, R. Polese and F. Morari
Department of Agronomy, Food, Natural resources, Animals and Environment (DAFNAE), University of Padua, Viale dell'Università 16, Padua 35020, Italy; johnny.moretto@hotmail.it

Abstract

To make decisions on variably applying fertilizers (i.e. Variable Rate Application, VRA) during durum wheat cultivation in Northern Italy, it is necessary to estimate the extant nitrogen within the crop biomass. Previously proven non-destructive methods of analyzing N content of plants utilize chlorophyll content and plant biomass measureable by reflected wavelengths emitted from the associated sensors. The objective of this work is to assess the ability of different combinations of vegetation indexes (VIs) derived from proximal and remote sensors in both an active and passive approach to describe the N uptake dynamic in durum wheat (*Triticum durum* Desf.) var. Biensur. A predictive model for future real-time assessment of N content at field scale by means of proximal sensing was finally tested.

Keywords: durum wheat, NDVI, remote sensing, proximal sensing, nitrogen uptake

Introduction

Durum wheat (*Triticum durum* Desf.) grown in the volatile weather conditions of Northern Italy, is subject to high levels of spatial and temporal variability that impact nitrogen (N) uptake providing the need for active variable rate application (VRA) (Stone *et al.*, 1996; Quebrajo *et al.*, 2015). Measuring the N content of a cropping system at field scale by means of biomass harvesting for lab analysis is both impractical and costly. Non-destructive sampling techniques for estimating plant N such as the use of vegetation indices (VIs) from proximal and remote sensors have been proven accurate and practical for agricultural fields (Chen, 2012; Lopresti *et al.*, 2015). Although these VIs in the form of normalized difference vegetation index (NDVI) have demonstrated the ability to quantify the N variability of wheat plants, there are many devices and instruments currently available that necessitate additional testing. Therefore, the primary objective of this research was to employ active and passive sensors of both remote and proximal sensing platforms, to measure NDVI of a durum wheat field throughout the growing season.

Materials and methods

The location used for this research was a rectangular 13.6 ha field located in Mira, Italy (45°22' N, 12° 8' E and 2 m a.s.l.) and cultivated with Durum wheat (*Triticum durum* Desf. var. Biensur) in the period November 2017- June 2018. The field was divided into three management zones – low, medium and high fertility – according to soil apparent electrical conductivity and other soil properties (Morari *et al.*, 2018). Each of the three management zones was equally subdivided into variable rate application (VRA) and conventional fertilization areas. VRA areas received an assortment of prescribed rates of N, while conventional areas received standard uniform rates typical for durum cultivation in NE Italy. In addition to these zones, four field-length by six meter wide strips: two nitrogen poor (no applied N) and two nitrogen rich (250 kg/ha) were used for calibration of the VIs (Figure 1).

DOI 10.3920/978-90-8686-888-9_1, © Wageningen Academic Publishers 2019

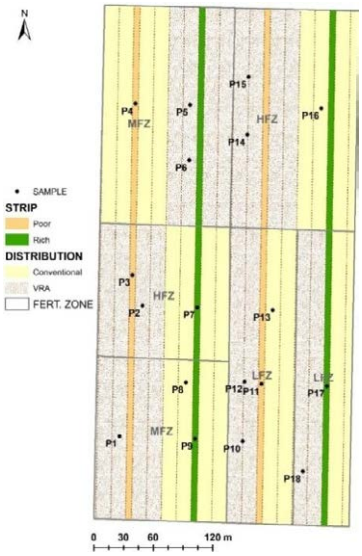


Figure 1. Study area with the overlain experimental scheme. 'Sample' represents the 18 sampling points, the 'Strip' refers to the N rich (rich) or no N added (poor) strips, the 'Distribution' refers to the (Conventional and VRA applications) and the 'Fert. Zone' divides the field into three different fertility zones: low (LFZ), medium (MFZ) and high (HFZ).

For the collection of NDVI, physiological data and biomass during the growing season, 1-m^2 sample plots were used sampled according to Table 1. These plots were placed in the immediate vicinity of 18 static point locations throughout the field. For remote sensing, sampling areas consisted only of the entire field at 10 m resolution. The plots were laid out and marked prior to each sampling event, with one plot at each of the 18 static points. All 1 m^2 plots, field boundaries and nearby corresponding static points were recorded each time with differential GPS (DGPS) from a Trimble Nomad equipped with a ProXT antenna and flagged appropriately. During each plot sampling a number of crop and soil parameters were measured: growth stage, plant height, VIs, leaf area index (LAI), chlorophyll content index (CCI), soil water content (SM), soil bulk density (BD), soil N mineral content.

Three instruments were used for VI estimation: 2 proximal and 1 remote. Tractor-mounted Ag Leader® technology included an In-Command computer display connected to a GPS antenna, equipped with differential correction, and two Opt-Rx sensors, was used for measuring NDVI. Opt-Rx was used to sample individual plots and the field in its entirety. The other proximal sensor was the 'Trimble GreenSeeker' (Trimble, 2010). This handheld device was used solely for sensing the 1-m^2 plots at 100 cm above the crop canopy.

Remote sensing entailed open access imaging from the Sentinel 2 satellites (European Space Agency). Sentinel 2 data had a spatial resolution of 10 meters with atmospheric correction, allowing for overall field assessment with NDVI but not accurate analysis of individual 1 m^2 plots.

LAI in each 1-m^2 plot was measured using a Sun Scan probe equipped with a Trimble Recon device (Delta-T Devices, n.d.) while chlorophyll content was quantified using a SPAD chlorophyll meter (Konica Minolat, n.d.). The phenological stage of the crops was determined by a visual assessment and coincided with an overhead and horizontal photo taken by a portable camera and plant height measurements. At the end biomass was collected, and dry biomass measured and forwarded to the lab for N analysis.

Table 1. Schedule of field operations.

Operation	Phenological stage	Date
1st Proximal/Remote Sensing Biomass collection, 18 points	tillering	03/28/18
2nd Proximal/Remote Sensing Biomass collection, 18 points	stem elongation	18/04/18
3rd Proximal/Remote sensing Biomass collection, 18 points	flagleaf	8/05/18
4th Proximal/Remote sensing Biomass collection, 18 points	anthesis	24/05/18
5th Biomass collection, 18 points	pre-harvest	26/06/18

Data analysis

Kriging method was used to provide NDVI field maps (Figure 2), evaluating the interpolation accuracy with the cross-validation method (ESRI ArcMap 10.2). Data collected in the 18 points were tested with the Moran test for spatial autocorrelation. Multiple regression was then applied to test factors affecting NDVI, considering the continuous factors of LAI, CCI and N uptake.

An empirical model describing N uptake according to NDVI and growing degree days (GDD) data was built using data collected in previous years (2011, 2012, and 2013) on the same field and wheat variety with Trimble Green Seeker. Proc-reg module (SAS 9.4) was used to estimate the coefficients of a linear model describing log-transformed N uptake according to the NDVI, accumulated growing degree days (GDD) and the interaction between the two variables. The model goodness was then tested with NDVI data measured in 2018.

Results

NDVI values obtained with both proximal (AgLeader OPT-Rx, GreenSeeker) and remote (Sentinel 2 satellites) sensors increased until the spike emergence and then decreased at the crop anthesis. Additional physiological parameters such as CCI followed the same pattern: 35.43, 45.31, 47.02 and 42.46, respectively. Conversely, average LAI values progressively increased until the crop anthesis (0.65, 2.47, 2.67 and 3.25) such as the dry biomass that peaked at 8.5 t/ha. On the same date total nitrogen uptake reached a maximum of 163.3 kg/ha in the N rich strip.

Interesting to note is that OptRx and Sentinel 2 satellites identify approximately the same spatial NDVI trend throughout the field (Figure 2). The remaining variation could be attributed to the differences in spatial resolution that are respectively 1 m (OptRx) and 10 m (Sentinel 2).

The multiple regression analysis with independent NDVI sensor values in relation to the dependent variable of N uptake showed the lowest Akaike information criterion (AICc) was attributed to both forms of proximal sensors over remote sensing shown in Table 2. Additional regression analysis tested the dependent variable of sensor NDVI values against the independent variables of LAI and CCI to show the significance of NVDI values from each sensing platform to the VI components as shown in Table 3. For all models performed, there was no spatial autocorrelation found through the Moran test.

The empirical model calibrated with the Green Seeker NDVI and N uptake values of years 2011-2013 showed the following linear equation:

$$\log N_{up} = -1.035 + (0.0015 \times GDD_{Acc}) + (3.77 \times NDVI) + (-0.0018 \times GDD_{Acc} \times NDVI) \quad (1)$$

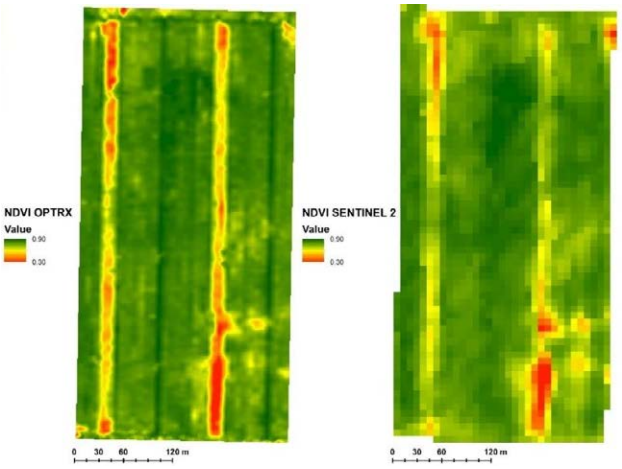


Figure 2. NDVI maps surveyed prior to the third fertilization of proximal sensing OptRx; AgLeader (NDVI OptRx – cell size 1×1 m; survey on 8 May 2018) and Satellite Sentinel 2 (NDVI Sentinel 2; cell size 10×10 m; survey 9 May 2018) respectively.

Table 2. Multiple regression for dependent N uptake.

Independent variable	Coeff.	Intercept	*P*-value	AICc	Moran test
NDVI Green Seeker	235.02	-92.8687	<0.0001	464	None
NDVI OPT-Rx	220.49	-82.1029	<0.0001	465.3	None
NDVI OPT-Rx Interpolated	241.24	-92.6522	<0.0001	470.1	None
NDVI Sentinel 2	209.03	-89.9994	<0.0001	493.2	None

Table 3. Multiple regression for independent NDVI components.

Dependent variable	CCI *P*-value	LAI *P*-value	AICc	Moran test
NDVI OPT-Rx interpolated	0.0006	<0.0001	-138.1	none
NDVI OPT-Rx	0.0005	<0.0001	-113.1	none
NDVI Green Seeker	<0.0001	0.0037	-107	none
NDVI Sentinel 2	0.2928	<0.0001	-99.8	none

The comparison of simulated and actual N uptake in 2018 (Figure 3) showed good predictive performances of Equation 1.

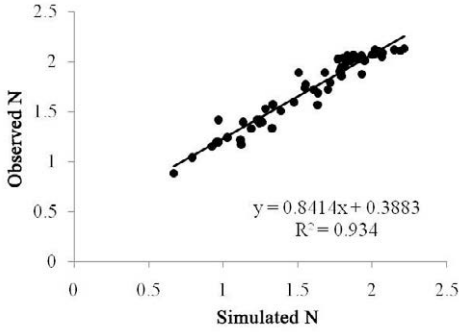


Figure 3. The graph above shows the simulated log transformed N uptake values of the deduced linear equation on the x-axis, with the observed N uptake values of 2018 on the y-axis.

Discussion

After completion of all 4 sampling events containing proximal and remote VI collection, it was observed that after the third sampling on 8/05/18, the VI values were no longer increasing in correlation to biomass due to crop maturity. Therefore, to produce a viable NDVI and N uptake correlation, only the first 3 sampling events were used in the statistical analysis.

It is evident that all forms of deployed proximal and remote sensing platforms for assessing VI either have or possess the potential to provide significant measurements of the N uptake dynamic of durum wheat ($P<0.0001$). The multiple regressions performed in SAS with independent variables set as NDVI values of the proximal and remote sensors used in relation to the dependent variable of lab analyzed N uptake showed proximal sensors as relatively more significant (lower AICc values than remote sensing). This may be explained by the additional regression models that associated the NDVI components of biomass and chlorophyll content quantified by LAI and CCI respectively. Proximal sensors contained significant values for both independent variables while the remote sensors were significantly associated with biomass values and not with chlorophyll content. This inability to significantly quantify chlorophyll content directly affects calculated NDVI values, perhaps as a result of the satellite's low spatial resolution (10 m).

For the establishment of an N uptake and NDVI correlation that is viable for more than one year, the three previous years of N uptake and Green Seeker measurements were used. By performing various models between the three years with common variables, it was observed that NDVI, accumulated GDD, and the interaction of the two, yields the most significant simulation of the log transformed N uptake values. By utilizing GDD, the NDVI values could be better associated across the three separate years. The resulting linear equation was then tested against 2018 NDVI values to produce simulated N uptake values similar to what was observed in the field by lab analysis. The significance of the simulated values against the observed values furthers the possibility of having a real time assessment within the field of N uptake derived solely from NDVI.

Conclusions

In conclusion, it was observed that while both proximal and remote sensors are viable platforms to quantify the N uptake dynamic of durum wheat in Northern Italy, active proximal sensors provide more accurate measurements of NDVI and subsequent N uptake. This was explained by their ability to derive values strongly correlated with both biomass and chlorophyll content of the targeted plant throughout the different phenological stages. Furthermore, by assessing the NDVI from the Trimble

Green Seeker and N uptake values of the three previous years, it was proven possible to formulate a model that was proved successful for accurately simulating N uptake values of additional years for this experimental site. These observations and yielded results validate the use of non-destructive NDVI platforms for quantifying the N uptake of cropping systems as to allow variable rate applications of fertilizers to be employed more efficiently.

Acknowledgements

This project was funded by Progetto AGER, n°2017-2194.

References

Chen, P., 2012. A Comparison of Two Approaches for Estimating the Wheat Nitrogen Nutrition Index Using Remote Sensing, 7, 4527-4548. https://doi.org/10.3390/rs70404527

Delta-T Devices. (n.d.). SunScan Canopy Analysis – Canopy Analyser – LAI – PAR. Retrieved August 8, 2018, from https://www.delta-t.co.uk/product/sunscan

Konica Minolat. (n.d.). A lightweight handheld meter for measuring the chlorophyll content of leaves without causing damage to plants. SPAD-502Plus. Retrieved from https://www.konicaminolta.eu/fileadmin/content/eu/Measuring_Instruments/ 2_Products/1_Colour_Measurement/6_Chlorophyll_Meter/PDF/spad-502plus_Feb2017_EN.pdf

Lopresti, M. F., Di Bella, C. M. and Degioanni, A. J., 2015. Relationship between MODIS-NDVI data and wheat yield: A case study in Northern Buenos Aires province, Argentina. Information Processing in Agriculture, 2(2), 73-84.

Quebrajo, L., Pérez-Ruiz, M., Rodriguez-Lizana, A. and Agüera, J., 2015. An Approach to Precise Nitrogen Management Using Hand-Held Crop Sensor Measurements and Winter Wheat Yield Mapping in a Mediterranean Environment. Sensors, 15(3), 5504-5517. https://doi.org/10.3390/s150305504

Stone, M. L., Solie, J. B., Raun, W. R., Whitney, R. W., Taylor, S. L. and Ringer, J. D., 1996. Use of Spectral Radiance for Correcting In-season Fertilizer Nitrogen Deficiencies in Winter Wheat. Transactions of the ASAE, 39(5), 1623-1631. https://doi.org/10.13031/2013.27678

Trimble, 2010. GreenSeeker ® RT100 Data Collection and Mapping System 2 GreenSeeker RT100 System User Guide. GreenSeeker User Manual. Retrieved from www.trimble.com

Integrating geospatial tools and a crop simulation model to understand spatial and temporal variability of cereals in Scotland

D. Cammarano[1*], J. Holland[1], B. Basso[2], F. Fontana[3], T. Murgia[4], C. Lange[1,5], J. Taylor[6] and D. Ronga[1,7]
[1]*James Hutton Institute, Errol Road, DD25BQ, United Kingdom; davide.cammarano@hutton.ac.uk*
[2]*Michigan State University, 307 Natural Science, East Lansing, MI, USA*
[3]*Universita' della Tuscia, Via S Maria in Gradi 4, 01100 Viterbo, Italy*
[4]*Universita' di Sassari, Via Enrico de Nicola 7, Sassari, Italy*
[5]*Ecole d'Ingénieurs de Purpan, 75 voie du TOEC, 31076 Toulouse, France*
[6]*ITAP, Irstea, University of Montpellier, Montpellier 34196, France*
[7]*Universita' di Modena e Reggio Emilia, Via Università 4, 41121 Modena, Italy*

Abstract

The objectives of this study were to: (1) quantify the factors causing spatial variability, and (2) design a proof of a concept for a support system that considers the impacts of temporal variability. The subdivision of a field based on six years of spatial yield data identified areas with different levels of stable and unstable zones. Overall, in the 0.0-0.30 m, the soil had between 40 to 60% across all the zones, while at depth, the clay content was about 20% and sand content 30 to 40%. The growing season rainfall for the period 1983-2018 was highly variable, ranging from 152 to 570 mm. The support system showed that there was a general trend where higher growing season rainfall (e.g. 566 mm) corresponded to higher yields, but this meant lower soil nitrogen content at harvest. However, the nitrogen leaching was not always lower and ranged between 20 to 70 kg N/ha. Results of this study will enable the development of an agronomic site-specific N management that would maximize the farm's net income and minimize the environmental footprint.

Keywords: spatial variability, temporal variability, barley, crop growth model, inter-annual variability

Introduction

Nowadays, the world is facing the biggest challenge of producing more food on the same or less land, and by reducing its environmental footprint while keeping farming profitable (Cammarano *et al.*, 2016). Precision agriculture (PA) offers a unique opportunity to achieve such trade-off, but its adoption for agricultural management has been slower than expected due to technological and socio-economic challenges (Miao *et al.*, 2018). The amount of any agricultural input applied (e.g. fertiliser) may be varied spatially within the field and temporally between years. In fact, for the same field, determination of the optimal amount of any agronomic input should consider site-specific soil properties, current seasonal crop growing conditions and interactions therein (Basso *et al.*, 2011). Spatial variability can be assessed routinely with a variety of tools (e.g. remote sensing, soil sensing) and the overlay of many thematic maps, such as for soil and crop yield, to divide the field into uniform management zones (Miao *et al.*, 2006). Several methods have been proposed to define such zones (Nawar *et al.*, 2017). However, an important step in taking the right agronomic decision is to quantify which factors cause the spatial and temporal variability in the field; with the latter component not receiving enough attention (McBratney *et al.*, 2005). Crop simulation models (CSM) can be used to consider the inter-annual weather impact on the soil-plant-atmosphere interactions (Jones *et al.*, 2003). They seek to simulate the effects and temporal interactions of water and nitrogen on crop growth as affected by weather and agronomic management (Jones *et al.*, 2003). They have been used in a variety of cropping systems and environmental conditions; and to

DOI 10.3920/978-90-8686-888-9_2, © Wageningen Academic Publishers 2019

quantify the temporal stability of management zones (Basso *et al.*, 2001; Koo and Rivington., 2005). Nitrate leaching from agricultural land is a consequence of over fertilization and the difficulties in matching crop demand with soil supply (Basso *et al.*, 2011). The European Union addressed the issue by defining nitrate vulnerable zones, which are areas where pollution of groundwater with nitrates is high (EC, 2000). In Scotland, farmers comply with such regulations by capping the amount and timing of N applications (STI, 2008). The objectives of this study were to: (1) quantify the factors causing spatial variability, and (2) use this information to design a proof of a concept for a support system that would consider the impacts of temporal variability.

Materials and methods

Measurements were taken on a commercial field (11 ha; 56° 33' N 3° 16' W; 50 m a.s.l.) in Scotland. The soil was classified as loam according to the USDA. Spring barley (cv. Concerto) was sown on 12th of April 2018 at 350 plants/m^2. In each zone (Figure 1), a transect of points was set and the soil was sampled at three depths (0-0.30; 0.30-0.60; 0.60-0.90 m) one month before sowing. Samples were analysed for bulk density, water retention characteristic, texture, mineral N, P, K, cation exchange capacity (CEC), organic matter and water content. Fertilization was uniformly distributed after emergence on the 1st of May 2018. Ammonium nitrate was applied for a total of 120 kg N/ha. Soil samples (for mineral N and soil water content), crop samples (for biomass and nitrogen content) and unmanned aerial vehicle (UAV) images were collected one day before sowing, three weeks after fertilization (23 May), at flowering (20 June) and harvest (16 Aug.). UAV data collected three weeks after fertilization was the only image used in this study.

The use of six consecutive years of winter/spring wheat and barley maps allowed the delineation of zones. The yield maps were generated by a combine harvester owned by the farmer. To define the zone, the approach developed by Maestrini and Basso (2018a,b) was used. Three stable zones were defined as High Stable (HYZ), Medium Stable (MYZ) and Low Stable (LYZ), while the unstable zones were named (UYZ) Unstable (Figure 1A).

The Decision Support System for Agrotechnology Transfer (DSSAT) v.4.7 was used as the CSM (Jones *et al.*, 2003). For this study, measured information of soil properties, initial soil water and soil N content were available. The crop coefficient for the cultivar Concerto was obtained by calibrating phenology, growth and yield on crop variety trials and on 'potential yield' experiments (irrigated

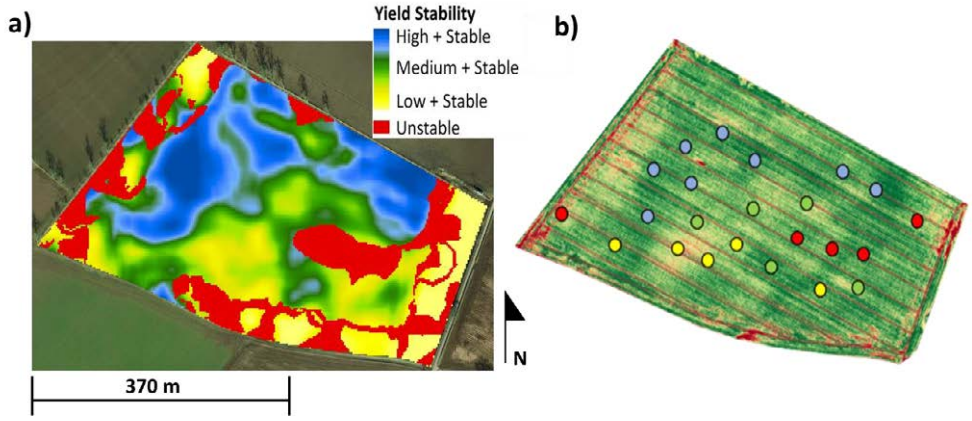


Figure 1. Spatial patterns of the (A) zones identified by overlaying six years of yield maps; and (B) drone image collected on the 23rd of May and the corresponding points sampled in the field.

and fertilized experiments). The calibrated model for this cultivar was evaluated on independent datasets across 4 experimental sites and 10 farms across Scotland (results not shown). The model was run using 36 years of historical weather data. Weather data for 2018 was available from the on-farm weather station while the historical 1980-2017 data were obtained from the NASA-Power website (power.larc.nasa.gov). Following the approach described in Basso *et al.* (2001, 2011), the DSSAT model was run for different yield stability zones with incremental N amounts (0N, 60N, 120N-farmers' practice, 180N). The model simulated yield and cumulative N leaching from sowing to harvest, and total soil N at harvest. The cumulative probability function of simulated yield and cumulative N leaching was calculated following the approach of Wallach *et al.* (2013). The coefficient of variation (CV) was used to quantify the amount of inter-annual variability for the simulated yield and the N leaching.

Results

While the HYZ was mostly concentrated in one portion of the field, the UYZ showed a spatial pattern being present at the edges of the field and the mid-lower portion (Figure 1A). The subsequent UAV image (NDVI), collected on the 23rd of May 2018 (22 days after N fertilization), showed how crop growth patterns corresponded with the zones defined in Figure 1A. The coloured points on Figure 1B indicate where samples were taken in each zone (Figure 1B).

At 0-0.30 m, most of the points were classified as loam soils according to the USDA (Figure 2). At 0.30-0.60 m and at 0.60-0.90 m, the different zones had texture that ranged from clay to sandy loam (Figure 2). Overall, in the 0.0-0.30 m, the soil had between 40 to 60% of sand and less than 20% of clay across all the zones, while at depth, the clay content was about 20% and sand content 30 to 40% (Figure 2). The LYZ and UYZ zones had the greatest variability in soil texture (Figure 2). Total growing season rainfall (April-Aug.) was 170 mm with most of the precipitation occurring before sowing and towards the end of the growing season (Figure 3). The plant available soil water content (PAWC) for the soil profile decreased for the period April-May due to 50% less rain with respect to the long-term average of 76 mm for the same period (Figure 3A). The HYZ showed PAWC value of 173 mm, while the MYZ, LYZ and UYZ showed values of 163, 170 and 164 mm, respectively (Figure 3A).

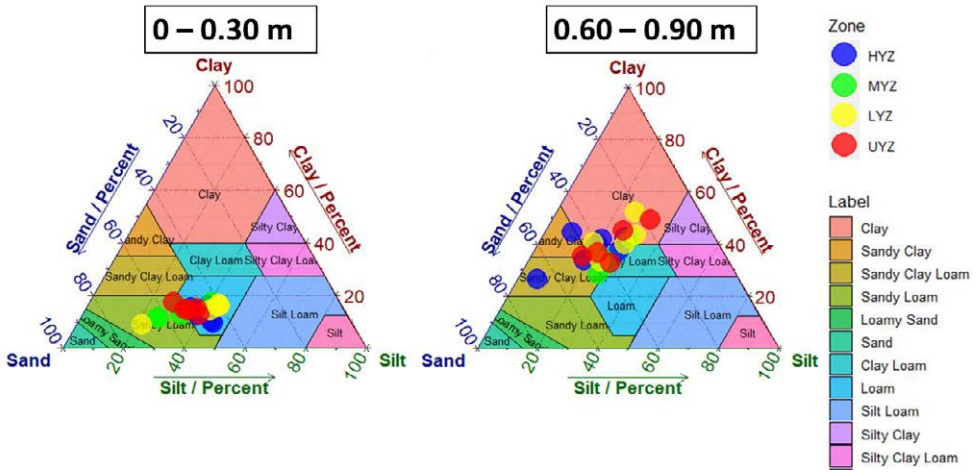


Figure 2. Soil texture triangle at two different depths, the coloured points represent the sampling taken at each zone.

The total soil nitrogen content (N) increased (especially in the UYZ) between sowing (April 12th) and the 22nd of May, a time in which fertiliser was added (May 1st) and no significant rainfall was recorded (Figure 3B). The UYZ had the highest amount of soil N during the whole growing season but, by harvest, a large amount was lost and it decreased to a total of 70 kg N/ha (Figure 3B).
The growing season rainfall for the period 1983-2018 was highly variable, ranging from 152 to 570 mm, while growing season minimum and maximum temperatures did not show high inter-annual variability (Figure 4A). The results of the simulations were presented as boxplots and showed that the HYZ was higher yielding than the other zones when the N fertilizer was between 0 and 120N (Figure 4B); at 180N the differences among the zones was minimal with an overall simulated yield of about 7,500 kg DM/ha. The inter-annual variability (1980-2018) of simulated yield and N leaching was also analysed in term of CV and is shown in Figure 4C. Overall, the simulated yield had a lower CV than N leaching with an overall 14% for yield and 33% for N leaching (Figure 4C). Figure 4D shows the relationship between the marginal yield increase per unit N applied versus the simulated N leaching. There was a higher marginal yield increase when fertilizer was applied and as the amount increased, the marginal yield return diminished. On the other hand, the amount of N leached increased with applied N for each zone, but it is more evident for the LYZ and UYZ (Figure 4D).
The relationship between the growing season rainfall, N leaching, total N in the soil at harvest, and grain yield is shown in Figure 5. Overall, there was a general trend where higher growing season rainfall (e.g. 566 mm) corresponded to higher yields, but this meant lower soil N content at harvest. However, the N leaching was not always lower and ranged between 20 to 70 kg N/ha (Figure 5).
At lower growing season rainfall (268 mm), the simulated yield was 3,964, 3,872, 3,645, and 3,786 kg DM/ha for the HYZ, MYZ, LYZ and UYZ, respectively (Figure 5). In this case, the N leaching was similar across the different zones (a mean of 25 kg N/ha) but the residual N at harvest was different, with 72 kg N/ha for the HYZ, 58 kg N/ha for the MYZ, 63 kg N/ha for the LYZ and 66 kg N/ha for the UYZ (Figure 5). When the growing season rainfall was 500 mm, the HYZ showed a simulated yield of 8,184 kg DM/ha, while the MYZ, LYZ and UYZ were 7,976, 7,961 and 8,110 kg DM/ha (Figure 5). The N leaching was 38, 37, 42 and 43 kg N/ha, for the HYZ, MYZ, LYZ and UYZ, respectively (Figure 5), but the residual soil N was about 13 kg N/ha for all zones (Figure 5).

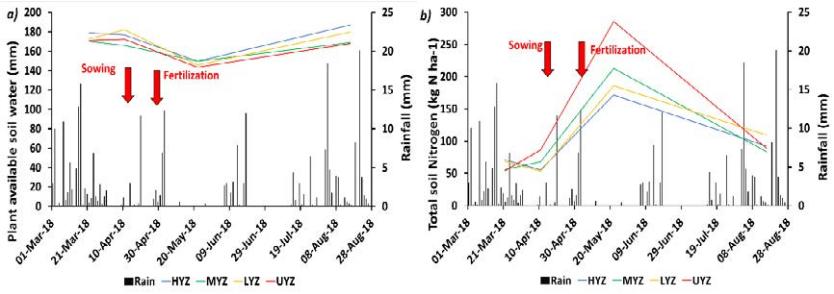


Figure 3. Growing season rainfall (black bar; secondary Y-axis) and (left) plant available soil water content and (right) total soil nitrogen for the different yield stability zones; High yielding zone (HYZ) (red line), medium yielding zone (MYZ) (green line), low yielding zone (LYZ) (yellow line) and unstable yielding zone (UYZ) (red line). The thick red lines indicate the time of sowing (12 April 2018) and N fertilization (1 May 2018).

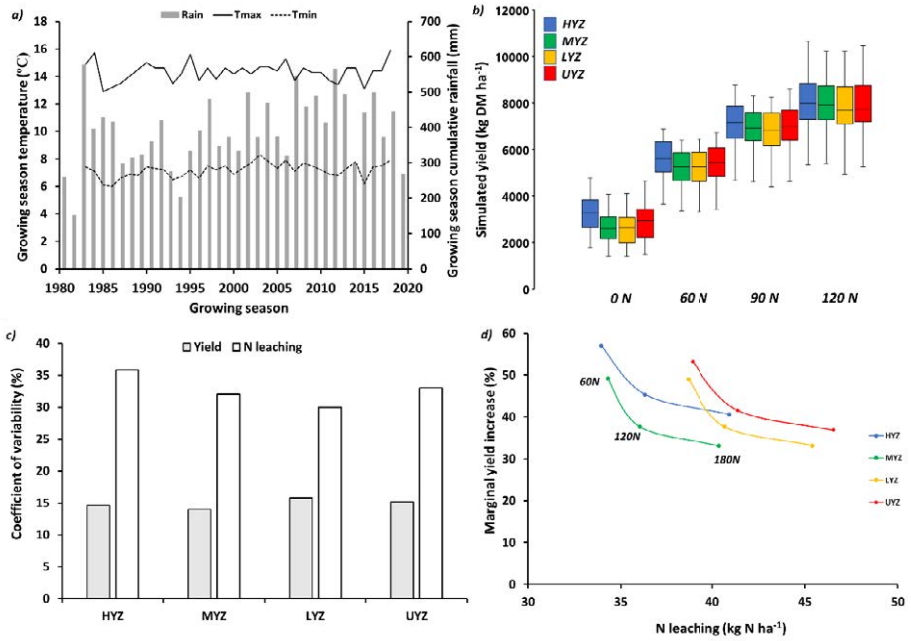


Figure 4. (A) values of cumulated growing season rainfall (grey bars), growing season maximum (full black line) and minimum (dotted black line) temperature; distribution of (B) simulated yields at 0N, 60N, 120N and 120N (C) coefficient of variation for the simulated yield (grey bars) and the N leaching (white bars) for each zone; and (D) relationship between marginal yield increase per unit of nitrogen applied and simulated N leaching.

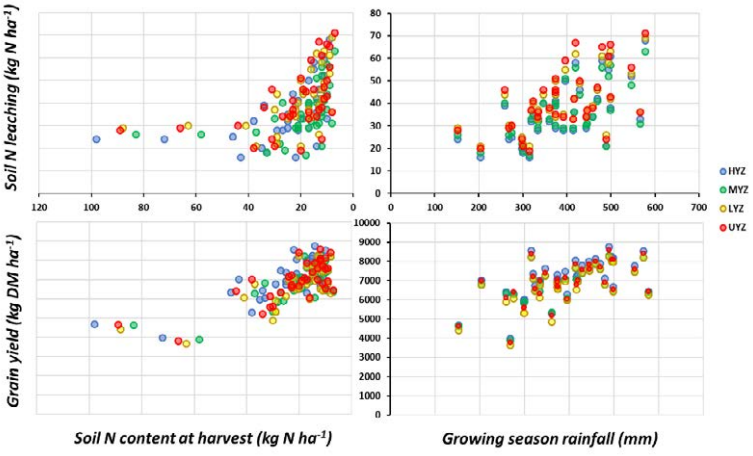


Figure 5. relationship between simulated N leaching, growing season rainfall, soil N content at harvest, and grain yield for the HYZ (blue dots), the MYZ (green dots), LYZ (yellow dots) and UYZ (red dots).

Discussion

There was a high degree of spatial variability within the field. However, the overlay of six years of yield maps allowed for discriminating stable and unstable yield zones in the field. The spatial variability of crop growing in the field was also confirmed by a UAV image taken 21 days after fertilization. The HYZ showed better growing conditions than other zones (Figure 1). Soil texture at depth below 0.30 m had an impact with respect to the water and N distribution, and bulk density results confirmed that below 0.40 m, the LYZ had very compact layers (with values of 1.8 g/cm^3; data not shown) while other zones had a lower bulk density (below 1.6 g/cm^3). This spatial pattern influenced the soil mineral N and soil water distribution and plant uptake where the UYZ tends to accumulate more N in the whole soil profile than other zones. The use of long-term weather data with a CSM demonstrated how the inter-annual variability affects crop growth and the environmental footprint of fertilization management. While 2018 was a very 'dry' season for Scottish growing conditions, other years have shown similar rainfall patterns with very 'dry' growing seasons (e.g. 1982 or 1994, Figure 4A). Results of this study regarding the impacts of growing season rainfall agree with the findings of Basso *et al.* (2012) conducted in a different environment. This indicates that even in such northern latitudes, growing season rainfall is an important determinant of grain yield for spring barley. The zones tended to respond differently to the inter-annual variability. However, the long-term simulations showed that there was a consistency between the map of the zones and the simulated yield. The HYZ was always higher yielding, even in drier years, while the UYZ was not stable high or low but tended to 'flip' between years. The HYZ was very responsive to rainfall and showed lower N leaching levels and low soil total N at harvest at higher rainfall amounts with respect to the other zones (Figure 5). It is interesting to notice that the simulated soil N at harvest for the dry years in each zone is in line with the observed data collected in 2018. For all the low rainfall growing seasons, there was low N leaching amount, and a high soil total N at harvest. This has important implications in managing N fertilization because, in the UYZ and LYZ, extra care should be taken in assigning a given amount of site-specific N. In fact, Figure 4D showed that these zones, fertilized with 120 kg N/ha like the whole field, are very prone to N leaching and, with respect to the EC regulations on Nitrate Vulnerable Zones, 60 to 90 kg N/ha would be the amount to apply (Figure 4D). Basso *et al.* (2011) defined tactical and strategic fertilization management amounts in each zone of a field using a CSM as decision support system. An integration of modelling results and on-farm experimentation will be planned for the next growing season (2019) with the aim of improving the efficacy of site-specific N fertilization.

Conclusions

The subdivision of the field into zones based on historic yields showed zones with different levels of yield potential and stability. This included an unstable zone where the historical yield tended to change from high to low values between years. The classification of yield into yield stability zones was confirmed by a subsequent UAV image, taken 20 days after N fertilization, and a crop simulation model run with long-term historical weather data. These results indicated the robustness of a yield-based zoning approach. Growing season rainfall, soil texture below 0.30 m and soil bulk density impacted the distribution of soil water and N and the crop growing conditions in each zone. By considering the spatial and temporal variability of each zone, an agronomic site-specific N management could be better tailored for each zone in order to maximize the farm's net income and minimize the environmental footprint. Results of this study showed that at 120 kg N/ha the simulated yield was higher. However, the marginal yield return per N applied tended to be lower and accompanied by a higher N leaching.

Acknowledgements

This project was funded by the Scottish Government Seedcorn Fund (18.12). We thank the Reid family for the availability in using their farm.

References

Basso, B., Fiorentino, C., Cammarano, D., Cafiero, G., Dardanelli, J., 2012. Analysis of rainfall distribution on spatial and temporal patterns of wheat yield in Mediterranean environment. European Journal of Agronomy 41 52-65.

Basso, B., Ritchie, J.T., Cammarano, D., Sartori, L., 2011. A strategic and tactical management approach to select optimal N fertilizer rates for wheat in a spatially variable field. European Journal of Agronomy 35, 215-222.

Basso, B., Ritchie, J.T., Grace, Pierce, F.J., Braga, R.P., Jones, J.W., 2001. Spatial validation of crop models for precision agriculture. Agricultural Systems 68, 97-112.

Cammarano, D., Roetter, R.P., Asseng, S., Ewert, F., Wallach, D., Martre, P., *et al.*, 2016. Uncertainty of wheat water use: Simulated patterns and sensitivity to temperature and CO_2. Field Crop Research 198, 80-92.

EC, 2000. Council directive of 23th October 2000 establishing a framework for Community action in the field of water policy (2000/60/EC). Official Journal of the European Union 22/12/2000 L327/1 to 73.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., *et al.*, 2003. The DSSAT cropping system model. European Journal of Agronomy 18, 235-265.

Koo, J., Rivington, M., 2005. Report on the Meta-analysis of Crop Modelling for Climate Change and Food Security Survey Climate Change. Agriculture and Food Security (CCAGS), Challenge Program, Denmark (available at www.ccafs.cgiar.org).

Maestrini, B., Basso, B., 2018a. Predicting spatial patterns of within-field crop yield variability. Field Crops Research 219, 106-112.

Maestrini, B., Basso, B., 2018b. Drivers of within-field spatial and temporal variability of crop yield across the US Midwest. Scientific Reports 8, 14833. doi: 10.1038/s41598-018-32779-3

McBratney, A., Whelan, b., Ancev, T., 2005. Future directions of precision agriculture. Precision Agriculture 6, 7-23.

Miao, Y., Mulla, D., Robert, P.C., 2018. An integrated approach to site-specific management zone delineation. Frontiers of Agricultural Science and Engineering. 5, 432-441.

Miao, Y., Mulla, D.J., Batchelor, W.D., Paz, J.O., Robert, P.C., Wiebers, M., 2006. Evaluating management zone optimal nitrogen rates with a crop growth model. Agronomy Journal 98, 545-553.

Nawar, S., Corstanje, R., Halcro, G., Mulla, D., Mouazen, A.M., 2017. Delineation of Soil Management Zones for Variable Rate Fertilization: A review. Advances in Agronomy 143,175-245.

Scottish Statutory Instruments (STI), 2008. The Action Programme for Nitrate Vulnerable Zones (Scotland) Regulation. Scottish Government, Edinburgh (www2.gov.scot/Publications/2012/03/4576)

Wallach, D., Makowski, D., Jones, J.W., Brun, F., 2013. Working with dynamic crop models: methods, tools, and examples for agriculture and environment. Amsterdam, the Netherlands: Elsevier Science, 504 pp.

Comparing satellite and high-resolution visible and thermal aerial imaging of field crops for precision irrigation management and plant biomass forecast

A. Chen*, V. Orlov-Levin, O. Elharar and M. Meron
MIGAL Galilee Research Institute, Kiryat Shmona 11016, Israel; assafc@migal.org.il

Abstract

RGB and thermal aerial imaging campaigns using unmanned aerial vehicles were deployed in the Upper Galilee of Israel to follow up and evaluate the uniformity and efficiency of cotton irrigation. Foliage intensity was enhanced by a green-red vegetation index (GRVI). Thermal images were used to measure plant and soil temperature and plant water stress over time. Sentinel-2 and Venus satellite imaging of the cropped fields were analyzed in order to distinguish irrigation uniformity issues using several-meter-scale spatial resolutions. Both GRVI and thermal product maps were found suitable for detecting and determining irrigation non-uniform patterns, as confirmed by field-based biomass measurements. Venus images were able to detect crop irrigation non-uniformities. This research demonstrates the efficient and complementary usage of RGB and Thermal high-resolution images for irrigation uniformity evaluation and management. These abilities could be combined with satellite imaging as a refined tool for large-scale general detection of irrigation uniformity issues.

Keywords: irrigation uniformity, green-red vegetation index (GRVI), lateral move irrigation, unmanned aerial vehicle (UAV), Venus satellite

Introduction

Precision irrigation can be defined as matching water application to the crop needs which is rarely uniform in space, time and amount. One of the main sources of crop growth variability in semi-arid climates is the lack of irrigation uniformity. Awareness of the extent and severity of possible uniformity problems requires spatial and temporal crop water status information in sufficient resolution and accuracy, attainable mainly by aerial sensing and imagery (Haghverdi *et al.*, 2015; Nahry *et al.*, 2011; Zhang *et al.*, 2011). New enabling technologies for precision irrigation information acquisition became affordable and popularized with the appearance in the open market of low cost, high performance small sized unmanned aerial vehicles (UAV) carrying high resolution digital camera and high resolution, fast revisiting, low cost satellite services in the visible (RGB), near infrared (NIR) and long wave (Thermal) infrared channels.
Spectral vegetation indices built from combinations of channels at various wavelengths allow for better information extraction from remotely sensed data because they reduce the effects of soil, view angle and topography, while enhancing the focus on the visibility of the vegetation (Hunt *et al.*, 2012). The objective of this research was to test the efficiency of monitoring the application uniformity of a lateral move irrigation machine by high resolution long wave thermal, infrared and visible light aerial mapping, using the green-red vegetation index (GRVI), against Venus and Sentinel-2 satellite imagery including visible light and near infrared indices. This is the first time (to the best of the authors' knowledge) that Venus satellite imaging ability is tested and validated (via comparison to high-resolution imagery) in recognizing irrigation uniformities in commercialized fields.

DOI 10.3920/978-90-8686-888-9_3, © Wageningen Academic Publishers 2019

Materials and methods

UAV imaging system

DJI (Shenzhen, China) Phantom 4 Advanced quadcopter UAV (model number: WM332A) was used as the RGB flying platform. The UAV is equipped with a built-in RGB camera with a 4,864×3,648 pixel 4K resolution CMOS sensor, a 24 mm (35 mm eq.) lens with field of view (FOV) of 84°, in a 3-axis stabilized gimbal (https://www.dji.com/phantom-4-adv/info). The UAV was flown using Pix4D (Lausanne, Switzerland) Capture pre-programmed flightpath control software. DJI Inspire-I UAV (model number: T601) was used as the thermal flying system, equipped with DJI Zenmuse-XT radiometric thermal camera (model number: ZXT01) with 640×512 pixel recording size, 32° FOV lens, fully stabilized gimbal control, and spectral band of 7.5-13.5 μm, with both RJPEG and TIFF file formats (https://www.dji.com/zenmuse-xt/info).

Flight campaigns

Flight campaigns (Table 1) were conducted at Kibbutz Hagoshrim lateral move irrigated commercial cotton field (33°11'57.75" N, 35°38'19.15" E), located in the Upper Galilee region in the northern part of Israel. The area has a Mediterranean climate, characterized by wet, mild winters with mean minimum and maximum temperatures of 7 and 14 °C, respectively, and hot, dry summers with mean minimum and maximum temperatures of 19 and 32 °C respectively. Annual winter rainfall is in the range of 400-600 mm, while summer crops utilize 80-120 mm of winter soil water storage for the initial growth periods. The cotton field size at Hagoshrim was 17 ha (850×200 m). Cotton sowing date: 11 March 2018; irrigation start: 25 May 2018 (75 days from sowing); harvest date 2 October 2018 (205 days from sowing).

Flight courses were created with the Pix4Dcapture software, which was also used to automatically pilot both the DJI Phantom 4 Advanced and DJI Inspire-I UAVs according to the flight path with front and side overlap percentage of 70% (Phantom 4 Advanced) and 90% (Inspire-I) in order to ease the task of mosaicking. All imagery was taken on full sun, cloudless days, around mid-day, in order to avoid changes in acquisition conditions (e.g. temperature, shade).

Data processing

The images collected in each flight campaign were mosaicked and georeferenced using the Pix4Dmapper software for the RGB images and Pix4DFields software for the thermal images. Whereas Pix4Dmapper proved to be adequate for mosaicking the RGB images, Pix4DFields performed better mosaicking of the thermal images in RJPEG format. It was impossible to mosaic

Table 1. Aerial scan campaigns details at Hagoshrim cotton field, Summer 2018.

Date	Time	Elevation (m)	Spatial resolution (m)	Days from sowing
RGB				
11/06/2018	13:30	100	0.043	92
17/06/2018	12:30	100	0.043	98
26/06/2018	13:00	100	0.043	107
29/07/2018	11:00	100	0.043	140
12/08/2018	11:30	100	0.043	154
Thermal				
17/06/2018	14:00	150	0.13	98
26/06/2018	14:00	150	0.13	107
29/07/2018	12:00	150	0.13	140

TIF format thermal images due to on-flight faulty automatic sensor calibration process which was impossible to disable. Furthermore, the camera has a long processing time during TIF format image acquisition which resulted in spatial gaps that hindered mosaicking. GRVI was calculated and mapped using the Image Analysis NDVI function of ArcGIS 10.5 (Esri, Redlands, CA, USA) with the Green replacing the NIR channel.

Sentinel-2 Level-2A atmospherically corrected images of the Hagoshrim test site from 20 June 2018 were acquired courtesy of the Copernicus Open Access Hub. Venus satellite Level-1 radiometrically corrected and top-of-atmosphere reflectance images of the Hagoshrim test site from 27 June and 12 August 2018 were acquired courtesy of CNES (French government space agency) ISA (Israeli space agency) and the scientific centre at the Ben-Gurion University, Israel. Several VIs were created and compared to the UAV images, in order to check whether it is possible to use Sentinel-2 and Venus satellite imaging (with spatial resolution of 10 and 5 m, respectively) to determine irrigation uniformity issues. The following VIs were checked: NDVI (Rouse *et al.*, 1973), green normalized difference vegetation index (GNDVI) (Gitelson *et al.*, 1996), and GRVI (Motohka *et al.*, 2010; Tucker, 1979). UAV images were compared to similar date satellite images for the sake of proper comparison (same day as Venus, and 3 days apart from Sentinel-2 imagery).

Field crop sampling

In order to validate the images taken by the UAV, it was essential to monitor plant growth and plant vigour *in situ*. The field was marked vertically in 12 different rows (i.e. segments) along the field (Figure 1), according to differences in plant vigour that were detected by images taken during a flight campaign on 29 July (Figure 2). According to these segments, plant samples were collected on August 15 along these rows. From each segment, 2 m row of plants covering 1.93 m^2 were cut from plant base, and its biomass was measured. Plant average height was also measured, as an indication for plant vigour and biomass. The same segments were sampled again for cotton yield on 2 October (date of harvest): similar 2 m rows of cotton bolls were harvested and their mass was measured. The average GRVI of each segment from the image of 29 July was determined using ArcGIS10.5, and compared to field measured plant biomass, plant height and plant yield.

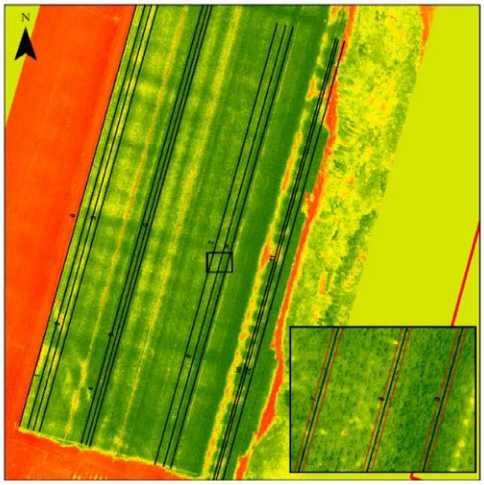


Figure 1. GRVI image of the cotton field in Hagoshrim on 29/07/2018 marked with the 12 segments chosen for field crop sampling. Enlarged frame indicating zoom-in to segments. GRVI values decrease from green to yellow to red.

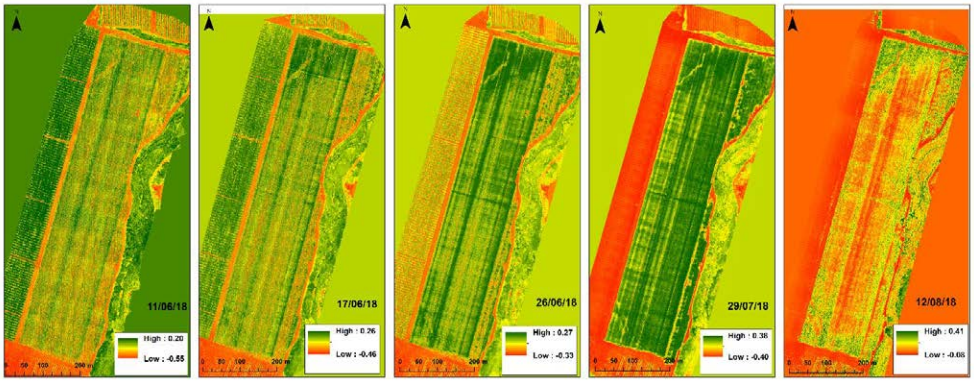


Figure 2. GRVI images of the cotton field in Hagoshrim with dates indicated next to each image. GRVI values decrease from green to yellow to red.

Results

GRVI images

As can be seen in Figure 2, throughout the whole campaign, some rows have dark green colour, resembling dense growth and vigour, while other lines are less green – indicating retarded or vigorous growth resulting from insufficient or over-irrigation. There is extreme variability in growth intensities between plant rows and along the rows. These differences are due to problems in sprinkler discharge rates uniformity in the lateral move system between the rows (as verified by the farmer), and in uneven discharge rates of the whole rig moving along the rows.

The sequence of GRVI images following temporally the growing season is in line with development from partial cover beginning of June to full cover in July, and decrease in vigour and maturation in August.

Thermal images

Considering the limitations of the Zenmuse XTR camera TIFF file format mosaicking, RJPEG format was used instead. Therefore, relative rather than absolute temperatures were obtained. As can be seen in Figure 3, this fact did not pose a problem in detecting differences in plant temperatures. However, mosaicking proved to be problematic in certain cases (Figure 3, 29/07/18). This is due to camera calibration issues, which is out of the scope of this paper.

Thermal images are in par with the GRVI image findings. As can be seen in Figure 3, there is direct correlation between the scanning methods in depicting the plant vigour differences resulting from irrigation non-uniformities. Soil temperatures appear in thermal images in the 50 °C range against the 30 °C ranges of foliage, therefore in partial crop cover, thermal imagery is expected to present similar results to GRVI, as seen in the first two dates. On the third date, the crop is in full cover and foliage temperatures show momentary water stress, rather than crop cover. Some of the between-row irrigation variability is expressed in increased soil water storage, and in colder canopy of less water stress, while other dense canopies are warmer.

Crop biomass and yield correlation to GRVI

As can be seen in Figure 4, there is relatively high (for field data) correlation between the GRVI and plant height (R^2=0.50) and plant biomass (R^2=0.35), whereas correlation to yield is low.

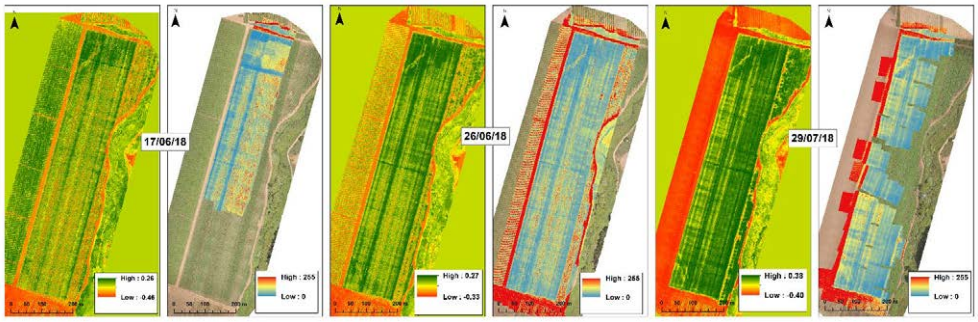


Figure 3. Thermal (right of pair) and corresponding GRVI images (left of pair) of the cotton field in Hagoshrim with dates indicated next to each image. Thermal images: red indicating hotter temperatures, blue colder. GRVI values decrease from green to yellow to red.

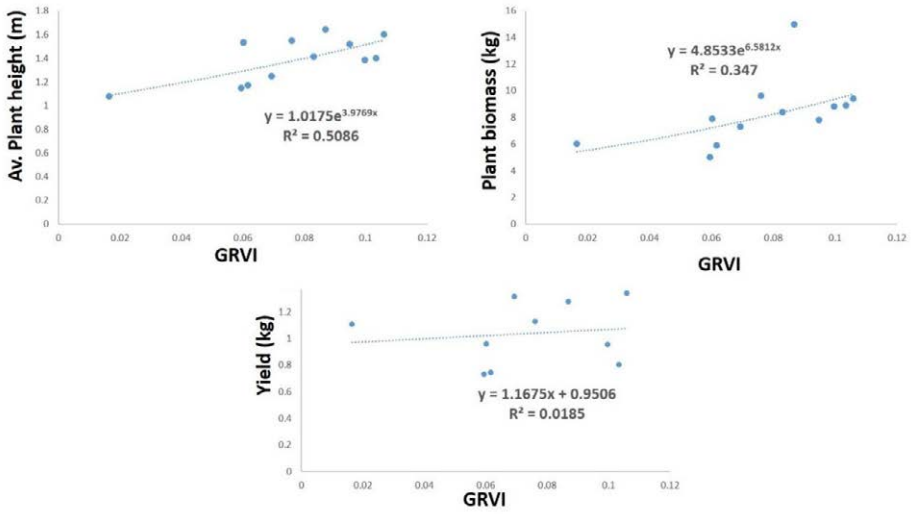


Figure 4. Plant height, biomass and seed cotton yield as a function of GRVI at 12 different locations according to UAV image of 29 July and field sampling on 15 August (biomass and height) and 2 October (yield).

Venus imagery

Mid-season image of the cotton field (Figure 5) at near full crop cover revealed that GRVI was able to roughly detect the main differences in plant biomass differences, compared to the UAV maps, while NDVI and GNDVI suffered from saturation. The image from August, of maturing crop towards the end of the growing season was less saturated and, is similar across all checked vegetation indices and on par with the UAV image portraying vegetation variability.

Sentinel-2 images

Even at sparser crop cover, only the GRVI index was partially able to detect the differences in plant vigour compared to the UAV image (Figure 6). The NDVI image was less saturated than the GNDVI but only resembles the pattern of the GRVI image.

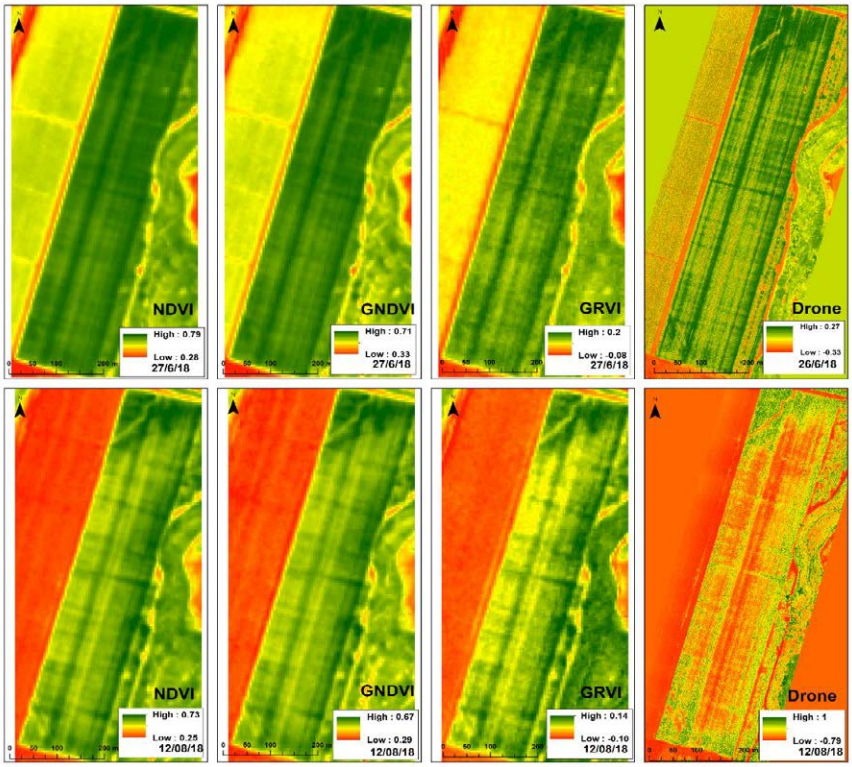


Figure 5. NDVI, GNDVI, and GRVI vegetation indices based on Venus imagery of 27/6/2018 (top) and 12/08/2018 (bottom) and the corresponding UAV GRVI image of the cotton field at 26/6/2018 and 12/08/2018, respectively.

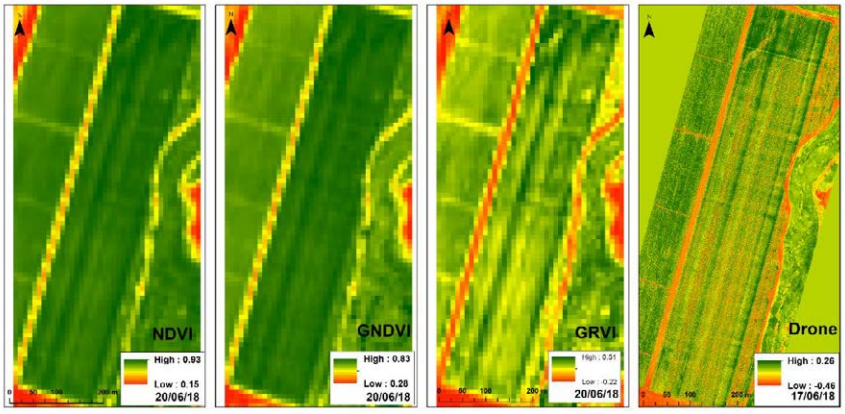


Figure 6. NDVI, GNDVI, and GRVI vegetation indices based on a Sentinel-2 imagery of the cotton field in Hagoshrim test site from 20/06/2018 and the corresponding UAV GRVI Image from 17/06/2018.

Discussion

Plant growth and vigour (defined by greenness and biomass) variabilities were apparent in the cotton field from the first UAV campaign, and intensified towards mid-season. The differences in plant height and biomass were visible to the naked eye, as 'waves' of high growing plant rows vs low rows. This variability across the lateral move rig can be attributed to water supply differences in the sprinkler discharge and/or their spatial spacing/distribution. Both the RGB and the thermal images of UAV were successful in depicting plant vigour variabilities. There is a trade-off between high spatial resolution UAV imagery with limited spatial resolution, low coverage capacity and high costs, and lower spatial resolution satellite imagery with extremely high capacity and relatively low or even no cost. In this respect, the Venus satellite RGB channels GRVI with 2-day revisit time and 5 m ground resolution provided reasonable detection ability of crop growth variability patterns for irrigation uniformity monitoring, even at crop cover conditions where NDVI and GNDVI and Sentinel 2 imagery became chlorophyll saturated (Daughtry *et al.*, 2000; Eitel *et al.*, 2009; Hunt *et al.*, 2011). Comparing field crop measurements of crop height, biomass and yield to Venus and Sentinel vegetation indices response at the various sampled segments, could be used as a tool to find which index correlates best to any type of field measurement and for the creation and calibration of new satellite-bands-based vegetation indices.

Conclusions

The success of Venus (and Sentinel-2 to a lesser extent) in depicting crop vigour or variability within a commercial field setting opens new opportunities for freely and highly available, large scale monitoring and identifying of irrigation malfunctions, and plant vigour in-field non-uniformities. Satellite imaging can be used as a 'first responder' in identifying areas suspected of non-uniformities, and report these areas for further inspection by the farmer. A UAV could be used to fine tune the satellite findings of suspected areas. The high temporal resolution of Venus of two days is highly adequate for early response to differences by identifying them early and allowing for a timely response.

Acknowledgements

Special thanks to Dr Onn Rabinovitz, Min. of Agriculture field crops advisor, for his valuable assistance.

References

Daughtry, C.S.T., Walthall, C.L., Kim, M.S., De Colstoun, E.B. and McMurtrey Iii, J.E., 2000. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sensing of Environment*, *74*(2), 229-239. https://doi.org/10.1016/S0034-4257(00)00113-9

Eitel, J.U.H., Long, D.S., Gessler, P.E., Hunt, E.R., & Brown, D.J., 2009. Sensitivity of Ground-Based Remote Sensing Estimates of Wheat Chlorophyll Content to Variation in Soil Reflectance. *Soil Science Society of America Journal*, *73*(5), 1715-1723. https://doi.org/10.2136/sssaj2008.0288

Gitelson, A., Kaufman, Y. & Merzlyak, M., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, *58*(3), 289-298. https://doi.org/10.1016/S0034-4257(96)00072-7

Haghverdi, A., Leib, B.G., Washington-Allen, R.A., Ayers, P.D., & Buschermohle, M.J., 2015. Perspectives on delineating management zones for variable rate irrigation. *Computers and Electronics in Agriculture*, *117*, 154-167. https://doi.org/10.1016/j.compag.2015.06.019

Hunt, E.R., Daughtry, C.S.T., Eitel, J.U.H., & Long, D.S., 2011. Remote sensing leaf chlorophyll content using a visible band index. *Agronomy Journal*, *103*(4), 1090-1099. https://doi.org/10.2134/agronj2010.0395

Hunt, E.R., Doraiswamy, P.C., McMurtrey, J.E., Daughtry, C.S.T., Perry, E.M., & Akhmedov, B., 2012. A visible band index for remote sensing leaf chlorophyll content at the Canopy scale. *International Journal of Applied Earth Observation and Geoinformation*, *21*(1), 103-112. https://doi.org/10.1016/j.jag.2012.07.020

Motohka, T., Nasahara, K.N., Oguma, H., & Tsuchida, S., 2010. Applicability of Green-Red Vegetation Index for remote sensing of vegetation phenology. *Remote Sensing*, *2*(10), 2369-2387. https://doi.org/10.3390/rs2102369

Nahry, A.H.E., Ali, R R., & Baroudy, A.A.E., 2011. An approach for precision farming under pivot irrigation system using remote sensing and GIS techniques. *Agricultural Water Management*, *98*(4), 517-531. https://doi.org/10.1016/j.agwat.2010.09.012

Rouse, J.W., Hass, R.H., Schell, J.A., & Deering, D.W., 1973. Monitoring vegetation systems in the great plains with ERTS. *Third Earth Resources Technology Satellite (ERTS) Symposium*, *1*, 309-317. https://doi.org/citeulike-article-id:12009708

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127-150. https://doi.org/10.1016/0034-4257(79)90013-0

Zhang, L., Clarke, M.L., Steven, M.D., & Jaggard, K.W., 2011. Spatial patterns of wilting in sugar beet as an indicator for precision irrigation. *Precision Agriculture*, *12*(2), 296-316. https://doi.org/10.1007/s11119-010-9177-2

Identifying yield stability and drivers of yield variability in cotton using multi-layered, whole-farm datasets

P. Filippi*, T.F.A. Bishop and B.M. Whelan
Precision Agriculture Laboratory, Sydney Institute of Agriculture, The University of Sydney, New South Wales, Australia; patrick.filippi@sydney.edu.au

Abstract

Particular areas within a cotton field often yield consistently high or low, but other areas fluctuate from season to season. Farmers have an idea of the cause of yield variability in cotton, but this is generally qualitatively assessed. The amount of spatial and temporal data that is available today is a promising opportunity to quantitatively identify areas that are constrained, and identify possible causes. This study collates yield monitor data from 16 different seasons, and publicly-available spatial data (digital soil maps, climate, etc.) from a large 13,000 ha cotton farm in Australia. The objectives were to (1) assess the spatial-temporal stability of cotton yield, and (2) use machine learning to model cotton lint yield to identify the most important driving factors of variability.

Keywords: cotton, yield variability, modelling

Introduction

Cotton (*Gossypium hirsutum*) crops can be highly variable, which is undesirable for growers. Cotton yield varies according to seasonal conditions, and at a variety of spatial scales, from within-fields to growing-regions. In Australian cotton-growing systems, farmers typically have a store of yield monitor data over many seasons. The temporal stability of this yield data can be analysed to identify areas within fields that are consistently low or high yielding, or are temporally unstable. This has been analysed for a variety of crop types in several studies (Lamb *et al.,* 1997; Blackmore, 2000; Blackmore *et al.,* 2003; Guo *et al.,* 2018). The temporal stability of yield can point to the nature of the constraint – whether it is permanent, or temporary.

As well as yield maps, cotton-growers frequently have auxiliary spatio-temporal data to augment this, such as soil test results, and apparent electrical conductivity (EC_a), and this data is becoming increasingly collected on-farm. Furthermore, publicly-available spatial and temporal environmental data are also becoming more available at finer spatial and temporal resolutions, and at declining costs (Filippi *et al.,* in press). The nature of these publicly-available datasets is diverse, and includes remotely sensed imagery, geophysical and climate data.

Cotton-growers generally have an idea of the cause of yield variability and the potential constraint to yield, but this is generally qualitatively assessed, whereas they need a more systematic approach to understand this. The wealth of on-farm and publicly-available datasets are a promising opportunity to more quantitatively assess and identify areas that are constrained, and detect the possible causes of those yield constraints. Yield is a function of the interaction between management, soil and weather, and many of these on-farm and publicly-available datasets represent these driving factors. There is often disconnect between these different data sources and they are rarely combined. The use of these data sources in isolation is often of limited value, but combining this wealth of on-farm and freely-available data helps to describe the conditions in which cotton crops are grown under, and subsequently allows a better understanding of the reasons for low or a high yield. This could enable farmers to rectify the causal factor of a yield constraint, or vary their inputs according to yield potential at a within-field scale.

DOI 10.3920/978-90-8686-888-9_4, © Wageningen Academic Publishers 2019

This study uses data from a very large cotton farm in Moree (-29.46441 S, 149.84511 E), northern New South Wales (NSW), Australia. This includes several seasons of cotton yield monitor data, and other spatial data that is freely-available, such as climate information, gamma radiometrics and soil maps. This study has two broad objectives. The first is to assess temporal yield stability of cotton by identifying areas within fields that are consistently low or high yielding, or are temporally unstable. This points to the nature of the constraint – whether it is a permanent constraint, or a temporary constraint. The second objective is to use machine learning to model cotton yield with the combined spatial and temporal data, with an aim to identify the most important driving factors of cotton yield variability.

Methods

Study area

The study area is ~13,000 ha in size, and flood irrigated cotton is the primary crop grown, with regular rotations of dryland wheat, barley, chickpea and sorghum. The average annual precipitation for the study area is 585 mm, with most of this received in the summer, which is the growing-season for cotton. The winter temperatures in Moree are cool to mild, and the summers are very hot. The farm is situated on an alluvial floodplain, and Grey and Brown Vertosols are the most common soil type (Isbell, 1996). Paleochannels are common, resulting in a high degree of spatial variability across the farm.

Datasets

A variety of spatial and temporal data collected on-farm, and publicly-available environmental and agricultural data for the whole study area were collated. The data were of varying spatial and temporal resolutions, and consisted of yield monitor data (collected on-farm), digital soil maps, gamma radiometric grids and climate data (publicly-available) (Table 1).

Cotton yield monitor data was available for multiple seasons, including 2002 to 2017. In total, data from 68 fields were used, with some fields having up to nine yield maps available during this period, whereas some fields had only one (Figure 1). The median number of yield maps for a field was five, with more than 300 individual field-year cotton lint yield maps. The yield monitor data was corrected using the total yields measured as a sum of total bales of cotton harvested per field. The yield data was then processed following the method of Taylor *et al.* (2007) to a 10 m resolution. This involved removing yield values below 0.5 b/ha and those above 20 b/ha, and then removing data points that are more than 2.5 standard deviations above and below the field mean.

Table 1. The data used, the spatial resolution, and the source.

Data type	Data description	Resolution	Source
Yield	Yield monitor data – cotton	10 m	On-farm
Gamma radiometrics	Potassium (K)	90 m	National (GADDS)
	Total dose (K, U, Th)	90 m	National (GADDS)
Digital soil maps	Clay content (0-50 mm)	90 m	National (SLGA)
	Clay content (0.30-0.60 m)	90 m	National (SLGA)
	Available water capacity (0-50 mm)	90 m	National (SLGA)
	Available water capacity (0.30-0.60 m)	90 m	National (SLGA)
Received rainfall	Total rainfall September-March	5,000 m	National (SILO)
Terrain attributes	DEM	30 m	National (SRTM)
	Slope	30 m	National (SRTM)

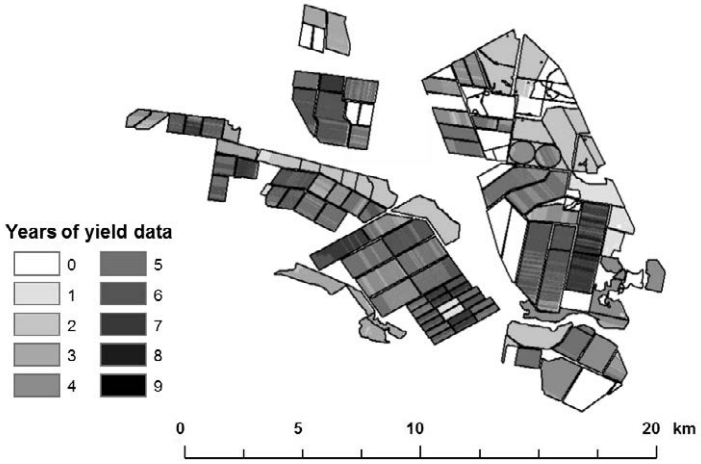


Figure 1. The number of seasons of cotton yield monitor data available for the study area.

Gamma radiometrics data for the study area at a spatial resolution of 90 m was obtained from the Geophysical Archive Data Delivery System – Geoscience Australia. Low-pass filtering was used on all radiometric products (Minty *et al.,* 2009). The digital elevation model (DEM) at 30 m resolution was obtained from the National Elevation Data Framework – Geoscience Australia. Digital soil maps of clay content and available water capacity (AWC) at 90 m resolution from the Soil Landscape Grid of Australia (SLGA) were also collated for this study (Grundy *et al.,* 2015). Total daily rainfall (mm) grids at 5,000 m resolution for the study area were obtained from SILO (Scientific Information for Land Owners). This was then aggregated into total rainfall received from September through to February to be used as predictor variables in the model. These months were aggregated because cotton in Australia is typically planted in October and harvested in May, and pre-sowing rainfall, and rainfall throughout the earlier stages of growth are important drivers of cotton lint yield.

Yield stability

Rather than using the raw yield values in bales per hectare, the z-score of the yield data was calculated for each grid point relative to all other points within that field, for each individual season. This standardises the yield data, and allows yields of different magnitudes to be compared between different seasons and fields, as it is all relative to the mean. The standard deviation of the z-scores from all years for a point was then calculated. In a similar process to that used by Blackmore (2000), these two criteria were then used to divide the points into three stability classes; low-yielding and stable, high-yielding and stable, and unstable. Table 2 describes the criteria used for classification for each location, which consisted of a cut-off value of 0 for the mean z-score, and a standard deviation of 1.2 from all z-scores. This standard deviation value is arbitrary, as there are no specific guidelines for this due to fluctuations in spatial and temporal variability of yield. Other studies have selected arbitrary thresholds for standard deviation, such as 200 kg/ha by Guo (2018) for cotton, and 800 kg/ha and 1000 kg/ha by Blackmore *et al.* (2003) for grains.

Space-time cube and modelling approach

All of this processed spatial and temporal data was then collated into a space-time cube (STC). A STC can be described as a large dataset where each row in the dataset possesses; spatial co-ordinates (x + y), time (season), yield and a large string of associated variables that relate to yield (soil data, weather, etc.). Each row will represent a spatial entity for a particular location, for a particular year

Table 2. Criteria used to classify areas into yield zones.

Yield zone	Mean Z-score	Standard deviation of all Z-scores
Low-yielding, stable	<0	<1.2
High-yielding, stable	>0	<1.2
Unstable	-	>1.2

or time point. Some spatial locations may have up to 9 years of data, where others may only have 1 (Figure 2). All data was extracted on a 15 m spatial grid for the study area.
Random forests (Breiman, 2001) were used in conjunction with the STC to create predictive models of crop yield. Data from all seasons and all fields were used in one model. Because we used data from 16 different years (2002-2017), and 68 different fields, a wide variety of seasonal conditions and soils/landscapes were represented in the model. While some studies aim to forecast crop yield with this wealth of yield and contextual data (e.g. Filippi *et al.,* in press), the aim of this modelling approach is to identify the most important variables that drive yield in space, and time.
To test the importance of different variables in the model, the mean decrease in accuracy was calculated. This shows the amount the random forest model prediction accuracy would decrease if that particular variable is excluded. The more important predictor variables possess a larger mean decrease in accuracy value. Lin's concordance correlation coefficient (LCCC) was used to test the model quality. All of the data processing and analysis was performed in the statistical software program R (R Core Team, 2017).

Results and discussion

Yield stability
Figure 2 shows the spatial and temporal cotton yield variability maps at a 15×15 m grid for the whole 13,000 ha study area. It is clear that some fields displayed higher levels of spatio-temporal yield variability than others. This figure shows that a large portion of the consistently low-yielding,

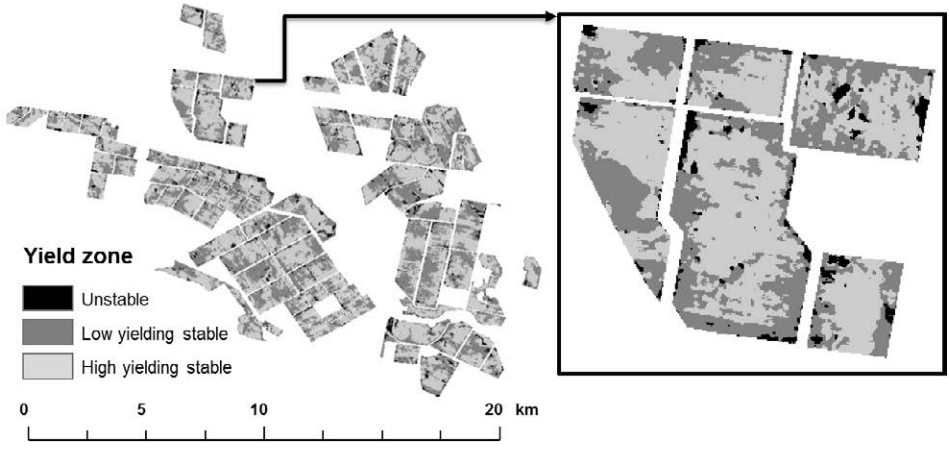


Figure 2. The yield stability zones for the whole study area, and a close-up of a subset of fields.

or unstable yield zones were located on edges of fields. This pattern has also been observed in other studies that have assessed spatio-temporal yield variability in cotton lint yield (e.g. Guo *et al.,* 2018). This is often due to exposure to weather and wind, a discrepancy in inputs (e.g. irrigation and fertiliser), or being located at the head or tail ditch. There were also some unstable yield zones located further away from field boundaries (Figure 2), which could be caused by a temporary constraint, such as disease. This could also be caused by a soil characteristic that responds contrarily in different seasons, such as a clayey soil yielding relatively low in a wet year, or a clayey soil yielding relatively high in a dry year. Areas that experienced consistently low yields often possess more permanent constraints, such as high levels of soil sodicity, which is a common problem for cotton-growing soils in Australia (Filippi *et al.,* 2018).

The relationship between temporal standard deviation and the temporal mean of the yield for each field is presented in Figure 3. Due to the large number of data points, it is difficult to see the quantities within each stability class; however, this can be seen in Figure 2. Our approach of using z-scores instead of raw data has its advantages, however, it is then difficult to compare the spatial variability of fields. The standard deviation of 1.2 from the z-scores was an arbitrary criteria used in the yield zone classification, and the optimisation of this will be the subject of future work. An overall less crisp approach to classification of yield zones will be explored in the future, such as fuzzy clustering. While most studies assess this spatio-temporal yield stability using a shorter time-series of data (e.g. Blackmore, 2000; Guo, 2018), this study used yield data from 16 different growing seasons, with up to 9 seasons of data available for some fields. This results in more representative yield zones, and is advantageous, as seasonal conditions considerably affect yield.

Yield modelling and drivers of cotton lint yield

The out-of-bag predictions for the model showed an LCCC of 0.96, suggesting that the variables included in the model represented spatial and temporal variability well. In the random forest model, the most important predictor variable was rainfall (Figure 4). This indicates the importance of seasonal variability in driving cotton yield. Predictor variables that represent more stable spatial variability, such as with DEM and gamma radiometrics K (potassium) were also deemed important. While the clay content and AWC digital soil maps from SLGA were less important predictors, these layers are correlated with each other, and thus the model would downplay their importance. It is anticipated that the importance of these would increase if some were removed.

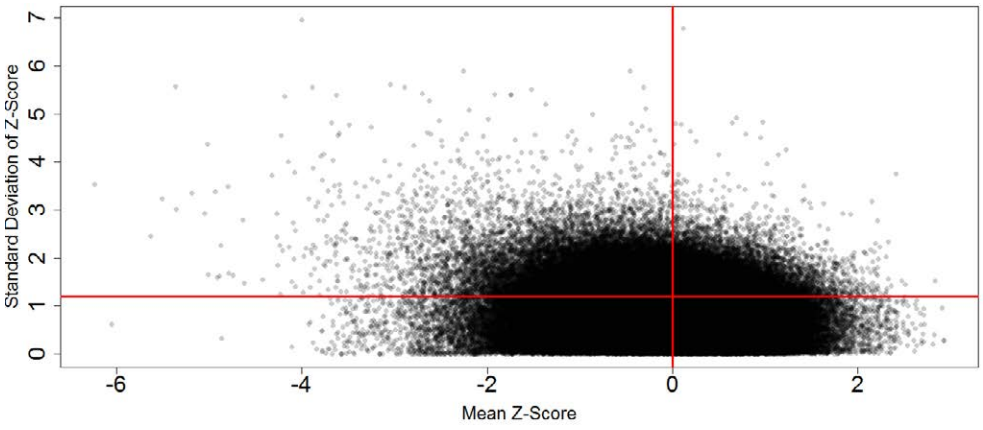


Figure 3. The mean z-score of each pixel plotted against the standard deviation of the z-scores. The lines show the separation between the different yield stability classes.

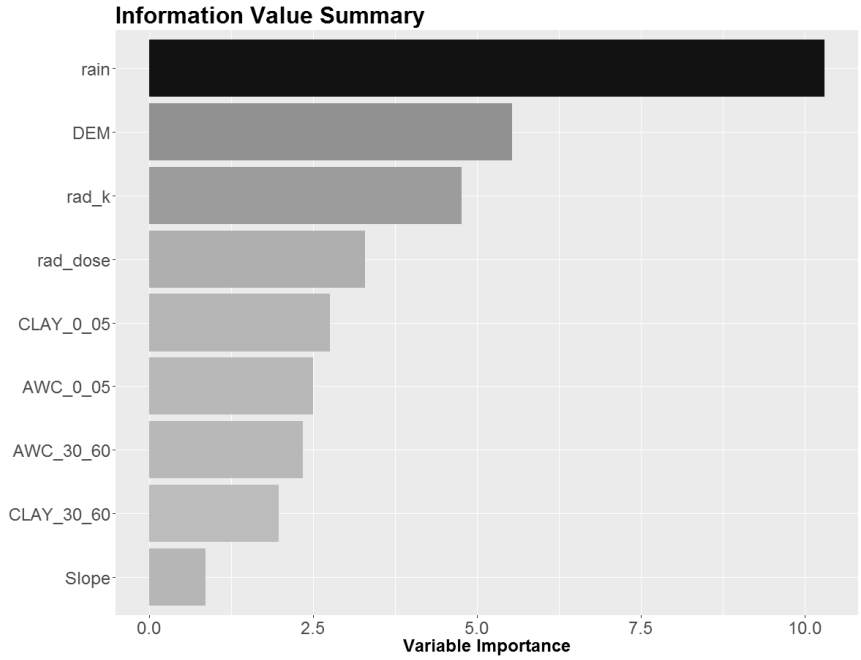


Figure 4. Variable importance plot for the yield model.

There are many more variables that could be included in the models to explain spatial and temporal yield variability. Future research will incorporate soil ECa data from an electromagnetic induction instrument (EMI), as this is available for the whole study area. This fine spatial resolution data (~10 m) and will provide finer-detail soil variability. This data needs to be calibrated/standardised when conjointly analysing this across many fields, as the signal is impacted by moisture content. Some studies have developed approaches for dealing with this challenge (Weller *et al.,* 2007). Rainfall was the only within-season variable included in the modelling approach. There are opportunities to include more, such as imagery from Landsat satellites at 30 m resolution at multiple time points throughout the season (e.g. early, mid, and late). This could give an indication of how early growth development of cotton impacts on final yield. Temperature, in the form of degree-days will also be included in future models. Degree-days are a useful tool in understanding cotton growth stage, and cotton yield is highly affected by temperature patterns (Viator *et al.,* 2005). It is also crucial that important management data is added in to the models, such as sowing date, and the amount and timing of irrigation and fertiliser application. This will give an understanding of the impact that grower's different management decisions have on cotton lint yield.

Additional techniques to identify areas that are constrained will also be explored in the future. One approach is to compare observed and predicted yield. An observed yield that is much lower than a predicted yield suggests that an unknown factor that is hard to include in the model is the cause (e.g. disease). Other approaches will include focussing on using quantile random forests to predict the potential yield (e.g. the upper 5%). This allows yield gaps to be identified, and areas within a field not reaching yield potential. The consistency of not reaching yield potential may point to the type of constraint – whether it is permanent (such as soil sodicity), or temporary (such as disease). There is also an opportunity to forecast cotton lint yield for the coming season using the wealth of spatial and temporal on-farm and publicly-available data collated in this study. Filippi *et al.* (in

press) developed an empirical approach to forecast yield using data from a large grain farm in Australia. This approach uses real, on-farm data and is a promising alternative to forecasting yield using simulation models such as OZCOT (APSIM) (Hearn, 1994), or CROPGRO-Cotton (DSSAT) (Thorp *et al.,* 2014). The current study used cotton yield data from 16 different growing seasons, and this will be advantageous in capturing the seasonal variability to improve yield predictions. These yield forecasts could assist growers in their pre-, and within-season management decisions. Future yield models will include other important spatial variables (e.g. on-farm ECa surveys) and within-season variables (e.g. remote-sensing, temperature). Alternative methods to identify yield gaps/constrained zones will also be explored. All of this will aid in identifying, and better understanding the causes of spatial-temporal yield variability. This further research will assist cotton farmers in easily identifying the spatial-temporal yield patterns of a field, and then using the outputs of the modelling to identify the actual causal factors of variability in space and time. This could lead to improvements in rectifying yield constraints, and input allocation.

Conclusions

This study analysed spatio-temporal stability of cotton yield by standardising yield data using z-scores, with the results showing various degrees of variability at the within-field scale across the 13,000 ha farm. Some fields had up to 9 years of yield data available, resulting in the classification of yield zones that are highly representative of spatio-temporal yield variability. The yield data from all fields and seasons was adjoined with corresponding spatial and temporal datasets, and this was then used to model cotton lint yield. The aim was to identify chief driving factors of yield in space and time, with the models revealing that seasonal variables (rainfall) were highly important predictors, as well as variables that represent spatial variability such as DEM and gamma radiometrics K.

Acknowledgements

Thanks to CRDC for funding, and PCT for providing access to the on-farm data.

References

Blackmore, S., 2000. The interpretation of trends from multiple yield maps. Computers and Electronics in Agriculture 26, 37-51.

Blackmore, S., Godwin, R.J., Fountas, S., 2003. The analysis of spatial and temporal trends in yield map data over six years. Biosystems Engineering 84, 455-466.

Breiman, L., 2001. Random forests. Machine Learning 45, 5-32.

Filippi, P., Cattle, S.R., Bishop T.F.A., Pringle, M.J., Jones, E.J., 2018. Monitoring changes in soil salinity and sodicity to depth, at a decadal scale, in a semiarid irrigated region of Australia. Soil Research 56, 696-711.

Filippi, P., Jones, E.J., Wimalathunge, N.S., Somarathna, P.D.S.N., Pozza, L.E., Ugbaje, S.U., *et al.,* in press. An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. Precision Agriculture. https://doi.org/10.1007/s11119-018-09628-4

Grundy, M.J., Viscarra-Rossel, R.A., Searle R.D., Wilson P.L., Chen, C., Gregory, L.J., 2015. Soil and Landscape Grid of Australia. Soil Research 53, 835-844.

Guo, W., 2018. Spatial and temporal trends of irrigated cotton yield in the Southern High Plains Agronomy 8, 298; doi:10.3390/agronomy8120298

Hearn, A.B., 1994. OZCOT: a simulation model for cotton crop management. Agricultural Systems 44, 257-299.

Isbell, R.F., 1996. The Australian soil classification. Melbourne, Australia: CSIRO Publishing

Lamb, J.A., Dowdy, R.H., Anderson, J.L., Rehm, G.W., 1997. Spatial and Temporal Stability of Corn Grain Yields. Journal of Production Agriculture 10, 410-414.

Minty, B., Franklin, R., Milligan, P., Richardson, M., Wilford, J., 2009. The radiometric map of Australia. Exploration Geophysics 40, 325-333.

R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org

Taylor, J.A., McBratney, A.B., Whelan, B.M., 2007. Establishing Management Classes for Broadacre Agricultural Production. Agronomy Journal 99, 1366-1376.

Thorp, K.R., Barnes, E.M., Hunsaker, D.J., Kimball, B.A., White, J.W., Nazareth, V.J., *et al.,* 2014. Evaluation of CSM-CROPGRO-Cotton for simulating effects of management and climate change on cotton growth and evapotranspiration in an arid environment. Transactions of the ASABE 57, 1627-1642.

Viator, R.P., Nuti, R.C., Edmisten, K.L., Wells, R., 2005. Predicting cotton boll maturation period during degree days and other climatic factors. Agronomy Journal 97, 494-499.

Weller, A., Zipprich, M., Sommer, M., Zu Castell, W., Wehrhan, M., 2007. Mapping clay content across boundaries at the landscape scale with electromagnetic induction. Soil Science Society of America Journal 71, 1740-1747.

Precision spraying by combining a variable rate application map with an on/off map

R. Hørfarter*, M.D. Thorsted, K. Stougård and H.V. Poulsen
SEGES, Agro Food Park 15, 8200 Aarhus N, Denmark; rih@seges.dk

Abstract

Most sprayers used in precision farming are only able to do variable rate application (VRA) of +/- 25% average dose. Consequently, the full benefit of precision spraying is not gained. However, combining the VRA function and the boom section on/off controller could have sprayers apply both none and average dose +/- 25% in the same field operation. How well a section on/off controller map and a VRA map would work in practice was tested. The results showed that it is possible for the terminal to read two different maps at the same time and there was a good correlation between the VRA map and the as-applied map logged by the sprayer.

Keywords: variable application, boom section controller, as-applied maps

Introduction

Precision agriculture (PA) can potentially introduce positive economic and environmental impacts by e.g. increasing yields while reducing lodging, nitrogen leaching and use of fertilizers, pesticides and growth regulators (Cassman, 1999). Studies have shown significant increase in economic savings as the resolution of spraying operations is increased from uniform spraying without swath control to uniform spraying with swath control and further to variable rate application (VRA)(Balafoutis *et al.,* 2017, Esau *et al.,* 2016, Tona *et al.,* 2018). However, most sprayers currently used in PA in Denmark are equipped with fixed orifice nozzles, which can only vary the application rate relatively little – commonly +/ 25% of the average rate. This means that it is not possible to apply no treatment in areas where there is no need and consequently, the full benefit from VRA is likely not gained.
One way to alleviate this problem could be to activate the boom section on/off control concurrently with the VRA function, so that the normal application rate of +/- 25% average rate is combined with a zero-application rate. To do this, a section control on/off map would have to be read simultaneously with – or as a combined section on/off and VRA map – by the controlling terminal. The section on/off control map would thus specify areas with no need for treatment, while the VRA map would specify areas in the field with treatment needs above zero.
During VRA spraying field operation, a difference between actual and target application rate may result for several reasons (Mangus *et al.,* 2017, Sharda *et al.,* 2011, 2013). One such source of error is the uniform rate of application along the length of the boom on most sprayers, which inevitable will result in application rates, read in the pixel at the current position of terminal, being applied to neighbouring pixels of the VRA map also. This effect will be exacerbated if resolution (pixel-size) of the VRA map is increased (smaller pixels) and the resolution of the treatment application is decreased (increased length of the boom). The magnitude of this effect on the accuracy of VRA operations is to the authors' knowledge largely unidentified. Comparison between the target application rate, read in the VRA maps, and sprayer logged position and application rate data, as-applied data, could provide some important information on this issue.
The first aim of this study was to test if it was possible to have a tractor terminal read a VRA map and a section control on/off map, or a combined map, and subsequently control a sprayer on this basis. The second aim was to investigate the correlation between actual applied rates, logged by the terminal on-the-go, and the target rates in the VRA map.

DOI 10.3920/978-90-8686-888-9_5, © Wageningen Academic Publishers 2019

Materials and methods

A total of four tests were carried out on three fields situated in Jutland, Denmark, on two different farms. All machinery and terminals used were owned by the farmers. The size of the fields was 34 ha (Field 1), 27 ha (Field 2) and 9 ha (Field 3). The treatment was variable rate application of growth regulation in winter oilseed rape carried out in September 2018.

Machinery
The sprayers used was an MGM trailer sprayer (MGM a.s., Tovární, Czech Republic) with 28 m boom and with a 3,000 Euro upgrade for precision farming and a self-propelled Amazone Pantera 4502 (Amazone, Hasbergen-Gaste, Germany) with a 36 m boom (Figure 1).

Maps
VRA maps were made using CropSAT web service (SEGES, Landbrug og Fødevarer, Skejby, Danmark) based on the actual biomass on the fields measured as NDVI (Figure 2, left). The treatment applied was growth regulator and accordingly the principles of assigning application rates: high biomass was assigned a high dose of growth regulation and low biomass assigned a low dose. The pixel size of the VRA maps was 20×20 m. From CropSAT the VRA maps were exported in the shapefile format and the projection was WGS84.
The boom section on/off control maps, steering the individual boom section of the sprayer, were made in the GIS program MapInfo Pro, Version 16.0.2 (Pitney Bowes Limited, Stamford, Connecticut, USA). The maps consisted only of a field polygon with holes in it (see Figure 2, right). The holes are



Figure 1. Left: the MGM trailer sprayer. Right: the Amazone Pantera 4502.

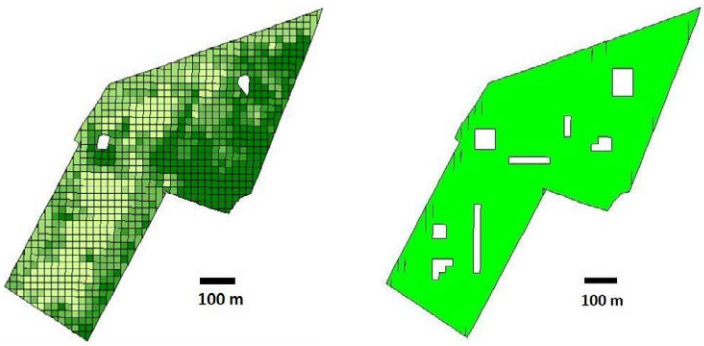


Figure 2. Left: the VRA map for growth regulation in Field 1 (39 ha) based on the vegetation index NDVI from Sentinel-2. Right: the boom section on/off control map for the same field.

theoretical areas with no need for application of growth regulation. When a boom section enters a hole in the map, the boom section automatically closes for the supply of liquid until it enters the solid map again, treating the holes as areas outside the boundaries of the field. In this test, the holes/areas of no treatment were relatively small and of different shapes to test if the sprayers could work correctly around such areas.
Combined maps, merged as one, of the boom section on/off control maps and VRA maps were produced in MapInfo. The areas of no treatment/holes in the section on/off maps were cut from the VRA maps by an Erase outside Target procedure (Dana, 2013) in MapInfo program removing portion of map outside specified target areas with treatment need above zero.
As-applied maps were produced in MapInfo from data logged by the terminal during a field operation. The terminal read the amount of liquid to be applied in its current position in the VRA map and logged the applied amount of treatment and the position and stored this data in a log file.

Tractor terminals
Both farmers had a Topcon (TopCon Positioning Systems, Inc., Tokyo, Japan) terminal, a model X30 and a X35 respectively, in their tractors and the tests were carried out using these terminals.

Loading maps in terminal
The VRA maps and the section on/off control maps were loaded in two different ways: the VRA maps were loaded according to the manual (Topcon, 2012), whereas the on/off maps had to be loaded as field boundary, making the terminal treat the white areas of no treatment need as outside the field boundary (Figure 2 right).

Analysis of compliance between VRA map vs as-applied map
Based on position and application rate data logged by the terminal during the field operation, as-applied maps were produced. In Figure 3, overlay of VRA map pixel-midpoints, the centre of each pixel in the VRA map, on as-applied maps is illustrated. The application rate logged in the positions

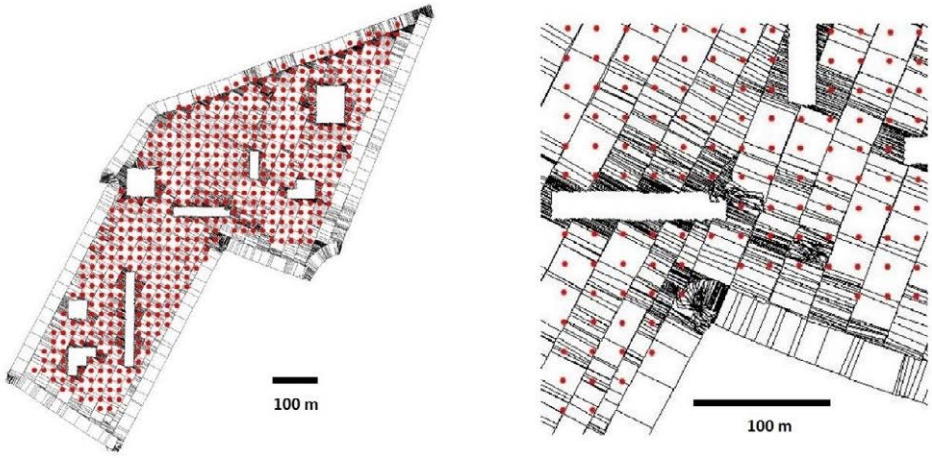


Figure 3. Left: overlay of VRA map pixel-midpoints (dots) and as-applied map from Field 1 (39 ha). Right: a zoom on a sub-section of same field. Serial rows of narrow rectangles show the sprayers path. The width of rows shows the actual active width of the boom along the sprayers path, and the length of the rectangles show the area along the sprayers path with unchanged application rate.

corresponding to the pixel-midpoints was extracted and plotted against the target application rate read from the VRA map. The level of compliance between as-applied rate data and the target rate data from the VRA map is inferred as a measure/indicator of application accuracy. Data from around the headland of the fields were omitted from these analyses, to focus within fields and around the introduced areas of no treatment-need.

Results

Loading maps

Concurrent loading of the two individual maps, VRA and section on/off control, to the tractor terminals was successful for all three tested fields. Visual inspection of the nozzles during the field operations identified that the section on/off control did shut down the treatment application in the areas specified by the section on/off map as areas of no treatment need/outside the boundaries of the field. In Figure 4, a picture acquired by drone 3 weeks after growth regulation treatment in Field 2 (27 ha) is shown. Several of the no treatment zones are evident as areas with brighter colour originating in a denser biomass.

Likewise, loading of the single combined map, VRA and section on/off control maps in one, to the terminals was seemingly successful. Yet, the visual inspection of the nozzles during the field operation revealed that in this case the terminals failed to recognize the areas of no treatment need/ outside the boundaries of the field.

VRA maps vs as-applied data

In Figure 5 is shown an example (Field 1) of an as-applied map based on data logged by the terminal during field operation and, for comparison, the corresponding VRA map. The black areas in the as-applied map concentrated around the edges of the introduced areas with no treatment need is a result of a high level of information logged in small areas, as the boom sections are instructed to open and close, thus the black areas are an artefact and not representative of actual application rates.



Figure 4. Left: the boom section on/off control map loaded to the terminal as a field boundary map for Field 2. Right: a drone photo of part of the corresponding field acquired three weeks after the growth regulation treatment.

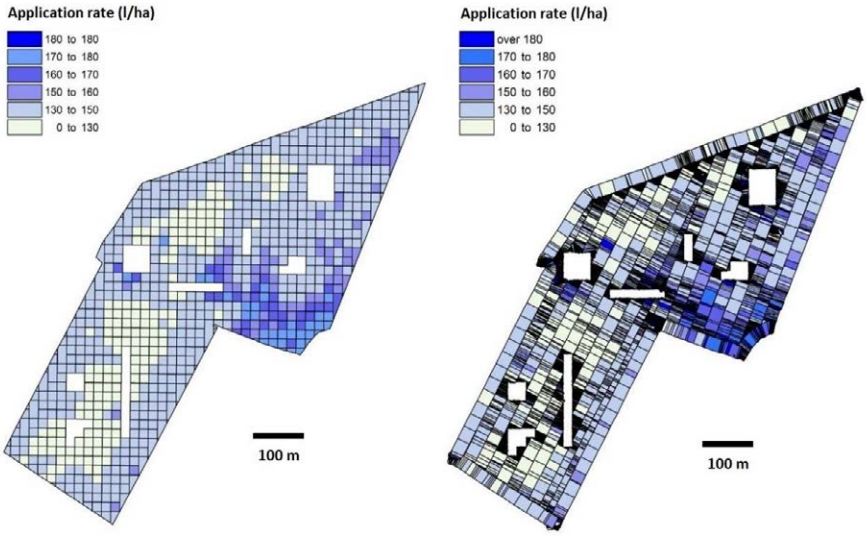


Figure 5. Left: the VRA map from Field 1 including areas from the on/off map with no treatment need specifying the target application rate. Right: the as-applied map-based data logged from the terminal controlling the sprayer.

Disregarding the black areas of the as-applied map, the color scale is even in the two maps, making for immediate comparisons between the target and the assigned application rates.
Plotting the as-applied rates recorded in mid-point of VRA map pixels against the corresponding VRA map rates revealed for all fields a high degree of compliance. In Figure 6 is shown a plot of the 550 rate-pairs from Field 1 – shown in Figure 5 above. The R^2 was found to be ~0.95. From the slope of the regression line, it appears that on average the as-applied rates were around 92% of the target rate specified in the VRA map.

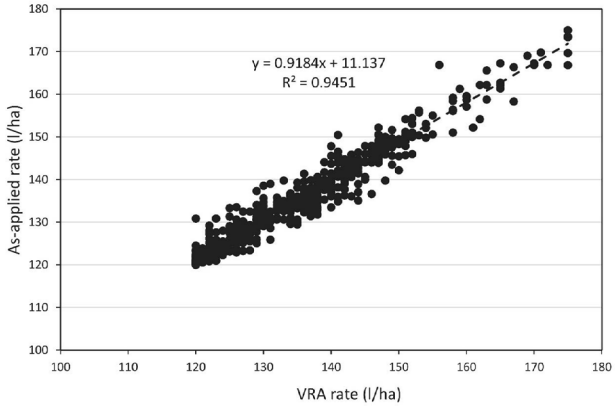


Figure 6. As-applied rates recorded in mid-point of VRA map pixels plotted against the corresponding VRA map rates. Data is from Field 1 sprayed by the MGM trailer sprayer with a 28 m boom. Equation representing the regression line (broken line) inserted.

Discussion

The test showed that it is possible to expand the range of sprayer application rates from the standard +/- 25% average rate to also include a zero-application rate for areas with no treatment need, using the equipment and procedures specified here. This result holds potential for gaining even more benefit from precision spraying. However, for the farmer to produce and load two separate maps is rather impractical and probably not a viable solution in practice. Unfortunately, the less tedious option of producing only one combined map for uploading to the controlling terminal proved not to be feasible – again, this applies to the specific equipment and procedures specified for this test. Further studies are required to test if the procedures for producing, uploading and reading maps used successfully here, applies more widely across different controlling terminals and machinery.
As-applied rates of treatment across a field will in practice, as mentioned in the introduction, deviate from the target rates specified in the VRA map. Plots of as-applied rates as function of target rates (Figure 6) provide an evaluation of the fit between these for a specific field test. The fit between target rates and as-applied rates will be perfect (slope=1) if the target rate is uniform across the VRA map. Consequently, the fit will decrease likely as a function of e.g. the specific variation across the VRA map and the resolution of both treatment application (boom length) and VRA map (pixel size) and therefore, the result inferred from Figure 6 is specific and cannot be extended to cover all sets of corresponding target and as-applied rates. Furthermore, the limit to the application rate of +/- 25% average dose puts a constraint on the potential difference between target and as-applied rates, thus affecting the result inferred from such plots. However, assuming the example of applying growth regulator to Field 1 in the current test is representative of the average spraying VRA situation, the general correlation between the treatment applied in the field and the target rates is relatively high and should not be a subject of concern in PA.
It is worth noting that in the analysis for Field 1, the average as-applied application rate read in the pixel mid-point position was only ~92% of the rate specified in the VRA map. If this discrepancy also applies to the total amount of treatment applied to the fields, basing estimates of total plant protection or fertilizers used on VRA maps may introduce, in this case, an over-estimation.

Conclusions

It was possible for tractor terminals (here Topcon terminals X30 and X35) to read a VRA map and a section on/off control map. This combination of maps opened the possibility for the sprayers to apply both the normal +/- 25% average dose and additionally zero litre per hectare on-the-go. This means that it is possible to get more benefit from precision spraying for the farmer and the environment as well. However, producing and uploading two maps is likely to be a barrier in practice.
A good correlation was found between the as-applied rates, logged during field operation, and the target application rates specified in the VRA map. This result may possibly be extended to most VRA operations, though this should be done cautiously.

Acknowledgements

This work was funded by the Danish Ministry of Environment and Food under the activity 'Partnership on precision spraying'.

References

Balafoutis, A., Beck, B., Fountas, S., Vangeyte, J., van der Wal, T., Soto, I., *et al.,* 2017. Precision Agriculture Technologies Positively Contributing to GHG Emissions Mitigation, Farm Productivity and Economics. Sustainability 9; 1339.

Cassman, K. G., 1999. Ecological intensification of cereal production systems: Yield potential, soil quality, and precision agriculture. Proceedings of the National Academy of Sciences of the United States of America 96: 5952-5959.

Dana, P.H., 2013. MapInfo Professional – user guide. The Geographer's Craft Project, Department of Geography, The University of Colorado at Boulder, USA.

Esau, T., Zaman, Q., Groulx, D., Corscadden, K., Chang, Y., Schumann, A., *et al.,* 2016. Economic analysis for smart sprayer application in wild blueberry fields. Precision Agriculture 17: 753-765.

Mangus, D.L., Sharda, A., Engelhardt, A., Flippo, D., Strasser, R., Luck, J.D., *et al.,* 2017. Analyzing the nozzle spray fan pattern of an agricultural sprayer using pulse width modulation technology to generate an on-ground coverage map. Transaction of the ASABE 60: 315-325.

Sharda, A., Fulton, J. P., McDonald, T. P., & Brodbeck, C. J., 2011. Real-time nozzle flow uniformity when using automatic section control on agricultural sprayers. Computers and Electronics in Agriculture 79: 169-179.

Sharda, A., Luck, J.D., Fulton, J.P., McDonald, T.P., Shearer, S.A., 2013. Field application uniformity and accuracy of two rate control systems with automatic section capabilities on agricultural sprayers. Precision Agriculture 14: 307-322.

Tona, E., Calcante, A., Oberti, A., 2018. The profitability of precision spraying on specialty crops: a technical–economic analysis of protection equipment at increasing technological levels. Precision Agriculture 19: 606-629.

Topcon, 2012. Operators manual. X30 Guidance and Auto Steering. © Topcon Precision Agriculture.

Understanding intra-field variation in N requirement for oilseed rape

S.L. Kendall[1*], K. Storer[2], R. Wade[1] and P.M. Berry[2]
[1]*ADAS, ADAS Gleadthorpe, Meden Vale, Mansfield, Nottinghamshire NG20 9PD, United Kingdom; sarah.kendall@adas.co.uk*
[2]*ADAS, ADAS High Mowthorpe, Duggleby, Malton, North Yorkshire YO17 8BP, United Kingdom*

Abstract

Two OSR chessboard experiments were set up in 2016 to understand factors controlling intra-field variation in N requirement. Each experiment consisted of 200-300 large plots each with one of four rates of fertiliser N applied in a chequerboard pattern. Data were kriged and N response curves were fitted to each plot. Economic optimum N (Nopt) rates were estimated and the 5th to 95th percentile varied by >100 kg N/ha in both experiments. Variation in Nopt was negatively related to soil N supply (SNS) (one site) and N uptake efficiency (both sites), and positively related to yield at Nopt (both sites). Of these factors SNS represents the most tractable approach for predicting intra-field variation in N requirement.

Keywords: oilseed rape, nitrogen fertiliser, nitrogen requirement, yield, variable rate

Introduction

Inorganic nitrogen (N) fertilisers are of major economic importance for oilseed rape production, often doubling crop yields (Berry and Spink, 2009). These yield increases help to meet the food demands of a growing global population, diminish expansion in the arable area and decrease associated negative impacts on the environment. However, fertiliser N can be responsible for over 75% of the greenhouse gas (GHG) emissions associated with crop production (Mahmuti *et al.*, 2009) and is also a significant source of water and air pollution (Misselbrook *et al.*, 2000). Oilseed rape (OSR) has a relatively high demand for N and has one of the lowest N use efficiencies of temperate arable crop types (Sylvester-Bradley and Kindred, 2009). There is therefore a great need to optimise N fertiliser applications to meet the crop's requirement as accurately as possible.
Current recommendations for nitrogen (N) fertiliser rates in the UK are based on a flat rate application for whole fields. For winter wheat, it has been demonstrated that N fertiliser requirement can vary substantially within a field (Kindred *et al.*, 2016). It is likely that N requirement of OSR also varies within fields. Possible evidence for this is observed in spatial variation of lodging within fields (Kendall *et al.*, 2017a). The size of intra-field variation in OSR N requirement needs to be quantified in order to justify developing technology to enable variable N application. Variation in fertiliser N requirement could be a consequence of variation in three main factors: (1) soil nitrogen supply (SNS); (2) crop N demand; or (3) N uptake efficiency. It is important to understand which of these factors are the key determinants of variation in N requirement to enable the development of reliable variable N rate technology. For example, if SNS is shown to be the key determinant then variable N rate technology should focus on sensing this parameter.
The aims of this paper were therefore to 1. Quantify the variation in fertiliser N requirement within a field and 2. Understand factors which cause variation in N requirement.

Materials and methods

Two large OSR chessboard experiments were set up in 2016 in the UK (Walkington, East Yorkshire (Lat 52.699290 Long 0.33821929) and Islington, Norfolk (Lat 53.811213 Long -0.52529949)). Each experiment consisted of 200-300 large plots (e.g. 12×12 m) with one of four rates of fertiliser N

DOI 10.3920/978-90-8686-888-9_6, © Wageningen Academic Publishers 2019

applied as liquid urea ammonium nitrate or solid ammonium nitrate in a chequerboard pattern, thus providing a range of crop N contents and canopy sizes to test (Figure 1). Using commercial application equipment to criss-cross the experimental area, N rate 1 (130 kg N/ha) was applied in one direction to one side of the tramline and N rate 0 (0 kg N/ha) on the other. In the perpendicular direction, N rate 2 (270 kg N/ha) was applied to one side of the tramline and again N rate 0 (0 kg N/ha) to the other. This sets up the chessboard pattern (a gingham pattern) of alternating squares of 0 kg N/ha (N0+N0) and 130 kg N/ha (N1+N0) plus 270 kg N/ha (N0+N2) and 400 kg N/ha (N1+N2) across the applied area (Figure 1). The size of each square depends on tramline width, a common 24 m tramline width giving 12 m squares. N applications were made in either two or three equal doses between late February and late April.

Each plot was combine-harvested directly using a Sampo small plot combine harvester (Sampo Rosenlew Ltd., Pori, Finland) with a vertical cutting bar and a cutting width of 2.3 m. This meant that a 2.3 m swath was taken from each 12-15 m wide plot. Combining for each plot was carried out in the same direction. Harvested plot lengths were measured immediately following harvest. A Dickey John GAC 2000 grain analysis computer (Church Industries, Minneapolis, MN, USA) was used to measure seed moisture content and this information was used to correct yields to 91% dry matter (DM).

Kriging was used in Genstat 16th edition (VSN International, 2013) or R (R Development Core Team (2008)) to interpolate measures at each N rate for each plot, so that for each plot measures were available for each N rate, even though each plot was conducted at only one N rate. Variogram analysis and ordinary kriging were performed within each N-rate to avoid assumptions of linear co-regionalisation. In this case, the finest lag distance for which variograms can be estimated was 24 m.

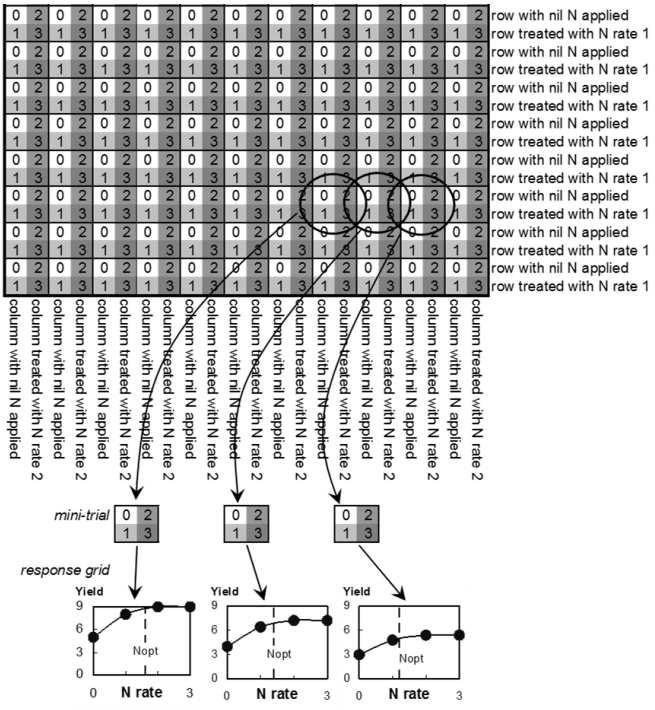


Figure 1. Components of the 'chessboard' tramline grid includes a mini-trial with N levels of 0, 1, 2 and 3, corresponding to 0, 130, 270 and 400 kg N/ha respectively.

Linear plus exponential curves were fitted to the kriged yield values for each plot (Equation 1) for OSR of the form:

$$Y = a + br^N + cN \quad (1)$$

Where Y is the yield (t/ha), and a, b, c and r are fitted constants. The economic N rate (Nopt) was determined from the fitted linear plus exponential parameters as follows:

$$Nopt = [\ln(k / 1000 - c) - \ln(b(\ln r))] / \ln r \quad (2)$$

Where k is the breakeven price ratio between fertiliser N (p/kg) and seed (p/kg). A breakeven ratio of 2.5 was used as this is a standard for fertiliser recommendations in the UK (AHDB, 2017). The yield at the optimum N rate (Yopt) was calculated from the fitted parameters using Equation 1. Total crop N uptake under nil N (a proxy for soil nitrogen supply (SNS)) was estimated from the nil N yields using the average seed N% and nitrogen harvest index (NHI) from the nil N plots to estimate crop N uptake under nil N conditions at the site. Similarly, N uptake at the optimum N rate was estimated using the mean seed N% and NHI data from the 130 and 270 N rates. The difference between the SNS and crop N uptake at the optimum was then used to estimate the fertiliser uptake efficiency at the optimum N rate for each plot. A Pearsons correlation analysis was used to understand relationships between parameters. Statistical analyses were carried out in Genstat 16th edition (VSN International, 2013).

Results

The N treatments caused clearly visible effects on the OSR crops (e.g. Figure 2A). The Nopt and Yopt are presented in Table 1 for the Walkington and Islington sites. At the Walkington site, there was a region in the middle of the trial area which had low yields at zero N fertiliser and this area coincided with a steep slope. The steeply sloping region had a high yield at high N rate (270 kg N/ha) and high Nopt. The estimated SNS ranged from 49-100 kg N/ha and the mean fertiliser recovery was estimated to be 26%, with a range of 15-38%. There was a negative relationship between estimated SNS and Nopt (r=-0.47, $P<0.001$), a positive relationship between Yopt and Nopt (r=0.54, $P<0.001$), no correlation between yield at 270 kg N/ha and Nopt, and a positive relationship between calculated N uptake at the Nopt and Nopt (r=0.54, $P<0.001$). There was also a negative relationship between estimated N uptake efficiency and Nopt (r=-0.73, $P<0.001$).

At the Islington site, the estimated SNS ranged from 39 to 100 kg N/ha and the mean fertiliser recovery was estimated to be 36%, with a range of 13 to 51%. There was no significant relationship between estimated SNS and Nopt, a positive relationship between Yopt and Nopt (r=0.51, $P<0.001$) and yield at 270 kg N/ha and Nopt (r=0.33, $P<0.001$). There was a positive relationship between calculated N uptake at the Nopt and Nopt (r=0.51, $P<0.001$). There was also a negative relationship between estimated N uptake efficiency and Nopt (r=-0.43, $P<0.001$).

Discussion

This study measured a wide range of N requirements within a field of OSR of over 200 kg N/ha in both experiments. Similar ranges of within field N requirement have previously been reported for winter wheat (Kindred *et al.,* 2016) and a wide range of N requirements was measured in grass (Berry *et al.*, 2017). The site at Walkington had a dip towards the top of the field which corresponded with the slightly higher yielding areas under zero N conditions, similarly there was a steep slope leading into the dip, which corresponded with the lower yielding area under nil N conditions. Interestingly, these trends were reversed under 270 kg N/ha, which illustrates that yields under low N conditions are not predictive of yields under high N fertiliser. A possible reason for this in this case was that

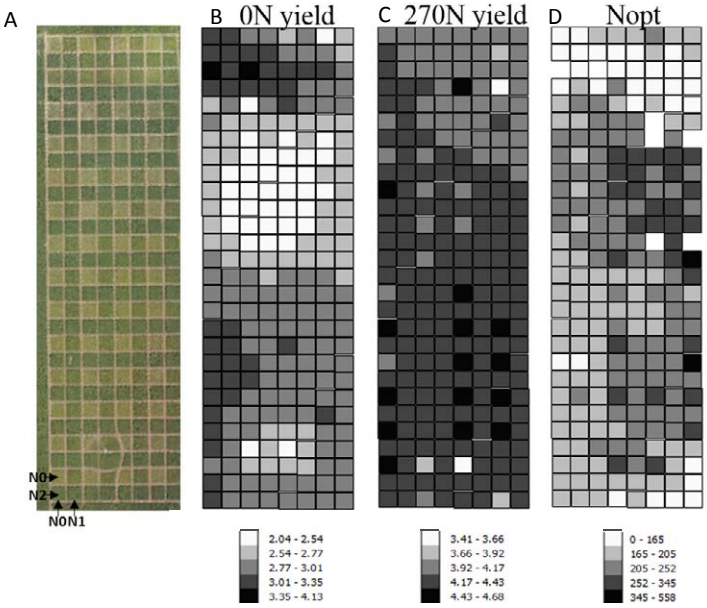


Figure 2. Walkington site. A. Aerial image of the oilseed rape crop including example of N application rates and directions (N0=nil N; N1=130 kg N/ha; N2=270 kg N/ha), B. Kriged yield under nil N conditions for each plot (t/ha at 91% DM), C. Kriged yield under 270 kg N/ha for each plot (t/ha at 91% DM) and D. Optimum N rate (Nopt) calculated using kriged yields at each of the four N rates for each plot (kg N/ha).

Table 1. Nopt and Yopt for the Walkington and Islington sites.

	Walkington	Islington
Average field Nopt (kg N/ha)	221	174
Yopt (t/ha at 91% DM)	4.15	3.81
Nopt range (kg N/ha)	115-370+	107-318
Nopt range (5^{th} to 95^{th} percentile) (kg N/ha)	140-297	121-245
Yopt range (t/ha at 91% DM)	3.6-5.2	3.17-4.37

applying high levels of N fertiliser to the dip at the top of the experimental area resulted in a crop with a super-optimal canopy which reduced yield (Berry *et al.*, 2011).

The SNS explained some of the variation in Nopt at the Walkington site, though the r^2 value was low at 0.22, with low SNS correlating with high Nopt. However, at the Islington site, variation in SNS showed no significant relationship with Nopt. At Islington, it is possible that a weak positive relationship between SNS and Yopt reduced the correlation between SNS and Nopt due to counteracting effects between these factors. The Yopt had a positive relationship with the Nopt at both Walkington and Islington (r^2=0.29 and 0.26 respectively). However, yield at a single high N rate of 270 kg N/ha did not correlate as well (r^2=zero and 0.11 respectively). This suggests that yield maps could be a useful factor for predicting N opt, but only if the Nopt has been used to achieve this yield.

It is likely that the variation in N requirement is driven by a number of factors, and therefore one parameter is unlikely to explain all of the variation. The Auto N project (Kindred *et al.*, 2016) found that different factors, including SNS, yield at Nopt and N uptake efficiency, were related to the Nopt in winter wheat at different sites. Of the factors which correlate with Nopt, SNS is the easiest to measure. This is particularly the case for OSR because often the majority of the SNS has been taken up into the crop canopy by the time that decisions on N fertiliser rates are taken. A range of methods can be used to estimate the N content of the crop such as using crop sensors, drone imagery and satellite imagery (Kendall *et al.*, 2017b), and these can be used as the basis for creating variable N application maps. Historic yield map data could be used as a proxy measure of crop N demand. However the pattern of yield variation is often inconsistent and may require several years of evidence before it can be used reliably. Additionally, this paper indicated that Yopt is the most important factor and most yield maps describe yield achieved at a single N rate. Fertiliser N uptake efficiency was negatively correlated with Nopt in both experiments and further research is required to understand factors which determine this to predict how it varies spatially within a field.

Conclusions

These chessboard trials demonstrated a large variation in fertiliser N requirement within a field, with a 5th to 95th percentile range of more than 100 kg N/ha at two sites in 2016. Variation in Nopt was negatively related to SNS at one site, negatively related to N uptake efficiency at both sites and positively related to yield at Nopt at both sites. Of these factors, measurement of SNS presents the most tractable approach for predicting intra-field variation in N requirement. Future research should focus on testing this at the field scale.

Acknowledgements

The authors are grateful to Innovate UK for part funding of project 101909 and to the project partners: CF Fertilisers UK Ltd, Hill Court Farm Research, Precise Crop Nutrition, Syngenta UK, Precision LandCare, Patchwork Technology and also to Fritzmeier.

References

AHDB., 2017. Nutrient Management Guide (RB209). Agricultural and Horticultural

Berry P.M. and Spink J.H., 2009. Canopy management and late nitrogen applications to improve yield of oilseed rape. Project Report No. 447, London, UK: HGCA. pp. 213.

Berry P.M., Roques S. and Clarke S., 2011. Predicting the optimum rate and timing of nitrogen fertiliser for winter oilseed rape. In Proceedings of 13th International Rapeseed Congress. Paris, France: International Consultative Research Group on Rapeseed. pp 285 http://gcirc.org/fileadmin/documents/Proceedings/IRCPrague2011vol1/crop%20management/285-288.pdf (last accessed March 2019)

Berry, P.M., Holmes, H.H. and Blacker, C., 2017. Development of methods for remotely sensing grass growth to enable precision application of nitrogen fertiliser. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture, Advances in Animal Bioscience 8 (2) 758-763.

Development Board, UK.

Kendall, S.L., Holmes, H.H., White, C.A., Clarke, S.M. and Berry, P.M., 2017a. Quantifying lodging-induced yield losses in oilseed rape. *Field Crops Research* 211 106-113.

Kendall, S.L., Storer, K. and Berry, P.M., 2017b. Measuring canopy size and N content in oilseed rape for variable plant growth regulator and nitrogen fertiliser application. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences 8 299-302.

Kindred, D.R., Hatley, D., Ginsburg, D., Catalayud, A., Storer, K., Wilson, L., *et al.,* 2016. Automating nitrogen fertiliser management for cereals (Auto-N). Cereals and Oilseeds Project No. 561, Stoneleigh, UK; Agricultural and Horticultural Development Board.

Mahmuti, M., West, J.S., Watts, J., Gladders, P. and Fitt, B.D.L., 2009. Controlling crop disease contributes to both food security and climate change mitigation. International Journal of Agricultural Sustainability 7 189-202.

Misselbrook, T.H., Van der Weerden, Y.J., Pain, B.F., Jarvis, S.C., Chambers, B.J., Smith, *et al.,* 2000. Ammonia emission factors for UK agriculture. Atmospheric Environment 34 871-880.

Sylvester-Bradley, R. and Kindred, D.R., 2009. Analysing nitrogen responses of cereals to prioritize routes to the improvement of nitrogen use efficiency. Journal of Experimental Botany 60 1939-1951.

Evaluation of a functional Bayesian method to analyse time series data in precision viticulture

C. Laurent[1,2,3*], M. Baragatti[4], J. Taylor[1], T. Scholasch[3], A. Metay[2] and B. Tisseyre[1]
[1]*ITAP, Université de Montpellier, Montpellier SupAgro, Irstea, France*
[2]*SYSTEM, Université de Montpellier, CIHEAM-IAMM, CIRAD, INRA, Montpellier SupAgro, France*
[3]*Fruition Sciences, Montpellier, France; cecile@fruitionsciences.com*
[4]*MISTEA, Université de Montpellier, Montpellier SupAgro, INRA, France*

Abstract

In precision agriculture, most studies focus on spatial crop variability whereas temporal variability and its role in decision-making is equally important. The classical methods for temporal analysis have limitations, potentially resulting in information loss. A novel method based on a Bayesian functional Linear regression with Sparse Steps functions (BLiSS method) is evaluated in this paper to investigate continuous influence analysis when working with time series data. The example of the influence of temperature on the number of clusters per vine during the year before harvest was considered as an example application. The evaluation of the BLiSS results was done by comparing identified critical time periods with traditional viticulture knowledge in the literature. It showed the relevance of the BLiSS method, highlighting already known results and identifying new critical time periods for yield elaboration.

Keywords: BLiSS, temporal analysis, temperature, yield, number of clusters

Introduction

In precision agriculture, most studies have focused on spatial crop variability (Oliver *et al.*, 2013) whereas temporal variability and its role in decision-making is equally important. Indeed, most indicators of production yield and quality are time-dependant. Understanding their evolution during the crop growth cycle and the factors that drive them is often an issue for better production management. This implies understanding the pattern of influencing factors over a whole production period, as well as the correlations between influences.

Historically, 'temporal' studies have focused on using climate variables at a few known key phenological stages or time steps. This has been done to simplify the analysis and permit the use of classical methods, such as least squares linear regression. However, these classical approaches have limitations: (1) they depend on choices of climate variables and timing, (2) it is often necessary to suppress data or to analyse only parts of a time series, and (3) times series temporality (observations correlated over time) is not considered. Therefore, information about temporal crop physiological regulation is potentially missed.

To advance temporal analysis in the crop production domain, this paper evaluates the potentialities of a new approach to study quantitative time series data. It uses a Bayesian functional Linear regression with Sparse Steps functions (BLiSS method) (Grollemund *et al.*, 2019). The principle of this method is to analyse the complete history of an explanatory variable to identify the periods during which it has an impact on the explained variable. It takes into account the data correlation over time i.e. it allows the analysis of a period's impact according to the impact of others periods. By overcoming the classical approaches' limitations, the BLiSS method could result in the discovery of new periods of influence of an explanatory variable on an explained variable in many agricultural sectors.

The aim of this paper is therefore to test the potentialities of the BLiSS method on a study case that is well understood in the literature. To this end, the impact of temperature history in the year

DOI 10.3920/978-90-8686-888-9_7, © Wageningen Academic Publishers 2019

before harvest on the number of clusters per vine has been chosen as a study case. Indeed, grape yield elaboration is an interesting case study in temporal terms as grapevine is a perennial crop, meaning its yield determining process covers two growth cycles (Howell, 2001; Vasconcelos *et al.*, 2009). These two years are noted *n-1* and *n*, the year *n* being the year of harvest (Figure 1). Thus, at least two growth cycles overlap and there is an interdependency between the different physiological processes leading to memory effects in vine physiology (Ravaz and Sicard, 1911; Keller *et al.*, 2010).

The number of clusters per vine (*CN*) is a major yield component, said to explain up to 60% of the final yield variability (Clingeleffer, 2001). However, understanding the timing of its elaboration and its sensitivity to external factors such as temperature is complicated by its interdependency with other physiological processes over time. Therefore, there is a challenge to determine all the periods during which any explanatory variable, such as temperature, has an impact on *CN* taking into account that the effects of these periods are correlated to many others periods' multivariable impacts. To do this, classical approaches have been carried out using key phenological stages (Buttrose, 1974; Pouget, 1981) or time steps (Guilpart *et al.*, 2014; Molitor and Keller, 2017). They have shown that temperature is one of the major factors driving grape fruitfulness (Buttrose, 1974; Srinivasan and Mullins, 1981; Petrie and Clingeleffer, 2005). In the literature, periods of phenological development during year *n-1* for one cycle are referred to by their concomitance with the phenological stages of year *n*. For example, the development of inflorescence primordia linked to the yield in year *n* starts with budbreak in year *n-1* (Vasconcelos *et al.*, 2009). Following this notation, a few hours of elevated temperature around budbreak of year *n-1* should be sufficient to positively impact inflorescences' or clusters' number per vine in year *n* (Buttrose, 1974; Srinivasan and Mullins, 1981; Pouget, 1981; Lobell *et al.*, 2006). Critical stages have also been highlighted around bloom (Durquety *et al.,* 1982; Molitor and Keller, 2017) of year *n-1*. Molitor and Keller (2017) found that high maximum temperatures around veraison and low average temperatures during maturation of year *n-1* favoured high yield. As a study case, this paper aims to assess whether using the BLiSS method to analyse the impact of the whole history of temperature in year *n-1* on *CN* permits the identification of critical periods previously described and well understood within literature. It also tests the potential detection of previously unknown critical periods.

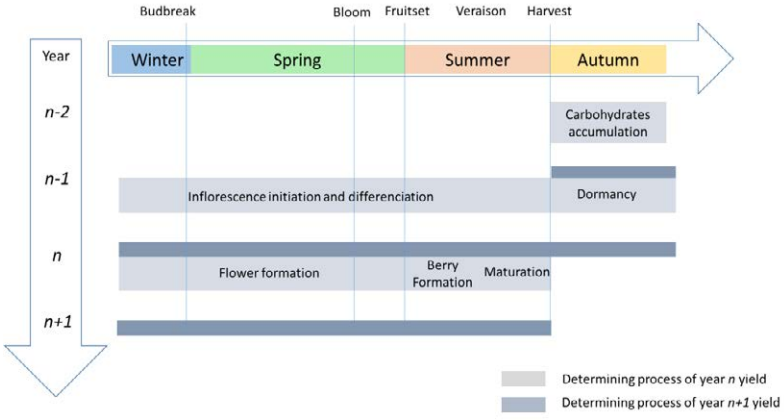


Figure 1. Key phenological stages and the temporal pattern of the grapevine yield determining process over two seasons (aligned for the northern hemisphere).

Materials and methods

Data description

Clusters number per vine data
Data were collected from a commercial Cabernet-Sauvignon (*Vitis vinifera* cv. Cabernet Sauvignon) vineyard field (~1 ha) in the Bordeaux region, France. The vines are short-pruned, trellised, non-irrigated, partially grass covered in the inter-row and planted at a density of 8,696 vines/ha. The *CN* was determined after fruitset on the same 100 vines in the vineyard from 2007 to 2018, with the exception of 2014. Note that depending on the seasonal conditions, cluster thinning is often performed in this vineyard, such that the final harvest differs from the amount of fruit set. The occurrence date of the principal phenological stages (budbreak, bloom, fruitset, veraison) was recorded by the vineyard manager according to the Gregorian calendar from 2006 to 2018.

Temperature data
Temperature data were collected by a local weather station (provided by DE.MET.E.R, Villenave d'Ornon, France) located in the wine estate from 2006 to 2018. Multiple measurements were observed, however, only daily minimum (*Tmin*) and average (*Tavg*) temperature (°C) are considered in this analysis as their impact on grape yield is well documented in the literature.
The *Tmin* and *Tavg* influences on *CN* have only been studied during the growing season of year *n-1* (from 1st of March to 25th of October). The potential impact of winter or the growing season of year *n* was not investigated in this preliminary study, but it could be included following the same steps. To reduce processing time, temperature was aggregated into 10-day periods. The *Tmin* therefore corresponded to the minimum daily temperature during each of these periods. The *Tavg* corresponded to the average of daily temperatures over the 10-day periods. It is expected that better software and increased computing power will remove the need for this pre-processing step in the near future.

Analysis by the BLiSS method

Theory
The BLiSS method (Grollemund *et al.*, 2019) allows quantitative data (variable *Y*) to be explained by functional data (variable *X*) using a linear functional regression model.

$$Y = \alpha + \int_0^T \beta(t)X(t)\,dt + \varepsilon \quad (1)$$

Where *Y* is the explained variable, α a scalar, *X* the explanatory variable, $\beta(t)$ the parameter function to be estimated and ε a residual error supposed to follow a normal distribution $N(0, \sigma^2)$. In this paper, *Y* corresponds to *CN* and *X* corresponds to *Tmin* or *Tavg*. BLiSS is a Bayesian method: it supposes that the parameter $\beta(t)$ is random, formulates an *a priori* estimation of the $\beta(t)$ distribution and updates it with the data to produce an *a posteriori* estimation of the $\beta(t)$ distribution. The BLiSS method investigates the estimated distribution support, i.e. all the *t* instants for which $\beta(t)$ differs from 0. The BLiSS method thus delivers two estimators: one of the $\beta(t)$ function and another of the $\beta(t)$ function support (here time). The latter allows periods during which *X* has an impact on *Y* to be identified.
The output of BLiSS is best represented in a graphical form (e.g. Figure 2). By way of explanation, Figure 2 features the *a posteriori* distribution of the BLiSS estimator of $\beta(t)$ as a line estimated over time. If the $\beta(t)$ estimator is null (= 0) during a certain period, it means that the variable *X* is not related to *Y*. If it has a positive value during a given period of time, *X* is positively correlated to *Y*. In contrast, if it has a negative value during a period of time, *X* is negatively correlated to *Y*. The

duration of the detected periods corresponds to their time length (distance on the x-axis). The magnitude of the detected periods corresponds to their estimated β(t) value (distance on the y-axis). The higher the magnitude, the greater the impact *X* has on *Y*. The BLiSS method also allows the correlation between *X* and *Y* to be estimated. Practically, it corresponds to the influencing period detection confidence. The colour code is a visual way to represent it. The darker the colour is (red on the extreme case), the more confident the correlation between *X* and *Y* during this period. In contrast, the lighter the colour is (white on the extreme case), the less confident the correlation.

Analysis

Each individual corresponds to a *CN* per year. The dataset was therefore composed of 5,368 individuals to be analysed. The analysis was performed using the package *bliss* version 1.0.0 (Grollemund, 2019) in R 3.5.1. The *Tmin* and *Tavg* time series were analysed independently.

Results interpretation

Assuming that the ecophysiological results reported in the literature and reviewed in the introduction represent the diversity of the classical method applications, the BLiSS method results were evaluated by comparison with the identified critical time periods. The literature review focused on temperature impact on the number of inflorescences or clusters per vine and on production yield during the year *n-1*.

Results

Figure 2A shows the influence of daily minimum temperature during the year *n-1* on *CN*. Four periods of influence were detected: (1) from mid-March to late April, (2) from mid-July to early August, (3) from early August to late September and (4) from late September to mid-October. The detection of each of these periods had an adequate degree of confidence (dark colour). Three out

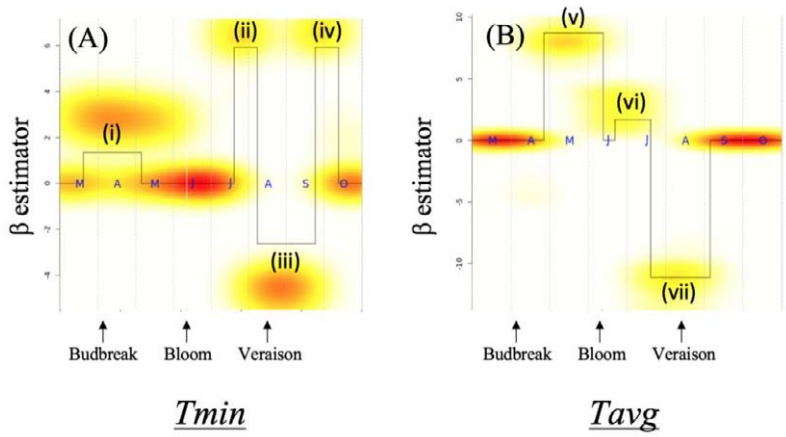


Figure 2. Analysis of possible periods of influence respectively of (A) daily minimal temperature (*Tmin*) and (B) average temperature (*Tavg*) on the number of clusters per vine from 1st of March to 25th of October of year *n-1*. The black line indicates the β estimator distribution. The colour gradient from white to red illustrates the density probability function of the β estimator distribution. The x axis labels *M,A,M,J,J,A,S,O* represent March, April, May, June, July, August, September, October respectively. Average dates of budbreak, bloom and veraison recorded on the commercial field from 2006 to 2018 vintages are indicated.

of these four periods (1,2 and 4) were positively correlated to the *CN* i.e. elevated *Tmin* during these periods favoured a high *CN*. In contrast, high *Tmin* during the (3) period favoured a lower *CN*. The duration of periods (1) and (3) exceeded one month whereas the period (2) and (4) lasted between two and three weeks. The absolute magnitude for each period gave an indication of the intensity of the temperature impact on *CN*. It was 6 for periods (2) and (4) and respectively 1.5 and 3 for periods (1) and (3).
Figure 2B shows the influence of daily average temperature on *CN* during year *n-1*. Three periods of influence were detected: (5) from late April to early June, (6) from late June to late July and (7) from late July to early September. The detection of each period had a lesser degree of confidence (lighter yellow) than periods (1) and (3) (Figure 1A). Two periods, (5) and (6), were positively correlated to *CN*, meaning high *Tavg* during these periods favoured a higher *CN*. In contrast, high *Tavg* during period (7) reduced *CN*. The duration of these three periods were between one month and a month and a half. The absolute magnitude of periods (5), (6) and (6) were 9, 2 and 11 respectively.

The BLiSS results are in agreement with literature
Period (1) (Figure 2A) was associated with the budbreak period in the Bordeaux region. It is in agreement with Pouget's (1981) results on the number of inflorescences per vine under experimental conditions and the findings of Lobbell *et al.* (2006) when studying night temperature around budbreak in California. Period (5) corresponded to the bloom period in the Bordeaux region. This agrees with the results of Durquety *et al.* (1982) on *CN* for Petit Manseng and the Molitor and Keller (2017) study on the impact of minimal temperature on production yield for Müller-Thurgau and Riesling in Alsace (North-East of France). In the latter case, the correlation between minimum temperature around bloom and yield was tested and found positive, as observed in period (5) (Figure 2B). Period (2) coincided with veraison in the Bordeaux region. Only Molitor and Keller (2017) reported this period studying daily *Tmax* effects on yield for Müller-Thurgau. The (3) and (7) periods are associated with the maturation period in the Bordeaux region. Molitor and Keller (2017) also reported a negative impact of minimum temperature on future yield potential during this period.

The BLiSS results allow further exploration of the data
Periods (4) and (6) have not been previously cited in the literature but they had a good degree of certainty according to the BLiSS method. They respectively correspond to a post-harvest and a fruitset period. Since the method correctly identified known periods of influence (previous section), the hypothesis is that these periods are also influential on *CN*. This shows the potential ability of BLiSS to highlight unknown periods of influence of temperature on *CN*. Unlike the classical method results, for which the duration of the resulting periods is determined by the pre-analysis choices, the periods detected by the BLiSS method were characterized by different durations. It is also interesting to note that the temperature influence during summer was characterized by several periods in Figure 2, whereas Molitor and Keller (2017) only reported one period of long duration. The BLiSS method had also refined the duration of known key influencing periods. The proposed alterations in duration and the identification of new periods of influence now need to be properly validated in a viticulture context. If true, these results have increased the knowledge and the temporal resolution of environmental effects on yield and consequent potential management.

Discussion

The BLiSS method: a novel and relevant method to analyse time series data
The BLiSS method allowed the detection of already confirmed periods of daily minimal and average temperature on *CN* and potential vineyard yield. This point validated the use of this method to analyse time series of agricultural data. Moreover, the method allowed the detection of previously unknown periods of influence or the refinement of the duration of known key periods. This demonstrated

the advantages of a linear functional model in a Bayesian framework for the temporal analysis of agricultural data.

Comment on the interpretation of the BLiSS results

Each period that was detected with the BLiSS method was characterized by a magnitude, a duration and a degree of confidence. There are therefore three criteria that can be used to compare the different periods. However, classifying the periods by decreasing order of influence required an integration of the information linked to both the degree of confidence and the magnitude. This may not be simple in every case. For example, a high magnitude but medium degree of confidence period cannot be prioritized in comparison with a high degree of confidence but medium magnitude period. Thus, the interpretation of the BLiSS results should be decomposed into the three criteria first and an expert knowledge of the field of study should then balance the interpretation. Clearer interpretation of the BLiSS outputs is an area for further development.

Potentialities for various study field analysis using the BLiSS

The BLiSS method allows further exploration of many types of datasets. Any variable characterized by time series data, which is a common case in agriculture, can be analysed with the BLiSS method. The method could also be applied to other variables evolving over different continuous supports such as wavelength or spatial coordinates.

While two univariate cases were presented here, several explanatory variables could be studied in a multivariable analysis to identify their potential combined effects as well as their individual (univariate) effects. For example, *Tmin* and *Tavg* have simultaneously been considered in a bivariate BLiSS analysis to distinguish their proper effects on *CN* (data not shown).

However, there are some limitations to the current BLiSS method. Firstly, the computation time has to be taken into account when designing the analysis. For example, the results presented in Figure 2 (derived from 5,368 individuals and using 27 10-day climatic periods) took approximately 4 hours on a machine of 4 CPU and 5 Gb memory. Secondly, it is not possible to know the portion of the Y variability explained by X at the moment. It is also not yet possible to include qualitative explanatory variables but this is planned to be implemented soon. In addition, it is important to remember that the BLiSS method is subject to the same limitations that any other linear analysis, such as disturbance due to the explanatory variables correlation.

Finally, the BLiSS method presents another option which has not been tested in this paper. Prior knowledge can be taken into account by the Bayesian-based method. In practice, this can help the method to detect relevant periods using *a priori* literature results or field experience, in the case of a small number of data for example.

Conclusions

The BLiSS method combines the advantages of both functional and Bayesian models to perform advanced temporal variability analysis. Functional analysis allowed a more complete and objective analysis of a data set, taking into account the explanatory variables histories and allowing detection of periods of influence in the time series. Bayesian analysis allowed a better uncertainty management using conditional probability. This approach can be used for further data exploration in many different fields, in agriculture or others. It is an interesting method to better study the effect of high temporal resolution variables, such as those generated in precision agriculture. The ability to correctly identify and manage key phenological stages in agriculture is an area that has often been overlooked in precision agriculture but is needed to improve the temporal-resolution of decision-making on farm.

Acknowledgements

The authors fully acknowledge the commercial estate that provided the data as well as the authors of the BLiSS method who spent time explaining its application.

References

Buttrose M. S., 1974. Climatic Factors and Fruitfulness in Grapevines, Horticultural Abstracts, 44, 319-326

Durquety P.M., Naude E. and Blanchard P., 1982. La Prévision de Récolte Sur Petit Manseng (*Vitis vinifera* L.) Basée Sur Les Courbes-Niveaux de Fertilité et Les Températures Durant Une Période Critique (Yield prevision on Petit Manseng cultivar (*Vitis vinifera* L.) based on fertility and temperature curves during a critical period), Progrès Agricole et Viticole, 99, 440-469

Grollemund P., Abraham C., Baragatti M. and Pudlo P., 2019. Bayesian Functional Linear Regression with Sparse Step Functions, Bayesian Analysis, 14, 111-135

Guilpart, N., Metay A., and Gary C., 2014. Grapevine Bud Fertility and Number of Berries per Bunch Are Determined by Water and Nitrogen Stress around Flowering in the Previous Year, European Journal of Agronomy, 54, 9-20

Howell G.S., 2001. Sustainable Grape Productivity and the Growth-Yield Relationship: A Review, American Journal of Enology and Viticulture, 52, 165-174

Keller M., Tarara J.M. and Mills L.J., 2010. Spring Temperatures Alter Reproductive Development in Grapevines. Australian Journal of Grape and Wine Research, 16, 445-544

Lobell D.B., Field C.B., Cahill K.N. and Bonfils C., 2006. Impacts of Future Climate Change on California Perennial Crop Yields: Model Projections with Climate and Crop Uncertainties, Agricultural and Forest Meteorology, 141, 208-218

Molitor D. and Keller M., 2017. Yield of Müller-Thurgau and Riesling Grapevines Is Altered by Meteorological Conditions in the Current and Previous Growing Seasons, OENO One, 50, 245-258

Oliver M., 2013. An overview of precision agriculture. In: Precision Agriculture for Sustainability and Environmental Protection. Abindon, UK: Routledge. pp 21-37.

Petrie, P. and Clingeleffer P., 2005. Effects of Temperature and Light (before and after Budburst) on Inflorescence Morphology and Flower Number of Chardonnay Grapevines (*Vitis vinifera* L.), Australian Journal of Grape and Wine Research, 11, 59-65

Pouget R., 1981. Action de La Température Sur La Différenciation Des Inflorescences et Des Fleurs Durant Les Phases de Pré-Débourrement et Post-Débourrement Des Bourgeons Latents de La Vigne. (Impact of temperature on inflorescences and flowers differentiation during pre- and post- budbreak stages of the grapevine latent buds.), Connaissance de La Vigne et Du Vin, 15, 65-79

Ravaz L. and Sicard L., 1911. L'effeuillage de La Vigne, (Leaf thinning for Grapevine), Annales de l'École Nationale d'agriculture de Montpellier, 11, 216-244

Srinivasan C. and Mullins M.G., 1981. Physiology of Flowering in the Grapevine – A Review, American Journal of Enology and Viticulture, 32, 47-63

Vasconcelos M.C., Greven M., Winefield C.S., Trought M.C.T. and Raw V., 2009. The Flowering Process of *Vitis Vinifera*: A Review, American Journal of Enology and Viticulture, 60, 411-434

Production gap analysis – an operational approach to yield gap analysis using historical high-resolution yield data sets

C. Leroux[1,2*], J. Taylor[2] and B. Tisseyre[2]
[1]*SMAG, Montpellier 34960, France; cleroux@smag.tech*
[2]*UMR ITAP, Montpellier SupAgro, Irstea, France*

Abstract

Yield potential has major interests in agronomic modelling such as for the computation of fertilizer recommendations. Most precise estimates of this potential can be generated with crop modelling approaches. However, such potential is often theoretical because it has never been obtained in the field and requires lots of input data to be computed, e.g. soil characteristics, management practices. In this study, a decomposition of this theoretical yield potential into a series of intermediate and more realistic potentials, denoted here production potentials, is proposed. This work comes along with a simple method relying on historical high-resolution yield datasets to compute site-specific production potentials and production gaps, i.e. the differences between actual yield and production potentials. Application of the methodology to a real database of yield mapping has proved of interest to identify areas with similar production potentials but different production gaps. Advantages and concerns regarding the use of such production potentials are discussed at the end of the study.

Keywords: production gap, production potential, within-field yield data

Introduction

Yield potential is key in agronomic modelling, especially when it comes to computing fertilizer recommendations or to selecting a relevant seeding density (Grassini *et al.*, 2015). This potential is also of particular interest for farmers and advisors because it provides an indication of the possible improvements in yield that can be obtained or the gap that can be filled if adequate management decisions are considered. Within the scientific community, such studies are referred to as yield gap analyses (Van Ittersum *et al.*, 2012, Silva *et al.*, 2016). Estimates of yield gaps are generally obtained via two different approaches. The first approach usually involves crop modelling and does not explicitly require any yield information once the models are calibrated. It can be argued that this approach provides the most reliable estimate of the 'total' yield gap as the models account for interactions among weather, soils and management to simulate the highest theoretical yield possible (under perfect production conditions). However, crop modelling requires the specification of many input parameters and current models are rarely, if ever, suitable for application at a site-specific scale. The second approach, in contrast, estimates potential yield, and the associated yield gap, from observed, objective yield measures, i.e. yields that have already been obtained. There have been several methods proposed for deriving these observation-base yield gaps including using boundary-functions, upper percentiles of yield distributions and maximum yields derived from one or a combination of experimental station yield, or growers contests (Van Ittersum *et al.*, 2012).
Most yield gap studies were intended to compute potential yields and resulting yield gaps at medium to very coarse scales, i.e. field, farm, regional or national level. Modelling approaches cannot yet be applied confidently at sub-field levels, in part because the input data, i.e. soil characteristics, weather, management practices have not been available at finer scales and in part because the models are not designed to account for the stochastic noise in these fine-scale measurements. Statistical approaches have utilized average yields at farm/field/regional scales, as these are widely available and have a long

DOI 10.3920/978-90-8686-888-9_8, © Wageningen Academic Publishers 2019

temporal sequence. The literature[1] shows that very few studies have aimed to perform site-specific yield gap analysis, despite much of the literature highlighting the potential for precision agriculture approaches to be used in yield gap analyses. This is a clear gap in knowledge as it has been shown multiple times that yield variations are very important at within-field scales. Nothing prevents the potential yield to be as, or maybe more, spatially variable than the actual yield itself.
The availability of combine yield monitors since the early 1990s means that long time-series (10+ years) of yield data are now available in many arable production systems. This within-field data is very valuable as it provides an objective and site-specific measure of the production that has been obtained across a field. Given the availability nowadays of these temporal site-specific data, a statistical approach to compute and analyze an operational site-specific yield gap, referred to as the production gap in this study, should be possible. This study provides a step towards this goal by deriving a site-specific production gap to help to make more informed management decisions at the within-field scale.

Production gap vs yield gap

To clarify the notions that will be introduced in this work, the yield gap concept is presented as a response to increasing availability of yield-limiting factors, e.g. water, nutrient, in a two-dimensional plot (Figure 1). In reality, this is a multi-dimensional interaction but, for simplification, a two-dimensional representation is given. The yield gap is usually defined as the difference between the actual yield (A) within the field and the potential yield (F) that the crop might reach under optimal growing conditions, i.e. when water and nutrients are not limited, and when the crop is not subjected to any abiotic or biotic stress (Van Ittersum *et al.*, 2012; Grassini *et al.*, 2015). Under those conditions, the crop growth is only controlled by temperature, lightning conditions, crops genetics and atmospheric CO_2. Following these conditions, a theoretical yield response can be

[1] A Boolean search of "precision agriculture" AND "yield gap".

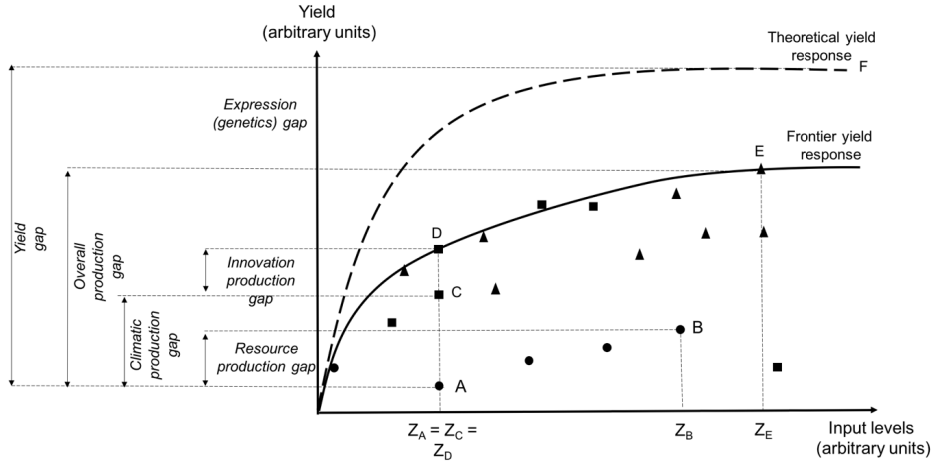


Figure 1. Decomposition of the theoretical potential yield into more realistic potentials. Inspired from Van Dijk *et al.* (2017). Points are indicative of observed yield obtained across space and time. Different shapes indicate different years. () actual yield, (B) agronomic production potential, (C) climatic production potential, (D) innovation production potential, (E) total production potential, (F): yield potential. The Z_i are different sets/levels of inputs.

derived through crop modelling (Figure 1, dashed line). This response curve generally forecasts higher yields than those that are actually observed in fields.

When considering empirical, rather than modelling approaches, yield and production gaps need to be determined by pairing and contrasting a target site, i.e. the site associated with A in Figure 1, with sites with similar production conditions. These sites are termed 'relatable' sites and need to be defined within the spatial scale of the analysis. The concept of 'relatable' sites is deliberately scalable in space. For example, they could be applied to: (1) mean farm yields and bench-marking at a regional scale, (2) mean field yields within a single or grouping of farming systems or (3) site/zone-specific yields at a within-field scale. Be aware that the following relational yield potentials and gaps described will vary across crops, growers, climatic zones or territories (Figure 1 is a schematic only). To provide more realistic and feasible estimates of yield gaps, some authors have proposed decomposing the yield potential (F) into a series of intermediate potentials (Van Dijk *et al.*, 2017; Silva *et al.*, 2016). The theoretical yield response (Figure 1, dashed line) is complemented by a more realistic boundary yield response (Figure 1, continuous line) that can be created with 'local' yield observations within a given area (Van Dijk *et al.*, 2017). Figure 1 is inspired from what was proposed by Van Dijk *et al.* (2017) but additional notions were included, more especially those that pertain to the production gap. All the yield potentials and the associated gaps plotted in Figure 1 are detailed below:

- [A] *Actual yield* (Y_A): Yield obtained at a site S_i for a given set of inputs Z_A in a given year y_m. It relies on management decisions, climatic conditions and crop stresses on yield.
- [B] *Agronomic production potential* (Y_{AP}): For a set of inputs Z_B (Z_A), this is the highest yield that has been obtained in y_m at a site relatable to S_i. Differences with A come from variations in the set of inputs and/or management activities.
- [C] *Climatic production potential* (Y_{CP}): For the level of inputs Z_C (Z_A), this is the maximum yield that has been obtained at a site relatable to S_i under differing climatic conditions than in y_m.
- [D] *Innovation production potential* (Y_{IP}): For the level of inputs Z_D (Z_A), this is the maximum yield that has been obtained at a site relatable to S_i with appropriate changes in technologies and management practices. Climatic conditions are the same as for C.
- [E] *Total production potential* (Y_{TP}): This is the maximum potential yield that has been observed at a site relatable to S_i with an optimal level of inputs Z_E and management (and favorable climatic conditions). Differences with the potential A come from differences in the set of inputs, management activities, and/or climatic conditions
- [F] *Yield potential* (Y_p): Maximum yield that could be obtained under optimal growing conditions, optimal inputs/management and in the absence of any stress (biotic or abiotic). This potential is theoretical and derived from crop modelling approaches.

Differences or gaps between these potentials can be understood as follows:

- *Resource production gap (B-A)*: Potential yield that could have been gained with better input management and management practices. In a site-specific crop management (SSCM) context, this could be derived using data derived from in-field experiments/decisions using variable-rate technologies.
- *Climatic production gap (C-A)*: When Z_A tends to be temporally uniform, observed differences in yield at site over time will be due to season-specific weather impacts (e.g. water supply via precipitation) considering relatively uniform management practices
- *Innovation production gap (D-C)*: When Z_A tends to be temporally uniform, it presents a plausible potential improvement in production via innovation in technologies or agronomic management (e.g. improve fertilizer type).
- *Overall production gap (E-A)*: Potential production gain for a given yield observation relative to all available data and with all options available in regards to improving input levels/efficiencies and improving technologies and management.

- Expression (genetic) gap (*F-E*): Classical difference between (1) potential yield as modelled by the best interactions between crop genetics, weather, soil and management, and (2) maximum yield observed.

Note that the gap between D and E is not defined as it is a combination of the previously defined production gaps. Here, the focus revolves around the overall production gap, and so for multiple reasons. Firstly, yield data sets provide yield potentials that are realistic and have already been observed, meaning that they could be attained again. Secondly, these data have not previously been well exploited in this context but offer a large opportunity for understanding site-specific production gaps/potentials. It is not claimed that the theoretical yield gaps/potentials are not of interest, but rather that the proposed site-specific production potentials require less input data to compute and are collected at and relevant to the scale of analysis (site-specific), which crop models are currently not. However, if the expertise and the relevant data and models are available for sub-field modelling, theoretical potentials might be equally valuable or preferable to calculate. In this case study, information on the amount of inputs applied is incomplete, thus it must be assumed that the set of inputs Z is not uniform over seasons. If it was, then the analysis would only be of the climatic and innovation production gaps.

Material and methods

Description of the multi-temporal within-field yield data

The study was carried out on a 31-ha field in the north of France near Evreux (WGS84 datum: E: 0.78°, N: 48.95°). The field is cropped in a wheat (2003-2005-2007-2009-2011-2013-2015) and canola (2004-2006-2010-2014) – flax (2008-2012) rotation, i.e. a 'first' wheat is grown every other year with a break crop in between. Nine years of yield mapping were available spanning the 2003-2015 period but only the wheat crop was considered in this study (six years) to minimize the need to standardize between different crop types in this exploratory work. More descriptive, spatial and temporal characteristics can be found in Leroux *et al.* (2018).

Yield map pre-processing

Within-field yield data were pre-processed in a similar way as in Leroux *et al.* (2018). Outlying observations were first filtered, and a grid composed of 10×10 m pixels, whose orientation followed that of the harvested rows, was then superimposed on the yield data. For each pixel of the grid, yield values were first averaged by year so as to obtain one yield value for each pixel and each year. The objective was to make sure that each year had the same influence in each pixel even if the number of observations falling into each pixel was different from year to year. Empty pixels in specific years due to missing yield observations were given the mean yield value over the years in the same pixel.

Computation and evaluation of the yield overall production gap

The estimated yield total production potential ($\widehat{Y_{TP}}$) is defined over each pixel *i* and is computed as the maximum yield obtained on the i^{th} pixel throughout the years:

$$\widehat{Y_{TP}}(i) = \max\left(Y^m(i)\right), \quad m \in \{1{:}\,M\} \tag{1}$$

Where $Y^m(i)$ is the i^{th} yield observation in year *m*, and M is the number of years in the historical yield database[2].

The year at which this potential is reached is denoted m' in the rest of the study. Once this potential is defined, the estimated annual overall production gap ($\widehat{Y_{pg}^{m}}$) can be easily computed over each

[2] In this case, the 'relatable' sites are simply constrained to the same pixel over time. However, there is no reason why Y^m could not be derived from a vector of yield values from 'relatable sites' in year *m*, instead of just the local pixel value (if sites with similar production conditions were known and the associated data was available).

pixel *i* for each year *m* as the difference between the total production potential and the actual yield (Equation 2).

$$\widehat{Y_{pg}^{m}}(i) = \widehat{Y_{TP}}(i) - Y^{m}(i)\,, \quad m \in \{1{:}\,M\} \tag{2}$$

where $\widehat{Y_{pg}^{m}}(i)$ is the overall production gap of the ith yield observation in year *m*, and M is the number of years in the historical yield database.

The annual overall production gap is interesting in itself as it can be used to evaluate the impact of specific growing conditions or management practices during a given year. However, to fill such a gap, there is a need to answer operational questions such as (1) how often is this potential reached or almost reached? or (2) how much deviation from that maximum is there in a given pixel over time? In this study, the second option is chosen and a multi-annual overall production gap $\widehat{Y_{pg}}$ is proposed to be computed over each pixel *i* using Equation 3 as:

$$\widehat{Y_{pg}}(i) = \frac{1}{\widehat{Y_{TP}}(i)} \times \frac{1}{M-1} \times \sum_{m=1, m\neq m'}^{M} \widehat{Y_{pg}^{m}}(i) \tag{3}$$

This indicator is expressed as a percentage so that it can be compared across varying situations. It is computed over M – 1 years because the year m' is not taken into account.

Results and discussion

Characterization of the within-field production gap over the field under study

Figure 2 details the annual spatial wheat yield patterns (2003-2015) on a common scale for all years. It is clear that some years are much more productive, e.g. 2009 and 2015, than others. Some yield patterns also exhibit a much wider range of variation, e.g. 2011, than others that seem to be much more homogeneous, e.g. 2005 and 2007. Spatial yield patterns are also relatively stable from year to year, especially in 2003, 2005, 2009, 2011 and 2015, with a lower productive area in the northern section of the field and a higher-yielding one in the centre and southern portions of the field (Leroux *et al.*, 2018).

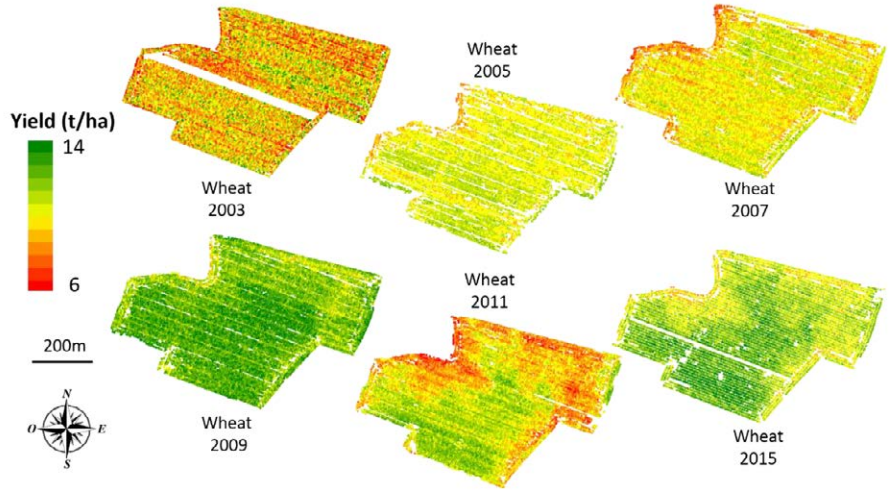


Figure 2. Within-field wheat yield maps of the field across the 2003-2015 period.

The yield total production potential and multi-annual overall production gap were computed from this historical database of yield mapping. Unsurprisingly, the production potential spatial pattern matches the yield spatial patterns in Figure 2. This operational potential is relatively variable over the field and ranges from less than 10 to more than 13 t/ha. This information should drive boundary decisions on maximum and optimum input levels and/or investigations into yield determining factors at each site. From a more general perspective, the total yield production potential in this field is strongly related to the yield obtained in the most productive year (2009), although it is unclear if a single year dominance like this is likely in other fields and production systems. The multi-annual overall production gap, on the contrary, exhibits a different spatial pattern that is not linked with the previously described yield spatial patterns (Figure 3). Areas of both high and low yield potential exhibit a low gap in production. There is however a central zone in the field where the production gap is relatively high (>15%) that incorporates both high and low yield potential areas. The multi-annual overall production gap brings in operational information to enable users to identify and manage areas that have the highest potential for improvement (or conversely, the highest area of risk). For example, it is clear that the northern section of the field is the least productive, but also the least variable. If the yield-limiting factor here is not manageable, it may mean that it is more valuable to put efforts into reducing the production gap in the centre of the field.

Further considerations

It must be stressed that the proposed production gap analysis is iterative in nature as many factors are likely to evolve across time, e.g. technologies, climate, variety. The intent of the production gap analysis is not to assess whether a technological change should be done but rather to say that with such technology or such variety, a given yield can be obtained. It remains possible (would be expected at some stage) that the actual yield of the current season outreaches the historical production potential. It is suggested here that the yield local production potential would simply be replaced by the yield of the current season, which would be seen as the new local production potential. In such a case, it should not be considered that the previously proposed production potential was a bad estimate but rather that as more years of yield mapping become available (best available data), estimates become more viable as they encompass a large range of growing conditions and can be derived from years with climatic conditions similar to current (or forecast) conditions. It is nonetheless acknowledged that the statement of best available data is questionable as crop models do exist and provide higher

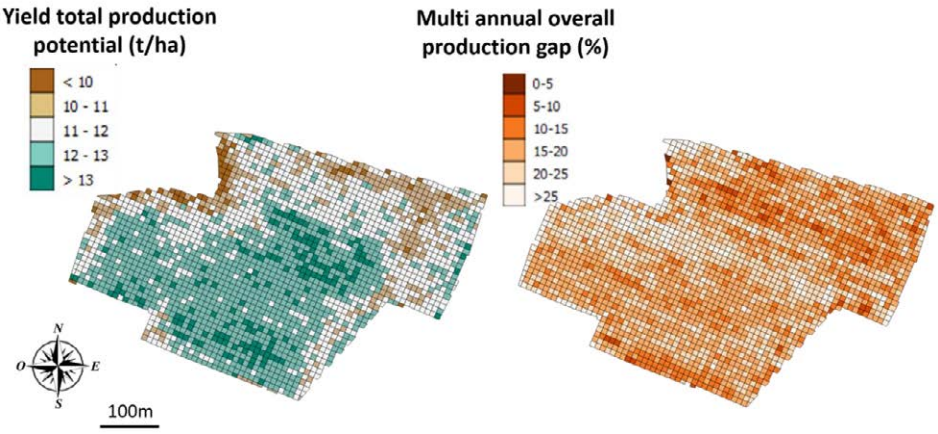


Figure 3. Within-field yield maps of the yield local production potential and multi-annual local production gap across the 2003-2015 period.

and certainly more accurate yield potentials. However, the strong advantage of historical yield maps is that they provide an operational potential, i.e. yield values that have already been obtained. The (site-specific) yield total production potential could be seen as a yield objective at the beginning of the cropping season. This potential should be revised during the growing season as uncontrolled environmental factors evolve, e.g. weather effects or disease pressures. One issue, however, with the proposed production gap analysis is that growers might obtain each year a yield that is close to their local maximum yield, but without knowing if their management practices are optimal. In fact, as this local maximum is seen as a production potential, growers cannot know if they can possibly outreach it. This information should be used to design and perform site-specific experiments to gain a better insight into the local yield response curve to inputs and management practices. In this study, specific attention was paid to the evaluation of the within-field yield total production potential, mainly because information on the amount of inputs applied or precise management practices is incomplete. Production potentials are interesting because they can be computed in terms of gross margins and gross margin gaps to help users evaluate the economic risk of making future management decisions. It must be clear that different yield production potentials would necessarily lead to different gross margin-related maps (Figure 1).

Conclusions

In this work, an operational approach relying on historical databases of yield mapping was proposed to help farmers position themselves in terms of their actual yield levels and their potential for improvement. The objective is to deliver maps of total production potential and multi-annual overall production gap. The proposed approach is iterative in nature as new yield maps will improve the definition of the yield production potentials and gaps. Future study will aim at merging the production gap analysis with economics to evaluate the risk for farmers of making future management decisions.

References

Grassini, P., van Bussel, L.G.J., van Wart, J., Wolf, J., Claessens, L., Yang, H. *et al.,* 2015. How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. Field Crops Research, 177, 49-63.

Leroux, C., Jones, H., Taylor, J., Clenet, A., Tisseyre, B., 2018. A zone-based approach for processing and interpreting variability in multitemporal yield data sets. Computers and Electronics in Agriculture, 148, 299-308

Silva, J.V., Reidsma, P., Laborte, A., & van Ittersum, M.K., 2016. Explaining rice yield gaps in Central Luzon, Phillippines: an application of stochastic frontier analysis and crop modelling. European Journal of Agronomy, 82, 223-241

Van Dijk, M., Morley, T., Jongeneel, R., van Ittersum, M., Reidsma, P., & Ruben, R., 2017. Disentangling agronomic and economic yield gaps: An integrated framework and application. Agricultural Systems, 154, 90-99.

Van Ittersum, M.K., Cassman, K.G., Grassini, P., Wolf, J., Tittonell, P., & Hochman, Z., 2012. Yield gap analysis with local to global relevance – A review. Field Crops Research, 143, 4-17.

Agriculture and digital sustainability: a Digitization Footprint

F. Marinello[1*], R.G.V. Bramley[2], Y. Cohen[3], S. Fountas[4], H. Guo[5], M. Karkee[6], J.A. Martínez-Casasnovas[7], D.S. Paraforos[8], L. Sartori[1], C.G. Sørensen[9], B. Stenberg[10], K. Sudduth[11], B. Tisseyre[12], G. Vellidis[13] and S.G. Vougioukas[14]
[1]*Department of Land Environment Agriculture and Forestry, University of Padova, Legnaro, Italy; francesco.marinello@unipd.it*
[2]*Commonwealth Scientific and Industrial Research Organization, Waite Campus, Urrbrae, Australia*
[3]*Agricultural Engineering Institute, Agricultural Research Organization, Israel, Rishon Le'Zion, Israel*
[4]*Department of Natural Resources Management and Agricultural Engineering, Agricultural University of Athens, Athens, Greece*
[5]*College of Land Science and Technology, China Agricultural University, Beijing, China P.R.*
[6]*Center for Precision and Automated Agricultural Systems, Washington State University, Prosser, USA*
[7]*Research Group in AgroICT & Precision Agriculture, University of Lleida, Spain*
[8]*Institute of Agricultural Engineering, University of Hohenheim, Stuttgart, Germany*
[9]*Department of Engineering, Aarhus Universitet, Aarhus, Denmark*
[10]*Department of Soil and Environment, Swedish University of Agricultural Sciences, Uppsala, Sweden*
[11]*US Department of Agriculture, University of Missouri, Columbia, USA*
[12]*Montpellier SupAgro, Irstea, University of Montpellier, Montpellier, France*
[13]*Crop & Soil Sciences Department, University of Georgia, Tifton, USA*
[14]*Department of Biological & Agricultural Engineering, University of California, Davis, CA, USA*

Abstract

The growing availability and capabilities of sensing and communication infrastructures such as monitoring stations, proximal and remote sensing technologies, geolocation systems, and standard communication protocols, along with apparently decreasing costs of the same technologies are pushing widespread collection, implementation, transmission and use of digitized information in agriculture. Such an uncontrolled process poses questions on sustainability of the virtual environment (i.e. a software-based space) where such processes take place. The aim of the present work is to introduce a digitization footprint (DF), which parameterizes the amount of digital information so as to quantify the specific or general use of digital or processing information in terms of volumes, time, efforts or costs invested for data storage, processing or transfer. Such a digitization footprint can be directly related to the availability and suitability of the digital resources in terms of costs (storage, transfer, processing, cloud computing), and speed (processing, upload, download), and can help define pathways for effective and widespread development.

Keywords: digitization, big data, data management, virtual environment

Introduction

In recent years, the original concepts underpinning precision agriculture have been further developed, taking extended, system-oriented advantage of the use of information and communication technologies in farm management to expand the toolkit available to adopters of precision agriculture (Wolfert *et al.*, 2017; Yost *et al.*, 2019). Indeed, new emerging technologies such as unmanned aerial vehicles, Internet of Things, cloud computing, artificial intelligence, and tracking systems, along with consolidated techniques such as ground sensing, remote sensing and communication technologies are extending the precision agriculture concept, enhancing the existing tasks of management and decision making influenced by the context, situation and location awareness (Kamilaris *et al.*, 2018). Additionally, Digital Twins (virtual, digital equivalents to physical objects that provide a thorough,

DOI 10.3920/978-90-8686-888-9_9, © Wageningen Academic Publishers 2019

realistic representation of the object and the context that this object is working in) that are already becoming available in other scientific disciplines, such as industrial informatics (Schluse *et al.*, 2018), are expected to have a profound impact on agriculture in the coming years. Therefore, agriculture is moving towards a new development phase, relying not only on location but also on data enhanced by context awareness and triggered by real-time events (Wolfert *et al.*, 2018; Paraforos *et al.*, 2017; Pezzuolo *et al.*, 2014).

Such a process is, in general, seen positively by the scientific community, as it also has positive impacts on sustainability. Indeed, it has the potential to optimize resources (agrochemical and fertilizers in general, machinery use, irrigation, etc.) with positive environmental and economic effects and to promote farmers' expertise through new skills and less labour intensive approaches.

On the other hand, the virtual environment is typically not included in sustainability considerations or calculations. As for the physical environment during the last century, the virtual one is now perceived as an infinite source and sink, mainly due to its relatively large size compared to the present economy and needs. This is clearly debatable, and will probably reach a more critical level in the next few years; however, questions on digital sustainability should already be posed.

The present work is aimed at introducing a so called *Digitization Footprint* (DF), which parameterizes the amount of digital information with the purpose of supporting virtual environment (a software-based space) sustainability analyses. Virtual environment (VE) sustainability is here defined as the ability to maintain VE related services at a needed rate or level. Such DF parametrization is thus intended to quantify the use of digital information and processing operations in terms of volumes, time, efforts or costs invested for data storage, processing or transfer. Such digitization can be directly related to the availability and suitability of the digital resources in terms of costs (storage, transfer, processing, cloud computing) or speed (processing, upload, download), etc.

Thus, this paper introduces a methodological approach or framework to encourage and support debate and recommendations for agriculture in the present and future digital era. Such an approach could and should form the basis for a reference framework that will better define and delineate the digitization footprint.

Proposed framework

As discussed by Van Evert *et al.* (2017), there are basically three categories of data intensive technologies that are available to farmers:

- Farm management information systems (FMIS), i.e. tools for collecting, processing, storing and disseminating data in the form needed to carry out farming operations and functions (Sørensen *et al.*, 2010a; Fountas *et al.*, 2015);
- Precision agriculture (PA), which is the scientific domain and management approach that deals with the targeting of inputs to crop production according to crop requirements on a localized basis (Stafford, 1996), or selective harvesting and product streaming (e.g. Bramley *et al.*, 2005), thus increasing the number of (correct) decisions per unit area of land per unit time, with associated net benefits (McBratney *et al.*, 2005);
- Automation and robotics, which deals with introduction of mechatronics and advanced computational methods at different levels of agricultural production (Amatya *et al.*, 2017; Le Flecher *et al.*, 2017).

For all of these categories, a common data framework can be defined, as shown in Figure 1. Three main groups form the proposed representation:

- Data sources, i.e. the 'places' where data are originally produced, computed or collected, such as proximal or remote sensors, global navigation satellite systems (GNSS), farmer surveys, etc.
- Data users, i.e. the terminals or devices which receive data and provide some kind of representation, synthesis or decision, such as farm computers, tractor universal terminals, smartphones, etc.

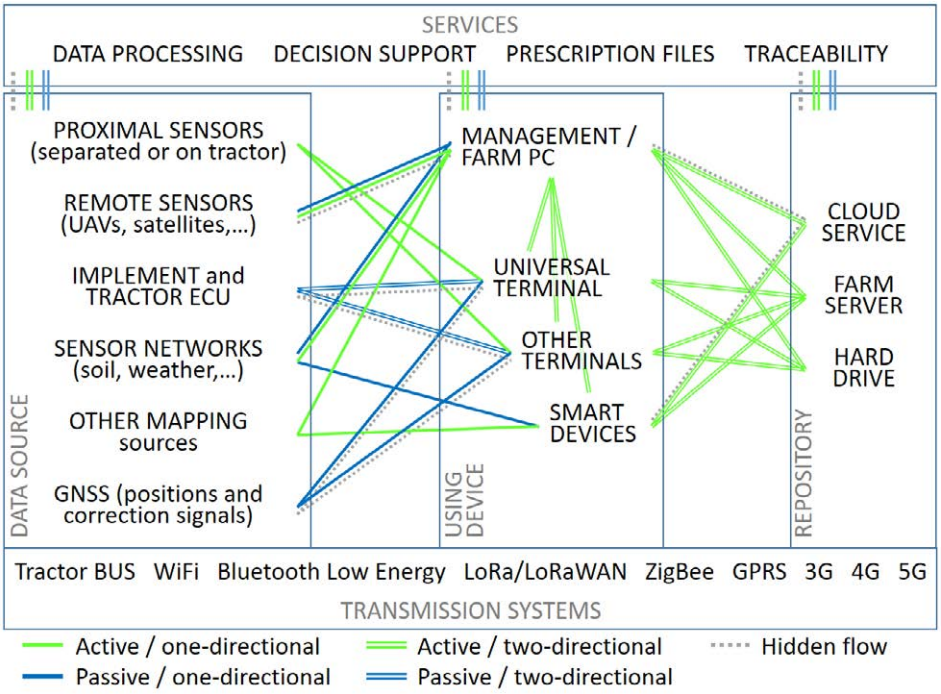


Figure 1. Reference data framework and representation of data flows of the Digitization Footprint concept.

- Data repositories, i.e. the 'places' where historical data as well as new data or synthesis parameters are temporarily stored. The time between collection and use can be very different, from a few seconds, minutes or hours to several weeks or even years in some cases.

The three groups are built on a large base, which encompasses different transmission methods, from cables (with BUS/ISOBUS lines) to wireless long distance means and, in some cases, high speed broadband communication networks such as 5G mobile networks.

In all of the cases, the systems constitute an enabling technology supporting data transfers between the groups. Across the top of the representation (Figure 1), there are services, such as traceability interfaces, zoning, definition of prescription maps, and other data processing applications based on software tools available online. The appearance of the lines indicates some characteristics of the data flow:

- Active flows are characteristic of transfers which take place as a consequence of a specific order or call;
- Passive flows are those generated independently of final users, e.g. data from the Sentinel 2 satellite, which are generated even when nobody is using them;
- One-directional flows are characteristic of those systems which have a well-defined direction of transmission, e.g. in the case of navigation satellite systems transferring data from the satellite to the GNSS unit;
- Two-directional flows are characterized by some kind of feedback from the user to the data source or repository;
- Hidden data flows are representative of information that is implemented to allow or track or correct transferred data, but that gives no information and is not specifically used by the farmer

(e.g. many values of the GNSS NMEA sentences, secure codes in the traceability systems, addresses, etc.).

The proposed data framework constitutes an attempt to develop a reference draft along the lines of information reference systems (e.g. Sørensen *et al.*, 2010b), which would be used in order to pose the basis for the two main concepts mentioned above:

1. The idea of a virtual environment, which of course is not new but is analysed here with respect to its sustainability.
2. The Digitization Footprint concept, which parameterizes the amount of generated, processed or transferred digital information.

Outcomes

A variety of indicators (ascribable to the three cited categories, FMIS, PA and Automation and Robotics) are giving evidence of clear changes in farming systems. In particular, a constant increase in the total amount of data being used is characterizing changing agricultural practices. Such an idea is supported by certain trends, which are recognizable both in farming practices and in technology evolution in general. Specifically, analysing field practices presented as case studies in published papers, a growing trend can be noticed during the last two decades of digitization (Miller *et al.*, 2017). At the same time, technologies have been also evolving, providing higher quality data year by year, at higher resolution and as a result, large volumes of data are being generated. By way of example, this can be noticed looking at the data volume per hectare provided by different satellite missions, as reported in Figure 2. The reported exponential trend highlights a 10-fold increase over 8 years.

This is clearly a promising and positive trend. However, it is apparent that the growth of technologies in agricultural applications is proceeding faster than their adoption by farmers. This is particularly evident in those countries where the access to technologies is limited by the unavailability of some enabling technologies, such as internet connection, reducing the ability of individual farmers to access or transfer large amounts of data (e.g. weather information, satellite maps).

Recent studies (Thompson *et al.*, 2016; ITU, 2016) show how many nations (mainly located in Africa and in the south parts of Asia) are in a condition where several hours or even several days would be needed to allow download of a single satellite map (even a Sentinel 2 map). This problem is not new (Lamb, 2017), and services have been made available where maps are processed online and only a few synthesis parameters are downloaded; clearly problems arise when generated data are

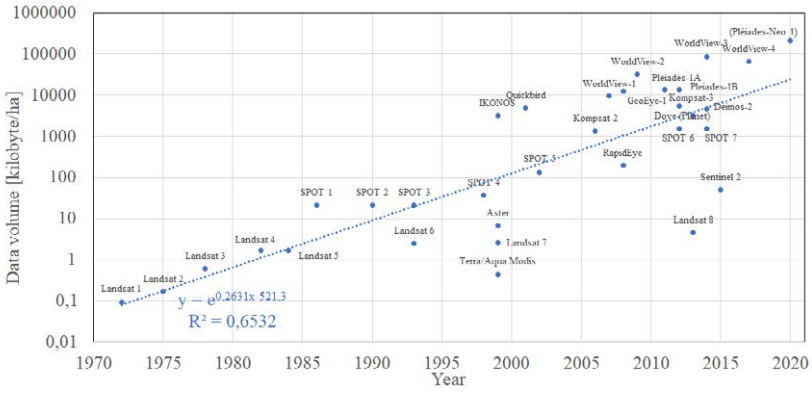


Figure 2. Trends in data volume per year made available by satellites used in agriculture, in the last 50 years (data retrieved from satellite manufacturers, managers or service providers' websites).

not proportional to available resources and skills. Indeed, another crucial issue is the ability of farmers to use these technologies and to interpret them in the new digital virtual environment. In developed countries, where the most technological advances are available, the bottleneck in their use is similarly the present capacity of farmers to manage technologies and interpret the data that are generated for decision support.

Virtual environment sustainability

The virtual environment should be considered on the basis of at least two elements: amount of data and time. With regard to the amount of data, the main issues are theorized to be related to the degree of data management infrastructure available per capita, including data storage, data processing, data transfer, and accessibility to cloud services. All of these services have to be available with a reasonably fast speed (or in timely fashion); indeed the value of data and related services often disappears after actions or decisions have been taken. The concept of reasonably fast speed or timing is not generalizable; however, for some agricultural processes it can be as low as a few days or even a few hours.

Goodland and Daly (1996) proposed a set of basic conditions for environmental sustainability. The main aspects of their work are briefly reported in Table 1, along with corresponding elements in the VE. In comparison, air, soil, water and other physical environment elements are replaced by features and attributes typically related to digital data such as storage, bandwidth or security.

The aim of this paper is to generate a discussion on the topics around these basic concepts and their implementation. Thus, the main questions to be posed with regard to virtual environment sustainability are:

1. Is it sufficient to limit virtual environment sustainability just to the amount of data and time taken or needed during a growing season?
2. Which other elements should be included?
3. Who are the players and the decision makers of the virtual environment?
4. What are the functional units of interest? Farm, area of production, volume of production, anything else?
5. Is life-cycle analysis (LCA) a viable approach for the analysis of the VE in terms of external or internal fluxes, type of farm (intensive, extensive), etc. over a given unit of importance?

This list is not exhaustive; certainly other questions will need to be defined and addressed.

Digitization Footprint

In the authors' point of view, a Digitization Footprint can be thought of as a decision support tool that could be used to support data management in relation to virtual environment sustainability. More practically, the DF should synthesize a set of indicators that quantify the impact of digitization processes (summarized in Figure 1) on farm-related activities. For this reason, a DF has to be

Table 1. Physical vs virtual environment from a sustainability perspective.

	Physical environment	Virtual environment
Natural/virtual capital	air, soil, water, forests,…	data storage, bandwidth, archives, algorithms,…
Non-renewable natural/virtual capital	natural capital, mineral resources	data storage
Non-substitutable natural/virtual capital	energy	time, missed or delayed data
Risks/threats	wastes	ownership, security

based on the digital information used, and on data processing operations, analysing data volumes, computational capabilities needed, download or upload time, and direct and indirect effects on management, among others. Extracted indicators can then be directly related to the availability and suitability of digital resources both in terms of costs (storage, transfer, processing, cloud computing) and speed (processing, upload, download). Additionally, the farmers' skills for the interpretation and implementation of virtual data have to be considered. Indeed, the farmer's training in the use of data is crucial for the success of the virtual environment implementation; otherwise technological advances will be ineffective. As an example, a DF can provide a basis for allowing comparison of the impact of a number of small farms organised as a co-operative with respect to a few independent large farms, or should trace development pathways reasonably implementable by farmers.
Following these considerations, the main questions to be posed with regard to the Digitization Footprint are:

1. Which parameters or indicators can better synthetize the DF?
2. What is the counterpart of DF at the farm level, which defines a maximum admissible DF threshold?
3. Is the DF somehow correlated to energy footprint (as a dependent or independent variable)?
4. How is the farmer's education related to the DF, and how much does it contribute to the use of the new technological advances?
5. Is there a way to quantify the *quality* of digital information and of related methodologies?
6. Should a DF counterpart be defined (e.g. a sort of *digitalization capability* index)?

Conclusions and perspectives

The present paper introduces the concept of the Digitization Footprint (DF) as a decision support tool that could be used to support data management in agriculture in relation to the virtual environment. To achieve this, a set of indicators is clearly needed to quantify the impact of digitization processes on farming and related activities. Such indicators have to be related to digitized data, at least in terms of costs (storage, transfer, processing, cloud computing) and time (processing, upload, download). Additionally, the quantified digitization level has to be adoptable, profitable, affordable and acceptable for the farmer (in relation to his or her education, skill and aptitude). To this end, DF analyses could support agricultural machinery manufacturers, sensor and instrumentation producers, service providers and policy makers in defining a process for effective and sustainable integration of IT and other impactful technologies at the farm scale.

References

Amatya, S., Karkee, M., Zhang, Q. and Whiting, M.D., 2017 Automated detection of branch shaking locations for robotic cherry harvesting using machine vision. Robotics 6 31.

Bramley, R.G.V., Proffitt, A.P.B., Hinze, C.J., Pearse, B. and Hamilton, R.P., 2005. Generating benefits from Precision Viticulture through selective harvesting. In: Stafford JV. (Ed) Proceedings of the 5th European Conference on Precision Agriculture. Wageningen, the Netherlands: Wageningen Academic Publishers. pp 891-898.

Fountas, S., Carli, G., Sorensen, C. G., Tsiropoulos, Z., Cavalaris, C. Vatsanidou, A., *et al.,* 2015. Farm management information systems: current situation and future perspectives. Computers and Electronics in Agriculture 115 40-50.

Goodland, R., Daly, H., 1996. Environmental sustainability: universal and non-negotiable. Ecological Applications 6 1002-1017.

ITU, International Telecommunication Union., 2016. Measuring the Information Society Report 2016, ISBN: 9789261214210. https://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2016/MISR2016-w4.pdf. (last accessed February 2019)

Kamilaris, A., Kartakoullis, A. and Prenafeta-Boldú, F. X., 2018. A review on the practice of big data analysis in agriculture. Computers and Electronics in Agriculture 143 23-37.

Lamb, D. W., 2017. Accelerating precision agriculture to decision agriculture: A review of on-farm telecommunications challenges and opportunities in supporting a digital agriculture future for Australia, ISBN: 9781921597756. https://www.crdc.com.au/sites/default/files/P2D telecommunications – UNE Final Report.pdf. (last accessed February 2019)

Le Flecher, E., Durand-Petiteville, A., Cadenat, V., Sentenac, T. and Vougioukas, S., 2017. Implementation on a harvesting robot of a sensor-based controller performing a u-turn. In Proceedings of the 2017 IEEE International Workshop of Electronics, Control, Measurement, Signals and their Application to Mechatronics, 7945895.

McBratney, A. B., Whelan, B., Ancev, T. and Bouma, J., 2005. Future Directions of Precision Agriculture. Precision Agriculture 6 7-23.

Miller, N. J., Griffin, T. W., Bergtold, J., Ciampitti, I. A. and Sharda A., 2017. Farmers' Adoption Path of Precision Agriculture Technology. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences 8 (2) 708-712.

Paraforos, D.S., Vassiliadis, V., Kortenbruck, D., Stamkopoulos, K., Ziogas V., Sapounas, A.A., *et al.*, 2017. Multi-level automation of farm management information systems. Computers and Electronics in Agriculture 142 504-514.

Pezzuolo, A., Basso, B., Marinello, F. and Sartori, L., 2014. Using SALUS model for medium and long term simulations of energy efficiency in different tillage systems. Applied Mathematical Sciences 8 (129-132) 6433-6445.

Schluse, M., Priggemeyer, M., Atorf, L. and Rossmann, J. 2018. Experimentable Digital Twins-Streamlining Simulation-Based Systems Engineering for Industry 4.0. IEEE Transactions on Industrial Informatics 14 1722-1731.

Sørensen, C.G., Fountas, S., Nash, E., Pesonen, L., Bochtis, D., Pedersen, S.M., *et al.,* 2010a. Conceptual model of a future farm management information system. Computers and Electronics in Agriculture 72 (1) 37-47.

Sørensen, C.G., Pesonen, L., Fountas, S., Suomi, P., Bochtis, D., Bildsøe, P. *et al.,* 2010b. A user-centric approach for information modelling in arable farming. Computers and Electronics in Agriculture 73 44-55.

Stafford, J. V., 1996. Essential technology for precision agriculture. In P. C. Robert, R. H. Rust, & W. E. Larson (Eds.), Precision Agriculture. Proceedings of the third International Conference on Precision Agriculture. Madison, WI, USA: ASA/SSSA/CSSA pp. 595-604.

Thompson, J., Sun, J., Möller, R., Sintorn, M. and Huston, G., 2016. Akamai's state of the internet: Q1 2016 report, edited by D. Belson.

Van Evert, F. K., Fountas, S., Jakovetic, D., Crnojevic, V., Travlos, I. and Kempenaar, C., 2017. Big Data for weed control and crop protection. Weed Research 57 218-233

Wolfert, S., Ge, L., Verdouw, C. and Bogaardt, M. J., 2017. Big Data in Smart Farming – A review. Agricultural Systems 153 69-80.

Yost, M.A., Sudduth, K.A., Walthall, C.L. and Kitchen, N.R., 2019. Public-private collaboration toward research, education and innovation opportunities in precision agriculture. Precision Agriculture 20 (1) 4-18.

Using the WOFOST crop growth model to assess within-field yield variability

A.C. Tagarakis[1*], G. Mimić[1], H.M. Vaessen[2], F. Rodriguez-Moreno[1], F.K. van Evert[2] and V. Ćirić[3]
[1]*University of Novi Sad, BioSense Institute, Dr. Zorana Djindjića 1, 21000 Novi Sad, Serbia; atagarak@biosense.rs*
[2]*Agrosystems Research, Wageningen University & Research, 6708 PB Wageningen, the Netherlands*
[3]*University of Novi Sad, Faculty of Agriculture, Trg Dositeja Obradovića 8, 21000 Novi Sad, Serbia*

Abstract

Usually crop models are run as point-based at field level. However, various soil properties may cause crop growth and yield to vary significantly at a smaller spatial scale than the field. Thus the objective of this study was to determine whether within-field variation in yield can be simulated when appropriate input data are available. A study was performed on a 64-ha maize field located in Vojvodina region (northern Serbia). The soil was characterized as Chernozem. The field was managed by the farmer at a sub-field level in 2017. Apparent electrical conductivity zones were used for targeted soil sampling and final yield was recorded by yield monitors installed on the two harvesters used to harvest the field. According to the results, field slope, water flow direction and accumulation were important yield driving factors. Spatially variable soil properties were introduced into the WOFOST crop model by estimating available water within the field, based on calculated water flow accumulation. Points were selected within management zones. Yield predicted by the model was correlated with the yield measured by the yield monitors.

Keywords: crop modelling, water flow accumulation, apparent electrical conductivity, soil texture, yield monitor

Introduction

Agricultural production systems models are used for different applications, such as yield gap analysis, resource use and efficiency, farmers' advice and food security (Holzworth *et al.*, 2015). From the beginning of their development, crop models sought to simulate crop growth, relating the factors and processes which influence plant development and biomass production. All the models were designed as point-based. The existence of high-quality, relevant agronomic data, originated from field observations, is of high importance for reliable crop yield simulations (Grassini *et al.*, 2015). WOFOST (WOrld FOod STudies) was one of the first crop simulation models that combined major components (weather variables, soil maps, agro-ecological zones with cropping calendars and crop cultivars) for applying a crop simulation for multiple locations (Boogaard *et al.*, 2014). Aggregations to higher spatial levels, i.e. grid or regional level, are obtained by averaging the simulated variables, weighted on the area of the specific soil type or land use cover, respectively. The model simulates yield at different production levels; potential yield (limited only by crop features and radiation, temperature and atmospheric CO_2 concentration) and water-limited production. Nutrient-limited growth is not included in the model since dynamic feedback between model and nutrients is missing. Over the last 25 years, WOFOST has been used operationally for crop yield prediction; new applications for precision and smart farming are expected in the future (De Wit *et al.*, 2019).
The concept of precision farming takes into consideration the within-field spatial and temporal variability with the scope to optimize yield through more precise application of inputs, minimizing the cost and the environmental impact and making traceability of the agricultural activities possible (Stafford, 2000). The relationship between yield variability and topographic attributes such as

DOI 10.3920/978-90-8686-888-9_10, © Wageningen Academic Publishers 2019

elevation, slope, curvature and flow accumulation has been previously explored (Kravchenko and Bullock, 2000). Maize yield showed a strong dependency on topography, with high correlation coefficients between yield and topographic indices that reflect field water availability, such as the wetness index and distance to flow accumulation lines (Marques da Silva and Silva, 2008). The interaction of within-field variability of topographic attributes and rain patterns is an important driver of yield (Maestrini and Basso, 2018). The variation in soil water retention properties was found to have an effect to yield variation (Hakojärvi *et al.*, 2013). However, none of the above-mentioned studies introduced within-field spatial variability into crop growth model.
In this study, the WOFOST crop growth model was used to predict yield at a within-field level. To achieve this, detailed spatial data were used concerning field properties, stressing the soil water flow accumulation as one of the most important factors affecting the yield.

Materials and methods

Experimental field
A 64 ha field was selected located in Vojvodina region (45.82947° N, 19.52431° E), one of the most important regions in terms of agricultural production in the Republic of Serbia. The field was cultivated with maize in 2017 and managed using variable rates of inputs based on 25 management zones delineated by the farmer according to advisor's suggestions (FMZs). The farmer was using center pivot irrigation system.

Yield monitoring and water flow accumulation
The harvest was logged by two Ag Leader (Incommand 800; AgLeader Technology, IA, USA) grain harvest monitors installed on each of the two grain maize harvesters used to harvest the field. The system consisted of a mass flow sensor (impact style) with an accuracy of approximately 4% (readings corrected by the integrated humidity sensor), and a RTK GNSS system (Trimble AGGPS 542; Trimble Inc., CA, USA) with an accuracy of 8 mm horizontally and 15 mm vertically. The yield shape file was transformed into a raster with a reasonably detailed resolution of 1 m^2 by bilinear interpolation in QGIS (QGIS Development Team, 2018). Similarly, the terrain elevation data provided by the GPS RTK were processed to obtain the digital elevation model (DEM) and then calculate the slope, orientation, curvature, flow direction and accumulation. Flow direction for each grid cell of the DEM was calculated using the direction of steepest descent to one of its adjacent or diagonal neighbors. Flow direction is reported as NODATA for any grid cell adjacent to the edge of the DEM domain, or adjacent to a NODATA value in the DEM. In flat areas, flow directions were assigned away from higher ground and towards lower ground using the method of Garbrecht and Martz (1997). The accumulation of flow was created in QGIS by the routine r.accumulate. It is a GRASS routine (GRASS Development Team, 2017), accessible through the corresponding plugin. The module only uses a flow direction map to trace and accumulate the amount of flow draining through and including each cell.

Soil apparent electrical conductivity
The mapping of the soil apparent electrical conductivity (ECa) was carried out using the EM38-MK2 (Geonics Limited, ON, Canada) at vertical dipole orientation. The measurement provided information of the soil spatial variability at two depths, 1.5 m and 0.75 m. For the georeferencing of the ECa measurements a RTK GNSS was used providing 20 mm accuracy. The raw soil ECa data were analyzed using Fuzzy Clustering by the MZA (Management Zone Analyst) software (Fridgen *et al.*, 2004) to delineate the field into three fuzzy cluster ECa zones.

Soil sampling

The soil ECa zones were used to perform directed soil sampling. Samples were taken in every ECa zone at six sampling points (repetitions) from the plow layer (0-300 mm) using soil probe. Particle size distribution was determined by the pipette method using tetrasodium pyrophosphate decahydrate as a dispersing agent.

Two or three locations in each FMZ were selected based on the flow accumulation parameter (high, low and/or medium accumulation categorized based on quantiles). Then, a 6 m buffer was created for each selected point. The average flow accumulation and measured yield were calculated in each buffer. In addition, the soil texture from analysis of the soil samples that were taken according to the fuzzy clustering ECa zones was assigned to each buffer.

Crop growth model

In this study, WOFOST version 7.1.3 was used. To properly run the model, it is necessary to have an appropriate crop, weather and soil data file (Boogaard *et al.*, 2014). For crop data, the MAG203.CAB file (Grain Maize 203) was used, since it is designed for maize and calibrated for the region of central France and northern Italy, where the experimental field is at the corresponding latitude and belonging to the same climate zone (Kottek *et al.*, 2006). Weather files were prepared based on the meteorological data observed in 2017 at the Sombor meteorological station, which is a part of the official observing network of the Republic Hydrometeorological Service of Serbia and the nearest to the field. Additional weather data were collected from the National Oceanic and Atmospheric Administration (NOAA, 2018). WOFOST does not provide a module to include irrigation in the system calculations. Therefore irrigation was introduced to the system as precipitation following the exact irrigation schedule and amount of applied water (25 mm). Total amount of water (precipitation + irrigation) was then multiplied with the flow accumulation factor, calculated for all the selected buffers in the field, resulting in 46 weather files. Based on the soil texture obtained by the soil sampling, parameters such as field capacity, wilting point and saturated hydraulic conductivity were calculated and presented in the soil files (three slightly different files).

Results and discussion

Observed yield

Observed yield showed significant variability at both field and sub-field level. There was a large yield gap between the areas inside and outside the irrigated area (Figure 1) demonstrating the significance of water availability in maize productivity.

In general, the areas with high elevation (Figure 2A) and the ones with high slope (Figure 2B) showed low water accumulation which resulted in lower yield. This is in agreement with the findings of Calvino *et al.* (2003), who found that most of the variation in yield (>84%) was accounted for by water availability. In this study, the relationship would be even more evident if the field was managed homogeneously, expressing in full the effect of soil and landscape variability in the final yield. However, due to the variable rate applications of fertilizer that occurred based on the FMZs, a new factor was introduced into the system that affected maize yield according to the soil and landscape spatial variability.

Management practices were homogeneous within each FMZ. However, yield zonal statistics (Table 1) indicated that within-zone spatial variability of yield was high. In fact, in many cases the within-zone variability was higher than the variability for the whole field. This result indicates that there is an important yield-driving factor that has not been taken into account in the FMZs. Yield spatial variability of the irrigated area, was significantly higher compared to the non-irrigated area (variance was 4.18 and 2.25 respectively; Table 1). This is explained by the extreme water scarcity in the non-irrigated areas due to severe drought conditions that occurred in summer 2017, which minimized the yield potential and thus range for variability of these areas.

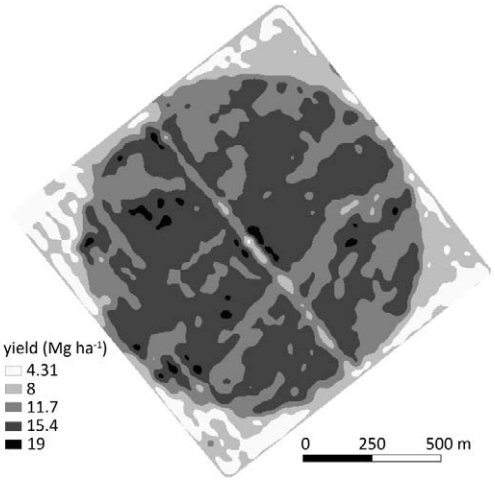


Figure 1. Yield map produced by interpolation of point data from the yield monitors.

Table 1. Zonal statistics of yield (Mg/ha) for each FMZ and for the total irrigated and non-irrigated area. N stands for the number of points within the zone. Zone IDs are shown in Figure 2C.

Irrigation	Zone	N	Range	Min.	Max.	Mean	STD	Var.
Irrigated	1	944	15.42	1.02	16.44	12.21	1.72	2.96
Irrigated	2	1749	15.82	0.65	16.47	12.25	2.20	4.82
Irrigated	3	1737	12.20	4.25	16.46	12.50	2.02	4.09
Irrigated	4	1334	13.48	2.98	16.46	11.27	2.16	4.68
Irrigated	5	3075	15.84	0.66	16.50	11.54	2.05	4.19
Irrigated	6	2642	15.82	0.64	16.46	12.00	2.55	6.52
Irrigated	7	2094	11.73	4.73	16.46	12.56	1.60	2.56
Irrigated	8	2592	15.64	0.85	16.50	12.49	1.56	2.44
Irrigated	9	2018	15.77	0.71	16.48	10.80	2.76	7.64
Irrigated	10	3499	9.30	7.17	16.48	12.34	1.63	2.65
Irrigated	11	1330	9.63	6.75	16.38	12.55	1.86	3.45
Irrigated	12	1223	12.22	4.24	16.46	12.86	1.40	1.96
Irrigated	13	1716	15.72	0.77	16.49	11.63	1.95	3.78
Irrigated	14	1203	15.16	1.17	16.33	11.42	1.78	3.15
Irrigated	15	2228	9.89	6.58	16.46	12.26	1.57	2.47
Irrigated	16	2097	13.07	3.40	16.47	11.28	1.50	2.26
Irrigated	17	2445	15.29	1.20	16.49	11.01	2.15	4.63
Total irrigated		33926	15.86	0.64	16.50	11.92	2.04	4.18
Non-irrigated	18	1013	11.57	0.71	12.28	4.58	1.86	3.44
Non-irrigated	19	1950	12.73	0.65	13.37	4.14	1.63	2.67
Non-irrigated	20	875	11.73	0.68	12.40	4.28	1.56	2.44
Non-irrigated	21	2227	10.22	0.66	10.88	4.78	1.21	1.47
Non-irrigated	22	2151	10.29	0.66	10.94	4.69	1.34	1.79
Total non-irrigated		8216	12.73	0.65	13.37	4.53	1.50	2.25

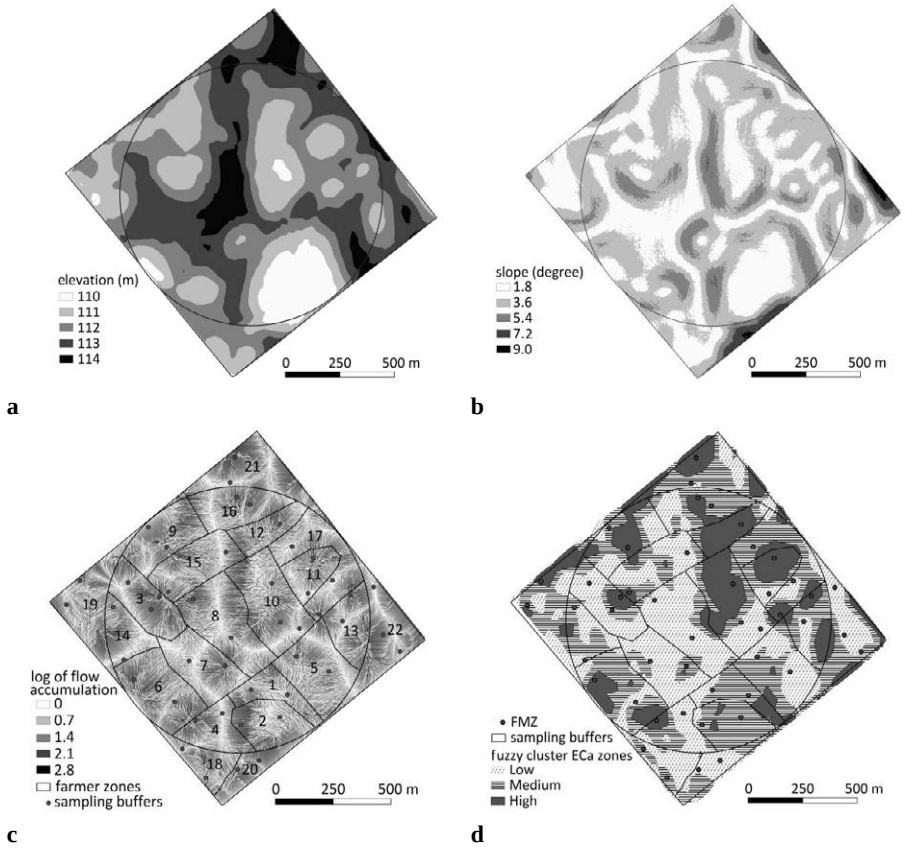


Figure 2. Maps of elevation (A), slope (B) and flow accumulation (log) (C) and ECa zones produced using fuzzy clustering (D) both overlaid with the farmer's management zones (FMZs) and the sampling buffers used for the analysis. Numbers inside map (C) denote the FMZ id.

The fuzzy cluster ECa zones produced by the analysis of the ECa measurements (Figure 2D), were similar to the elevation (Figure 2A), slope (Figure 2B) and flow accumulation (Figure 2C) maps. This indicates that soil water content was a key parameter affecting soil electrical conductivity. Therefore, the study was focused on the yield driven by the flow accumulation on FMZs basis.

Simulated yield

The average flow accumulation was imported in WOFOST and the yield predicted by the system was compared with the average measured yield per buffer (Figure 3). The relationship between flow accumulation and model predicted yield was described by a quadratic plateau model, showing that the effect of water availability was high at the low yielding areas and decreased after a certain level as the maximum yield potential of the field as calculated by the system was reached (Figure 3B). Yield predicted by the WOFOST model was well correlated to the measured yield (R^2=0.79, RMSE=4.2, D-Index=0.78; Figure 3A). Statistics of the linear regression between measured and simulated yield are presented in Table 2. Results indicate the high importance of the flow accumulation as a factor which provides more accurate yield prediction at a subfield level, when using crop modelling systems.

Table 2. Results of the linear regression between measured and simulated yield.

	Regression coefficients	*P*-value	95% confidence level interval	
			lower limit	upper limit
slope	0.948	0.000	0.800	1.097
constant	-3.190	0.000	-4.832	-1.549

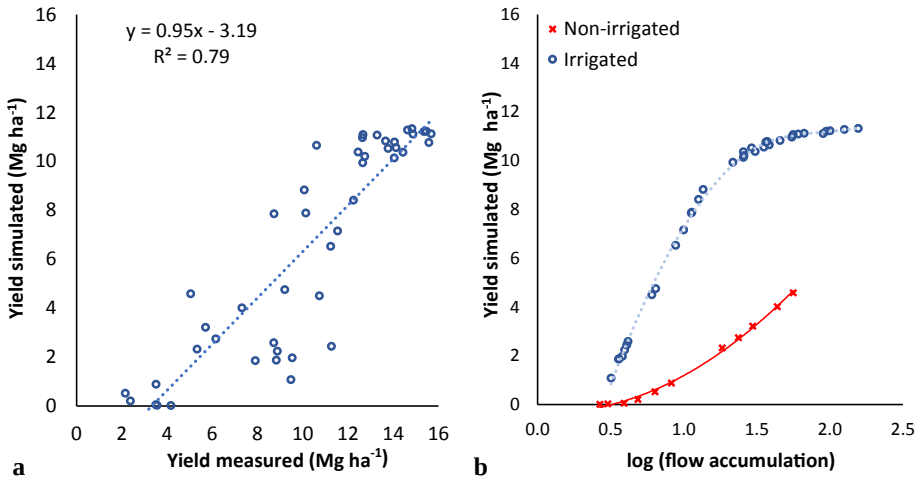


Figure 3. Relationship between measured yield and yield simulated by WOFOST crop model (A) and between the logarithm of flow accumulation and yield simulated by WOFOST crop model (B) for the 46 selected areas of the analysis (6 m buffers).

In addition, this result shows that even though such systems are designed to run simulations on field and regional level, they show great potential to be used on a sub-field level if taking into account the within-field variability. To do so, the key factors introducing spatial variability in the field, such as slope and flow accumulation, need to be introduced into the modelling process. However, WOFOST model underestimated measured yield (constant=3.2 Mg/ha; Table 2), due to clear limitation of the model, which does not include a component to simulate the effect of nutrients on the crop growth (De Wit *et al.*, 2019).

Conclusions

The current study focused on the implementation of field spatial properties into a crop modelling system, evaluating the potential to provide yield estimation at sub-field level, with positive results. Yield predicted at zone level by the WOFOST model was correlated with the actual yield measured by the yield monitors. The proof of concept was confirmed indicating that if the spatial variation of the key yield-driving parameters is imported in the system calculations, accurate prediction of yield can be achieved at within-field level using crop modelling systems. Water availability driven

by the slope and flow accumulation was proven essential in the description of field variability which, in turn, explained the spatial variability of maize yield.

Acknowledgements

This project was funded by ANTARES project (no. 664387) that has received funding from the European Union's Horizon 2020 research and innovation programme.

References

Boogaard, H.L., de Wit, A.J.W., te Roller, J.A. and Van Diepen, C.A., 2014. WOFOST Control Centre 2.1; User's guide for the WOFOST Control Centre 2.1 and the crop growth simulation model WOFOST 7.1.7. Wageningen (Netherlands), Alterra, Wageningen University & Research Centre.

Calvino, P. A., Andrade, F. H., and Sadras, V. O., 2003. Maize Yield as Affected by Water Availability, Soil Depth, and Crop Management. Agronomy Journal 95 275-28.

De Wit, A., Boogaard, H., Fumagalli, D., Janssen, S., Knapen, R., van Kraalingen, D. *et al.,* 2019. 25 years of the WOFOST cropping systems model. Agricultural Systems 168 154-167 https://doi.org/10.1016/j.agsy.2018.06.018.

Fridgen, J.J., Kitchen, N.R., Sudduth, K.A., Drummond, S.T., Wiebold, W.J. and Fraisse, C.W., 2004. Management zone analyst (MZA): software for subfield management zone delineation. Agronomy Journal 96 100-108.

Garbrecht, J. and Martz, L. W., 1997. The Assignment of Drainage Direction Over Flat Surfaces in Raster Digital Elevation Models. Journal of Hydrology 193 204-213.

GRASS Development Team, 2017. Geographic Resources Analysis Support System (GRASS) Software, Version 7.2. Open Source Geospatial Foundation. Electronic document: http://grass.osgeo.org

Grassini, P., van Bussel, L.G.J., Van Wart, J., Wolf, J., Claessens, L., Yang, H. *et al.,* 2015. How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. Field Crops Research 177 49-63.

Hakojärvi, M., Hautala, M., Ristolainen, A. and Alakukku, L., 2013. Spatial and temporal yield variation in three different clay soil fields. IFAC Proceedings Volume 46 (18) 196-201.

Holzworth, D.P., Snow, V., Janssen, S., Athanasiadis, I.N., Donatelli, M., Hoogenboom *et al.,* 2015. Agricultural production systems modelling and software: Current status and future prospects. Environmental Modelling & Software 72 276-286.

Kottek, M., Grieser, J., Beck, C., Rudolf, B. and Rubel, F., 2006. World Map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift 15 259-263.

Kravchenko, A. N. & Bullock, D. G., 2000. Correlation of corn and soybean grain yield with topography and soil properties. Agronomy Journal 92 75-83.

Maestrini, B. and Basso, B., 2018. Drivers of within-field spatial and temporal variability of crop yield across the US Midwest. Scientific Reports 8:14833 1-9.

Marques da Silva, J.R. and Silva, L.L., 2008. Evaluation of the relationship between maize yield spatial and temporal variability and different topographic attributes. Biosystems Engineering 101(2) 183-190.

NOAA, 2018. National Oceanic and Atmospheric Administration. https://www.ncdc.noaa.gov/ghcnd-data-access (last accessed November 1st 2018)

QGIS Development Team, 2018. QGIS Geographic Information System, Geospatial Foundation Project https://qgis.org/en/site (last accessed February 14th 2019)

Stafford, J.V., 2000. Implementing precision agriculture in the 21st century. Journal of Agricultural Engineering Research 76(3) 267-275.

Empowering farmers by resolving the trust and legal issues emerging from precision farming

L. Wiseman[1*] and J. Sanderson[2]
[1]*Griffith Law School, Griffith University, Nathan, 4111, Queensland, Australia; wiseman@griffith.edu.au*
[2]*USC Law School, USC Australia, Sippy Downs QLD 4556, Australia*

Abstract

While many laws potentially influence the ownership, control of and access to ag data, there are many reasons behind farmers' reluctance to engage in the sharing of their farm data. One key concern, the lack of transparency and clarity in data licences, has led to mistrust in the ag data relationship. While voluntary ag data codes of practice address some of the concerns, the recent expansion of privacy laws has been more effective in highlighting the responsibilities that data aggregators have in relation to the data they manage. To build trust in ag data sharing, farmers need to be empowered with tools and knowledge to help them change the current imbalance in their ag data relationships.

Keywords: licences, transparency, codes, privacy, trust

Introduction

Farmers' concerns about the way in which their farm data is being collected, managed and shared have now become a vital part of the dialogue of how best to realise the potential of precision (or smart farming) technologies. One of the first detailed surveys that specifically highlighted farmers' attitudes and concerns about sharing their farm data was done by the American Farm Bureau in 2015 (American Farm Bureau Federation, 2016). Since then there has been increasing attention focussed on these issues (Bronson and Knezevic, 2016; Wolfert *et al.*, 2017; Kosior, 2018; Wiseman and Sanderson, 2017). At an international level, the focus on farmers' concerns about data sharing has also been raised in the context of promoting food security through the opening of access to agricultural data in developing nations. Of note is the recent work on the broader ethical, legal and policy aspects of open data affecting smallholder farmers that began in 2018 by the Global Open Data Agriculture and Nutrition (GODAN), Global Forum on Agricultural Research (GFAR) and the Technical Centre for Agricultural and Rural Cooperation ACP-EU (CTA) and the Kuratorium für Technik und Bauwesen in der Landwirtschaft (KTBL) (Global Open Data Agriculture and Nutrition (GODAN) *et al.*, 2017).

This paper argues that all farmers would benefit from a cultural shift in the way that the relationships that govern agricultural data and information generated on farm is thought about. Of vital importance to farmers is the fact that their farm data and the risks and liabilities associated with its aggregation are managed well. While some argue that the incremental risks of making farm and agricultural data more available might appear very small (given how much agricultural data is already in the public domain), incentives and trust on the part of the data contributors, i.e. the farmers, nevertheless have to be maintained (Australian Productivity Commission, 2017). Of key concern to farmers, as reiterated and confirmed by a whole of agricultural industry survey undertaken as part of the Accelerating Precision Agriculture to Decision Agriculture (Zhang *et al.*, 2017) in Australia in 2017, is the lack of transparency and clarity around issues such as data ownership, portability, privacy, security, trust and liability. This survey involved, amongst other matters, ascertaining attitudes towards the sharing of farm data by 1000 producers and farmers across all seventeen agricultural industries in Australia (Zhang *et al.*, 2017). Put simply, the results revealed that these issues lie at the very heart of the lack of trust that farmers experience when

DOI 10.3920/978-90-8686-888-9_11, © Wageningen Academic Publishers 2019

their farm data is collected, aggregated and shared. The fact that the current legal and regulatory frameworks around the collection, management, sharing and use of agricultural data are at best ad hoc and piecemeal and, at worst, absent altogether, facilitates these concerns expressed by farmers. This paper suggests that from these learnings, fundamental and systematic changes are needed to the way that farmers' data are currently being managed by agribusinesses, researchers and governments. The reason for this is that the nature of sources of farm data and the way in which analytical techniques are rapidly evolving means that the data being harvested from farms is moving away from any effective control by individual farmers. Agribusinesses, as much as governments and researchers rely on the willingness of farmers – the source of so much of the agricultural data – to continue to trust the way in which they manage agricultural data collection and use.

With the vast push for open data in agriculture, there is now, and will be in the future, a need for continued community acceptance and trust of farmers in the handling of their data by agribusinesses, researchers and governments. For farmers to 'have a sound basis for believing in the integrity and accountability of entities (public and private) handling data, they need to feel that they have some control over how their own data is used and by whom, and an inalienable ability to choose to experience some of the benefits of these uses themselves.' (Australian Productivity Commission, 2017) This will also have the added benefit of promoting better understanding of the potential community-wide benefits of agricultural data use and reuse.

Farmers' concerns

Of key concern to many farmers is the general lack of transparency and clarity around issues such as data ownership, portability, privacy, security, trust and the allocation of risk and liability. Farmers' concerns about these issues can be traced back to the impact that the expansive terms of use in the data licences have upon them when they enter into agreement to adopt smart or digital farming technologies. The concerns over the terms of use that are found in the overly complex data licences that govern their farm data can be grouped into three major themes: (1) A lack of transparency about the terms of use in the data licences, particularly in relation to who may have access to the data that is being shared; (2) Inequality of bargaining power in data licences; and (3) A lack of benefit-sharing of the value of the data collected between farmers and agribusinesses. Each of these will be examined in turn before turning to some strategies that are aimed at resolving these concerns.

A lack of transparency about the terms of use in the data licences, particularly in relation to who may have access to the data that is being shared

One of the major concerns identified by farmers is the lack of transparency over the terms of the data licences that govern the use of their agricultural data. Australian farmers who were surveyed expressed concern about the lack of information they were given about data ownership, control and sharing prior to entering contracts with agri-businesses. Overall, the majority (74%) of respondents did not know much about the terms and conditions relating to data collection in their agreement with service providers. Specifically, 47% of farmers surveyed said they had no understanding and an additional 27% said they had little understanding of the terms and conditions of data licence agreements before signing up to a new software or service, particularly where the service is provided online (Zhang *et al.*, 2017; p. 42).

The fact that so many farmers are unaware of the terms that govern the ownership and use of, and access to, their data indicates that farmers do not understand the implications of what they are signing, they are often unaware of how much control the agricultural service provider is asserting over their data or the extent to which their farm data is being shared and traded. This has significant consequence for the agricultural industry as it forms the basis for the lack of trust that farmers have towards some of the new digital services and products on offer.

Like many licences embedded in software enabled devices today, most data licences involve the use of a 'click wrap' agreement (where the click of an 'I agree' icon, the downloading of an app or the starting of an engine signifies consent to the terms and conditions of a licence), and this is often the way farmers enter into and agree to data licences for smart farming technologies. Despite the ease at which we can agree to various terms and conditions, the data licences that are embedded in digital agricultural technologies are generally complex standard-form licence agreements that are generally non-negotiable and presented on a 'take it or leave it' basis when the technology is adopted. The terms of use of the technology are therefore agreed to either at the time of downloading an app or turning on a machine. Each different agricultural industry has an industry-focused range of commercial relationships. The more vertically integrated the industry, the tighter the contractual relationships tend to be. This is when knowledge of all of the terms of the licence becomes an important issue, as it is when the contract between the farmer and the agricultural technology provider is formed.
Often a data licence will also provide links to other policy documents such as the agricultural technology provider's privacy policy. In some cases, it is the privacy policy rather than the terms of use of the data licence that outlines who may have access to the data generated under the agreement. What has been recognised as 'one of the most potentially pernicious practices with data is the onward trade or disclosure of data to third parties, leaving individuals unaware of who knows what about them and their business. The damage is often not so much in monetary terms but in the feeling of exploitation. This has great capacity to undermine social licence over time, if misused.' (Australian Productivity Commission, 2017; p. 18).

Inequality of bargaining power

The digital revolution that has occurred through the adoption of precision agriculture on farms has introduced a new type of commercial on-line relationship into the agricultural supply chain. Whereas in the past, many commercial relationships concerning the delivery of goods and services on farms arose from face to face interactions with agronomists and advisers who were well known and trusted by generations of farming families, this has now changed. Precision agricultural technologies have brought a new way of transacting with agri-business. Adopting new smart farming technologies and services on farm is not unlike joining a social media platform. The simple 'click' of a mouse when downloading software or the starting of a machine is now the way that farmers enter into a myriad of online transactions for their digital goods and services. Most precision agricultural technologies now embedded in smart farming machinery and programs are governed by typical software licences. These are without doubt, overly complex, peppered with legal jargon and difficult for lay people to understand. As opposed to the discussions that precede face to face transactions with local advisers and agribusinesses, what this means is that often farmers now find themselves downloading services and software without any prior detailed discussion of the terms of use of the data that is being gathered by that particular technology.
The presentation of standard, pre-printed online contracts on a 'take it or leave it basis', reflects a clear inequality of bargaining power between the contracting parties. This is often referred to as the digital divide – a divide between those who contribute the data and those who control, aggregate and share the data. The power imbalance between data contributors and data aggregators is evidenced by the inability of farmers to negotiate the standard terms of the large agri-business' data licences that govern the agricultural technology, and is well accepted (Carbonell, 2016). In addition to the fact that one party sets all of the terms and the other has very little choice but to agree, the size, scale and multinational operations of many of the large agricultural technology providers compounds this digital divide. This is particularly the case with even more new entrants into the agricultural supply chain and market by technology providers with little or no experience in agriculture.
The fact that many large ag technology providers are also foreign owned, is another important factor when examining the level of trust and confidence that farmers have in the way their farm data is

currently being managed. Often the licence agreements that contain the terms of use that govern the data sharing will be governed by the law of the country where that company is registered and this creates uncertainty over the level of protection afforded to farmers. For example, farmers may not have the benefit of consumer or other legislation of their own country, where the dispute is taken to an overseas jurisdiction.

A lack of benefit-sharing between farmers and agribusinesses

Farmers are also concerned that others are getting greater value from their farm data than they are. Though the issue of who gains the most benefit from agricultural data sharing is one that remains unclear. Many farmers are concerned that advisers or agri-businesses derive the greatest benefit from their data. While recognising the value added by third-party aggregators, farmers recognise that they are the original contributors of their agronomic data, and are thus responsible for the integrity and quality of the data that is later aggregated and analysed. Farmers are concerned that businesses are making money from their data without sharing any of the benefits with the original data contributors. For example, the majority of Australian farmers surveyed who responded (67%) that they would not be comfortable if service providers make profit out of their data without sharing the profit with them (Zhang *et al.*, 2017; p. 75). Some 38% of farmers surveyed were extremely concerned that businesses are making money from their data without sharing it with the producer (Zhang *et al.*, 2017; p. 45).

While there were some examples of co-operative benchmarking exercises that returned benefits to member farmers, many farmers expressed concern that the real value of sharing data was not yet being realised. While it is often said that improved knowledge, products and services is the real benefit of digital agriculture being returned to farmers, many farmers said they were uncomfortable about the current business models where they contribute their data for nothing but then pay full cost for the services delivered from the aggregated data (Zhang *et al.*, 2017).

Addressing farmers' concerns

When thinking about strategies to improve the equity and transparency of ag data licences, it is useful to consider the number of, and inconsistency between, data licences entered into by farmers. The sheer number of transactions, each with their own terms of use makes the issue of the ownership and the sharing of data even more complex. Some of the contractual relationships that farmers enter that result in the collection and collation of agricultural data include: chemical/fertiliser suppliers; broader service providers (e.g. telecommunications, sensors, soil testing, drones, etc., agri-technology/agri-business providers, and third parties and professional advisers (e.g. agronomists, contractors and advisers). Once the breadth of farm transactions and the complexity of the contractual relationships are understood, farmers can be empowered by the provision of tools, knowledge and training about data ownership, control and access rights. This knowledge will equip them to engage in dialogues with agribusinesses, researchers and agricultural stakeholders. There are already a number of tools available to farmers, for example, lists of questions that should be asked before entering into data licences to satisfy themselves of the trustworthiness of the agribusinesses or technology providers (Pennsylvania Farm Bureau, 2015). Yet further work is needed on providing additional support to farmers, both on-line and in face to face forums, on their ability and their right to ask questions of their technology suppliers and to better understand and negotiate the terms of use of their data.

There have been attempts to address the inequality of bargaining power in the agricultural digital divide in Australia, where there has been a ramping up of competition law protections for agricultural matters. For example, there is now agricultural expertise within the Australian Consumer and Competition Commission, with the appointment of an Agricultural Commissioner. This role focuses on a range of potential anti-competitive practices with the agricultural sector. Legislative support

has also been afforded to small farming businesses through amendments made to unfair terms legislation within consumer protection laws aimed at 'small businesses' that employ fewer than 20 persons, which would incorporate a large number of farming enterprises (Wiseman, 2016). This reform does indicate that the practice of using standard form contracts comes with a responsibility on the part of the larger businesses to ensure that their contractual terms are transparent and fair in the sense that they do not go beyond what is necessary to protect their interest and that they do not create a significant imbalance in the parties' rights and obligations. These reforms again go some way towards empowering farmers in the way in which they enter into contractual relationships about their farm data.

Another strategy to improve ag data contracting practices, largely supported by farming organisations and industry bodies, has been the adoption and uptake of voluntary codes of practice.

The emergence of voluntary agricultural data codes of practice

Since 2014, voluntary agricultural data ('ag-data) codes of practice have been developed to not only help develop 'good' data management practices, but also build trust in the way ag-data is managed (Sanderson *et al.*, 2018). Ag-data codes of practice have emerged in part because of the increasing realisation of the potential benefit and value of ag-data: with many decisions and processes along the whole agri-food supply chain, from paddock to plate, being data enabled and data driven (Wolfert *et al.*, 2017; Kamilaris *et al.*, 2017; Coble *et al.*, 2018). Ag-data codes of practice are implemented beyond legal and policy mandates and attempt to both harness the benefits of ag-data as well as protect farmers and producers.

More specifically, data codes tend to focus on the key areas giving rise to mistrust ie consent, disclosure, security and transparency around ag-data practices. For example, under the New Zealand 'Farm Data Code of Practice' ('NZ Data Code') 2014, organisations agree to disclose their practices and policies around data rights, data processing and sharing, and data storage and security. In the United States, the American Farm Bureau Federation's 'Privacy and Security Principles for Farm Data' sets out data principles for agricultural technology providers including that 'access and use of farm data should be granted only with the affirmative and explicit consent of the farmer.' (American Farm Bureau Federation, 2014b) The 2018 Copa Cogeca EU Code on Agricultural Data Sharing by Contractual Agreement ('EU Code') attempts to define key concepts and sets out general principles for sharing agricultural data including that '[t]he collection, storage and usage of the collected agricultural data can only occur once the data originator has granted their explicit, express and informed permission via contractual arrangement'(Copa-Cogeca, 2018). Other countries, such as Australia, are also looking to introducing a whole of industry ag data governance framework, which may or may not include an ag-data code of practice.

It is important to give credit to the impact and attention that these codes have had on what is best practice in ag data management; however, while they remain voluntary, they may not be able to achieve the seismic shift that is needed within the industry to regain trust in data relationships.

The rise of privacy rights

In contrast to the voluntary ag-data codes of practice, the implementation of the EU's General Data Protection (GDPR) on 25 May 2018 has without doubt, turned the world's attention on the way data aggregators are managing the personal data they collect. The informed consent of the individual, the right of portability and the right to be forgotten are now required to be integral considerations in data management practices.

While it is not exactly clear what impact the GDPR will have on the management of agricultural data (given that some of the data will involve personal information but some may not), it has certainly shone a spotlight on the need for increased responsibility by data aggregators towards

data management. The key features of informed consent, the right of portability and the right to be forgotten introduced by the GDPR are also being implemented into a range of privacy or data protection laws in other countries. The Australian Government, for example, has committed to the introduction of a broad Data Sharing and Release regime, including a consumer data right which will give consumers similar rights. The proposed Consumer Data Right will be rolled out industry by industry, with the agricultural industry waiting in line behind banking, energy and telecommunications.

With the ramping up of privacy and data rights for individuals across the world, now is the time for an open dialogue with farmers, agribusinesses, researchers and Governments about what constitutes best practice in agricultural data governance.

Conclusions

This paper concludes that the best way to address farmers' concerns about the way in which their data is currently managed is to empower farmers (and the agricultural industry more broadly) with knowledge and tools that will enable them to engage in discussions about how to achieve best practice in ag data governance. It is possible, that with the right knowledge and tools, farmers and their agricultural stakeholders can alter the social architecture of their ag data relationships to achieve a more open, transparent and balanced relationship. While empowering farmers with knowledge and skills about what good data management looks like, the question is, what role can the broader agricultural industries play in offering a genuine two-way support of farmers' continuing willingness to supply their data, the crucial input to agricultural business development, research and public policy? Continued engagement in not only digital literacy but also data and legal literacy will assist in the re-skilling of the agricultural industry. Knowledge and enhancement of consumer protection rights – such as unfair terms in contracts safeguards, transparency of the terms of collection, management and use of agricultural data, and effective management of risk – will enhance the operation of the agricultural industry as a whole. This would mean that farmers would no longer be perceived as just sources of data, but they would rank equally with the key agricultural data aggregators – agri-businesses, researchers and governments – in being able to trade and use their data. Trust and acceptance of good data management practices within agricultural and farming communities will be vital in achieving a re-calibration of the ag data relationship and thus will benefit the broader ag data supply chain infrastructure.

Acknowledgements

The authors would like to thank our anonymous reviewers for the constructive feedback that helped improve our paper.

References

American Farm Bureau Federation, 2014 b. Privacy and Security Principles for Farm Data. https://www.fb.org/issues/technology/data-privacy/privacy-and-security-principles-for-farm-data (last accessed November 2018).

American Farm Bureau Federation, 2014a. American Farm Bureau Survey Shows Big Data Use Increasing, Big Questions Remain. https://www.fb.org/newsroom/american-farm-bureau-survey-shows-big-data-use-increasing-big-questions-rem (last accessed January 2019).

American Farm Bureau Federation, 2016. Farm Bureau Survey: Farmers Want to Control Their Own Data. American Farm Bureau Federation, Washington, D.C., USA https://www.fb.org/newsroom/farm-bureau-survey-farmers-want-to-control-their-own-data (last accessed November 2018)

Australian Productivity Commission, 2017. Data Availability and Use, Final Report, https://www.pc.gov.au/inquiries/completed/data-access#report, (last accessed March 2019).

Bronson, K., Knezevic, I., 2016. Big Data in food and agriculture. Big Data & Society 3 (1), 1-5.

Carbonell, I. 2016. The ethics of big data in agriculture. Internet Policy Review 5 (1) https//doi.org/10.14763/2016.1.405

Copa-Cogeca, 2018. EU Data Code of Conduct on agricultural data sharing by contractual agreement, https://copa-cogeca.eu/img/user/files/EU%20CODE/EU_Code_2018_web_version.pdf (last accessed December 2018).

Global Open Data for Agriculture and Nutrition (GODAN), Global Forum on Agricultural Research (GFAR) and the Technical Centre for Agricultural and Rural Cooperation ACP-EU (CTA) and the Kuratorium für Technik und Bauwesen in der Landwirtschaft (KTBL), 2017. The Ethical, Legal and Policy aspects of data sharing affecting small lot farmers. http://www.fao.org/e-agriculture/news/highlights-e-consultation-ethical-legal-and-policy-aspects-data-sharing-affecting-farmers (last accessed November 2018).

Kamilaris, A., Kartakoullis, A., and Prenafeta-Boldú, F., 2017. A Review on the Practice of Big Data Analysis in Agriculture. Computers and Electronics in Agriculture 143 23-37.

Kosior, K., 2018. Digital Transformation in the Agri-Food Sector- Opportunities and Challenges. Scientific Annals of the Association of Agricultural and Agribusiness Economists, 20 (2), 98-104. https//doi.org/105604/01.3001.0011.8122

New Zealand Farm Data Standards, 2016. Building Consistency- Common Data Vocabularies promoting efficient data exchange. New Zealand Farm Data Standards. http://www.farmdatastandards.org.nz (last accessed December 2018).

Pennsylvania Farm Bureau (PFB), 2015. Ponder These Nine…Before You Sign. https://www.pfb.com/images/issues/Ponder-These-Nine.pdf (last accessed November 2018).

Sanderson, J., Wiseman, L., Poncini, S., 2018. What's behind the Ag-data logo? An examination of voluntary agricultural data codes of practice. International Journal of Rural Law and Policy, 7 November 2018 https://epress.lib.uts.edu.au/journals/index.php/ijrlp/article/view/6043 (last accessed November 2018).

Wiseman, L. and Sanderson, J., 2017. The Legal Dimensions of Digital Agriculture in Australia: An Examination of the Current and Future State of Data Rules Dealing with Ownership, Access, Privacy and Trust. Griffith University, USC Australia and Cotton Research and Development Corporation, Australia. https://www.crdc.com.au/sites/default/files/P2D%20Legal%20Dimensions%20-%20Griffith%20USC%20Final%20Report.pdf (last accessed November 2018).

Wiseman, L., 2016. Changes to contracts laws could give small farming businesses more control of data and innovation, The Conversation. http://theconversation.com/changes-to-contract-laws-could-give-small-farming-businesses-more-control-of-data-and-innovation-69275 (last accessed January 2019).

Wolfert, S., Ge, L., Verdouw, C., Bogaardt, M.J., 2017. Big Data in Smart Farming – A review. Agricultural Systems 153, 69-80.

Zhang, A., Baker, I., Jakku, E. and Llewellyn, R., 2017. Accelerating precision agriculture to decision agriculture: The needs and drivers for the present and future of digital agriculture in Australia. A cross industries producer survey for the Rural R&D for Profit 'Precision to Decision' (P2D) project. CSIRO and Cotton Research and Development Corporation, Australia, 45, Fig 40. https://www.crdc.com.au/sites/default/files/P2D%20producer%20survey%20-%20CSIRO%20Final%20Report.pdf (last accessed November 2018).

Section 2
Precision horticulture

3D point clouds from UAV imagery for precise plant protection in fruit orchards

M. Hobart*, M. Schirrmann and M. Pflanz
Leibniz Institute for Agricultural Engineering and Bioeconomy, Max-Eyth-Allee 100, 14469 Potsdam, Germany; mhobart@atb-potsdam.de

Abstract

Precise application of plant protection products (PPPs) is a challenging task for fruit growing. In future, the application should be adapted to the leaf area for improving the safety of food, user and environment. This requires site-specific information of fruit wall geometries. To achieve this, structural tree parameters were determined in advance by using autonomously flying robots (UAV) and photogrammetric methods. Results showed that it is possible to extract tree heights from 3D point clouds of an apple orchard using an individually designed UAV flight campaign with oblique camera views but it still fails to resolve fine shoots of the trees.

Keywords: fruticulture, precise plant protection, unmanned aerial vehicle (UAV), 3D-surface tree model

Introduction

With regard to their plant structure and morphological development, most fruit orchards show considerable spatial variability after a few years of cultivation (Aggelopoulou *et al.*, 2010). Individual trees or fruit wall sections differ in terms of their crown volume, leaf area, age and physiological development. An adaptive application of pesticides in orchards would, however, preserve biological diversity, ensure user protection and enhance food quality and environmental safety (Herrington *et al.*, 1981).
Although site-specific approaches have already been implemented in field crop production and different sensors are used to collect site-specific data to understand and react to the intra-field variability (Dammer *et al.*, 2009; Schirrmann *et al.,* 2011), only a few concepts exist for precision fruticulture. Manual methods are usually used in orchard management to collect site-specific data of trees. However, this is labour-intensive and expensive. Moreover, the empirical models created this way do not adequately reflect real plant structure. Consequently, sensor-based methods have been tested to automatically collect data from individual trees and fruit walls (Zarco-Tejada *et al.*, 2014; Mejia-Aguilar *et al.*, 2015). This data can be measured online or offline with sensors and can be used then to control the PPP application of the sprayer in orchards site-specifically (Walklate *et al.*, 2011; Gil *et al.*, 2007; Llorens *et al.*, 2011).
In the online approach, this has been achieved by ultrasound systems, infrared cameras or LiDAR systems, which were used to detect geometric properties such as leaf wall area, crown volume or plant spacing (Gil *et al.*, 2007; Balsari *et al.*, 2009; Llorens *et al.*, 2011; Walklate *et al.*, 2011; Pflanz *et al.*, 2013). In general, the control is carried out by connecting and disconnecting individual or several boom sections of the sprayer. Because the individual amount of PPP are not known before the application, economic and ecological advantages are diminished, when disposing the residuals. In addition, treatment safety is directly linked to the flawless functioning of the sensors (Pflanz *et al.*, 2013).
Another way for precise plant protection would be an offline approach in which the site-specific data were collected separately ahead of the spraying event. Since data collection is being done before PPP application, required quantities of PPP can be accurately calculated before application and

DOI 10.3920/978-90-8686-888-9_12, © Wageningen Academic Publishers 2019

therefore residual amounts will be reduced. This approach would also match the amount of PPP per unit volume better to the targeted structural parameter. This might also help reduce the technical complexity at the sprayer because, except GNSS, sensing technology is not needed with the map based offline approach. Furthermore, online sensing can deliver unobserved errors, whereas with the offline approach maps can be checked for flaws in advance, which increase treatment security. One feasible offline approach of generating application maps could be done by UAVs. UAVs can be equipped with different sensors providing spatial information with high resolution useful for agricultural applications (Schirrmann *et al.*, 2016). At low altitudes, high-resolution aerial imagery can be recorded that enables the calculation of 3D surface information of the ground. To achieve this, images need to be taken with large overlaps. This ensures that the same location on the surface is viewed from different angles, which enables photogrammetry to calculate 3D point information for this location. Current photogrammetry software is designed to meet the vast requirements of UAV imagery and uses structure from motion (SfM) algorithms to calculate 3D point clouds from hundreds of images in reasonable time. In order to minimize the risk of incorrect maps reducing treatment security, several flights during a growing season should be conducted.

The aim of this study was to create 3D point clouds and extract tree height information from fruit walls with UAV imagery and photogrammetry in an apple orchard already at the beginning of the growing season. For this purpose, point clouds must be so highly resolved that the creation of application maps for precise crop protection in fruit growing is possible. The characterization of the tree structure at early dates during the season is meaningful because the first PPP measures are carried out even before all leaves are developed. The characterization of the trees on early dates is, however, challenging for UAV photogrammetry because the trees are represented as very fine structures resulting from first leaves, blossoms and small shoots in the aerial images, which only have low contrast to the background.

Materials and methods

Test site

First field trials were carried out in late April in 2018 in Bavendorf, Germany (N47° 46' 9.588' E9° 33' 31.716'). The test site belongs to the Kompetenzzentrum Obstbau-Bodensee (KOB), labelled as Quartier 15.1. It consists of 336 adult 'Kanzi' apple trees (6 rows with 56 trees each), of which 83 trees were chopped off and 20 were cut at half of their height. It is placed in an area of about 0.11 ha. The spacing between tree rows was 3.2 m and, within the rows, trees were planted with an interval of 1 m. This orchard was chosen because of its high heterogeneity in tree heights. Also there were numerous gaps between the trees as shown in Figure 1.

Flight campaign

The flight campaign was planned with the UAV MK Mobile software (Cis GmbH, Rostock, Germany), which was specifically designed to meet the requirements of the UAV. With this software, waypoints

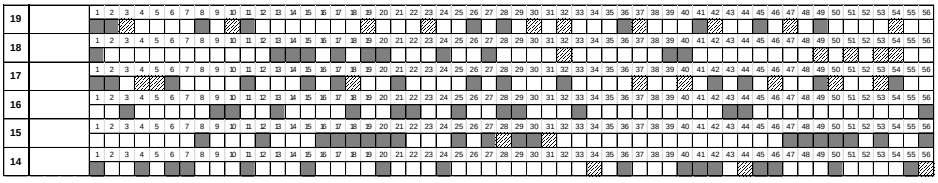


Figure 1. Trees in Q15.1 full grown (white-69.3%), cut at half (shaded-6.0%) and erased (darkgrey-24.7%) as of March 2018.

were transferred to the UAV that directed the UAV with GNSS along a user-defined flight path. The flight campaign involved three different flight paths – a contour flight path and two detailed flight paths. The contour flight was planned around the test area. With a flight speed of 1 m/s and 272 m length the flight took about 5 min. The contour flight was an overview flight and the camera angle was set in such a way that the outermost row of trees in the upper part of the images could just be seen. Therefore the camera angle was set to 9° against nadir. The forward overlap of the pictures was then about 87%. The detailed flights were conducted to take nearby images of the trees by looking into the rows. The flight path was planned directly over the trees with an altitude of 10 m. The camera was aligned to record the neighbour tree row in the lower part of the image. Therefore an angle of 20° against nadir was set. The forward overlap of the pictures was then about 74%. For every detailed flight, the UAV travelled a distance of 861 m at 1 m/s, which summed up to 15 min each flying time.

Image acquisition, georeferencing and LiDAR reference measurements

UAV imagery was acquired with an Oktokopter (CiS GmbH, Rostock, Germany) mounted with a consumer grade RGB camera (Alpha-6000, Sony, Tokyo, Japan). The system had a take-off weight of less than 2 kg and the battery was capable of maintaining flight times up to 30 min. To take sharp and steadily angled imagery, a 2-axis gimbal was used. The calculated ground sample distance (GSD) was about 2.3, 2.6 and 3.2 mm (sensor size 4,000×6,000 px, format APS-C) for the first, second and third tree rows, respectively. The focal length was fixed at 16 mm.

For point cloud georeferencing, seven plates were installed throughout the orchard with good visual appearance from above and their co-ordinates were recorded with a GNSS-RTK system (HIPer Pro, Topcon, Tokyo, Japan). The UAV point cloud validation was done by manual tree height measurements and LiDAR measurements along the rows 18 and 19. The LiDAR measurements were carried out using a laser scanner (LD MRS 400001, Sick, Germany) mounted on a tractor at a height of 2 m. The tree rows were scanned laterally with a fixed scanning angle of 0°. The LiDAR scanner was pulled with a driving speed of 0.4 m/s. The sampling rate was 25 Hz, resulting in a resolution of 1.6 cm along the track.

Data analysis

For calculating the 3D point clouds, the photogrammetric software Photoscan (Agisoft LLC, St. Petersburg, Russia) was used. The software uses SfM to estimate a 3D point cloud from the overlapping aerial images. In the first step, image matching was performed to align the images, estimate the exact camera positions and generate a sparse 3D point cloud (Key point limit: 40,000). The options 'high' and 'generic' were used here. In the second step, a dense, multi-view stereo reconstruction of the aligned image set was conducted to generate a dense 3D point cloud. The options 'high' quality and 'mild' depth filtering were used in this step.

An automatic extraction of the tree heights from the 3D point cloud was conducted for this study based on the following procedure. First the point cloud was classified into ground and tree points. For this, a progressive morphological filter after Zhang *et al.* (2003) was used. The morphological operator works directly at the point cloud level which was time-demanding. Therefore, the point cloud was randomly resampled to 1,000,000 3D points before subjecting it to the morphological classifier. After classification, ground and tree points were separated. Using a nearest neighbour (NN) search with KD-trees with k=200, the resampled tree points were retrieved from the original point cloud. The separated tree points were used to estimate the tree top location using a local maximum filter also directly at the point cloud level. To estimate reliably the tree height, a reference point on the ground was needed. One idea is to use the nearest neighbour of the tree top point to the ground point. However, the point cloud for the ground is heavily undulating. Therefore, a smoothing with a locally-weighted polynomial regression was used on the points of the ground cloud in order to estimate a reference ground surface with a spatial resolution of 10 mm. This enabled the estimation

of the tree height by calculating the distance between the tree top points and the nearest point found on the reference ground surface using a KD-tree NN search. The quality of the 3D point clouds was verified along two tree rows with manual tree height measurements and compared visually with the high quality LiDAR measurements.

Results

Figure 2D shows an image of the orchard of our test site. It is not only a 2D orthophoto, it shows a point cloud consisting of over one hundred million 3D points. They were computed from 531 images taken during the UAV flight campaign from an oblique camera view using SfM photogrammetry. Since the images are taken from above, the whole terrain of the orchard was easily mapped. The gaps visible towards the edges of the point cloud are a result of missing images in those areas in which point locations on the ground are not fully represented by multiple viewing angles anymore that were sufficient to estimate 3D points. In the point clouds, especially the ground is very detailed as can be seen in Figure 2A. Even subtle nuanced features and terrain shapes are represented in the point cloud. The ground reference surface estimated with a lowess smoother is shown as white points (Figure 2C). Regarding the trees, leaves, blossoms and even the trunk beneath is recognizable and due to the use of an RGB camera, it can be easily differentiated by its colour.
In comparison, Figure 3 shows two rows of the UAV point cloud from a side view together with a point cloud resulting from the high resolution, ground-based LiDAR measurements. Of course, the LiDAR point clouds are more detailed than the UAV point cloud because they are measured directly from the side towards the tree row with a distance of only 2 m maximum. For example, the

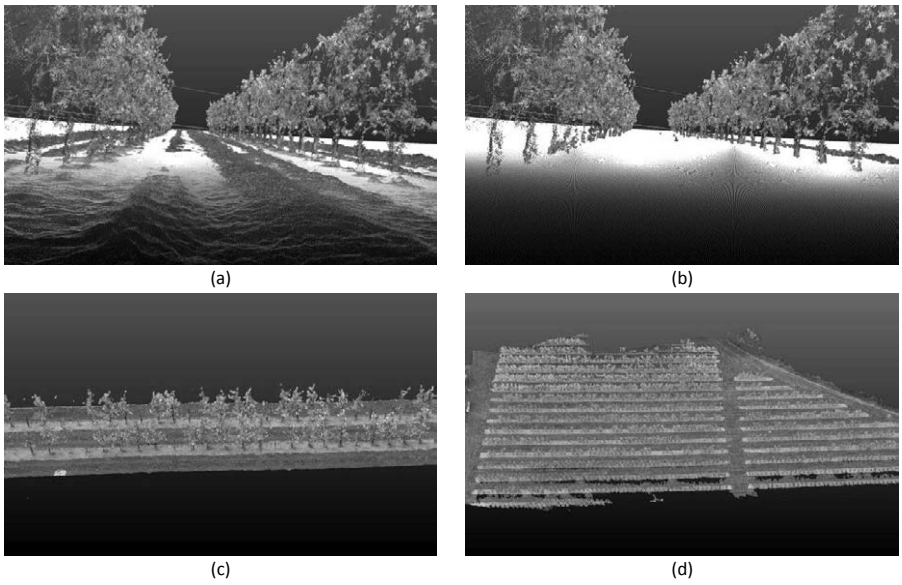


Figure 2. (A) Shows the UAV point cloud emphasizing the detailed and undulating ground. The white points are showing the ground reference surface estimated with a lowess smoother. (B) The same view with the ground removed instead showing the ground reference surface. (C) The two tree rows that were investigated in more detail in this study showing the estimated tree tops as coloured points. (D) The full point cloud as calculated from the UAV flight campaign and photogrammetry.

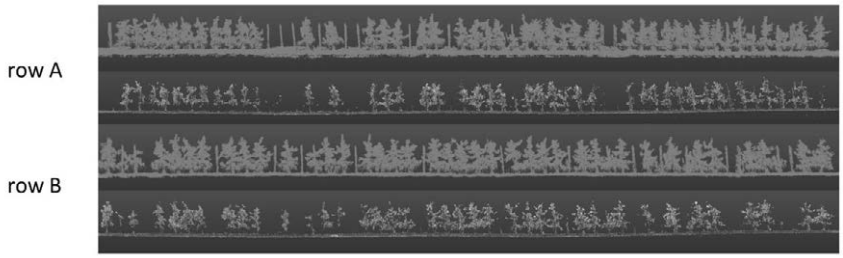


Figure 3. The two tree rows A and B in side view. In the upper half the LiDAR point cloud is shown and below the tree row from the UAV photogrammetry.

poles that were put to each tree were not resolved in the UAV point clouds. However, the side view clearly shows that the UAV point cloud sufficiently displays all the gaps within the tree rows. This would enable site-specific gap control for a precise PPP application, given as spatial information to the sprayer. Another advantage of the UAV point cloud is that it will generate 3D points with RGB colour. The RGB colour is an additional source of information that could help to separate e.g. trees from the ground and also provide information on e.g. the vitality of the trees.

In Figure 4A, the scatter plot shows the relationship between manual measurements and the estimated tree heights. The tree heights were derived from the estimated tree tops shown in Figure 2C as red-to-blue coloured 3D points over the tree crowns and the ground reference surface (Figure 2B). The correlation was highly significant with an r^2 of 0.86 (Pearson) for LiDAR measurements and r^2=0.60 for tree heights estimated from UAV point clouds (Figure 4A and 4B).

This shows that is possible to extract tree heights from the UAV point cloud with an automatic procedure to certain extent. However, it was problematic to resolve very fine details of the apple trees such as fine shoots when they were still in a leafless state. In this case, the LiDAR ground data was slightly better, which resulted in the stronger relationship with the manual height measurements. The problematic of fine shoot estimation will be further investigated with additional studies to refine UAV image acquisition, flight campaign and image processing to enable the point cloud generation with UAV on even earlier dates in the season for apple orchards.

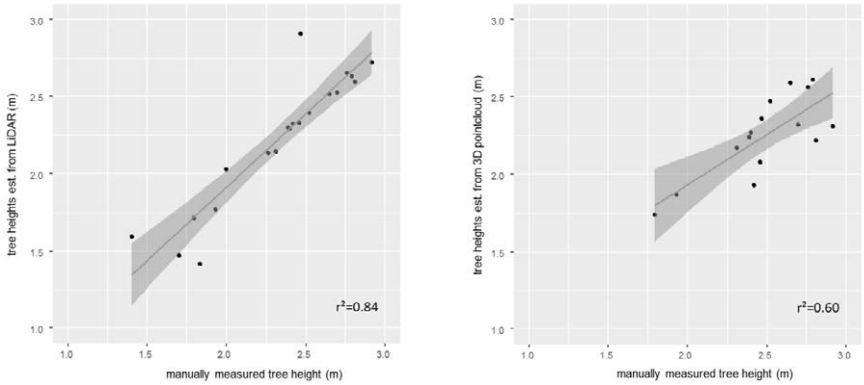


Figure 4. Scatter plot showing the linear regression between tree heights as calculated from (A) the LiDAR point cloud and (B) the UAV point cloud in regression with manual measurements in the orchard. The light gray area represents the confidence interval of 95%.

Conclusions

This study showed the successful generation of 3D point clouds from an apple orchard during blossom stage (late April) using UAV imagery and SfM photogrammetry. To achieve this, UAV images were acquired with specifically designed flight campaigns and oblique camera views. It was possible to extract tree heights from the UAV point cloud with an automatic procedure. These values were significantly correlated with manually measured tree heights in the orchard. In comparison with a tree annotation map of the orchard (Figure 1), the UAV point cloud adequately covered all gaps where trees were completely grubbed or pruned. However, it was difficult to resolve very fine details of the apple trees such as small shoots. This will be investigated further with subsequent studies for apple orchards to refine UAV image acquisition, flight campaign and image processing to enable the point cloud generation from trees even earlier during the growing season and in juvenile orchards.

References

Aggelopoulou, K. D., Wulfsohn, D., Fountas, S., Gemtos, T. A., Nanos, G. D., & Blackmore, S., 2010. Spatial variation in yield and quality in a small apple orchard. Precision Agriculture, 11 (5), 538-556. doi:10.1007/s11119-009-9146-9

Balsari, P., Marucco, P., & Tamagnone, M., 2009. A crop identification system (CIS) to optimise pesticide applications in orchards. Journal of Horticultural Science & Biotechnology, 113-116.

Dammer, K.-H., Thöle, H., Volk, T., & Hau, B., 2009. Variable-rate fungicide spraying in real time by combining a plant cover sensor and a decision support system. Precision Agriculture, 10 (5), 431-442. doi:10.1007/s11119-008-9088-7

Gil, E., Escola, A., Rosell, J. R., Planas, S., & Val, L., 2007. Variable rate application of plant protection products in vineyard using ultrasonic sensors. Crop Protection, 26 (8), 1287-1297. doi:10.1016/j.cropro.2006.11.003

Herrington, P. J., Mapother, H. R., & Stringer, A., 1981. Spray retention and distribution on apple trees. Pesticide Science, 12 (5), 515-520. doi:10.1002/ps.2780120508

Llorens, J., Gil, E., Llop, J., & Escolà, A., 2011. Ultrasonic and LIDAR Sensors for Electronic Canopy Characterization in Vineyards: Advances to Improve Pesticide Application Methods. Sensors, 11 (12), 2177-2194. doi:10.3390/s110202177

Mejia-Aguilar, A., Tomelleri, E., Vilardi, A., & Zebisch, M., 2015. UAV based tree height estimation in apple orchards: potential of multiple approaches. In European Geosciences Union General Assembly Conference Abstracts (Vol. 17, p. 7082).

Pflanz, M.; Ralfs, J.-P.; Pelzer, T., 2013: Site-specific plant protection using precise canopy gap detection. In: Garcia, E.M.; Manterola, L.V.; Perez, F.J.; Adell, P.C.; Figueroa, C.G. (Eds.): 12th Workshop on Spray Application Techniques in Fruit Growing; Book of abstracts; Valencia, Spain: Editorial Universitat Politècnica de València, pp 68-70.

Schirrmann, M., Gebbers, R., Kramer, E., & Seidel, J., 2011. Soil pH Mapping with an On-The-Go Sensor. Sensors, 11 (12), 573-598. doi:10.3390/s110100573

Schirrmann, M., Giebel, A., Gleiniger, F., Pflanz, M., Lentschke, J., & Dammer, K.-H., 2016. Monitoring Agronomic Parameters of Winter Wheat Crops with Low-Cost UAV Imagery. Remote Sensing, 8 (9), 706. doi:10.3390/rs8090706

Walklate, P. J., Cross, J. V., & Pergher, G., 2011. Support system for efficient dosage of orchard and vineyard spraying products. Computers and Electronics in Agriculture, 75 (2), 355-362. doi:10.1016/j.compag.2010.12.015

Zarco-Tejada, P. J., Diaz-Varela, R., Angileri, V., & Loudjani, P., 2014. Tree height quantification using very high resolution imagery acquired from an unmanned aerial vehicle (UAV) and automatic 3D photo-reconstruction methods. European Journal of Agronomy, 55, 89-99. doi:10.1016/j.eja.2014.01.004

Zhang, K., Chen, S. C., Whitman, D., Shyu, M. L., Yan, J., & Zhang, C., 2003. A progressive morphological filter for removing nonground measurements from airborne LIDAR data. IEEE transactions on geoscience and remote sensing, 41 (4), 872-882.

Calculating the water deficit spatially using LiDAR laser scanner in an apple orchard

N. Tsoulias[1,2*], D.S. Paraforos[3], S. Fountas[2] and M. Zude-Sasse[1]
[1]*Leibniz-Institut für Agrartechnik und Bioökonomie (ATB), Dept. Horticultural Engineering Max-Eyth-Allee 100, 14469 Potsdam, Germany; ntsoulias@atb-potsdam.de*
[2]*Agricultural University of Athens, Department of Natural Resource Management and Agricultural Engineering, Iera odos 75, 11855 Athens, Greece*
[3]*Universität Hohenheim, Institute of Agricultural Engineering, Technology in Crop Production, Garbenstr. 9, 70599 Stuttgart, Germany*

Abstract

Water use efficiency can be improved by optimized irrigation scheduling. In this study, a spatially resolved analysis of water deficit was carried out in an apple orchard (*Malus×domestica* 'Gala'). The apparent soil electrical conductivity (EC_a) was measured and a light detection and ranging (LiDAR) laser scanner system was used for estimating the leaf area per tree (LA_{LiDAR}). The LA_{LiDAR} correlated positively with EC_a in low (r=0.74), mid (r=0.86), and high (r=0.68) areas. The monthly values of the reference water balance model (WB_{RF}) based on adjusted FAO values were 36.3 mm lower than values of the model implementing LA_{LiDAR} (WB_{LiDAR}) in low EC_a regions. With LiDAR-derived plant data, land managers can improve water scheduling in fruit orchards.

Keywords: electrical conductivity, horticulture, laser scanner, leaf area, water needs

Introduction

The adoption and development of precision horticulture technologies during the last decade has enabled the observation of temporal and spatial variability within orchards (Zude-Sasse *et al.*, 2016). Particularly, the spatial data of soil and plant characteristics can be the key for more accurate irrigation scheduling of tree water needs.
Soil mapping in orchards can be based on apparent soil electrical conductivity (EC_a) (Aggelopoulou *et al.*, 2013; Tagarakis *et al.*, 2018). Beside the soil texture, EC_a can serve as an indicator for soil moisture content, due to the fact that soil properties affect the water content, when measuring at field capacity (Hedley *et al.*, 2010). Furthermore, the ability of plants to extract water is affected by the soil texture and water potential, which may vary spatially within the orchard. This parameter is defined as the readily available water content in the root zone (RAW) and utilized as a threshold for the irrigation water needs (Jensen and Allen, 2016). A widely acknowledged method to estimate the tree water requirements is the combination of potential evapotranspiration (ET_0) with the crop coefficient (K_c) and the water stress coefficient (K_s) (Allen *et al.*, 1998). The fact that orchards have a significant area of bare soil between and within trees rows makes the soil evaporation an essential factor in crop evapotranspiration (ET_c) calculation.
However, tree water needs are affected also by morphological and physiological parameters, such as the developmental stage, height and leaf area. A LiDAR terrestrial laser scanner was used to calculate the leaf area in apple, pear, and vines pointing out a strong correlation between the leaf area index (LAI) and the number of points obtained with LiDAR (Sanz-Cortiella *et al.*, 2011; Sanz *et al.*, 2013; Arnó *et al.*, 2013). However, the use of LiDAR data was never approached in water balancing.
The objectives of the present study were to (1) investigate the impact of EC_a on the LA_{LiDAR} in an apple orchard; (2) calculate the water balance, utilizing the measured fruit developmental stage,

DOI 10.3920/978-90-8686-888-9_13, © Wageningen Academic Publishers 2019

spatial variability of EC_a and LA_{LiDAR}; and (3) analyze the difference in water needs considering three zones of low, medium, and high EC_a values for setting the irrigation threshold.

Material and methods

Site description

The experiment was conducted in 36 rows (235 m) of *Malus×domestica* 'Gala' in a commercial apple orchard located in fruit production area near Brandenburg (latitude: 52.606° N, longtitude: 13.817° E), Germany, during 2018. The trees (n=5,400) were trained to grow as slender spindles on 2 horizontal parallel wires spaced 3.5×1 m. The orchard was located on slope of 3% with northeast orientation. A drip irrigation system was installed along the rows.

EC_a measurement and guided soil sampling

A Wenner electrode configuration and data logger (4-Point light, LGM, Schaufling, Germany) with four equidistantly spaced electrodes in a straight line at the soil surface was used for the calculation of soil resistivity. The depth of the measurement was fixed at 0.25 m, which corresponds to the main root system. The measurement took place every fourth tree and every second row (n=425). Soil sampling was carried out in 10 locations guided by the range of ECa data. From the results of soil texture analysis, the field capacity (FC) and the wilting point were derived based on the proposed values for typical soils in Germany (Sponagel *et al.*, 2005).

LiDAR measurements

A LiDAR system was mounted on a tractor and driven parallel to the rows of the orchard at constant speed (1 km/h) and at 2.5 m distance from the row axis (Figure 1). The 2D-LiDAR sensor in this study (LMS511 pro model, Sick, Düsseldorf, Germany) had an angular resolution of 0.1667°, scanning frequency of 100 Hz, scanning range of 80 m and scanning angle of 270°. The height of the laser sensor (1.6 m) was approximately half the maximum height of the trees (3.5 m).

A real time kinematic global navigation satellite system (RTK-GNSS) (AgGPS 542, Trimble, Sunnyvale, USA) was placed vertically next to the LiDAR to georeference each scan of the LiDAR. Furthermore, an inertial measurement unit (IMU) (MTi-G-710,XSens, Ans Enschede, the Netherlands) was used to provide roll, pitch, and yaw information needed to correct sensor readings for deviation due to the movement of the tractor. The IMU worked at the same frequency

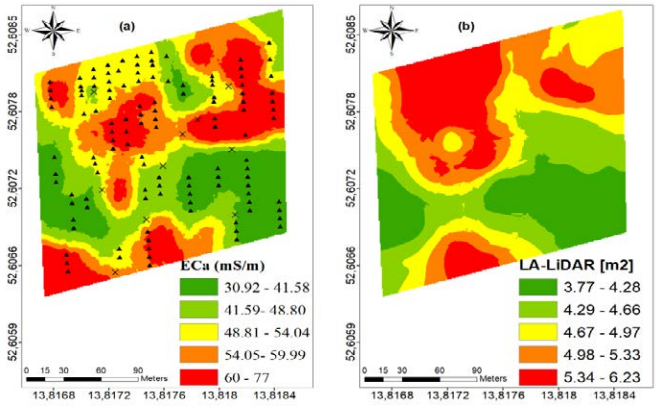


Figure 1. Spatial distribution of (A) EC_a [mS/m] and (B) LA_{LiDAR}. In ECa map the sample locations are illustrated with triangles whereas the soil samples with crosses.

with LiDAR, whilst the RTK-GNSS worked at 5 Hz. Thus, a software was developed in Visual Studio (Microsoft, USA) for all data to be timestamped, synchronized and saved on computer in a txt file. During the measurement, the scanner was pointed at a right angle to the forward direction of the tractor. Therefore, each row could be scanned from both sides when driven up and down a row. Measurements were obtained every second row of trees. Thus, 40 trees were selected and analyzed in each EC_a level (low, mid, high) during the harvest period. The LiDAR points per tree (PPT) were converted to LA based on the linear relationship that was found between these parameters. The relationship was described by a linear model, with slope 6.8250e-05 and 1.6 intercept. Therefore, the PPT for each tree in the orchard was interpreted into LA using the above model (LA_{LiDAR}).

Manual measurement of LA

The leaf area per tree was obtained from 6 representative trees manually (LA_m) to compare with the LiDAR results (LA_{LiDAR}). Each tree was defoliated manually, whereas the leaves have been counted, weighted and categorized according to their size in small, middle and large. The leaves were scanned and their surface estimated by counting the green pixels.

Water balancing

A weather station (IMT 280, Pessl, Austria) located inside the orchard, recorded air temperature (T), relative humidity (RH), wind speed (u), and solar radiation (R_n) in 15 minute intervals each day. These data can be retrieved from the following website: http://technologygarden.atb-potsdam.de. The calculation of ET_c [mm/d] was estimated with and without LiDAR data (Equations 1 and2):

$$Et_{c,RF} = (K_s K_{cb,RF} + K_e)ET_0 \quad (1)$$

$$ET_{c,LiDAR} = (K_s K_{cb,LiDAR} + K_e)ET_0 \quad (2)$$

where ET_0 is potential evapotranspiration, K_s is the soil water stress coefficient, $K_{cb,RF}$ is the tabulated basal crop coefficient (Allen *et al.*, 1998), $K_{cb,LiDAR}$ is the basal crop coefficient based on LA_{LiDAR}, and K_e is the soil surface evaporation coefficient. The ET_0 was calculated using the Penman-Monteith equation (Allen *et al.*, 1998), and the coefficients were estimated according to the methods that have been proposed earlier (Allen *et al.*, 1998; Jensen *et al.*, 2016). The crop evapotranspiration for the tabulated and measured values are described by $ET_{c,RF}$ and $ET_{c,LiDAR}$, respectively (Equation 1 and 2). The K_{cb} values were set according to the growth stages of apple trees. For considering, soil evaporation in the water balance model, the K_e for the mid and the last growth stage was calculated according to previous work (Paço *et al.*, 2014). The bud break occurred at the end of April, full bloom on 18 June, end of cell division on 18 July, harvest at 28 September, and leaf drop during the last of October. For the estimation of ET_c, the $K_{cb,RF}$ from bud break until full bloom ($K_{cb,ini,,RF}$) was set at 0.8, from full bloom till the end of cell division ($K_{cb,mid,RF}$) at 1.2, and at 0.85 from harvest till leaf drop ($K_{cb,end,RF}$) (Allen *et al.*, 1998; Table 12). For the estimation of $ET_{c,LiDAR}$, the $K_{cb,ini}$ was set at 0.8, whereas for the following stages the $K_{cb,LiDAR}$ values were adjusted to the local crop parameters and daily climatic conditions (Allen *et al.*, 1998) and implemented into the following equations (Equations 3 and 4):

$$K_{cb,mid} = K_{c,min} + (K_{cb,full,mid} - K_{c,min})(1 - e^{-0.7LA_{LiDAR}}) \quad (3)$$

$$K_{cb,end} = K_{c,min} + (K_{cb,full,end} - K_{c,min})(1 - e^{-0.7LA_{LiDAR}}) \quad (4)$$

Where $K_{cb,full,mid}$ and $K_{cb,full,end}$ were estimated based on daily weather conditions. The minimum K_c for bare soil ($K_{c,min}$) was set at 0.15 (Allen *et al.*, 1998). The K_s was estimated by the following

equations (Equations 5-7), only when the water depletion in the root zone at the end of day i ($D_{r,i}$) [mm] exceeded the readily available water content (RAW).

$$K_S = TAW - D_{r,i} / TAW - RAW \quad (5)$$

$$TAW = 1000(\Theta_{FC}-\Theta_{WP})Z \quad (6)$$

$$RAW = TAW\ p \quad (7)$$

TAW is total plant available soil water in the root zone considering the suggested root depth Z of 1 m from FAO-56; RAW, the readily available soil water content for the reference field-uniform root depth of 1 m. The RAW was estimated daily for low, mid and high EC_a regions, considering the average fraction of TAW that can be used by the plant in the root zone (p); Θ_{FC} is the volumetric water content of the soil sample at field capacity [m^3/m^3]; Θ_{WP} is the volumetric water content at wilting point of the soil samples [m^3/m^3]. As was mentioned before, the last two factors have been specified according the indicative values of Θ_{FC} and Θ_{WP} for Germany (Sponagel *et al.*, 2005) (Equation 6). The $ET_{c,RF}$ and $ET_{c,LiDAR}$ were implemented in a water balance model for calculating the daily water needs for tabulated (WB_{RF}) and LiDAR (WB_{LiDAR}) values [mm] in low, mid, and high EC_a regions of the orchard (Equations 8 and 9). P is the daily precipitation [mm].

$$WB_{RF} = ET_{c,RF} - P \quad (8)$$

$$WB_{LiDAR} = ET_{c,LiDAR} - P \quad (9)$$

Data processing

The orchard was divided in 28 rectangular cells. The mean value of EC_a and LiDAR measurements was then calculated in each cell. Subsequently, ordinary Kriging was utilized to interpolate the EC_a spatial variability. Quintile classification was then used to group Kriged values of ECa into a thematic map. The analysis was carried out in ArcGIS (Version 10.2.2. Redlands, USA). The calculation of water balance using the Geisenheim model was carried out in Matlab (Version R2017a, Mathworks Inc., Natick, MA, USA).

In the case of the LiDAR sensor, the data were filtered to eliminate, considering only those detections with a maximum distance of 4 m from the sensor. Therefore, the generated point clouds of each side of the tree were merged using the iterative closest point (ICP) algorithm (Vazquez-Arellano *et al.*, 2018a) and the points which fell on the ground were erased with the random sample consensus (RANSAC) algorithm based on Garrido *et al.* (2015). Furthermore, the tree stem position was defined based on Vazquez-Arellano *et al.* (2018b) methodology whereby the position of the stem was used as the centre base in order to create a cylinder with radius equal to the half distance between each stem position (radius=0.52). The points inside the cylinder were counted for each tree individually. Finally, a linear regression was used to define the relationship between the LA_m and LA_{LiDAR}.

Results and discussion

The results from soil samples according to the guided soil sampling revealed that the soil texture was loamy sand. The highest EC_a values were located in the northern extreme of the orchard while the lowest values in the middle section of the orchard (Figure 1A). The values of EC_a varied between 30.92 to 48.8 mS/m characterized as low, between 48.81 to 54.04 mS/m as mid, and between 54.05 to 77 mS/m as high EC_a zones.

A significant correlation was revealed between the LA_m and the PPT from the six defoliated trees (r^2=0.87). The LA_{LiDAR} revealed a similar spatial distribution with EC_a (Figure 1B). This was also

described by a significant correlation between the EC_a and the PPT (Table 1). A result which indicates that trees in areas of high EC_a can have larger canopy surface compared to low EC_a areas. Frosts occurred from April to May. Precipitation remained low during the full bloom. In June, rains above 30 mm were recorded, representing an extremely dry year for this area (295 mm). The air relative humidity fluctuated from 40 to 85%, between full bloom and harvest. The temperature descended, while the precipitation remained below 8 mm during harvest and leaf drop.

The highest water demands were depicted during harvest in low EC_a areas with 78.7 mm for WB_{RF} and 115 mm for WB_{LiDAR} (Figure 2). The WB_{RF} dispersion around the mean was enhanced in comparison with the WB_{LiDAR}, especially in July in mid EC_a areas (C.V.%=60.52). However, none of the models showed significant variation between May and June in all EC_a regions.

The results of WB_{RF} and WB_{LiDAR} illustrate the same patterns in all EC_a regions (Figure 2). A steep drop of water needs occurred in May due to increased precipitation. In low EC_a areas, the water needs according to WB_{LiDAR} overpassed the threshold (RAW_{low}) in early June in comparison with mid and high EC_a in the end of this month. Moreover, WB_{RF} revealed 20.2 mm and 23.3 mm less water need from WB_{LiDAR}, resulting no water stress during May, in mid and high EC_a areas, respectively. Furthermore, in August and September, a difference of approximately 36.3 mm was found between WB_{LiDAR} and WB_{RF}, in low EC_a areas. A steady deficiency of water needs were illustrated after harvest, nevertheless both models remained above their thresholds. It is well known that plants

Table 1. Pearson correlation between apparent soil electrical conductivity (EC_a) and LA_{LiDAR}. n = the number of samples. Correlation at 0.01 level (**) and at 0.05 (*) level of significance.

	n	Low EC_a	Mid EC_a	High EC_a
LA_{LiDAR} [m^2]	28	0.74*	0.86**	0.68**

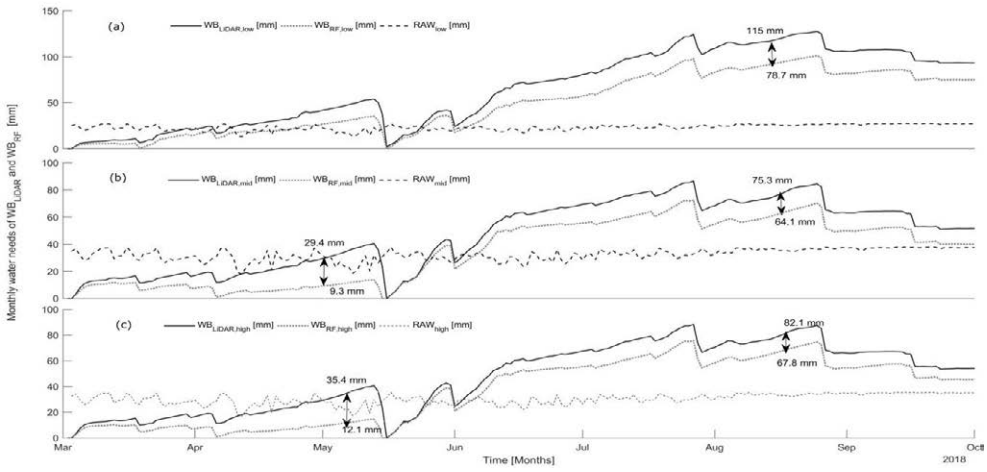


Figure 2. Water needs [mm] according to water balance model for tabulated values from FAO (WB_{RF}) and considering the leaf area per tree measured by means of LiDAR (WB_{LiDAR}) in low (A), mid (B), and high (C) regions of soil electrical conductivity (EC_a). Readily available water content (RAW) in low (RAW_{low}), mid (RAW_{mid}) and high (RAW_{high}) EC_a region were set as thresholds, respectively.

reduce transpiration rate under drought stress due to physiological responses such as stomatal closure and subsequent reductions in photosynthesis rate. Furthermore, water stress adversely affects the physiological and nutritional development of crops, leading to reduced biomass, yield, and quality of crops (Rossini *et al.*, 2013). Also, the variation in rooting depth with soil texture can affect the water needs spatially (Pérez-Pastor *et al.*, 2014). In this study, the results revealed that trees in low EC_a areas had reduced LA and enhanced water stress. Slightly enhanced water needs were pointed out by WB_{LiDAR}, however the WB_{LiDAR} did not differ significantly from the WB_{RF} in all months.

Conclusions

In this study, the estimation of LA by means of LiDAR was confirmed. Furthermore, an impact of LA and soil patterns on water needs was found in the apple orchard. The spatial dependence of LA on the EC_a region was revealed. In low EC_a areas the highest water demand was depicted, indicating 36.3 mm difference from the FAO model. LiDAR-based plant data are potentially useful for optimizing and predicting monthly water needs spatially within orchards.

References

Aggelopoulou, K., Castrignanò, A., Gemtos, T., & De Benedetto, D., 2013. Delineation of management zones in an apple orchard in Greece using a multivariate approach. Computers and Electronics in Agriculture, 90, 119-130.

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Rome, Italy: FAO 300(9), D05109.

Arnó, J., Vallès, J. M., Llorens, J., Sanz, R., Masip, J., Palacín, *et al.*, 2013. Leaf area index estimation in vineyards using a ground-based LiDAR scanner. Precision Agriculture, 14(3), 290-306.

Garrido, M., Paraforos, D. S., Reiser, D., Vázquez Arellano, M., Griepentrog, H. W., & Valero, C., 2015. 3D maize plant reconstruction based on georeferenced overlapping LiDAR point clouds. Remote Sensing, 7(12), 17077-17096.

Hedley, C. B., Bradbury, S., Ekanayake, J., Yule, I. J., & Carrick, S., 2010. November. Spatial irrigation scheduling for variable rate irrigation. In Proceedings of the New Zealand Grassland Association (Vol. 72, pp. 97-102). New Zealand Grassland Association.

Jensen, M. E., & Allen, R. G., 2016. Evaporation, Evapotranspiration, and Irrigation Water Requirements. Reston, VA, USA: American Society of Civil Engineers.

Paço, T. A., Pôças, I., Cunha, M., Silvestre, J. C., Santos, F. L., Paredes, P., *et al.*, 2014. Evapotranspiration and crop coefficients for a super intensive olive orchard. An application of SIMDualKc and METRIC models using ground and satellite observations. Journal of Hydrology, 519, 2067-2080.

Pérez-Pastor, A., Ruiz-Sánchez, M. C., & Domingo, R., 2014. Effects of timing and intensity of deficit irrigation on vegetative and fruit growth of apricot trees. Agricultural water management, 134, 110-118.

Rossini, M., Fava, F., Cogliati, S., Meroni, M., Marchesi, A., Panigada, *et al.*, 2013. Assessing canopy PRI from airborne imagery to map water stress in maize. International Society for Photogrammetry and Remote Sensing. Remote Sensing 86, 168-177.

Sanz, R., Rosell, J. R., Llorens, J., Gil, E., & Planas, S., 2013. Relationship between tree row LIDAR-volume and leaf area density for fruit orchards and vineyards obtained with a LIDAR 3D Dynamic Measurement System. Agricultural and forest meteorology, 171, 153-162.

Sanz-Cortiella, R., Llorens-Calveras, J., Arnó-Satorra, J., Ribes-Dasi, M., Masip-Vilalta, J., Camp *et al.*, 2011. Innovative LIDAR 3D dynamic measurement system to estimate fruit-tree leaf area. Sensors, 11(6), 5769-5791.

Sponagel, H., Grottenthaler, W., Hartmann, K.J., Hartwich, R., Jaentzko, P., Joisten, H., *et al.*, 2005. Bodenkundliche Kartieranleitung, Bundesanstalt für Geowissenschaften und Rohstoffe und den geologischen Landesämtern in der Bundesrepublik Deutschland Hannover.

Tagarakis, A. C., Koundouras, S., Fountas, S., & Gemtos, T., 2018. Evaluation of the use of LIDAR laser scanner to map pruning wood in vineyards and its potential for management zones delineation. Precision Agriculture, 19(2), 334-347.

Vázquez-Arellano, M., Paraforos, D. S., Reiser, D., Garrido-Izard, M., & Griepentrog, H. W., 2018b. Determination of stem position and height of reconstructed maize plants using a time-of-flight camera. Computers and Electronics in Agriculture, 154, 276-288.

Vázquez-Arellano, M., Reiser, D., Paraforos, D.S., Garrido-Izard, M., Burce, M.E.C., Griepentrog, H.W., 2018a. 3-D reconstruction of maize plants using a time-of-flight camera. Computers and Electronics in Agriculture. 145, 235-247.

Zude-Sasse, M., Fountas, S., Gemtos, T.A., Abu-Khalaf, N., 2016. Applications of precision agriculture in horticultural crops – review. European Journal of Horticultural Science 81: 78-90.

Section 3
Precision viticulture

Comparison of water potential and yield parameters under uniform and variable rate drip irrigation in a cabernet sauvignon vineyard

I. Bahat[1,3,4,5*], Y. Netzer[4,5], A. Ben-Gal[6], J.M. Grünzweig[3], A. Peeters[2] and Y. Cohen[1]
[1]*Institute of Agricultural Engineering, Agricultural Research Organization (Volcani Center), Rishon LeZion, Israel; idanba1@gmail.com*
[2]*TerraVision Lab, Midreshet Ben-Gurion, Israel*
[3]*Institute of Plant Sciences and Genetics in Agriculture, The Robert H. Smith Faculty of Agriculture, Food & Environment, The Hebrew University of Jerusalem, Rehovot, Israel*
[4]*Department of Chemical engineering, Biotechnology and Materials, Ariel University, Ariel, Israel*
[5]*Department of Agriculture and Oenology, Eastern R&D Center, Israel*
[6]*Department of Soil, Water and Environmental Sciences, Agricultural Research Organization Gilat Research Center, Israel*

Abstract

An experiment in variable-rate drip irrigation (VRDI) was aimed at reducing variance of midday stem water potential (SWP) and in yield parameters by applying VRDI in a highly variable vineyard. During 2018, irrigation was separated into VRDI and uniform irrigation (UI) blocks. Each block was delineated to 10 management cells, and results of 2018 season were compared to 2017 season under UI. Standard error (SE) of SWP in the last 3 measurements before harvest decreased in the VRDI block in 2018 compared to 2017 by 14-44%. In contrast, in the UI block, SE in 2018 was higher by 11-42% compared to 2017. SE of fruit yield showed a similar trend. Applying principles of precision irrigation may lead to a more homogeneous vineyard in the parameters described above and to improved wine quality.

Keywords: vineyard, variable rate drip irrigation, midday stem water potential, management cell, deficit irrigation

Introduction

In-field plant heterogeneity in commercial vineyards is a key factor that affects fruit yield and red wine quality. Nonetheless, wine growers assume uniformity and apply uniform agro-technical and irrigation management within each plot. Precision irrigation may improve homogeneity of vine water status, fruit yield and wine quality. A few recent studies applied VRDI in vineyards (McClymont *et al.*, 2012; Nadav and Schweitzer, 2017; Sanchez *et al.*, 2017). These studies differed in zonation methodology and how irrigation decisions were made for each zone, but in all of them yield variability decreased. Mcclymont *et al.* (2012) divided the vineyard into management zones (MZ) that differed in size and shape using multi-year NDVI maps and a canopy temperature map. In comparison, Sanchez *et al.* (2017) used a 'blind' zonation based on regular polygons of equal sizes. Despite the differences in zonation approach, Mcclymont *et al.* (2012) assumed relative stability of the MZ condition over years, i.e. a zone with lower vigour in one year was assumed to likely exhibit less vigour in other years. This assumption led to VRI decisions based on conditions in the previous years. In contrast to this, several studies have shown that MZs are dynamic and can change from year to year and vary even during a single season (Cohen *et al.*, 2017; O'Shaughnessy *et al.*, 2015). In Sanchez *et al.* (2017), the in-season dynamic was addressed to some extent, as NDVI maps were used to govern irrigation in each zone. Deficit irrigation is a common and necessary agricultural practice in vineyards in order to achieve optimal fruit quantity and quality suitable for wine production (Acevedo-opazo *et al.*, 2010; Munitz *et al.*, 2017; Romero *et al.*, 2016). Wine

DOI 10.3920/978-90-8686-888-9_14, © Wageningen Academic Publishers 2019

quality can be affected by regulated deficit irrigation (RDI) which is dynamic and can be changed during the growing season. Applying deficit irrigation necessitates water status monitoring measures and tools. SWP was demonstrated to be a comprehensive indicator of early water deficit in vines (Choné *et al.*, 2001), and an accurate indicator for vine irrigation management especially for deficit irrigation management (Girona *et al.*, 2006). Based on the hypothesis that reducing water status and yield variability would improve wine quality as a larger fraction of grapes mature uniformly, the capability of SWP-based VRDI in a highly variable commercial vineyard to decrease water status and yield variability was examined.

Materials and methods

Mevo-Beitar vineyard

During 2017-2018, a field experiment was conducted in a highly variable 2.4 ha *Vitis vinifera* cv. *'Cabernet sauvignon'* vineyard, established in Mevo-Beitar (31°43' N; 35°06' E) in the Judean hills region in Israel. Vines were planted in 2011 in a northwest-to-southeast direction with a vine and row spacing of 1.5 m and 3.0 m, respectively (Figure 1). Prior to the irrigation experiment, soil and vine spatial variability was mapped and quantified by various measurements: (1) ECa (apparent electrical conductivity) mapping using Veris-3000 (VERIS technologies, Oregon, USA); (2) Elevation (676-700 m asl), slope (1-22%) and aspect (N, NW, SW, and S); (3) Soil depth (80 sample points throughout the vineyard; 0.35-1.60 m); (4) Trunk diameter (all 5,521 vines measured; 11-72 mm).

Variable-rate drip irrigation design

The VRDI system was installed in spring 2018. Each irrigation zone was independently irrigated and was controlled by an electric valve with a separate water meter, and all of the valves were controlled by a remote irrigation controller (Dream 2, Talgil, Israel). 'Blind' zonation based on regular polygons of equal sizes was used to divide the vineyard into 20 management cells (MC, 30×30 m each; Figure 1). There were 11 valves: 1 for each of the 10 MCs in the western VRDI block and one for the entire eastern UI block.

Experimental design, water status monitoring and irrigation management

MCs were individually monitored during the 2017 and 2018 growing season. Field measurements included SWP and LAI in 6 vines per MC every 1-2 weeks. SWP was measured prior to irrigation at solar noon (from 12:00 to 14:30), using portable pressure chamber (Arimad 3000, MRC, Holon Israel) according to the procedures described by Boyer (1995). Drip irrigation in the 2017 season

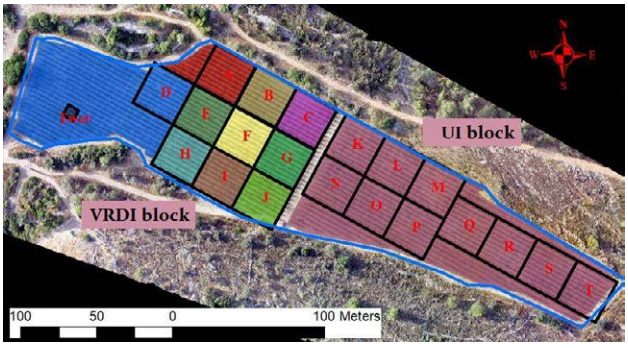


Figure 1. An aerial image of the Mevo-Beitar commercial vineyard overlaid with the management cells and the uniform irrigation and VRI blocks.

was uniform for all 20 MCs, based on crop coefficients and reference evapotranspiration (ET_o), calculated according to the ASCE Penman-Monteith equation, and corrected by SWP measurements. In 2018, the vineyard was divided into Western (10 MCs) VRDI and Eastern (10 MCs) UI blocks. Irrigation decision management in the vineyard in 2018 was according to an irrigation model which considers temporal ET_o and LAI to calculate crop evapotranspiration (ET_c; Munitz *et al.,* 2016). ET_o was extracted from regional meteorological station (Tzuba), 5,920 m distance from the vineyard. In this study, a novel approach to irrigation management was adopted by setting an optimal preliminary SWP curve for all phenological stages that was suitable for the cultivar, the region and winery expectations (Munitz *et al.,* 2017). Management cells in the western block (A-J) were variably irrigated according to their individual weekly SWP results and as such, crop coefficient could range from 0 to 50% ET_c while all MCs in the eastern block (K-T) were irrigated uniformly according to the block average SWP. The vineyard was harvested by a mechanical weighing harvester ('Weight online system', Zvi Weinberg & Sons LTD, Israel assembled on New Holland, Braud 9040), and yield maps were created.

Data analysis

The major variability parameter was SE. Yield maps were processed using ArcGIS10.3 (ESRI, Redlands, CA, USA). Statistical analysis of data variance was carried out using JMP Pro 14.0.0 (SAS Institute, Cary, NC, USA).

Results

Seasonal ET_o (Apr-Sept) was 858 mm in 2017 and 824 mm in 2018. Annual precipitation was 369 mm in 2017 and 452 mm in 2018. Differences were mainly due to a rainy spring in 2018 that affected the growing season. Irrigation amounts varied accordingly and were 98 mm/season in 2017 and significantly lower in 2018 (32.7 mm/season) due to the application of the VRDI regime, higher initial soil water content and lower climate demands (Figure 2).

SWP in 2017 (under UI) showed high in-field variability and ranged from -0.87 to -1.77 MPa at harvest. Figure 3 and 4 shows the change in SWP in MCs with two extreme water stress conditions in the VRDI block and in the entire eastern UI block in 2017 and in 2018, respectively. In 2017, under uniform irrigation MC 'C' suffered from drought stress throughout the whole season, with the exception of the harvest period, in which it reached the desirable SWP range. In contrast, at most measurement dates, MC 'D' had higher (less negative) SWP values than the upper SWP limit. The eastern block, showed an intermediate trend. In 2018, the eastern block under uniform irrigation, followed the targeted SWP range in the beginning of the irrigation period, and then showed high water availability till the end of the season. Similarly MC 'D', which was irrigated individually showed a similar trend. In contrast, MC 'C', with the aid of higher amount and more frequent irrigation applications, generally followed the targeted range throughout the entire season.

The standard error (SE) of SWP in the last three measurements before harvest in 2018 decreased by 14-44% in the VRDI block compared to 2017, which was under unified irrigation. In contrast, in the western block which was under uniform irrigation in both seasons, SE of SWP in 2018 was higher by 11-42% compared to 2017 (Figure 5). The SE of fruit yield increased in both blocks, yet the increased magnitudes were very different. In the UI block, SE of fruit yield increased by 30.4% from 2017 to 2018 while in the VRDI block by only 4.9% (Figure 6). The yield and wine quality parameters are under analysis.

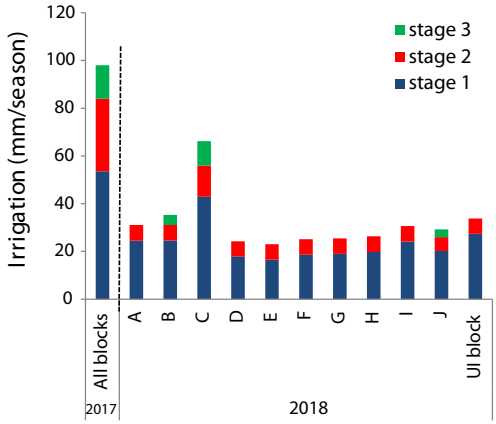


Figure 2. Irrigation amounts applied in 2017 for the whole vineyard and in 2018 for every management cell in the VRDI block and for the whole UI block. The irrigation amounts were divided by berry phenological stages (I – fruit set to cluster closure, II – cluster closure to end of veraison, III – veraison to harvest).

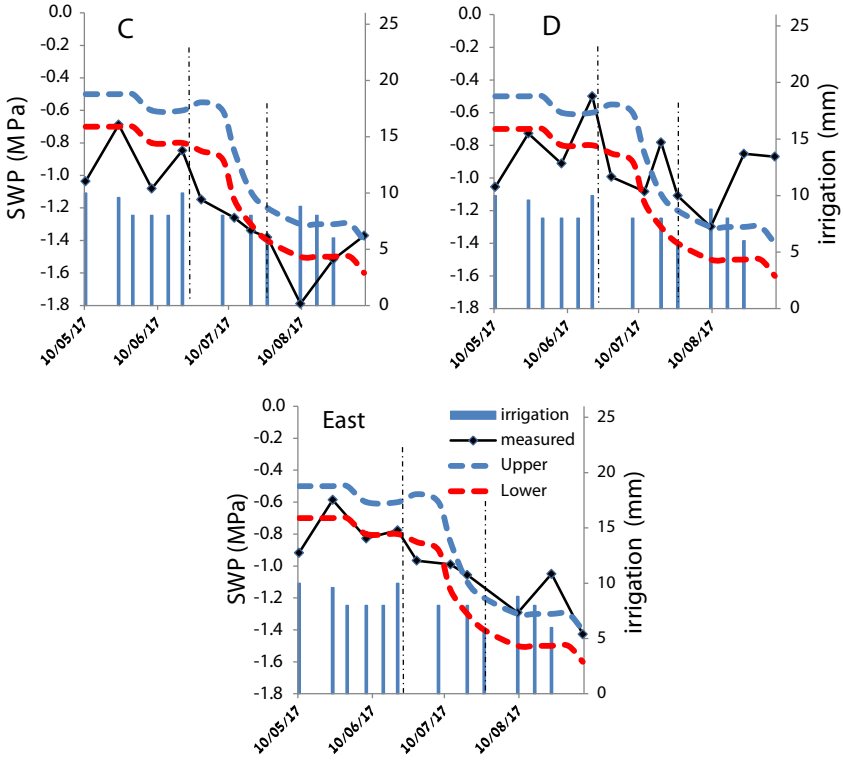


Figure 3. Seasonal trend of SWP and irrigation amounts throughout 2017 season in MC 'C' and 'D' in the VRDI block (western) and in the whole UI block (eastern). The measured SWP changes are compared to upper and lower target SWP curves which were implemented in 2018.

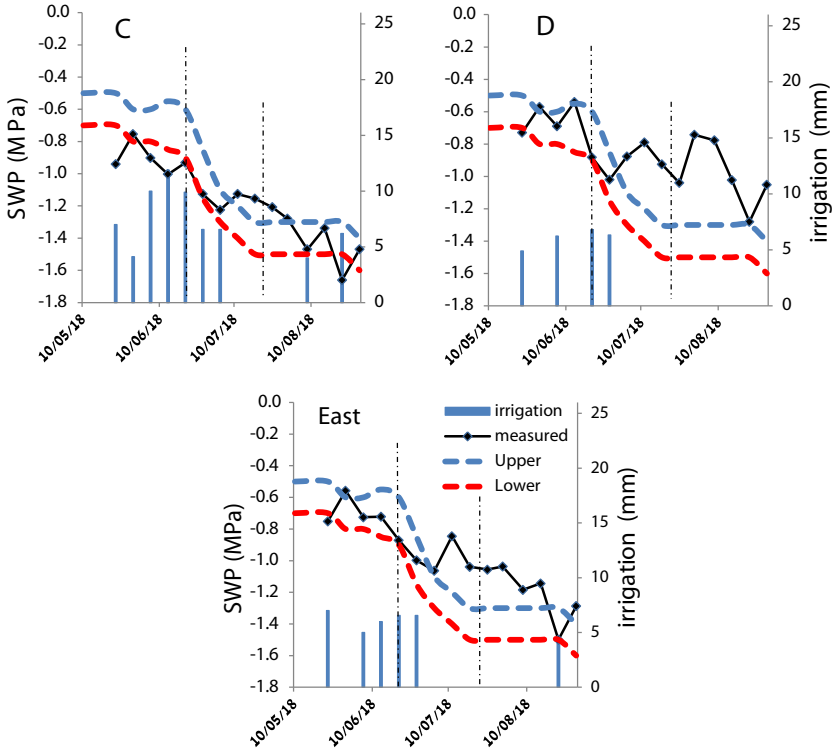


Figure 4. Seasonal trend of SWP and irrigation amounts throughout 2018 season in MC 'C' and 'D' in the VRDI block (western) and in the whole UI block (eastern). The measured SWP changes are compared to upper and lower target SWP curves which were implemented in 2018.

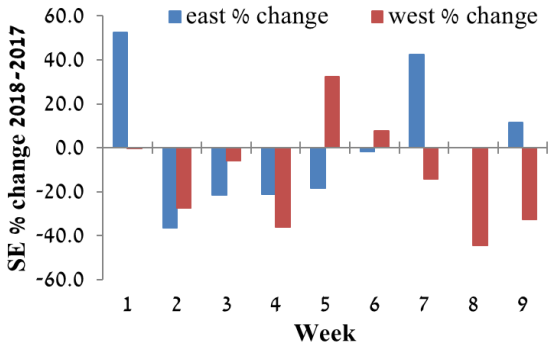


Figure 5. The relative change in SE of SWP from 2017 to 2018.

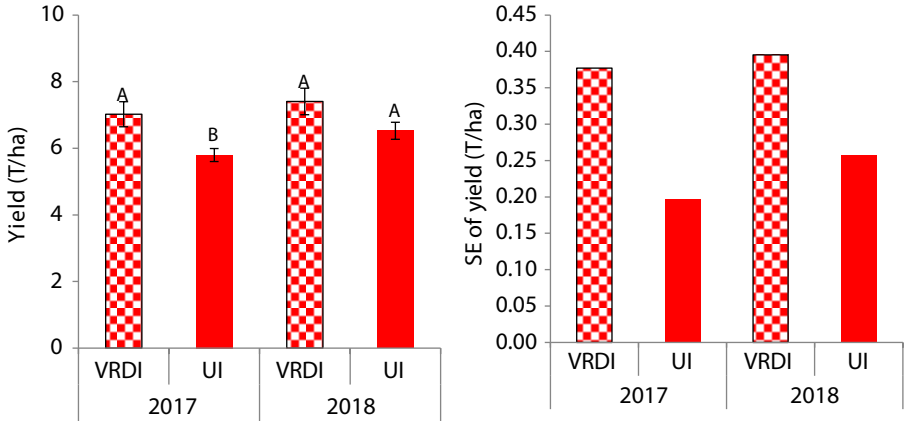


Figure 6. Yield and SE of yield in VRDI block vs UI block at 2017-2018.

Discussion

Quantifying the impact of VRDI on a commercial scale in order to decrease water status and yield spatial variability is not trivial. Eliminating the effects of intra-field spatial variability and inter-season temporal variabilities is complex. Also, potential impact could be expected to become apparent only after several seasons because of the fact that fruit development extends over 2 seasons. Due to differences in temporal distribution of rain, the irrigation regime was substantially different in 2018 as total irrigation was 33 mm compared to 98 mm applied in 2017. Monitoring the spatial variability on a fine scale in the two blocks throughout two very different seasons enabled the quantification of the VRDI effect. Variability of SWP towards the end of the season, was decreased in the VRDI block, while in the UI block, it was either similar or increased. Similar studies (McClymont *et al.*, 2012; Sanchez *et al.*, 2017) have shown that yield variability decreased under VRDI. The fruit yield variability in this study, increased under the two irrigation regimes in the second season. Yet, while it was slightly increased in the VRDI block, it majorly increased in the UI block. To apply VRDI based on SWP on a commercial scale, a substitute for the tedious SWP measurements should be established. Ground and aerial thermal images, interpreted for crop water stress index (CWSI) maps (Cohen *et al.*, 2017) were shown to be a sensitive indicator of SWP (Möller *et al.*, 2007; Matese *et al.*, 2018). As part of a larger study, 13 thermal images from a UAV were acquired above the vineyard simultaneously with SWP measurements. The data indicated strong correlations in most phenological stages with relatively stable linear models. In 2019, the VRDI will be governed by thermal images to test the ability to map plant water status variability to an acceptable and applicable accuracy.

Conclusions

Applying principles of precision irrigation based on field heterogeneity, SWP thresholds and irrigation models may lead to a more homogeneous vineyard in terms of water status and certain yield parameters. It is assumed that this homogeneity is a key attribute for better red wine quality control as fruit maturation and berry development are more adequate.

Acknowledgements

This project was funded by chief scientist fund of the Ministry of Agriculture in Israel. We wish to thank Dror Dotan, David Twito and Avi Yehuda from 'Carmel Winery' and Ben Hazut, Matan Golomb and Doron Kleiman for their accurate field measurements.

References

Acevedo-opazo, C., Ortega-farias, S., & Fuentes, S., 2010. Effects of grapevine (*Vitis vinifera* L.) water status on water consumption, vegetative growth and grape quality : An irrigation scheduling application to achieve regulated deficit irrigation. Agricultural Water Management, 97 (7), 956-964.

Boyer, 1995. Measuring the water status of plants and soils. San Diego, CA, USA: Academic Press.

Choné, X., van Leeuwen, C., Dubourdieu, D., & Guadilleare, J. P., 2001. Stem Water Potential is a Sensitive Indicator of Grapevine Water Status. Annals of Botany 87, 477-483.

Cohen, Y., Alchanatis, V., Saranga, Y., Rosenberg, O., Sela, E., & Bosak, A., 2017. Mapping water status based on aerial thermal imagery: comparison of methodologies for upscaling from a single leaf to commercial fields. Precision Agriculture, 18 (5), 801-822.

Girona, J., Mata, M., Del Campo, J., Arbonés, A., Bartra, E., & Marsal, J., 2006. The use of midday leaf water potential for scheduling deficit irrigation in vineyards. Irrigation Science, 24 (2), 115-127.

Matese, A., Baraldi, R., Berton, A., Cesaraccio, C., Di Gennaro, S. F., Duce, P., *et al.,* 2018. Estimation of Water Stress in grapevines using proximal and remote sensing methods. Remote Sensing, 10(1), 1-16.

McClymont, L., Goodwin, I., Mazza, M., Baker, N., Lanyon, D. M., Zerihun, A., *et al.,* 2012. Effect of site-specific irrigation management on grapevine yield and fruit quality attributes. Irrigation Science, 30(6), 461-470.

Möller, M., Alchanatis, V., Cohen, Y., Meron, M., Tsipris, J., Naor, A., *et al.,* 2007. Use of thermal and visible imagery for estimating crop water status of irrigated grapevine. Journal of Experimental Botany, 58(4), 827-838.

Munitz, S., Netzer, Y., & Schwartz, A., 2017. Sustained and regulated deficit irrigation of field-grown Merlot grapevines. Australian Journal of Grape and Wine Research, 23 (1), 87-94.

O'Shaughnessy, S. A., Evett, S. R., & Colaizzi, P. D., 2015. Dynamic prescription maps for site-specific variable rate irrigation of cotton. Agricultural Water Management, 159, 123-138.

Romero, P., García García, J., Fernández-Fernández, J. I., Muñoz, R. G., del Amor Saavedra, F., & Martínez-Cutillas, A., 2016. Improving berry and wine quality attributes and vineyard economic efficiency by long-term deficit irrigation practices under semiarid conditions. Scientia Horticulturae, 203, 69-85.

Sanchez, L. A., Sams, B., Alsina, M. M., Hinds, N., Klein, L. J., & Dokoozlian, N., 2017. Improving vineyard water use efficiency and yield with variable rate irrigation in California. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences, 8(02), 574-577.

Use of an integrated model of water consumption as a decision support system for scheduling regulated deficit irrigation in a vineyard

J. Bellvert*, M. Mata, X. Vallverdú, C. Paris and J. Marsal
Efficient Use of Water in Agriculture, Institute of Agri-Food Research and Technology (IRTA), Av. Alcalde Rovira Roure 191, 25198 Lleida, Spain; joaquim.bellvert@irta.cat

Abstract

While the agronomic and economic benefits of the adoption of regulated deficit irrigation (RDI) strategies have long been established in red winegrape varieties, the spatial variability in water requirements across a vineyard limits its practical application. This study demonstrated an integrated methodology for conducting precision irrigation management in a commercial vineyard through the use of a vine water consumption model and remote sensing imagery. The case study was conducted in a vineyard of 52 irrigation sectors with the varieties Tempranillo, Cabernet Sauvignon and Syrah. This methodology allowed the winery to adopt a RDI strategy, and, despite differences in water requirements between irrigation sectors, vines approached the same pre-defined levels of water stress. Maximum differences in the total amount of water applied between irrigation sectors were 84.1 and 125.9 mm respectively for 2016 and 2017, indicating important water savings compared to the fact that vines had been watered homogeneously.

Keywords: viticulture, irrigation, crop model, remote sensing

Introduction

In viticulture, maximum incomes are not achieved with maximum yield but in maintaining a certain balance between yield and wine quality. It is well-known that this yield/wine quality balance can be achieved with the adoption of regulated deficit irrigation (RDI) strategies (Girona *et al.,* 2006). However, to adopt the desired levels of water stress accurately at field level is a very complicated task. If water stress is excessive, there is a risk of a very significant productive and quality reduction. On the other hand, if insufficient, the intended wine quality levels are not achieved. In addition, spatial variability in water requirements across a vineyard limits the efficient use of irrigation. Therefore, to succeed with the practical application of RDI strategies at field level, it is necessary to first solve the following bottlenecks: (1) to determine within-field soil spatial variability, since not all soils retain or provide water to the plant at the same rate; (2) to adapt the secondary irrigation network according to the natural variability of soils, in order to be able to irrigate each sector in a differentiated way; and (3) to know the water stress thresholds that need to be adopted in each variety in order to achieve the best combination of yield/quality. These thresholds have already been described for some varieties (Girona *et al.,* 2006, Basile *et al.,* 2011, Bellvert *et al.,* 2016). The new challenge is to use a series of procedures in an integrated manner that can individually resolve these bottlenecks. Without a clear link between them, they are not commercially practical. This new methodology should be properly integrated and could be based on the use of geographic information systems, remote sensing, crop simulation models and irrigation management. These integrated procedures must also consider the criteria of the seasonal sensitivity of vines to water stress, and to have a reasonable cost/benefit ratio. Therefore, the first goal of this study was to design a decision-oriented vine consumption model for scheduling irrigation, capable of determining the amount of water to apply when a RDI strategy is imposed. The second objective was to evaluate the performance of the model in a commercial vineyard together with the help of remote sensing data; and finally, to

DOI 10.3920/978-90-8686-888-9_15, © Wageningen Academic Publishers 2019

evaluate the effects of conducting differential irrigation management in terms of water efficiency and spatial variability, as well as to explore new advances in new remote sensing techniques.

Materials and methods

The model

The vine water consumption model developed in this study is based on CropSyst (Stöckle *et al.*, 2003), with recent modifications for deciduous trees (Marsal *et al.*, 2013, Marsal *et al.*, 2014) (Figure 1). The model is written in Matlab® (MATLAB 2014b, The MathWorks, Inc., Natick, Massachusetts, USA) and allows the user to choose either a full or a regulated deficit irrigation (RDI) strategy. The latter is carried out throughout Ψ_{stem} measurements and defining Ψ_{stem} thresholds at different phenological stages. The necessary inputs are the following: (1) weather data on a daily basis; (2) day of the year (DOY) corresponding to vine bud-break. These parameters are used to simulate vine phenology, which was defined by using the accumulated degree days needed to fulfil each vine development phase requirements; (3) number of berries per m^2, which is set as default; (4) fraction of canopy light interception (f_{IPAR}), which can be either simulated or manually introduced. Daily f_{IPAR} (f_{IPARd}) was calculated by using the hourly model of light interception of Oyarzun *et al.* (2007), and was used to calculate the crop coefficient for total canopy cover (K_{cfc}); and (5) midday stem water potential (Ψ_{stem}). Daily values of Ψ_{stem} (Ψ_d) were obtained as integration of the diurnal course using a polynomial relationship which related midday Ψ_{stem} and Ψ_d.

Soil evaporation (E) was empirically estimated, being a function of the reference evapotranspiration (ET_o), irrigation doses, phenology and f_{IPARd}. Maximum plant hydraulic conductance (C_{max}) was calculated according to an analogue Ohm's law for a full canopy cover, and is a function of the

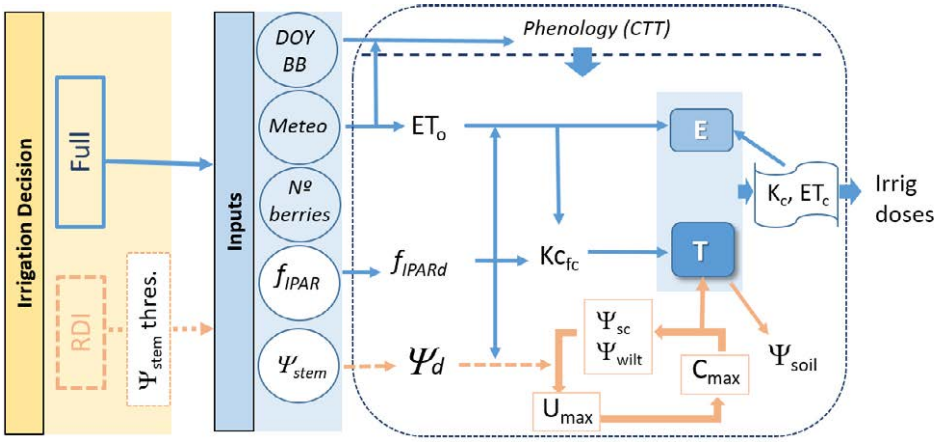


Figure 1. Diagram for the inter-relation between vine-water relation parameters of the model. Abbreviations: Full and RDI = full and regulated deficit irrigation strategies, respectively; DOY BB = day of the year at bud-break; Meteo = weather data, no. berries = number of berries per m^2, which, as default, is considered to be 900 berries·m^{-2}; f_{IPAR} and f_{IPARd} = measured and daily fraction of intercepted radiation, respectively; Ψ_{stem} and Ψ_d = midday and daily stem water potential, respectively; CTT = cumulative thermal time; ETo = reference evapotranspiration; K_{cfc} = full canopy crop coefficient; Ψ_{sc} and Ψ_{wilt} = water potential that does not limit transpiration and at wilting point, respectively; Ψ_{soil} = soil water potential; C_{max} = maximum vine hydraulic conductance; U_{max} = maximum water uptake; T = vine daily transpiration; E = soil evaporation; K_c = irrigation crop coefficient; IrrigDoses = irrigation doses of water applied.

maximum water uptake of the crop (U_{max}), the lowest plant water potential that does not limit transpiration (Ψ_{sc}) and soil water potential at field capacity (Ψ_{fc}). Ψ_{wilt} corresponds to the plant water potential when transpiration approached zero. These parameters were used to calculate vine transpiration (T) and the irrigation doses on a weekly basis.

The case study

The study was carried out during the growing seasons 2016-2017 in a 100 ha commercial vineyard located in Raïmat (41°39'43.88"N, 0°30'59.27"E) (Lleida, Spain). The vineyard consisted of 52 irrigation sectors ranging between 0.6 to 2.6 ha with the varieties *Cabernet Sauvignon* (CAB) (38), *Tempranillo* (TMP) (12) and *Syrah* (SYR) (2). The area has a typical Mediterranean climate with approximately an average annual rainfall from 230 to 300 mm and an annual accumulated reference evapotranspiration (ET_o) from 1,100 to 1,400 mm. An airborne remote sensing image acquired in July of the previous growing season was used to determine the spatial variability of the vegetative canopy vigour using the normalized difference vegetation index (NDVI). The camera had four independent narrow bandwidth spectral filters of 20 nm width (FMHW) at the specific band centre of 450 (blue), 550 (green), 680 (red), and 780 (near-infrared) nm. The image resolution was 2,048×2,048 pixel with 14-bit digitization and equipped with a 24-28 mm fixed focal length lens yielding an angular field of view (FOV) of 17°. Flight altitude was at ~1,500 m above ground level obtaining images at 0.50 m spatial resolution. The NDVI map was used to classify the vineyard in three different categories (Low, Medium and High) u
sing a *K-means* clustering classification analysis with the software ArcMap (version 10.5, ESRI Inc. Redlands, CA, USA) (Figure 2). Each irrigation sector was then arbitrarily classified according to the predominant NDVI category. Seven 'smart points' which visually corresponded to the most representative location by variety/category were also identified. In each 'smart point', vine physiological and structural measurements (inputs of the model) were conducted every one to two weeks and were introduced into the model in order to calculate the irrigation doses. The f_{IPAR} was measured using a portable ceptometer (Accupar Linear PAR, Decagon Devices, Inc., Pullman, WA, USA). Ψ_{stem} was measured at noon with a pressure chamber (model 3005; Soil Moisture Equipment Corp., Santa Barbara, CA, USA) according to the recommendations of McCutchan and Shackel (1992). Irrigation decisions were then conducted independently in each irrigation sector, according to their classification category. The volume of water applied was also measured on a weekly basis in each location using digital water meters.

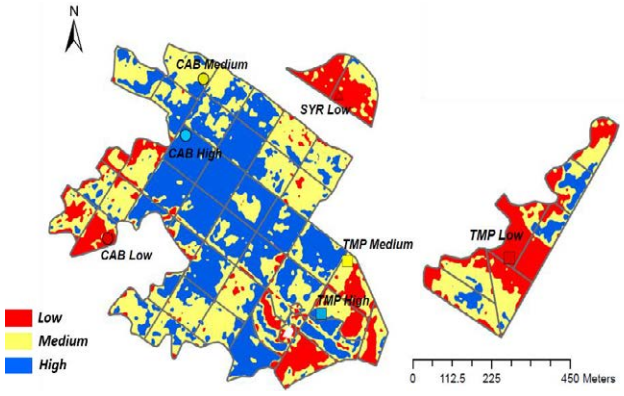


Figure 2. Spatial variability of a vineyard, based on a NDVI map (July 2015), classified in three categories: Low, Medium and High. Squares, circles and triangles show the location of the 'smart points'.

Definition of the irrigation strategies

Irrigation scheduling consisted in adopting RDI strategy for all varieties. To achieve it, Ψ_{stem} thresholds were pre-defined bearing in mind optimal values provided by previous studies in those varieties. In 2016, Ψ_{stem} during the phenological stage II (fruit-set to veraison) was defined at -0.7 MPa for all varieties, while for the stage III (veraison to fruit-set), thresholds were defined at -1.2, -1.1 and -1.0 MPa for CAB, TMP and SYR, respectively. In 2017, pre-defined Ψ_{stem} were more negative. Respectively for CAB, TMP and SYR, thresholds were -0.8, -0.9 and -0.8 MPa (stage II), and -1.2, -1.1 and -1.0 MPa (stage III).

Evaluation of the spatial variability

The evaluation of vineyard spatial variability after carrying out a differential precision irrigation management was analysed with time-series of Landsat-8 NDVI maps. Time-series of Landsat-8 NDVI and the fraction of reference evapotranspiration (ETrF) (ETrF is obtained as the ratio of the computed ET_{inst} from each pixel to the instantaneous tall crop evapotranspiration, computed from weather data) were downloaded from the Google evapotranspiration application EEFlux (https://eeflux-level1.appspot.com), which uses the mapping evapotranspiration at high resolution with

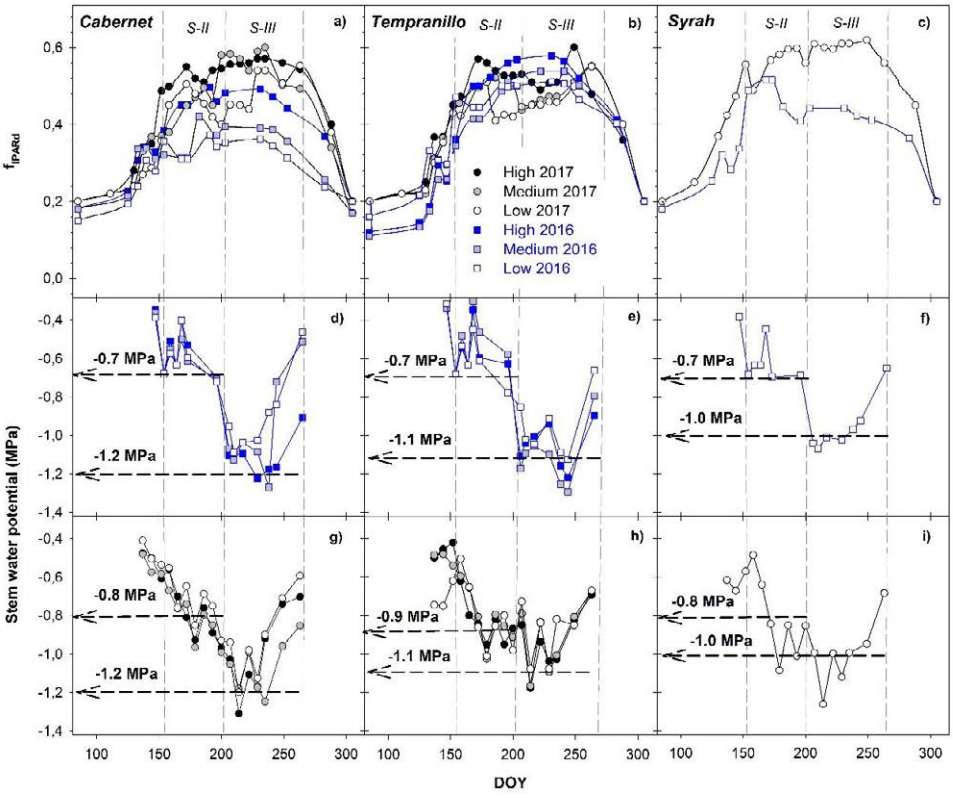


Figure 3. Seasonal pattern of daily fraction of intercepted radiation (f_{IPARd}) (A-C) and midday stem water potential (Ψ_{stem}) (D-I) measured in each 'smart point' during the growing seasons 2016 and 2017. Dotted arrows indicate the pre-defined Ψ_{stem} thresholds, and dotted lines represent the phenological stages.

internalized calibration (METRIC) model to estimate the evapotranspiration (ET). In this study, ETrF was converted to relative values using data of the growing seasons 2015 to 2017.

Results and discussion

Figure 3 provides an insight into the seasonal pattern of f_{IPARd} and Ψ_{stem} at the 'smart points' during 2016 and 2017. Seasonal differences in f_{IPARd} were evident for both years. In 2016, f_{IPARd} of TMP was higher than the other two varieties. CAB and TMP also showed significant differences between categories, indicating in both cases that f_{IPARd} of the *High* category trend to be higher than others. In 2017, f_{IPARd} of CAB increased with respect to the previous growing season, and differences between categories were also detected. f_{IPARd} of SYR was higher during 2017, which attained a maximum value of 0.62. Figure 3D-I also show the seasonal trend of midday Ψ_{stem} with the pre-defined intended thresholds at each phenological stage. It should be pointed out that despite the clear differences in f_{IPARd} among categories, this study was able to lead the Ψ_{stem} to similar values throughout the season according to the pre-defined Ψ_{stem} thresholds, and independently of their water demand. This was achieved carrying out a differential irrigation management.

Results showed differences in the total amount of water applied between varieties and NDVI categories (Table 1). Generally, TMP received the highest amount of water in all the irrigation sectors, together with SYR in 2017. This higher amount of water might be explained by the following two reasons: (1) in 2016, f_{IPARd} of TMP was the highest, implying a higher water demand, and (2) in 2017, Ψ_{stem} of TMP were slightly higher than those of CAB. The amount of water applied in Syrah during 2017 was proportional to f_{IPARd}. The coefficient of variation (C_V) of water applied between irrigation sectors was 10.6 and 20.8% respectively in 2016 and 2017, which represents a difference in the total amount of water applied of 84.1 (2016) and 125.9 (2017) mm between the maximum and minimum categories. Analysing differences between categories of the same variety, maximum differences were found in 2017, which reached 30.2 and 33.5 mm for CAB and TMP, respectively. These results exemplify important water savings, highlighting the necessity to conduct a precise irrigation management according to differences in water demands among irrigation sectors. Table 1 also shows the accumulated actual (T_a) and potential (T_p) transpiration for each category during the growing season. T_p was linearly dependent of f_{IPARd}, with a coefficient of determination (R^2) of 0.71. On the other hand, the crop water stress index (CWSI), which was derived as ($1-T_a/T_p$), showed a significant regression with the seasonal average Ψ_{stem} ($\Psi_{stem} = -7.856 \times CWSI^2 + 3.899 \times CWSI - 1.088$, $R^2=0.88$).

The spatial variability within and between irrigation sectors was also analysed using time-series of NDVI. The C_V showed that after two consecutive years of applying a differential irrigation management, C_V slightly decreased from 31.1 to 28.3% in CAB, from 28.0 to 26.1% in TMP, and from 16.9 to 9.5% in SYR. Although these differences may not be significant, it seems there is a trend, which may be pronounced for longer-term studies.

Differences in f_{IPARd} between varieties and irrigation sectors were also seen in NDVI maps. Figure 4 shows Landsat-8 NDVI and ETrF maps of the vineyard at two different phenological stages for years 2015 to 2017. Despite the regressions between NDVI and ETrF having a significantly high R^2, there is a still a percentage of the ETrF not explained by the vegetative canopy growth. This is because it is possible to find winegrapes with a high vegetative growth (measured by NDVI), but under water stress. According to this, when NDVI is used to estimate yield in heterogeneous crops, it may happen that the regression is not as good as regressing yield with ETrF. In fact, in this study only the regression between ETrF and yield was significant ($R^2=0.50$, $P<0.001$) (data not shown).

Table 1. Analysis of differential irrigation management. Irrigation is the accumulated water applied (mm); T_a and T_p correspond to the accumulated actual and potential simulated transpiration (mm); CWSI (crop water stress index) was calculated as ($1-T_a/T_p$); Ψ_{stem} is the average daily stem water potential during the growing season; Cv is the coefficient of variability of the NDVI. Rainfall during the irrigation period (1st May-25th September) was 57.9 and 117.6 mm for 2016 and 2017, respectively.

	Irrigation (mm)	T_a (mm)	T_p (mm)	CWSI	Ψ_{stem} (MPa)	NDVI	C_V NDVI (%)	
2015								
CAB High	-	-	-	-	-	0.62±0.02 a	22.5	31.1
CAB Medium	-	-	-	-	-	0.59±0.01 ab	29.2	
CAB Low	-	-	-	-	-	0.56±0.02 a	35.4	
TMP High	-	-	-	-	-	0.65±0.00	-	28.0
TMP Medium	-	-	-	-	-	0.62±0.01	35.1	
TMP Low	-	-	-	-	-	0.58±0.03	17.2	
SYR Low	-	-	-	-	-	0.53±0.08	16.9	
2016								
CAB High	262.0	243.7	383.6	0.37	-0.70	0.54±0.01 a	17.2	27.5
CAB Medium	282.3	201.8	309.3	0.35	-0.69	0.51±0.01 ab	27.7	
CAB Low	279.8	183.4	283.3	0.35	-0.68	0.49±0.01 a	31.7	
TMP High	313.0	268.3	422.6	0.37	-0.76	0.57±0.00	-	24.3
TMP Medium	327.5	244.5	383.1	0.36	-0.70	0.54±0.01	33.0	
TMP Low	313.0	259.6	396.3	0.34	-0.68	0.50±0.02	13.1	
SYR Low	243.4	260.2	371.3	0.29	-0.64	0.49±0.02	9.36	
2017								
CAB High	226.6	279.6	434.6	0.36	-0.69	0.57±0.02	29.2	28.3
CAB Medium	200.0	244.2	395.3	0.38	-0.74	0.53±0.01	25.6	
CAB Low	196.5	244.5	373.8	0.35	-0.66	0.54±0.01	29.5	
TMP High	322.1	327.9	416.2	0.21	-0.61	0.54±0.00	-	26.1
TMP Medium	288.6	301.3	384.9	0.22	-0.62	0.54±0.01	33.8	
TMP Low	255.2	279.5	359.4	0.22	-0.61	0.48±0.01	8.3	
SYR Low	322.4	305.8	477.5	0.35	-0.70	0.49±0.05	9.5	
C_V 2015 (%)	-	-	-	-	-	-	32.0	
C_V 2016 (%)	10.6	13.6	13.7	6.6	-	-	28.6	
C_V 2017 (%)	20.8	11.1	9.9	25.9	-	-	28.4	

Conclusions

Precision irrigation scheduling of winegrapes could benefit from a model like the one presented in this paper. Based on different winegrape physiological parameters, the model calculates the doses of water to be applied in order to reach a desired level of water stress. This study demonstrated the feasibility of conducting differential irrigation management in a commercial vineyard, which entailed varying the amount of water applied between irrigation sectors of 10.6 and 20.8% respectively in 2016 and 2017, and important water savings.

One of the shortcomings of using this approach may be the need to go to the vineyard and measure f_{IPAR} and Ψ_{stem}. Further improvements of this model will be related to the exploration of assimilating

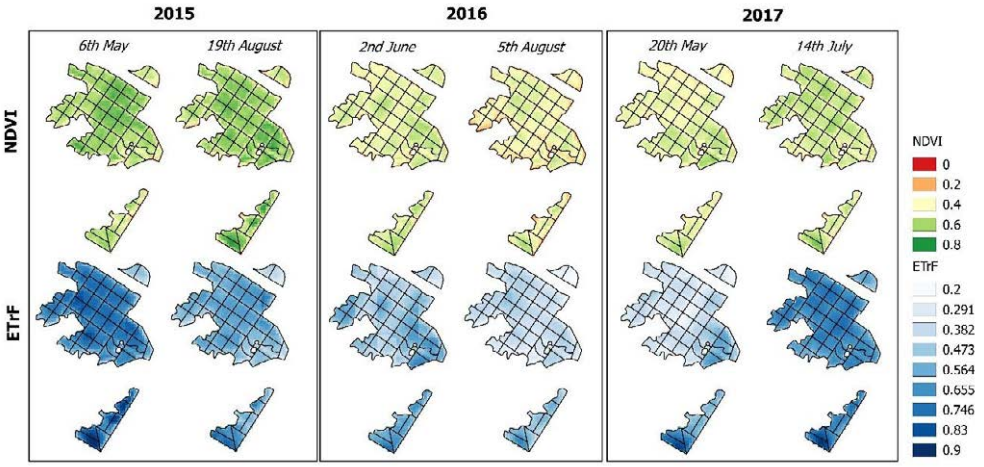


Figure 4. Maps of NDVI and relative fraction of reference ET (ETrF), obtained from the platform EEFLUX using the energy balance METRIC model.

into the model estimates of CWSI ($1 - T_a / T_p$) through remotely sensed energy balance models and through the use of biophysical variables such as fAPAR (fraction of absorbed photosynthetically active radiation) and LAI (leaf area index) derived from Sentinel-2.

Acknowledgements

This research was supported by the Operational Group 'VINECO: Profitability of applying new technologies to enhance irrigation efficiency in a 100-ha organic and conventional vineyard' (16.01.01) funded by the Catalan Department of Agriculture (DARP) and the European Association for the Innovation. We also thank Ms Emily Buerer for her contribution in the revision of the paper.

References

Basile, B., Marsal, J., Mata, M., Vallverdú, X., Bellvert, J., Girona, J., 2011. Phenological sensitivity of Cabernet Sauvignon to water stress: Vine physiology and Berry composition. American Journal of Enology and Viticulture 62 452-461.

Bellvert, J., Marsal, J., Mata, M., Girona, J., 2016. Yield, must composition, and wine quality responses to preveraison water deficits in sparkling base wines of Chardonnay. American Journal of Enology and Viticulture 67(1) 1-12.

Girona, J., Mata, M., del Campo, J., Arbonés, A., Bartra, E., Marsal, J., 2006. The use of midday leaf water potential for scheduling deficit irrigation in vineyards. Irrigation Science 24 115-127.

Marsal, J., Girona, J., Casadesús, J., Lopez, G., Stöckle, C.O., 2013. Crop coefficient (Kc) for apple: comparison between measurements by a weighing lysimeter and prediction by CropSyst. Irrigation Science 31 455-463.

Marsal, J., Johnson, S., Casadesús, J., Lopez, G., Girona, J., Stöckle, C., 2014. Fraction of canopy intercepted radiation relates differently with crop coefficient depending on the season and the fruit species. Agricultural and Forest Meteorology 184 1-11.

McCutchan, H., Shackel, K.A., 1992. Stem water potential as a sensitive indicator of water stress in prune trees (*Prunus domestica L. cv French*). Journal of the American Society of Horticultural Science 117 607-611.

Oyarzun, R.A., Stöckle, C.O., Whiting, M.D., 2007. A simple approach to modeling radiation interception by fruit-tree orchards. Agricultural and Forest Meteorology 142 12-24.

Stöckle, C.O., Donatelli, M., Nelson, R., 2003. CropSyst, a cropping system simulation model. European Journal of Agronomy 18 289-307.

Deep learning for in-field image-based grapevine downy mildew identification

J. Boulent[1,2*], M. Beaulieu[2], P.-L. St-Charles[2], J. Théau[1] and S. Foucher[2]
[1]*Department of Applied Geomatics, Université de Sherbrooke, 2500 Boul. de l'Université, Sherbrooke, QC J1K 2R1, Canada; justine.boulent@usherbrooke.ca*
[2]*Vision and Imagery Team, Computer Research Institute of Montréal, 405 Ogilvy Avenue #101, Montréal, QC H3N 1M3, Canada*

Abstract

Since the early 2010's, great progress has been made in image classification through deep learning and convolutional neural networks (CNNs). This allows us to consider the development of automatic high-precision scouting tools for agricultural use. The objective of this project is the automatic identification of grapevine downy mildew – *Plasmopara viticola*. To assess the potential of a commercial prototype, several wine-growing regions and grape varieties were studied. Images were acquired under in-field conditions. The selected CNN architecture used for developing this model was ResNet. Multiple depths of ResNet have been tested to determine which one achieves the best score for this study case and for the amount of data collected. The accuracies reached by five different depths were surprisingly close. The best accuracy is achieved by a ResNet with 18 layers, with 95.48% on an independent test set.

Keywords: automatic disease identification, grapevine downy mildew, convolutional neural networks, deep residual learning, network depth

Introduction

Automatic identification of crop pests and diseases using visible range images is considered to be challenging, especially when it comes to an application under in-field conditions. Barbedo (2016) highlighted some factors making such automatic identification with standards techniques unfeasible under these conditions: (1) the busy backgrounds and the lack of clear border around symptoms making segmentation difficult; (2) the variation of image capture conditions; (3) the fact that the same disease may produce several symptoms and that, conversely, the same symptoms may be caused by different diseases; and (4) the fact that several problems can affect the same plant simultaneously. To address these challenges, innovative image processing techniques must be used. Convolutional neural networks (CNNs) are a type of network used mostly in computer vision problems. They achieve good results in many classification and detection applications and they are now being investigated for diagnostic aid in crop pests and diseases control. Sladojevic *et al.* (2016) and Mohanty *et al.* (2016) were among the first to publish extensive studies on multiple disease detection on different crop types with CNNs. They achieved, respectively, 96.3 and 99.4% accuracy rate. However, these performances were obtained on image sets acquired under conditions that do not match those found on the field. In fact, Mohanty *et al.* (2016) reported a drastic drop in performance (31.69%) while testing images acquired under different conditions compared with those used in training. To help the model to generalize and operate under field conditions, some studies proposed to use image datasets that reflect the complexity of a real agricultural setting. Fuentes *et al.* (2017) trained a model that recognizes nine different types of tomato diseases and pests with an accuracy of 85.98%. DeChant *et al.* (2017) implemented a system able to identify northern leaf blight infected maize plants with an overall accuracy of 96.70%. Both authors demonstrated the usefulness of images acquired directly from the field.

DOI 10.3920/978-90-8686-888-9_16, © Wageningen Academic Publishers 2019

In a similar fashion, this study uses a set of field images from several phenological stages, disease development levels and grape varieties. The purpose of the set is to train CNNs for identification of grapevine downy mildew – *Plasmopara viticola*. This disease is found throughout the world's vineyards and leads farmers to use high doses of fungicides to limit its propagation and to avoid heavy crop losses (Ambiaud, 2012). With an automatic tool for downy mildew identification, farmers could have a better understanding of the phytosanitary status of their vineyard and could therefore set up optimised and more sustainable pesticide management.

Another aspect of this work is the investigation of neural network architectures with an emphasis on the impact of network depth on the classification accuracy. It is assumed that the depth of networks – i.e. the number of layers with parameters – leads to higher level features. In return, these features improve the aim of discriminating between more complex elements and are less sensitive to large irrelevant variation like background content (LeCun *et al.*, 2015). In consequence, deep architectures should give better accuracy compared to shallow ones. However, the number of parameters to estimate gets larger as the network gets deeper and this pushes the network to overfit if insufficient data are used during the training phase. This is a real problem in the context of crop diseases and pests identification because datasets are usually small and unevenly annotated, meaning that some classes can have few labels for training. Also among other things to consider is the operational level of the system to provide farmers with quick diagnosis. In this case, smaller networks may be proven more efficient. Nonetheless, the selection of a viable architecture for real time application depends on both factors: accuracy and speed.

The objective of this work was to determine the optimal network depth for this study case, i.e. the one that provides the highest accuracy on the test set – demonstrating its ability to generalize. There exists a broad variety of architectures to choose from, but the study focuses on the Residual Network (ResNet, He *et al.*, 2015). A performance comparison between the five depths of ResNet for grapevine pest and disease identification was done together with an evaluation of the ability of a single model to work on several grape varieties, in several wine-growing regions, under various acquisition conditions and on a wide range of symptoms.

Materials and methods

Data acquisition

Images used to train and evaluate the models were acquired in 2017 and 2018, between June and September. The 2017 acquisitions took place in the vineyard of Agriculture and Agri-Food Canada's experimental farm in Frelighsburg (45°2'60'N, 72°51'36'W), located in the Montérégie wine-growing region of Canada. The acquisitions covered three grape varieties: Chancellor, Geisenheim and Frontenac. The 2018 acquisitions were carried out in these plots but also in two other French wine-growing regions: Cognac and Languedoc-Roussillon. Data were acquired around the cities of Cognac (45°41'24'N, 0°19'12'W), Carcassonne (43°12'36'N, 2°19'48'E) and Limoux (43°2'60'N, 2°12'36'E) in commercial vineyards. The acquisitions covered three grape varieties: Ugni Blanc, Chardonnay and Chenin.

The photos were taken using standard cameras in the field, either by hand or using a 2 m pole, which allowed a more global and elevated view of the vines. Leaves, stems, grapes and background were included in the photos (without prior arrangement), under various weather conditions and exposures. The only requirement was that the vines should be free of water drops. These acquisitions produced a database of 1,192 images.

Dataset

Two separate datasets (for training and testing) were created based on the captured images. These datasets are based on different plots. This geographical separation provided independence between the training and test sets. The training set contained 899 images, from which 80% were used for

training the models and 20% for validating their performance. The test set contained 293 images from other plots. However, the geographical separation was impossible for images acquired in 2017 in Montérégie. Therefore, a random separation into 80% for training and 20% for testing was performed for that subset. The training set contains photos of Ugni Blanc, Chancellor, Geisenheim and Frontenac varieties. The test set contains photos of these four grape varieties plus Chenin and Chardonnay. The proportion of samples for each location and grape variety is described in Table 1. The images were then annotated manually by adding points in salient image regions in order to label them according to their local content. Even if the main goal is the identification of downy mildew, seven labels were defined, which include other phytosanitary problems. This was done to avoid confusion between these problems and also to assess the discrimination capacity of the CNNs. Those labels were defined by their plot's phytosanitary state: Downy Mildew (DM), Japanese Beetle damage (JB, *Popillia japonica*), Deficiencies (D), Healthy (H), Vine Leaf Blister (VLB) caused by the mite *Colomerus vitis*, Dry Plant Organ (DPO) and Background (B). Several deficiencies, including phosphorus, magnesium and iron, were grouped under the same class due to the lack of samples to label them individually. The DM class includes samples from several disease stages (sporulation, oil spots and drying) and levels of contamination.

Samples of each class were created from 318 by 318 pixels regions of interest for each annotated point. The training set was made up of 6,055 samples, the validation set of 1513 samples and the test set of 1,883 samples (Table 2). Figure 1 illustrates the intra-class diversity of the 7 classes with some examples for each of them. Data augmentation was performed on the training set with random rotation, flip and zoom transformations. All the samples were resized to 224×224 pixels and normalised.

Training and validation phases

This study focused on ResNet networks (He *et al.*, 2015). This is now an established architecture thanks to its performance and fast convergence. The architecture comes in different depth configurations: shorter ones (18 layers), longer ones (34 and 50 layers), and very deep ones (101 and 152 layers). Deep residual learning was introduced to overcome the degradation problem as observed by He *et al.* (2015). The degradation problem corresponds to the saturation and rapid degradation of the training accuracy observed with the increasing network depth. To improve convergence while having

Table 1. Samples origin and grape varieties per dataset and class (in %). The percentage distribution between grape varieties in Frelighsburg is broadly estimated.

	Downy Mildew		Japanese Beetle		Deficiencies		Healthy		Vine Leaf Blister		Dry Plant Organ		Back-ground	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
Carcassonne *Chardonnay*	0	29	0	0	0	51	0	5	0	0	0	33	0	13
Cognac *Ugni Blanc*	67	29	0	0	100	4	48	14	100	76	48	29	38	19
Frelighsburg *90% Chancellor 5% Geisenheim 5% Frontenac*	33	39	100	100	0	18	52	77	0	0	52	29	62	66
Limoux *Chenin*	0	3	0	0	0	27	0	4	0	24	0	9	0	2

Table 2. Samples per dataset and class.

Label	Train	Validation	Test
Downy Mildew	2,038	510	451
Japanese Beetle	330	82	227
Deficiencies	400	100	248
Vine Leaf Blister	142	35	42
Dry Plant Organ	94	23	45
Healthy	2,837	709	716
Background	214	54	154
Total	6,055	1,513	1,883

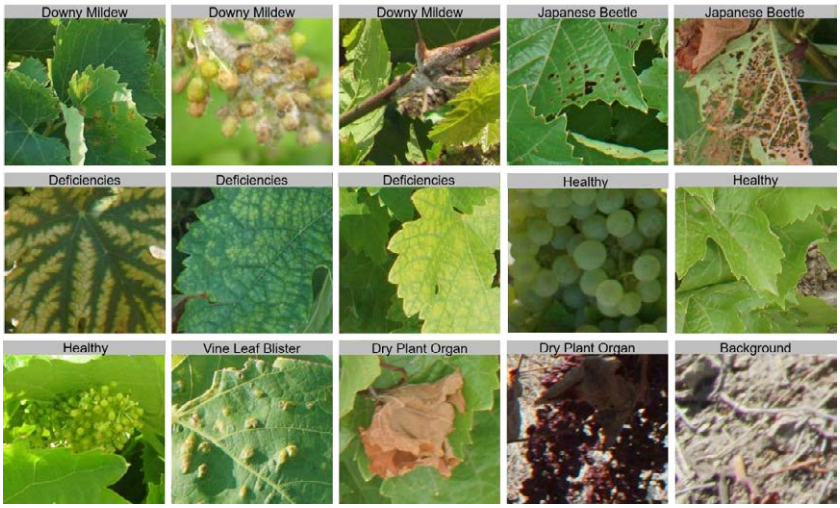


Figure 1. Illustration of classes and their diversity.

very deep architectures, the network integrates shortcut connections. These connections provide an identity map from a given layer to a deeper one and whose value is added to the output of the deeper layer. This allows the intermediate layers of a block to learn a residual function referring to the block input and alleviate the degradation problem. The network's number of parameters ranges from roughly 11 M for ResNet-18, to over 58 M for ResNet-152.

To compare their performance on the vine disease classification problem, the five architectures were trained with the same hyperparameters. The learning rate was initialized at 0.0005 and decreased by 90% every 10 epochs. This small value was selected due to the use of transfer learning. The batch size was 32. The optimizer used was Adam (Kingma and Ba, 2015). The networks were trained for 30 epochs.

The training phase was initialized with ImageNet pre-trained weights to allow transfer learning. The goal of transfer learning is to use knowledge learned from tasks for which a large amount of labelled data is available to improve performance in settings where less data is available. This is possible because of the hierarchical nature of the CNNs features. The first layers represent general features, likes corners, edges and colour combinations. The deeper layers represent complex and class-specific

structures (Zeiler and Fergus, 2013). The features from the first layers then can be used to describe any object of study. And a partial or total re-training enables the weights to be fine-tuned for the target dataset. This technique reduces processing time and leads to more generalizable models. In this study, all layers were fine-tuned.

Each model was trained five times to ensure the reproducibility of the results. For each training session, the seeds determining the initial state of the random number generators have been fixed. Random values were used among other things here to determine the ordering of the training samples and the augmentation operations to apply to these samples. Five seeds have been fixed with distinct arbitrary values. Each network was trained successively with these five seeds. This guarantees that the performance differences are due to the networks depth differences and not to random initialization collateral effects. For each ResNet, the seed that provides the best overall accuracy on the validation set was kept, and the five best models were then evaluated and compared with the test set.

The processing was implemented with a publicly available deep learning framework[3] based on PyTorch[4] developed at the Computer Research Institute of Montréal. Figure 2 gives an overview of the processing setup.

Results and discussion

The 25 training sessions resulted in similar and high validation accuracies, fluctuating from 97.28 to 98.48%. The same seeds did not produce the best model for the five depths. The best epoch value for each depth according to the five different seeds varied from 0.46 to 0.96%, demonstrating the importance of running several training sessions. The best epoch value on the validation set for the five depths varied from 98.11% for ResNet-34 to 98.48% for ResNet-18. A larger disparity occurred on the accuracy achieved on the test set, reaching from 93.73% for ResNet-101 to 95.48% for ResNet-18 (Table 3). The ResNet-18 achieved the best results. This performance remains relative to this study case and to the parameters specified for training. Indeed, the five depths produced almost similar performance. Given the marginal differences between the results, different seeds could lead to a different ranking. However, the good performance of a low depth network indicates that low-level features are sufficient for this classification task. Adding more layers did not result in better features or discrimination. These findings are very promising on an operational level, as a shallow model allows faster executions with limited computing power.

[3] Link to the framework used: https://github.com/plstcharles/thelper.
[4] https://pytorch.org.

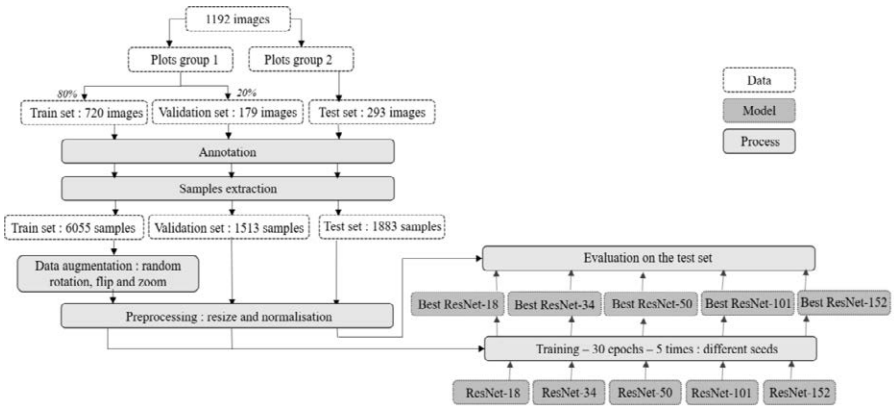


Figure 2. Processing setup overview.

Table 3. Accuracy of the best models for each ResNet (%).

	ResNet-18	ResNet-34	ResNet-50	ResNet-101	ResNet-152
Best model epoch	30	28	28	24	29
On validation set (%)	98.48	98.11	98.24	98.14	98.27
On test set (%)	95.48	94.47	95.43	93.73	94.10

The classification report shows that three classes had very good performance scores (F1-Score <96%): DM, JB and H (Table 4). The model performed less well for two of the seven classes: VLB and D, with a respective F1-Score of 88.00 and 89.08%. For VLB, this can be explained by the small number of samples for training. Also, 14.29% of misclassified VLB were classified as H, the class with the biggest number of false positives (Figure 3). This amount of false positive in H may be due to the diversity covered by this class. In fact, depending on the grape varieties and phenological stages, the leaves can have colours ranging from yellow-green to dark green, with undulating to smooth textures. A healthy leaf and a leaf in early stage of VLB contamination have close characteristics, both in terms of colour and texture which could explain in part the observed confusion (Figure 4A). For D class, there is a

Table 4. Classification report from ResNet-18 evaluation on the test set.

	Precision (%)	Recall (%)	F1-score (%)	Support count
Downy Mildew	96.85	95.34	96.09	451
Japanese Beetle	99.53	94.27	96.83	227
Deficiencies	92.98	85.48	89.08	248
Healthy	95.32	99.58	97.40	716
Vine Leaf Blister	100.00	78.57	88.00	42
Dry Plant Organ	93.48	95.56	94.51	45
Background	90.53	99.35	94.74	154
Weighted average	95.56	95.49	95.43	1,883

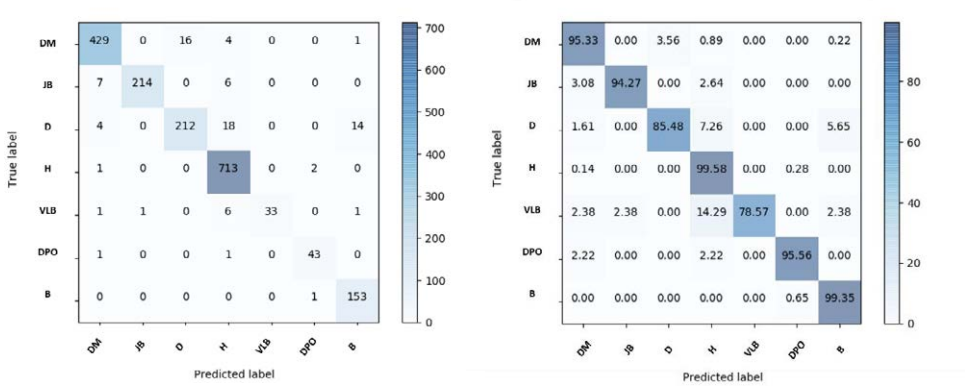


Figure 3. Original (left) and normalised (right) confusion matrices (%) from ResNet-18 evaluation on the test set.

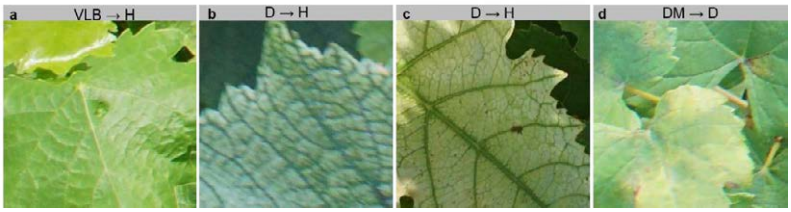


Figure 4. Examples of wrong predictions. True Label → Predicted Label. VLB = Vine Leaf Blister; DM = Downy Mildew; D= Deficiencies; H = Healthy.

lack of samples for each deficiency in its different stages. In fact, it was found that deficiencies samples with intense and advanced symptoms were present only in the test set. Therefore, the model had never been confronted with samples with similar characteristics. From the 36 misclassified samples of class D, 23 were advanced symptoms (Figure 4B and 4C). DM class achieved a F1-Score of 96.09%. Some false negatives can be explained by spots with similar patterns found on D class (Figure 4D). The capacity to explain some of the errors gives the positive sense that they can be corrected, in part by adding more images covering those more complex cases.

The model was also able to process unknown grape varieties. Chardonnay and Chenin were only shown in the evaluation phase. 12.8% of Chardonnay samples were misclassified but 59% of these errors correspond to those presented in Figure 4B and 4C. As for Chenin samples, only 3.14% were misclassified.

Finally, only 3.52% of the misclassified samples were not leaves. Even if leaves represent the vast majority of annotations, this underlines the capacity of the model to process all the vine's organs and different backgrounds.

Conclusions

In this study, through the CNNs, most of the challenges making the automatic identification of pests and diseases unsuitable under in-field conditions (see Introduction) seem to have been handled. Due to CNNs ability to manage busy backgrounds and foliage complexity, prior segmentation was not necessary and images of unarranged plants were processed. Next, the images collected over several days, at different hours and under different lighting conditions and poses were all processed by this model. The classification errors did not seem to be linked to these factors. The variety of disease symptoms over all plant organs could also be managed – the variety of grapevine downy mildew symptoms being a good example to assess this aspect. Going even further, this study demonstrates that CNNs can identify diseases and pests on several grape varieties, with different vine training systems, at different phenological stages.

The model produced here is not yet fully operational. In particular, it should be adapted to process complete images and not just pre-selected samples. It should also handle cases where a single plant may be affected with multiple phytosanitary problems. Given the advances in other related fields, CNNs will undoubtedly be able to meet these challenges. The fact that networks with low depth such as ResNet-18 produce such good results is very encouraging for operational deployment. CNNs are undeniably a technique that will allow the commercial development of tools for the automatic identification of pests and crop diseases. With such tools, it can be expected that scouting will be less painful for farmers and that they will be able to use it to change their phytosanitary practices to more sustainable and environmentally friendly ones.

Acknowledgements

This research project was funded by the Centre de Géomatique du Québec (CGQ), Microdrones and Mitacs. We also thank the Ministère de l'Economie, de la Science et de l'Innovation (MESI) of the province of Québec, Agriculture and Agri-Food Canada, the Chambre d'Agriculture de l'Aude, the Fédération Départementale de défense contre les Organismes Nuisibles (FEDON) de l'Aude as well as all the owners and wine growers who gave us access to their vineyards and/or guided us in the image acquisition process.

References

Ambiaud, E., 2012. Pratiques phytosanitaires dans la viticulture en 2010 (Phytosanitary practices in viticulture in 2010). Agreste, Ministère de l'agriculture et de l'alimentation, 288, 1-8.

Barbedo, J., 2016. A review on the main challenges in automatic plant disease identification based on visible range images. Biosystems Engineering. 144 52-60.

DeChant, C., Wiesner-Hanks, T., Chen, S., Stewart, E.L., Yosinski, J., Gore, M.A., *et al.*, 2017. Automated Identification of Northern Leaf Blight-Infected Maize Plants from Field Imagery Using Deep Learning. Phytopathology. 107 1426-1432.

Fuentes, A., Yoon, S., Kim, S.C., Park, D. S., 2017. A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition. Sensors. 17 2022.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep Residual Learning for Image Recognition. In IEEE Conference on Computer Vision and Pattern Recognition. pp 770-778.

Kingma, D., Ba, J., 2015. Adam: A Method for Stochastic Optimization. In: Ithaca, NY, USA, International Conference on Learning Representation.

LeCun, Y., Bengio, Y., & Hinton, G., 2015. Deep learning. Nature, 521 436-444.

Mohanty, S.P., Hughes, D.P. and Salathé, M. 2016. Using Deep Learning for Image-Based Plant Disease Detection. Frontiers in Plant Science. 7 1419.

Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D. and Stefanovic, D., 2016. Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. Computational Intelligence and Neuroscience. 2016 11. http://dx.doi.org/10.1155/2016/3289801

Zeiler, M.D., Fergus, R., 2013. Visualizing and Understanding Convolutional Networks. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (Eds.) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, 8689.

Monitoring site-specific spraying in vineyards from a prescription map obtained with a UAV

J. Campos*, M. Gallart, J. Llop and E. Gil
Universitat Politècnica de Catalunya, Department of Agro Food Engineering and Biotechnology, Esteve Terradas, 8, 08860 Castelldefels, Spain; javier.campos@upc.edu

Abstract

This work presents the development of an actual site-specific spraying map of a variable rate sprayer prototype, which works following a prescription map obtained with a UAV. Trials were conducted in Torrelavit (NE Spain) in a representative vineyard of 3 ha of variety Merlot. Prescription map was obtained by merging the canopy data obtained with an multispectral camera mounted on an UAV, and the information provided by the app DOSAVIÑA®. This prescription map was then uploaded into a variable sprayer prototype, obtaining an actual application map, after the spray process. The actual application map allowed traceability data in all treated areas and calculated pesticide savings versus conventional spray application. Results showed a saving of around 20% of application when compared with conventional spray application.

Keywords: variable rate application, UAV, vigour map, actual map

Introduction

Considering the specific case of 'specialty crops', the most important factors to ensure an efficient spray application process are canopy characteristics such us canopy dimensions and trellis system (Rosell and Sanz, 2012; Salcedo *et al.*, 2015).
Several studies have been carried out with the aim of adapting the volume application rate to canopy dimensions (Balsari *et al.,* 2008; Gil *et al.*, 2014; Miranda-Fuentes *et al.*, 2016). All of them determined that canopy measurements were the key factor for a successful spray application.
The high degree of variability in the vine canopy characteristics within the field, increases the difficulty to obtain a general solution. When a uniform canopy structure is assumed for the whole vineyard, differences in the total amount of pesticide arriving at the canopy can occur, which reduces the effectiveness of spray application.
Recent advances in hardware and software, global navigation satellite systems, vegetation sensors and remote sensing offer opportunities to obtain canopy measurements to be used for variable application technologies (Zaman and Salyani, 2004). In the case, of variable application technology based on prescription maps, this canopy characterization has to be performed before the application. Vogel *et al.* (2005) evaluated a modified conventional boom sprayer for variable application of herbicide based on prescription maps. The system was capable of combating weeds of corn and soybean crops. Michaud *et al.* (2008) developed a variable application prototype based on maps, obtained from aerial spectral images, to combat weeds in a cranberry crop. Variable rate application technology based on prescription maps is widely used in field crops but is not used in 3D crops. For this reason, the aim of this research was to develop a variable rate application (VRA) system for vineyard sprayers based on canopy vigour maps obtained with remote sensors.

DOI 10.3920/978-90-8686-888-9_17, © Wageningen Academic Publishers 2019

Materials and methods

Experimental site
The trials were carried out on a representative vineyard of c.a. 3 ha in Torrelavit, Barcelona, Spain (X:392234, Y:4587843, ETRS89 UTM31). Vineyards of variety Merlot were trained in trellis system *Double Royat* at 1.2 m distance on the row, and 2.2 m between rows. Trials were conducted at growth stage 75-77 at BBCH scale (Meier, 1997) and 31-32 at E-L system (Coombe, 1995).

Data acquisition for canopy characterization
The vineyard was overflown with an unmanned aerial vehicle (UAV) loaded with a multispectral camera (RedEDGE, Micasense, Seattle, USA) equipped with five bands (R, G, B, RedEdge and NIR). Flight was conducted at 95 m above ground level (AGL). Overlapping zones were adjusted at 80% in the sense of flight and 60% in the transverse sense.

Adapted sprayer for variable rate application
The prototype was mounted on a trailed cross flow air sprayer (Saher, Barcelona, Spain) with a 1000 l tank and an axial fan with a diameter of 800 mm. In order to follow a prescription map, the sprayer was equipped with: (1) One pressure sensor GEMS 1200 series (Gems Sensors & Controls, Plainville, USA) to allow the adjustment of required pressure according to the prescription map; and (2) Electronic controller (Estel S.L., Barcelona, Spain), including GNSS receiver with a frequency of 1 Hz, a touchscreen and an automatic section controller. The function of the system was to determine the position of the sprayer in the field detected by the GNSS receiver, to calculate the desired volume rate, based on the previously uploaded prescription map, and to modify the working pressure in order to obtain the adjusted nozzle flow rate.

Decision Support System to determine optimal volume rate
A decision support system (DSS) was used to determine the optimal volume rate based on the canopy characteristics. The DSS used was DOSAVIÑA® (UMA-UPC, 2018). DOSAVIÑA® (Gil *et al.*, 2011) allows determination of the most accurate volume rate based on a modified version of the leaf wall area method (LWA) (Walklate and Cross, 2012).

Methodology of the process
The whole process for variable rate application based on vigour maps is shown in Figure 1. A orthophotomap was obtained from spectral images acquired with the UAV. The orthophotomap was radiometrically calibrated using four grayscale standards (22, 32, 44, and 51% grayscale reflectance) placed in the field during the flight, to transform grayscale 12-bit digital numbers to reflectance values. This new data was used to calculate the normalized differential vegetation index (NDVI) (Rouse *et al.,* 1973) (Equation 1).

$$\text{NDVI} = \frac{NIR - RED}{NIR + RED} \quad (1)$$

Due to the vineyard growth in rows, the image was segmented by an NDVI threshold in order to eliminate weeds, shadows and soil between rows.
Once the NDVI threshold was applied, an inverse distance weighting interpolation (IDW) was performed to generate a continuous NDVI map. Final processing consisted of a value clustering in 2 NDVI levels, which were later smoothed by performing a neighbor median filtering to produce the final vigor map. The process was executed using the QGIS software (QGIS Development Team, 2016). Once the vigour map was created with the two different zones, both were located in the field, corresponding to low, and high canopy vigour. In order to obtain the recommended volume rate using DOSAVIÑA® (Gil *et al.*, 2011), 15 manual measurements of canopy height and width were randomly taken in both zones to have a complete canopy characterization (Table 1).

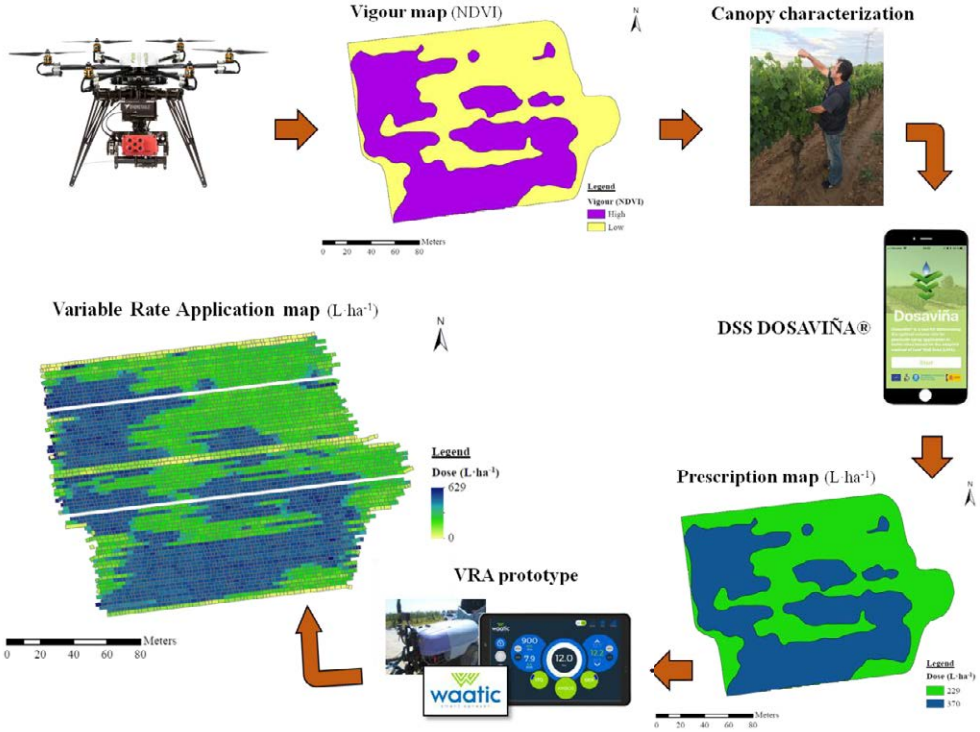


Figure 1. Scheme of the whole process: from UAV vigour map to actual variable rate application map.

Then, the values obtained were introduced into the georeferenced canopy vigour map using QGIS software (QGIS Development Team, 2016) to obtain the georeferenced prescription map.
The prescription map was transferred via USB to the touch screen previously described and installed into the sprayer. Furthermore, it was uploaded into the system information concerning working parameters for each vigour zone (Table 2). The number and nozzle type were maintained constant in all cases, using hollow cone nozzles Albuz ATR (Albuz Saint-Govain, Evreux, France).
Once the parameters were loaded on the touch screen, the spraying started. From the start, the system recorded information every second concerning the sprayer position in the field, the applied volume rate and the adjusted working pressure. At the end, the system generated an actual application map which was downloaded from the touch screen via USB.

Table 1. Canopy characterization and volume rates following *Dosaviña®*.

Vigour	Canopy height (m)	Canopy width (m)	LWA (m^2_{canopy}/ha)[1]	Volume (l/ha)
Low	0.79	0.43	7,182	230
High	1.06	0.53	9,636	370

[1] LWA calculated for a row distance of 2.2 m.

Table 2. Specific working conditions for each vigour zone.

Vigour	Volume (l/ha)	Nozzle type	No. nozzles	Pressure (kPa)	Fwd. speed (km/h)
Low	230	ATR Brown + ATR Yellow	2 + 6	5.8	6.8
High	370	ATR Brown + ATR Yellow	2 + 6	15.6	6.8

Comparison of prescription vs actual application map
In order to evaluate the precision of the system, the prescription and actual maps were compared using QGIS software (QGIS Development Team, 2016). For the comparison process, a random net of *c.a.* 56,000 points was developed. For every value of expected value '*r*', 11 intervals of tolerance were assigned (each representing an increase of 5% compared with the previous one). The defined intervals ranged from zero to 50% deviation. Each point was compared and quantified for its coincidence between obtained value (actual) '*P*' and expected value '*r*'. In addition, a determination was made as to whether the *P*-value was inside the calculated range [*r*-i, *r*+i]. Once all the points were compared, the percentage of coincidence was also calculated. Finally, in order to visualize the level of accuracy of the actual spray application map, a specific interpolation process based on the inverse distance weighted (IDW) method was applied.

Results and discussion

Developed maps
This subsection presents and discusses the maps generated during the process. The sequence of the maps obtained was as follows: (1) NDVI map; (2) canopy vigour map; (3) prescription map; and (4) actual application map (Figure 2).

NDVI map
The spectral images obtained from the multispectral camera mounted on the UAV was used to generate the NDVI map (Figure 2A). This map shows the condition of the vegetation.

Canopy vigour map
Once the NDVI map was developed, it was treated in order to obtain two different zones in the vineyard. The zones were plotted on the map (Figure 2B). According to the values obtained, the total area of the field (2.8 ha) was distributed as follows: 50.9% (1.42 ha) for low canopy vigour and 49.1% (1.37 ha) for high canopy vigour.

Prescription map
Taking the canopy vigour map as the starting point, 15 complete manual characterizations of the canopy were made in each zone. Table 1 shows the results obtained per zone. These measures were used to determine the optimal volume rate using the DSS DOSAVIÑA® (UMA-UPC, 2018). In this case, the corresponding values were 230 l/ha for low canopy vigour, and 370 l/ha for high canopy vigour. From that, the intended prescription map was generated (Figure 2C).

Actual variable application map
Once the prescription map was loaded into the touch screen installed on the sprayer, the spray application process started. During the process, the system recorded information concerning the sprayer position in the field, the applied volume rate, and the adjusted working pressure, at a frequency of 1 Hz. Following further processing of the saved data, the actual application map was

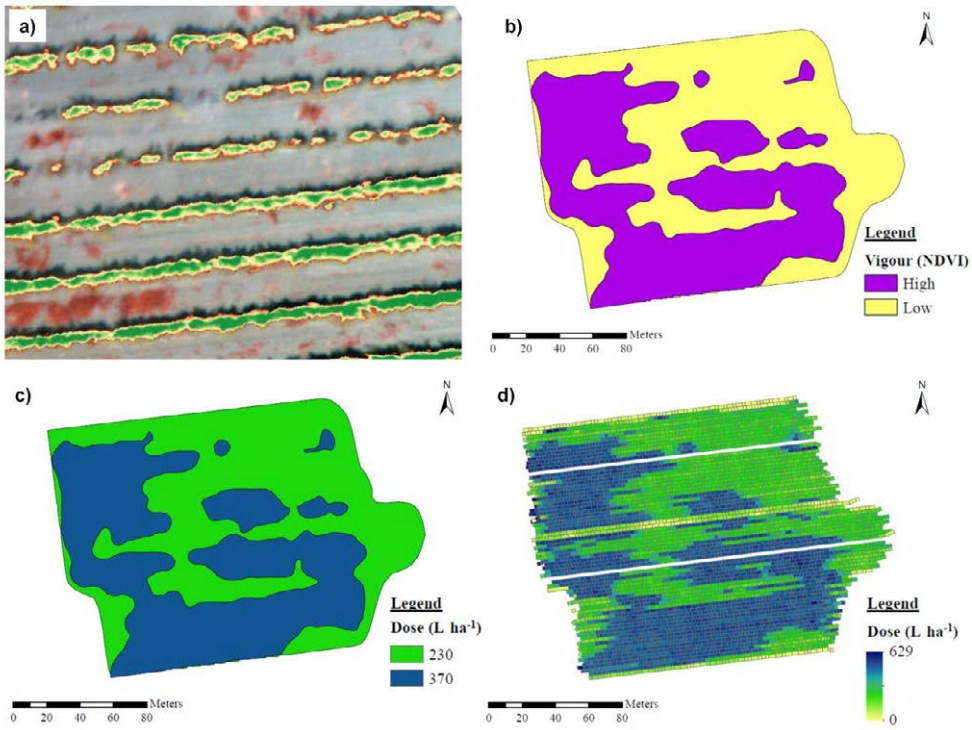


Figure 2. (A) raw NDVI map; (B) canopy vigour map; (C) prescription map; (D) actual variable application map.

generated (Figure 2D). A detailed analysis of this map facilitates explanation of certain characteristics. The white lines observed in the actual application map (Figure 2D) correspond to two row lines that were not sprayed due to driver failures.

Quantification of the accuracy of the system

In order to quantify the correspondence between prescription and actual map, a range of eleven different thresholds was established, from 0 to 50% tolerance. The most restrictive threshold (0%) measured the percentage of points (56,000) where there was no difference between the intended and actual application rate. In this case, only 3.5% of points were classified as successful points. At the opposite extreme, the highest tolerance (50%) quantified the percentage of points where variations of ±50% of applied volume was detected. In this case, 92.1% of points were classified as successful points. Assuming that from a practical point of view a tolerance of ±10% could be allowed, the system was able to classify 75% of the points as successful points. Figure 3 shows the accuracy distributed around the field, classified according to the established threshold level. The dark zones on the maps indicate the areas where the accuracy of the system exceeded the established thresholds. The main percentage of dark zones corresponds to transition zones, where the sprayer was ordered to modify the working pressure. As the results show, in terms of the comparison between the current and the prescription map, the variable rate application method obtained good accuracy.

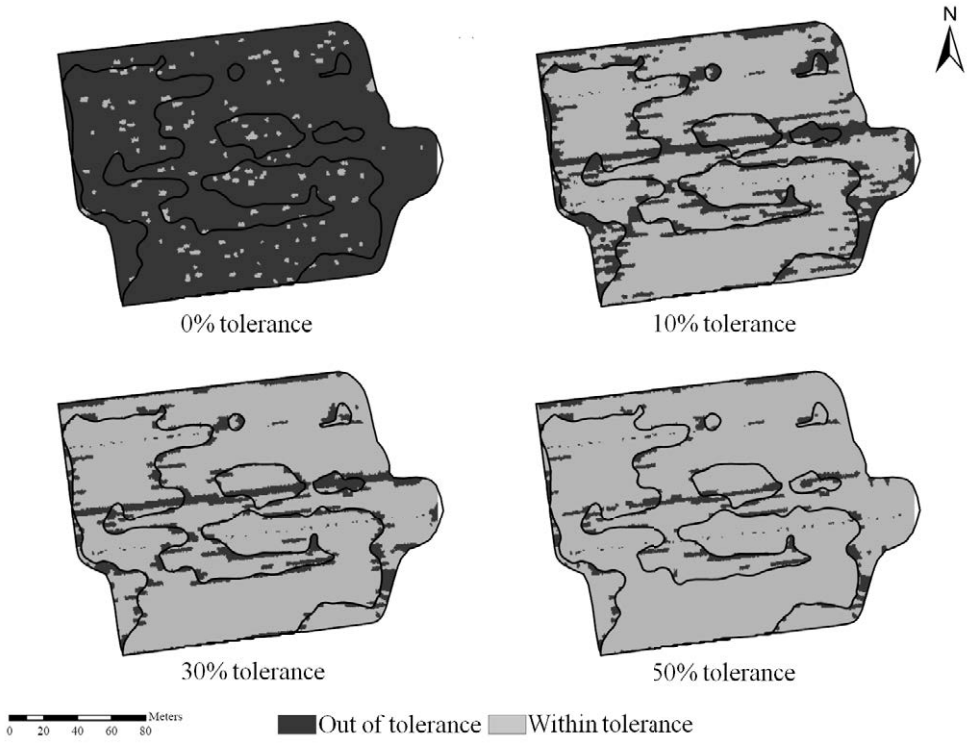


Figure 3. Spatial distribution of accuracy (intended vs actual application).

Quantification of savings

The variable application process based on maps was compared with a conventional spray application process based on a constant volume of 350 l/ha, the normal volume used by farmers in the area. For those two cases, the amount of water, the total application time assuming 45 minutes for the filling of the machine and the hypothetical amount of active material, assuming a copper concentration of 0.4% (400 g/hl) were calculated (Table 3). The results show a positive effect of the use of the variable rate application. Despite not obtaining a saving in the time used, the amount of liquid consumed and the amount of active ingredient were reduced by 17.9% using the variable rate application.

Table 3. Quantification of potential savings of the site-specific sprayer for conventional application and variable rate application (VRA).

	Volume (l/ha)	Total volume (l)	Volume savings (%)	No. of filling tanks	Time (h)	Time savings (%)	a.i. (kg)	a.i. savings (%)
Conventional	350	950	0.0	1	2.8	0.0	3.8	0.0
VRA	230/370	780	17.9	1	2.8	0.0	3.12	17.9

Conclusions

In this research, a prototype of a variable rate application sprayer based on prescription maps was tested. According to the results, the major findings were:
The system is able to read a prescription map and modify the working parameters (working pressure) depending on the position of the sprayer in the field. However, it is necessary to mention that the modification of working pressure could affect physical properties of delivered spray and therefore drift risk and spray coverage.
Good accuracy was obtained with the system (demonstrated by comparing the intended and actual application maps), with assumed tolerances of around 10% deviation.
The combination of vigour maps with Dosaviña® (DSS) to determine the optimal volume rate can reduce by around 20% the amount of water wasted in a spray application process and the amount of pesticide, as long as the concentration of the product is maintained.

References

Balsari, P., Doruchowski, G., Marucco, P., Tamagnone, M., Van de Zande, J., and Wenneker, M., 2008. A system for adjusting the spray application to the target characteristics. Agricultural Engineering International: CIGR Journal 10 1-11

Coombe, B.G., 1995. Growth stages of the grapevine: Adoption of a system for identifying grapevine growth stages. Australian Journal of Grape and Wine Research 1, 104-110

Gil, E., Gallart, M., Llorens, J., Llop, J., Bayer, T. and Carvalho, C., 2014. Spray adjustments based on LWA concept in vineyard. Relationship between canopy and coverage for different application settings. International Advances in Pesticide Applications. Aspects of Applied Biology 122 25-32

Gil, E., Llorens, J., Landers, A., Llop, J. and Giralt, L., 2011. Field validation of DOSAVIÑA, a decision support system to determine the optimal volume rate for pesticide application in vineyards. European Journal of Agronomy, 35 (1), 33-46; doi:10.1016/j.eja.2011.03.005

Meier, U., 1997. BBCH-Monograph. Growth stages of plants – Entwicklungsstadien von Pflanzen – Estadios de las plantas – Développement des Plantes. Berlin and Wien, Germany and Austria: Blackwell Wissenschaftsverlag. 622 p

Michaud, M., Watts, K.C., Percival, D.C., and Wilkie, K.I., 2008. Precision pesticide delivery based on aerial spectral imaging. Canadian Journal of Biosystems Engineering. 50: 2.9-2.15

Miranda-Fuentes, A., Llorens, J., Rodriguez-Lizana, A., Cuenca, A., Gil, E., Blanco-Roldán, G.L., *et al.,* 2016. Assessing the optimal liquid volume to be sprayed on isolated olive trees according to their canopy volumes. Science of the Total Environment, 568 (2016), 269-305

QGIS Development Team, 2016. QGIS Geographic Information System. Open Source Geospatial Foundation. URL http://qgis.osgeo.org

Rosell, J. R., and Sanz, R., 2012. A review of methods and applications of the geometric characterization of tree crops in agricultural activities. Computers and Electronics in Agriculture, 81, 124-141

Rouse, J. W., Haas, R. H., Schell, J. A. and Deering, D. W. 1973. Monitoring vegetation systems in the Great Plains with ERTS, Third ERTS Symposium, NASA SP-351 I, 309-317

Salcedo, R., Garcerá, C., Granell, R., Moltó, E. and Chueca, P., 2015. Description of the airflow produced by an air-assisted sprayer during pesticide applications to citrus. Spanish Journal of Agricultural Research, 13 (2), 2-8

UMA-UPC. 2018. DOSAVIÑA®. Decision Support System for determining the optimal volume rate in vineyard. URL: https://dosavina.upc.edu (last accessed August 2018)

Vogel, J.W., Wolf, R.E. and Dille, A., 2005. Evaluation of a Variable Rate Application System for Site-Specific Weed Management. Paper No. 051120, St Joseph, MI, USA: ASAE

Walklate, P. and Cross, J., 2012. An examination of Leaf-Wall-Area dose expression. Crop Protection, 35, 132-134

Zaman, Q.U. and Salyani, M., 2004. Effects of foliage density and ground speed on ultrasonic measurement of citrus tree volume. Applied Engineering in Agriculture. 20: 173-178

Investigation on LiDAR-based indicators for predicting agrochemical deposition within a vine field

A. Cheraiet[1,2*], M. Carra[1], A. Lienard[2], S. Codis[2], A. Vergès[2], X. Delpuech[2] and O. Naud[1]
[1]*ITAP, Univ. Montpellier, Irstea, 34196 Montpellier, France; anice.cheraiet@vignevin.com*
[2]*IFV, 34196 Montpellier, France*

Abstract

One strategy to reduce usage of crop protection agrochemicals in viticulture is to adjust the spray rate according to local vegetation characteristics. Acquisitions of mobile 2D LiDAR data and deposition rate measurements with a face-to-face sprayer were made on plant blocks in 5 fields and at 4 dates in a Mediterranean vineyard. Two indicators were calculated: the leaf wall area by points and the tree area index. Univariate deposition prediction models using either indicator were estimated from data acquired in 2 fields, and evaluated with regards to data from 3 other fields. In this experiment with a 3 m spatial resolution, the model based on LWApts was found a little more accurate.

Keywords: measurement, ground-based 2D LiDAR, temporal and spatial variability, crop structure, tree area index, leaf wall area, spraying deposits

Introduction

The current regulatory context, as well as the very high societal demand for a reduction in the use of pesticides, leads to reconsideration of the entire plant protection process, particularly in viticulture. Among the questions asked is how to express the registered doses of plant protection products (PPP). Indeed, the system for expressing doses of phytosanitary products currently in force in France in viticulture is based on a fixed dose, defined per hectare of ground surface area. In this context, a system for expressing doses of plant protection products that would explicitly take into account the evolution of the structure of the plant to be protected would be an important step toward precision agriculture. In this perspective, the development of precision spraying practices and technologies that take into account the characteristics of the plant is an area that has been identified (Berk *et al.,* 2016). In recent years, interest in laser sensors (2D LiDAR) in viticulture has increased considerably. Experiments in vineyards have been reported by Rosell-Polo *et al.* (2009) and Bastianelli *et al.* (2017, 2018), among others. Vegetative indicators can then be calculated from the LiDAR point clouds, in order to provide a simpler expression of the complex structure of vegetation. Subsequently, these vegetative indicators can be used for proposing and fitting PPP deposit prediction models and thus calculating PPP doses that reach the vegetation to protect. The leaf wall area by points (LWApts) proposed by Bastianelli *et al.* (2017) is one such indicator. It differs from the classical and manually acquired leaf wall area (LWAm) in the sense that surfaces from the leaf wall that that have no leaf are discarded. LWApts is a geometric representation of vegetation and can be considered a high resolution variant of LWAm. One of the reasons to study this indicator is that the less precise yet analogous LWAm will be used for the homologation tests of the new PPP for vines in Europe from 2020. One other indicator is tree area index (TAI), which takes a stochastic approach and deviates from a geometric interpretation (Walklate *et al.,* 2002). It is recognized as a robust indicator with good deposit prediction quality for an axial fan sprayer on a 50 m spatial scale in orchards. An adaptation of the TAI calculation method for vineyards has been suggested by Arnó *et al.* (2012). The main objective of this paper is to assess the ability to predict the normalized average amounts of spray deposition in the vegetation canopy at a fine spatial dimension (3 m) from the two 2D LiDAR based indicators TAI and LWApts, considering a face to face spraying technology. Face to

DOI 10.3920/978-90-8686-888-9_18, © Wageningen Academic Publishers 2019

face spraying technology relies on arms disposed on each side of the canopy, each arm bearing a set of nozzles aligned in a vertical plane. To do this, two univariate empirical deposition prediction models using the TAI and LWApts indicators were calibrated using experimental data acquired on two plots (each having a different grape variety) in 2017 at a spatial resolution scale of 3 m. These two models were then evaluated in terms of robustness and predictive quality on 3 other plots (with other grape varieties) of the same vineyard.

Materials and methods

Site

A vineyard with fields of different varieties with contrasting vigour was chosen for experimentation in 2017. It is located in Montpellier (Hérault, France, Lat. 43°31'55.9'N, Long. 3°51'50.3'E) and is characteristic of a vineyard in the south of France, both in terms of the grape varieties that constitute it and the way in which its vines are grown. Row-spacing was 2.5 m. Spray deposition measurements and 2D LiDAR characterization of vegetation were performed on five fields with five different varieties (Aranel, Marselan, Syrah, Caladoc and Petit Verdot) and at four dates (termed T1, T2, T3 and T4). These dates correspond to following growth stages: separate bunches (T1: 05/05/2017), flowering (T2: 19/05/2017), green pea stage (T3: 13/06/2017) and bunch closure (T4: 24/07/2017). Spatial monitoring of vegetation characteristics was carried out on a subsample of the fields. For each date and each field, a 15 m long plot, composed of 16 vine trees, was chosen. The 15 m long plots were different from one date to the other to ensure that the spray measurement on the collectors were not contaminated by previous tests. However, places of similar vigour in the field were chosen so that the growth dynamics were consistently represented. For building the model, the choice was made to use the two fields that were most different in terms of vegetation cover structure in the study area. These two fields (Aranel and Petit Verdot) were selected on the basis of the vegetative indicator LiDAR LWApts (Bastianelli *et al.,* 2017).

Measurements of average standardized spray deposition

In each of the 15 m long plots, 2 series of 3 consecutive homogeneous vines were sampled for deposition measurement. On each vine stock, 0.001 m^2 PVC collectors were positioned on the leaves inside the canopy according to a profile perpendicular to the row according to a cell grid 0.4 m high and 0.1 m wide. Tartrazine, a food coloring agent (E102), was mixed with water and sprayed with a face-to-face Precijet sprayer (Tecnoma brand, Epernay, France) equipped with TXA800067VK nozzles (Teejet brand, Wheaton, USA). Mean deposit on collectors was assessed and normalized by total collector area and dose rate per hectare as in Bastianelli *et al.* (2018). Spray deposition rates were expressed in nanogram. decimeter square of leaf per 1 g sprayed per hectare (ng/dm^2 per 1 g/ha). For each spraying date, sprayer parameters (number and direction of diffuser and nozzles) were adapted to the vegetation following good agricultural practice.

LiDAR 2D data collection

A 2D LiDAR sensor LMS 100 (Sick brand, Waldkirch, Germany) coupled with a RTK GNSS (Teria GSM, Vitry-sur-Seine, France) and an Effibox processing unit (Effidence brand, Romagnat, France) were mounted on a quad-bike. The quad's forward speed was 5 km/h. The LiDAR 2D data were acquired at a frequency of 50 Hz, with an angular range of 270°, at an angular resolution of 0.5°. A full scan with 541 values (distances from LiDAR to vegetation) was thus obtained every 0.027 m along drive direction. Once aggregated, the LiDAR scans form 3D point clouds of the studied vine rows. These 3D point clouds were used to describe the structure and architecture of the vegetation cover on each 15 m row sections. LiDAR measurements were taken on each section to calculate the TAI indicators (Arnó *et al.,* 2014) and LWApts (Bastianelli *et al.,* 2017) every 3 m forward, yielding 5 values for each 15 m plot.

LWApts indicator

The LWApts indicator differs from the PLWA expressed by Del-Moral-Martinez *et al.* (2015) because its calculation takes into account the entire width of the row and was designed not to require high accuracy for the positioning sensor. The LWApts (expressed in m^2/ha) is obtained by adding the elementary surfaces $Sveget_i$ of the points in the scan i in the rows on each side of the LiDAR sensor and dividing this sum by the total floor area occupied by the rows over the considered length l (l=3 m in this paper).

$$LWA_{pts} = \frac{\sum_{i=1}^{N} Sveget_i \times 2 \times 10000}{l \times W} \quad (1)$$

With i ϵ {1, 2, ..., ..., N} and N the number of scans acquired by the LiDAR;
$Sveget_i$: the elementary vegetation area calculated for the scan i (in m^2) by assigning a rectangular area, surface $Sveget_i^j$ associated with each point j using the following equation:

$$Sveget_i^j = \frac{(d_i + d_{i+1}) \times d_x}{2} \quad (2)$$

($d_i + d_{i+1}$): the distance between the current point j and the previous one, j-1;
l: the length of the vine row (in m) on which the measurements were acquired;
W: the distance between rows (in m);
10,000 m^2 per ha;
2: considered both sides of the row to be treated;
d_x: the distance between two consecutive Lidar scans.

TAI indicator

The process of obtaining TAI is explained in detail in Walklate *et al.* (2002). It is based on the probability of interception of light at different depths along the conical light beam of the LiDAR. The method used to calculate TAI in this article is similar to the method explained in detail in Arnó *et al.* (2012) and was coded by Bastianelli *et al.* (2017), who coded the LWApts as well. The TAI obtained from Equation 3 is formulated as the ratio between the area detected by the LiDAR and the ground surface. A scanning length (Δx) of 3 m was chosen in this study based on the results obtained by Arnó *et al.* (2016). Thus, each TAI value was calculated with 666 laser beams emitted for each angle of the LiDAR (from 0° to 270°), with a forward speed of 5 km/h. Thus for each section of length, Δx was estimated from the laser beam interception rate of each grid cell using the following formula:

$$TAI = -\frac{\Delta\theta}{W} \sum_{k=1}^{K} \sum_{j=1}^{J_k} r_j \cdot \delta_{k,j} \cdot \ln\left(1 - \frac{\Delta n_{k,j}}{n_{k,j}}\right) \quad (3)$$

With W: the distance between rows (in m);
$\Delta n_{k,j}$: the number of laser beams intercepted in the polar cell (ρ_j, θ_k);
$n_{k,j}$: the number of laser beams that reach and pass through the polar cell;
r_j: the radial distance between the polar cell and the sensor position (in m);
$\delta_{k,j}$: a function that detects the presence or absence of foliage in each cell. It takes the value of 1 if ($\Delta n_{k,j} / n_{k,j}$) >0.01 and the value of 0 if ($\Delta n_{k,j} / n_{k,j}$) <0.01 otherwise;
Δθ: angular increment (rad).
All interception points in the canopy were projected in different vertical planes parallel to the vine rows. The overall projected cross-sectional area was studied in cylindrical co-ordinates (ρ, θ, x) and projected into a polar grid. The size of the grid cells is defined by Δρ, Δθ and Δx on the 3 axes of the cylindrical co-ordinates. The dimensions of the selected cells were with an angular increment Δθ set at 3°, a radial increment Δρ set at 100 mm. Each polar cell of the grid is defined by its radial

distance, r_j with jε {0.1, ..., J}, and its angle, θ_k with kε {0.1, ..., K}. J is the number of intervals Δρ between 0 and W. K is the number of intervals Δθ between 0° and 270°. For each polar cell (j, k), the number of laser beams entering the polar cell, $n_{k,j}$, and the number of intercepted beams, Δn_{kj}, inside the polar cell, can then be calculated.
The TAI indicator design is based on the assumption that the probability of laser beam transmission in the vegetation can be approximated by the Poisson probability law when cell depth (Δρ) is small enough and the spatial distribution of leaves in the vegetation canopy is considered random.

Calibration and validation of local model of deposition using indices

Calibration
Empirical prediction models were calculated from experimental data collected on the Aranel and Petit Verdot plots.
The empirical models for predicting the deposition of vegetative indicators that have been defined are univariate. Because of collinearity between LWApts and TAI, as observed from the data, bivariate models using these 2 indicators would not be appropriate. The equations given in Table 1 were obtained from data by regression using a power function. The power function provided better robustness and prediction quality than logarithmic and simple linear regressions (data not shown).

Validation
The data acquired on the other 3 fields of the vineyard (Caladoc, Syrah and Marselan) were used to evaluate the local empirical models of deposit-TAI and deposit-LWApts prediction. To do this, a comparison between observed and simulated deposition was performed for each of the two prediction models using a 1:1 linear regression.

Results and discussions

Description of data

Deposition
Spray deposition measurements were used to assess the amount of PPP deposited on vegetation by a given sprayer for different areas of vegetation. Figure 1 shows the empirical density curves of the average deposition values per vine stock (3 consecutive vine trees) observed at the vineyard level for a given date. The trend is towards a decrease in the average deposit values as a function of time and consequently of vine growth. Due to the different management methods and vine architecture

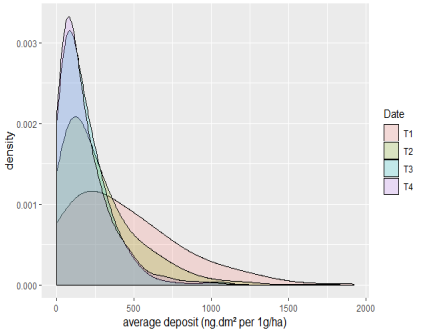


Figure 1. Empirical density curves of average deposition values as a function of the date (T1, T2, T3, T4) of spraying.

from one field to another, it can be observed that the deposits are very variable from one plot to another at a given date. On average, this inter-field variability is between 22 and 25% by date (not shown in Figure 1). High values of average deposits corresponded to less vigorous vines, which resulted in lower values of vegetative indicators LiDAR TAI and LWApts.

LiDAR

The dynamics of the LWApts and its variations from the average within the 5 plots are presented in Figure 2A. The LWApts increases almost linearly from bud break (T1) to green pea stage (T3), the date of the first trimming, to stagnate at around 8,500 m^2/ha. Overall, the earlier varieties (Aranel, Marselan and Caladoc) had the highest LWApts values up to trimming (T3), while the later developing Petit Verdot plot did not catch up and had lower LWApts values over the entire growing season. Between plots, LWApts differences are around 1,500 m^2/ha for a given date. This showed that there were real differences between plots during the growing season. Nevertheless, the plateau at the end of the (T4) growing season meant that the LWApts was saturated, which made it less informative on the evolution of vegetation at the end of the season, unlike the leaf area index (LAI), which continues to increase despite mechanical trimming (Johnson and Pierce, 2004). It was found that regardless of the date of acquisition of the vegetation cover measurements (T1 to T4), the Aranel and Petit Verdot fields were the most different in terms of vegetation cover structure. On average over the entire growing season, the Aranel fields had a value of LWApts (in m^2/ha) 22% higher than that of the Petit Verdot fields (Figure 2A). The wide range of vineyard management practices (in France as well as in Europe) means that there is a great diversity of vegetation structures. The LWApts depends on the row-spacing, the height of the vegetation and its porosity and not on its thickness. Thus, at different thicknesses, the same value of LWApts can be obtained. Using the leaf hedge area measured manually (LWA.manual) or by a LiDAR sensor (LWApts) as an indicator of the dose to be applied has a potential risk of over- or under-dosing depending on the thickness of the foliage to be protected (Rüegg *et al.*, 2001). This has potential consequences on the biological effectiveness of PPP.

The evolution of TAI on the 5 fields of the estate is presented in Figure 2B. It should be noted that, unlike the LWApts, the TAI indicator did not seem to saturate at T4. It was also observed that, regardless of the plot, the further into the season, the greater was the variability of the TAI indicator. This can be explained by the increasing density of the vegetation. With denser foliage, the LiDAR laser beams have a lower probability of penetrating inside the foliage and the majority of the signal is directly returned to the sensor. There is therefore a masking effect of vegetation, reducing the

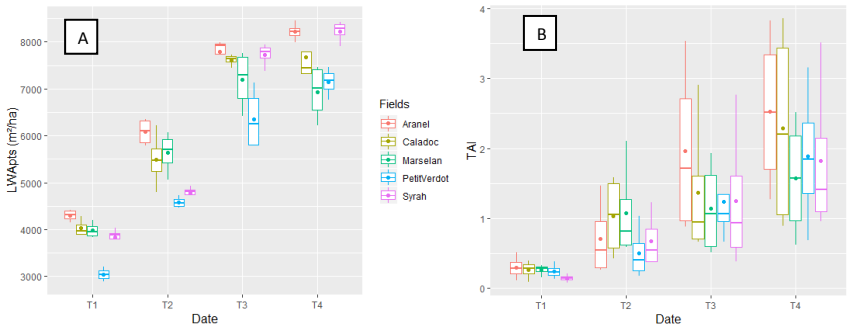


Figure 2. Evolution of the LiDAR LWApts (A) and TAI (B) indicators by field over the entire growing season (T1 to T4). A mechanical trimming operation of the vine is carried out in T3 except on the Syrah plot.

quantity of information perceived by the LiDAR, which degrades the quality of the LiDAR indicators measured on larger canopies (Berk *et al.*, 2016).

Calibration results

Table 1 shows that at a resolution scale of 3 m (set of 3 vine trees), the quality of the average deposition forecast by the LiDAR indicators TAI and LWApts for a side by side sprayer is very good. The R^2 of these relations is of 0.8 and 0.9 for the TAI and LWApts models respectively. This is consistent with the conclusions of Bastianelli *et al.* (2017). However, this previous study established empirical prediction models at a larger resolution scale (15 m) than in the present paper (3 m). The shape of the regressions (power type) obtained with the TAI and LWApts models is similar to that obtained in Siegfried *et al.* (2007). This similarity reinforces the good correlation between LWApts and TAI mentioned in Arnó *et al.* (2012, 2014), although the calculation of TAI in this study differs slightly. The results show that the predictive quality of these two empirical deposition prediction models is maintained at a finer spatial dimension (3 m). It seems that 3 m may be a relevant spatial scale to adjust doses according to vegetation architecture, for sprayers that would have real-time control features.

Validation results

Figure 3A and 3B show that the prediction quality of the two empirical models applied to the other 3 plots of the vineyard estate was acceptable with a R^2 of 0.65 and 0.88 for the LWApts (Figure 3A) and TAI (Figure 3B) models respectively. In addition, it was considered that a prediction was better if it reduced uncertainty with the observed mean deposition.

The RMSEp criterion was used to measure this uncertainty. For ease of reading, only the average RMSEp of the 3 plots over the entire growing season for the two prediction models were represented

Table 1. Equations, correlation coefficient (R^2) and root mean square error (RMSE) of empirical average deposit rates prediction models.

Prediction models	Model equation	R^2	RMSE (ng/dm^2 per 1 g/ha)
TAI-model	$y = 25.59x^{-0.421}$	0.8	120
LWApts-model	$y = 4E+08x^{-1.622}$	0.9	65

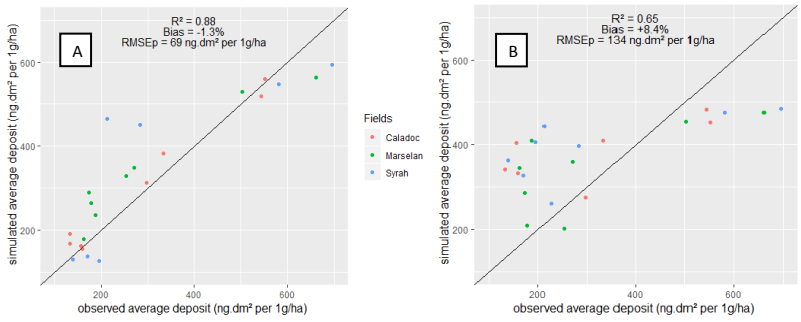


Figure 3. Comparison of the overall performance of the local empirical prediction models LWApts (A) and TAI (B) against the mean deposition rate measurements. Model quality is assessed using the determination coefficient (R^2), bias (%) and mean square error (RMSEp) on the combined Marselan, Caladoc and Syrah plots throughout the growing season (T1, T2, T3, T4).

in Figure 3A and 3B. The result was that the higher the quality of the prediction (R^2), the more accurate the prediction model, as confirmed by the entire literature on the subject. There is an RMSEp value of 69 ng/dm^2 per 1 g/ha for the LWApts model and 134 ng/dm^2 per 1 g/ha for the TAI model. This is visualised in Figure 3. The LWApts prediction model offered a better prediction quality than the TAI model. This was in line with the results of Bastianelli *et al.* (2018) which demonstrated a better quality of deposition prediction for a face-to-face sprayer with the LWApts model by comparing empirical prediction models calculated from experimental data collected in 2016 in a set of fields that were not used for the construction of the model and for a different year. It should be noted, however, that the bias value was negative for the LWApts-based prediction model (-1.3%) but positive for the TAI-based prediction model (+8.4%). This indicates that the LWApts prediction model under-estimates and the TAI prediction model over-estimates the average surface deposition over the entire vegetation period. This might be explained by local variations of interception points distribution along the 3 m traverse on which TAI was calculated. A modelling approach with 'worst case' risk management should encourage under-estimation rather than over-estimation of deposits. Over these 3 m, the TAI overall takes into account the vigorous parts and the weak parts of the vegetation cover. The less vigorous parts contribute to lower TAI values and thus to higher values for the expected deposition rates. Yet, the deposition sampling was carried out on parts of the vines that are rather vigorous so that collectors can be placed on the whole extent of vegetation cover.

Conclusions

A lever to optimise the use of phytosanitary inputs in viticulture is to take into account the structural characteristics of the vegetation in order to compute an appropriate dose, reducing phytosanitary treatments. In this study, the strategy was to implement a model for predicting spray deposition on vegetation based on vegetative cover parameters. Thanks to its ability to scan a large amount of vegetation at a high-resolution, LiDAR 2D is a powerful tool for acquiring 3D vegetation point clouds and calculating indicators that have meaning down to a small scale. In this study, a scale of 3 m with dose modulation was investigated. The results obtained from data gathered on several fields of the same estate showed that the empirical model that allows prediction of the average spray deposition of a typical face-to-face sprayer with more precision and robustness is the prediction model based on the LWApts. New developments in sprayer equipment make it possible to consider using such prediction models to modulate doses on an intra-parcel scale, which would allow a potential reduction in the quantities of PPP to be applied. Further studies could examine if such a univariate model can be extrapolated from estate to estate, or be enhanced to a multivariate model.

Acknowledgements

Research partially funded under CASDAR ArchiTechDoseViti and EcoPhyto LIDARECOSE projects. We thank the colleagues of the UMT ECOTECH. Thanks to James Taylor for methodological advice in writing the article. Work supported by the French National Research Agency under the Investments for the Future Program ANR-16-CONV-0004.

References

Arnó, J., Escolà, A., & Rosell-Polo, J. R., 2016. Setting the optimal length to be scanned in rows of vines by using mobile terrestrial laser scanners. Precision Agriculture, 18 (2), 145-151. doi:10.1007/s11119-016-9451-z.

Arnó, J., Escolà, A., Masip, J., & Rosell-Polo, J. R., 2014. Influence of the scanned side of the row in terrestrial laser sensor applications in vineyards: practical consequences. Precision Agriculture, 16 (2), 119-128. doi:10.1007/s11119-014-9364-7.

Arnó, J., Escolà, A., Vallès, J. M., Llorens, J., Sanz, R., Masip, J. *et al.,* 2012. Leaf area index estimation in vineyards using a ground-based LiDAR scanner. Precision Agriculture, 14 (3), 290-306. doi:10.1007/s11119-012-9295-0.

Bastianelli, M., de Rudnicki, V., Codis, S., Naud, O., 2018. Assessing models from Lidar based vegetation indicators for predicting spraying deposit amounts in a set of vine estates in France, In: Proceedings International Advances in Pesticide Application Conference, Aspects of Applied Biology, 137:375-384, Brighton, UK.

Bastianelli, M., Rudnicki, V., Codis, S., Ribeyrolles, X., Naud, O., 2017. Two vegetation indicators from 2D ground Lidar scanner compared for predicting spraying deposits on grapevine, In: Proceedings EFITA-WCCA Conference, Montpellier, France.

Berk, P., Hocevar, M., Stajnko, D., & Belsak, A., 2016. Development of alternative plant protection product application techniques in orchards, based on measurement sensing systems: A review. Computers and Electronics in Agriculture, 124, 273-288. doi: 10.1016/j.compag.2016.04.018.

Del-Moral-Martínez, I., Arnó, J., Escolà, A., Sanz, R., Masip-Vilalta, J., Company-Messa, J., *et al.,* 2015. Georeferenced Scanning System to Estimate the Leaf Wall Area in Tree Crops. Sensors, 15(4), 8382-8405. doi:10.3390/s150408382.

Johnson, L. F., & Pierce, L. L., 2004. Indirect Measurement of Leaf Area Index in California North Coast Vineyards. HortScience, 39 (2), 236-238. doi:10.21273/hortsci.39.2.236.

Rosell Polo, J. R., Sanz, R., Llorens, J., Arnó, J., Escolà, A., Ribes-Dasi, M., *et al.,* 2009. A tractor-mounted scanning LIDAR for the non-destructive measurement of vegetative volume and surface area of tree-row plantations: A comparison with conventional destructive measurements. Biosystems Engineering, 102 (2), 128-134. doi: 10.1016/j.biosystemseng.2008.10.009.

Rüegg, J., Siegfried, W., Raisigl, U., Viret, O., Steffek, R., Reisenzein, H., *et al.,* 2001. Registration of plant protection products in EPPO countries: current status and possible approaches to harmonization. EPPO Bulletin, 31 (2), 143-152. doi:10.1111/j.1365-2338. 2001.tb00983. x.

Siegfried, W., Viret, O., Huber, B., & Wohlhauser, R., 2007. Dosage of plant protection products adapted to leaf area index in viticulture. Crop Protection, 26 (2), 73-82. doi: 10.1016/j.cropro.2006.04.002.

Walklate, P-J., Cross, J-V., Richardson, G-M., Murray, R-A., Baker, D-E., 2002. Comparison of different spray volume deposition models using LIDAR measurements of apple orchards. Biosystems Engineering, 82 (3), 253-267. doi:10.1006/bioe.2002.0082.

Missing plant detection and biomass estimation from 3D models generated from UAV in a vineyard

A. Matese[1*], P. Cinat[1], Y. Romboli[2], A. Berton[3] and S.F. Di Gennaro[1]
[1]*Institute of Biometeorology, National Research Council (CNR-IBIMET), Via G. Caproni 8, 50145 Florence, Italy; a.matese@ibimet.cnr.it*
[2]*Department of Agricultural, Food and Forestry Systems (GESAAF), University of Florence, Piazzale delle Cascine 24, 50144 Florence, Italy*
[3]*Institute of Clinical Physiology, National Research Council (CNR-IFC), Via Moruzzi 1, 56124 Pisa, Italy*

Abstract

In the last decade, there has been a progressive diffusion of UAVs (Unmanned Aerial Vehicles). A computational 3D model of a vineyard can be assembled by combining low-cost UAV, RGB cameras and automated 3D reconstruction algorithms. The aim of the study was to compare two different methods of image processing in order to provide best performance for biomass estimation and missing plant detection from a 3D model of the vineyard. Both methods showed good overall accuracy, with an RMSE lower than 0.5, of biomass estimation and missing plant detection but the second method, although more precise, needed more computational time.

Keywords: UAV, precision viticulture, 3D modelling, missing plants, biomass

Introduction

Vineyards are characterized by high heterogeneity due to structural factors mediated by topography and soil characteristics, and to non-structural factors, mediated by crop practices. Remote sensing technologies have been successfully used for vineyard monitoring and could be useful to characterize vineyard variability. Unmanned aerial vehicles (UAV) provide high flexibility of use, low operational costs and very high spatial resolution (Matese *et al.*, 2015). RGB sensors mounted on UAVs are capable of providing high-resolution images that can be processed to build digital surface models (DSMs), using three-dimensional (3D) reconstruction software based on stereo vision or structure from motion (SfM) algorithms (Padua *et al.*, 2017). Using these methods, a large set of applications can be undertaken such as biomass monitoring (Bendig *et al.*, 2014, 2015; Matese *et al.*, 2017), volume characterization (Ballesteros *et al.*, 2015) and early-season crop monitoring (Torres-Sánchez *et al.*, 2014; de Castro *et al.*, 2018). Many authors reported that the use of SfM from UAV-images may produce a 3D point cloud similar to one obtained acquiring data with a LiDAR. The RGB point cloud has a point density depending on the image ground resolution and overlap level, but with a consistently lower cost than a LiDAR one. These advantages have led to an increasing interest in this technology and in the last few years, several studies utilized dense point clouds from SfM in vineyards with different applications (Ballesteros *et al.*, 2015; Mathews and Jensen, 2013; Weiss and Baret, 2017).
Missing plant management in a vineyard is a critical issue that can be managed by new technologies. Most vineyards have a lower number of vines per hectare than the calculated number due to missing vines from different causes such as disease, winter injury or mechanical damage. As a consequence, farmers lose a significant percentage of potential vineyard production. The simplest approach to identifying missing plants would be to detect areas not covered by canopy along the row. Unfortunately, vertical aerial photography is often not able to identify the actual situation under the top part of the canopy, and in the case of absence of a plant, the neighbouring plants can extend their shoots and foliage up to occupy the adjacent free space. Using a surface approach, the estimation of height and area from UAV measurements do not denote tangible errors because the

DOI 10.3920/978-90-8686-888-9_19, © Wageningen Academic Publishers 2019

ground-based measurements have been derived by applying the conventional geometric equation that considers the trees as ellipsoid forms, which can produce inexact ground estimations. Torres-Sánchez *et al.* (2015) reported that 3-D products derived in their study reconstructed the irregular shape of the tree crown, which hypothetically allows better estimation of tree volume than those derived from ground measurements. A step forward could be made using a methodology typical for forestry application (Wallace *et al.*, 2014), where the 3D dense cloud is reconstructed as an object shape from a set of unorganized points. Using these methods, not only the tree crown shape was estimated, but also the entire canopy. The aim of this study was to implement an unsupervised and integrated procedure of biomass estimation and missing plants detection, using both the 2D-surface and 3D-alphashape methods.

Materials and methods

Experimental site

The research was undertaken in 2018 in an experimental plot within a 3 ha vineyard (43.00°N, 11.26°E) located in Montalcino, Tuscany, central Italy, planted in 1973 with Sangiovese cv spur-pruned cordon trellis system. This vineyard was chosen to evaluate the methodological approach under extreme and different conditions. In fact, the vineyard presents irregular structure with different plant ages and spacing due to high number of replacements during the years (Figure 1). The dimension of the experimental plot was 50×150 m including 15 rows.

Remote sensing platform and data

Flight campaigns were performed using an open-source UAV platform consisting of a modified multi-rotor Mikrokopter (HiSystems GmbH, Moomerland, Germany). A universal camera mount equipped with three servomotors allowed accurate image acquisition through compensation of tilt and rolling effects. The RGB camera was a Sony Cyber-shot DSC-QX100 RGB camera (Sony Corporation, Tokyo, Japan) with a 20.2 megapixel CMOS Exmor R sensor and a Carl Zeiss Vario-Sonnar T lens. Flight campaigns were carried out on 13 September 2018 through 2 consecutive surveys to monitor both sides of the canopy with a 45° oblique camera with respect to the vertical at the ground and 45° with respect to the direction of advancement along the row. The flights were conducted at 20 m above ground level at midday, yielding a ground resolution of 0.005 m/pixel. The RGB camera was set at 0.2 s triggering time with automatic exposure.

Ground measurements

Ground-truth field data measurements for model validation consisted of measuring height and thickness of 3 vines within 5 polygons (15 sampling vines), while field data for missing plant validation were visually counted within the same polygons.



Figure 1. Experimental plot selected for data analysis.

Data processing

RGB images acquired by UAV were mosaicked using Agisoft Photoscan Professional Edition 1.3.3 (Agisoft LLC, St. Petersburg, Russia). A polygon mesh was computed from the dense 3D point cloud generated and the pixel values of each image were then projected onto the mesh to create an orthomosaic. When combined with GNSS positions, this process allowed the creation of a high-resolution orthophoto and a digital elevation model (DEM) of the experimental site.

2D approach – surface model method

Starting from the DEM originated by Agisoft, a uniform polygon grid was then uniformly generated to isolate, in each vine-row, ten plants. The vine spacing was 0.9×3.0 m and, then, each polygon grid was generated starting from the middle point between two rows with regular spacing of 3.0×9.0 m (Figure 2A). Moreover, vines and soil were separated with a thresholding approach. The '*imopath*' function of Matlab 2016 was applied to the DEM to mitigate the effect of the terrain slope. Then, the Otsu's thresholding technique was used to distinguish between soil and vines, generating a logical mask of the complete field. At this point, the geometric features of vines (vine height dh and mean thickness δ_m) in each polygon grid were extrapolated as follows. The DEM of a representative region is shown in Figure 3. White pixels identify soil and black pixels identify vine rows.

The sum of black pixels (logical indexing equal to 0) was computed for each column, defining the vine thickness function $\delta(x)$ (in pixels) along the region length. The region length corresponded to the vine extent along the row. This sum was then multiplied for the real spacing associated with each pixel dp and the cosine of the angle θ indicating the vine slope with respect to the horizontal axis of the original image. The function $\delta_p(x) \times dp \times \cos(\theta)$ represents the distribution of vine thickness with respect to the central axis of the vine row, represented on the bottom of Figure 3. Consequently, the zones where $\delta_p(x)=0$ identifies the missing plants. Computationally speaking, missing plants are identified considering all points with $\delta_p(x)<0.1$ m. The remaining set of points was considered as vine thickness. The mean thickness was computed as the mean value of $\delta_p(x)$ avoiding missing plants $\delta_m = <\delta_p(x)>$ considering only values with $\delta_p(x)\geq 0.1$. With the same approach, the mean altitude of vines and soil was computed for each row of the image, the mean value of the height of vines or soil. The vertical distribution of vine and soil altitude is shown on the right of Figure 3. The mean vine-elevation $h_m^{(v)}$ was considered as the maximum value of $h^{(v)}(y)$. The mean soil-height $h_m^{(s)}$

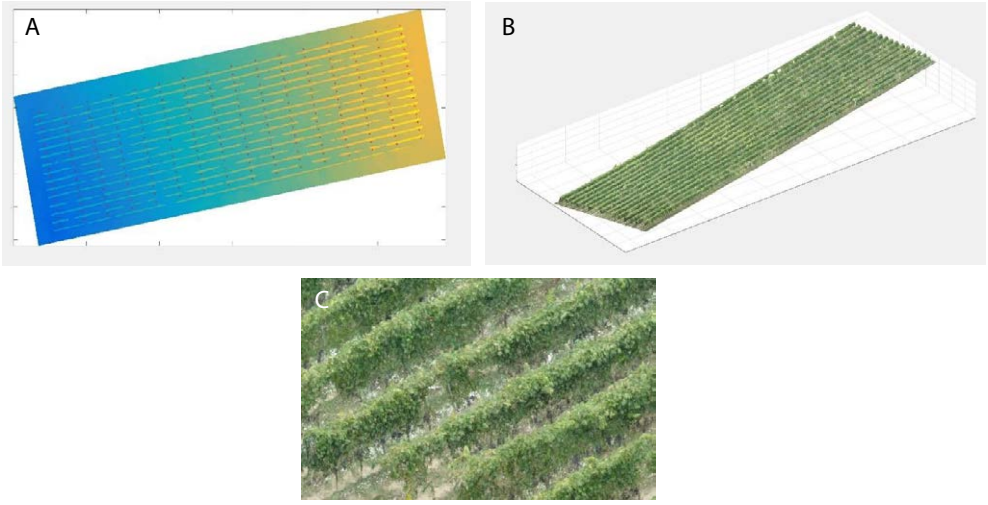


Figure 2. (A) DEM with polygon grid; (B) 3D dense cloud; (C) a detail of the 3D dense cloud.

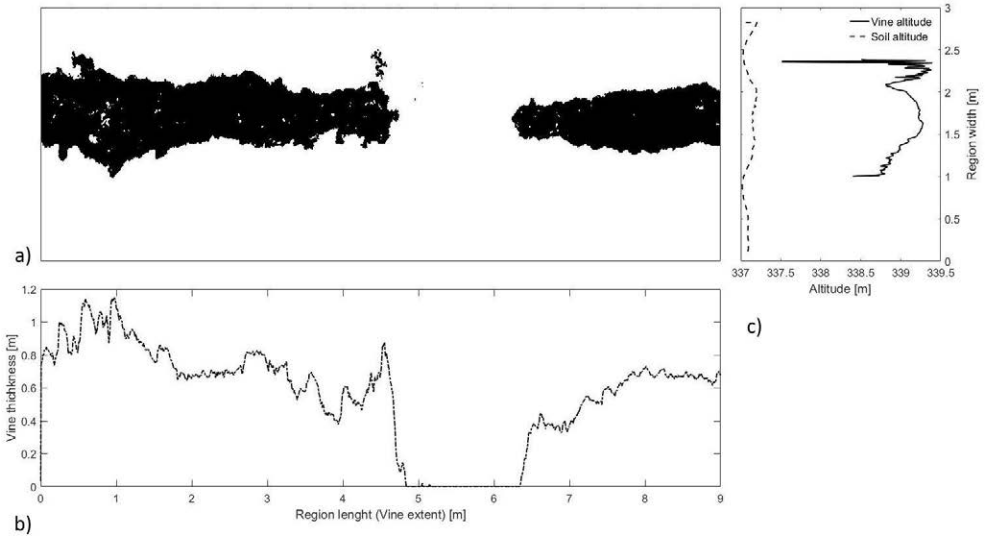


Figure 3. Digital elevation model of vines within a polygon grid and its geometric features. Vine length (B) and thickness (B) are extrapolated through the binarized image (A), mediating the values on the image columns. The soil and vine altitude are obtained mediating respective pixels on horizontal rows (C).

was extracted as the minimum value of $h^{(s)}(y)$, obtaining the vine height as $dh = h_m^{(v)} - h_m^{(s)}$. Then, 1 m was subtracted from dh considering the height of the cordon. At this point, it was possible to estimate the biomass associated with the complete region V_{DEM} and a mean plant V_p. The volume associated with the mean plant was: $V_p = dh \times dL \times \delta_m$ where dL=0.9 is the vine-length defined by the vine spacing of the vineyard. The volume associated with the DEM was $V_{DEM} = dh \times dA$, where dA is the area of the vines computed by multiplying the total number of black pixel N_{px} to the area associated with each pixel, i.e. $dA = N_{px} \times (dp \times \cos(\theta))^2$. The number of missing plants in the region was identified as: $N_{bio} = n_p - \frac{V_{DEM}}{V_p}$ where n_p is the number of plants in the region, equal to 10 according the vine spacing of the vineyard. In addition, missing plants can be estimated in another way. From Figure 3, it is also possible to estimate the extension of missing plants, since it is the extension L_v of the line with missing plant thickness $\delta_p(x)$ <0.1 m. The number of missing plants is then: $N_{ext} = \frac{L_v}{\ell_p}$ where ℓ_p is the length of a plant in the vineyard, equal to 0.9 m according to the vine spacing.

3D approach – alphashape

Starting from the 3D dense cloud generated by Agisoft, an alphashape (Edelsbrunner and Mucke, 1994) or volume that envelopes the set of 3D points must be obtained to estimate volumes.

Alphashape allows the reconstruction of an object's shape, namely α-shape, from a set of unorganized points. The parameter α is used to tune the 'tightness' of the shape around the points. For a very large value of α, the shape is equivalent to the convex hull. For a very small value of α, the α-shape forms holes and pockets with the shape clustering around the original points. Matlab '*alphashape*' function was used to create a bounding volume that envelops a set of 3-D points. Then, it is possible to manipulate the α-shape object to tighten or loosen the fit around the points to create a non-convex region and perform geometric queries (Ribeiro *et al.*, 2017).

Firstly, the canopy points of the 3D dense cloud were extracted for each polygon grid using a '*planeModel*' object to construct a parametric plane model (Figure 4). Then, the '*alphashape*' function

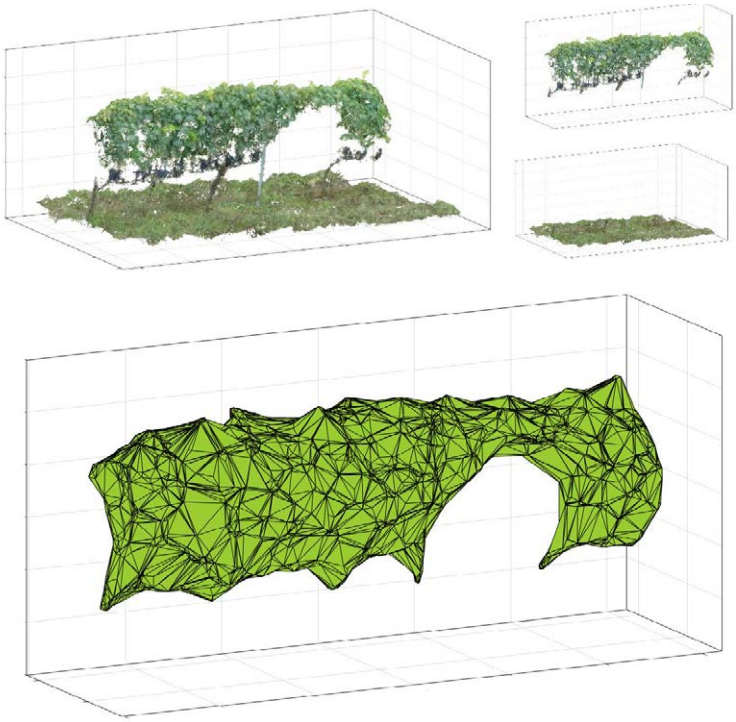


Figure 4. 3D dense cloud and α-shape processing of one polygon grid.

was applied for each polygon using an α value of 0.3. The function allowed both the calculation of the α-shapes and volume estimation. Finally, a volume map was obtained as the union of the volume maps of all sampled rows. To obtain the volume map of a row, the volume value of each section was projected onto the segment joining the starting and ending co-ordinates of the row.

Results and discussion

Model validation

Table 1 presents the results of the comparison between vine height (h_estim), thickness (t_estim) and missing plants (mp_estim) estimated by the proposed method and measured (h_meas, t_meas,

Table 1. Model validation using 3 vines per 5 polygons.[1]

Polygon	Row	h_meas	h_estim	t_meas	t_estim	mp_meas	mp_estim
1	1	1.93	1.91	0.55	0.51	5.00	5.00
2	5	2.03	1.95	0.60	0.68	0.00	0.00
3	10	1.93	1.94	0.63	0.65	0.00	0.00
4	14	2.07	2.05	0.65	0.67	0.00	0.00
5	15	2.10	1.89	0.69	0.62	4.00	3.00

[1] RMSE (h)=0.10; RMSE (t)=0.05; RMSE (mp)=0.45.

mp_meas) in the field during the flight campaign. In general, the model assessed with good accuracy the ground-truth measurements. These variables are the most used by farmers for in-field volume measurements, being non-destructive and easy to acquire. Therefore, these variables were chosen to validate the model.

Biomass estimation

The volume of the vines within each polygon grid was calculated using the 2D and 3D method and highlighted in Figure 5 and 6, respectively. The tendency of the volumes in the different polygons is similar using the two methods and the values range from about 1 m^3 with high missing plants presence to over 8 m^3 without missing plants and very large vine thickness. The 2D method tends to over-estimate the volume because it builds a polygon given by height and thickness, while the 3D can model the whole shape of the canopy providing smaller average but probably more consistent values, considering the canopy structure heterogeneity.

Missing plant detection

The number of missing plants within each polygon grid was calculated using the 2D method and 3D method and are highlighted in Figure 7 and 8 respectively. In general, a large number of missing plants were detected using the 3D method, probably because the 2D method was not able to identify the real situation under the top of the canopy and frequently neighbouring plants can extend their shoots and leaves to cover the free space of a missing plant. This vineyard had many missing plants: in fact, the results showed up to 8 missing plants in a polygon that should contain 10 plants. The 3D method provided higher performance in canopy segmentation, reconstructing not only the canopy

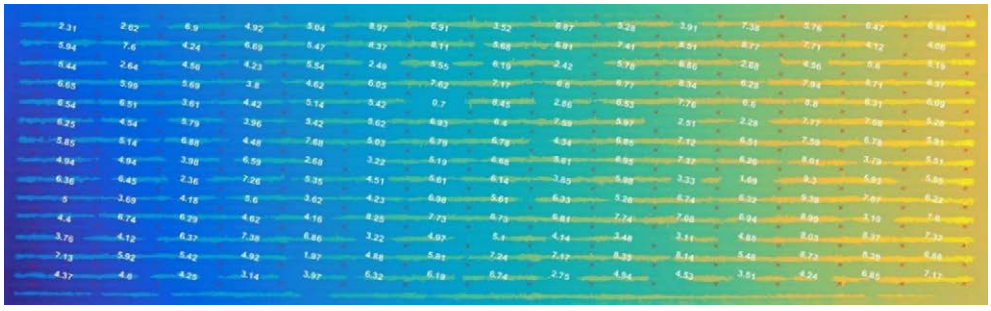


Figure 5. Volume (m^3) for each polygon grid obtained from 2D method.

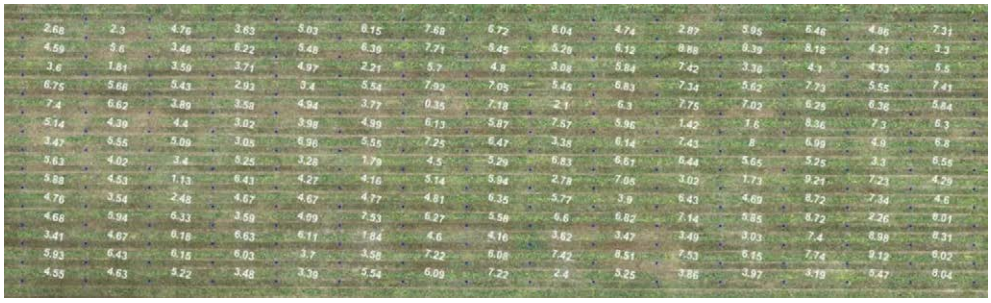


Figure 6. Volume (m^3) for each polygon grid obtained from 3D method.

shape, but also trunk and cordon of dead vines. Very high ground resolution and elevated overlap were necessary to reconstruct an accurate dense cloud able to provide such results. Considering the low flight height and forward speed needed to obtain enough dense cloud quality, the UAV battery autonomy was the main issue limiting the flight time and thus the surface covered in a survey. The large amount of data acquired to be processed was another critical aspect, and image processing requires computers to be equipped with a larger working memory to reconstruct dense clouds and perform image analysis, especially following 3D methodology.

Conclusions

This study confirmed the feasibility of rapid assessment of the biomass volume using different approaches based on the SfM algorithm applied to high resolution RGB images with large overlap and acquired by a UAV platform. A secondary task, represented by an accurate identification of missing plants within the rows, able to detect also in the case of single dead vines, where only the cordon is present, partially covered by adjacent ones. These methods provided thematic maps related to biomass and missing plants with the aim of supporting the farmer in canopy management in order to achieve the desired vine balance. Another potential application could be the optimization of the re-planting process, better quantifying the order of new vines from a nursery and allowing fast location of each re-planting site. This research described a rapid and objective tool for the farmer to promptly identify canopy management strategies and drive replanting decisions.

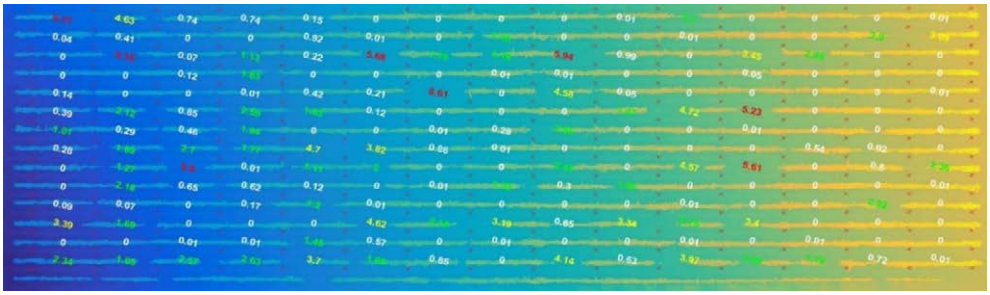


Figure 7. Missing plants (number) for each polygon grid from 2D method.

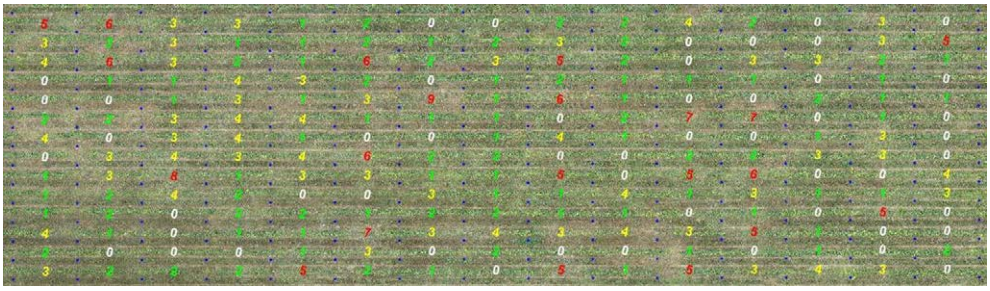


Figure 8. Missing plants (number) for each polygon grid from 3D method.

References

Ballesteros, R., J. F. Ortega, D. Hernández, and M. Á. Moreno, 2015. Characterization of Vitis vinifera L. Canopy Using Unmanned Aerial Vehicle-Based Remote Sensing and Photogrammetry Techniques. American Journal of Enology and Viticulture, 66, 120-129.

Bendig, J., A. Bolten, S. Bennertz, J. Broscheit, S. Eichfuss, and G. Bareth, 2014. Estimating Biomass of Barley Using Crop Surface Models (CSMs) Derived from UAV-Based RGB Imaging. Remote Sensing 6 (11): 10395-10412.

Bendig, J., K. Yu, H. Aasen, A. Bolten, S. Bennertz, J. Broscheit, *et al.,* 2015. Combining UAV-Based Plant Height from Crop Surface Models, Visible, and Near Infrared Vegetation Indices for Biomass Monitoring in Barley. International Journal of Applied Earth Observation and Geoinformation 39: 79-87.

De Castro, A.I.; Jiménez-Brenes, F.M.; Torres-Sánchez, J.; Peña, J.M.; Borra-Serrano, I.; López-Granados, F., 2018. 3-D Characterization of Vineyards Using a Novel UAV Imagery-Based OBIA Procedure for Precision Viticulture Applications. Remote Sensing, 10, 584.

Edelsbrunner, H. and Mücke, E.P., 1994. Three-dimensional alpha shapes. ACM Transactions on Graphics 13 (1), 43-72

Matese, A., Di Gennaro, S.F. and Berton A., 2017. Assessment of a canopy height model (CHM) in a vineyard using UAV-based multispectral imaging. International Journal of Remote Sensing, 38:8-10, 2150-2160.

Matese, A., Toscano, P., Di Gennaro, S.F., Genesio, L., Vaccari, F.P., Primicerio, J., *et al.,* 2015. Intercomparison of UAV, Aircraft and Satellite Remote Sensing Platforms for Precision Viticulture. Remote Sensing 7 2971-2990.

Mathews, A. J. and Jensen, J. L. R., 2013. Visualizing and quantifying vineyard canopy LAI using an unmanned aerial vehicle (UAV) collected high density structure from motion point cloud. Remote Sensing, 5 (5), 2164-2183

Pádua, L., Vanko, J., Hruška, J., Adão, T., Sousa, J.J., Peres, E., *et al.,* 2017. UAS, sensors, and data processing in agroforestry: A review towards practical applications. International Journal of Remote Sensing, 38, 2349-2391.

Ribeiro, A., Bengochea-Guevara, J., Conesa-Muñoz, J., Nuñez, N., Cantuña, K. and Andújar, D., 2017. 3D monitoring of woody crops using an unmanned ground vehicle. In J.A. Taylor, D. Cammarano, A. Prashar, A. Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences, 8(2), 210-215.

Torres-Sánchez, J., J. M. Peña, A. I. De Castro, and F. López-Granados, 2014. Multi-Temporal Mapping of the Vegetation Fraction in Early-Season Wheat Fields Using Images from UAV. Computers and Electronics in Agriculture 103: 104-113.

Torres-Sánchez, J., López-Granados, F., Serrano, N., Arquero, O., Peña, J.M., 2015. High-Throughput 3D Monitoring of Agricultural-Tree Plantations with Unmanned Aerial Vehicle (UAV) Technology. PLoS ONE 10(6): e0130479.

Wallace, L., R. Musk, and A. Lucieer, 2014a. An assessment of the repeatability of automatic forest inventory metrics derived from UAV-borne laser scanning data, IEEE Transactions on Geoscience and Remote Sensing, 52 (11):7160-7169

Weiss, M., and Baret, F., 2017. Using 3D point clouds derived from UAV RGB imagery to describe vineyard 3D macro-structure. Remote Sensing, 9 (2), 111.

Combining target sampling with route-optimization to optimise yield estimation in viticulture

B. Oger[1,2*], P. Vismara[2,3] and B. Tisseyre[1]
[1]*ITAP, Univ. Montpellier, Montpellier SupAgro, Irstea, France; baptiste.oger@supagro.fr*
[2]*MISTEA, Univ. Montpellier, Montpellier SupAgro, INRA, France*
[3]*LIRMM, Univ. Montpellier, Montpellier SupAgro, CNRS, France*

Abstract

This paper describes a new approach for yield sampling in viticulture. It combines approaches based on auxiliary information and path optimization to offer more consistent sampling strategies, integrating statistical approaches with computer methods. To achieve this, groups of potential sampling points, comparable according to their auxiliary data values are created. Then, an optimal path connecting several points, one from each group of potential sampling points and minimizing the route distance is created. This part is performed using constraint programming, a programming paradigm offering tools to deal efficiently with combinatorial problems. The paper presents the formalization of the problem, as well as the tests performed on real fields. Results show that combining target sampling and path optimization can reduce by 45% the average sampling circuit length compared to previous methods based on auxiliary data while being almost equivalent in yield prediction error.

Keywords: sampling optimization, yield estimation, model sampling, normalised difference vegetation index, constraint programming

Introduction

In order to optimize harvest organization, prepare the winemaking process and establish commercial strategies, the wine industry needs to know the yield of each vine field. Ideally, yield has to be estimated a few days before harvest with a relative error of less than 10%. Although models have been developed to forecast the yield at the regional level (Cristofolini and Gottardini, 2000), their results were not precise enough to manage logistic issues linked to harvest operations at the farm and winery level. Therefore, precise estimation of vine field yield always requires fruit sampling and counting. This must be carried out quickly at a time when the workload at harvest is critical. Practical constraints limit the number of sampled sites per field. Therefore, yield estimation is based on a low number of sampling sites (4/5 per field) where yield components (number of clusters, number of berries per cluster, mean berry weight) are manually measured. Due to these practical constraints and the high within-field variability of grape yield, the small number of observation results in high errors in yield estimation (generally around 20 to 30%).
Recent works (Carillo *et al.,* 2016) have shown the interest of integrating auxiliary data to improve sampling strategies and yield estimation for perennial crops. Among possible auxiliary information, vegetation indices derived from multispectral airborne images is of great interest since they can be used to characterise the spatial variability of several fields; in one acquisition, with a high spatial resolution (<1 m.) and at an optimal date. In viticulture, Carrillo *et al.* (2016) showed the potential of normalised difference vegetation index (NDVI) to drive target sampling of the main grape yield components (e.g. bunch number, berry weight) to improve yield estimation. They demonstrated the value of using NDVI information to determine relevant within-field sampling sites selection based on the distribution of NDVI values.

DOI 10.3920/978-90-8686-888-9_20, © Wageningen Academic Publishers 2019

Although interesting, the methodology proposed by Carrillo *et al.* (2016) presents a significant drawback. Indeed, it does do not take into consideration the relative position of the sites to be sampled, and the fact that vine fields are structured in rows. This peculiarity implies that rows cannot be crossed, leading to sampling plans optimized in terms of prediction but potentially unrealistic in terms of sampling routes and resulting travelled distance (and time) for the operator.
This paper proposes a new approach to optimally design within-field sampling routes which take into account the spatial organisation of the crop (rows) and spatial location of sampling sites. The originality of the approach, called constraint sampling, is to combine statistical and computer methods. It can be decomposed into two steps. In the first step, potential sampling sites are sorted into different groups according to their auxiliary data value in a similar way to traditional targeted sampling. The second step finds an optimal route that passes through one sampling site from each group. A constraint programming solver is used to build an optimal route in terms of travelled distance. This kind of solver has already been used in precision viticulture to solve the differential harvest problem (Briot *et al.,* 2016).

Materials and methods

Sampling sites and selection principles

The purpose of *constraint sampling* is to select N sampling sites constituting a sampling route in the vine field. Accounting for classical sampling practices in viticulture, N will vary between 5 and 10. It is assumed that there is a finite number of sites on the plot where sampling can be carried out, these sites are called potential sampling sites. For instance, in the data presented in this article, a potential sampling site is defined every 15 m. The sampling sites are then chosen from the list of potential sampling sites. In order to be able to apply selection methods based on auxiliary data, an NDVI value must be associated with each potential sampling point. For each potential site, the method assumes that: (i) the co-ordinate of the potential site; (ii) the row that the potential site belongs to; and (iii) the corresponding NDVI values are known.
The method requires a distance matrix to be computed. This matrix gives the distance between each couple of potential sampling sites. This distance must take into account the structure of the vineyard. It corresponds to the shortest walking distance between two sites. If the points are in the same row, it corresponds to classical Euclidian distance. If they belong to different rows, the distance is computed considering that the practitioner has to leave the row, reach the desired rows passing by all extremities of intermediate rows and finally reach the targeted sampling site. As rows have two extremities, two different distances can be computed and the shortest one is kept.
To consider auxiliary data efficiently, the idea of *model sampling* proposed by Carillo *et al.* (2016) was considered. The *model sampling* approach aims at calibrating a linear regression which relates the yield to auxiliary data (NDVI). This author also shown that yield components, especially berry

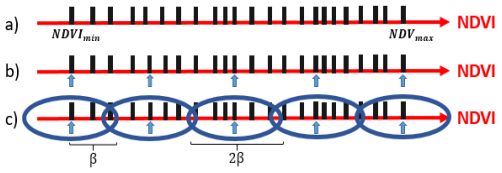


Figure 1. Choice of potential sampling sites (for n=5 sampling sites) based on NDVI values with method (iii) adapted from Kennard and Stone (1969); (A) Distribution of observed NDVI values (each black dash represents one NDVI value). (B) Selection of values corresponding to a potential sampling (arrows) and (C) Groups of potential sampling sites built to account for the distribution of NDVI values, the width of groups is controlled by β.

weight, were linearly related to NDVI. Therefore, sampling sites can be selected representatively to build this linear model. This model is then used to estimate yield using all available high-resolution NDVI data.

The approach proposed in this paper relies on the following principle. Potential sampling sites are split into N homogeneous groups. Once groups are formed, one element from each group is selected in order to optimize the length of the route connecting all these points. It will ensure a good repartition of selected points by accounting for auxiliary data distribution, which is a key element to build a linear model with very few points (Kennard and Stone, 1969). Three different ways to create these groups were tested:

i. The first method relies on quantiles. If K is the number of potential sampling points, then each group has K/N elements. A first group will contain the K/N potential sampling points with the lowest NDVI values. A second one the K/N potential sampling sites with NDVI values just above those of the first group and the last group contains the K/N highest NDVI values.
ii. The second method uses the K-means algorithm. This clustering algorithm is efficient for partitioning K elements into N groups. The main principle of K-means is to minimize the difference between points in the same groups and thus maximise the difference between a group's mean.
iii. The last method is derived from the Kennard and Stone (1969) approach. As described by these authors, this approach selects elements by iteratively choosing a new site that is the furthest from the sites already selected in terms of auxiliary data (Figure 1B). The approach was adapted to create groups, one centred around each value selected by Kennard and Stone approach. A parameter called β, expressed as a percentage of the NDVI range, set the width of the groups (Figure 1C). Using this method, groups may overlap with each other or, on the contrary, some sites may not belong to any group. This depends on the number of groups (N) and their length (β). Figure 1 illustrates the method with potential sampling sites projected on an NDVI axis.

The second step of the approach consists in selecting one sampling site per group. These N sampling sites must be all different and have to form the shortest possible sampling circuit. There are numerous possible choices to select these sampling sites and many ways to order them to form a circuit. It is therefore a highly combinatorial optimization problem. Constraint Programming is one of the programming paradigms able to deal with such problems. It aims at solving a problem expressed as a set of variables and a set of constraints on these variables. Such a problem is called a Constraint Satisfaction Problem (CSP). A Constraint Solver is used to find a solution to the problem that satisfies all the constraints. The efficiency of these solvers relies on the implementation of many methods such as filtering, which allows quick detection of combinations of values that do not lead to an optimal solution. The interest of constraint solvers lies in their ability to address many types of constraints. Without going into small detail, let S = {1,…,K} be the set of potential sampling sites and $\{G_i\}_{i\varepsilon(1,\ldots,N)}$ the set of groups covering S, formed in the previous step. For decomposition (i) and (ii) all groups are disjoint and $\{G_i\}_{i\varepsilon(1,\ldots,N)}$ is a partition of S. P_i is defined as the selected site for group G_i. The first constraint imposes that all P_i must be different (this constraint is immediately satisfied in the case of methods (i) and (ii)). P_0 represents the point of departure and arrival; it is a fixed parameter representing the initial position of the practitioner. The length of the optimum route passing through all the $P_{i\varepsilon(0,\ldots,N)}$ must be a minimum. This is a particular case of the vehicle routing problem (VRP) where the goal is not to find a Hamiltonian tour (visiting once every site) but a tour covering only a subset of sites. Recent work about the *WeightedSubCircuit* constraint (Vismara *et al.,* 2018) has proposed a filtering algorithm that is well adapted to address this type of situation. All these constraints and variables constitute the constraint satisfaction problem. An instance of this problem is built from each dataset and solved with the solver in order to get an ordered set of sites that form a sampling circuit. The program returns the list of sampling sites, the order in which they are visited and the associated distance.

Yield estimation

The aim is to estimate Y, the average grape weight (GW) per vine. For each site selected by the sampling method (s ε selected), GW(s), the observed grape weight per vine value, is available. A linear model linking the NDVI to GW is built from these sites (Equation 1):

$$\widehat{GW}(s) = a \times NDVI(s) + b \quad (1)$$

For a given site s, $\widehat{GW(s)}$ represents an estimate of GW(s). The parameters a and b are obtained from a linear regression on the N sites selected by the sampling method.

With S = {1,…,K} being the full set of potential sampling sites available, $\widehat{\overline{GW}}$, the estimate of $\overline{GW}$, can be computed from the model using all these potential sites (Equation 2):

$$\hat{Y} = mean_{s \in S}(\widehat{GW}(s)) \quad (2)$$

Estimation error

The estimation error is a deviation from the actual yield value (Y), expressed as a percentage (Equation 3).

$$Error\ (\%) = \frac{|Y - \hat{Y}|}{Y} \quad (3)$$

Reference methods

The method is compared to two references:

- A conventional *random sampling* where the N sampling sites are randomly selected. $\widehat{\overline{GW}}$ is directly estimated from the mean of observed GW values. Here, represent the N sampling sites chosen randomly (Equation 4). *Random sampling* represents what is generally done in practice in terms of yield estimation

$$\hat{Y} = mean_{s \in selected}(GW(s)) \quad (4)$$

- A *model sampling* whose method principles have been described by Carillo *et al.* (2016). Sampling sites are chosen according to NDVI values. One site is randomly selected for each of the N NDVI quantiles. *Model sampling* uses a model based on the NDVI/yield relationship, as described in Equation 1. The main difference is that *model sampling* does not consider the spatial position of the selected sampling sites.

To compare the length of the routes between the different methods, the optimal route between the selected sites must be computed for the reference methods. This is done with the Concorde TSP solver. As for *constraint sampling*, the route includes the starting point of the practitioner (P_0).

Experimental data

The data used to test the method came from INRA Pech-Rouge (Narbonne, France). The experiment and the database were detailed by Carrillo *et al.* (2016). It is briefly summarised hereafter. NDVI values from 9 different vine fields were considered. All of them are non-irrigated and exposed to Mediterranean climate with precipitation occurring during spring and with hot and dry summers. The characteristics of each plot are shown in Table 1.

NDVI values were derived from a 1 m. resolution multi-spectral image taken 31 August 2008 by Avion Jaune (Montpellier, Hérault, France). The spectral regions captured in the images were: blue (445-520 nm), green (510-600 nm), red (632-695 nm) and near-infrared (757-853 nm). From these, 1 m square image pixels, aggregation method described in Carillo *et al.* (2016) was used to obtain 9 m square image pixels reducing the effect of canopy discontinuity and bare soil on measured values.

Table 1. Description of the 9 fields.

Field (Id)	Area (ha)	Variety	Row spacing (m)	Vine spacing (m)	Potential sampling sites
P22	1.72	Syrah	2.5	1	45
P63	1.33	Syrah	2.5	1	42
P65	0.69	Syrah	2.5	1	33
P76	1.14	Carignan	2.25	1.5	37
P77	1.24	Syrah	2.5	1	19
P80	0.54	Syrah	2.5	1	40
P82	1.15	Syrah	2.5	1	53
P88	0.85	Syrah	2.25	1.5	21
P104	0.81	Carignan	2.25	1.5	23

NDVI was finally computed from processed images. Mechanical or chemical weeding was performed over the inter-row spacing; therefore, row cover crop did not affect NDVI values.
Sampling sites were selected regularly over the fields with measurement made on each node of a 15 m^2 width sampling grid. At each node, yield components [bunch number per vine (BuN) and bunch weight (BuW)] were measured in 2009. Each site was considered as 5 consecutive vines in the row. BuW was estimated at harvest by weighing 10 bunches (2 bunches per vine) also randomly taken from the same 5 consecutive vines. BuN was determined by counting all bunches of the 5 consecutive vines of each sampling point. Grape weight per vine (GW) was then calculated from BuW and BuN. The distance between vines along the row was 1 or 1.5 m. Data were associated with the spatial co-ordinates of the central vine. The final data base was a set of 313 sites over the 9 different fields. The number of sites per field varied from 19 to 45 sampling sites. Each site was then characterized by GW as field parameter and NDVI values. NDVI value was assigned to each site as the mean of 4 pixels corresponding to a square of 36 m^2.
Actual yield values measured at harvest are not available. For each field, the average of all available measured GW values is then used as the reference yield value (Y) when computing estimation error (Equation 3).

Implementation
The core of the approach was written in java and used the Choco solver (Prud'homme *et al.,* 2016). The calculations to obtain the distance matrix, groups of individuals classified according to their NDVI value, estimation errors and route distance were made with R.
As explained in the description of the constraints, the approach presented here takes into account the starting point of the practitioner (P_0) which is included in the sampling circuit. Varying the starting point thus changes the sampling route. In order to increase the number of situations tested, this starting point was positioned on different ends of row across the vineyards. The approach was then applied to 86 situations instead of 9.

Results and discussion

Evaluation of sampling strategies
Figure 2A and 2B show the results of estimation errors and sampling route distance observed for the different sampling methods. Remember that 'i-quantile', 'ii-kmeans', 'iii-kennard β=10' and 'iii-kennard β=15' refer to the *constraint sampling* methods (i.e. methods that account simultaneously for auxiliary data distribution and distance between sampling points) while *model sampling* and

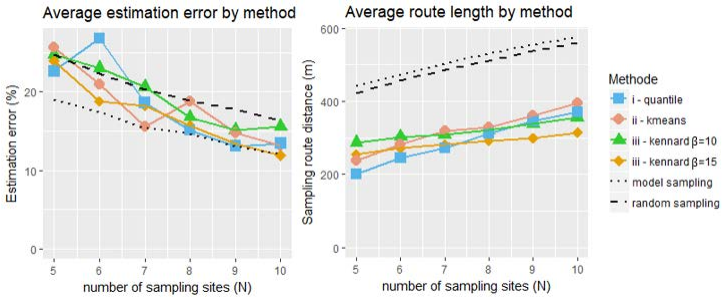


Figure 2. (A) (Estimation error %) and (B) (Sampling route distance): Results and comparison to reference samplings.

random sampling refers to methods that account only on auxiliary data distribution. 'iii-kennard β=10' and 'iii-kennard β=15' are based on the same approach with different group widths (β=10% & β=15%). Results for the different starting points are averaged for each field and then all together to give the same weight to each field.

Figure 2A shows that all the methods follow the same trend with decreasing error as number of samples increases. This result is logical, and consistent with the literature.

As Carrillo *et al.* (2016) have already shown, taking into account auxiliary data (*model sampling*) slightly improves the quality of yield estimation compared to a *random sampling*. Despite higher variability in the observed error, the integration of constraints does not increase estimation errors, the methods (i), (ii), (iii) allow, in most cases, to maintain lower errors than *random sampling*. Kennard and Stone decomposition with β=15 may be the best option when creating the N groups, the results could match those of *model sampling* on most of the cases. Note however that observed errors with constraint methods are higher than for *model sampling* (i.e. without constraints). This result may be logical considering that the addition of the constraints may lead to the choice of less optimal sites within the groups of potential sites. Also, the irregularity of the curves associated with *constraint sampling* can be explained by a smaller number of experiments. These curves are based on 86 results compared to the 1,000 repetitions considered for reference methods, resulting in a higher variability.

Figure 2B clearly illustrates the gains brought by *constraint sampling* in terms of travel distance across the vineyard. Logically, the travelled distance within the plot increases linearly with N, the number of sampling sites visited. The four curves representing *constraint sampling* are at the same level, with a reduced distance of about 45% compared to *model sampling* and *random sampling*. Overall, this method offers a good compromise between the quality of the estimate and the travel constraint on the plot.

Applying the approach to new data could consolidate the results presented here. It would also be interesting to test the method with plots of vines having different characteristics (shape, size), under different cultivation practices (weed management between rows) or with different auxiliary data available (e.g. historical yield).

This is a first model that can still be improved. Increasing the number of usable auxiliary variables or allowing the method to adjust directly as the first sites are selected for instance, could improve the accuracy and quality of the results.

Computation times

Computation times increase with the number of possible combinations. The higher K and N (the number of potential sampling sites and the number of sites to be selected respectively), the longer it will take. The way groups are created also affects the computational time. For instance, when using

the Kennard and Stone approach to build a group, an increase of the β parameter (group width) can consistently increase computation times. In general, for plots with N<8, the computation times are in the order of a second. It took about a few hours in the most complex cases.

Conclusions

The methodology presented in this paper described a new approach for yield sampling in viticulture. The originality of the approach comes from the association of a previously published method based on auxiliary data and optimisation algorithms to propose relevant sampling routes in term of estimation error and travelled distance. While the *model sampling* principle guides sampling choice considering auxiliary information, optimisation through constraint programming ensures the relevancy of the chosen route in term of walking distance for the practitioner. Results presented here are of course preliminary results.
As available time is often the principal constraint for growers, they tend to rely on random samplings limited to a small part of the vineyard. Integrating spatial aspect accounting for travelling constraints is a key element to propose new methods that are relevant for field application. Further tests should be considered to confirm these first results and identify the limitations of the approach.

Acknowledgements

This work was supported by the French National Research Agency under the Investments for the Future Program, referred as ANR-16-CONV-0004 (#Digitag).

References

Briot, N., Bessiere, C., Tisseyre, B. & Vismara, P., 2015. Integration of Operational Constraints to Optimize Differential Harvest in Viticulture. In J Stafford (Ed.), Proceedings Of the 10th European Conference on Precision Agriculture (ECPA 2015), the Netherlands: Wageningen Academic Publishers. pp 487-494.

Carrillo, E., Matese, A., Rousseau, J., & Tisseyre, B., 2016. Use of multi-spectral airborne imagery to improve yield sampling in viticulture. Precision Agriculture, 17(1), 74-92.

Cristofolini, F. & Gottardini, E., 2000. Concentration of airborne pollen of Vitisvinifera L. and yield forecast: a case study at S.Michele all'Adige, Trento, Italy. Aerobiologia, 16(1), 125-129.

Kennard W., and Stone L. A., 1969. Computer Aided Design of Experiments. Technometrics, 11, 137-148.

Prud'homme C., Fages J.G. & Lorca X., 2016. Choco Documentation. TASC, INRIA Rennes, LINA CNRS UMR 6241, COSLING S.A.S, http://www.choco-solver.org

Vismara P. & Briot N., 2018. A Circuit Constraint for Multiple Tours Problems. In Proceedings of 24th Int. Conference on Principles and Practice of Constraint Programming (CP 2018). Lecture Notes in Computer Science. 11008, 389-402.

Assessment of grapevine yield and quality using a canopy spectral index in white grape variety

M. Sozzi[1*], A. Kayad[1], D. Tomasi[2], L. Lovat[2], F. Marinello[1,3] and L. Sartori[1,3]
[1]*Department of Land, Environment, Agriculture and Forestry, University of Padova, Viale dell'Università 16, Legnaro (PD), Italy; marco.sozzi@phd.unipd.it*
[2]*CRA-VIT Centro di Ricerca per la Viticultura, Via XXVIII Aprile 26, Conegliano (TV), Italy*
[3]*NEOS srl, Spin-off of the University of Padova, Piazzetta Modin, 12, Padova (PD), Italy*

Abstract

Two vineyards were chosen as a study area to analyse the relation between the spectral sensor measurements and grape quality. Both fields were investigated through an active spectral sensor to retrieve the NDVI. Results were related to total soluble solid, pH, titratable acidity and yeast assimilable nitrogen measured on grape samples. The results showed a good correlation (r=0.74) between quality data and NDVI at veraison, and lower correlation (r=0.64) between quality and NDVI in late ripening. According to this study, a proximal sensor can be attached to a vehicle to rapidly assess grape quality before harvesting in white grape variety.

Keywords: precision viticulture, grape yield, grape quality

Introduction

In the last decade, precision viticulture technologies have become widespread in the most renowned winemaking regions due to high standard quality needs and the continuous development of this sector. Precision viticulture enables enhancement of the oenological features of grapevine through variable rate applications (Llorens *et al.,* 2010) and differential harvesting (Briot *et al.,* 2015). Indeed, precision viticulture tools enable farmers and winemakers to see how terroir features change according to within-field variability (Bramley and Hamilton, 2007). Moreover, innovative solutions of precision viticulture lead to a cost reduction, improving crop quality, yield and environmental sustainability (Matese and Di Gennaro, 2015). Besides, within-field variability knowledge is used to define stable management zones, according to the potential yield of each point (Arnó *et al.,* 2009). Stable management zones can be defined by measuring different crop and soil characteristics. Growing attention in wine production is being paid to grape quality, which is now studied as a source of variability. Assessment of yield quality, in terms of total soluble solids, titratable acidities or other oenological features, to define the harvest moment is often based on manual sampling of grapes (Urretavizcaya *et al.,* 2014). Manual sampling is a time-consuming task and it is affected by a relatively low accuracy, especially if the definition of the sample points is not based on already detected field variability.
On the other hand, optical sensors are widely used in variability detection for precision viticulture, both proximal and remote (De Benedetto *et al.,* 2013). Proximal sensors are capable of detecting vegetative features and fruit characteristics. Sensing of vegetative characteristics often utilizes indices like the normalised difference vegetation index (NDVI), which returns information on canopy biomass and plant health. Associated sensors are usually tractor-mounted to combine field-operation and surveys. In fruit monitoring, spectrometry is used to detect the ripening stage, chemical compounds concentration and technological features. These sensors need to be calibrated in the laboratory to retrieve specific fruit features. Saranwong *et al.* (2004) successfully developed a technique to predict eating quality of fruit using spectrometry. Nicolaï *et al.* (2007) reviewed near-infrared spectrometry for fruit quality, stating that typically a root mean square error of prediction

DOI 10.3920/978-90-8686-888-9_21, © Wageningen Academic Publishers 2019

of 1° Brix can be achieved. In precision viticulture, the use of a handheld spectrophotometer with integrated GNSS is spreading, to monitor grape maturation (Geraudie *et al.,* 2010). Unfortunately, spectrometers for fruit monitoring need a short distance between sensor and target. Consequently, they cannot be used to quickly map all the vineyard, and few samples are generally not completely representative of the field variability. A correlation between yield, grape quality and vegetation indices was found by Lamb *et al.* (2008). Quality parameters and there relation with yield and vegetation index were studied by Bramley and Hamilton (2004) in Cabernet Sauvignon. NDVI was used by Martinez-Casasnovas *et al.* (2012) to delineate zones with different wine grape properties for selective harvesting in Syrah and Cabernet Sauvignon. According to the latter works, relation between vegetation indices and grape chemical properties was deeply investigated in red grape varieties (Syrah and Cabernet Sauvignon), finding good relation due to colour and phenolic content of the grape. The present work studied the assessment of grape quality characteristics using a commercial proximal sensor to produce a vegetation index in white grape varieties at different phenological stages (veraison and ripening).

Materials and methods

Two vineyards located in Northern Italy were chosen for testing the possibility to use a spectral sensor to measure grape quality characteristics. These vineyards are cultivated with a white grape variety (Vitis vinifera cv. 'Glera') which is destined to become sparkling wine and it represents the most cultivated cultivar in this part of Italy. Both vineyards were planted between 2011 and 2015 with a plant density of 3,700 plants/ha. Vines were trained by spurred cordon with two green pruning events through the season. Yield quality assessment and vegetation index survey occurred during 2018 crop season. Two surveys with GreenSeeker™ crop sensing system (Trimble Navigation Limited, Sunnyvale, CA, USA) were carried out during veraison (25 July, Figure 1A) and in late ripening (12 September). GreenSeeker is an active portable spectrophotometer capable of measuring reflectance at 656 nm and 774 nm, which are used to automatically compute the NDVI values. Surveys were performed using an all-terrain vehicle carrying the sensor with a DGPS antenna (Trmble Juno T41) with an accuracy of ±1 m and a sampling frequency of 1 Hz. Sensor was installed on the all-terrain vehicle with a distance of 1 m from the canopy, to avoid the background detection (FOV 60°). Georeferenced data were collected with a spatial resolution of 2.7×1 m (according to row and inter-row spacing). Both sampling (veraison and ripening) were performed with an average speed of 4 km/h.

Yield data were collected with a grape harvester integrating load cells under the discharge conveyor belt. Grapes were harvested at technological maturity stage (24-25 September, Figure 1B) in order to achieve the best result in the wine-making process. Yield data are expressed in terms of percentage of the mean value. Yield, veraison NDVI and ripening NDVI were analysed and cleaned with ArcGIS Pro 2.2.4 (Esri, Redlands, CA, USA) and then were interpolated using ordinary kriging with a lag size of 2.5 m. Grape samples were collected before harvesting at 56 points chosen according to stratified sampling design based on the veraison NDVI survey (Figure 1B). NDVI at veraison were analysed with MZA (Management Zone Analyst, University of Missouri-Columbia) to find homogeneous zone. Inside each zone, a random sampling was applied choosing ~5 points/ha. Grape quality parameters were evaluated in terms of total soluble solid (TSS), pH, titratable acidity (TA) and yeast assimilable nitrogen (YAN) through physico-chemical laboratory testing. Table 1 reports descriptive statistics of NDVI, yield and quality parameters at the sample points. Yield quality samples were collected randomly from different part of bunches and at different height of the canopy (both sides of a row).

Based on these datasets, a statistical analysis of correlation and regression was carried out to evaluate the possibility of assessing grapevine yield and quality using proximal spectral sensing. Linear regression based on a uniform level of yield was studied to estimate quality parameters.

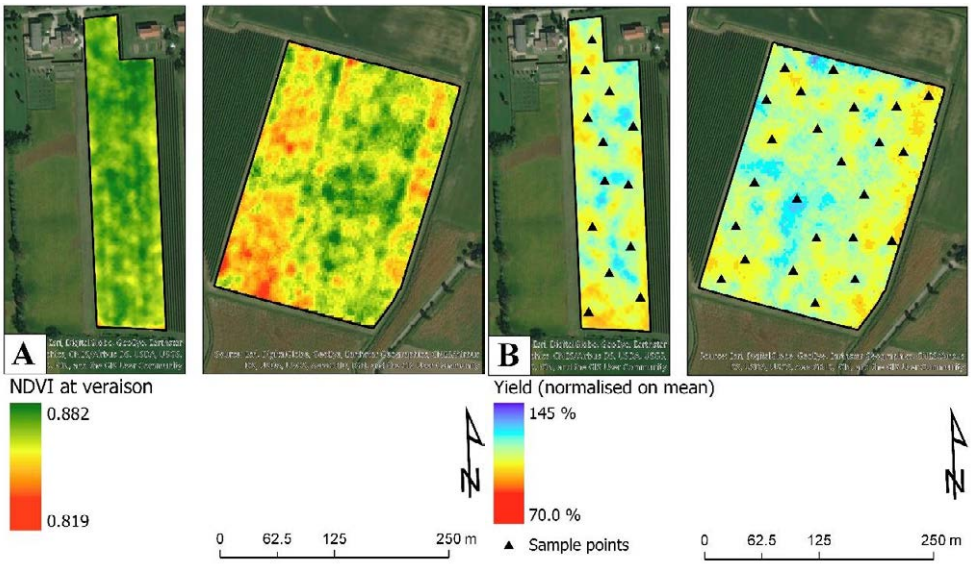


Figure 1. NDVI at veraison and yield maps.

Table 1. Descriptive statistics of the parameters.

	NDVI veraison	NDVI ripening	Yield (%)	TSS (%)	pH	TA (g/l)	YAN (mg/l)
Mean	0.859	0.786	98.9	14.9	3.17	6.84	157
Standard deviation	0.0086	0.0198	9.05	0.862	0.0582	1.07	72.0
CV (%)	1.00	2.52	9.15	5.80	1.84	15.59	45.92
Range	0.041	0.105	48.0	4.4	0.23	4.26	260.5
Kurtosis	0.733	2.52	0.941	0.821	-0.460	-0.429	-0.814
Skewness	-0.687	-1.36	0.389	0.440	0.163	0.548	0.435

Distribution of raw yield values and associated NDVI values were grouped into classes in intervals of 2.5%. Yield production classes were used to classify NDVI and quality parameters in order to analyses correlation at uniform yield classes. NDVI data were extracted from interpolated map, due to the different spatial resolution of the yield quality survey.
Then, sample points feature inside each class were averaged. Statistical analysis was performed on these classes (15) taking advantages of Statgraphics Centurion XVII software (Statgraphics Technologies, Inc., Warrenton, VA, USA).

Results and discussion

The first step was to analyse the correlation coefficient (Pearson coefficient, *r*) between all the parameters considered. The highest correlation value was found between veraison NDVI and yield (Figure 2), as demonstrated by Hall *et al.* (2011) in grapevine and in other crops by Kayad *et al.* (2016). Veraison NDVI showed a negative correlation with TSS (-0.74).

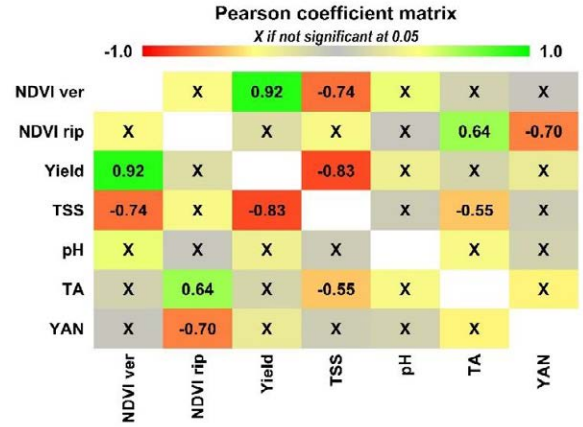


Figure 2. Correlation between all considered parameters.

In addition, ripening NDVI was correlated with TA and YAN with correlation coefficients of 0.64 and -0.70. In yield quality parameters, a weak correlation was found between TSS and TA (-0.55). Yield showed a strong negative correlation with TSS (-0.83). In the cases where a significant correlation was found, regression analysis was completed using NDVI as an explanatory variable of quality parameters. R^2 squared of 55.3% was found in the regression model of TSS based on veraison NDVI (Figure 3). Ripening NDVI showed an R^2 of 41.1 and 49.6% for TA and YAN. To study the possibility of early detection of quality parameters at veraison, regression analysis based on NDVI was carried out even if the correlation was not significant. In the regression model based on veraison NDVI, the maximum value of YAN (250.35 g/l) was considered as an outlier and excluded from the regresion, due to the high distance from the fitting model calculated without that point (4.07 σ). Accordingly, the final regression model for YAN showed an R^2 of 45.1% (Figure 4). The Regression models for pH and TA based on veraison NDVI were not sufficiently strong, confirmed by a statistically insignificant correlation.

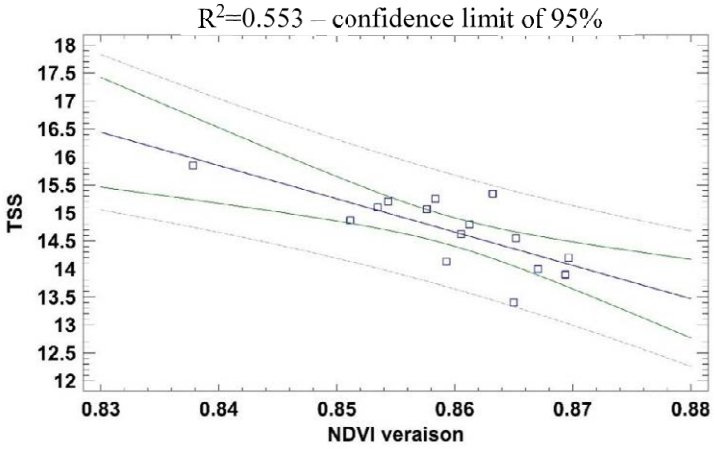


Figure 3. Linear regression model based on veraison NDVI for TSS.

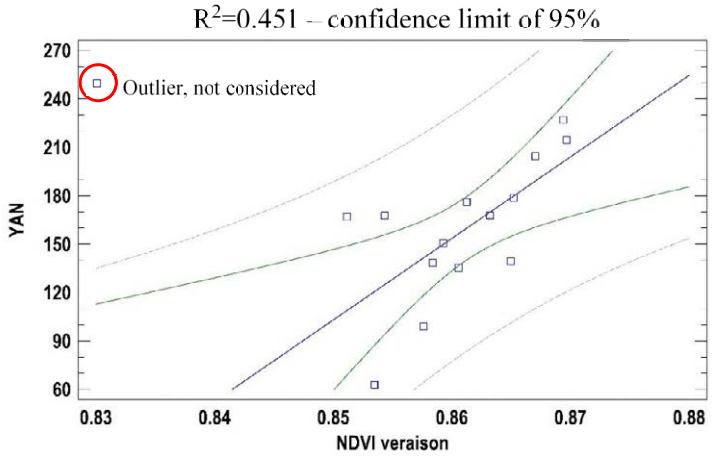


Figure 4. Linear regression model based on veraison NDVI for YAN.

The linear regression model based on veraison NDVI for TSS showed a high value of R^2 and narrow confidence limits (Figure 3). In contrast, the linear regression model based on veraison NDVI for YAN, showed a lower R^2 value and wider confidence limits (Figure 4) even when excluding an outlier from the analysis.
Future studies could more deeply explore the correlation between NDVI and grape quality in white varieties, to better understand physiological processes which affect this correlation. The relation between YAN and NDVI is an interesting parameter which should be investigated in the future.

Conclusions

The present study evaluated the possibility of assessing quality parameters using a proximal spectral sensing. The NDVI retrieved from a GreenSeeker™ sensor showed a good correlation with TSS, TA and YAN. In contrast, no correlation was found between NDVI and pH. Variability in TSS and YAN can be partially predicted from the NDVI at veraison whereas TA and YAN can be partially predicted from the NDVI in late ripening. From a practical point of view, assessment of quality parameters during veraison is more useful than assessment during ripening because it affords more time to manage field operations (e.g. green pruning) and thus improve or maintain the quality features required for the winemaking. Similar studies performed on red grape variety showed higher value on correlation and more performant regression model (Bramley and Hamilton, 2004; Lamb *et al.,* 2008; Martinez-Casasnovas *et al.,* 2012). Polyphenols and anthocyanins content of red grape variety strongly affect the spectral response of vineyard to vegetation indices especially during the veraison. These chemical compounds are not contained in white grape variety, leading to a weaker relation between vegetation indices and grape quality. According to the result of this work, winegrowers might be able to take advantage of spectral sensors to assess sugar content and yeast assimilable nitrogen in white grape variety with a NDVI survey during veraison.

Acknowledgements

This project was developed with the support of Cantina sociale di Cona, Cona (VE) and Centro di Ricerca per la Viticultura, Conegliano (TV).

References

Arnó, J., Martínez Casasnovas, J. A., Ribes Dasi, M., & Rosell, J. R., 2009. Review. Precision viticulture. Research topics, challenges and opportunities in site-specific vineyard management. Spanish Journal of Agricultural Research, 7 (4), 779-790.

Bramley, R. G. V., & Hamilton, R. P., 2004. Understanding variability in winegrape production systems. Australian Journal of Grape and Wine Research, 10 (1), 32-45.

Bramley, R. G. V., & Hamilton, R. P., 2007. Terroir and precision viticulture: Are they compatible? Journal Internationale des Sciences de la Vigne et du Vin, 41 (1), 1-8.

Briot, N., Bessiere, C., Tisseyre, B., & Vismara, P., 2015. Integration of operational constraints to optimize differential harvest in viticulture. In J V Stafford (Ed.) Proceedings of the 10th European Conference on Precision Agriculture, Precision Agriculture '15. the Netherlands: Wageningen Academic Publishers. pp 487-494.

De Benedetto, D., Castrignano, A., Diacono, M., Rinaldi, M., Ruggieri, S., & Tamborrino, R., 2013. Field partition by proximal and remote sensing data fusion. Biosystems Engineering, 114 (4), 372-383.

Geraudie, V., Roger, J. M., & Ojeda, H., 2010. Développement d'un appareil permettant de prédire la maturité du raisin par spectroscopie proche infra-rouge (PIR) (Development of a device to predict the ripeness of grapes by near-infrared spectroscope). Revue française d'Oenologie, 240 (1), 2-8.

Hall, A., Lamb, D. W., Holzapfel, B. P., & Louis, J. P., 2011. Within-season temporal variation in correlations between vineyard canopy and winegrape composition and yield. Precision Agriculture, 12 (1), 103-117.

Kayad, A. G., Al-Gaadi, K. A., Tola, E., Madugundu, R., Zeyada, A. M., & Kalaitzidis, C., 2016. Assessing the Spatial Variability of Alfalfa Yield Using Satellite Imagery and Ground-Based Data. PLOS ONE, 11(6).

Lamb, D. W., Weedon, M. M., & Bramley, R. G. V. 2008. Using remote sensing to predict grape phenolics and colour at harvest in a Cabernet Sauvignon vineyard: Timing observations against vine phenology and optimising image resolution. Australian Journal of Grape and Wine Research, 10 (1), 46-54.

Llorens, J., Gil, E., Llop, J., & Escolà, A., 2010. Variable rate dosing in precision viticulture: Use of electronic devices to improve application efficiency. Crop Protection, 29 (3), 239-248.

Martinez-Casasnovas, J. A., Agelet-Fernandez, J., Arno, J., & Ramos, M. C., 2012. Analysis of vineyard differential management zones and relation to vine development, grape maturity and quality. Spanish Journal of Agricultural Research, 10 (2), 326-337.

Matese, A., & Di Gennaro, S. F., 2015. Technology in precision viticulture: a state of the art review. International Journal of Wine Research, 7, 69-81.

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., *et al.,* 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Postharvest Biology and Technology, 46 (2), 99-118.

Saranwong, S., Sornsrivichai, J., & Kawano, S., 2004. Prediction of ripe-stage eating quality of mango fruit from its harvest quality measured nondestructively by near infrared spectroscopy. Postharvest Biology and Technology, 31 (2), 137-145.

Urretavizcaya, I., Santesteban, L. G., Tisseyre, B., Guillaume, S., Miranda, C., & Royo, J. B., 2014. Oenological significance of vineyard management zones delineated using early grape sampling. Precision Agriculture, 15 (1), 111-129.

Assessment of vineyard trimming and leaf removal using UAV photogrammetry

J. Torres-Sánchez[1*], D. Marín[2], A.I. De Castro[1], I. Oria[2], F.M. Jiménez-Brenes[1], C. Miranda[2], L.G. Santesteban[2] and F. López-Granados[1]
[1]*Institute for Sustainable Agriculture, Campus Alameda del Obispo, Avda. Menéndez Pidal s/n, 14004 Córdoba, Spain; jtorres@ias.csic.es*
[2]*Universidad Pública de Navarra, Dpto. Agronomía, Biotecnología y Alimentación, Edificio Los Olivos, 31006 Pamplona, Spain*

Abstract

Canopy management operations, such as trimming and leaf removal, are among the most relevant agricultural practices in viticulture. This research presents an innovative procedure for evaluating them in a vineyard using two phases: (1) generation of photogrammetric point clouds using UAV technology; (2) analysis of the point clouds through object-based image analysis (OBIA) algorithms. A UAV was flown over a vineyard before and after trimming and leaf removal treatments with different intensities. The OBIA algorithm developed was fully automatic. Results showed the feasibility of the detection of the trimming and leaf removal treatments, and the possibility of discriminating among different treatment intensities. Furthermore, the detected projected area of vineyard was well correlated with the field measurements of the leaf area.

Keywords: point cloud, OBIA, vine architecture, 3D, canopy management

Introduction

Canopy management comprises some of the most relevant agricultural practices in viticulture, and includes operations such as shoot trimming or tipping, shoot thinning and leaf removal. These operations contribute to improve bunch microclimate, balance the leaf to fruit ratio, and also ease the compliance of other crop practices such as spraying and hand harvesting (Smart and Robinson, 1991). Among canopy management practices, shoot trimming and leaf removal are probably the most prevalent ones, being performed at least once a year in most vineyards around the world. For growers and field managers, it is important to be sure trimming is carried out in a proper and timely way. However, this supervision demands a visual inspection of the whole vineyard, which is time-consuming and laborious.

Unmanned aerial vehicles (UAVs) are efficient and affordable tools increasingly used in agriculture. They are able to fly at low altitudes taking images with high overlap using low cost cameras. This permits the generation of 3D models representing the crops by using photo-reconstruction techniques or artificial vision (Nex and Remondino, 2013).

The UAV-based 3D models have been recently used in woody crops, for example to calculate tree height and crown area and volume in olive orchards (Torres-Sánchez *et al.,* 2015), and to quantify the impact of different pruning (Jiménez-Brenes *et al.,* 2017); to characterize the canopy architecture of vineyards (Burgos *et al.,* 2015; de Castro *et al.,* 2018a); or to estimate the biomass of poplar plantations (Peña *et al.,* 2018).

Processing the large amount of data embedded in 3D models requires the implementation of robust and automatic image analysis procedures. In the last decade, object-based image analysis (OBIA) techniques have reached high levels of automation and adaptability to ultra-high spatial resolution images and yield better results than conventional pixel-based methods (Blaschke *et al.,* 2014). Recent

DOI 10.3920/978-90-8686-888-9_22, © Wageningen Academic Publishers 2019

research has shown successful applications of OBIA techniques in agricultural scenarios (de Castro *et al.,* 2018b; Mathews, 2014; Souza *et al.,* 2017).
In this study, a UAV equipped with a conventional camera was flown at 30 m height over a commercial vineyard before and after two trimming treatments that differed in the height at which the cutting element was fixed. The images from the UAV were used for the generation of 3D point clouds representing the vineyard, and an OBIA algorithm was used for measuring the height and volume of the vineyard. The specific objectives of this research were: (1) detection of trimming and leaf removal treatments; (2) discrimination of the trimming and leaf removal intensity.

Materials and methods

Study site

The experiment was carried out in a commercial vineyard located in La Rioja (42.366° N, 2.096 W, WGS84) Spain, representative of the growing practices in the region. Planting distance was 2.8×1 m, the variety was Tempranillo, and the vines were trellised as a vertical-shoot positioned unilateral spur-pruned cordon. Soil management included a natural cover crop, mown twice during the spring. The cover crop presented very low development at the moment of the measurements due to scarce summer rainfall.

Canopy management treatments

Shoot trimming and leaf removal operations were performed on 20th June 2018 in order to determine the potential of UAV photogrammetry to detect them and to discriminate their intensity. Regarding shoot trimming, two 6 m-long vine row segments were applied, differing in the height at which the cutting element was fixed. In the low intensity treatment, trimming was performed by cutting shoots at 0.80 m from the main branch. In the high intensity trimming, the cutting was done at ca. 0.45-0.50 m from the main branch. Concerning leaf removal treatments, a similar scheme was applied. All the leaves in the basal part of the shoot were removed, irrespective of whether they belonged to the main shoot or to laterals. Two leaf removal intensities were chosen, low implying leaf removal in the lower 0.40 m of the shoot, whereas high implied leaf removal in the lower 0.70 m.

UAV flights

The remote images were acquired with a commercial off-the-shelf camera, model Sony ILCE-6000 (Sony Corporation, Tokyo, Japan) mounted in a quadcopter model MD4-1000 (microdrones GmbH, Siegen, Germany). The camera has a 23.5×15.6 mm APS-C CMOS sensor capable of acquiring 24 megapixel (6,000×4,000 pixels) spatial resolution images, and is equipped with a 20 mm lens. The UAV can be either manually operated by radio control (1000 m control range) or execute user-defined flight routes autonomously by using its autopilot in combination with its Global Navigation Satellite System (GNSS) receiver. The UAV is battery powered and can load any sensor weighing up to 1.25 kg. The flight route was designed with forward and side overlaps of 91 and 60%, respectively; and flight altitude was 30 m.

Point cloud generation

The 3D point cloud was generated by using the Agisoft PhotoScan Professional Edition software (Agisoft LLC, St. Petersburg, Russia) version 1.4. The process was fully automatic; the only user intervention was the manual localization in the software of 6 ground control points taken in the field with a real time kinematic (RTK) GPS. The estimated accuracy of GNSS-RTK system is 0.02 m in planimetry and 0.03 m in altimetry. The resulting point cloud was stored in '.las' format, a public file format for the interchange of 3D point cloud data. The processing parameters of the software were the same as the ones used in by de Castro *et al.,* 2018a).

Point cloud OBIA analysis

The OBIA algorithm for the detection and characterization of grapevine rows was developed with the eCognition Developer 9 software (Trimble GeoSpatial, Munich, Germany). The only input in this process is the point cloud. The algorithm does not need any user intervention and is divided into the following steps automatically executed:

- DTM generation*:* the point cloud was segmented using a chessboard pattern with 2 m side. For the DTM generation, a new image layer was created with 2 m resolution, where each square is assigned the average height of the 15% of points with lower height, that are supposed to be points belonging to the soil.
- Vineyard classification: the point cloud is segmented using a chessboard pattern with 0.1 m side. All the squares whose points have a height over the DTM higher than 0.7 m were classified as vineyard. A new point cloud containing the points from the areas classified as vineyard was created. This new point cloud did not include any point belonging to the soil.
- Point cloud slicing: the new point cloud was divided in slices with 0.1 m height. Considering that the point cloud was previously divided in 0.1 m squares, it follows that the vine point cloud was divided in 3D pixels (voxels) with 0.1 m side. A new image layer was created with a resolution of 0.1 m, storing in every pixel the number of voxels containing points belonging to the vine that had the same *x,y* co-ordinates as the pixel. Consequently, knowing the voxel volume ($0.1 \times 0.1 \times 0.1$ m^3) and the number of voxels sharing *x,y* co-ordinates with every pixel, it is possible to know the volume occupied by the vines in every pixel.
- Vine row segmentation: the vine rows were divided in slices with a user-defined length, which in this case was 0.1 m.
- Vine segments characterization: for every one of the vine segments, the volume is calculated taking into account the number of voxels including vine points that are under the vine slice. The maximum height of every segment is calculated by comparing the point height with the DTM height. And the segment projected area is extracted from the classification performed in the second step of the algorithm.

The vineyard slices were grouped for performing the statistical analysis and the validation by using the field data. The groups for studying the discrimination of treatment and intensities were formed by 10 slices (1 m) since this was the approximate length of the vine plants; they are called 'vine group' through the rest of the manuscript. The validation of the OBIA estimations was done comparing them with the measured foliar area in the 6 m row segments, i.e. grouping 60 vine slices.

Statistical analysis were done with JMP software (SAS, Cary, NC, USA). Mean comparisons among variables for different intensities or moments (after and before the treatments) were performed using Student's t-test with a level of significance of 5%.

Field measurements of foliar area

Total leaf area was determined for each of the 6-vine replicates before and after the compliance of canopy management operations. In order to determine leaf area before canopy management operations, a three-step procedure protocol based on allometric relationships was followed (Miranda *et al.,* 2017). First, a sample of 50 leaves gathered from each field was used to establish a relationship between individual leaf area (LA) and secondary leaf vein length (L_{SN}) using ImageJ software (Schneider *et al.,* 2012), allowing a very good estimation of individual leaf area ($LA = 0.51L_{SN}^2 - 2.21L_{SN} + 8.67$, $R^2>0.99$). That relationship was applied to estimate shoot leaf area (LA_S) of 40 shoots by measuring shoot length (L_S), and the length of the secondary veins of all their leaves, obtaining a relationship to estimate it ($LA_S = 0.179L_S^2 + 1.87L_S + 196.1$, $R^2=0.92$). Finally, L_S was measured for all the shoots in the 6-vine replicates, and leaf area calculated. After canopy management operations, all the leaves removed were picked, transferred to the laboratory, and their LA estimated using image analysis with ImageJ software.

Results and discussion

The proposed workflow generated 3D point clouds that accurately represented the vineyard. Figure 1 shows the comparison between the transversal section of one of the treated areas after and before the trimming. It can be seen that the height of the vineyard was reduced by the canopy management operation. Although the accuracy of the height detection has not been assessed in this work, previous studies have demonstrated the accuracy of UAV-derived 3D models for quantification of vine height (de Castro *et al.,* 2018a). Furthermore, the height of the vines was significantly different before and after the treatment. On the other hand, there were no significant differences in 'vine groups' maximum height after and before the leaf removal treatment.

The detected volume was significantly different after and before the canopy management operations with no regard to the treatment or the intensity. This change in volume can be seen in Figure 2. For trimming and leaf removal, the volume after the treatment was significantly lower when it was applied with high intensity.

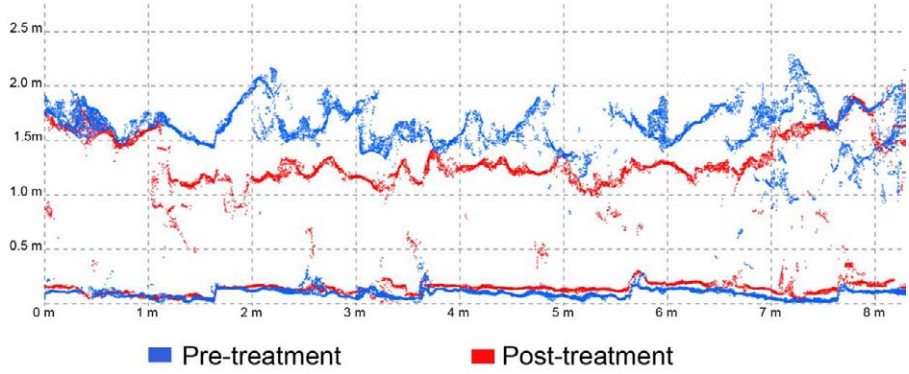


Figure 1. Transversal section of the pre- and post-trimming point clouds along one vine row segment with high intensity trimming. Points in the lower part of the image represent the soil.

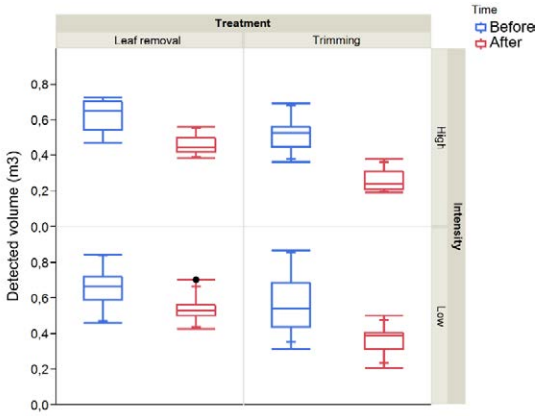


Figure 2. Boxplots for the volume before and after the treatments. The data come from the 'vine groups' (groups of 10 slices).

These results imply that the application of an automatic OBIA algorithm to the 3D point cloud obtained with a UAV equipped with a conventional camera can be used to evaluate if trimming and leaf removal treatments have been performed on a vineyard. The evaluation of volume changes is useful for detecting both treatments, while height changes are only detectable in trimming treatments. The comparison between the detected projected areas of the treatment repetitions (groups of 60 slices) with its measured leaf area (Figure 3) reached a coefficient of regression of 0.85. This comparison was evaluated using the data from both treatments after and before their application.

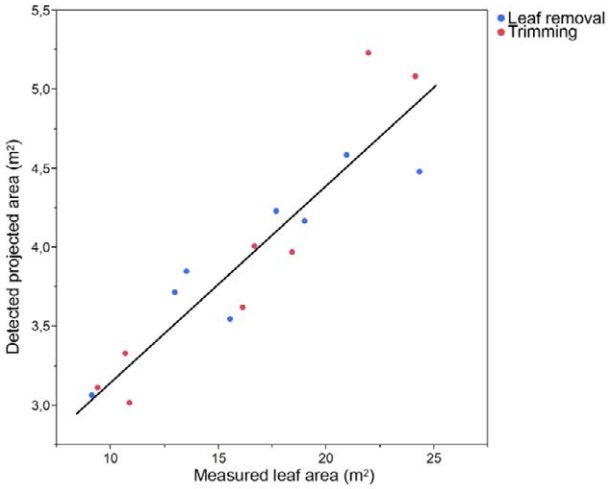


Figure 3. Graphical comparison of the projected area of every treatment repetition and its measured foliar area. The data come from the treatment repetitions (groups of 60 slices).

Conclusions

A UAV equipped with a conventional camera was flown at 30 m height over a commercial vineyard before and after two trimming and leaf removal treatments executed at different intensities. 3D point clouds representing the vineyard were generated with the UAV imagery, and an OBIA algorithm was developed for analysing them. The OBIA algorithm was fully automatic, generating accurate estimations of every vineyard canopy height, projected area, and volume in a very efficient way. Results showed the feasibility of the detection of the trimming and leaf removal treatments, and the possibility of discriminating among different treatment intensities. Furthermore, projected area of the vineyards was well correlated with the on-ground field measurements of the leaf area. This work highlights the high potential of UAV imagery and the OBIA algorithm for a remote evaluation of the compliance of field operations. The information derived can also be used for generating canopy maps that would be useful to design site specific spraying in the context of precision viticulture.

Acknowledgements

This research was supported by the Innovation & Economic Development Department of the Government of Navarre (DECIVID, Ref: 0011-1365-2017-000113, project co-financed with EU funds from the ERDF). D.M. is currently beneficiary of a postgraduate scholarship funded by Universidad Pública de Navarra (FPI-UPNA-2016). Research of A.I. de Castro was financed by the Juan de la Cierva Incorporación Program (IJCI-2015-24811).

References

Blaschke, T., Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., *et al.,* 2014. Geographic Object-Based Image Analysis – Towards a new paradigm. ISPRS Journal of Photogrammetry and Remote Sensing, 87, 180-191. doi:10.1016/j.isprsjprs.2013.09.014

Burgos, S., Mota, M., Noll, D., & Cannelle, B., 2015. Use of Very High-Resolution Airborne Images to Analyse 3d Canopy Architecture of a Vineyard. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 40(3), 399-403.

de Castro, A. I., Jiménez-Brenes, F. M., Torres-Sánchez, J., Peña, J. M., Borra-Serrano, I., & López-Granados, F., 2018a. 3-D Characterization of Vineyards Using a Novel UAV Imagery-Based OBIA Procedure for Precision Viticulture Applications. Remote Sensing, 10(4), 584. doi:10.3390/rs10040584

de Castro, A. I., Torres-Sánchez, J., Peña, J. M., Jiménez-Brenes, F. M., Csillik, O., & López-Granados, F., 2018b. An Automatic Random Forest-OBIA Algorithm for Early Weed Mapping between and within Crop Rows Using UAV Imagery. Remote Sensing, 10(2), 285. doi:10.3390/rs10020285

Jiménez-Brenes, F. M., López-Granados, F., de Castro, A. I., Torres-Sánchez, J., Serrano, N., & Peña, J. M., 2017. Quantifying pruning impacts on olive tree architecture and annual canopy growth by using UAV-based 3D modelling. Plant Methods, 13, 55. doi:10.1186/s13007-017-0205-3

Mathews, A. J., 2014. Object-based spatiotemporal analysis of vine canopy vigor using an inexpensive unmanned aerial vehicle remote sensing system. Journal of Applied Remote Sensing, 8(1), 085199. doi:10.1117/1.JRS.8.085199

Miranda, C., Santesteban, L. G., Escalona, J. M., Herralde, F. D., Aranda, X., Nadal, M., *et al.,* 2017. Allometric relationships for estimating vegetative and reproductive biomass in grapevine (Vitis vinifera L.). Australian Journal of Grape and Wine Research, 23(3), 441-451. doi:10.1111/ajgw.12285

Nex, F., & Remondino, F., 2013. UAV for 3D mapping applications: a review. Applied Geomatics, 6(1), 1-15. doi:10.1007/s12518-013-0120-x

Peña, J. M., Castro, A. I. de, Torres-Sánchez, J., Andújar, D., Martín, C. S., Dorado, J., *et al.,* 2018. Estimating tree height and biomass of a poplar plantation with image-based UAV technology. AIMS agriculture and food, 3, 313-326. doi:10.3934/agrfood.2018.3.313

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nature Methods, 9(7), 671-675.

Smart, R., & Robinson, M., 1991. Sunlight Into Wine: A Handbook for Winegrape Canopy Management. Broadview, Australia: Winetitles Media.

Souza, C. H. W. de, Lamparelli, R. A. C., Rocha, J. V., & Magalhães, P. S. G., 2017. Mapping skips in sugarcane fields using object-based analysis of unmanned aerial vehicle (UAV) images. Computers and Electronics in Agriculture, 143(Supplement C), 49-56. doi:10.1016/j.compag.2017.10.006

Torres-Sánchez, J., López-Granados, F., Serrano, N., Arquero, O., & Peña, J. M., 2015. High-Throughput 3-D Monitoring of Agricultural-Tree Plantations with Unmanned Aerial Vehicle (UAV) Technology. PLoS ONE, 10(6), e0130479. doi:10.1371/journal.pone.0130479

Section 4
Precision crop protection

On-board colour imaging for the detection of downy mildew

F. Abdelghafour[1,2*], F. Rançon[1,2], B. Keresztes[1,2], C. Germain[1,2] and J.-P. Da Costa[1,2]
[1]*Univ. Bordeaux, IMS UMR 5218, 33405 Talence, France; florent.abdelghafour@ims-bordeaux.fr*
[2]*CNRS, IMS UMR 5218, 33405 Talence, France*

Abstract

Conventional crop protection practices in viticulture tend to overuse phytosanitary inputs. That is mainly due to the combination of precaution principles and a lack of insight regarding the spatial variability of diseases in vineyards. This paper aims to present and evaluate a computer vision method designed to detect foliar symptoms of downy mildew from proximal images acquired with an embedded RGB camera. The method is based on the parametric modelling of the colour and structure of downy mildew symptoms. This method was trained and validated on images taken in a Merlot Noir vineyard in the Bordeaux area. The proposed method provided satisfying detection rates (recall), though further work has still to be done to reduce false positives occurrence (precision).

Keywords:, proximal sensing, disease detection, parametric classification, texture, colour

Introduction

Cryptogamic fungal diseases are a major threat for perennial crops, including grapevines. Without proper control, these diseases can infect large areas very swiftly and cause substantial reduction in yields and winemaking potential. Downy mildew, caused by the pathogen *Plasmopara viticola*, is one of the most common and devastating pathogens affecting vineyards. In order to preserve viable yields and grape quality, growers rely mainly on spraying with fungicides. Despite the development of alternative biological control methods, chemical crop protection is still required. Crop protection practices in viticulture are based on regular applications of pesticides following a precautionary principle consisting of preventative spraying plans. Phytosanitary operations are implemented taking into account climatic risks, phenological stages and according to threat indices. Fungal threat indices are currently estimated by visual inspection from sparse samplings in vineyards. As a result, each plot is only described by a uniform value that does not account for spatial variability. With the growing costs of field operations, the collective willingness to reduce their detrimental impact on health and environment, the implementation of more efficient spraying practices becomes necessary.

Precision spraying is one of the most promising techniques to reduce the use of pesticides while preserving the economic interests of farmers. This is the ability to control the flow-rate and the spatial distribution of pesticide inputs across the field according to the site-specific characteristics of the crop (health status and foliage volume), in contrast to a uniform application of field treatments (Tona *et al.*, 2017). A possible way to implement precision spraying could consist in selectively targeting pesticide application where plants exhibit symptoms over a previously defined threshold (Oberti *et al.*, 2016). Epidemics present typical patch structures evolving around discrete foci, especially during the early stages of an infection's development (Everhart *et al.*, 2013). Assuming a reliable assessment of infected foci, it is possible to control the disease establishment and its spreading thanks to surrounding buffer areas while significantly reducing the total amount of pesticide applied (West *et al.*, 2003).

This paper aims at presenting and evaluating a computer vision method designed to detect and map foliar symptoms of downy mildew. The method is based on the joint parametric modelling of colour and texture of downy mildew symptoms, healthy organs and a set of confounding factors including abiotic necrosis, chlorosis and hail damage. A similar scheme has been applied in Abdelghafour *et al.* (2018) for the segmentation of grapevine organs in proximal images. Here the method has been adapted in order to distinguish portions of infected tissues at various severity stages from healthy vines.

DOI 10.3920/978-90-8686-888-9_23, © Wageningen Academic Publishers 2019

Material and methods

Plant material

The plant material used for image acquisition came from two 0.2 ha plots, with 1.20 m row spacing and planted with the red wine grape variety 'Merlot Noir' in 'Le Domaine de la Grande Ferrade' (44°47'31.07'N,0°34'40.38"O), a public experimental facility in Bordeaux area (INRA, French National Institute of Agricultural Research). The plots were grown under conventional standards and managed with a limit in terms of phytosanitary inputs and fungicides set to 50% of the legal dose. At the end of July 2018, the plots were extensively photographed weekly seeking examples of healthy vines and examples of vines showing early and late symptoms (phenological stages between BBCH 75 and 79) (Lorenz *et al.*, 1995).

Instrumentation

The imaging system used for acquisitions was composed of a 5 Mpx industrial Basler Ace RGB camera (acA2500-14gc GigE, Basler AG, Ahrensburg, Germany) with a 55° horizontal field of view lens. To overcome the weather and time dependent variations of illumination in outdoor environments, the imaging system included a high-power 58GN xenon flash (Neewer speedlite 750ii, Neewer Technology Co., Ltd, Guangdong, China) used with a short exposure time (250-300 µsec). All the components were powered by a 12 V battery. The device was equipped with an on-board industrial computer that controlled simultaneously the triggering of the camera and the flash, and stored the acquired image data. The computer was built around a low consumption 4-core ARM chip, robust to vibrations and watertight. The device (Figure 1A) was mounted on a vineyard tractor at 0.7 m above ground and at 0.5 m from the target (Figure 1B). At this distance, each 2,592×2,048 image covered a 1.3 m^2 area, which enabled the capture of a vine and its full canopy at a resolution of 4 px/mm (Figure 1C). Acquisitions were adapted to the work-rate speed in vineyards (3-8 km/h).

Image data

On the acquired images, symptoms of downy mildew appeared mostly as small patches of 'oilspots' that were generally contained in 32×32 px patches (Figure 2A). Some less common symptoms appeared as thin bands at the edges of a leaf (Figure 2B), or as bigger patches containing several coalesced spots (Figure 2C). The symptoms of downy mildew exhibited on grape bunches usually appeared as undeveloped and rotten parts of berries. While unitary infections were relatively small (Figure 2B), larger infections could cover up to a whole bunch (Figure 2E).

Image processing pipeline

The image processing pipeline is meant to detect the presence of downy mildew symptoms among a wide range of organs and textures (leaves, fruits, stems, necrosis, etc.). It hinges on two major central steps: (1) local properties are computed for the pixel of interest and (2) they are submitted to a decision (i.e. classification) process. The local properties are a combination of colour and texture information captured into a single feature vector, as done by Abdelghafour *et al.* (2018) for the



Figure 1. Instrumentation for image acquisition: device enclosure (A), device mounted on a farm tractor (B) and an example of image acquired by the device (C).

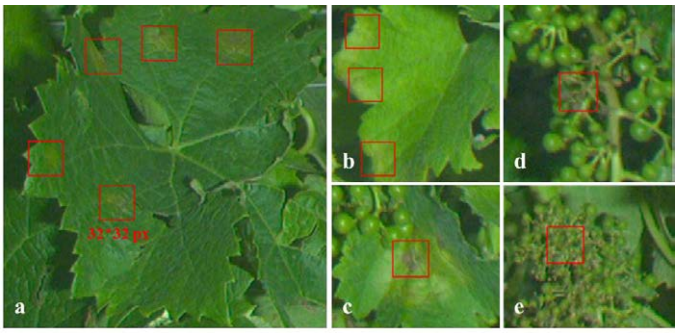


Figure 2. Examples of downy mildew symptoms on vine leaves (A, B and C) and grape bunches (D and E). Red boxes size is 32×32 pixels.

segmentation of grapevine organs. The classification was done using a Bayesian decision process based on the prior probabilistic modelling of the feature vectors from a learning database (Bouman and Sauer, 1993). The decision process can be applied at any pixel position within an image and, when applied on the whole image, is able to produce classification maps at the scale plant (Figure 3).

Joint colour-texture based Bayesian classification of pixels

Local texture information was captured here using the Local Structure Tensor, i.e. LST. The LST is a reference tool developed by Knutsson (1989) that extracts geometric information in local patterns within grayscale images. It is commonly defined as the local covariance of gradients (Rosu *et al.*, 2017). At each pixel position (i,j), the LST is a 2×2 SPD matrix the eigenvectors and eigenvalues of which provide information about local orientation, anisotropy and contrast. In (Rosu *et al.*, 2017), probabilistic models in the log-Euclidean (LE) space were proposed to handle LSTs in Bayesian frameworks. More recently, Abdelghafour *et al.* (2018) proposed the LEEST (Log-Euclidean Extended Structure Tensor) representation which enhances LST data with colour information by concatenating RGB triplets to gradient vectors before computing the LSTs and mapping them into in the LE space. The feature-based method used here is based on an alternative representation called CELEST (Colour Extended Log-Euclidean Structure Tensor). The latter consists in concatenating spatially filtered RGB triplets to the logarithms of the eigenvalues (γ_1, γ_2) of the LSTs. In any position (i,j), the color and texture information is thus summarized by a 5-uplet: $Y(i,j) = [\log\gamma_1, \log\gamma_2, R, G, B]$.

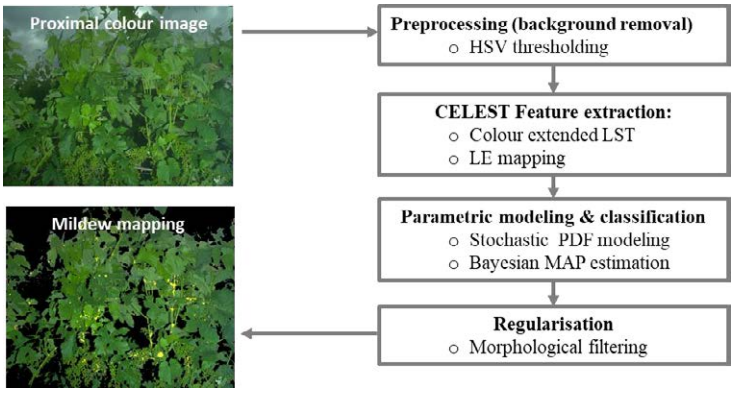


Figure 3. Image processing methodology.

Decision process: classification of pixels

Maximum a posteriori probability estimation (MAP)
In order to assess the class λ that a pixel (i,j) belongs to, from the observation of the feature vector Y(i,j), a Bayesian Maximum a Posteriori (MAP) decision process is applied. It consists in finding the class λ that maximizes $p(\lambda \mid Y)$ i.e. the probability of λ knowing Y:

$$\lambda_{opt} = \arg\max_{\lambda \in L} p(\lambda \mid Y) = \arg\max_{\lambda \in L} f(Y|\lambda)p(\lambda), \quad (1)$$

where $f(Y|\lambda)$ is the likelihood of the feature vector Y in the class λ, p(λ) is the *a priori* probability of class λ and L is the set of all possible classes. p(λ) can be interpreted as the relative proportion of class λ in the pixel sets submitted to the decision process. Regarding $f(Y|\lambda)$, a Gaussian distribution with mean μ_λ and covariance matrix Σ_λ is assumed here:

$$f(Y|\lambda) = f(Y|\mu_\lambda, \Sigma_\lambda) = \frac{1}{2\pi^{5/2}{\Sigma_\lambda}^{1/2}} e^{-\frac{1}{2}(Y-\mu_\lambda)^t \Sigma_\lambda^{-1}(Y-\mu_\lambda)} \quad (2)$$

p(λ), μ_λ and Σ_λ for all classes λ have to be learned prior to the decision process during a learning phase, thanks to representative samples of feature vectors.

Learning phase
For the purpose of detecting downy mildew symptoms in real proximal images of grapevines, it was decided to split each image in six classes to be learned. The six classes are leaves, grape bunches, stems, downy mildew symptoms, confounding factors (necrosis) and leaf edges. These classes depict all the organs represented on the images either they are exhibiting healthy-usual properties or abnormal properties i.e. infections, damage or stress.

Results

The results presented were obtained from 13 images with a total of about 2.6×10^6 labelled pixels following a cross-validation process. To produce each validation-classification map, a model was computed for the six classes from the 12 remaining images that constituted the learning set. The limited size of the image set used here is explained by the time required to manually label the pixels in order to obtain efficient leaning and validation sets. Classification performance are presented as confusion matrices with two machine learning metrics: precision and recall (Table 1, 2 and 3). In order to illustrate the potential of this method outside of the validation set, classifications on full images are presented in Figure 4 and examples of successful detections, recurring omissions, confusions and errors are presented in Figure 5.
Table 1, 2 and 3 present the classification results produced on the validation dataset. Three approaches are presented. In the first approach, the occurrence of mildew is considered as probable as the occurrence of any other class (Table 1) i.e. $p(\lambda) = \frac{1}{L}$. Under this assumption, the process was able to retrieve 72% of the symptoms, but with many false positives, present especially on damaged parts of shoots (Figure 5) that results in a 28% precision. However, the presence of a symptom is a rare event compared to normal healthy state. While it is valid to assume the proportions and spatial repartition of the different organs, it is not possible to estimate a priori the proportion of leaves infected by mildew. The proposed solution was setting a reasonable occurrence rate of mildew which is materialised by uneven *a priori* probabilities. For an example of a 5% *a priori* probability for the mildew (Table 2), the precision was improved to 43%, but only 44% of the symptoms were detected. For the 2% approach (Table 3), the precision improved to 52%, but with a recall rate of 33%.

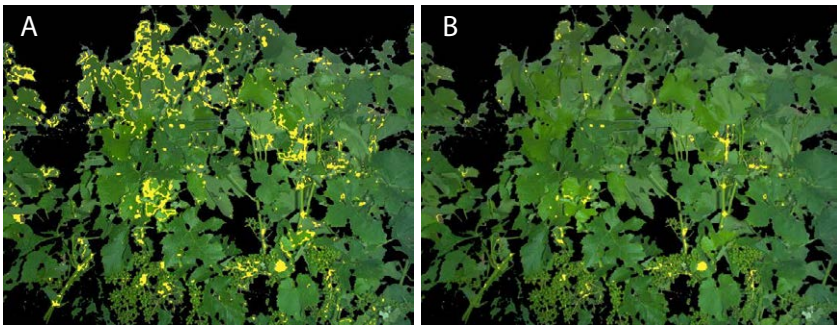


Figure 4. Classification map of downy mildew symptom on a full vinestock, with equal probability for all classes (a), with an a priori probability about the abundance of symptoms arbitrary set to 5%.

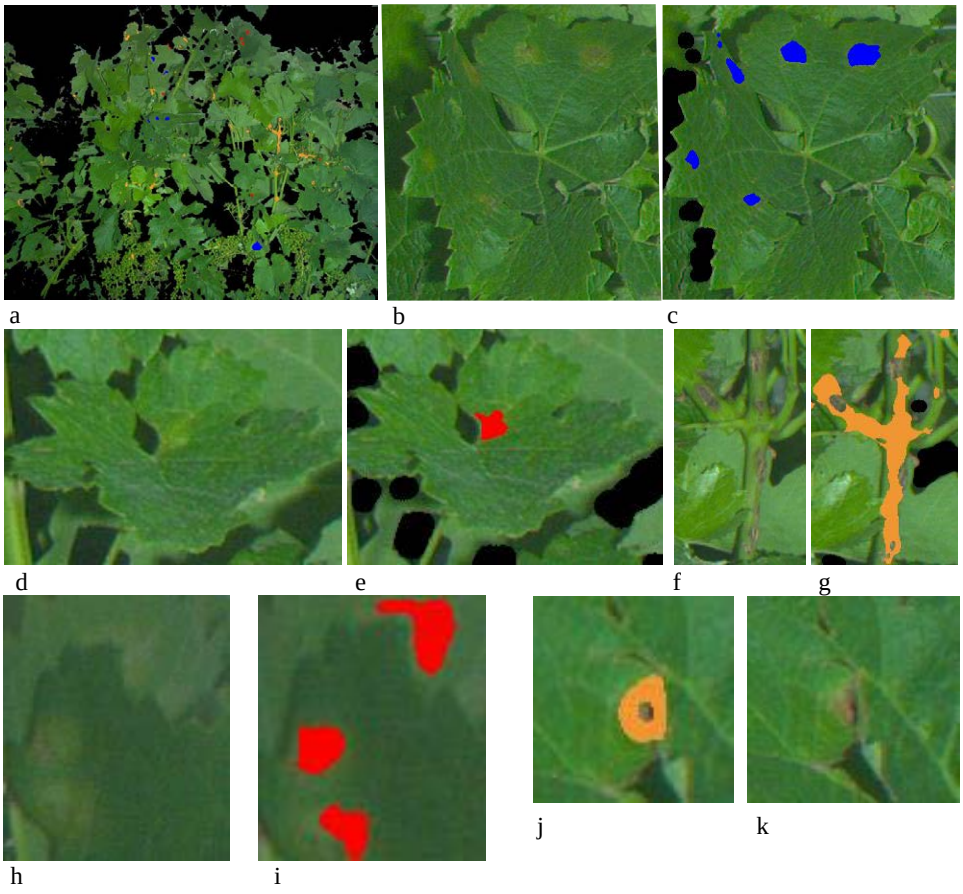


Figure 5. Examples of local decisions compared to real status, in a full image (A). Details: blue=true detection (B,C), red=false negative (D,E; H,I), orange=false positive (F,G; J,K)

Table 1. Confusion matrix and performance metrics for equal a priori probabilities.

		Prediction						Total	Recall
		Mildew	Leaves	Bunches	Stems	Necrosis	Leaf edges		
Truth	Mildew	81,739	8,773	9,667	10,716	1,517	1,083	113,495	72%
	Leaves	14,842	303,404	8,570	13,684	64	20,968	361,532	84%
	Bunches	13,990	8,080	324,389	11,685	506	27,524	386,174	84%
	Stems	178,381	52,381	242,003	644,171	12,549	161,860	1,291,345	50%
	Necrosis	1,185	0	86	979	47,071	27	49,348	95%
	Leaf edges	691	16,693	74,817	27,639	3,081	347,970	470,891	74%
	Total	290,828	389,331	659,532	708,874	64,788	559,432		
	Precision	28%	78%	49%	91%	73%	62%		

Table 2. Confusion matrix and performance metrics with a priori probability for mildew <5% of leaf surface (assuming symptoms can be abundant).

		Prediction						Total	Recall
		Mildew	Leaves	Bunches	Stems	Necrosis	Leaf edges		
Truth	Mildew	49,497	39,190	10,337	13,219	945	307	113,495	44%
	Leaves	828	348,229	4,462	4,225	0	3,788	361,532	96%
	Bunches	2,203	18,441	353,261	7,153	64	5,052	386,174	91%
	Stems	62,301	307,797	241,512	583,036	5,796	90,903	1291,345	45%
	Necrosis	1,239	39	127	2,322	45,595	26	49,348	92%
	Leaf edges	3	115,689	76,886	26,348	278	251,687	470,891	53%
	Total	116,071	829,385	686,585	636,303	52,678	351,763		
	Precision	43%	42%	51%	92%	87%	72%		

Table 3. Confusion matrix and performance metrics with *a priori* probability for mildew <2% of leaf surfaces (assuming symptoms are scarce).

		Prediction						Total	Recall
		Mildew	Leaves	Bunches	Stems	Necrosis	Leaf edges		
Truth	Mildew	37,760	45,120	11,438	17,745	993	304	113,360	33%
	Leaves	355	348,852	4,474	4,242	0	3,744	361,667	96%
	Bunches	1,066	18,433	354,024	7,516	64	5,071	386,174	92%
	Stems	32,758	317,813	245,074	599,351	5,855	90,494	1,291,345	46%
	Necrosis	676	45	132	2,683	45,786	26	49,348	93%
	Leaf edges	0	117,026	76,769	26,185	278	250,633	470,891	53%
	Total	72,615	847,289	691,911	657,722	52,976	350,272		
	Precision	52%	41%	51%	91%	86%	72%		

In the previous classification results, erratic classifications appear on recurring abnormal objects. With a better consideration of these marginal textures, with different models containing more samples for confounding factors, the results could improve. Figure 5 presents examples of symptoms which are usually successfully detected (in blue), the common false positives (in orange) and the cases where the method fails to detect symptoms (in red). Typical 'oilspots' are easily detected (Figure 5B,C), whereas early symptoms (Figure 5D,E) and symptoms on the abaxial side (Figure 5H,I) are harder to detect and often ignored. Some mildew-like textures can be exhibited around necrosis, particularly on stems where a pale yellow discolouration presenting properties as mildew symptoms can appear on their edges (Figure 5F,G) and (Figure 5J,K). In this case it often leads to false detections. The case of the stems is atypical because some necrosis due to hail damage followed by a mildew infection appear on the stems and became quickly indiscernible from mechanical wounds or old dried fungal infections.

Discussion

Judging from both the results computed on the validation dataset and from the graphical outputs computed on whole images, it is not trivial to find a compromise between recall rates and precision at this point. By restricting the maximum probability of occurrences of symptoms, the precision improves but the recall rates drop below acceptable values. However, considering the difference in proportions of healthy leaves and infected leaves, just a few false positives on leaves impair severely the precision metric. The precision could be improved by extending the features for a more discriminative potential. It could also be considered to maximise the recall rate regardless of a precision below 50% and discard the false positives with post-processing methods relying on different properties. For instance, since most of the false positive appear when necrosis occur on stems, a specific morphological post treatment could be designed to eliminate them.

Conclusions

In this paper, a Bayesian method primarily designed for the segmentation of well-structured and healthy classes of organs has been applied to the detection of sporadic and randomly distributed anomalies and infections. The method was able to extract the local distinctive properties of the infections and differentiate them among the numerous and variable textures in the image. While being able to retrieve most of symptoms or at least parts of them, the method still needs improvements in terms of precision. Numerous anomalies are confused with infections. In order to discard these confounding factors more various and structured classes of anomalies should be modelled. It can also be considered to apply post-processing methods to eliminate the most frequent causes of false positive detection.

Acknowledgements

This work was supported and funded by the French Ministry of Agriculture student grant. Authors would like to thank the French National Institute for Agricultural Research (INRA, UMR SAVE 1065 and UMR EGFV 1287) and the French National Institute of Vines and Wine (IFV) who made available the vineyards and the farm equipment and who also provided data and expertise regarding the monitoring of agronomic parameters of the plots under study.

References

Abdelghafour F., Rosu R., Keresztes B., Germain C. & Da Costa J-.P., 2018. Joint structure and colour based parametric classification of grapevine organs from proximal images through several critical phenological stages. In Proceedings of the 14th International Conference on Precision Agriculture. https://www.ispag.org/Proceedings.

Bouman, C., & Sauer, K., 1993. A generalized Gaussian image model for edge-preserving MAP estimation. IEEE Transactions on Image Processing, 2 (3), 296-310. doi:10.1109/83.236536

Everhart, S. E., Askew, A., Seymour, L., and Scherm, H., 2013. Spatio-temporal patterns of pre-harvest brown rot epidemics within individual peach tree canopies. European Journal of Plant Pathology, 135, 499-508.

Knutsson, H., 1989. Representing local structure using tensors. In Proceedings of the 6th Scandinavian Conference of Image Analysis, Oulu, Finland: Pattern Recognition Society of Finland. pp. 244-251.

Lorenz, D.H., Eichhorn, K.W., Bleiholder, H., Klose, R., Meier, U., Weber, E., 1995. Growth Stages of the Grapevine: Phenological growth stages of the grapevine (Vitis vinifera L. ssp. vinifera) – Codes and descriptions according to the extended BBCH scale. Australian Journal of Grape and Wine Research 1 (2) 100-103

Oberti, R., Marchi, M., Tirelli, P., Calcante, A., Iriti, M., Tona, E., *et al.,* 2016. Selective spraying of grapevines for disease control using a modular agricultural robot. Biosystems Engineering, 146, 203-215.

Rosu, R., Donias, M., Bombrun, L., Said, S., Regniers, O., Da Costa, J.-P., 2017. Structure Tensor Riemannian Statistical Models for CBIR and Classification of Remote Sensing Images. IEEE Transactions on Geoscience and Remote sensing, 5 (1):248-260.

Tona, E., Calcante, A., Oberti, R., 2018. The profitability of precision spraying on specialty crops: a technical–economic analysis of protection equipment at increasing technological levels. Precision Agriculture 19, 606-629. https://doi.org/10.1007/s11119-017-9543-4

West, J. S., Bravo, C., Oberti, R., Lemaire, D., Moshou, D., and McCartney, H. A., 2003. The potential of optical canopy measurement for targeted control of field crop diseases. Annual Review of Phytopathology, 41, 593-614.

Image-based assessment of hyperspectral reflectance characteristics of plants in the field: Lessons Learned

J. Behmann[1*], D. Bohnenkamp[1] and A.-K. Mahlein[2]
[1]*INRES-Plant Diseases and Plant Protection, University of Bonn, Nussallee 9, 53115 Bonn, Germany; jbehmann@uni-bonn.de*
[2]*Institute for Sugar Beet Research (IFZ), Holtenser Landstraββe 77, 37079 Göttingen, Germany*

Abstract

Hyperspectral sensors are of interest for applications within the agricultural sector. For measuring reflectance characteristics of a crop stand is aimed to be measured, hyperspectral cameras are often applied as the high variability within the crop stand can only be captured using sensors with high spatial resolution under a suitable significant measurement distance. However, hyperspectral cameras provide only reproducible results if normalization, measurement principle and experimental layout are coordinated. In this article, a hyperspectral measurement platform for field experiments is introduced and particular problems with the resulting adaptations are considered.

Keywords: measurement platform, hyperspectral camera, spectrometer, data processing, sensor dynamics

Introduction

Various studies have shown that hyperspectral data can be used to detect and differentiate symptoms of various fungal diseases (Mahlein *et al.*, 2012; Lowe *et al.*, 2017). These studies have usually been conducted under well-controlled conditions in laboratories to exclude potentially interfering factors (Kuska *et al.*, 2015, Thomas *et al.*, 2017). However, for practical agriculture, the field scale is the most relevant application, as only at this scale a reliable detection and identification of disease symptoms can make a valuable contribution to a more targeted and thus more efficient use of plant protection products.
Collecting hyperspectral imaging data in the field makes high demands on the measurement platform, since control, data administration and power supply must be combined depending on the sensors used. The objectives of the article are (1) to present a ground-based measurement platform to perform hyperspectral imaging in the field; (2) to highlight critical points and how they can impair the measurement quality; and (3) to outline solutions and resulting changes of the platform.

Measurement platform and used sensors

The measurement platform based on a square steel construction with four wheels has been designed to be applied on 3 m wide experimental plots (Figure 1). As an autonomous unit, it combines sensors for reflectance characteristics as well as localization systems, power supply and control of the sensors from a control laptop. All sensors are variable in height by an aluminium profile construction so that the sensor platform can be adapted to the growth stages of the plants and enables a constant distance between sensors and crop canopy. With a weight of around 150 kg, the construction approaches the limit of platforms without steering, which can still be moved by two people. On the other hand, implementing moving axes or motorization would have made the structure even more complex and heavier.
Power is provided by two lead-gel batteries (each 22 Ah) with 12 V car outlets. For the linear drive a 24 V supply is provided by transformer. Three laptops controlled the sensors and stored the data. By

DOI 10.3920/978-90-8686-888-9_24, © Wageningen Academic Publishers 2019



1 RTK-GPS (Leica GPS 1200)
2 Control laptop with sun protection
3 RGB-camera (Canon 6D)
4 FODIS (Specim Ltd.)
5 ASD Fieldspec 4
6 ASD-optics (ASD Fieldspec 4)
7 ASD-reflectance standard 99%
8 Reflectance standard of HS-cameras 50%
9 Specim V10E
10 Rikola hyperspectral camera
11 Linear Stage for Specim V10 movement

Figure 1. Composition of the measurement platform with annotation of the different sensors and components.

combining fully automatic measurements and the possibility to adjust measurement parameters via local networks, a flexible reaction to the changing lighting situation due to cloud cover is possible. As a primary hyperspectral sensor Rikola Hyperspectral Camera (Senop, Oulu, Finland) and Specim V10E line camera (Specim, Spectral Imaging, Oulu, Finland) were used. The Rikola camera is a spectral scanner. Specific phenomena and dynamic scenarios arise as the camera records single bands in rapid succession one after the other (Honkavaraa *et al.*, 2017). The motion required for the Specim cameras was realized by a linear axis. The V10e was equipped with a "fiber optic diffuse irradiance sensor" (FODIS) allowing to illuminate a part of the sensor via a fibre optic cable with ambient light and, by doing this, to capture the incoming light situation (Homolova *et al.*, 2009). In addition, a spectrometer ASD FieldSpec 4 (Malvern Panalytical, Malvern, UK) observing the light intensity in the spectral range 350-2,500 nm was installed to measure ambient light by continuously observing a reflectance standard. To enhance the orientation within the spectral images a high resolution RGB camera Canon 6D (Canon Inc., Tokyo, Japan) with a 20 MP CMOS sensor was used. The position was continuously determined by a differential GNSS system Leica 1200 (Leica Geosystems AG, Heerbrugg, Switzerland) with an accuracy of 20 to 50 mm using a radio-connected reference station at a distance of approximately 1 km. It enabled later synchronization with the data of other optical measurements via a time stamp. All measurements were carried out automatically or were triggered via the control computer.

Measuring procedure and spectral normalization

In the vegetation season 2016/2017 the platform measured a trial with 32 winter wheat plots, each 3×7 m. The plot trial was designed with an east-west orientation, which was followed for the first three measurements. Later measurements were oriented north-south. This adaptation was required by the necessity of measurements at midday and sun orientation. At this time of day drop shadows of the instruments and the carrier platform fell into the measurement field and particularly on the reflectance standard. After changing the orientation, this problem was solved but the work load moving the platform increased distinctly due to the lack of driving lanes.

The measuring sequence was standardized for all plots. The measurement platform was positioned at a defined position and the measurement were conducted without moving it. First, the Rikola camera was positioned by the linear stage at a defined point and the measurement were triggered. Then the measurement of the Specim camera was conducted by collecting image frames while moving the camera at a defined speed of 12.5 mm/s from a start to an end point by the linear stage. The frame rate has been set to 10 frames/s allowing integration times of up to 30 ms. Simultaneously the synchronously moving RGB camera took five consecutive shots at 15 second intervals. The GPS sensor and the ASD FieldSpec 4 were continuously logging data at 1 and 15 s interval, respectively. Subsequently, the hyperspectral images were normalized. The normalization describes the calculation of the reflectance R_{img} out of the measured light intensity I_{img} per band. In addition, the dark current DC of the specific sensor as well as the irradiation intensity of the sun I_{ref} must be recorded. The different possibilities to determine I_{ref} are outlined in the following. According to Geladi (2007), the reflectance is calculated by Equation 1.

$$R_{img} = \frac{I_{img} - DC}{I_{ref} - DC} \quad (1)$$

To determine the required irradiation intensity $I_{ref,}$ three independent systems were used to allow a mutual validation: (1) an ASD FieldSpec 4 recording continuously the light intensity reflected by a spectral reference panel with a known reflectance of approximately 99%; (2) the FODIS of Specim V10E detects the sunlight and directs it to the sensor of Specim V10E via a fibre optic cable (Homolova *et al., et al.*, 2009). There, a part of the sensor (approx. 25 pixel) is illuminated parallel to the measurements and this signal is stored together with the hyperspectral images. In addition, (3) in the hyperspectral cameras field of view a 50% reflectance standard was installed. Because a reflectance of 50% is comparable to the reflectance of healthy leaves in near-infrared, no adjustment of the exposure time is necessary. The exposure time can be chosen depending on the demands of the crop characteristics and thus optimally exploit the sensor dynamics.

The results showed the complexity of the lighting situation. There was no agreement between two of the methods. Theoretically, the ASD Fieldspec 4 delivers the most accurate results but the characteristic sensor response curve of the different sensors (ASD Fieldspec 4, Specim V10E, Rikola camera) do not match so that unexplained reflectance values were determined during the normalization in the case of distinct brightness fluctuations. The FODIS has the great advantage that during the measurement changing illumination conditions are registered and the sensor response is identical to the Specim V10e. However, test measurements showed that non-linear oscillations of the spectrum were superimposed on the intensity signal, which were particularly noticeable when comparing FODIS measurements at different times (Figure 2). The deviations from the mean FODIS signal (each normalized to mean =1) are presented. The non-linear effects and the oscillations especially are visible, especially from 450-700 nm, which cannot be reduced to brightness fluctuations or a spectral shift of the irradiation intensity. The relative differences showed spectral patterns/oscillations that could not be explained by changes of environmental light. One possible explanation would be interference in the optical cable or at the transition to the sensor. The observation of the spectral reference panel with a known reflectance of approximately 50% in the field of view of the two cameras during the actual measurement of the cameras represents the most direct form of recording the irradiation intensity in comparison with the two other approaches. As a result, the hyperspectral images obtained were normalized using this approach.

Phenomena reducing the measurement quality

A problem that was clearly underestimated in the design phase was the wind and the accompanying continuous movement of the crop. With the growth of the plants the problem increases and at canopy closure, the stand moved almost continuously. Since both sensors are not snapshot cameras, the

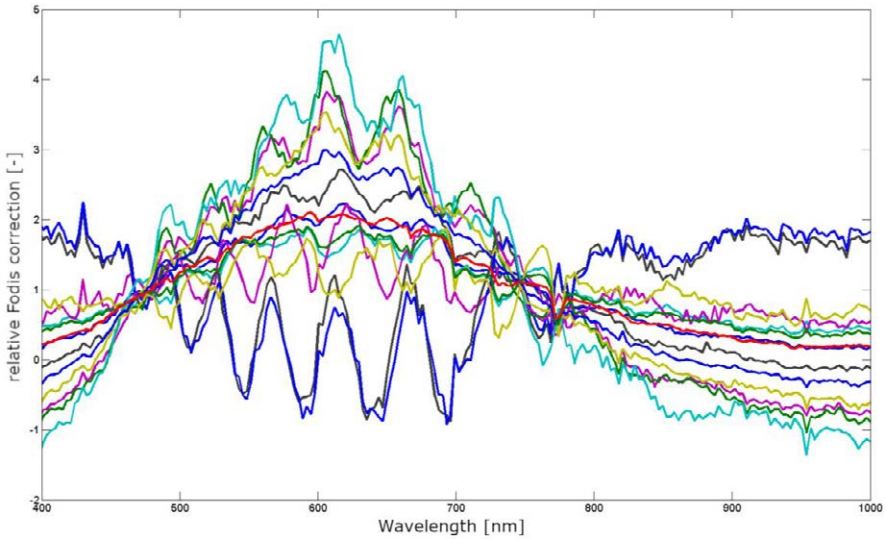


Figure 2. Single FODIS readings, captured within four hours under changing cloud conditions. The random line colouring has been selected to enhance the separability of the different plots.

single image information is recorded with a small temporal offset. With the Specim camera, the single lines are measured one after the other and the movement distorts the spatial image, whereas the spectral information remains valid. In manual mode, the Rikola camera uses a USB 2.0 connection for data transport and takes about 45 s to record the image with 120 spectral bands. If the observed leaves were moved by the wind during this measurement time, the spectrum obtained consisted of the reflectance information from different leaf areas and could be severely distorted. This effect, which could not be corrected due to the unknown 3D geometry and 3D movement with spatial resolution, prevented both manual and automatic evaluation. By averaging the entire hyperspectral image, this effect can be mostly eliminated, but then the whole spatial resolution is lost and the data that is obtained is comparable to a simple spectrometer.

Another dominant factor is in regard to the exploitation of sensor dynamics. The different leaf levels require very different exposure times. The leaf tips of the uppermost leaf level already reached saturation of the sensor at more than 1 ms exposure time, whereas the lower leaf levels did not reach saturation even at 20 ms and more due to drop shadows. Since the choice of the exposure time was based on the brightest object, the exposure time was much lower than required for the lower leaf levels and thus a very low, noisy signal was received in these regions.

Adaptations

The majority of the problems that have influenced the measurement quality can be attributed to three factors: (1) canopy movement due to wind; (2) extreme variability of reflectance intensity in the stand; and (3) insufficient registration of the irradiation intensity to normalize hyperspectral data. These factors cannot be completely eliminated, but their influence can be significantly reduced by improving the measurement platform, sensors and experimental layout.

Based on these experiences different adaptations were carried out. The measuring platform was equipped with a cotton-based diffuser, which shields the measuring range from sunlight, but allows diffuse lighting (Figure 3). This reduces the problem of intensity peaks at the leaf tips and provides

a certain degree of wind protection. In addition, the diffuse light simplifies the normalization by reflectance panels, since the measured reflectance information is less dependent on the angle of incoming light. Furthermore, the measurement mode of the Rikola hyperspectral camera has been adapted so that the hyperspectral images were stored on an internal memory card, which significantly reduced the observation time between the two adjacent bands. With a stronger north-south alignment of the experimental layout in conjunction with the diffuser, the problem of drop shadows within the stand and in particular by the measurement system itself was reduced. Figure 4 visualizes a hyperspectral image of a winter wheat canopy infested by yellow rust and obtained by the improved system platform (Figure 3) in the field experiment 2017/2018 with the adapted orientation. Different leaf layers are imaged with high spatial and spectral resolution. The spectral image quality was characterised by a low noise level except for the high noise level in the spectral border regions.

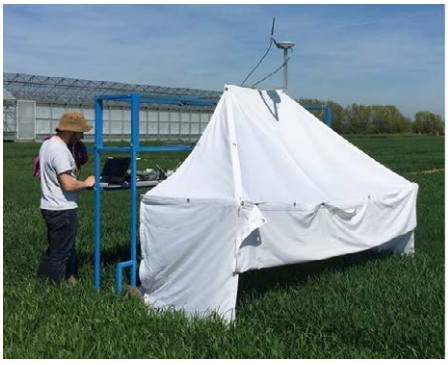


Figure 3. Improved version of the field platform for hyperspectral imaging, including a cotton based diffuser enhancing the image quality.

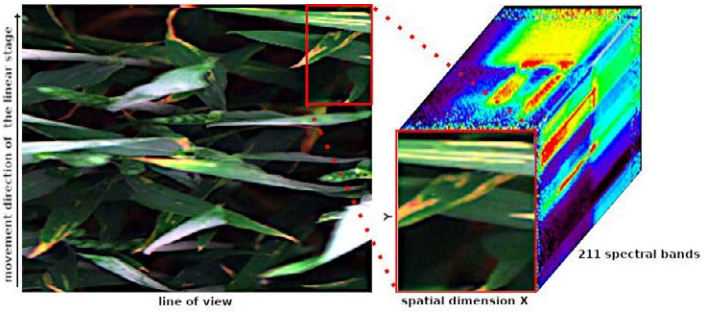


Figure 4. Visualization of hyperspectral image samples showing symptoms of yellow rust (left) obtained in the field in June 2018. The image covers approximately 200×200 mm of the wheat canopy.

Conclusions

Measuring hyperspectral characteristics in the field is a challenging task that cannot be performed by out-of-the-box sensor systems. Reliable and reproducible measurements depend on intelligent

selection of sensors, measurement platform, spectral normalization approach and experimental layout of the field trial. Multiple measurement systems have been developed worldwide but typically such a system is developed starting from scratch to meet the specific demands of the investigated crop.

Acknowledgements

This work was funded by BASF Digital Farming. We would like to thank Sabine von Tiedemann for internal editing and proofreading of the manuscript.

References

Geladi, P. L., 2007. Calibration standards and image calibration. In: Techniques and applications of hyperspectral image analysis. In: Grahn, H., & Geladi, P. (Eds.). Techniques and applications of hyperspectral image analysis. pp. 203-220, Chichester, West Sussex, UK: John Wiley & Sons.

Homolova, L., Alanko-Huotari, K., & Scheapman, M. E., 2009. Sensitivity of the ground-based downwelling irradiance recorded by the FODIS sensor in respect of different angular positions. In: IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing. WHISPERS'09, pp. 1-4.

Honkavaara, E., Rosnell, T., Oliveira, R., & Tommaselli, A., 2017. Band registration of tuneable frame format hyperspectral UAV imagers in complex scenes. ISPRS Journal of Photogrammetry and Remote Sensing, 134, 96-109.

Kuska, M., Wahabzada, M., Leucker, M., Dehne, H. W., Kersting, K., Oerke, E. C., *et al.*, 2015. Hyperspectral phenotyping on the microscopic scale: towards automated characterization of plant-pathogen interactions. Plant Methods, 11(1), 28.

Lowe, A., Harrison, N., & French, A. P., 2017. Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. Plant Methods, 13(1), 80.

Mahlein, A. K., Steiner, U., Hillnhütter, C., Dehne, H. W., & Oerke, E. C., 2012. Hyperspectral imaging for small-scale analysis of symptoms caused by different sugar beet diseases. Plant Methods, 8(1), 3.

Thomas, S., Kuska, M.T., Bohnenkamp, D., Brugger, A., Alisaac, E., Wahabzada, M., *et al.*, 2018. Benefits of hyperspectral imaging for plant disease detection and plant protection: a technical perspective. Journal of Plant Disease Protection, 125, 5-20.

Implications of dose-response relationships of herbicide droplet applications for leaf-specific weed control in leeks

N. Koukiasas[1], J. Martinez-Perez[1], R.A. Pilgrim[2], S. Sanford[2] and A.J. Murdoch[1*]
[1]*School of Agriculture, Policy and Development, University of Reading, Earley Gate, Reading, RG6 6AR, United Kingdom; a.j.murdoch@reading.ac.uk*
[2]*Concurrent Solutions LLC, 1216 Wilkins Road, Benton, KY 42025, USA*

Abstract

Weed control in UK field vegetables faces the problem of producing and maintaining high yields while herbicide active ingredients and amounts applied are reducing. These problems, which are unlikely to be solved with new herbicides or the use of GM crops, call for a paradigm shift in weed control. The concept of leaf-specific weed control was tested. Individual weed seedlings were controlled in the glasshouse by applying a single droplet of herbicide and dose-response relationships were quantified. In the field, manually-applied droplets achieved high levels of weed control in leeks, prevented yield loss and reduced herbicide inputs by up to 77%. Because of the high value of the crop and the higher yields associated with ultra-precise droplet application, it would appear to be economical to apply these droplets using a robotic weeder.

Keywords: glyphosate, glufosinate-ammonium, 2,4-D, droplet application, leaf-specific weed control

Introduction

Weeds and their control is one of the most important factors in maintaining yields and quality of field vegetable crops. Typical weed control methods for field vegetables include use of pre-planting tillage followed by pre-emergence herbicide and one or more of post-emergence selective herbicide, inter-row cultivation and hand-weeding (Slaughter *et al.*, 2008). Herbicides are an efficient and economic method of weed control. However, improper or inappropriate use of herbicides, when they are applied as an overall spray, risks adverse effects on human health and the environment. These impacts along with pressures from consumers and policy makers to lower pesticide inputs, have resulted in strict regulation including the EC Regulation No. 1107/2009, the Water Framework Directive (2000/60/EC) and the Sustainable Use Directive (2009/128/EC). As a result, several herbicides have been withdrawn and it is much more difficult for new chemicals to gain approval (Hillocks, 2012). The predicament is worse for vegetable growers because they rely on limited and old herbicides that require a lot of funding and effort in order to keep them in the market (Fennimore *et al.*, 2014).
Weed control of vegetables, therefore, needs a paradigm shift from conventional spraying. Efficacy needs to be maintained or improved while reducing herbicide usage. Genetically-modified, herbicide-tolerant crops (GM-HT) may offer an alternative, but whole fields still need to be sprayed. Moreover, growing GM-HT crops is not currently permitted in the UK and, even if it were, seed companies would focus on broad-acre crops rather than vegetables because of the relatively small market for vegetable seeds. Other approaches include site-specific weed control e.g. by spraying weed patches or, plant-specific weed control e.g. by identifying individual weed plants in the field and then spot-spraying them using precisely targeted applications (Christensen *et al.*, 2009). Drop-On-Demand (DOD) application to individual weed plants is close to the leaf-specific approach to weed control described here. Although no DOD systems are commercially available, efficacy of weed control by DOD treatment with glyphosate in research trials in fields of carrots, was 73%-95% (Utstumo *et al.*, 2018).

DOI 10.3920/978-90-8686-888-9_25, © Wageningen Academic Publishers 2019

The study presented is part of a project which explores an engineering solution to weed control as an alternative to spraying fields with selective herbicides, spot-spraying or using GM crops. The system, which is under development by Reading University, UK and Concurrent Solutions LLC, USA, is designed to apply single herbicide droplets (1-2 µl) at an appropriate concentration to the leaves of unwanted plants (Pilgrim and Murdoch, 2008). The concept is that a machine vision system locates individual green leaves, classifies each leaf as being crop or weed. Individual weed leaves are then targeted and treated with a single droplet of a broad-spectrum, translocated herbicide (e.g. glyphosate). No herbicide should, therefore, directly hit the soil and none should be applied to the crop. The use of much larger droplets than for conventional spraying avoids the risk of drift and appropriate formulation should minimise spatter.

In this paper, dose-response studies are reported in which doses relative to commercial rates were applied as single droplets to weed leaves. For proof of concept in the field, herbicide droplets were applied manually to weed leaves in a field growing leeks. The hypothesis tested is that the recommended rate of herbicide applied as a single droplet will reduce weed biomass by over 90%. In the field, the hypotheses tested were that (1) the efficacy of leaf-specific weed control by herbicide droplets is not significantly lower than that achieved from the hand-weeding or current herbicide spraying methods; (2) crop yield is not significantly less than for hand weeding; (3) leaf-specific weed control reduces the amount of herbicide applied by around 80%; and (4) a sequence of droplet applications is required over the growing season to control late emerging weeds to ensure that the crop remains weed-free during the critical period for weed control.

Materials and methods

Glasshouse trials

Dose-response trials were carried out under glasshouse conditions at the facilities of the Crop and Environment Laboratory at Reading University during summer 2018. Droplets containing glyphosate, glufosinate-ammonium or 2,4-D were applied to seedlings of *Chenopodium album*, *Senecio vulgaris* and *Poa annua*. In addition to the herbicides described in Table 1, a combined treatment of 2,4-D and glufosinate-ammonium was applied as two droplets on the same leaf. For the other herbicides, one droplet (1-2 µl) was applied to a single leaf except in the case of *P. annua* where droplets were applied at the growing point of the weed. The volume of droplets and amount of herbicide was calculated according to the ground cover of the seedlings and the minimum label recommendation for spray application. The adjuvant AS 500 SL (Z.P.H Agromix, Niepołomice, Poland), was included in all treatments at 1% concentration. All trials were randomized complete blocks with ten replicates and herbicide doses ranged from 1/128 to 4-times the recommended rates. Application of the three control treatments included the use of droplets containing water only,

Table 1. Details of herbicides for the dose-response trials. The 'recommended' dose is for a targeted droplet application to a *Chenopodium album* seedling with 721 mm^2 ground cover.

Active ingredient (a.i.)	Product	A.i. concentration in product (g/l)	Recommended product dose (l/ha)	'Recommended' dose (µg/a.i. seedling)
Glyphosate	Roundup® Biactive GL	360	1.5	38.9
Glufosinate-ammonium	Harvest®	150	3	32.5
2,4-D	Depitox®	500	1.4	50.5
2,4-D & glyphosate	Kyleo®	160 & 240	3	86.5

water with adjuvant (1%) and undiluted product. Deionised water was used to prepare all solutions. Serial dilutions were prepared so that the correct amount of herbicide per seedling could be applied. Seedling mean ground covers when treated were 721 mm^2 (4-6 leaves), 223 mm^2 (2-3 leaves) and 1,054 mm^2 (2-3 tillers visible) for *C. album*, *S. vulgaris* and *P. annua*, respectively. Fresh and dry weights of the weed seedlings were estimated three weeks after droplet application. In order to estimate dry weights, fresh seedlings were oven-dried for 48 h at 80 °C. Biomass data were fitted to the four-parameter log-logistic model which can be specified as a non-linear function given by:

$$y = \frac{c + (d - c)}{[1 + \exp[b\ (logx - loge)]]} \quad (1)$$

where y is the biomass, b is the relative slope around parameter e, e is the ED_{50} which corresponds to the dose causing 50% reduction of biomass. c and d are the lower and upper limits of the curve respectively. Analyses were carried out using the open source statistical software R, version 3.3.2 and the add-on package 'drc' (Ritz *et al.*, 2015).

Field trials

Field experiments were carried out with leeks during spring/summer 2017 and 2018 at Sonning Farm near Reading, UK. Leek seedlings were transplanted to the field at the 3-leaf stage with 0.4 m between rows and 0.2 m between plants. Fertilizer was applied on the day of transplanting using sulfur (SO_3) and nitrogen (N) at the rates of 50 kg/ha and 100 kg/ha respectively. An automated drip irrigation system was used in order to water plants individually. The natural weed infestation comprised *Senecio vulgaris*, *Matricaria recutita*, *Chenopodium album*, *Poa annua*, *Capsella bursa-pastoris* and *Solanum nigrum*. The experimental design was a randomised complete block with three replications of nine treatments for the 2017 trial and ten treatments in 2018. In addition to weedy (unweeded) and weed-free (hand-weeded) control plots, other controls included a broadcast spray of a pre-emergence herbicide (Stomp Aqua®, 455 g/l, pendimethalin) at a rate of 2.9 l/ha or a post-emergence herbicide (Buctril®, 255 g/l, bromoxynil) at 1.5 l/ha. The pre-emergence treatment was applied five days before planting and the post-emergence four weeks after planting the crop. In 2018, a further treatment received both pre- and post-emergence herbicides. Droplets containing 9 µg of glyphosate or 7.5 µg of glufosinate-ammonium were applied manually to every weed leaf bigger than 100 mm^2, using a pipette. Droplet applications commenced 2 and 3 weeks after planting the leeks for the 2017 and 2018 trials, respectively, with the last application taking place 12 weeks after planting. Droplet treatments were applied either 5 or 10 times over the 12-week period. Glufosinate-ammonium droplets were only applied weekly because of the herbicide's limited translocation. Leeks and weeds were harvested from the treated central area of the plots 14 weeks after planting. Leeks were trimmed to a height of 0.34 m and then weighed as for commercial sale. Weed dry weights (above ground fresh biomass only) was determined as above. Yield and weed control data were analysed by ANOVA using GenStat 17th Edition.

Preliminary economic analysis was carried out in which the value of the crop and the cost of weed control (material + application cost) were estimated for each herbicide treatment for both years. The economic value of the crop was based on the average wholesale market prices for 2018 (Brigham, 2018). Information regarding the material cost of herbicides came from personal communications with industry representatives. Application cost for spraying herbicides was assumed to be £12.50/ha (Redman, 2018). The gross margin was estimated by subtracting the weed control cost from the crop value. Further assumptions were: (1) the cost of an applicator platform including a maintenance contract would be £25,000 over five years (S. Sanford, personal communication, December 2018); (2) an area of 4 ha could be treated per day and would need to be retreated every 14 days; (3) a platform could operate for 10 days every fortnight and could, therefore, treat a maximum of 40 ha over 14 days, and, to build in a safety margin; (4) the platform is only used for 40 ha per year. The platform cost per hectare was, therefore, assumed to be £ 125/ha/year (i.e. £5,000/40 ha).

Results

In the glasshouse, *C. album* seedlings were effectively controlled when the recommended rates of the herbicides used were applied as a single droplet to a single leaf, producing increasingly phytotoxic effects with increase of dose (Figure 1). These effects were quantified by fitting Equation 1 to the dry weight data to produce dose-response curves as exemplified for *C. album* (Figure 2). Interestingly, the same relative levels of efficacy were observed for both glyphosate and glufosinate-ammonium, the ED_{50} values for each herbicide corresponding to $1/6^{th}$ of their respective recommended doses (Figure 1B, C). A dose of 36.6 µg (±19.9, SE) per seedling was needed to achieve a 90% level of control (ED_{90}) for both herbicides (Figure 2). A slightly lower ED_{50} was observed when the mixture of 2,4-D and glyphosate was applied in comparison compared to the application of glyphosate on its own (Figure 1D). Glyphosate was more potent than the other herbicides when applied to seedlings of *Senecio vulgaris* (ground cover: 223 mm^2) with an ED_{90} of 0.24 µg (±0.04, SE). The second-best efficacy was achieved from the mixture of 2,4-D and glyphosate with a dose of only 1.33 µg (±0.30, SE; $1/20^{th}$ of the recommended dose) being required to achieve 90% control of *S. vulgaris*. The larger *P. annua* seedlings needed higher doses than *S. vulgaris* and *C. album* to achieve the same efficacy. Efficacies of glyphosate and glufosinate-ammonium were also the same for *P. annua*, a dose of 79.9 µg (±15.5, SE) being needed to reduce weed biomass by 50%.

In the field in both years, the treatment where droplets containing 9 µg of glyphosate were applied approx. weekly (i.e. 10 times over 12 weeks) reduced weed dry biomass by 99%. The same levels of weed control were achieved when this treatment was applied fortnightly (i.e. only 5 times) in 2018. Yields achieved from the weekly glyphosate droplet treatment were the highest in the plots treated with herbicides (droplets or overall spray) and were not significantly lower than the yields from the hand-weeded controls. However, yields for the fortnightly application of glyphosate droplets were significantly lower than from the weed-free control (63% of weed-free).

The weekly glufosinate-ammonium droplet treatment reduced weed biomass by 91% or more in both years. Trimmed leek yields from these glufosinate-ammonium treated plots were not significantly lower than the 39 t/ha in the hand-weeded control in 2018. In 2017, however, yields in this treatment were only 65% of the 42.2 t/ha hand-weeded control yield. A pre-emergence spray of pendimethalin reduced weed dry biomass by 42 and 47% in the 2017 and 2018 trials, respectively and crop yields were significantly lower in both years. When both pre and post-emergence herbicides were applied in 2018, the efficacy of weed control was still only 64%.

Regarding the amounts of herbicide applied over the whole growing season, the weekly glyphosate droplet treatment was the equivalent of 340 g and 299 g glyphosate/ha in 2017 and 2018, respectively, approximately 75% less than spraying with pendimethalin (1,320 g/ha). The fortnightly application of glyphosate droplets in 2018 reduced the herbicide inputs by 71 and 77% relative to pre-emergence and the pre+post-emergence sprays, respectively. Weekly application of glufosinate-ammonium droplets reduced total herbicide usage by 51 and 68% in 2017 and 2018 trials, respectively.

The preliminary economic analysis suggested that the droplet treatments could be highly profitable. The average gross margin of the fortnightly droplet treatments was estimated at £30,246/ha, much more than the combined pre+post-emergence treatment (£18,502/ha). Allowing for the platform cost of £125/ha/y for the platform, then the estimated additional profit of the droplet treatments above that for the combined pre- and post-emergence treatment would be £11,619/ha for the 2018 trial.

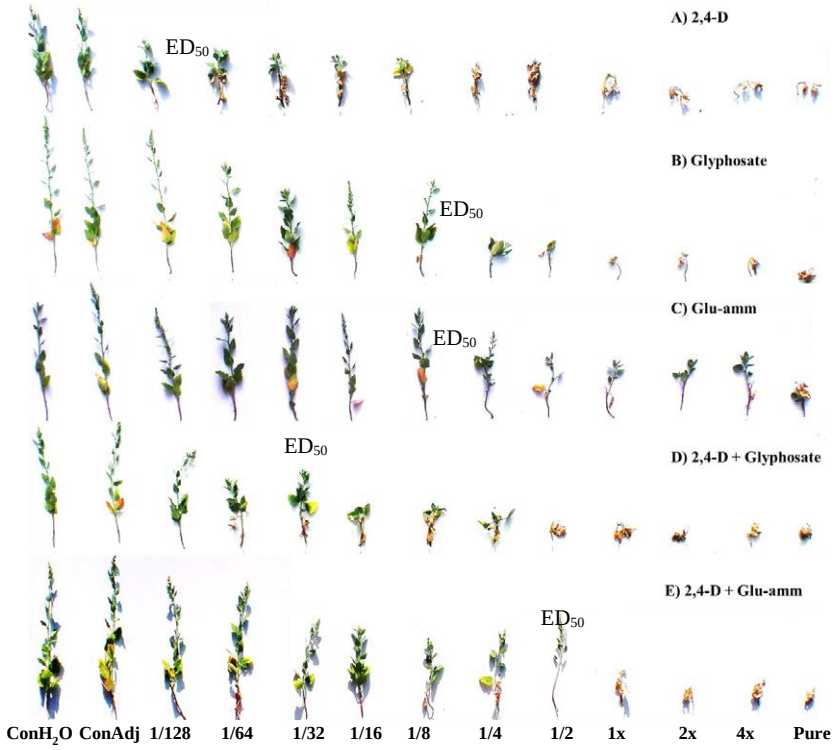


Figure 1. *Chenopodium album* seedlings three weeks after applying droplets containing different concentrations of (A) 2,4-D, (B) glyphosate, (C) glufosinate-ammonium, (D) 2,4-D plus glyphosate and (E) 2,4-D with glufosinate-ammonium. Doses of a.i. are expressed relative to the recommended (1x) rate (Table 1). In (E), separate droplets of the two products were applied to the same leaf. For the controls, droplets of deionized water ($ConH_2O$), 1% adjuvant (ConAdj) or undiluted herbicide (Pure) were applied. Concentrations giving 50% reduction in dry biomass (ED_{50}) are shown.

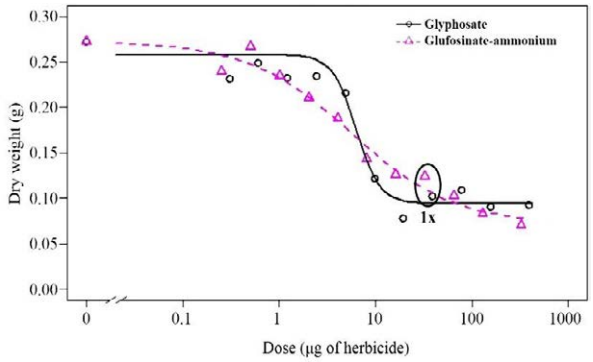


Figure 2. Dry weight of *Chenopodium album* seedlings as a function of herbicide dose, applied as one droplet per seedling. Recommended rates (1×) were 38.9 µg for glyphosate and 32.5 µg for glufosinate-ammonium. Parameters of lines fitted by Equation 1 are in Table 2.

Table 2. Parameters of herbicide dose-response curves (Equation 1) for *Chenopodium album* in Figure 2.

Herbicide active	Parameter (±SE)			
	b (slope)	c (lower limit)	d (upper limit)	e (ED_{50})
Glyphosate	3.54 (±1.59)	0.09 (±0.01)	0.26 (±0.01)	6.29 (±0.85)
Glufosinate-ammonium	0.81 (±0.23)	0.07 (±0.02)	0.27 (±0.01)	5.66 (±2.41)

Discussion

In this study, the concept of weed control by applying a single droplet of herbicide to weed leaves was proven in the glasshouse and in the field in a leek crop. Dose-response studies have quantified the amount of herbicides needed to control *C. album*, *S. vulgaris* and *P. annua* seedlings. The relative efficacy of control of *C. album* (4-6 leaves) when droplets of either glyphosate or glufosinate-ammonium were applied was the same, the ED_{50} for both being about $1/6^{th}$ of the recommended dose, equivalent to 90 and 75 g a.i./ha, respectively. In glasshouse spraying trials with *C. album* plants about twice the size of those here (i.e. 8-10 leaves), Tharp *et al.* (1999) found the same ED_{50} for glyphosate and glufosinate although slightly over twice as much was needed (200 g/ha). As regards, the amount required for each plant for droplet applications, Mathiassen *et al.* (2016) reported that 3.7 µg of glyphosate per seedling reduced the fresh weight of *C. album* seedlings (2-leaf stage) by 50%, which compares well with the 5.1 µg here at the ED_{50}.The efficacy of weed control reported here is, therefore, at least as good as that expected if the same doses were applied using a conventional spray. Moreover, a major improvement in application technology is that the amount of herbicide applied to each plant is known whereas, when spraying, only the average per unit land area is known. In the field, the crop must be kept weed-free during the critical period for weed control, which, for transplanted leeks, is from 1 to 12 weeks after planting (Tursun *et al.*, 2007). Not surprisingly, therefore, a sequence of droplet applications was needed to keep the crop free of weeds and achieve yields that did not differ significantly from the hand-weeded controls. Multiple treatments are desirable to control weeds which emerge later in the season, but especially for any larger weeds which might not be controlled satisfactorily by droplets. Moreover, since one aim is to avoid applying herbicide to soil, seedlings too small to target accurately (say <100 mm^2) could be treated on a subsequent visit to the field. In this study, it was found that fortnightly applications during the critical period achieved high levels of weed control provided a systemic herbicide like glyphosate is used and if one droplet is applied to each leaf over 100 mm^2. This overall concept has also been shown to be successful in fields with cabbages (Murdoch *et al.*, 2017) although because cabbage is a more weed suppressive crop without a clear critical period only three treatments were needed to achieve high yields and approximately 95% control of weeds.

The reductions in herbicides used per unit area are an important environmental consideration and the droplet treatments reduced the a.i. required by 68-82%/ha compared with pre-emergence and combined pre- and post-emergence sprays. Although much lower amounts of herbicides were applied per hectare, those applied to individual weeds were within the range of label recommendations. Concerns about individual plants surviving low doses and herbicide resistance are not, therefore, relevant.

The preliminary economic analysis calculated the net profitability of an automated droplet applicator platform for weed control in leeks. The high profitability (£11,619/ha) compared with spray

applications using commercial herbicides derives from the high value of the crop. A single value of the crop has been used (£1.22/kg). However, leek prices change seasonally and are categorized according to stem diameter and then sold as pre-packed, loose product or discarded as non-marketable. Finally, development of software and hardware (including a prototype applicator) have been in progress for several years and are funded by AHDB (Murdoch, 2018).

Conclusions

Dose-response studies in the glasshouse proved that single droplet application on to small seedlings was a highly efficient method of weed control. Droplets containing 36.6 µg of glyphosate or glufosinate-ammonium controlled *C. album* seedlings. In leek crops, a sequence of applications of glyphosate or glufosinate-ammonium achieved high levels of weed control, no yield penalty and reduced applications by over 70% compared to conventional pre- and post-emergence sprays. Economic analysis demonstrated that because of the high crop value, applications could increase profits by over £11,000/ha per year, despite £5,000 per year for the application system cost.

Acknowledgements

Our thanks to our co-funders (AHDB Horticulture, the Douglas Bomford Trust, University of Reading, Edith Mary Gayton Trust), the technical staff at the Crops Research Unit in Sonning Farm for help setting up field trials and Victoria Wickens, Jennifer Wickens and Shwan Hama Rash for help with droplet applications. We are also grateful to Hammond Produce Ltd. for supplying leeks for transplanting, Herbiseed Ltd. for providing the weed seeds, Dr Zenon Woznica for providing a sample of the adjuvant AS 500 SL and Bayer Crop-Science Ltd. for supplying a sample of Harvest®.

References

Brigham, J., 2018. Wholesale fruit and vegetable prices. Dept of Environment, Food & Rural Affairs, UK. https://www.gov.uk/government/statistical-data-sets/wholesale-fruit-and-vegetable-prices-weekly-average (last accessed October 2018).

Christensen, S., Søgaard, H. T., Kudsk, P., Nørremark, M., Lund, I., Nadimi, E. S. *et al.,* 2009. Site-specific weed control technologies. Weed Research 49 (3) 233-241.

Fennimore, S. A., Hanson, B. D., Sosnoskie, L. M., Samtani, J. B., Datta, A., Knezevic, S. Z. *et al.,* 2014. Field Applications of Automated Weed Control: Western Hemisphere. Dordrecht, the Netherlands: Springer. pp 151-169.

Hillocks, R., 2012. Farming with fewer pesticides: EU pesticide review and resulting challenges for UK agriculture. Crop Protection 31 (1) 85-93.

Mathiassen, S. K., Kudsk, P. and Lund, I., 2016. Adjuvants for single droplet application of glyphosate. In: Proceedings of 11th International Symposium on Adjuvants for Agrochemicals, edited by ISAA Society, pp. 131-137.

Murdoch, A., 2018. 'eyeSpot' – Leaf specific herbicide applicator for weed control in field vegetables. Annual Report of Project CP 134., Stoneleigh, UK: AHDB.

Murdoch, A., Koukiasas, N., De La Warr, P. and Price-Jones, F., 2017. Precision targeting of herbicide droplets potentially reduces herbicide inputs by at least 90%. Aspects of Applied Biology 135 39-44.

Pilgrim, R. and Murdoch, A. 2008. A weed growth-rate model to support plant-specific agriculture. Aspects of Applied Biology 84 309-314.

Redman, G., 2018. The John Nix Pocketbook for Farm Management 2018. 48th ed. Melton Mowbray, UK: Agro Business Consultants

Ritz, C., Baty, F., Streibig, J. C. and Gerhard, D., 2015. Dose-Response Analysis Using R. PLOS ONE 10(12) e0146021.

Slaughter, D. C., Giles, D. K. and Downey, D., 2008. Autonomous robotic weed control systems: A review. Computers and Electronics in Agriculture 61(1) 63-78.

Tharp, B. E., Schabenberger, O. and Kells, J. J., 1999. Response of annual weed species to glufosinate and glyphosate. Weed Technology 13 (3) 542-547.

Tursun, N., Bükün, B., Karacan, S. C., Ngouajio, M. and Mennan, H., 2007. Critical period for weed control in leek (*Allium porrum* L.). HortScience 42 (1) 106-109.

Utstumo, T., Urdal, F., Brevik, A., Dørum, J., Netland, J., Overskeid, Ø., *et al.,* 2018. Robotic in-row weed control in vegetables. Computers and Electronics in Agriculture 154 36-45.

Exploring the factors affecting spatio-temporal variation in grapevine powdery mildew

R. Melyon-Delage[1], B. Bois[2], S. Zito[2], M. Rega[2], G. Garin[1*] and A. Caffarra[1]
[1]*SAS iTK, Cap Alpha, Avenue de l'Europe, 34830 Clapiers, France; guillaume.garin@itkweb.com*
[2]*CRC, Biogéoscience, Université de Bourgogne Franche Comté, 2 Bld Gabriel, 21000 Dijon, France*

Abstract

The aim of the project was to explore the role of environmental/agronomic factors involved in powdery mildew pressure variation in time and at the plot and regional scale. To do so, a series of variables were monitored in 10 different vineyards, in the Beaune region (Bourgogne) of France, over two years. These factors included agronomic descriptors and weather variables. Weather data were acquired at the plot scale. Disease symptoms were observed weekly on leaves and grapes, highlighting inter-plot variation in disease indicators. The factors that most impacted this variability were relative humidity, rain, leaf wetness, vigor and phenology. A more in-depth study of the interactions between these factors will help to disentangle the complex effects of the environment on powdery mildew inter-plot heterogeneity.

Keywords: correlation, heterogeneity, humidity, leaf wetness, local effects

Introduction

Powdery mildew (caused by *Erysiphe necator*) is one of the most common and harmful diseases of grapevine. Each year, this polycyclic fungus goes through sexual and asexual reproductive cycles starting in spring (primary infection). It is only after a number of cycles after the primary infection that the first symptoms are noticeable and it is often too late to effectively treat the fungal attack (Wilcox, 2003; Calonnec *et al.*, 2008).

If its optimal conditions are met, a new generation can be completed in less than a week (Wilcox, 2003). These explosive germinations combined with the fact that symptoms are visible only once the disease is well established, make its control difficult (Pearson and Gadoury, 1987). In addition, the spatial distribution of the disease is often heterogeneous between neighboring plots, with higher disease pressure in certain places that can be considered as disease 'hotspots' (Calonnec *et al.*, 2009; Mammeri *et al.*, 2014). The position of hotspots can vary over the years, even if some plots consistently present a higher vulnerability over time.

This spatio-temporal heterogeneity makes it difficult to obtain accurate prediction by epidemiological models that are fed by meteorological variables coming from weather stations that are not in close proximity to the vineyards or are spatialized over large cell grids. Indeed, local disease pressure might be related to micro-climatic factors but could also be the result of local environmental/agronomic variables that are rarely considered by models (Delp, 1954; Sall, 1980; Caffi *et al.*, 2011), such as the type of soil, agricultural practices and the interaction with the plant: e.g. the growth, the physiology and genetic specificities of the variety.

To date, the effects of local factors on disease pressure have not been analyzed in depth, and the triggers of powdery mildew hotspots have not been clearly identified. However, an exploration of these effects is of the foremost importance to improve the understanding of the disease and to support the development of predictive models of powdery mildew. Such models could play an important role in the optimization of phytosanitary management. Therefore, the general aim of this study was to explore the role of the factors involved in powdery mildew pressure variation in time and space

DOI 10.3920/978-90-8686-888-9_26, © Wageningen Academic Publishers 2019

at a local scale. To do so, a series of environmental and agronomic variables in 10 different vineyard plots near Beaune (Bourgogne) were monitored over the growing seasons of 2017 and 2018.

Methods

The project focused on powdery mildew dynamics in Burgundy, France over two years (2017-2018). Each year, a phase of field observations was followed by a phase of data analysis. Statistical correlation and multivariate analyses were conducted to explore the relationships between environmental variables and disease pressure. This paper presents the description of the data collected in 2018 (for more details on 2017, see Weens, 2017) and the multivariate and correlation analyses conducted on the joined 2017-2018 datasets.

Study sites

Observations were conducted on 10 plots in the Côte d'Or, all located in the same area near Beaune so that local effects could become evident. The plots were not treated to avoid the confounding effects of pesticides on disease pressure. Each plot was characterized with its soil type, its variety, its training method, its cover crop, whether it was in the plain or on the hills and its orientation (plot name codes and description is given in Table 1).

Observations

Disease pressure monitoring was carried out in the field through visual notations from April to July. At the same time, meteorological factors and plant-related factors such as vigor and phenological stages were measured (Table 2). These variables were suspected to contribute to disease heterogeneity, as suggested by previous literature. They can be divided into:

1. Meteorological variables, either measured at reference stations at 2 m above soil on each plot, and also in the canopy (micro-climate) on three specific plots (CRC_20, CRC_21, CRC_28). They included, temperature, relative humidity, rain, leaf wetness (only for the three micro-climate stations in the canopy).
2. Agronomic (plant and 'situation') variables: phenology, plant vigor (measured as weekly growth in twig length and width), variety, training method, cover crop, plain or hills, orientation year (2017 or 2018).
3. Disease variables: percentage of leaves and grapes affected by powdery mildew.

Statistical analysis

The disease progression curves on grapes and leaves of 2018 were first analyzed visually to describe the symptom dynamics over time. Secondly, correlation and multivariate analyses were performed. The data was used to calculate indicators (see below) and organized to compute correlation matrices. The correlation coefficients (Pearson) obtained in this way were used to select a set of 'meaningful indicators' highly correlated to disease occurrence that were then used in a principal component analysis (PCA). In order to carry out these analyses, static descriptive indicators of the plots (such as grape variety) or synthesis indicators for dynamic variable such as temperature were used.
The indicators in the statistical analysis were the following:

1. Meteorological indicators: Average daily temperature, minimum and maximum temperature, relative humidity and leaf wetness. These variables were averaged either over the month of May (typical time of the year for primary contaminations) or weekly with a specific focus on the three weeks before the week of the first symptoms: 15th-21st of May (week 1), 22nd-28th of May (week 2) and 29th of May-4th of June (week 3).

Table 1. Plot characteristics.

Plot name	Situation	Altitude (m)	Orientation	Variety	Pruning	Soil	Cover crop	Year
CRC_07	Plain	224	North-East	Chardonnay	Cordon Simple	Clay Loam	Yes	2018
CRC_08	Plain	275	East	Pinot Noir	Guyot	Clay Loam	No	2017, 2018
CRC_20	Hill	326	East	Chardonnay	Cordon Simple	Clay Loam	No	2017, 2018
CRC_21	Hill	334	South-West	Chardonnay	Cordon Simple	Clay Loam	No	2018
CRC_22	Hill	309	South	Chardonnay	Cordon Simple	Clay Loam	No	2017, 2018
CRC_23	Hill	279	North-East	Chardonnay	Cordon Simple	Clay Loam	Yes	2017, 2018
CRC_25	Hill	320	East	Chardonnay	Cordon Simple	Clay Loam	No	2017, 2018
CRC_26	Hill	362	West	Pinot Noir	Cordon Simple	Clay Loam	No	2017, 2018
CRC_27	Hill	267	South-East	Pinot Noir	Cordon Simple	Clay Loam	No	2017, 2018
CRC_28	Plain	232	East	Chardonnay	Cordon Simple	Clay Loam	Yes	2017, 2018

Table 2. Observed variables, frequency and method of observation.

Measured variable	Unit	Frequency	Device / protocol
Temperature	°C	hourly	HOBO MX2302 External Temperature/RH Sensor Data Logger (Onset®)
Rain	mm	hourly	Rain gauge: RainWise (RainWise Inc.) with HOBO Pendant® Event Data Logger (Onset®)
Relative humidity	%	hourly	HOBO MX2302 External Temperature/RH Sensor Data Logger (Onset®)
Leaf wetness	%	hourly	Microstation HOBO The Leaf Wetness Smart Sensor (Onset®)
Vigor, length	cm	weekly	Measuring tape
Vigor, thickness	cm	weekly	Caliper
Phenological stage	BBCH	weekly	BBCH scale
Disease	%	weekly	Visual observation

2. Agronomic indicators:
   - plant related: plant vigor (average weekly growth), date of phenological phases;
   - Disease related: first date of symptom onset, final percentage of affected leaves or grapes (on the 4th of July);
   - plot related: variety, training method, cover crop, plain or hills, orientation;
   - year.

Results

Disease dynamics

Unlike 2017, Year 2018 was characterized by a moderate powdery mildew pressure for Burgundy. The disease developed relatively late (first symptoms between the end of May and the beginning of June over the region). For the plots, disease pressure ranged between weak and moderate/strong with a maximum of 30% symptoms in July before the beginning of phytosanitary treatments (Figure 1). The date of first observable symptoms was clearly heterogeneous, varying between the 4 and 23 of June (CRC_22 vs CRC_28 on grapes). Dynamics of symptom progression also varied, with some plots showing a faster rate of disease increase than others. Interestingly, whereas final disease pressure was higher in 2018 than 2017, the date of first symptoms (first week of June) was similar for the two years under study.

Correlations

Correlation matrices were created to explore the correlation between disease indicators and agronomic and meteorological indicators. The most correlated agronomic variables with the date of disease onset were year (the pressure of powdery mildew was much weaker in 2017 compared to 2018), vigor and date of budburst ($P<0.05$). Final disease pressure showed looser correlation with the set of agronomic variables, with at most a 50% correlation with an earlier flowering date ($P<0.05$). Meteorological variables showed higher rates of correlations with disease indicators compared to agronomic variables. The variables that were mostly correlated with the date of disease onset on grapes were temperature (positively), relative humidity and leaf wetness (negative correlation) and rain (all $P<0.05$). In particular, the best predictor of the absence /presence of symptoms in mid-June (separating a group of plots with early infections from another group with late infections) appeared to be relative humidity. Indeed, in late May, two groups of plots emerge either with either high (70-80%) or low (50-60%) relative humidity. These groups differ by the year of observation which appears to be the most important explicative factor of infection earliness. Average leaf wetness duration in 2018, showed a negative correlation with the date of disease onset on leaves for the three plots where it was measured. Final disease pressure was also positively correlated to leaf wetness both on leaves and grapes.

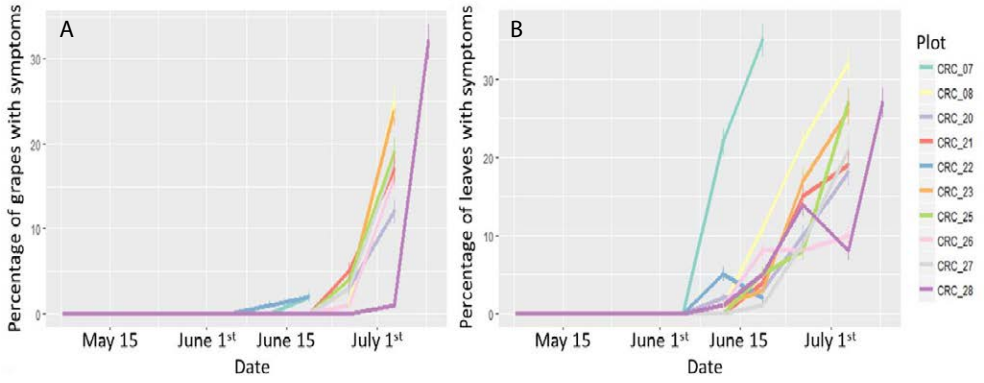


Figure 1. Dynamics of powdery mildew on grapes (2A) and leaves (2B) in 2018.

PCA

The data were further explored with PCA analysis for both years (Figure 2) and for 2018 only (Figure 3). Only the indicators that were found to be highly correlated to disease pressure in the correlation analysis (Figure 2, 3) were considered. Since leaf wetness and relative humidity were positively correlated, and leaf wetness was measured only on 3 data points over a total of 19, only relative humidity was kept in the PCA. Even if grape variety did not show a strong correlation in 2018, it was integrated due to its effect in 2017 (Weens, 2017).
The first two axes describe nearly 70% of the inertia of the graph. The first axis is correlated with the vigor and date of flowering, these 2 variables being negatively correlated. The second axis is

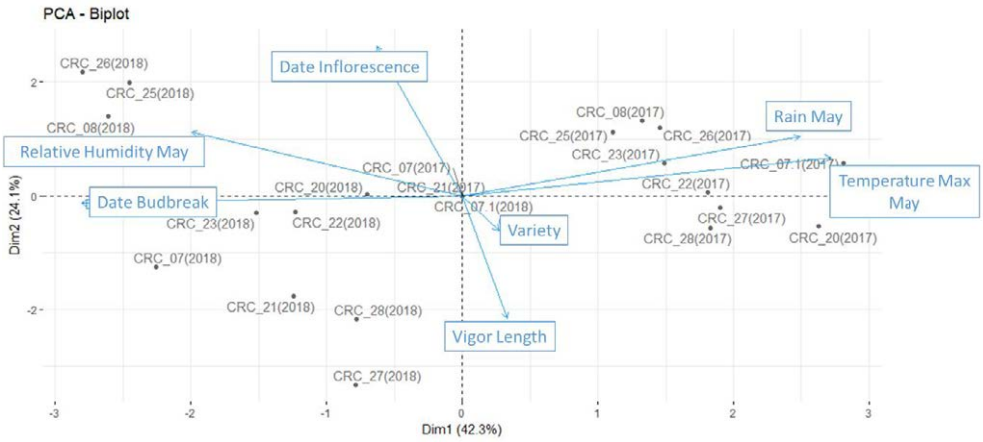


Figure 2. PCA for both years with the variables that were found to be highly correlated to disease pressure.

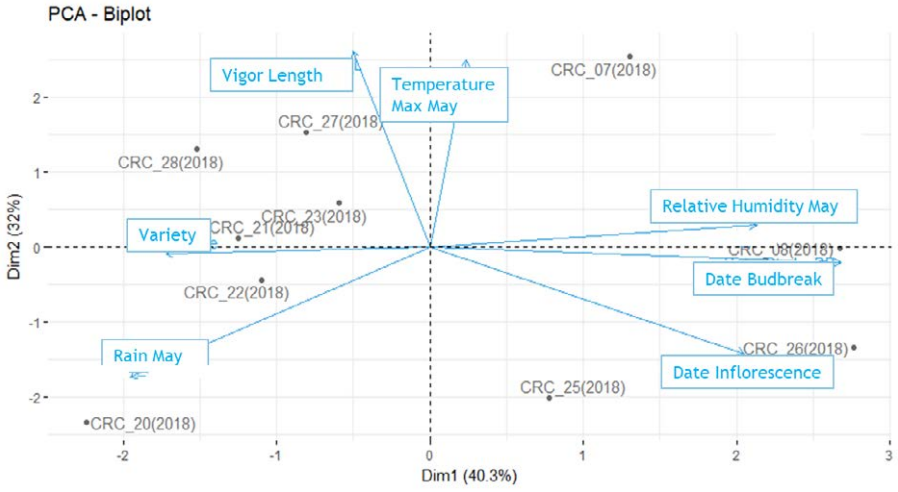


Figure 3. PCA for 2018 with the variables that were found to be highly correlated to disease pressure.

correlated mainly with budbreak date, total rainfall and maximum temperature. Relative humidity also contributes to the description of this axis. The position of plots varies according to the year (2018 on the right and 2017 on the left). CRC_20, CRC_27 and CRC_28 were characterized by heavy rains and high maximum temperatures as well as being more vigorous than the other plots. These 3 plots were the only ones that showed powdery mildew in 2017.
For 2018, a group made up of CRC_21, CRC_27, CRC_23 and CRC_28, having high disease pressure stands out. It is characterized by marked vigor and early flowering and budburst, low humidity. A second group stands out: CRC_08, CRC_25 and CRC_26, with little disease. These plots are characterized by late budburst and high relative humidity. In contrast, they were characterized by heavy rains in May.

Discussion

The correlation analysis highlights a strong year effect, and stronger correlations of powdery mildew with meteorological than with agronomic indicators. Indeed, disease biology in general is known to be strongly affected by weather variables like temperature, humidity and leaf wetness, which also have a large inter-annual variability (Rapilly, 1991). In the case of powdery mildew, temperature hastens germination and promotes spore production, and relative humidity and leaf wetness promotes infections (Delp, 1954; Chellemi and Marois, 1991a; Chellemi and Marois, 1991b).
These effects were confirmed by the analysis, with positive correlations between temperature and humidity and disease indicators. The strength of these correlations depended on the period in which the weather indicator was calculated, which suggest that local and temporal vulnerability to the disease is caused by local and temporal weather conditions. For example, the occurrence of symptoms after Mid-June was observable only in plots that at the end of May had recorded low average relative humidity. The fact that relative humidity at the end of May is also related to the earliness of the first symptoms on grapes but not to the final intensity of the disease suggests that humidity is involved more in the beginning of the infection than its progression.
Almost all indicators of temperature were positively correlated to disease onset on grapes, but not on leaves, implying that warmer temperatures might correspond to later grape infection. This might seem counter-intuitive, given the promoting effect of temperature on spore release (Chellemi and Marois, 1991b). However, this inhibiting effect might not be directly due to temperature (which over the study period always stayed in a favorable range for powdery mildew) but to relative humidity, which is highly negatively correlated to temperature. Interestingly, relative humidity during the month of May was very low (50%) in 2017, but there were later grape infections. Humidity might have acted as a 'limiting factor' for infections in 2017 and determined the strong year effect observed in the study.
The promoting effect of humidity and free water on leaves on disease could be confirmed by the highly significant correlation between leaf wetness and disease onset on leaves. Indeed, whereas some caution must be exerted due to the small size of the leaf wetness dataset (3 data points), this factor seems to explain not just final disease pressure but also the date of onset of leaf symptoms.
The only agronomic factors that appeared to be related to powdery mildew are the year of observation and phenological timing and vigor. These two plant variables are both related to growth. Phenological timing has already been indicated as a factor affecting powdery mildew (Gadoury *et al.*, 2003; Caffarra *et al.*, 2012). On the other hand, vigor is known for favoring diseases in general, as it produces denser canopies, more susceptible tissue, and more humid conditions for the disease to develop (Calonnec *et al.*, 2013). In the study, vigorous plots had later dates of grape infection, contrary to expectations. However, as vigor is positively related to temperature and temperature is negatively related to humidity, it is impossible to say whether the relationship between vigor and disease onset is causal, or whether it is due to the covariance of indicators. Interestingly, the PCA groups more heavily diseased plots in the area of the plot in direction of increasing plant vigor.

This difference might be due to the fact that multivariate analysis decorrelates covarying variables and shows plot similarities according to a group of descriptors rather than highlighting correlation coefficients between single variables.
Moreover, budburst date and flowering date were respectively positively and negatively related to disease onset date on grapes, implying an earlier onset of grape symptoms on early flushing and late flowering plots. This could suggest that a longer budburst-flowering period increased the vulnerability of plots and earlier grape infections.

Conclusions

The present study highlighted, at least for 2017-2018, a stronger correlation of powdery mildew with climate and plant growth than with static plot descriptors. The high degree of temporal (year effect) and spatial (plot effect) heterogeneity of powdery mildew corresponded to an equally high degree of variability in local weather conditions during the period of primary infections. Such variability and the high correlations with local powdery mildew pressure encourages a more precise weather monitoring on plots and inside the canopy. Relative humidity and leaf wetness appeared as the most closely correlated variables to powdery mildew onset and pressure. Undoubtedly these results need to be further confirmed and quantified through more extended surveys, but they indicate interesting directions for the improvement of predictive models of powdery mildew, that are actually unable to predict disease hotspots. Ultimately the ability to build precise tools to optimize phytosanitary management depends on understanding of the environmental effects underlying the physiology of this disease.

References

Caffarra, A., Rinaldi, M., Eccel, E., Rossi, V. and Pertot, I., 2012. Modelling the impact of climate change on the interaction between grapevine and its pests and pathogens: European grapevine moth and powdery mildew. Agriculture, Ecosystems & Environment, 148, 89-101.

Caffi, T., Rossi, V., Legler, S. E. and Bugiani, R., 2011. A mechanistic model simulating ascosporic infections by Erysiphe necator, the powdery mildew fungus of grapevine: dynamic model for ascosporic infections. Plant Pathology 60, 3, 522-531.

Calonnec, A., Burie, J-B., Langlais, M., Guyader, S., Saint-Jean, S., Sache, I. *et al.*, *et al.*, 2013. Impacts of plant growth and architecture on pathogen processes and their consequences for epidemic behaviour. European Journal of Plant Pathology. 135, (3), 479-497

Calonnec, A., Cartolaro, P., Chadoeuf, J., 2009. Highlighting Features of Spatiotemporal Spread of Powdery Mildew Epidemics in the Vineyard Using Statistical Modeling on Field Experimental Data. Phytopathology. 99, (4), 411-422.

Calonnec, A., Cartolaro, P., Naulin, J.-M., Bailey, D. and Langlais, M., 2008. A host-pathogen simulation model: powdery mildew of grapevine. Plant Pathology. 57, (3), 493-508.

Chellemi, D. O., Marois, J. J., 1991a. Development of a demographic growth model for Uncinula necator by using a microcomputer spreadsheet program. Phytopathology. 81, (3), 250-254

Chellemi, D. O., Marois, J. J., 1991b Sporulation of Uncinula necator on grape leaves as influenced by temperature and cultivar. Phytopathology. 81, (2), 197-201

Delp, C. J., 1954. Effect of temperature and humidity on the grape powdery mildew fungus. Phytopathology. 44, 615-626

Gadoury, M. O., Robert, C., Ficke, S. A., Wilcox, W. F., 2003. Ontogenic Resistance to Powdery Mildew in Grape Berries. Ecology and Population Biology. 93, (5), 547-555

Mammeri, Y., Burie, J.B., Langlais, M. and Calonnec, A., 2014. How changes in the dynamic of crop susceptibility and cultural practices can be used to better control the spread of a fungal pathogen at the plot scale? Ecological Modelling. 290, 178-191.

Pearson, R.C and Gadoury, D.M., 1987. Cleistothecia, the Source of Primary Inoculum for Grape Powdery Mildew in New York. Phytopathology. 77, (11), 1509.

Rapilly, F., 1991. L'épidémiologie en pathologie végétale: mycoses aériennes (Epidemiology in plant pathology: aerial mycosis). Paris, France: Editions Quae (INRA), 318 pp. 2-7380-0297-8

Sall, M.A., 1980. Epidemiology of Grape Powdery Mildew: A Model. Phytopathology. 70, (4), 338

Weens V., 2017. L'hétérogénéité de l'oïdium de la vigne – Etude des facteurs influençant son développement. (Hetereogeity of grapevine powdery mildew. Study of the factors affecting its development). Mémoire de fin d'étude, supervised by Bois, B., Caffarra, A. and Garin, G., Nancy, France: Université de Lorraine.

Wilcox, W. F., 2003. Effects of humidity on the development of grapevine *powdery* mildew. Phytopathology 93, 1137-1144

Identifying the *Fusarium* spp. infestation in winter wheat based on RGB imaginary

G.G. Peteinatos[1*], M. Sökefeld[1], J. Machleb[1], K. Cambel[2] and R. Gerhards[1]
[1]*Department of Weed Science (360b), Institute of Phytomedicine, University of Hohenheim, Otto-Sander-Straße 5, Stuttgart, 70599, Germany; g.peteinatos@uni-hohenheim.de*
[2]*Queens University Belfast, Institute for Global Food Security, Stranmillis Road, Belfast, BT9 5AG, United Kingdom*

Abstract

Infestation from *Fusarium* spp. in winter wheat (*Triticum aestivum* L.) fields has increased over recent years. This increase is attributed to the rotation with maize and increased adoption of reduced tillage practices. In these systems, if fungicides are not applied at an appropriate time, the infestation can develop quickly and the wheat can be contaminated with mycotoxins. Current test procedures cannot identify mycotoxin contents before harvest. In the current experiment, six different winter wheat cultivars have been used to monitor the *Fusarium* spp. infestation in Southern Germany in 2018. Wheat ears were harvested during the medium milk stage, along with visual estimations of the infestation level. These ears were measured with an RGB camera, to identify the infested wheat ears, thus creating an infestation index.

Keywords: disease detection, head blight, HSV, mycotoxin content correlation

Introduction

One of the most important pathogens in wheat fields is head blight, which can be caused by a number of *Fusarium* species. Infestation from *Fusarium* spp. in winter wheat fields increased during recent years, due to rotation with maize and reduced tillage practices (Bernhoft *et al.*, 2012). If there is no timely fungicide application, and moist, rainy weather, then the infestation can develop quickly. Apart from the yield reduction that head blight can cause, the biggest impact is the quality degradation due to the presence of toxins in the product (Desjardins, 2006). The mycotoxins produced by the *Fusarium* species include deoxynivalenol (DON), nivalenol (NIV) and HT-2 toxin (HT-2) (Bernhoft *et al.*, 2012). These can result in the production of zearalenone (ZEN) later on if the harvest is delayed. Grains with mycotoxin content can be dangerous if consumed by humans or animals (Trenholm *et al.*, 1994; Gao and Yoshizawa, 1997; Desjardins, 2006). For many toxins, (e.g. DON, ZEN, Fumonisins) there are already established acceptable toxic levels that should not be surpassed if the product is meant for edible use (Parliament and EU, 2007). It is therefore imperative to eliminate the presence of mycotoxins in the food chain. Yet the timely application of a fungicide can be quite challenging since the rainy and moist weather which is favourable for fungal development is prohibitive for the application of fungicides. Based on external factors such as the weather conditions, the infected grains can accumulate mycotoxin content (Bauriegel and Herppich, 2014). Especially in humid and high precipitation years the infestation can be quite high. In those years there is a higher probability of a delayed harvest. Even in a reduced infestation level with the aid of fungicides, the remaining infected ears can still accumulate a significant amount of mycotoxins making the complete product useless. Different methods, like chlorophyll fluorescence, visual and hyperspectral imaging, have been proposed as techniques for identifying *Fusarium* spp. infestations in the field (Bauriegel *et al.*, 2011; Dammer *et al.*, 2011; Bauriegel and Herppich, 2014; Barbedo, 2016). Yet, there is still a lack of rapid and comprehensive *Fusarium* spp. detection methods. Furthermore, there has not been any correlation between the *Fusarium* spp. infestation measured in the field with the mycotoxin content

DOI 10.3920/978-90-8686-888-9_27, © Wageningen Academic Publishers 2019

of the harvested grain. Due to legislative requirements, samples of all edible grain are tested for a possible mycotoxin contamination only during harvest or post-harvest with time-consuming and expensive laboratory tests (e.g. HPLC, serological rapid tests, Fast-DON-ELISA-test, etc.).(Bauriegel and Herppich, 2014).
The aim of this study was to examine the possibility of a robust, efficient and non-destructive method to determine the mycotoxin concentrations caused by *Fusarium* spp infestations in wheat fields.

Materials and methods

Field experiment
Six different winter wheat cultivars have been used to monitor the *Fusarium* spp. infestation in a field experiment located in Southern Germany during the 2018 cultivation period. The cultivars used, ranged from being almost resistant to *Fusarium* spp. infestation till mildly sensitive (Table 1). The field was located near the town Helmstadt (49°20'44.8'N 8°57'52.7'E) and was under a governmental fungicide efficiency monitoring program. It was selected for its high risk of getting a *Fusarium* spp. infestation.
The previous cultivar was maize, and the cultivation treatments were performed in a way to sustain a high infestation probability. For example, leftover biomass from the previous crop (maize) was not removed or integrated into the ground prior to the winter wheat cultivation. Fertilization, weed management, etc. were performed according to the rules and practices of good, professional farming. Seeding was done at 19th of October 2017 with 320 seeds per m^2. The field soil was a sandy loam. A treatment with two fungicide applications was performed, along with an untreated control, all placed in 40×3 m plots. The first application was made on 17th of May 2018 when the first flag leaf was visible (Z 39 to 51) based on Zadoks *et al.* (1974) scale with 100 g/ha epoxiconazol, and 100 g/ha fluxapyroxad, while the second application was performed in the middle of anthesis for all cultivars (Z63) with 125 g/ha tebuconazol, and 125 g/ha prothioconazo. During the medium milk stage (Z 75) the ears contained in 4 random locations of 0.25 m^2 each were harvested from every plot (Figure 1). Images were taken from these ears in front of a black gridded background, and afterward, these ears were counted, and the wet and dry biomass were measured. Furthermore, the mycotoxin content of the harvested grains, particularly DON was measured after the harvest. On the same date as the ear gathering, a visual estimation of the *Fusarium* spp. presence and the percentage of the infestation level of the complete plot was undertaken by an experienced observer in 10 randomly selected spots in each plot.

Table 1. The susceptibility for head blight, the classified maturity group and the number of ears per m^2 harvested for the winter wheat varieties used in the field trial.

Variety	Head blight susceptibility group[1]	Maturity group[2]	Number of ears per m^2
Ambello	3	3	1,147
Chiron	3	5	1,107
Moschus	3	5	1,043
Sokal	3	3	1,373
Spontan	3	5	1,007
Viki	2	5	1,290

[1] Rating from 1 to 9, smaller values are resistant to the infestation, higher values = sensitive cultivars.
[2] Rating from 1 to 9, small numbers belong to early maturing cultivars, big numbers are for late maturity.



(a) (b)

Figure 1. (A) Harvesting of the ears in the field. (B) Ear placement in a defined grid for image acquisition.

Sensor

Optical images were gathered from the ears on a black background with a commercial Canon IXUS 125 HS (Canon Inc., Tokyo, Japan) camera. The camera contains a 1/2.3 type back-illuminated CMOS sensor of 16M resolution. It has a focal ratio between f/2.7 and f/5.9. and a focal length of 21.5 mm /24 mm providing 24 mm at the 35 mm format equivalent. The camera was used in an automatic mode, allowing itself to choose the acquisition parameters (white balance, ISO speed, intelligent image stabilization). Each image captured was 4,608×3,456 pixels containing 42 ears, resulting in an area of ~0.4M per ear and an image resolution of 0.2 mm/pixel. The images were taken at a height of 0.6 m above the ground, and they were taken under natural light conditions, without the aid of any artificial light source or the use of a flash.

Image analyses

Each RGB (Red/Green/Blue) image captured from the camera was separated per gridded ear, resulting in 42 sub-images, one per ear. All further procedures and analyses were conducted on each of those sub-images. For each sub-image, the over-saturation was removed and the colour contrast was improved between the healthy and the infected regions of the ears (Figure 2). Afterward both the original and the enhanced images were transformed in the hue saturation value (HSV) colour space. In the HSV colour space, the colour representation of each object is not correlated with the variation of its brightness (Yang *et al.*, 2015). A mask was created from the hue values of the original image in order to remove the background and any possible noise and distortion inside the image. The Value values were also integrated into the mask to remove shadows on both images, as proposed by Yang *et al.* (2015). Afterward, based on the difference of the hue values from the original and the enhanced image, a second mask was created presenting only the significantly infected regions of the ear for a manual visual assessment and comparison. An infestation ratio was created, defined as the ratio of infected pixels versus the total amount of ear pixels. This ratio was compared with the visual estimation done in the field and the respective mycotoxin content (DON) measured after harvest (Figure 2).

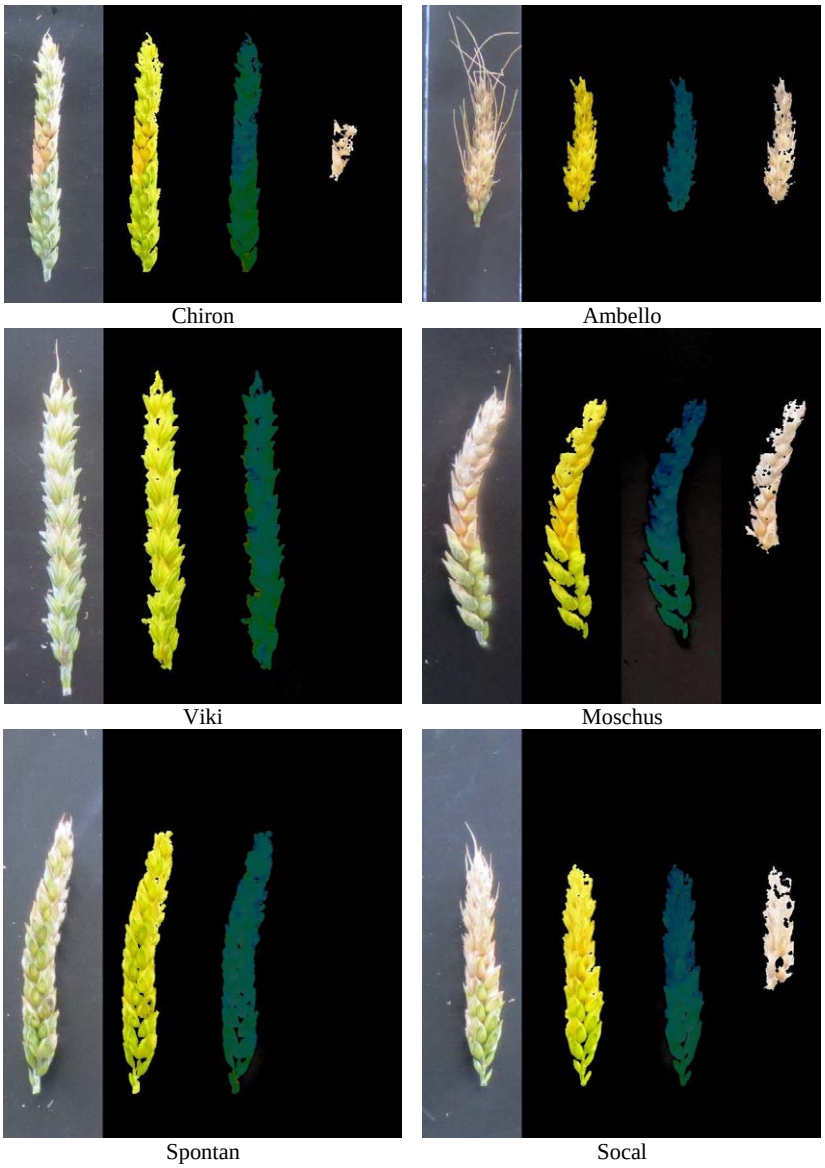


Figure 2. Images from wheat ears of different cultivars in none to low infestation levels (left) and medium to high infestation levels (right). Each image contains the following sectors from left to right: (A) the original image, (B) enhanced image, (C) infestation probability (green is healthy, while blue can be infested) (D) the region identified as infested.

Results and discussion

Even though 2018 was considered to be a dry year in Northern Europe with damage monitored in wheat fields due to draughtiness, the experimental field showed a significant *Fusarium* spp. infestation. All cultivars demonstrated an infestation at the untreated controls (Figure 2), while the

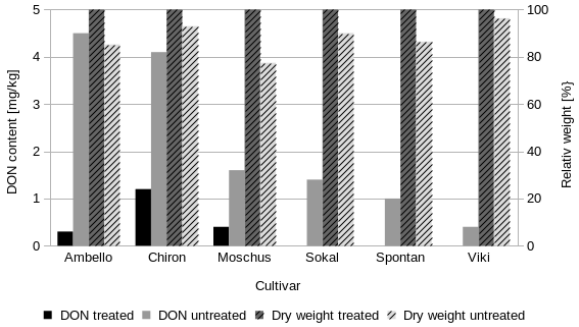


Figure 3. The deoxynivalenol (DON) content (mg/kg) measured on the grains during harvesting and the dry biomass of the harvested ears at the medium milk stage. The darker bars belong to the treated with the fungicides plots of all cultivars and the more light bars are for untreated control per cultivar for all the six cultivars.

cultivars Ambello, Chiron and Moschus had also an infestation at the treatment where fungicides were applied. On the untreated control treatments, only the cultivars Viki and Spontan were measured with a DON content below the legislative threshold of 1,250 mg/kg (European Parliament, 2007), and Spontan was marginally below the threshold (Figure 3).

A good relationship was found between the visual estimation and the infestation ratio calculated from the ear images (r=0.8) (*P*=0.05). A similar relationship was demonstrated for the prediction of the DON content of the plots (r=0.81) (Figure 4). It should be noted though that there is a trend of overestimating the infestation ratio in zero and low infestation levels. This agrees with the findings of Dammer *et al.* (2011) and Bauriegel and Herppich (2014). The harvesting of the ears was performed at the optimum for measurement and identification stage. Even though anthers and sterile heads caused by growth disturbances are reported to be classified as diseased (Bauriegel and Herppich, 2014), this was not identified as a problem in the current experiment. Yet it should be acknowledged that the samples consisted of high-resolution images of individual ears, without background noise or any other form of distortion. The current experiment was the first step to prove the viability of the current methodology and its applicability. Trying to identify and quantify the visual symptoms originating from fungal diseases can be quite laborious and challenging (Barbedo, 2016). The simplicity of the method can be proposed as an addition or alternative to visual infestation

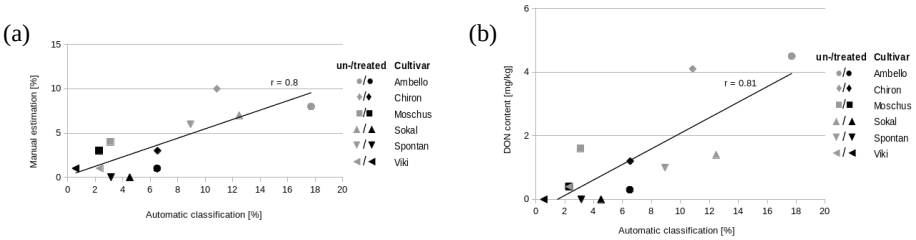


Figure 4. (A) Relationship between the estimated and automatic classification infestation levels (%) for the different cultivars. The grey icons represent the control treatments, while the black icons are for the treatments where fungicides were applied. (B) Relationship between the DON content (mg/kg) and the infestation level (%) identified by automatic classification for the different cultivars.

assessments since it can be fast, and can gather and classify a significant amount of data in a short time. If manually isolated infected ears can be pictured, or a robust background noise reduction can be provided for the in-field ear detection the proposed method can act as a useful tool for farmers and consultancy agencies for identifying and mapping *Fusarium* spp.

Conclusions

Using RGB images for identifying *Fusarium* spp. infected ears, can be achieved in the medium milk stage. The results were correlated with the DON content. Even though the method has been tested only under laboratory conditions, with no noise background, the simplicity of the method can make it a viable candidate as an in-field procedure. More work needs to be done in that direction in order to use this methodology for surveillance and monitoring.

Acknowledgements

The authors would like to thank Mr. Gerd Münkel from Landratsamt Rhein-Neckar-Kreis who set up and executed the field trial, for permitting us gathering samples in the LTZ fungicide field experiment and Mrs. Brita Allgayer, Mrs. Susanne Brüx and Mrs. Foteini Tzioumerka for their assistance during the data sampling. This work was funded by the EIT Food project 'Separating MycOtoxin-contaminated Wheat grains using Precision Farming technologies'.

References

Barbedo, J.G.A., 2016. A review on the main challenges in automatic plant disease identification based on visible range images. Biosystems Engineering 144, 52-60.

Bauriegel, E., Giebel, A., Herppich, W.B., 2011. Hyperspectral and Chlorophyll Fluorescence Imaging to Analyse the Impact of Fusarium culmorum on the Photosynthetic Integrity of Infected Wheat Ears. Sensors 11, 3765-3779.

Bauriegel, E., Herppich, W.B., 2014. Hyperspectral and chlorophyll fluorescence imaging for early detection of plant diseases, with special reference to Fusarium spec. infections on wheat. Agriculture 4, 32-57.

Bernhoft, A., Torp, M., Clasen, P.-E., Løes, A.-K., Kristoffersen, A.B., 2012. Influence of agronomic and climatic factors onFusariuminfestation and mycotoxin contamination of cereals in Norway. Food Additives & Contaminants: Part A 29, 1129-1140.

Dammer, K.-H., Möller, B., Rodemann, B., Heppner, D., 2011. Detection of head blight (Fusarium ssp.) in winter wheat by color and multispectral image analyses. Crop Protection 30, 420-428.

Desjardins, A.E., 2006. Fusarium Mycotoxins: Chemistry, Genetics, And Biology. St Paul, MN, USA: American Phytopathological Society.

Gao, H.-P., Yoshizawa, T., 1997. Further study on Fusarium mycotoxins in corn and wheat from a high-risk area for human esophageal cancer in China. JSM Mycotoxins 45, 51-55.

Parliament, E., EU, C. of the, 2007. Commission Regulation 1126/2007 of 28 September 2007 amending Regulation (EC) No 1881/2006 setting maximum levels for certain contaminants in foodstuffs as regards Fusarium toxins in maize and maize products. Official Journal of the European Union L 255, 1-4.

Trenholm, H.L., Thompson, B.K., Foster, B.C., Charmley, L.L., Hartin, K.E., Coppock, R.W., *et al.*, 1994. Effects of feeding diets containing Fusarium (naturally) contaminated wheat or pure deoxynivalenol (DON) in growing pigs. Canadian Journal of Animal Science 74, 361-369.

Yang, W., Wang, S., Zhao, X., Zhang, J., Feng, J., 2015. Greenness identification based on HSV decision tree. Information Processing in Agriculture 2, 149-160.

Zadoks, J.C., Chang, T.T., Konzak, C.F., 1974. A decimal code for the growth stages of cereals. Weed Research 14, 415-421.

Marrying futuristic weed mapping with current herbicide sprayer capacities

G.J. Somerville[1*], R.N. Jørgensen[2], O.M. Bojer[3], P. Rydahl[3], M. Dyrmann[2], P. Andersen[4], N.-P. Jensen[5] and O. Green[6]
[1]*Department of Agroecology, Aarhus University, Forsøgsvej 1, 4200 Slagelse, Denmark; gs@agro.au.dk*
[2]*Department of Engineering, Aarhus University, Finlandsgade 22, 8200 Aarhus N, Denmark*
[3]*IPM Consult ApS, Stenlille, Denmark*
[4]*Datalogisk A/S, Noerre Alslev, Denmark*
[5]*I-GIS A/S, Risskov, Denmark*
[6]*Agro Intellingence ApS, Aarhus, Denmark*

Abstract

For the last 25 years in Denmark, there has been an emphasis on both improving weed control, and reducing the use of pesticides. Analysis at the field level has achieved a reduction in herbicide rates of 20-40%, using uniform rates and typically available herbicide application machinery (sprayers). A winter wheat field of ~8.9 ha was photographed at a rate of 100 photos/ha. The resulting weed maps were then used to build herbicide spray maps. The advantage of these spray maps is that they were specifically designed to be used with currently available sprayers, using only a global navigation satellite system and boom-wide spray changes. This study predicted a further 40% reduction in herbicide use (below the already reduced recommendations discussed in paragraph 1 above), whilst still maintaining a high level of weed control.

Keywords: application maps, automated weed recognition, herbicide mixtures, integrated weed management, 'IPMwise'

Introduction

Several universities and companies now have, or are developing, the capability to analyse high-resolution photos via machine learning (Berge *et al.*, 2012; Dyrmann, 2017; Franco *et al.*, 2017; Ørum *et al.*, 2017). Finely detailed weed maps of within-field weed populations are becoming a reality for precision agriculture, creating the opportunity for significant improvements in integrated weed management (IWM) (Christensen *et al.*, 2009; Gutjahr and Gerhards, 2010; Hamouz *et al.*, 2013; Johnson *et al.*, 1995).
Detailed weed maps of within-field weed populations are becoming more common. This increased knowledge of weed populations is then being used when selecting the herbicides and rates to be used in each field, leading to a 20-40% reduction in herbicide use, when compared to local best management practices (Sønderskov *et al.*, 2015). These 'pesticide action plans' use a Danish decision support system (DSS) for integrated weed management (IWM). The latest on-line tool, called 'IPMwise' (and its predecessor systems) have undergone field validation trials in Denmark, Norway, Germany, Spain and Latvia (Montull *et al.*, 2014; Sønderskov *et al.*, 2014; Sønderskov *et al.*, 2015; Vanaga and Zarina, 2008). IPMwise integrates the eight general IPM-principles as defined by the EU in Directive 2009/128/EC. Typically, one common herbicide mixture and uniform dose is recommended for each field, and all requirements for agronomical robustness have been examined in the above cited studies. This implies that any sprayer independent of technology level may realise similar savings when utilizing cost-efficient photograph acquisition and automated weed identification. Christensen *et al.* (2009) named this spatial treatment resolution as level 4, utilising the uniform treatment of the whole field. However, further herbicide reductions will require spatial

DOI 10.3920/978-90-8686-888-9_28, © Wageningen Academic Publishers 2019

on/off nozzle control and/or variable herbicides and dosages, thereby moving to spatial treatment resolution level 3, which treats weed patches or subfields with clusters of weed plants.

Berge *et al.* (2007) concluded that on average, boom width had little influence on the potential herbicide reductions, and in the longer term (i.e. several years and/or larger fields) boom width had no influence on herbicide savings. This is in line with Franco *et al.,* (2017) who concluded 'the simplest and most profitable increase in precision comes from turning the full meter boom sprayer on (and later off) every time it passes a weed patch. Such an increase in precision reduced the herbicide cost for controlling thistles in barley by 35-60%'.

The next significant herbicide saving potential (assuming a single tank mixture is used) is to turn the sprayer boom off when the cumulative weed pressure is below a predetermined threshold. Gutjahr *et al.* (2012) showed the average untreated area when patch spraying in 11 fields was 37% (between 2 and 89%). Wiles (2009) calculated an average untreated area of 34.5% in 16 corn fields. However, efficient herbicides differ for broad-leaved weeds and grass weeds. Wiles (2009) showed that when herbicide use included having two different herbicides for patch spraying (instead of one uniform application) this reduced the herbicide use (g/ha) nearly 10-fold. Further increasing the number of herbicides (above two) didn't increase the net gain significantly. Hence, it seems that the majority of herbicide savings may be obtained by patch spraying with two separate herbicides.

The aim of this work was to use a case field to explore the herbicide and cost savings possible from using a single (two passes) and dual herbicide tank mixture (one pass) to perform full boom width patch spraying. For both the single and dual herbicide tank mixtures the dosage was varied laterally whilst spraying, which is within the limited capabilities of many sprayers already in use within the agricultural sector.

Materials and methods

A field map which was usable with currently available agricultural sprayers was created using re-examined data from a field previously analysed in Rydahl *et al.* (2018). They had analysed a farmer's field of 8.92 ha grown with winter wheat in Funen, Denmark in the autumn of 2016. Their photos were taken using a machine vision camera with a flash mounted on an ATV, described as field A in Laursen *et al.* (2017). Each photo covered 0.25 m^2, and they were taken approximately 10×10 m apart. In total, 862 photos were taken across the whole field (Figure 1), and the location of each

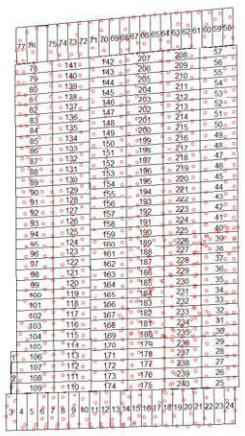


Figure 1. Red circles show the location of each individual photo, taken in autumn 2016. The sequential numbers within most of the 243 polygons suggest a route for the sprayer.

image was recorded with a global navigation satellite system (GNSS). A total of 7,794 images of single weeds were extracted from the 862 photos using a trained convolutional neural network, and then analysed for their weed species (N.B. some weeds were classified to the genera level (Rydahl *et al.*, 2018)) and growth stage. Eleven different weed species were identified, with the most common weed size being 2- and 4 leaves, when counting both cotyledon- and true leaves (Rydahl *et al.*, 2018).
A field map which could be useful with commonly available machinery, was designed using 36 m (a common sprayer width) as the main width for all field divisions. In addition, a typical sprayer route was used, involving an outer circuit, followed by a back and forth spraying pattern, to completely encompass the field (Figure 1). The resulting 36 m wide paths were divided laterally into polygons (approximately rectangular in shape), so that each polygon was 36 m wide, and approximately 10 m long. This created 243 polygons, which were then overlaid on the photograph locations. This meant that each polygon typically contained three-four photos, with a range from zero (3 polygons were empty) to nine. (Figure 1). A path for the sprayer was then determined, using the central points of each polygon (Figure 2A).
The photos within each polygon were analysed as a group, to calculate an overall density and growth stage of each weed species, independently for each polygon. The density of each species was averaged within each polygon. Growth stage within each polygon was based on the 75 percentile for each species, within each polygon. Three polygons which did not contain any photos were assigned intermediate values from nearby polygons.
The most suitable herbicides (to a maximum of four) to use within this field were calculated using the density and 75 percentile on growth stage for each species in the field as a whole. The dose rates were then calculated by IPMwise (Rydahl *et al.*, 2017; Rydahl *et al.*, 2018) individually for each polygon using the additive dose model (ADM) (Jensen and Kudsk, 1988; Streibig *et al.*, 1998) to produce a spray map of application rates for each herbicide, to each polygon. In this first instance (case one) it was assumed that the sprayer gave infinite control over the amount applied to each polygon (Figure 2A and 2B). In case two, a spray map was developed which used both herbicides together in a tank mix, with individual rates (of the same tank mix) determined for each polygon (Figure 3). For case two, adjustments of the spray rate were done within a more realistic ±50%, which is possible using

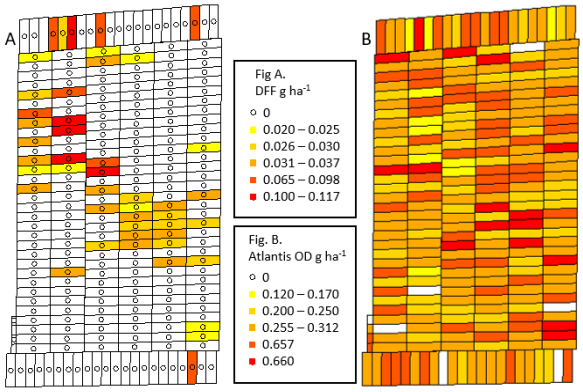


Figure 2. Case one. IPMwise determined application rates for each herbicide individually. Each polygon is 10 m by 36 m. (A) The herbicide DFF (in l/ha) to a maximum of 0.117 l/ha. A total of 26 levels were recommended by IPMwise, which are summarised here into six graduated classes. A GNSS path for the sprayer using the central points of each polygon is illustrated using open circles. (B) The herbicide Atlantis OD (in l/ha), to a maximum of 0.66 l/ha. A total of 28 levels were recommended by IPMwise, which are summarised here into six graduated classes.

Danfoil atomizer technique (personal communication, December 6th, 2018, Martin Dyhrman Sørensen, Danfoil A/S). Costing were calculated for the three alternative spray maps and compared to the use of an optimum uniform rate of application (thereafter called the 'standard rate') which was also calculated by IPMwise (Table 1). The standard rate was 0.177 l/ha diflufenican (DFF) and 0.66 l/ha Atlantis oil dispersion (OD).

Results

In case one, where the herbicides were applied separately (Figure 2A and 2B), there was a 53% saving in financial outlay for the cost of the herbicides, when compared to the standard rate applied across the field (Table 1). The use of DFF was reduced by a huge 92.5%, whereas Atlantis OD was reduced by a still respectable 44.5%.
Case two, involving a tank mix of DFF and Atlantis OD (Note 1), reduced the use of both herbicides by 40.3% (when compared to the standard rate across the field). In addition, because all the application rates were within ±50% of a median value, a standard sprayer could be used.

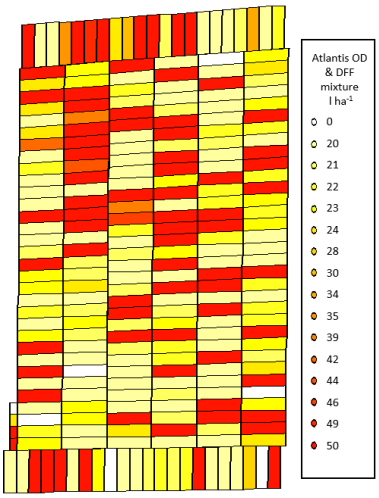


Figure 3. Case two. IPMwise determined application rates for the two herbicides, applied as a tank mix. Each polygon is 10 m by 36 m. Variable application rates for the tank mix (in l/ha, when 50 l contains 0.117 l DFF and 0.66 l Atlantis OD).

Table 1. Herbicide costs in Denmark for case one and two, compared with the standard rate (assuming 1 DKK=0.134 €).

	DFF (55.61 €/l)	Atlantis OD (44.22 €/l)	Tank mix
Standard rate (IPMwise, uniform application)	87.80 €	260.33 €	348.13 €
Assuming complete rate control (case one)	6.60 €	144.53 €	Total 151.13 €
Tankmix, with ±50% rate control (case two)	-	-	207.83 €

Note 1

- DFF: 500 g/l diflufenican, 1 TFI=0.2 l/ha
- Atlantis OD: 2 g/l iodosulfuron-methyl-Na + 10 g/l mesosulfuron-methyl, 1 TFI=0.9 l/ha
- Tankmix: each 50 litres contains 0.117 l DFF and 0.66 l Atlantis OD

Discussion

In case one, the greatest herbicide reductions were obtained, with a 53% saving in the financial cost of the herbicides. The additional investment for the farmer, of the cost of surveying the fields will scarcely impact this saving. To achieve an average photo density of 99 photos/ha the estimated cost would be from €2.40 to €6.60/ha (Laursen *et al.,* 2017). However, in case one, the two herbicides were applied at different rates to different polygons, and a standard sprayer would not be able to achieve this without completing (almost) two circuits of the field, with a different herbicide in the tank for each circuit. This extra time on the sprayer would significantly impact on the savings obtained from using less herbicide. In addition, the currently calculated rates assume both herbicides are applied together, meaning a rate recalculation may be necessary, for separated applications (Streibig *et al.*, 1998).

Case one could also be fulfilled by using a sprayer mounted with two independently controlled tanks, which could apply both herbicides at once, or, by using a direct injection system. However if this level of complexity (with two tanks) was included in the sprayer, it may also be coupled with individually controlled nozzles, which would mean that a spray map with much smaller polygons would be more useful (Rydahl *et al.*, 2018).

Case two, using a tank mix, reduced the use of both herbicides by 40.3%, when compared to a full rate application across the field. This tank mix solution is slightly more expensive and uses more herbicide (when compared to case one), however the whole operation is achieved with just one tank mix, and one circuit of the field, which would mean substantial savings in time and fuel when compared to the two circuits required in case one. Furthermore, the areas receiving more DFF herbicide are often surrounded by areas needing near maximum rates of Atlantis OD, so a little overlap will help cover possible weeds missed by the limited range of the photographs.

The use of a 10 m grid gave 3-4 photos per 360 m^2 polygon. This system was designed to both maximise accuracy (Kolářová and Hamouz, 2016), and to allow a minimum distance between spray volume changes of 10 m. In future, it may be possible to adopt a flexible sampling policy, where parts of a field with more heterogeneity are sampled more frequently (Smartt, 1978). In order to maximise the accuracy of spray mapping, further research could look at whether 3-4 photos per 360 m^2 polygon is the best use of photo number, or polygon size. However, as spatial heterogeneity varies for different weed species (Rydahl *et al.*, 2018), an ideal grid and polygon size may be difficult to determine

Conclusions

A further 40% reduction in herbicide use (below the already reduced IPMwise recommendations) can be obtained by using maps designed for currently utilised tank mix sprayers in a wheat field in Denmark. These saving are obtained using only GNSS and varying the liquid flow rate within the spray nozzle limitations or by changing the sprayer forward speed, which is within the limited capabilities of many sprayers already in use within the agricultural sector.

Acknowledgements

The work was founded by the Innovation Fund Denmark via the RoboWeedMaPS project (J.nr.6150-00027B). Helpful feedback on the manuscript was provided by Solvejg Kopp Mathiassen.

References

Berge, T. W., Fykse, H., Aastveit, A. H., 2007. Patch spraying of weeds in spring cereals: simulated influences of threshold level and spraying resolution on spraying errors and potential herbicide reduction. Acta Agriculturae Scandinavica Section B-Soil and Plant Science 57, 212-221.

Berge, T., Goldberg, S., Kaspersen, K., Netland, J., 2012. Towards machine vision based site-specific weed management in cereals. Computers and Electronics in Agriculture 81, 79-86.

Christensen, S., Søgaard, H. T., Kudsk, P., Nørremark, M., Lund, I., Nadimi, E. S., *et al.*, 2009. Site-specific weed control technologies. Weed Research 49, 233-241.

Dyrmann, M., 2017. Automatic detection and classification of weed seedlings under natural light conditions. PhD thesis. Syddansk Universitet. Det Tekniske Fakultet

European Parliament: Council of the European Union, 2009. Directive 2009/128/EC of the European Parliament and of the Council of 21 October 2009 establishing a framework for Community action to achieve the sustainable use of pesticides (Text with EEA relevance). BG, ES, CS, DA, DE, ET, EL, EN, FR, IT, LV, LT, HU, MT, NL, PL, PT, RO, SK, SL, FI, SV, pp. 71-86.

Franco, C., Pedersen, S. M., Papaharalampos, H., Ørum, J. E., 2017. The value of precision for image-based decision support in weed management. Precision Agriculture 18, 366-382.

Gutjahr, C., Gerhards, R., 2010. Decision rules for site-specific weed management. Precision Crop Protection-the Challenge and Use of Heterogeneity. Cham, Switzerland: Springer, pp. 223-239.

Gutjahr, C., Sökefeld, M., Gerhards, R., 2012. Evaluation of two patch spraying systems in winter wheat and maize. Weed research 52, 510-519.

Hamouz, P., Hamouzová, K., Holec, J., Tyšer, L., 2013. Impact of site-specific weed management on herbicide savings and winter wheat yield. Plant, Soil and Environment 59, 101-107.

Jensen, P., Kudsk, P., 1988. Prediction of herbicide activity. Weed research 28, 473-478.

Johnson, G., Mortensen, D., Martin, A., 1995. A simulation of herbicide use based on weed spatial distribution. Weed Research 35, 197-205.

Kolářová, M., Hamouz, P., 2016. Introduction to weed mapping methodology. Atlas of Weed Mapping, 425-427.

Laursen, M. S., Jørgensen, R. N., Dyrmann, M., Poulsen, R., 2017. RoboWeedSupport-sub millimeter weed image acquisition in cereal crops with speeds up till 50 km/h. Paper accepted for 19th international conference on precision agriculture. Kyoto, Japan April, pp. 27-28. International Journal of Agricultural and Biosystems Engineering 11 (4) 317-321. https://waset.org/publications/10007027/roboweedsupport-sub-millimeter-weed-image-acquisition-in-cereal-crops-with-speeds-up-till-50-km-h

Montull, J., Soenderskov, M., Rydahl, P., Boejer, O., Taberner, A., 2014. Four years validation of decision support optimising herbicide dose in cereals under Spanish conditions. Crop Protection 64, 110-114.

Ørum, J. E., Kudsk, P., Jensen, P. K., 2017. Economics of site-specific and variable-dose herbicide application. Precision Agriculture: Technology and Economic Perspectives. Cham, Switzerland: Springer Nature. pp. 93-110.

Rydahl, P., Bojer, O. M., Jorgensen, R. N., Dyrmann, M., Andersen, P., Jensen, N.-P., *et al.*, 2018. Spatial variability of optimized herbicide mixtures and dosages. International Conference on Precision Agriculture. International Society of Precision Agriculture, pp. 1-14. https://www.ispag.org/proceedings/?action=abstract&id=5040&search=years

Rydahl, P., Jensen, N.-P., Dyrmann, M., Nielsen, P.-H., Jørgensen, R. N., 2017. RoboWeedSupport-Presentation of a cloud based system bridging the gap between in-field weed inspections and decision support systems. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences 8, 860-864.

Smartt, P., 1978. Sampling for vegetation survey: a flexible systematic model for sample location. Journal of Biogeography 1978, 43-56.

Sønderskov, M., Fritzsche, R., de Mol, F., Gerowitt, B., Goltermann, S., Kierzek, R., *et al.,* 2015. DSSHerbicide: weed control in winter wheat with a decision support system in three South Baltic regions–field experimental results. Crop Protection 76, 15-23.

Sønderskov, M., Kudsk, P., Mathiassen, S. K., Bøjer, O. M., Rydahl, P., 2014. Decision support system for optimized herbicide dose in spring barley. Weed technology 28, 19-27.

Streibig, J. C., Kudsk, P., Jensen, J. E., 1998. A general joint action model for herbicide mixtures. Pesticide Science 53, 21-28.

Vanaga, I., Zarina, L., 2008. Evaluating herbicides in a range of doses for integrated plant protection in Latvia. Zemdirbyste-Agric 95, 22.

Wiles, L., 2009. Beyond patch spraying: site-specific weed management with several herbicides. Precision Agriculture 10, 277-290.

Use of simulations to study herbicide site-specific spraying

S. Villette[1*], T. Maillot[1], J.P. Douzals[2], G. Jones[1], J.N. Paoli[1] and J.P. Guillemin[1]
[1]*Agroécologie, AgroSup Dijon, INRA, Univ. Bourgogne, Univ. Bourgogne Franche-Comté, 26, bd Docteur Petitjean, 21000 Dijon, France; sylvain.villette@agrosupdijon.fr*
[2]*UMR ITAP, Irstea Montpellier SupAgro, Univ. Montpellier, 361 Rue Jean-François Breton, BP 5095, 34196 Montpellier cedex 05, France*

Abstract

Site-specific weed management is one of the potential ways to reduce the chemical amount required to control crop weeds. Sources of optimization were investigated with spray width from an individual nozzle or a boom section width and spray deposition patterns were applied to different spatial coverage rates of weeds. The objective was to quantify herbicide saving possibilities as well as potential under-application areas. Simulations were performed with a collection of virtual weed infestation maps and an actual infestation map derived from aerial images. Results issued from the computation were expressed in terms of three indicators: (1) the ratio between herbicide amount in the case of a site-specific application and the amount deposited in the case of an application on the full field surface area; (2) the proportion of weed coverage on which the application rate is higher than 85% of the rated dosage; (3) the proportion of the weed surface on which the deposited application rate is lower than 50% of the rated dosage. The interest in reducing the section width depends on weed spatial distribution and aggregation degree. Moreover, in the case of uniformly dispersed small patches and the use of narrow sections (1 to 4 nozzles) equipped with triangular spray pattern nozzles, attention should be put to the increase of under-dosing areas.

Keywords: sprayer, site-specific weed management, boom section control, spray pattern

Introduction

Reducing herbicide use, without lowering productivity, is going to remain challenging in future years. Site-specific weed management could help in reducing herbicide amount by triggering spraying only where weeds are present in a field (Timmermann *et al.,* 2003, Weis *et al.,* 2008). From a technological point of view, various weed detection and spraying control systems have been available for several decades (Stafford and Ambler, 1994, Lee *et al.,* 1999, Nordmeyer, 2006, Gerhards and Oebel, 2006, Gutjahr *et al.,* 2012). However the adoption rate of these technologies by farmers is still extremely low. The uncertainty regarding herbicide reduction, weed control efficiency and return on investment are the main brakes. The width of the section control directly affects the cost and the performance of the sprayer. On the one hand, the technology cost increases when the width of the individual application sections decreases from a boom section of half dozen nozzles to individual nozzles. On the other hand, the herbicide amount required in the field is expected to decrease with the reduction of section width as it improves the spatial resolution of the application. Recently, Franco *et al.* (2017) addressed the problem of selecting the appropriate weed precision spraying technology. These authors carried out simulations on two small virtual spreading areas infested with many scattered circular patches or few significant circular patches respectively. In their simulations, the spray pattern of the smaller independent boom section was a rectangular shape of 1 m width. Their findings showed that the potential gains not only depend on spraying resolution but also on the density and the distribution of weed patches.
The aim of the present work was to investigate the influence of spray deposit patterns, section widths, weed infestation rates and weed spatial distributions on herbicide reduction in the case

DOI 10.3920/978-90-8686-888-9_29, © Wageningen Academic Publishers 2019

of site-specific weed management. To overcome the limits of practical experiments and to explore the impact of these parameters, virtual experiments were carried out using numerical simulations.

Materials and methods

In this study, herbicide applications were computed combining infestation maps and targeted spraying simulations. Simulations were implemented using Matlab® software.

Virtual weed infestation map

Assuming that weeds frequently grow in patches, virtual infestation maps were built using a random selection of weed patch locations on virtual fields as bands. These bands were of 72 m width and 20,000 m long, computed with a spatial resolution of 0.025×0.025 m (e.g. pixels of 6.25 cm^2). The band width corresponds to a central sprayer pass with a working width of 24 m and two adjacent swaths. The length of the band was chosen long enough so that the assessment indicators computed in the study reached stabilized values. Weed patches were modelled with ellipse surfaces defined by their width (w_e) and length (l_e). Ellipses were oriented so that their major axes were aligned with the travelled direction of the virtual sprayer. This patch shape and orientation were chosen to take into account seed dispersal due to natural processes, combine harvesting and tillage processes. A set of virtual field bands were computed for various targeted weed coverage rates set at 5, 10, 20, 30, 40 and 50%; various ellipse lengths of 0.5, 1, 4, 8, 16 and 32 m; and various minor axis to major axis ratio of ellipses of 0.2, 0.5 and 1 (ellipses correspond to circles when the ratio is 1). Thus a set of virtual weed maps were computed representing various infestation rates and aggregation degrees. In this study, the elliptical patches correspond to the areas resulting from a weed detection process and requiring herbicide application. The detection process is assumed to perfectly localize infested areas regardless of weed densities within the areas.

Actual weed infestation map

Aerial images of a sugar beet field were used to provide one realistic weed infestation map. The sugar beets, sown with a row spacing of 50 cm, were at an early growth stage corresponding to the post-emergence herbicide application period. The field was infested by thistle (*Cirsium arvense* L.) from seedling to rosette stage, wild buckwheat (*Fallopia convolvulus* (L.) Löve) at 4-leaf stage, and rye-grass (*Lolium multiflorum* Lam) at tillering stage. The images were captured in May 2015 by Airinov Company with a MultiSPEC 4C sensor mounted on a UAV. The spatial resolution of the orthoimage was 6 mm/pix. Crop rows were used to discriminate crop from weeds using an algorithm based on spatial information and described in Louargant *et al.* (2018). Figure 1 shows the result of image processing. The binary image of weed location was then rescaled so that the final resolution of the weed infestation map was 0.025 m/pix. A set of 10 field bands was derived from the weed infestation map so that each field band was centred on the sprayer tramlines (manually detected) spaced at 36 m. The mean value of the weed coverage was 5.5%.

Targeted spraying simulation

In this study, the spraying simulator considered a virtual boom equipped with nozzles spaced every 0.5 m. The spray boom was divided in independently controlled sections including a proportional number of nozzles to the boom width (constant nozzle spacing). Simulations were implemented with sections of 1, 2, 4, 6, 12, 24 and 48 nozzles when the working width was 24 m and 1, 2, 4, 6, 9, 12, 18, 36 and 72 when the working width was 36 m. For herbicide application, each section was switched on only where weeds (at least one pixel) were present under the 'weed detection width' of the considered section. For example, in the case of independently controlled nozzles, the 'weed detection width' and the 'spray width' were defined as depicted in Figure 2. Since nozzles were spaced every 0.5 m, the weed detection width of a section composed of N_N nozzles was always

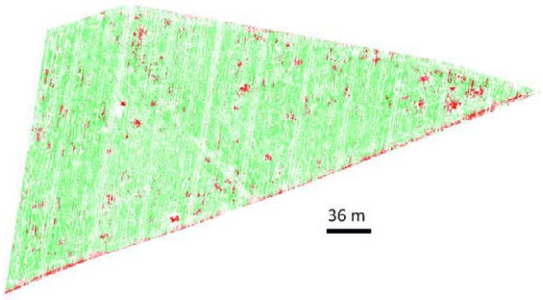


Figure 1. Orthoimage of the sugar beet field after weed detection: crop in green and weed in red (light and dark grey for printed version).

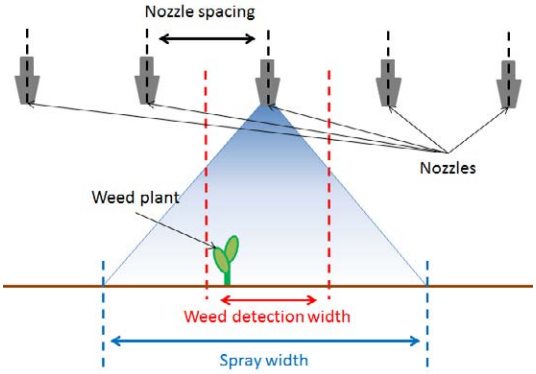


Figure 2. Weed detection width and section spray width in the case of nozzles controlled individually.

N_N×0.5 m while the spray width depended on the nozzle spray pattern. Thus, the spraying simulator was designed to take into account the overlapping between adjacent nozzles, adjacent sections or adjacent sprayer passes. An instantaneous application was assumed, without any time lag of the control valve inducing a smoothed herbicide rate transition in the travelled direction.

Regarding the transverse direction, three different theoretical spray patterns were simulated: (1) a rectangular pattern with a spraying width of 0.5 m (corresponding to a perfect even spray nozzle distribution with a nozzle at 0.6 m high); (2) a triangular pattern with a spraying width of 1 m (corresponding to a flat fan spray of a reduced top angle nozzle of 80 to 90° at 0.6 m high); and (3) a triangular pattern with a spraying width of 1.5 m (corresponding to a flat fan spray of a nozzle top angle of 110° to 120° at 0.6 m high). Figure 3 presents the spray patterns of individual nozzles (i.e. application rate profile perpendicular to the travelled direction) and the resulting application rate distribution for a section of 4 nozzles.

The shape of each of the three spray patterns was chosen so that: (1) the herbicide amount applied by each nozzle (computed as the integral of the spray pattern) was set at 1 in arbitrary units; (2) the mean value of the application rate obtained after overlapping was set at 1 in arbitrary units. Thus, the application was at the rated dosage when the value of the application map pixels was 1 in arbitrary units. The application maps derived from spraying simulations had the same spatial resolution as infestation maps (i.e. 0.025×0.025 m).

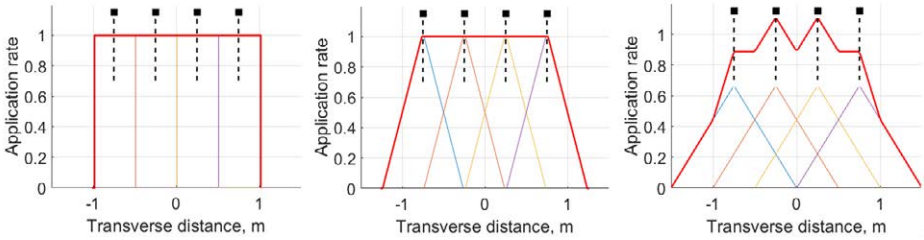


Figure 3. Nozzle spray patterns and transverse application rates (in Arbitrary Unit) after overlapping (bold) for a section of 4 nozzles with rectangular pattern of 0.5 m width (left), triangular pattern of 1 m width (middle) and triangular pattern of 1.5 m width (right). Black squares and dashed lines correspond to the nozzles and their vertical axes.

Assessment indicators

Some output parameters and indicators were defined to assess the reduction of herbicide use and the application efficacy. These parameters were computed on the area corresponding to the central pass of the sprayer (whose working width was 24 m for virtual maps and 36 m for actual map). This area was noted A_T. The parameters used in this study were defined as follows:

- Q_S (in l) is the amount of herbicide used in the case of a site-specific application;
- Q_T (in l) is the amount of herbicide required in the case of a uniform application at the rated dosage on the full field surface area;
- A_W (in m^2) is the weed coverage area;
- $A_{W>\alpha\%}$ (in m^2) weed area on which the application rate is higher than α% of the rated dosage;
- A_S (in m^2) is the sprayed area (area subjected to real herbicide deposit);
- $A_{S<\beta\%}$ (in m^2) is the sprayed area on which the application rate is lower than β% of the rated dosage.

These were used to compute the following dimensionless ratios: A_S/A_T, Q_S/Q_T, $A_{W>85\%}/A_W$ and $A_{S<50\%}/A_S$. In this paper, the thresholds α and β were arbitrarily set at 85 and 50% to provide simple indicators related to the proportion of weed area receiving an application rate considered high enough to be lethal and, on the other hand, the proportion of sprayed area receiving sub-lethal dosage and thus presenting conditions that may contribute to potential development of herbicide resistance. The value of the threshold α was chosen lower than the minimal value of the application rate obtained after overlapping (Figure 3).

Results and discussion

Results derived from virtual infestation weed map

For constant values of weed coverage rate, section width and spray pattern, simulations demonstrated that the sprayed area ratio A_S/A_T only depends on the width of the patches regardless of their lengths in the travel direction. This can be explained by: (1) the number of switched-on sections that were only depending on the width of the patches and (2) by the fact that the cumulated surface of patches is constant for an infestation rate. This can be used in further studies to reduce the number of required simulations.

Figure 4A and 4B introduce the evolution of the sprayed amount ratio (Q_S/Q_T) as a function of the weed coverage, the boom section width and the type of spray pattern. All values represented in Figure 4A and 4B consist of the superimposition of points corresponding to the type of spray pattern. The explanation is that the amount applied by each type of nozzle (computed as the integral of the spray pattern) is the same and the mean value of the application rate obtained after overlapping is also the same for the three cases.

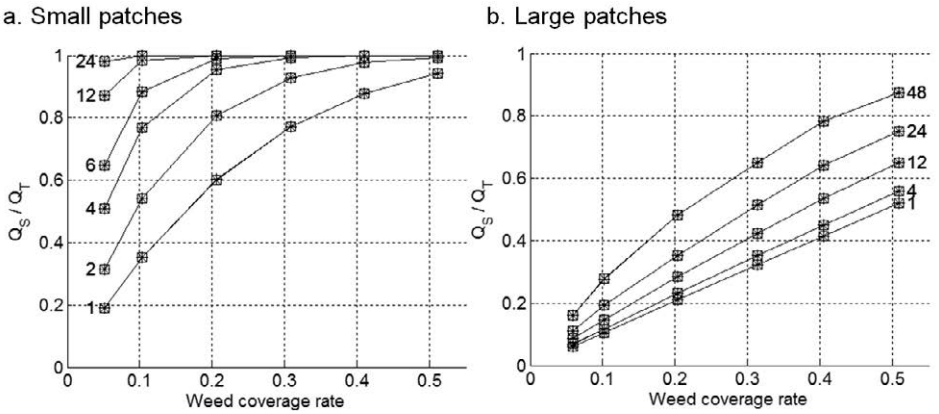


Figure 4. Sprayed amount ratio with respect to weed coverage rate for small ellipse (l_e=1 m, w_e=0.2 m) patches (A) and large ellipse (l_e=32 m, w_e=16 m) patches (B); for three spread patterns: rectangle of 0.5 m width *, triangle of 1 m width × and triangle of 1.5 m width +. The section width (expressed in number of nozzles) is indicated next to the curves.

Figure 4A and 4B also illustrate the changes in the sprayed amount ratio Q_S/Q_T for two different patch sizes. Whatever the situation, the ratio Q_S/Q_T obviously increases when both the section width and the weed coverage rate increase. The reduction of herbicide use is more important when patches are larger. Moreover, the change in the sprayed amount ratio with respect to the infestation rate is affected by the size of the patches. For example, in the case of small patches (l_e=1 m, w_e=0.2 m) with a low weed coverage rate of 5%, the required amount of herbicide is very sensitive to the spray section width. The amount is only 20% with 1-nozzle sections, 50% with 4-nozzle sections and almost 90% with 12-nozzle sections. Moreover, in the case of small patches, the required amount of herbicide rapidly increases with the infestation rate. It approximately rises to 60 and 95% for 1 and 4-nozzle sections respectively when the coverage rate is 20%. In the case of large patches (l_e=32 m, w_e=16 m), the reduction of herbicide use is more important. Moreover, the changes in the sprayed amount ratio (related to the different parameters) are not the same. In this case, the difference observed in the ratio Q_S/Q_T between the use of narrow spray sections or large spray sections is highly reduced. For example, when the weed coverage rate is 10%, the herbicide required amount is 10% with 1-nozzle sections, 15% with 12-nozzle sections and less than 20% with 24-nozzle sections. Additionally, the range of change in the ratio Q_S/Q_T decreases with the patch widths. For example, when the weed coverage rate reaches 50%, the ratio Q_S/Q_T is approximately 52% with 1-nozzle sections, 63% with 12-nozzle sections and 75% with 24-nozzle sections.

In the case of a perfect rectangular spray pattern (as depicted in Figure 3), the application rate corresponds to the rated dosage on the entire sprayed area. Consequently, the proportion of weed surface area on which the application rate is higher than 85% of the rated dosage ($A_{W>85\%}/A_W$) is 100% and the proportion of sprayed surface receiving less than half the rated dosage ($A_{S<50\%}/A_S$) is 0%. In the case of a triangular spray pattern, which is more representative of conventional flat-fan nozzles, the sprayed area is necessarily wider than the one obtained with a theoretical rectangular pattern and the application rate is lower than the rated dosage in the case of limited or no overlapping. This is illustrated in Figure 5.

Thus, in the case of small patches (l_e=1 m, w_e=0.2 m), a significant proportion of weed area can receive an insufficient rate to ensure weed control (Figure 5) while a significant proportion of a sprayed area presents a low application rate that can yield to resistance condition development. For example, with large triangular spray pattern (i.e. 1.5 m), the proportion of weed area on which

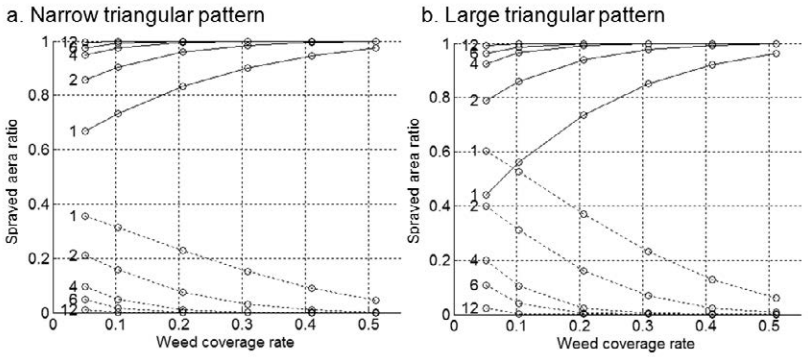


Figure 5. Analyse of the sprayed area with respect to the weed coverage rate, in the case of small patches (l_e=1 m, w_e=0.2 m) and triangular spread patterns with a spray width of 1 m (left) and 1.5 m (right): $S_{W>85\%}/S_W$ (continuous line), $S_{S<50\%}/S_S$ (dashed line). The section width (expressed in number of nozzles) is indicated on the left of the curves.

the application rate is higher than 85% of the rated dosage ($A_{W>85\%}/A_W$) can drop under 50% when nozzles are controlled individually. Considering the same spraying situation, the proportion of sprayed area receiving less than half the rated dosage ($A_{S<50\%}/A_S$) reaches 60%.

Results on an example of actual infestation map

Simulations of targeted applications with various section widths had also been carried out with the infestation map derived from aerial images of a sugar beet field. Findings, presented in Figure 6, confirm results and trends already observed with virtual maps. The shape properties of the three spray patterns used in this study leads to the same herbicide savings. Thus, the ratio Q_S/Q_T only depends on the section widths. This ratio is obviously minimal when nozzles are controlled individually but it is worth highlighting that splitting the 36 m boom into 8 independent sections of 9 nozzles (i.e. section working width of 4.5 m) already drops the herbicide use to 40% of the amount required for full application. Significant proportion of weed and sprayed areas is exposed to under-applications when narrow sections are used (Figure 6).

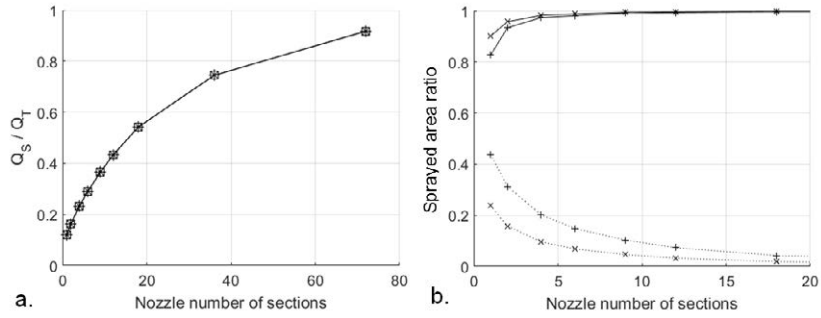


Figure 6. (A) Sprayed amount ratio with respect to the section width (expressed in nozzle number) for three spray patterns: rectangle of 0.5 m width (*), triangle of 1 m width (×) and triangle of 1.5 m width (+). (B) Sprayed area ratio $A_{W>85\%}/A_W$ (continuous line) and $A_{S<50\%}/A_S$ (dashed line) for triangular spray pattern of 1 m (×) and 1.5 m (+) width.

In this study, the longitudinal rate transitions (i.e. in the travelled direction), due to non-instantaneous on-off switching of sprayer sections are not taken into account. Moreover infestation maps are assumed to provide the exact location of weeds present in the field. Taking into account rate transitions and potential weed positioning errors should increase the chemical amount ratio Q_S/Q_T presented in this study. This should also lead to decreases in the weed area correctly controlled and increase the area exposed to sub-lethal rate especially in the case of narrow spray sections or individual nozzle control.

Conclusions

The use of virtual weed infested maps and spraying simulations appears as a relevant way to study factors that affect herbicide savings in site-specific weed control. The study demonstrated that the herbicide reduction achieved by the width reduction of independent application sections is highly affected by the infestation rate and the weed spatial distribution (i.e. patch sizes). Consequently, the choice of very narrow application sections can be relevant in the case of small patches randomly distributed, but moderate width sections also appear as a good choice when weeds are aggregated in larger patches. Otherwise the study demonstrates that in the case of small patches and very narrow sections (i.e. composed of one or few nozzles) the emphasis should also be put on the sprayed area exposed to sub-lethal rate where weed-control failures or herbicide-resistance development could occur.

Acknowledgements

This project was funded by the National Superior Institute of Agronomic Sciences, Food and Environment (AgroSup Dijon), the Agroecology Joint Research Unit (Agroécologie, AgroSup Dijon, INRA, Univ. Bourgogne Franche-Comté) and the French National Research Institute of Science and Technology for Environment and Agriculture (Irstea). Authors also thank AIRINOV Company for providing the aerial image of the sugar beet field presented in this paper.

References

Franco, C., Pedersen, S.M., Papaharalampos, H. and Ørum, J.E., 2017. The value of precision for image-based decision support in weed management. Precision Agriculture 18(3), 366-382.

Gerhards, R. and Oebel, H., 2006. Practical experiences with a system for site-specific weed control in arable crops using real-time image analysis and GPS-controlled patch spraying. Weed Research 46(3), 185-193.

Gutjahr, C., Sokefeld, M. and Gerhards, R., 2012. Evaluation of two patch spraying systems in winter wheat and maize. Weed Research 52(6), 510-519.

Lee, W.S., Slaughter, D. and Giles, D., 1999. Robotic weed control system for tomatoes. Precision Agriculture 1(1), 95-113.

Louargant, M., Jones, G., Faroux, R., Paoli, J.-N., Maillot, T., Gée, C. *et al.,* 2018. Unsupervised Classification Algorithm for Early Weed Detection in Row-Crops by Combining Spatial and Spectral Information. Remote Sensing 10(5), 761.

Nordmeyer, H., 2006. Patchy weed distribution and site-specific weed control in winter cereals. Precision Agriculture 7(3), 219-231.

Stafford, J.V. and Ambler, B., 1994. In-field location using GPS for spatially variable field operations. Computers and Electronics in Agriculture 11(1), 23-36.

Timmermann, C., Gerhards, R. and Kühbauch, W., 2003. The Economic Impact of Site-Specific Weed Control. Precision Agriculture 4(3), 249-260.

Weis, M., Gutjahr, C., Ayala, V.R., Gerhards, R., Ritter, C. and Schölderle, F., 2008. Precision farming for weed management: techniques. Gesunde Pflanzen 60(4), 171-181.

Section 5
Proximal and remote sensing of soil and crop

Design of a portable sensor suite for real-time monitoring of crop water stress index in maize breeding plots

O.E. Apolo-Apolo*, M. Pérez-Ruiz, J. Martínez-Guanter and G. Egea
Aerospace Engineering and Fluids Mechanics Department, University of Seville, Ctra. Sevilla-Utrera km 1, 41013 Seville, Spain; eapolo@us.es

Abstract

Field-based high throughput phenotyping techniques using remote sensors mounted on terrestrial platforms (HTPPs) have emerged as a promising tool for plant breeders. Among the several phenotypes and crop characteristics that can be monitored, crop temperature is a variable of great interest as it provides indirect information on crop water status. However, its absolute value is affected by the prevailing environmental conditions. There has been great interest in developing thermal indices, such as crop water stress index (CWSI) needed to normalize the crop temperature against the changing climatic conditions. Therefore, the aim of this research was to test the feasibility of new artificial wet and dry reference surfaces that would be suitable to be mounted on HTPPs for CWSI estimation. Paper-made hemispherical artificial surfaces were constructed using three RGB colours and three replications per colour for each artificial surface (wet and dry). Plastic-made structures printed in the lab with a 3D printer were used to hold and to moisten the hemispherical paper-based surfaces. Temperatures of artificial wet and dry hemispherical surfaces were evaluated against well-watered maize canopies and senescent maize plants, respectively.

Keywords: plant phenotyping, drought, water stress, canopy temperature

Introduction

Although tolerance levels and crop responses to drought stress can be different depending on the plant species and the environmental conditions, plant breeding programs are mainly focused on selecting phenotypes with enhanced abiotic and biotic stress resistance and, especially, on identifying cultivars with high yields under water stress conditions (Bita and Gerats, 2013). However, the production of new cultivars with increased yield capacity in water stressed environments requires an evaluation under field conditions of the genotypes previously selected in the laboratory (Bita and Gerats, 2013). These field trials, in which a large number of cultivars are screened, are time-consuming and frequently use crop yield as the only variable for selecting the best crop varieties (Makanza *et al.,* 2018).

In this context, field-based high throughput phenotyping techniques using remote sensors mounted on either aerial or terrestrial platforms have emerged as a promising tool for plant breeders (Andrade-Sanchez *et al.,* 2014; White *et al.,* 2012). Among the various phenotypes and crop characteristics that can be monitored, crop temperature is a variable of great interest to evaluate the crop response to water stress (Jones *et al.,* 2018; Martínez *et al.,* 2017). Canopy temperature provides indirect information on crop water status, as it is tightly related to crop transpiration through the evaporative cooling effect of the latter. In spite of the interest in this variable, its absolute value is very uninformative due to the great influence that the environmental conditions (e.g. solar radiation, air temperature, relative humidity, etc.) have on crop temperature (Leinonen and Jones, 2004; Payero and Irmak, 2006). This is why since the early 1980s, there has been great interest in developing thermal indices to normalize the crop temperature against the changing climatic conditions (Maes and Steppe, 2012). The crop water stress index (CWSI) has been the most widely used.

DOI 10.3920/978-90-8686-888-9_30, © Wageningen Academic Publishers 2019

This index is calculated from canopy-air temperature differences and from two reference baselines (wet and dry baselines) that are empirically determined for each crop species (Idso *et al.,* 1981). The uncertainty of using existing reference baselines for CWSI computation is high, as radiation level and solar height affect the baselines shape and canopy temperature must be measured under the same solar radiation conditions of those prevailing when the reference baselines were determined (Maes and Steppe, 2012). In order to deal with this constraint, some scientists have proposed the use of artificial surfaces to be used as wet (e.g. leaves sprayed with water) and dry (e.g. leaves covered with petroleum jelly) references to be measured at the time that canopy temperature is also measured (Jones *et al.,* 2018; Meron *et al.,* 2010). This methodology would not be feasible to be used with terrestrial field-based high-throughput phenotyping platforms, as it would not be practical to establish wet and dry references in each plot of a plant breeding program (Bellvert *et al.,* 2014). Therefore, the aim of this research was to test the feasibility of new artificial wet (WR) and dry (DR) reference surfaces that would be suitable for mounting on terrestrial phenotyping platforms for CWSI estimation in maize breeding programs.

Materials and methods

Evaluation of materials and colours for the reference surfaces

Plant leaves have a specific architecture to meet a good transpiration process (Köstner, 2001). Finding artificial surfaces that reproduce maize leaf behaviour under dry and wet reference conditions requires a number of preliminary tests to get the right surface thickness and colour. Among the various materials assessed (wool, cotton, cloth, paper, etc), paper was selected (Figure 1) because of its ability to stay wet under field conditions. Regarding the colour of the reference surfaces, three RGB colours per artificial surface (wet and dry) were assessed (Figure 1). Spectral reflectance measurements were performed with a portable spectrometer (Unispec DC, PP Systems, Haverhill, MA, USA) in each colour and compared to spectral reflectance measurements of healthy maize leaves.

Shape design of artificial wet and dry reference surfaces

Several approaches to obtaining the reference temperatures for CWSI calculation have been proposed in recent decades (Maes and Steppe, 2012). Those who propose the use of artificial reference surfaces use flat reference surfaces (Jones *et al.,* 2018). These kind of surfaces do not represent the range of radiation conditions experienced in a crop canopy (Jones *et al.,* 2009). In this study, dome-shaped artificial surfaces were constructed using 0.001 m^2 of paper. Moreover, plastic-made structures were built with a 3D printer to hold and to moisten the hemispherical paper-based surfaces.

Data acquisition system for temperature measurement

The dome temperature was measured using two inexpensive infrared thermal sensors (MLX90614 from Melexis, Ypres, Belgium): (1) a downward-facing sensor (10° field of view) to measure the outdoor temperature and; (2) an upward-facing sensor (90° field of view) to measure the indoor artificial surface (Figure 2). Melexis infrared thermal sensors have a resolution of 0.02 °C and an accuracy of ±0.5 °C from 0 to 50 °C. The sensors used an Arduino-based platform with a RTC (DS3231) and MicroSD (40811_M000523) shield for data collection (Figure 2).

Crop canopy temperature measurements were collected by means of two infrared thermometers (IR120, Campbell Scientific Ltd., Shepshed, UK) connected to a CR1000 data logger (Campbell Scientific Ltd., Shepshed, UK). Both thermometers were deployed above the canopy on a 4 m mast. Air temperature, relative humidity and vapor pressure were measured with a sensor (ATMOS 14, Decagon, Pullman, WA; USA). All data were recorded every 10 min.



Figure 1. RGB colour values. RGB (32, 66, 34) 01 colour, RGB (58, 14, 81) 02 colour and RGB (159, 173, 88) 03 colour.

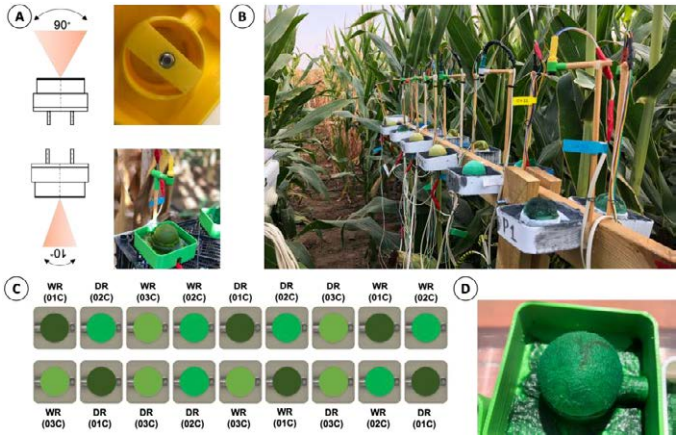


Figure 2. Melexis sensor with both field of views (A), artificial references in the field (B), illustration of the randomized design (C) and detail of wet dome (D).

Field site and experimental design

The trial was conducted in an experimental maize field (*Zea mays* L.) during the 2018 growing cycle. The field was located 10 km from Seville (lat. 37°27'N, long. 5°58'O). The climate of the study area is Mediterranean with an average rainfall and average annual temperature of 565 mm and 18.5 °C, respectively. To assess the behaviour of the artificial reference surfaces, a random design with three replications per RGB colour was made. The wet reference surfaces were kept wet by a water tank that fed the paper domes through a wick made of the same kind of material. Artificial wet and dry surfaces were evaluated against canopy temperature measurements of well-watered and senescent maize plants, respectively. The measurements in well-watered leaves were made for 20 days and on senescent leaves for 5 days. However, in this paper only DOYs (213, 214, 215, 245, 246 and 247) respectively are published, since in these days it is certain that leaves are completely green in the well-watered treatment.

Results and discussion

Characterization of spectral reflectance of artificial reference surfaces and maize canopy

Different surface features objects reflect or absorb the sun's electromagnetic radiation in different ways. The reflectance properties of a surface depend on the material and its physical and chemical

composition, the surface roughness as well as its geometry. These differences in reflectance make it possible to identify different materials by analyzing their spectral reflectance signatures. In Figure 3, spectral reflectance of the artificial surfaces assessed (both under wet and dry conditions) and the crop canopy (well-watered and senescent) are shown.
The three colours evaluated generated spectral differences of reflected radiation. However, 01 Colour (WW) presented the greatest similarity with the spectral signature of a well-watered maize canopy. In the case of dry reference surfaces, the 03 Colour (WS) had a similar behaviour to a senescent maize canopy.

Correlation between external and internal temperatures on hemispherical surfaces
To decide which temperature (internal or external) represents better crop temperature, both temperatures were measured and compared (Figure 4). Preliminary results showed that both temperatures are highly correlated (R^2: 0.99 and R^2: 0.97) for both (wet and dry) artificial references, respectively. In order to prevent the hemispherical surfaces from possible shading of the external IR thermometers, the option of using the dome internal temperature seems to be more appropriate than measuring the external one.

Performance of artificial reference surfaces to represent maize canopy temperatures
The correlation between artificial reference surfaces and maize canopy temperatures are shown in Figure 5. The wet reference surfaces showed a high degree of correlation (R^2 values over 0.90) with

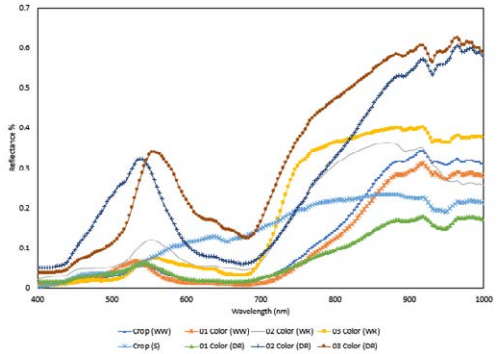


Figure 3. The spectral reflectance curves for each colour and treatment.

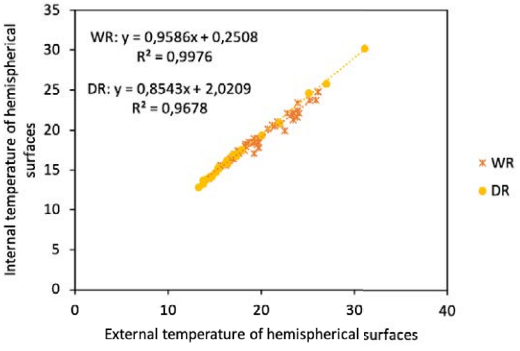


Figure 4. Relationship between external and internal temperatures of the hemispherical surfaces.

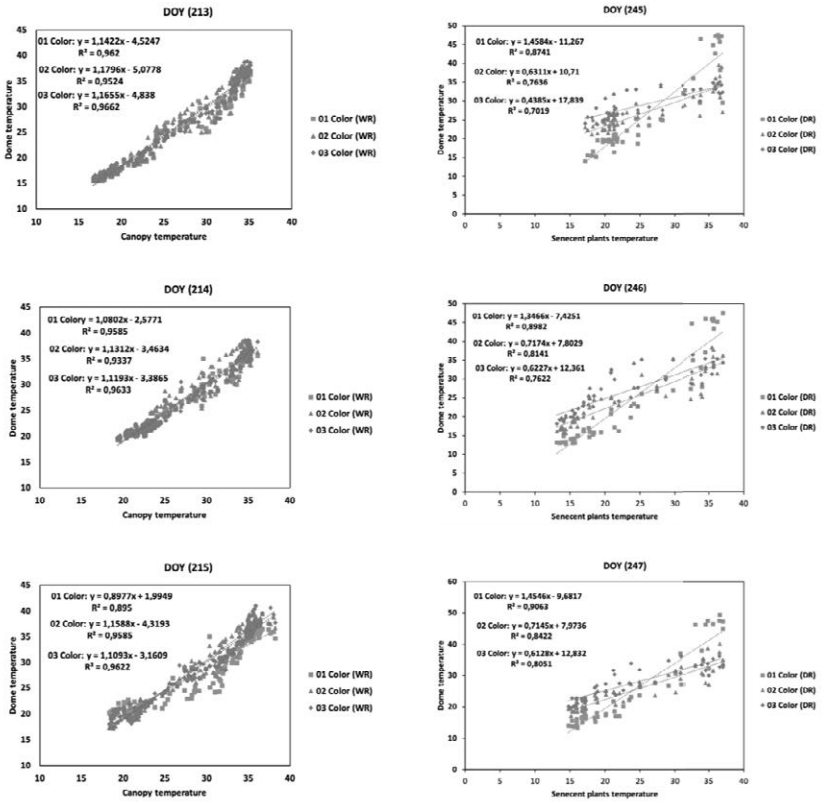


Figure 5. Relationships between artificial reference temperatures and crop temperatures. (left panels) Wet reference surface against well-watered maize canopy; (right panels) Dry reference surface against senescent canopy. Half-hourly data for three complete summer days are shown in both cases.

canopy temperature of well-watered maize plants. No clear differences between surface colours were observed in this relationship. Dry reference surfaces showed also a promising potential to represent the temperature of non-transpiring maize canopies (R^2 values over 0.7), as previously pointed out by Jones *et al.* (2018). However, surface colour had a higher impact on this relationship, 02 colour being the colour that best represented the temperature of wilting maize canopies. Despite these promising results, further research is needed to evaluate the performance of the hemispherical surfaces under a broader range of environmental conditions (solar radiation, vapor pressure deficit, etc.).

Conclusions

The artificial surfaces developed in this work to simulate the thermal behaviour of well irrigated and non-transpiring maize canopies have shown promising results during the sampled days. Of the three colours evaluated, colour 02 showed the best performance to represent both, 'wet' and 'dry' crop conditions. Further research is needed to assess the performance of the hemispheric surfaces under a broader range of environmental conditions.

Acknowledgements

This work was made possible through the support of the Spanish Ministry of Economy and Competence (PROJECT: AGL2016-78964-R). Moreover, the authors want to thank the Predoctoral Research Fellowship granted to the first author by the University of Seville (call IV.3 2017 of its own RDI program).

References

Andrade-Sanchez, P., Gore, M. A., Heun, J. T., Thorp, K. R., Carmo-Silva, A. E., French, A. N., *et al.,* 2014. Development and evaluation of a field-based high-throughput phenotyping platform. Functional Plant Biology, 41 (1), 68-79.

Araus, J. L., Slafer, G. A., Royo, C., & Serret, M. D., 2008. Breeding for yield potential and stress adaptation in cereals. Critical Reviews in Plant Sciences, 27 (6), 377-412.

Bellvert, J., Marsal, J., Girona, J., & Zarco-Tejada, P. J., 2014. Seasonal evolution of crop water stress index in grapevine varieties determined with high-resolution remote sensing thermal imagery. Irrigation Science, 33 (2), 81-93.

Bita, C. E., & Gerats, T., 2013. Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops. Frontiers in Plant Science, 4(July), 1-18.

Idso, S. B., Jackson, R. D., Pinter, P. J., Reginato, R. J., & Hatfield, J. L., 1981. Normalizing the stress-degree-day parameter for environmental variability. Agricultural Meteorology, 24(C), 45-55.

Jones, H. G., Hutchinson, P. A., May, T., Jamali, H., & Deery, D. M., 2018. A practical method using a network of fixed infrared sensors for estimating crop canopy conductance and evaporation rate. Biosystems Engineering, 165, 59-69.

Jones, H. G., Serraj, R., Loveys, B. R., Xiong, L., Wheaton, A., & Price, A. H., 2009. Thermal infrared imaging of crop canopies for the remote diagnosis and quantification of plant responses to water stress in the field. Functional Plant Biology, 36 (11), 978-989.

Köstner, B., 2001. Evaporation and transpiration from forests in Central Europe-relevance of patch-level studies for spatial scaling. Meteorology and Atmospheric Physics 76, 69-82.

Leinonen, I., & Jones, H. G., 2004. Combining thermal and visible imagery for estimating canopy temperature and identifying plant stress. Journal of Experimental Botany, 55 (401), 1423-1431.

Maes, W. H., & Steppe, K., 2012. Estimating evapotranspiration and drought stress with ground-based thermal remote sensing in agriculture: a review. Journal of Experimental Botany, 7 (3), 20-23.

Makanza, R., Zaman-Allah, M., Cairns, J. E., Magorokosho, C., Tarekegne, A., Olsen, M., *et al.,* 2018. High-throughput phenotyping of canopy cover and senescence in maize field trials using aerial digital canopy imaging. Remote Sensing, 10 (2) 330.

Martínez, J., Egea, G., Agüera, J., & Pérez-Ruiz, M., 2017. A cost-effective canopy temperature measurement system for precision agriculture: a case study on sugar beet. Precision Agriculture, 18 (1), 95-110.

Meron, M., Tsipris, J., Orlov, V., Alchanatis, V., & Cohen, Y., 2010. Crop water stress mapping for site-specific irrigation by thermal imagery and artificial reference surfaces. Precision Agriculture, 11 (2), 148-162.

Payero, J. O., & Irmak, S., 2006. Variable upper and lower crop water stress index baselines for corn and soybean. Irrigation Science, 25 (1), 21-32.

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., *et al.,* 2012. Field-based phenomics for plant genetics research. Field Crops Research, 133, 101-112.

Characterisation of fungal diseases on winter wheat crop using proximal and remote multispectral imaging

R. Bebronne[1], A. Michez[2], V. Leemans[1], P. Vermeulen[3], B. Dumont[4] and B. Mercatoris[1*]
[1]*Biosystems Dynamics and Exchanges, TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech, Liège University, Passage des Déportés, 2, 5030, Gembloux, Belgium; benoit.mercatoris@uliege.be*
[2]*Forest Management, TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech, Liège University, Passage des Déportés, 2, 5030, Gembloux, Belgium*
[3]*Walloon Agricultural Research Centre, Valorisation of Agricultural Products Department, Food and Feed Quality Unit, Chaussée de Namur, 24, 5030, Gembloux, Belgium*
[4]*Plant Sciences, TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech, Liège University, Passage des Déportés, 2, 5030, Gembloux, Belgium*

Abstract

Winter wheat fungal diseases, responsible for high yield losses, can be assessed by computer vision to increase phenotyping performance. This study aims to compare multispectral imagery based on remote and proximal sensing for disease detection. Wavelength selection was achieved by ANOVA and stepwise regression. Prediction of disease severity was performed by means of an artificial neural network based on proximal sensing data. Septoria tritici blotch (STB) requires proximal measurements, but stripe and brown rusts can be detected from UAVs and from the ground. Prediction results obtained gave R^2 of 0.55 and 0.57 for STB and stripe rust respectively.

Keywords: wheat, fungal diseases, multispectral, unmanned aerial vehicle sensing, proximal sensing

Introduction

Recent study indicates Belgium is the fifth biggest consumer of phytopharmaceutical products for crop management (Service Public de Wallonie, 2018). The use of pesticides has been reduced by half compared to 1995 but has remained stable since 2010 and shows no sign of further decrease (Service Public de Wallonie, 2018). Pesticides are used to prevent the infection of stripe rust, brown rust and Septoria tritici blotch (STB) which are common diseases of winter wheat caused respectively by *Zymoseptoria tritici*, *Puccinia striiformis* and *Puccinia recondita* (Bodson *et al.*, 2017). The strategy to reduce pesticide application in wheat management relies on: (1) the selection and evaluation of varieties resistant to biotic stress; and (2) the spraying regulation depending on severity of disease infection. Conventional assessment is based on visual or manual observations, but these methods suffer from limited phenotype data sets and bias between assessors. These problems can be resolved by use of automated phenotyping platform which can be exploited both for agronomic trial and production fields (Fiorani and Schurr, 2013).
Developing a phenotyping device for early disease detection and to quantify plant health status for wheat has been reported in the literature by means of spectral imaging which has been identified as a key technology for plant phenotyping (Li *et al.*, 2014). In 2004, Moshou *et al.,* detected stripe rust in wheat by means of proximal sensing with high success rate, highlighting wavebands centred on 543 nm, 630 nm, 750 nm and 861 nm. By using a spectroradiometer in proximal sensing, Krishna *et al.* (2014) developed a model able to quantify stripe rust severity in wheat. For STB, Yu *et al.* (2018) accurately discriminated healthy from infected wheat based on hyperspectral proximal measurements.
Most disease detection studies focus on proximal measurements of reflectance in the field (Bravo *et al.*, 2003; Odilbekov *et al.*, 2018) or in a laboratory (Devadas *et al.*, 2009) but few of them consider

DOI 10.3920/978-90-8686-888-9_31, © Wageningen Academic Publishers 2019

aerial spectral measurements for disease detection (Nebiker *et al.*, 2016; Huang *et al.*, 2007) which is more adequate for large-scale area such as production field. Moreover, disease detection studies are generally based on artificial disease inoculation leading to field infection occurrence at early growing stages. A combination of stresses can affect winter wheat simultaneously and belatedly in the growing season, complicating spectrally-based detection efforts.
This study aimed to integrate ground-based and aerial multispectral approaches to characterise the three major fungal diseases affecting winter wheat, i.e. stripe rust, brown rust and STB, under natural conditions. The goal was to select the wavebands which could allow the differentiation between healthy and diseased wheat for the two scales and to investigate their potential for disease severity regression on proximal sensing data.

Materials and methods

Field experiment

The experiments took place in fields dedicated to larger agronomic trials on winter wheat, *Triticum aestivum* L., designed in microcrops (1.8×6 m) to study the effects of wheat variety, fungal treatment and sowing date on crop yield. The experiments were monitored during the 2017 growing season by aerial remote sensing and during the 2018 growing season by proximal sensing. Both fields were located in Lonzée, Belgium (50° 32' 58" N and 4° 44' 08" E). In 2017, the approach based on unmanned aerial vehicle (UAV) focused on six varieties (Albert, Anapolis, Edgar, Mentor, Reflection and RGT Reform) sown on October 25 under two types of fungal treatments, i.e. no treatment (0T) and two treatments (2T) on May 10 and May 31. Three replications were studied. The experiment of 2018 involved five varieties (Alcides, Benchmark, Edgar, RGT Reform and Triomph) sown on November 16 under 0T and 2T on May 8 and May 24 with four replications. For each experiment, the varieties were chosen based on their disease sensitivity to allow measurement of different levels of infection.

Disease visual assessment

Severities of STB, stripe rust and brown rust were visually assessed at the leaf level. Rust score is a degree of severity based on the Cobb's modified scale (Peterson *et al.*, 1948). STB severity represents the proportion of the leaf affected by symptoms. Five plants were collected for each microcrop. Disease score for each microcrop was the mean of leaf disease score. In 2017, 2T microcrops were not visually assessed and were considered as healthy (disease severity=0) due to sufficient application of pesticides.

Aerial sensing acquisition, image treatment and statistical analysis

An octocopter UAV (X frame type) was flown over the 2017 agronomic trial on June 1, June 15 and July 7 at 100 m above ground level (AGL) at a speed of 5 m/s. The Parrot Sequoia camera (Parrot SA, Paris, France) covering green (550 nm), red (660 nm), red edge (735 nm) and near infrared (790 nm) wavelengths was mounted on the UAV. The full width at half maximum was 40 nm for red, green, near infrared (NIR) and 10 nm for the red edge. The sensor allowed the computation of true reflectance imagery by means of an upward looking sunshine sensor measuring incident lighting and the use of grey calibration plate before each flight survey. Pix4D suite 3.1 (Pix4D SA, Lausanne, Switzerland) was used to perform a photogrammetric 3D reconstruction of the multispectral imagery using the 'Ag Multispectral' predefined workflow. The images acquired by the multispectral sensor were used to derive orthomosaic reflectance maps (0.1 m ground sampling distance, GSD) in the four wavelengths (green, red, near-infrared and red edge). Five white targets were precisely georeferenced using precision differential GPS (Real Time Kinematic GPS, ±0.03 m mean XYZ accuracy) in order to ensure geometrical and positioning accuracy of the reflectance maps.

At each date, averaged reflectance values were compared between 0T and 2T microcrops. Mean reflectance comparison was achieved by analysis of variance (ANOVA) carried out for each waveband on each date separately. A Kolmogorov-Smirnov test was applied on the samples to guarantee their normality.

Proximal sensing acquisition, image treatment and disease prediction

A multispectral acquisition prototype acquired images at 1 m above the canopy. The system included a monochromatic 12 bit 1.3 Mpixel camera CMOS BCi5 (C-Cam technologies, Leuven, Belgium) combined with optical filters placed on a motorised wheel to place each filter successively in front of the camera (Figure 1). Eleven wavebands were measured centred from 450 to 950 nm by step of 50 nm and with a bandwidth of 40 or 50 nm. To take into account the variations of illumination, each image was normalised by adapting the integration time of the camera based on the reflectance measured on a silicon white reference placed in the field of view of the camera.

For a scene, 11 images were acquired and overlaid to obtain a single multispectral image on which plant-soil segmentation was performed. Since the image normalisation only considered global illumination, an artificial neural network (ANN) was trained to account for light contrast inside the scene of interest by classifying pixels under shady, sunny or cloudy conditions (Figure 2).

The mean reflectance was computed for the four groups of plant pixels (mean sunny, mean shady, mean cloudy and mean of total plant pixels). Texture analysis was performed by means of co-occurrence matrix on all plant pixels from which four parameters were extracted: correlation, contrast, energy and homogeneity (Baraldi and Parmigianni, 1995). Considering the 11 layers of a multispectral image, a scene was then characterised by 88 features. A bidirectional stepwise regression was used to establish a model involving the most discriminative features and allowing waveband selection. Based on this feature selection, an ANN was trained and validated using 80% of the dataset and the model was tested on the remaining 20% images. This machine learning method was chosen for its capacity to resolve complex problems and to deal with a small amount of data (Hornik *et al.*, 1989).



Figure 1. Proximal multispectral acquisition system. On the left, the camera and the motorised filter wheel. On the right, the full platform supporting the acquisition system.

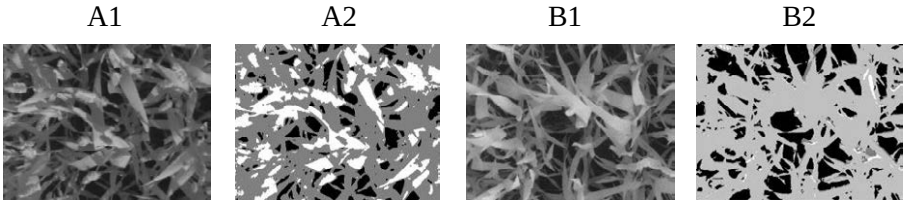


Figure 2. Grey images of winter wheat acquired (A1, B1) and the same images after segmentation (A2, B2) in four classes: soil (black), shady (dark grey), sunny (white), cloudy (light grey).

Results and discussion

Aerial sensing approach

On June 1, 2017, due to absence of rust on all 0T microcrops except the variety Reflection infected by stripe rust, the other microcrops were only infected by septoria with severities ranging from 4 to 13%. The statistical analysis performed showed no significant mean difference between 0T and 2T (healthy) microcrops (Table 1). STB emerges on the lower leaves (Bahat *et al.*, 1980) which makes it harder to detect from above. Indeed, the aerial nadir perspective induces occlusion on infected leaves below the crop canopy. In order to peek through the canopy and collect unmixed data from lower leaves, the spatial resolution needs to be finer which raises interest for proximal detection of STB.

On June 15, all five varieties were healthy regardless of their treatment. The lowest leaves were not taken into account due to their necrosis. However, ANOVA highlighted, by increasing *P*-value, red edge, NIR, green and red as significant (Table 1). Even though STB was absent of the crop, the changes in reflectance values appeared due to accelerated plant senescence originating from previous STB infection. The detection of STB by UAVs is delayed in time.

Between June 15 and July 7, brown rust developed on 0T microcrops up to severities ranging from 76 to 97%. On July 7, differences between 0T and 2T means were more significant than on June 15 for all wavebands (Table 1). This level of significance is probably due to the combination of the damage caused by septoria earlier in the growing season and the strong infection by brown rust. The most significant difference was still observed in the red edge band. However, the red band which showed the least significance on June 15 was the second most significant on July 7 due to symptoms of brown rust characterised by brown to orange spores.

Variety Reflection was reported separately from the other varieties as it was the only variety infected by stripe rust during the acquisition period (31, 54 and 93% respectively on the three dates of acquisition). It was also infected by septoria but only on June 1 with 21% severity. It was shown that septoria induced no significant mean difference in early infection. However, the ANOVA applied

Table 1. *P*-values of the mean reflectance difference between infected and healthy microcrops for the three dates and varieties Albert, Anapolis, Edgar, Mentor and RGT Reform. Interaction between variety and health status being significant on July 7 for green waveband, analysis of variance has been split into five for every variety.

P-value	June 1	June 15	July 7
Green	0.24	0.018	[0.045-6.95E-07]
Red	0.28	0.033	3.18E-06
Red edge	0.18	0.0002	1.07E-06
Near infrared	0.98	0.0025	4.85E-06

Table 2. *P*-values of the mean reflectance difference between infected and healthy microcrops for the three dates and variety Reflection only.

P-value	June 1	June 15	July 7
Green	0.094	0.029	0.0077
Red	0.0059	0.0068	0.0054
Red edge	0.23	0.021	4.94E-05
Near infrared	0.0014	0.015	7.77E-04

on June 1 highlights NIR (P=0.0014) and red (P=0.0059) as significantly different for the Reflection variety (Table 2). This difference can be attributed to stripe rust infection. From June 1 to July 7, stripe rust severity kept growing but significance of the red waveband remained stable while the green and red edge bands became increasingly significant. It was hypothesised that stripe rust infection can be detected early due to its yellow/orange spores reflecting light in the red band. Necrosis and cellular structure modifications induced by the fungus lead to changes in NIR (P=7.77E-04) and red edge (P=4.94E-05) bands where differences were the most significant on July 7.

Proximal sensing approach

An increase in spatial resolution by means of proximal sensing allowed severity regression based on multispectral imagery (Figure 3). Wavelengths of interest (Table 3) were located in the green (500 and 550 nm) and in the red/red edge (650 and 700 nm) bands. Wavelengths in the NIR band were excluded from the model despite being identified as key wavelengths for septoria detection due to the influence of water in this band (Yu *et al.*, 2018). Prediction tends to underestimate STB severity especially for highly infected microcrops. Images at high severity were acquired on the last acquisition date (June 21) at growth stage 75 (Zadoks *et al.*, 1975) where infected lower leaves were obstructed due to crop density. To improve the regression, shady features could help to detect these lower leaves. However, these features were excluded from the model since at advanced stages almost all the leaves are shaded except the flag leaves. The isolation of infected lower leaves could be assessed by means of sensors measuring depth like stereoscopic systems.

These variables are used by the ANN for a regression characterised by a predicted R^2 of 0.55 and a RMSE of 4.35% (Figure 3).

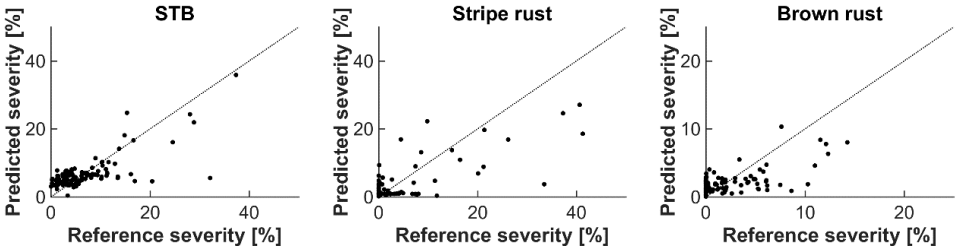


Figure 3. Regression results for STB, stripe rust and brown rust between predicted and measured severity.

Table 3. Features selected by stepwise regression for STB, stripe rust and brown rust detection sorted by increasing order of significance.

STB		Stripe rust		Brown rust	
Parameters	λ (nm)	parameters	λ (nm)	parameters	λ (nm)
Contrast	700	energy	700	correlation	800
Homogeneity	650	homogeneity	500	homogeneity	600
Contrast	550	energy	600	homogeneity	650
Contrast	500	mean sunny	850	homogeneity	450
Energy	650	mean sunny	900	energy	800
Mean total	700			energy	950

Wavelengths entered in the stepwise model for stripe rust detection in this work cover the spectrum used (Table 3). The waveband centred on 600 nm corresponds to the yellow/orange colour of stripe rust spores. In proximal sensing, Moshou *et al.* (2004) identified wavebands at 543±10 nm and 861±10 nm as the most discriminating by stepwise regression while Krishna *et al.* (2014) identified by principal component analysis (PCA) wavelengths at 428 nm and 672 nm. In this study, wavebands centred on 500 nm and 700 nm were the two most discriminative, 850 nm was also highlighted. There seems to be no consensus on proper waveband selection which implies that wavelengths selected are dataset specific. However, it is important to note that stepwise regression and PCA do not include highly correlated variables in the model which is the case for wavebands being close to each other. These variables were used by ANN to give regression results of predicted R^2 of 0.57 and predicted RMSE of 5.37% (Figure 3).

Brown rust detection is assessed mainly based on textural features computed from wavelengths in the NIR and in the red bands (Table 3) corresponding respectively to the physiological changes induced by the disease and the colour of the brown rust spores. Reflectance at wavelengths from 550 to 700 nm increased from the first days of inoculation (Kuska *et al.*, 2015). In the literature, Leaf Rust Disease Severity Index 2 (LRDSI2) developed by Ashourloo *et al.* (2014) reached an accuracy of 86.5% for leaf rust detection and required wavebands centred in 455 nm and 605 nm, which were both part of the model in this work.

Brown rust remained in 2018 at very low severity levels (<13%) which gave results of lower quality with a predicted R^2 of 0.41 and RMSE of 2.40% (Figure 3). This disease should be studied on a highly infected field in order to observe a wider range of infection. The co-development of the two rusts would be very challenging to discriminate.

Conclusions

This study compared multispectral imagery based on remote and proximal sensing for winter wheat fungal disease detection. The remote approach remains limited to late STB detection. In contrast, proximal sensing improved detection using wavelengths in the red, red edge and green bands. For stripe rust, the stepwise regression model highlighted NIR (850 nm and 900 nm) and red (600 nm and 700 nm) wavebands which were also retained with significant mean difference for UAV-based approach. The coherence of wavelengths between remote and proximal sensing along with the very significant NIR and red mean difference in UAV sensing at 31% severity fosters the potential for stripe rust detection. For brown rust, no specific wavebands of interest were clearly identified by ANOVA in UAV-sensing, making the link between both scales difficult. As for stripe rust, UAV-sensing showed great potential for brown rust detection and should be tested on wheat infected at lower severity. In proximal sensing, red and NIR were significant but brown rust infection was too low to obtain good regression results.

UAV imagery, due its ability to detect rusts should be used for management as the larger resolution allows a quick overview of fields. The measurements at leaf level by proximal sensing could be used in plant phenotyping experiments to have access to information about STB and to go further than a simple detection for the three diseases. Proximal sensing, by leaf segmentation and texture analysis, has potential for severity prediction of septoria and stripe rust. The results can be used for health status discrimination but accurate quantification would require an improvement of the multispectral imaging device, especially by setting up a co-registration of the images.

Acknowledgements

This research has been performed in the framework of PHENWHEAT project financially supported by the Agriculture, Natural Resources and Environment Research Direction of the Public Service of Wallonia, Belgium (Project D31-1385).

References

Ashourloo, D., Mobasheri, M. R., & Huete, A., 2014. Developing two spectral disease indices for detection of wheat leaf rust (Puccinia triticina). Remote Sensing 6 4723-4740.

Bahat, A., Gelernter, I., Brown, M. B. & Eyal, Z., 1980. Factors affecting the vertical progression of Septoria leaf blotch in short-statured wheats. Phytopathology 70 (3) 179-184.

Baraldi, A. & Parmiggiani, F., 1995. An investigation of the textural characteristics associated with Gray Level Co-occurrence Matrix statistical parameters. IEEE Transactions on Geoscience and Remote Sensing 33 (2) 293-304.

Bodson B., De Proft M. & Watillon B., 2017. Livre Blanc Céréales. Edition février 2017 (White paper cereals. February edition 2017).

Bravo, C., Moshou, D., West, J., McCartney, A., & Ramon, H., 2003. Early disease detection in wheat fields using spectral reflectance. Biosystems Engineering. 84 (2), 137-145.

Devadas, R., Lamb, D. W., Simpfendorfer, S., & Backhouse, D., 2009. Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves. Precision Agriculture 10 459-470.

Fiorani, F., & Schurr, U., 2013. Future Scenarios for Plant Phenotyping. Annual Review of Plant Biology 64 267-291.

Hornik K., Stinchcombe M. & White H., 1989. Multilayer feedforward networks are universal approximators. Neural networks 2 (5) 359-366.

Huang, W., Lamb, D. W., Niu, Z., Zhang, Y., Liu, L., & Wang, J., 2007. Identification of yellow rust in wheat using in-situ spectral reflectance measurements and airborne hyperspectral imaging. Precision Agriculture 8 (4-5) 187-197.

Krishna, G., Sahoo, R. N., Pargal, S., Gupta, V. K., Sinha, P., Bhagat, S. *et al.,* 2014. Assessing wheat yellow rust disease through hyperspectral remote sensing. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Hyderabad, India, 40 1413-1416

Kuska, M., Wahabzada, M., Leucker, M., Dehne, H. W., Kersting, K., Oerke, E. C. *et al.,* 2015. Hyperspectral phenotyping on the microscopic scale: Towards automated characterization of plant-pathogen interactions. Plant Methods 11 (1) 1-14.

Li, L., Zhang, Q., & Huang, D., 2014. A review of imaging techniques for plant phenotyping. Sensors 14 (11) 20078-20111.

Moshou, D., Bravo, C., West, J., Wahlen, S., McCartney, A., & Ramon, H., 2004. Automatic detection of 'yellow rust' in wheat using reflectance measurements and neural networks. Computers and Electronics in Agriculture 44 (3) 173-188.

Nebiker, S., Lack, N., Abächerli, M., & Läderach, S., 2016. Light-weight multispectral uav sensors and their capabilities for predicting grain yield and detecting plant diseases. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Pragues, Czech Republic, 41 963-970.

Odilbekov, F., Armoniené, R., Henriksson, T., & Chawade, A., 2018. Proximal Phenotyping and Machine Learning Methods to Identify Septoria Tritici Blotch Disease Symptoms in Wheat. Frontiers in Plant Science 91-11.

Peterson, R. F., Campbell, A. B., & Hannah, A. E., 1948. A diagrammatic scale for estimating rust intensity on leaves and stems of cereals. Canadian Journal of Research 26 (5) 496-500.

Service Public de Wallonie, 2018. Programme wallon de réduction des pesticides 2018-2022 (Walloon program of pesticides reduction 2018-2022). http://diantonio.wallonie.be/files/PWRP_II.pdf. (last accessed August 2018)

Yu, K., Anderegg, J., Mikaberidze, A., Karisto, P., Mascher, F., McDonald, B. A. *et al.,* 2018. Hyperspectral Canopy Sensing of Wheat Septoria Tritici Blotch Disease. Frontiers in Plant Science 9 1-17.

Zadoks, J. C., Chang, T. T., Konzak, C. F., & Fryer, J. D., 1974. A decimal code for the growth stages of cereals. Weed Research 14 415-421.

Assessing infield temporal and spatial variability of leaf water potential using satellite imagery and meteorological data

O. Beeri*, R. Pelta, T. Shilo, J. Raz and S. Mey-tal
Manna-Irrigation, Gvat-Yagur 1, Gvat 3657900, Israel; ofer.beeri@manna-irrigation.com

Abstract

The goal of this study was to examine whether the combination of meteorological data along with satellite imagery can be used to represent the spatial and temporal changes in cotton leaf water potential (LWP). Daily plant measurements showed a decrease in LWP after each irrigation event and then a gradual increase until the next irrigation event. However, in most events, the lowest LWP was on the day after irrigation and sometimes even two days after. These changes generally correlated to seasonal meteorological changes but were not sensitive to the daily LWP changes. In contrast, spectral indices based on satellite imagery were able to map infield LWP variability and after calibration enabled the estimation of LWP with an RMSE of 0.26 MPa and R^2 of 0.72, regardless of sensor type.

Keywords: precision irrigation, crop water stress, Landsat, Sentinel-2, leaf water potential

Introduction

Precision irrigation is one of the main method to improve water-use-efficiency (WUE) in agriculture crops, whether by decreasing irrigation amounts or by increasing yield, or both. One of the methods for determination of the precise irrigation amounts during the growing season is measuring the crop water potential using a pressure chamber as a monitoring tool (Naor, 2010; Shackel *et al.*, 1997). Yet, since this method requires unique equipment, expertise and is labor intensive, it is rarely used in commercial agriculture monitoring and irrigation decision making. To overcome this issue, it was found that crop water potential highly correlates with temporal changes in evapotranspiration (ET) or vapor pressure deficit (VPD) allowing estimation of the crop water stress by meteorological measurements from a local weather station (Idso, 1982; Kopyt and Tsadok, 2015; McCutchan and Shackel, 1992). On the spatial scale, remote sensing imagery has been used to estimate crop water stress by plant sensor (Kullberg *et al.*, 2017), aerial (Cohen *et al.*, 2017) or satellite (Van Beek *et al.*, 2013) platforms. The goal of the current study was to assess infield temporal and spatial variability of leaf water potential (LWP) using satellite imagery and meteorological data in cotton.

Methods

The study was carried out at Yagur farm in the North of Israel (32°01'N; 35°45'E), in four commercial Pima (*Gossypium barbadense* L.) cotton fields (Figure 1). Cotton was sown on March 11, 13, 18 and 27, while irrigation started on May 28, 30, 28 and 31, in plots A, B, C and D, respectively. The irrigation was applied twice a week, applying a stress strategy that led to a weekly average shortage of -3.5 mm during the period of study (3-June to 28-July 2018).

Meteorology
Weather data was downloaded from the Israeli Meteorology Service (https://ims.data.gov.il/ims, last access 5-Dec-2018), for Haifa-refineries station located 4-5 km from the study area (Figure 1A). A 10-minute interval data of temperature (T) and relative humidity (RH) was used for vapor-pressure-deficit (VPD) calculations. Daily summaries of minimum, maximum and average were

DOI 10.3920/978-90-8686-888-9_32, © Wageningen Academic Publishers 2019

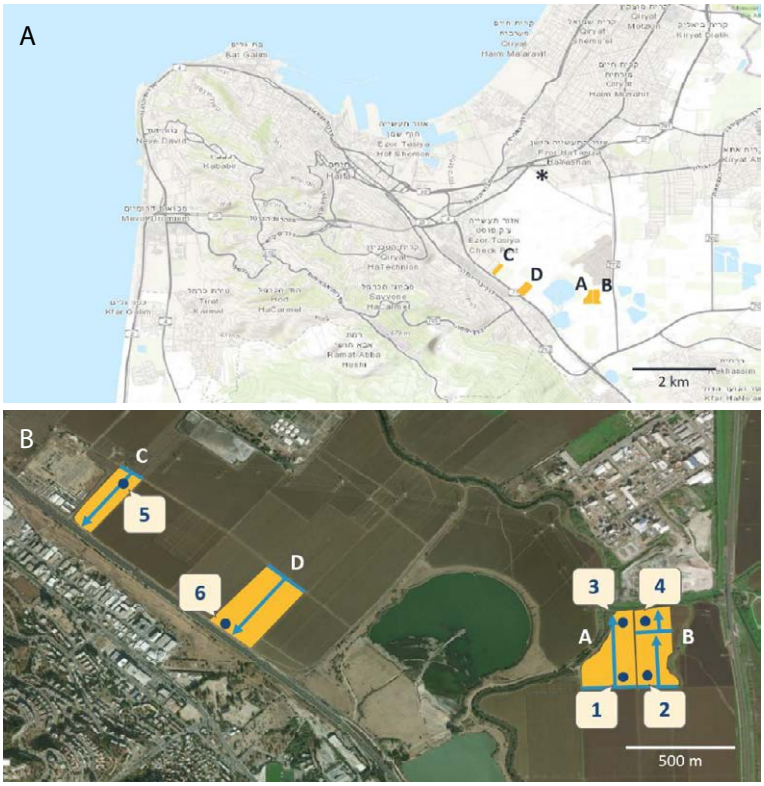


Figure 1. The study sites. (A) Fields (orange polygons) and the adjacent weather station (asterisk) and (B) The same fields and the measurement sites (blue dots). The blue lines and arrows represent drip irrigation lines and the direction of flow.

calculated for T, RH and VPD enabling the calculations of reference evapotranspiration according to Hargreaves and Samani (1985) (ET_0hs).

Crop measurements

Crop measurements were conducted daily, five days a week, over a period of six weeks, to cover a total of 31 days at six sites (Figure 1B). Each measurement included leaf water potential (LWP) and leaf area index (LAI) as described below. Sites 1, 2 and 4 were set close to the beginning of the drip line while site 3 was set as far as possible (350 m down the drip line) (Figure 1B). Sites 5 and 6 presented the extreme cases of very wet and very dry areas, respectively, according to the Tasseled Cap Wetness map (Beeri *et al.*, 2018) plotted at the beginning of the study (see satellite imagery), respectively. Therefore, these two sites were utilized to generate LWP estimations, which were validated by sites 1-4. Each site was located inside a 30×30 m Landsat pixel, at least 75 m into the center of the field. For LWP measurement, eight leaves were sampled randomly from each site, at midday (13:00-15:00 local time), stored in a dark plastic envelope inside a cooling box, and measured no more than 15 minutes after clipping. LWP was measured with a pressure chamber (model 655, PMS Instrument Company, Albany, OR, USA). LAI was measured using a plant canopy analyzer (SunScan model SS1, Delta-T Devices, Cambridge, England). For each site, LAI was averaged based on 200 samples covering five strips of approximately 5×1.93 m each.

Satellite imagery

Sentinel-2 imagery from 5-June, 10-June, 15-June, 25-June, 5-July and 25-July 2018 were downloaded from the European Space Agency datahub (https://scihub.copernicus.eu/dhus/#/home, last access 7-Jan-2019). Landsat- 8 images from 7-June, 16-June, 2-July, 9-July and 25-July-2018 and Landsat-7 from 24-June were downloaded from the US Geological Survey website (https://earthexplorer.usgs.gov/, last access 7-Jan-2019). All images were downloaded in Level-2a, i.e. in surface reflectance values.

For each image, four crop water stress spectral indices were calculated (Table 1), where each index had been found to be highly correlated to cotton LWP in previous studies (Beeri *et al.*, 2018).

Statistical analysis

Meteorological data, namely, maximum T, maximum VPD and ET_0hs were linearly regressed vs LWP. To assess the ability of the above spectral indices in LWP estimations, linear regressions between crop measurements and the spectral information from sites 5 and 6 were used as calibration to estimate LWP in sites 1-4. For the purpose of this step, only images that were cloud-free for all sites were used and included Sentinel-2 from 25-June, 5-July and 25-July and Landsat-8 images from 7-June, 2-July and 9-July. These estimations were compared to crop measurements according to four evaluation metrics: bias, root mean square error (RMSE), normalized RMSE (nRMSE) and coefficient of determination (R^2) that were calculated as follows:

$$\text{Bias} = \sum(Si - Mi) / n \quad (1)$$

$$\text{RMSE} = [\sum Si - Mi)^2 / n]^{1/2} \quad (2)$$

$$\text{nRMSE} = RMSE / (maximum\ (Mi) - minimum\ (Mi)) \quad (3)$$

where Si, Mi and n are estimation, measurement and number of observations, respectively.

Results

Daily changes in cotton LAI and LWP measurements presented a clear view of growth and water stress increase, as expected in highly managed deficit irrigation. Yet, even in this pattern, the effect of the distance from the irrigation source is well represented by the differences between the LAI and LWP in sites 1 (75 m from drip irrigation source) and 3 (350 m from the source) (Figure 2A and 2D), compared to sites 2 and 4 that were similarly distant from the start of irrigation line (Figure 2B and 2E). The highest need for precision irrigation is in mid-June, when LWP already showed

Table 1. Crop water stress spectral indices used in this study.

#	Index name	Equation[1]	Source
1	Tasseled Cap Wetness	B2 × 0.14 + B3 × 0.18 + B4 × 0.33 + B5 × 0.34 – B6 × 0.62 – B7 × 0.42	Kauth & Thomas, 1976
2	NDWI	(B5 – B6) / (B5 + B6)	Hardisky *et al.*, 1983
3	WSI1	(B6 + B7) / B5	Beeri *et al.*, 2018
4	GBNI	(B3 – B2) / (B3 + B2)	Beeri *et al.*, 2018

[1] The bands in the equations (i.e. B2, B3, etc.) refer to Landsat-8 bands and they can be applied to other sensors according to the corresponding wavelength of each band.

differences between the sites (Figure 2D), while LAI values were the same for both sites (Figure 2A). Only on June 24th, almost 4 weeks after irrigation had begun, the differences in LAI measurements were clear, while LWP was already significantly different on June 17th (see also Figure 3A).
Irrigation events generally decreased cotton water stress (Figure 3) as expected (McCutchan and Shackel, 1992). Furthermore, on five cases out of eleven, the water stress was substantially lower even two days following the irrigation event (Figure 3A, E, F). In contrast, only in one out of eleven cases, LWP had increased following irrigation (Figure 3B, site 1).
Daily changes in cotton LWP were related to meteorology parameters, although R^2 values indicated only moderate correlation (Table 2).
These relationships did not seem as strong indicators for LWP change inside Field A (sites 1 and 3) and the daily changes in specific days as presented in mid-June and mid-July (Figure 4).
This infield water stress variability was assessed by the Landsat and Sentinel imagery. It was evident particularly in Field A, where the differences between sites 1 and 3 were observed a few weeks after the beginning of irrigation (Figure 5).
The assessment of LWP estimations using satellite imagery were reasonably accurate (Table 3) and did not differ between the two sensors. Furthermore, the Tasseled Cap Wetness index (Table 1, Equation 1) achieved the highest accuracy.
The resulting RMSE for that index are within the range of cotton LWP as estimated with aerial thermal imagery, 0.09-0.32 (Cohen *et al.*, 2017) or with World-View satellite imagery, 0.32-0.39 (Van Beek *et al.*, 2013). The 13-17% errors, presented by the nRMSE, are in the range of remote sensing estimations for crop water stress (Kullberg *et al.*, 2017), crop height or LAI (Anderson *et al.*, 2004).

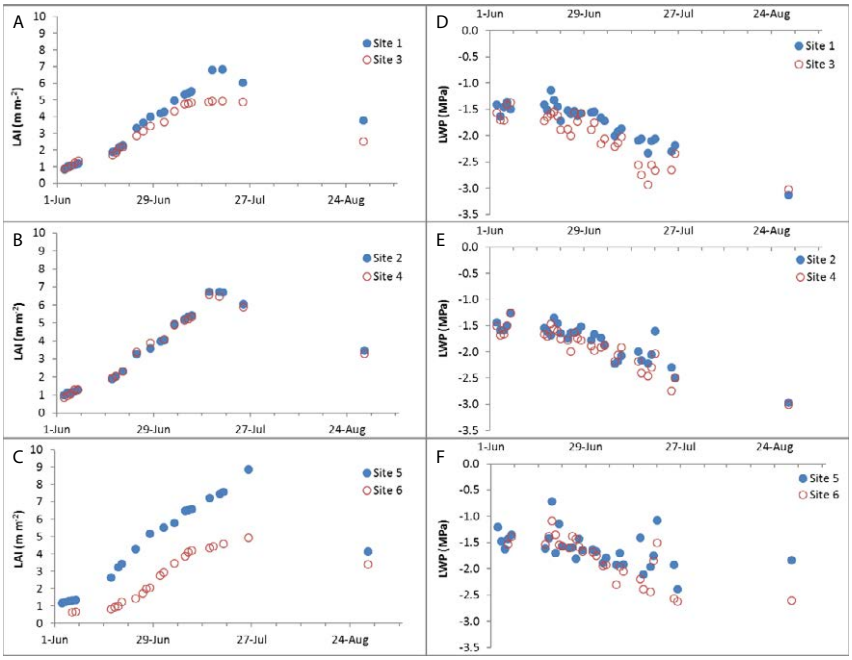


Figure 2. LAI (A-C) and LWP (D-F) measurements for sites 1 and 3 (A, D), sites 2 and 4 (B, E) and sites 5 and 6 (C, F).

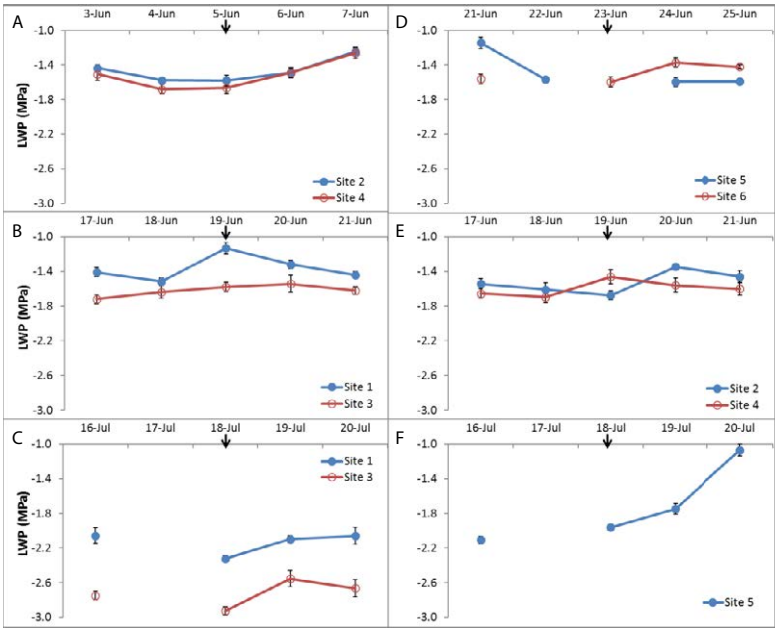


Figure 3. Daily changes in cotton LWP, two days prior and two days after irrigation. Arrow represents the irrigation event. Bars represent mean ± standard error.

Table 2. Regression parameters between cotton LWP and meteorological parameters.

	Site	Slope	Intercept	R^2
T_{max}	1	-0.10	1.38	0.59
	2	-0.10	1.34	0.59
	3	-0.13	2.26	0.57
	4	-0.12	1.83	0.70
	5	NS		
	6	0.14	-2.55	0.63
VPD_{max}	1	-0.21	-1.20	0.40
	2	-0.22	-1.23	0.45
	3	-0.29	-1.27	0.40
	4	-0.28	-1.22	0.57
	5	NS		
	6	0.29	1.07	0.46
ET_0hs	1	-0.29	-0.24	0.43
	2	-0.32	-0.15	0.52
	3	-0.41	0.09	0.44
	4	-0.36	-0.10	0.53
	5	NS		
	6	0.38	-0.14	0.45

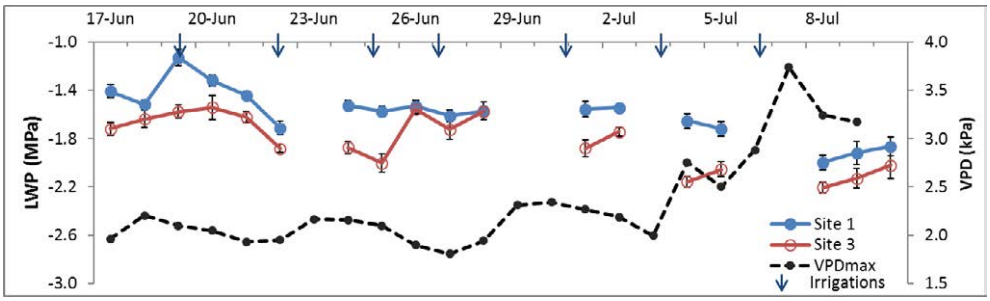


Figure 4. Daily changes of LWP and VPD_{max} in Field A. Blue arrows indicate irrigation events. Bars represent mean ± standard error.

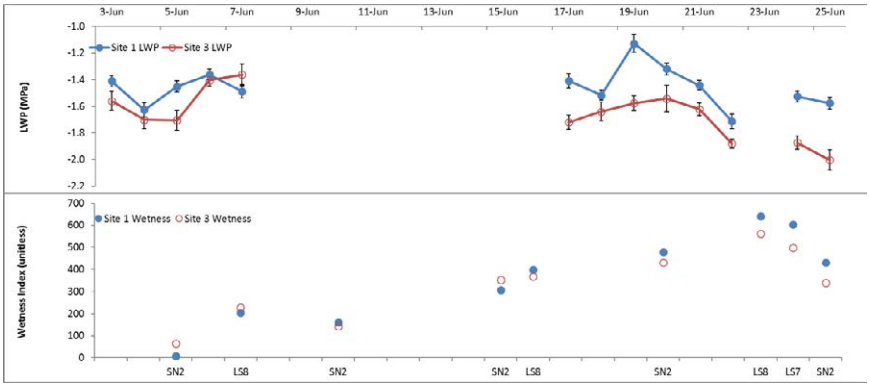


Figure 5. LWP measurements and Tasseled Cap Wetness index values for sites 1 and 3 during the beginning of the irrigation season. Bars represent mean ± standard error.

Conclusions

Cotton water stress significantly changes during the season, both temporally and spatially, especially in the case of deficit irrigation strategy. The irrigation itself decreased this stress not only on the day of the irrigation but also for the next day or two. The maximum temperature and VPD are also correlated to the temporal changes, but the significant changes between consecutive days cannot be detected accurately by these parameters as measured from a weather station 5 km from the field. Improving the spatial variability of the actual climate data by using daily thermal imagery, could increase this accuracy. The spatial variability in crop water stress can be detected by satellite imagery. However, it requires calibration to wet and dry crop LWP measurements. Future research directions should include these estimations, together with high-resolution daily surface temperature estimations, to detect infield variability based on Landsat/Sentinel-2 spectral images and to predict day-to-day variability according to meteorology data.

Table 3. Evaluation metrics of crop LWP [MPa] measurements and imagery estimations for sites 1-4. These values are the averages, after calibrating each image specifically to the measurements that were performed on that day.

Sensor	Index	Wetness	NDWI	WSI1	BGNI
All	Bias	0.078	0.055	0.058	0.035
7 images	RMSE	0.263	0.334	0.429	0.595
	nRMSE	15%	19%	24%	33%
	R^2	0.72	0.65	0.54	0.21
SN2	Bias	0.218	0.212	0.226	0.175
3 images	RMSE	0.299	0.293	0.301	0.426
	nRMSE	17%	16%	17%	24%
	R^2	0.76	0.71	0.62	0.58
LS8	Bias	-0.026	-0.062	-0.068	-0.318
4 images	RMSE	0.234	0.362	0.504	0.695
	nRMSE	13%	20%	28%	39%
	R^2	0.83	0.74	0.63	0.09

References

Anderson, M. C., Neale, C. M. U., Li, F., Norman, J. M., Kustas, W. P., Jayanthi, H., *et al.*, 2004. Upscaling ground observations of vegetation water content, canopy height, and leaf area index during SMEX02 using aircraft and Landsat imagery. Remote Sensing of Environment 92, 447-464.

Beeri, O., Mey-tal, S., Raz, Y., Rud, R., Pelta, R., 2018. Detecting variability in plant water potential with multi-spectral satellite imagery, in: The 14th International Conference on Precision Agriculture. https://www.ispag.org/Proceedings (last access 7-Jan-2019).

Cohen, Y., Alchanatis, V., Saranga, Y., Rosenberg, O., Sela, E., Bosak, A., 2017. Mapping water status based on aerial thermal imagery: comparison of methodologies for upscaling from a single leaf to commercial fields. Precision Agriculture. 18, 801-822.

Hardisky, M. A., Klemas, V., and Smart, R. M., 1983. The influence of soft salinity, growth form, mad leaf moisture on the spectral reflectance of Spartina alterniflora canopies. Photogrammetric Engineering & Remote Sensing. 49, 77-83.

Hargreaves G. H., Samani Z. A., 1985. Reference crop evapotranspiration from temperature. Applied Engineering in Agriculture. 1, 96-99.

Idso, S. B., 1982. Non-water-stress baselines: A key to measuring and interpreting plant water stress. Agricultural Meteorology. 27, 59-70.

Kauth R. J., Thomas, G.S., 1976. The tasseled cap – a graphic description of the spectral-temporal development of agricultural crops as seen in Landsat, in: Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, West Lafayette, Indiana: Purdue University. pp. 41-51.

Kopyt, M., and Tsadok, S., 2015. Everyday monitoring of stem water potential for decision making in irrigation of grapevines. Grapegrower & Winemaker. 618, 44-46.

Kullberg, E. G., DeJonge, K. C., Chavez, J. L., 2017. Evaluation of thermal remote sensing indices to estimate crop evapotranspiration coefficients. Agriculture Water Management. 179, 64-73

McCutchan, H., Shackel, K. A., 1992. Stem-water potential as a sensitive indicator of water stress in prune trees (*Prunus domestica* L. cv. French). Journal of the American Society for Horticultural Science. 117, 607-611.

Naor, A., 2010. Irrigation scheduling and evaluation of tree water status in deciduous orchards, Horticultural Reviews. 32, 111-165.

Shackel, K. A., Ahmadi, H., Biasi, W., Buchner, R., Goldhamer, D., Gurusinghe, S., *et al.*, 1997. Plant water status as an index of irrigation need in deciduous fruit trees. HortTechnology. 7, 23-29

Van Beek, J., Tits, L., Somers, B., Coppin, P., 2013. Stem water potential monitoring in pear orchards through worldview-2 multispectral imagery. Remote Sensing. 5, 6647-6666.

Characterisation of cereal morphology through proximal stereo vision under contrasting nitrogen inputs

S. Dandrifosse[1*], A. Bouvry[1], V. Leemans[1], B. Dumont[2] and B. Mercatoris[1]
[1]*Biosystems Dynamics and Exchanges, TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech, Liège University, Passage des Déportés 2, 5030, Belgium; sebastien.dandrifosse@uliege.be*
[2]*Plant Sciences, TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech, Liège University, Passage des Déportés 2, 5030, Belgium*

Abstract

This study aimed to determine the capacity of a proximal stereo vision system to characterise the architecture of winter wheat and spring barley canopies. Images were acquired by a nadir pair of cameras, in two fields dedicated to fertilisation trials (Gembloux, Belgium). Soil and plants were separated through support vector machines. Dense 3D point clouds were computed, describing canopy height. The study of mean leaf angles and leaf angle distributions revealed different dynamics for wheat and barley but little effect of fertilisation. Due to its low cost and its compactness, such a 3D vision system is suitable for farm applications.

Keywords: stereo vision, wheat, barley, morphology, leaf angle

Introduction

Nitrogen is a key nutrient to ensure cereal crop growth. Nowadays, for precision agriculture, many imaging approaches focus on evaluating the above ground plant development or the plant nitrogen status to adapt fertilisation. These indirect measurements are usually based on crop reflectance at specific wavelengths. The proposed approach aims to directly extract morphological plant traits from images to assess the crop growth and understand the effects of nitrogen inputs on plant morphology. Some conventional morphological traits such as plant height, tiller density, spike density and flag leaf length can be manually measured in the field. The originality of this contribution is to highlight image-based measurements of traits that cannot be easily measured by manual means. Plant ratio, which is the proportion of plant in a scene is firstly studied. Secondly, the mean angle between the leaves and the ground, called the mean tilt angle (MTA), is analysed. Thirdly, leaf angle distribution (LAD) describing the repartition of leaf angles is considered. Few studies have focused on LAD despite its importance (Zou and Mõttus, 2017). MTA and LAD condition the fraction of intercepted sunlight. Moreover, leaf inclinations influence the impact of phytosanitary product droplets (Massinon *et al.*, 2014). The characterisation of leaf angles may contribute to the choice of spraying features adapted to crops and weeds (droplets size, speed, ejection angle, nozzle type, etc.) and to the development of precision spraying. Finally, the last trait of interest is canopy height. Computer vision methods measure the height of all plant elements visible in the image while manual measurements only focus on specific organs.

The direct high-throughput measurement of these traits requires 3D imaging. For precision agriculture applications, the 3D imaging method has to be robust to environmental conditions, affordable and compact. Active 3D methods fail to meet these criteria. Time of Flight cameras need to illuminate the scene which is suitable for indoor measurements but is an issue under natural sunlight (Perez-Sanz *et al.*, 2017). Light Detection And Ranging (LiDAR) is expensive, needs to be coupled with a RGB camera to provide colour information and the resolution depends on the scanning time which is an issue in windy conditions (Gibbs *et al.*, 2017; Qiu *et al.*, 2018). For these

DOI 10.3920/978-90-8686-888-9_33, © Wageningen Academic Publishers 2019

reasons, the binocular stereo vision method has been chosen. This passive method only relies on two static RGB cameras which is a very simple, cheap, flexible and compact set-up.
The goals of this study were (1) to record the dynamics of plant ratio, MTA, LAD and canopy height by stereo vision for wheat, *Triticum aestivum* L., and barley, *Hordeum vulgare* L., and (2) to assess the effect of varied nitrogen managements on those traits.

Materials and methods

Field experiments and image acquisition

Images were acquired during the 2018 spring from two fields dedicated to fertilisation trials. The winter wheat field was located in Lonzée, Belgium (50° 32' 58" N and 4° 44' 08" E) and sown on October 13, 2017 with *Edgar* variety. Four nitrogen fertilisation practices (27% ammonitrate) were tested in the trial. Nitrogen was applied at tillering and stem elongation stages (respectively April 02 and April 21), as 0-0, 30-90, 0-60 and 90-30 kg N/ha. The spring barley field was located in Gembloux, Belgium (50° 33' 53" N and 4° 42' 37" E) and sown on March 21, 2018 with *Planet* variety. Nitrogen was applied at seedling and stem elongation stages (respectively April 14 and May 13) as 0-30, 45-15, 90-30 and 135-45 kg N/ha. Both fields were divided into 1.8'6 m micro-plots.
Wheat images were acquired on April 11, April 23, May 02, May 16 and May 24. Barley images were acquired on April 27, May 07, May 14, May 22 and May 28. The acquisition hours were between 10:30 a.m. and 03:00 p.m. The cameras were perpendicular to the ground and kept at 1 meter above the canopy. The platform was immobile when capturing frames. No shadowing panel was used as one of the goals was to develop a method robust to sunlight. Four images were acquired per micro-plot with a colour resolution of 12 bits per pixel per colour component to take full advantage of the hardware. The cameras were GO-5000C-USB (JAI, Kushima, Japan) equipped with a 2,560'2,048 CMOS sensor and a RGB Bayer filter. The size of each camera was 29'29'41.5 mm. The objectives were Kowa LM16HC with a focal length of 16 mm. The length of each objective was 54.2 mm and the diameter was 43 mm. The iris aperture and the focus were set to F2.8 and 1 m, respectively. Optical axes of the two cameras were parallel and spaced 50 mm apart. The total cost for the two cameras and the two objectives was around 2,000 euro (November 2017). The camera pair were calibrated using a 9'6 checkerboard (square side of 40 mm) and the Matlab application Stereo Camera Calibrator. The calibration error was 0.32 pixels.

Image segmentation and plant ratio measurement

Image processing was performed with Matlab R2016a. Soil and leaf pixels were classified using Support Vector Machine (SVM) trained with components of RGB and HSV color spaces. The training data set comprised 7,000 pixels and 3,000 other pixels were selected for validation. A 5'5 median filter was applied on binary images to remove noise. Segmentation directly provided plant ratio.

3D point cloud construction and height measurement

In order to reduce the noise induced by stereo matching and minimise computation time, the images were binned from 2,560'2,048 to 1,280'1,024 pixels by pixel value averaging. The spatial resolution was on average 0.56 mm per pixel for binned images. Stereo matching between left and right images was performed using Semi-Global Block Matching (SGBM) algorithm from Hirschm (2007). The output of this step was a depth map. Matching parameters are presented in Table 1. An implementation of the algorithm proposed by Yun (2012) was used to fill holes in the depth map. The ground-plant distance (plant height) was computed for each plant pixel as the difference between the camera-wheat distance (from depth map) and the camera-ground distance. The output was a plant height map. Figure 1 illustrates segmentation and height map computation steps. It was then possible to build a 3D point cloud on the basis of this height map and of the calibration parameters.

Table 1. Parameters of the SGBM stereo matching algorithm. Block size and disparity range condition the matching between homologous pixels in the two images. Uniqueness ratio, contrast and texture threshold are post filtering parameters.

Parameter	Block size	Disparity range	Uniqueness ratio	Contrast threshold	Texture threshold
Value	5	automatically adjusted	5	0.8	0.0002

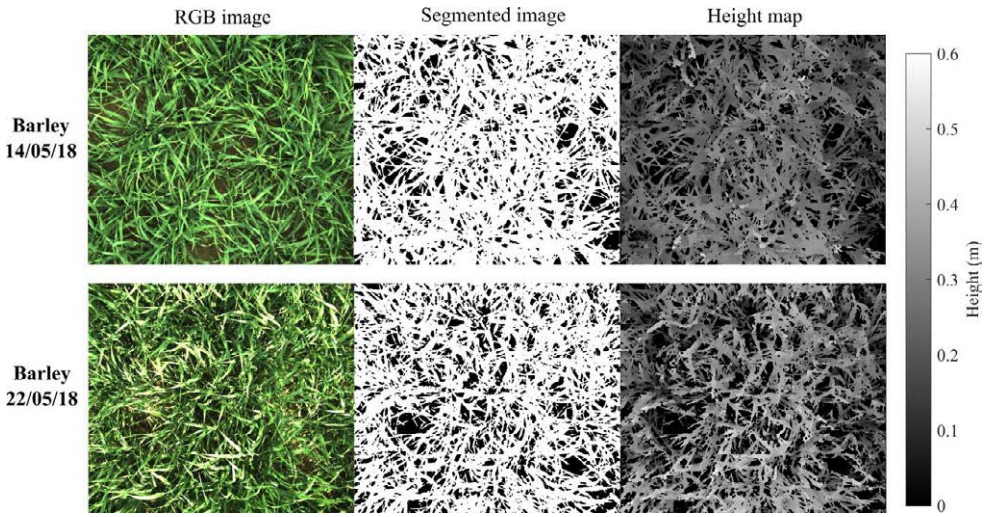


Figure 1. Segmented images (barley in white and soil in black) and height maps.

Angle measurement

A method to determine leaf angle on the basis of a dense point cloud relies on reconstructing the leaf surface by Delaunay triangulation and then computing the inclination of individual triangles (Leemans *et al.*, 2013; Müller-Linow *et al.*, 2015). MTA was extracted by averaging the triangle inclinations on leaf pixels of the scene. However, the surface triangulation suffered from a stair-step effect leading to non-representative LAD. To better estimate LAD, the proposed approach consisted of automatically sampling independent leaf zones of 16'16 pixels (for binned images). To avoid selecting zones riding on two different leaves, an edge detection Canny filter was applied to detect leaf boundaries. A leaf zone was considered valid if two criteria were met: (1) all pixels presented a plausible height value (ranging between the ground and 0.6 m below the stereoscopic device); and (2) the zone did not contain any leaf edge. If the number of zones satisfying these criteria was inferior to a threshold of 10, the algorithm started the process over by searching for smaller zones such as 8'8 and 4'4 pixels. For each zone, a plane was fitted on the point cloud. The angle of the leaf face was the angle of the plane. In comparison, the approach employed by Hosoi *et al.* (2009) to compute wheat LAD required visual selection of leaves.

Results and discussion

Stereo matching

Nowadays, one of the main reported drawbacks of stereo vision for precision agriculture applications is the heavy computation time for stereo matching (Perez-Sanz *et al.*, 2017; Qiu *et al.*, 2018). This may particularly be an issue for real-time measurements. Pixel binning helped to reduce the computation time by ten times. Nevertheless, the computation time is not the only limiting factor for real time applications. The computation of a proper exposure time during image acquisition is also crucial, especially for sunny field conditions.

Another drawback imputed to stereo vision in field conditions is the difficulty of performing stereo matching due to natural sunlight and the homogeneous texture of leaves (Perez-Sanz *et al.*, 2017). Müller-Linow *et al.* (2015) reported wrong height values for barley leaves because of homogeneous surface properties. Thanks to the methodology proposed here, acceptable height values (ranging between the ground and 0.6 m below the stereoscopic device) were found for 83.5% of wheat pixels and 87.9% of barley pixels.

Segmentation and plant ratio

A unique SVM-based classifier was trained to isolate wheat and barley leaves for all the acquisition dates during the growing season. The classification accuracy computed on an independent validation data set including both species and different environmental conditions was 98.5%. This underpins that a segmentation using machine learning combining RGB and HSV color spaces is robust to cereal species and natural light conditions. Figure 2 shows the dynamics of plant ratio for wheat and barley according to four fertilisation practices. Overall, unfertilised plots presented a significantly lower plant ratio but there were no significant differences between other fertilisation practices.

It is noticed that plant ratio can also be obtained by measuring canopy transmittance, as performed by spectrometers. Imaging however presents the advantage of also providing the repartition of soil and plant elements and their morphology. For instance, criteria based on the size and the shape of visible soil patches could be exploited to evaluate spatial leaf clumping or identify crop rows.

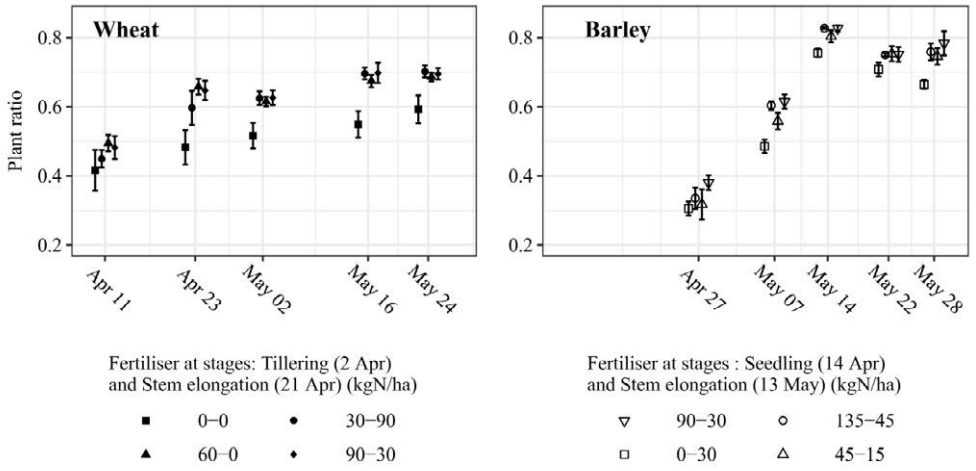


Figure 2. Dynamics of plant ratio. The graphs present plant ratio and its standard deviation for the different micro-plots.

Mean tilt angle and leaf angle distribution

A fully automated image treatment method has been developed to compute MTA and LAD. The dynamics of MTA are shown in Figure 3 for both species. Small differences were observed for MTA between the different fertilisation practices.

For each date and crop, LAD was not significantly different for the four fertilisation practices. Figure 4 presents the most contrasting difference. For this reason, the rest of the study focused on the comparison by date of the average LAD of the crops without differentiation of nitrogen treatments (Figure 5). *Edgar* wheat presented overall a plagiophile LAD characterised by most common angles at 65° while *Planet* barley presented an erectophile LAD characterised by a maximal frequency for nearly vertical leaves. Moreover, the dynamics of the two crops were different. During the season, wheat leaves tended to be less tilted whereas barley leaves tended to be more tilted.

It is noticed that the acquisition hour may have impact on LAD measurement. This hypothesis needs to be verified in a further study dedicated to a smaller number of plots captured at different hours and dates.

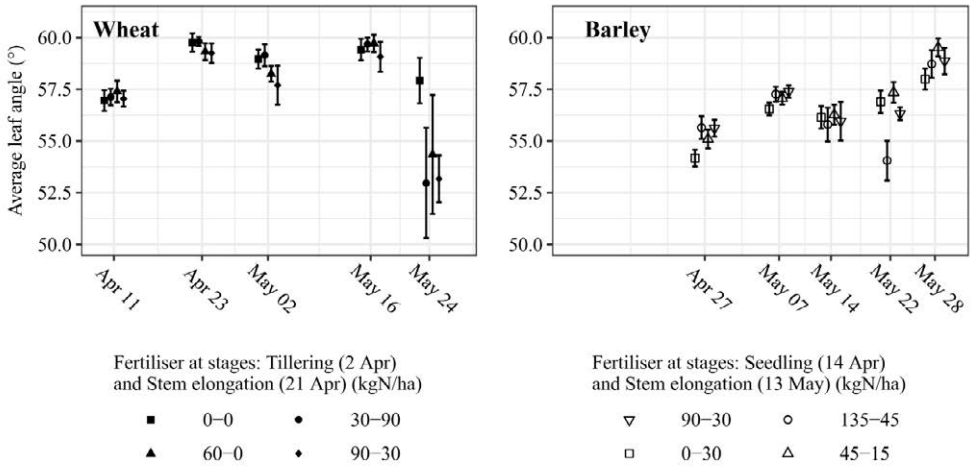


Figure 3. Dynamics of mean tilt angle of leaves. The graphs present MTA and its standard deviation for the different wheat and barley micro-plots.

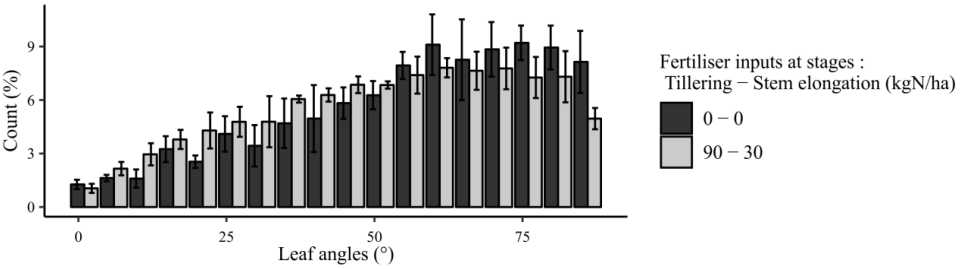


Figure 4. Leaf angle distribution of wheat measured on May 24 for two contrasting nitrogen treatments.

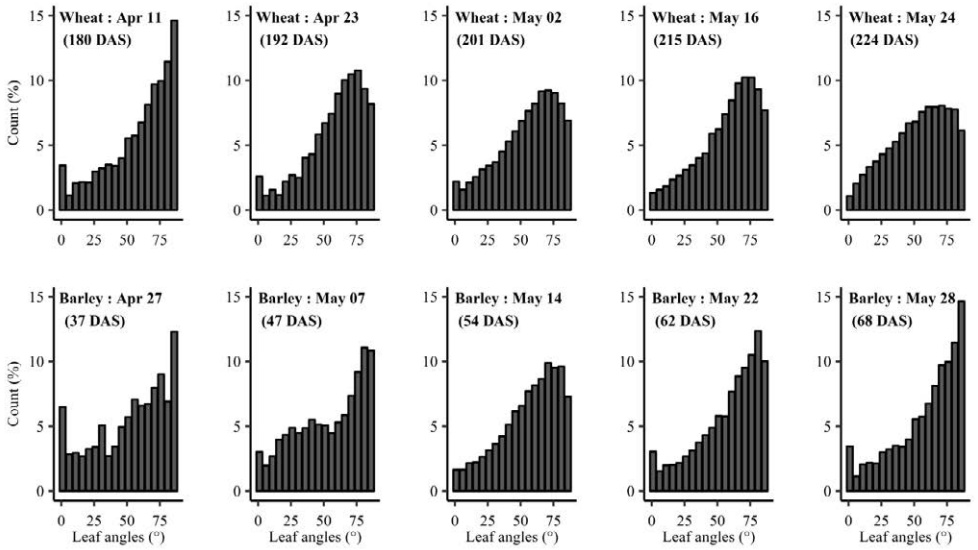


Figure 5. Dynamics of leaf angle distribution. The dates are translated in days after sowing (DAS).

Canopy height

The dynamics of canopy height, described as the 95th percentile of plant pixel height in a scene, are shown in Figure 6. For the early growing stages of wheat, there were no significant differences between the nitrogen treatments. Differences appeared on May 24 and may have resulted from smaller flag leaves for the unfertilised micro-plots. For barley, height differences due to fertilisation increased as the crop grew. At booting stage (May 28), the plants fertilised with 0-30 kg/ha were significantly smaller. Dynamics of wheat and barley height from tillering to booting stage were different.

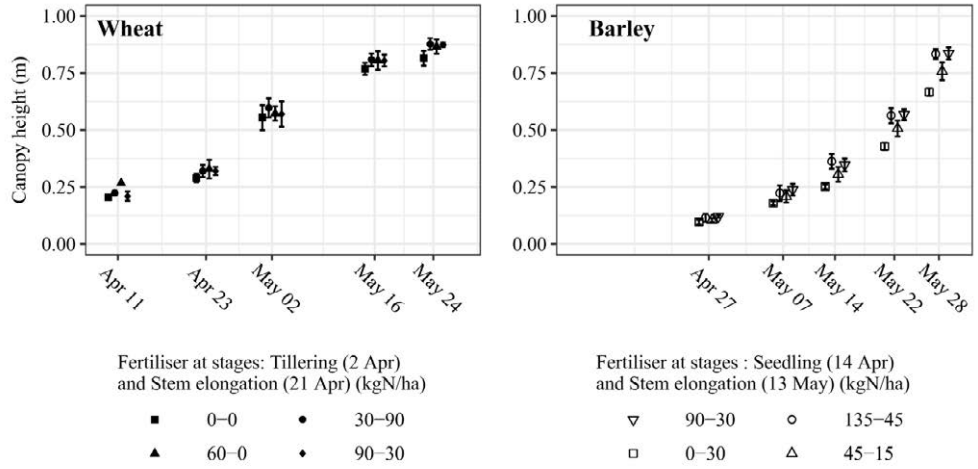


Figure 6. Dynamics of canopy height. The graphs present 95th percentile of plant element heights and its standard deviation for wheat and barley.

Validation of angle and height measurements

Validation of angle and height measurements is a tricky issue since equivalent manual reference measurements are complex, time-consuming or even unrealistic. As example, a manual measurement of leaf angle as the angle between sheath and blade is not representative of the average angle of curved leaves. Manual height measurements also need to focus on specific plant organs while, in contrast, stereo vision takes into account all the captured plant zones. For instance, the height of the last node can significantly differ from the global height provided by stereo vision which takes all the leaves into account.

Another way to validate the method would have been to assess the quality of distance measurement by stereo vision. The morphological traits are indeed obtained from the stereo-based depth map by geometric computations. However, evaluating the accuracy of distance measurement in field conditions is not obvious. It depends on stereo matching, which is itself strongly influenced by the texture of leaves and the lighting conditions, the latter rendering laboratory testing non-representative.

As a consequence, even if the in-field measurements seem coherent (especially regarding nitrogen inputs) and that the error on distance measurement for test set-ups is generally inferior to 0.5 mm, it was not possible to strictly validate the values of the canopy structure parameters determined by stereo vision. Nevertheless, the low variability of angle and height measurements despite natural field heterogeneities indicates that the method is relatively precise, regardless of its absolute accuracy. Although the absolute values may be biased, the method still allows comparisons between different treatments for similar acquisition conditions.

Conclusions

The developed stereo vision method was able to precisely measure plant ratio, MTA, LAD and canopy height for both barley and wheat crops, even under strong sunlight. In contrast to methods based on spectral reflectance properties, 3D imaging directly provides crop morphology. Fertilisation had little effect on leaf angles but affected plant ratio and canopy height. The price and the compactness of the sensor make it an interesting tool to equip phenotyping platforms or tractors. Moreover, the combination of colour images and depth information offers the perspective to extract many other plant traits such as flag leaf morphology or spike density, making the stereoscopic camera a really versatile sensing device.

Acknowledgement

This study was funded by the Walloon region (project D31-1385 and FRIA grant).

References

Gibbs, J. A., Pound, M., French, A. P., Wells, D. M., Murchie, E., and Pridmore, T., 2017. Approaches to Three-Dimensional Reconstruction of Plant Shoot Topology and Geometry. Functional Plant Biology 44(1) 62-75.

Hirschm, H., 2007. Stereo Processing by Semi-Global Matching and Mutual Information. IEEE Transaction on Pattern Analysis and Machine Intelligence 30(2) 328-341.

Hosoi, F., Nakai, Y., and Omasa K., 2009. Estimating the Leaf Inclination Angle Distribution of the Wheat Canopy Using a Portable Scanning Lidar. Journal of Agricultural Meteorology 65(3) 297-302.

Leemans, V, Dumont B., and Destain, M.-F., 2013. Assessment of Plant Leaf Area Measurement by Using Stereo-Vision. In: International Conference on 3D Imaging (IC3D), edited by IEEE, Piscataway, New Jersey, USA, pp. 1-5.

Massinon, M., Boukhalfa, H., and Lebeau, F., 2014. The Effect of Surface Orientation on Spray Retention. Precision Agriculture 15(3) 241-54.

Müller-Linow, M., Pinto-Espinosa, F., Scharr, H., and Rascher, U., 2015. The Leaf Angle Distribution of Natural Plant Populations: Assessing the Canopy with a Novel Software Tool. Plant Methods 11(1) 1-16.

Perez-Sanz, F., Navarro, P.J., and Egea-Cortines, M., 2017. Plant Phenomics: An Overview of Image Acquisition Technologies and Image Data Analysis Algorithms. GigaScience 6(11) 1-18.

Qiu, R., Wei, S., Zhang, M., Li, H., Sun, H., Liu, G., *et al.,* 2018. Sensors for Measuring Plant Phenotyping: A Review. International Journal of Agricultural and Biological Engineering 11(2) 1-17.

Yun, L., 2012. Traitement et Analyse d'Images Stéréoscopiques avec les Approches du Calcul Générique sur un Processeur Graphique (Stereoscopic Image Processing and Analysis with Generic Computing Approaches on a Graphics Processor). Thesis, Ecole polytechnique de Montreal, Canada, 174 pp.

Zou, X., and Mõttus, M., 2017. Sensitivity of Common Vegetation Indices to the Canopy Structure of Field Crops. Remote Sensing 9(10) 1-15.

Early detection of corn and sunflower stress induced by chemical spraying

S. Gad[1], Y. Edan[1], T. Sandovsky[1], I. Harary[2], T. Nacson[3], E. Kosover[4], A. Levi Bar Shalom[4] and V. Alchanatis[3*]
[1]*Department of Industrial Engineering & Management, Ben-Gurion University of the Negev, Beer Sheva, Israel*
[2]*Evogene, Rehovot, Israel*
[3]*Institute of Agricultural Engineering, Agricultural Research Organization (ARO), Volcani Center, Rishon-LeZion, Israel; victor@volcani.agri.gov.il*
[4]*AgPlenus, Rehovot, Israel*

Abstract

Herbicides inhibit plant growth by affecting different bio-chemical pathways or bio-physical states. Stress that was induced by inhibiting lipid metabolism or photosynthesis on corn and sunflower was detected using leaf spectral reflectance. Leaf spectral signatures were measured at six time points during two weeks by a spectro-radiometer. Phenotypes were evaluated by an expert at the same time points. Spectral processing models were used to compare the spectral response of a control group to treated plants using statistical T-tests. Results revealed that spectral reflectance can be used to detect the stress induced by inhibiting photosynthesis on corn and sunflower at the same time that the first visual symptoms appeared. Also, early detection can be accomplished with a lower confidence two days before visual symptoms appear.

Keywords: herbicides, spectrometry, leaf, detection

Introduction

Different organisms reflect or absorb the sun's electromagnetic radiation in different ways. By analyzing the spectral reflectance at suitable wavelengths, changes in the plant mechanisms can be derived. Multispectral and hyperspectral sensors are commonly used for early disease detection and zone delineation for site-specific treatment of the disease. Spectral data analyses derive vegetation parameters and identify diseases using indices like normalized difference vegetation index (NDVI) and the red inflection point (REIP) which are based on selected wavelengths (Lowe *et al.,* 2017).
Application examples include detection of: (1) Zn stress in Bahia grass using NDVI and ratio vegetation index (RVI) indices (Schuerger *et al.,* 2003), (2) Sunn pest/cereal pest stress in wheat (Genc *et al.,* 2008) using NDVI and (3) leaf rust (*Puccinia triticina*) in wheat using a leaf rust disease severity index (LRDSI) (Ashourloo *et al.,* 2014). Moreover, a relationship was discovered between the concentration of a pollutant named Phenanthrene, to several plant characteristics such as chlorophyll density and live stem density, by comparing the red edge area and their first derivative of the reflectance spectra (Zhu *et al.,* 2014). In another work, NDVI was used to estimate biomass index and resulted in improved vineyard management (Drissi *et al.,* 2009).
The objective of this work was to estimate the potential of using leaf spectral reflectance to detect inhibition of two leaf mechanisms: lipid metabolism and photosynthesis. This article presents results of the analysis that aimed to detect the above plant stresses at an early stage, using spectral sensing. Successful completion of this task can lead to delineation of management zones to be treated with variable rate application systems.

DOI 10.3920/978-90-8686-888-9_34, © Wageningen Academic Publishers 2019

Materials and methods

Experimental design

Corn (cultivar VIVANI 700) and sunflower plants (cultivar Jerusalem Dwarf Yellow Spray) were grown in pots in a controlled greenhouse (25±2 °C day 20±2 °C night, flood irrigation of tap water + fertilizer NPK 5:3:8 + 6, 2 ml/l at 50% water content) for a period of one month from the seeding time (sunflower-15.10.18 and corn- 25.10.18). The plots were arranged on two tables, one for each crop (Figure 1). Every plot received a number and a bar code for sampling management.

Three seeds were planted in each pot, which were thinned out into one plant. Three days after emergence, the plants were sprayed with different herbicides. Twenty-five different chemical applications were applied to 250 plots/pots which comprised five repetitions for each herbicide type, in a randomized design. Spectral measurements were conducted on 14 treatments (different chemical applications), 140 plants, due to sampling time constraints. A plant physiologist selected these treatments, to reflect a variety of modes/mechanism of action (MOA) of the herbicides (the primary biochemical or biophysical interference imposed by an herbicide that leads to lethality).

This paper focuses on analysis of seven treatments belonging to two modes of action: 'lipid metabolism (treatments T_9, T_{11} and T_{12}) and 'Inhibition of photosynthesis (treatments T_{14}, T_{15} and T_{16}). These treatments were compared to the control group (treatment T_1).

Spectral acquisition system

Spectral reflectance of each leaf was recorded by a spectroradiometer (ASD FieldSpec 4 hi-res, ASD Inc. Malvern Panalytical, Boulder, Colorado, USA) in the range of 350 nm-2,500 nm, optical spectral resolution of 3 nm in the VNIR range and 8 nm in the SWIR range.

Data acquisition:

Spectral measurements were conducted between 09:30 to 15:30 local time (UTC+2) through autumn 2018, at six time points (TP): TP1-04.11.18, TP2-06.11.18, TP3-08.11.18, TP4-11.11.18, TP5-13.11.18 and TP6-15.11.18.

Each plant was placed on the greenhouse table under direct sunlight, with the ASD fiber probe aimed at a leaf from a distance of a few centimeters without shading it. White reference was sampled whenever a change in illumination conditions occurred or every fifteen minutes. Four consecutive series of 10 spectra were acquired for each leaf, and the average was saved for later analysis.

Since the sprayed chemicals affect differently mature or young leaves, both young and mature leaves were sampled from each plant. Figure 2 shows the leaves that were sampled and Figure 3 shows



Figure 1. Experimental layout.

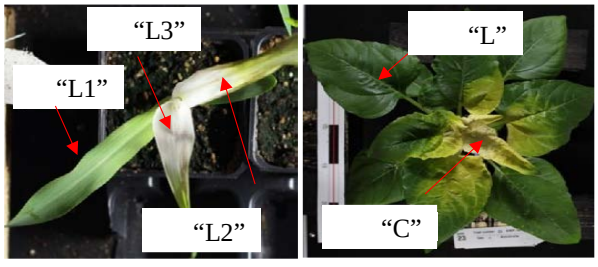


Figure 2. Sampling points of corn and sunflower leaves. The stress effect on the leaf growth and the sampled leaves is visible on the leaves.

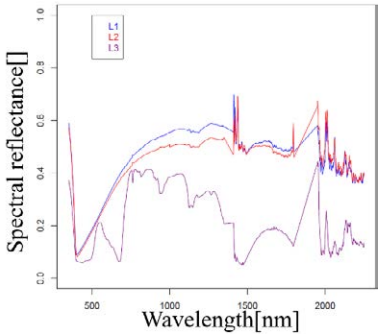


Figure 3. Spectral signature of three corn leaves at the same time point.

the spectral signatures of leaves L1, L2 and L3 at the same time point. In corn plants, three leaves were sampled starting with one leaf at the first time point, two at the second time point and three leaves from the third time point and on (in some treatments, a third leaf did not grow due to the chemical effect). In sunflower plants, one mature leaf and one young leaf were measured at every time point (Figure 2).

Eleven phenotypes were visually evaluated on a 1-6 scale by an expert agronomist at each time point: Pheno 1- Necrosis (spots), Pheno 2- Burning, Pheno 3- Bleaching, Pheno 4- Chlorosis, Pheno 5- Epinasty (curling), Pheno 6- Inhibited growth, Pheno 7- Wilting, Pheno 8- Disturbed apical bud, Pheno 9- Abnormal pigmentation (dark leaf), Pheno 10- Abnormal pigmentation (anathocyanins) and Pheno 11- Disturbed gravitropism.

Analysis of spectral data

Water vapor atmospheric absorption windows (1,350 nm-1,410 nm, 1,800 nm-1,950 nm and 2,250 nm-2,500 nm) were removed from the spectral data. The correlations between every wavelength reflectance to every phenotype were calculated for each treatment and leaf during all time points. T test was performed to determine significant differences between the groups, to select the most significant wavelength.

All possible combinations of normalized difference indices, NDI, were calculated (Equation 1) for each pair of wavelengths (W_i, W_j) for each treatment, leaf and time point:

$$NDI_{ij} = \frac{R_i - R_j}{R_i + R_j} \quad (1)$$

where R_i and R_j are reflectance values at wavelengths W_i and W_j respectively.

In order to find an indicator of stress in the plant for every treatment, each index was compared to its corresponding index in the control group using the statistical T test with a Bonferroni correction. All analysis was conducted using RStudio software version 1.0.153 (www.rstudio.com).

Results and discussion

Pre-processing of the reflectance data

Qualitative analysis revealed that the variance of different plants belonging to the same group was large, and that it would be difficult to attribute the changes in the spectral reflectance to the chemical treatments only. Therefore, a first derivative and smoothing was performed on the data using a procedure defined by Savitzky-Golay with a filter order of 2 and a filter length of 11 wavelengths (Savitzky and Golay, 1964).

Phenotyping

Time point 1 (TP1) was a baseline measurement, before chemicals spraying, resulting in a scale of 1 for all the phenotypes and implying no visual effect. The first phenotypes to arise (at time point 2) were (1) *Necrosis* and (2) *Burning* among a few plants from treatments T_{14} and T_{16} both in corn and sunflower (Table 1). Phenotypes: (5) *Epinasty (curling)*, (9) *Abnormal pigmentation (dark leaf)* and (11) *Disturbed gravitropism* did not change throughout the experiment period, therefore, they were not included in the analysis. Table 1 shows the time-points when the expert noticed visual changes for each treatment. For treatments T_{11}, T_{14}, T_{15} and T_{16} the first visual changes were noticed at TP4, TP2, TP3, TP2 respectively for both crops. For corn plants, in T_9, the first-time phenotype appeared at TP3 and for T_2 it appeared at TP4. These treatments had no effect on the sunflower plants. Table 1 also shows that phenotypes appeared earlier for the MOA of inhibition of photosynthesis.

To determine the wavelength that is best correlated with each phenotype, the correlation coefficient (in absolute value) between the smoothed derivative of the reflectance spectra and the phenotype score was calculated. Figure 4 and 5 show an example and Table 2 summarizes all phenotypes.

The last row in Table 2, indicates the correlation when every visual change was considered. Using the wavelength with the highest correlation with phenotype average, a linear regression prediction model for corn was built. The data were divided into training and test groups (80 and 20% accordingly). The accuracy of the lipid metabolism prediction model was 0.64 and for inhibition of photosynthesis, the accuracy was 0.79. Using the same wavelength, additional linear prediction models were built, where the data from TP(i) were used to predict the average phenotype score at TP(i+1). The prediction accuracy was 0.26 for lipid metabolism and 0.67 for Inhibition of photosynthesis

The mean and standard deviation of the spectra first derivative of each treatment was plotted against the control group for all time points. Due to multiple comparisons, a Bonferroni repair was conducted applying the P value compared to 5% divided by the number of comparisons (1,692). For instance, Figure 6 shows the comparison in the mature leaf of T16 at TP2 and TP3.

Figure 6A shows that in TP2 there was no significant difference (at a significance level of 95%) between treated and control plants. Figure 6B shows that in TP3, there was a difference between the groups at a significance level of 95% with *P*-value of $2.914383e^{-07}$ for wavelength 531 nm. From this time point onward, the differences remained significant. At TP4, TP5, TP6, the biggest difference was obtained for bands 568 nm, 648 nm and 582 nm respectively. The same procedure (visualization and T test) was performed for all sampled plants and the results are summarized in Table 3. Table 3 shows that the MOA of inhibition of photosynthesis was detected earlier than lipid metabolism.

Table 1. Appearance of plant phenotypes. T_x is the treatment number. The phenotype number written in brackets.

MOA	Plant	TP1	TP2	TP3	TP4	TP5	TP6
Inhibition of photosynthesis	Sunflower		T_{14}(2)	T_{14}(2,6)	T_{14}(2,6)	T_{14}(2,6)	T_{14}(2,6)
				T_{15}(2,6)	T_{15}(2,6)	T_{15}(2,6)	T_{15} (2,6)
			T_{16}(1,2)	T_{16}(2,6)	T_{16}(2,6)	T_{16}(2,6)	T_{16} (2,6)
	Corn		T_{14}(2)	T_{14}(6,7)	T_{14}(2,6)	T_{14}(2,6)	T_{14} (6,7)
				T_{15}(1,6)	T_{15}(2,6)	T_{15}(2,6)	T_{15} (3,6)
				T_{16}(2,6)	T_{16}(2,6)	T_{16}(2,6)	T_{16} (6,7)
MOA	Plant	TP1	TP2	TP3	TP4	TP5	TP6
Lipid metabolism	Sunflower				T_{11}(4)	T_{11}(7)	
	Corn			T_9(1,6)	T_9(1,4,6)	T_9(4,6,7)	T_9(6,7)
					T_{11}(1,6)	T_{11}(4,6,7)	T_{11}(6,7)
					T_{12} (6)	T_{12}(6,10)	T_{12} (6,8)

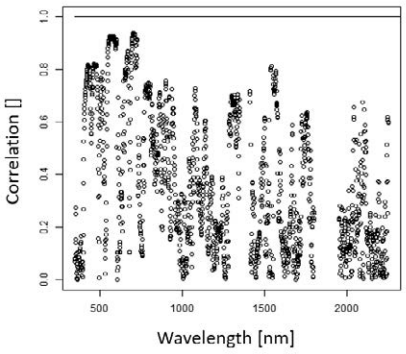


Figure 4. Single wavelength correlation between L1 first derivative of reflectance to Pheno6 for T14.

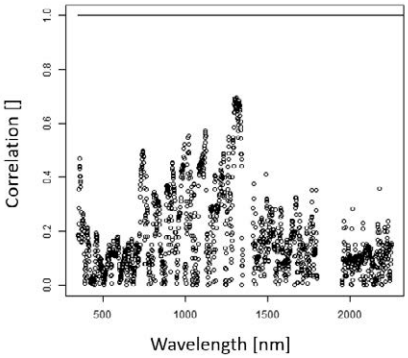


Figure 5. Single wavelength correlation between L1 first derivative of reflectance to Pheno1 for T9.

Table 2. Highest correlations received for corn using the derivative at a single wavelength.

	MOA 1- lipid metabolism			MOA 2-inhibition of photosynthesis		
Phenotype	Wavelength [nm]	Correlation [%]	Leaf	Wavelength [nm]	Correlation [%]	Leaf
Necrosis	1,303	34	L1	356	32.2	L1
Burning			L1	801	53.8	L1
Bleaching			L1	1994	56.9	L1
Chlorosis	548	59	L1			L1
Epinasty			L1			L1
Inhibited growth	553	68.1	L1	569	85.9	L1
Wilting	553	78.9	L1	587	46.1	L1
Disturbed apical bud	1,969	55.8	L1			L1
Abnormal pigmentation (dark leaf)			L1			L1
Abnormal pigmentation (anthocyanins)	1,558	69.9	L1			L1
Disturbed gravitropism.			L1			L1
Average all phenotypes	553	69.7	L1	569	88.2	L1

NDI indices

NDI indices were constructed and compared to the control group. Figure 7 shows a comparison of the mature leaf of treatment 14 through all NDI combinations during all time points. In this sequence, the gap between the groups was observed at time point 3 (the time when significant difference appeared). This example resulted in 1,131 significant indices that were able to separate treatment 14 leaf L1 from the control leaf L1. These indices include combinations of wavelengths 692 nm-750 nm that belong to the red edge.

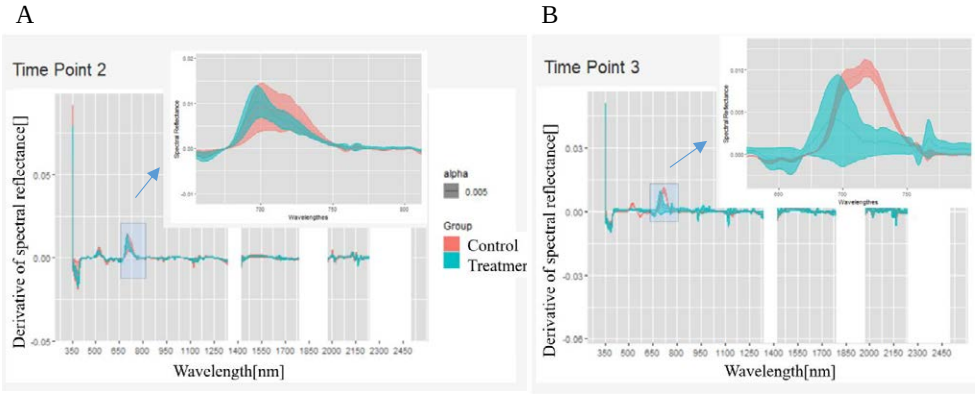


Figure 6. First Derivative spectral data of T16 and control plants at (A) TP2 and (B) TP3.

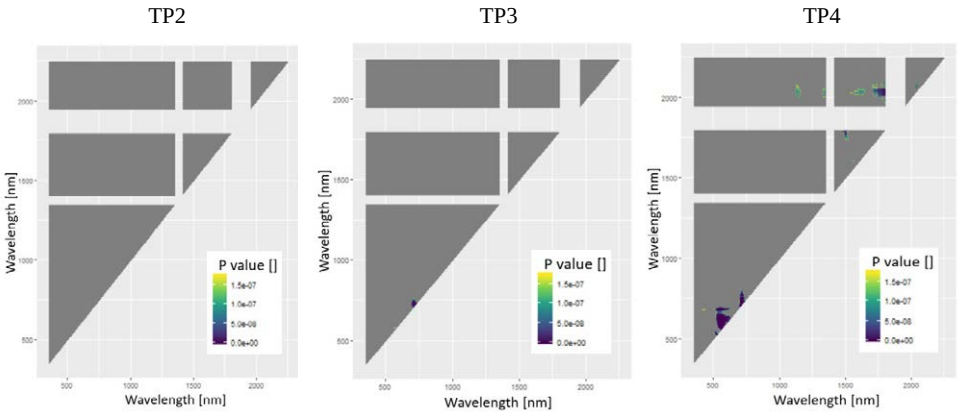


Figure 7. NDI maps demonstrate the difference between treatment 14, leaf 'L1' in corn plants to control 'L1' corn plants through three time points.

Conclusions

Inhibition of photosynthesis was successfully detected using leaf spectral reflectance at the same time when visual phenotypes appeared. The change was detected using the first derivative of a single wavelength and by NDI indices. The first NDI indices which detected the effect of the stress, consisted of pairs of close wavelengths. Linear regression predicting visual changes in corn revealed prediction accuracy of 0.64 and 0.79 for lipid metabolism and photosynthesis inhibition respectively. Moreover, linear regression for predicting the visual changes two days after the spectral measurement resulted in accuracies of 0.26 and 0.67 for lipid metabolism and photosynthesis inhibition respectively. Inhibition of lipid metabolism was detected using spectral reflectance in corn plants after visual symptoms were present. In sunflower, visual symptoms appeared only at the end of the sampling period, and they were not expressed in leaf spectral reflectance. The effect of both MOAs on leaf spectral reflectance included the spectral signature at the red edge area, as well as wavelengths in the SWIR range. This was a preliminary analysis to determine significant areas of the spectrum which can be used to detect specific plant stress. However, sampling frequency should be increased near the sensitive time (time point 3); this may enable earlier detection.

Acknowledgements

This work was partially supported by the Innovation Authority Phenomics Project number 580653293 and by the Ben-Gurion University Rabbi W. Gunther Plaut Chair in Manufacturing Engineering.

References

Ashourloo, D., Mobasheri, M. R., & Huete, A., 2014. Developing two spectral disease indices for detection of wheat leaf rust (Pucciniatriticina). Remote Sensing, 6 (6), 4723-4740. doi:10.3390/rs6064723

Drissi, R., Goutouly, J. P., Forget, D., & Gaudillere, J. P., 2009. Nondestructive measurement of grapevine leaf area by ground normalized difference vegetation index. Agronomy Journal, 101 (1), 226-231. doi:10.2134/agronj2007.0167

Genc, H., Genc, L., & Turhan, H., 2008. Vegetation indices as indicators of damage by the sunn pest (Hemiptera: Scutelleridae) to field grown wheat. African Journal of Biotechnology 7 (2), 173-180. doi:10.4314/ajb.v7i2.58347

Lowe, A., Harrison, N., & French, A. P., 2017. Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. Plant Methods, 13 (1), 1-12. doi:10.1186/s13007-017-0233-z

Savitzky, A., & Golay, M. J. E., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry, 36 (8), 1627-1639. doi:10.1021/ac60214a047

Schuerger, A. C., Capelle, G. A., Di Benedetto, J. A., Mao, C., Thai, C. N., Evans, M. D., *et al.,* 2003. Comparison of two hyperspectral imaging and two laser-induced fluorescence instruments for the detection of zinc stress and chlorophyll concentration in bahia grass. Remote Sensing of Environment, 84. doi:10.1061/(ASCE)EE.1943-7870.0001402.

Zhu, L., Chen, Z., Wang, J., Ding, J., Yu, Y., Li, J., *et al.,* 2014. Monitoring plant response to phenanthrene using the red edge of canopy hyperspectral reflectance. Marine Pollution Bulletin, 86 (1-2), 332-341. doi:10.1016/j.marpolbul.2014.06.046.

Hyperspectral imaging application under field conditions: assessment of the spatio-temporal variability of grape composition within a vineyard

S. Gutierrez, M.P. Diago, J. Fernandez-Novales and J. Tardaguila*
Televitis Research Group. University of La Rioja, 26007 Logroño, Spain; javier.tardaguila@unirioja.es

Abstract

This work illustrates how hyperspectral imaging, carried out from a moving ground vehicle, can be used under field conditions for assessing grape composition and evaluating its spatio-temporal variability within a vineyard. To analyse the distribution of the grape composition parameters, measurements were carried out at four different dates during the ripening season, at which total soluble solids (TSS) and anthocyanin concentrations analyses were also performed. Prediction models were generated for the two parameters, yielding 10-fold cross validation determination coefficients (R^2) of 0.92 for TSS and 0.72 for anthocyanin concentrations (RMSEs of 1.444 °Brix and 0.272 mg/g berry, respectively). The results were used for mapping the evolution and distribution of grape composition in time-four dates-and space-the vineyard plot, illustrating the suitability of this technology for the in-field monitoring of the vineyard.

Keywords: new sensing technologies, non-invasive sensors, mapping, total soluble solids, anthocyanins, in-field monitoring

Introduction

Hyperspectral imaging (HSI) is a powerful technology currently used for many food and agricultural applications (Park and Lu, 2015; Sun, 2010). HSI combines the advantages of regular imaging, producing a high amount of data in an image, also gathering spatial information; and the high-dimensional information that is provided by spectroscopy. Several traits can be monitored using hyperspectral imaging. One of the most promising applications of this technology is for the estimation of different fruit composition parameters. Hyperspectral cameras have been used for the prediction of total soluble solids (TSS) in apple (Tian *et al.,* 2018), mango (Rungpichayapichet *et al.,* 2017) or grape (Piazzolla *et al.,* 2017). Anthocyanin concentration (AC) has also been monitored using hyperspectral imaging in grapes (Diago *et al.,* 2016; Zhang *et al.,* 2017) and raspberry (Rodríguez-Pulido *et al.,* 2017). These studies demonstrate that these chemical parameters can be successfully estimated from spectroscopy in different spectral ranges and modelled using different machine learning techniques.

Recent studies have demonstrated the possibility of deploying ground-based hyperspectral cameras under field conditions (Gutiérrez *et al.,* 2018, 2019b; Wendel *et al.,* 2018). Many challenges for this deployment have been addressed, due to the new uncontrolled conditions hyperspectral cameras are subjected to. Illumination compensation techniques have been developed to reduce the necessity of using references (Wendel and Underwood 2017a); while techniques for extrinsic parametric calibration have also been reported to take into account scene geometry (Wendel and Underwood 2017b). The use of hyperspectral imaging especially acquired on-the-go from ground vehicles allow the acquisition of high-resolution data. This feature is very useful when precision mapping is desired, as this can be used as a powerful tool for growers in the decision-making process.

TSS and AC are two important grape composition parameters useful for the monitoring of the optimal harvest time (Bramley *et al.,* 2011) and determining wine quality (Iland, 2004). Hand-held devices have been developed for the quick characterization of TSS, still sampling is required if a

DOI 10.3920/978-90-8686-888-9_35, © Wageningen Academic Publishers 2019

high area is sought to be measured. In the case of AC, its analysis require of complex wet chemistry methodologies (Iland, 2004), with the need of specialized personnel and high cost materials. A potential non-destructive monitoring system based on hyperspectral imaging from ground vehicles would be very valuable for the quick characterization of large areas within a vineyard, to generate maps illustrating the spatial variability of these two parameters.
The objective of this work was to develop a system for mapping the spatio-temporal variability of vineyard grape composition using on-the-go hyperspectral imaging acquired from a ground vehicle during several dates within the grape ripening period.

Materials and methods

Figure 1 shows the experimental pipeline followed in this study. The study was conducted in a commercial Tempranillo vineyard (*Vitis vinifera* L.) located in Ábalos, La Rioja, Spain (Lat. 42° 34' 45.7", Long. -2° 42' 27.78", asl 628 m). Grapevines were planted on rootstock R-110 in 2010 and trained to a vertical shoot positioned (VSP) trellis system. Row spacing was 2.2×1.0 m, inter and intra-row, with Northeast-Southwest orientation. From the vineyard, three different rows were selected and, within each one of them, 12 blocks containing five plants were marked for data acquisition. Figure 2 displays the design of the experimental layout in the field.
The acquisition of hyperspectral images was carried out on-the-go using a Resonon Pika L VNIR hyperspectral imaging camera (Resonon, Inc., Bozeman, MA, USA), working in the range of 400-1000 nm. The Pika L camera works using a push broom scanning system with 300 pixels of spatial resolution at a spectral resolution of 2.1 nm (300 bands). The camera was installed on an all-terrain vehicle (ATV) (Trail Boss 330, Polaris Industries, MN, USA), pointing perpendicular to the canopy, at 2.0 m distance. Using an 8 mm focal length lens, the recording line cast upon the plants projected a field of view of 36.5°, covering a length of 1.32 m from top to bottom. Special attention was paid so as to assure that the grapevines fruiting zone lay completely within the recording line. Measurements were performed on the southeast side of the canopy.
At each hyperspectral acquisition, depending on the light conditions, the camera's configuration parameters (frames per seconds and integration time) were adapted to maximize the received signal while avoiding any spectral saturation in the sensor. Before block measurement, and to convert the raw spectra to reflectance values, white reference acquisition was performed by manually presenting a Spectralon® with a reflectance value over 95% and at the same distance and position as the canopy. Block measurement was carried out maintaining an ATV's speed of 5 km/h.

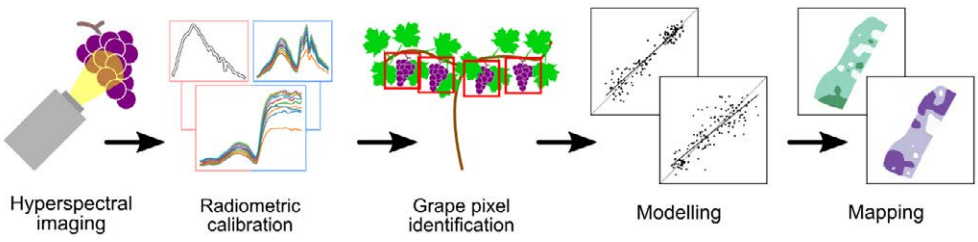


Figure 1. Experimental pipeline followed in this study. On-the-go hyperspectral imaging was performed under field conditions, followed by radiometric calibration using a white reference. Afterwards, an automated process for grape pixel identification was designed, to enable the development of machine learning prediction models. Finally, mapping was carried out for the spatio-temporal analysis of grape composition in the vineyard.

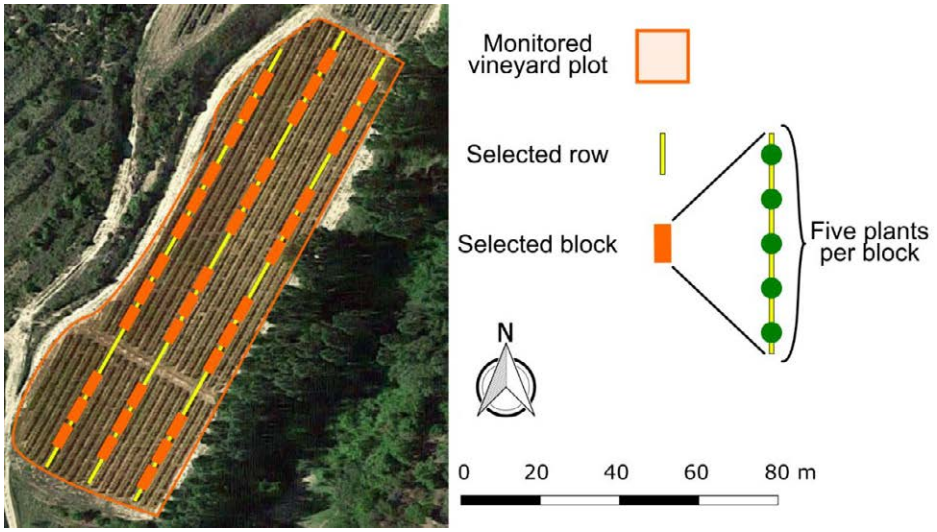


Figure 2. Design of the experimental layout in a commercial vineyard.

To develop a model for the prediction of grape composition at different grape phenological stages, data acquisition was performed at four different times through the ripening period, from veraison to harvest. Therefore, having four data acquisition dates, three grapevine rows and 12 blocks per row, a total of 144 measurements were taken. Each block measurement comprised the imaging of the five plants with the hyperspectral camera and immediately afterwards the grape berries were sampled to have them analysed for their composition values of TSS and anthocyanins, as reference values. Grape berries were manually picked by sampling groups of 10-15 visible berries from the exposed bunches in the canopy. Total soluble solids (TSS) and total anthocyanin concentration (AC) were determined using refractometry and the Iland method (Iland, 2004), respectively, and assigned to each block.

A hyperspectral image processing was designed for the automated selection of grape berries' spectra. This procedure, similarly to Spectral Angle Mapper (Dennison *et al.,* 2004), consisted of the comparison of each image pixel (spectrum) with a previously taken grape berry signature spectrum. If the correlation between both spectra surpassed an R^2 score of 0.75, the pixel is considered as grape pixel. Support vector machines were used for the development of the prediction models, validated with a 10-fold stratified cross validation. For further details on these points, the reader is referred to (Gutiérrez *et al.,* 2019a).

The prediction results from the models were bound to their corresponding block in the vineyard plot for the generation of grape composition maps. The mapping for each measurement stage was carried out using QGIS 2.19 (Free Software Foundation, Boston, MA, USA) and multilevel b-spline interpolation (Lee *et al.,* 1997). Interpolation was used to illustrate a better coverage of all points within the vineyard.

Results and discussion

The regression plots from the 10-fold cross validation are shown in Figure 3 for TSS (A) and AC (B). The cross-validation results from TSS modelling returned a R^2 of 0.89 (RMSE=1.444 °Brix). As it can be seen in Figure 3A, samples were well distributed along the regression line, and a clear cluster was found at the top right part of the plot. This cluster, that gathers those samples with higher

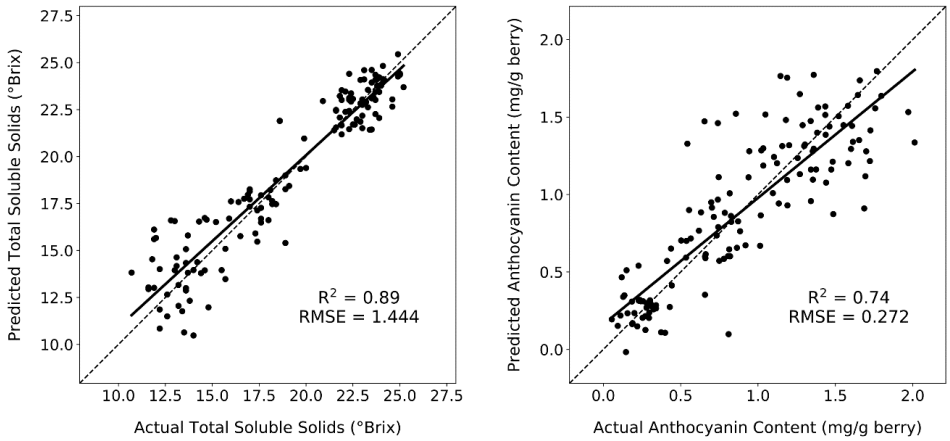


Figure 3. Correlation plots for the 10-fold cross validation of total soluble solids (A) and anthocyanin content (B) modelling. Dashed line corresponds to 1:1.

TSS values, correspond to samples measured on the last two days of the experiment, in which grape berry sugar content is higher. Samples from 10 to 20 °Brix exhibited a higher level of dispersion in both actual and predicted values, meaning that more variability can be found from the beginning to mid-ripening. Still, the average response would suggest that a spectral TSS prediction model can deliver satisfactory precision and accuracy for its estimation at any phenological stage between veraison and harvest. A slightly higher precision was found when predicting higher values of TSS (top-right part of Figure 3A), as higher variability in actual TSS was present at lower values.

The regression plot from AC cross validation (Figure 3B), with R^2 of 0.74 (RMSE=0.272 mg/g berry), showed improved precision (i.e. higher fit to the regression line) for predicting samples with low AC – bottom left part of the plot – than for those with a large anthocyanin content – top right area. This suggests that this methodology can be successfully applied early in the season (post-veraison), when AC values are very low, and a larger amount of fruit sample is required to assess the AC content using wet chemistry methods, to yield estimated values of AC with very good precision.

These results confirm the suitability of spectroscopy for the estimation of grape composition reported by other studies under laboratory and field conditions (Barnaba *et al.,* 2014; Bellincontro *et al.,* 2011; Musingarabwi *et al.,* 2016; Urraca *et al.,* 2016). Although this technology is not as accurate as laboratory measurements, it is good enough and suitable for quick determination of these parameters. Moreover, the non-destructive nature of hyperspectral imaging and the fact that it can be performed in the field at a commercial speed to provide a huge amount of data are also relevant advantages towards its commercial implementation. The addition of an on-the-go approach allow for the quick monitoring of large areas in a vineyard but still acquiring high resolution data, that can be translated into more detailed distribution maps.

Figure 4 shows eight distribution maps for each one of the two grape composition parameters at each measurement date.

As can be seen in Figure 4A,B, the TSS evolution started earlier at the southern part of the vineyard. At 18 September (Figure 4C), the north and mid zones of the vineyard also exhibited grapes with the highest °Brix values, while at 28 September (Figure 4D) TSS accumulation was mostly completed, with more steady values in almost all zones of the plot. At the end, TSS values from the grape bunches were in most cases above 23.2 °Brix, as it can be expected at so late a stage.

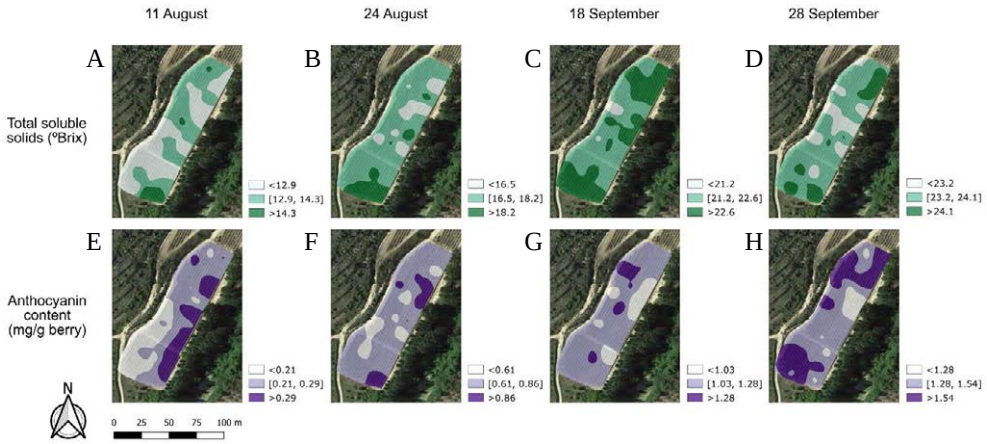


Figure 4. Spatio-temporal maps displaying the predicted distribution of grape composition within the vineyard plot.

Analysing the evolution of AC (Figure 4E to 4H, bottom maps), the first date (Figure 4E) stands out by the higher values in the eastern rows, while in both 24 August (Figure 4F) and 18 September (Figure 4G), the grapes underwent a more balanced anthocyanin accumulation. At the end of the campaign, the AC values still exhibited a high variability, with zones in the north and south above 1.54 mg/g berry and some areas in the middle part of the vineyard with a berry anthocyanin content below 1.28 mg/g berry.

The high-resolution information from on-the-go hyperspectral imaging enabled the generation of grape composition distribution maps. This brings useful data for the farmer for the delineation of zones for setting different harvesting dates (in the case of TSS) or assigning grapes to different quality level wine-making procedures.

Conclusions

The deployment of on-the-go hyperspectral imaging has been possible for the mapping of two key grape composition parameters. This technology, after overcoming many of the challenges of its usage under field conditions, allows for the acquisition of high-resolution data that, after modelling, can be translated into grape composition maps. These maps become a very useful tool for grape growers in the decision-making process, as efficiency in harvesting and wine production is increased.

Acknowledgements

SG would like to acknowledge the research funding FPI grant 299/2016 by Universidad de La Rioja, Gobierno de La Rioja. MD is funded by the Spanish Ministry of Science, Innovation and Universities, with a Ramon y Cajal grant RYC-2015-18429.

References

Barnaba, F.E., Bellincontro, A. and Mencarelli, F., 2014. Portable NIR-AOTF spectroscopy combined with winery FTIR spectroscopy for an easy, rapid, in-field monitoring of Sangiovese grape quality. Journal of the Science of Food and Agriculture 94 1071-1077.

Bellincontro, A., Cozzolino, D. and Mencarelli, F., 2011. Application of NIR-AOTF spectroscopy to monitor Aleatico grape dehydration for Passito wine production. American Journal of Enology and Viticulture 62 256-260.

Bramley, R.G. V, Le Moigne, M., Evain, S., Ouzman, J., Florin, L., Fadaili, E.M., *et al.,* 2011. On-the-go sensing of grape berry anthocyanins during commercial harvest: development and prospects. Australian Journal of Grape Wine Research 17 316-326.

Diago, M.P., Fernández-Novales, J., Fernandes, A.M., Melo-Pinto, P. and Tardaguila, J., 2016. Use of visible and short-wave near-infrared hyperspectral imaging to fingerprint anthocyanins in intact grape berries. Journal of Agriculture and Food Chemistry 64 7658-7666.

Gutiérrez, S., Fernández-Novales, J., Diago, M.P. and Tardaguila, J., 2018. On-the-go hyperspectral imaging under field conditions and machine learning for the classification of grapevine varieties. Frontiers of Plant Sciences 9 1102.

Gutiérrez, S., Tardaguila, J., Fernández-Novales, J. and Diago, M.P., 2019a. On-the-go hyperspectral imaging for the in-field estimation of grape composition. Australian Journal of Grape and Wine Research 25 (1) 127-133.

Gutiérrez, S., Wendel, A. and Underwood, J (2019b) Ground based hyperspectral imaging for extensive mango yield estimation. Computers and Electronics in Agriculture 157 126-135.

Iland, P., 2004. Chemical analysis of grapes and wine. Athelstone, Australia: Patrick Iland Wine Promotions PTY LTD,.

Lee, S., Wolberg, G. and Shin, S.Y., 1997. Scattered data interpolation with multilevel B-splines. IEEE Transactions on Visualization and Computer Graphics 3 228-244.

Musingarabwi, D.M., Nieuwoudt, H.H., Young, P.R., Eyéghè-Bickong, H.A. and Vivier, M.A., 2016. A rapid qualitative and quantitative evaluation of grape berries at various stages of development using Fourier-transform infrared spectroscopy and multivariate data analysis. Food Chemistry 190 253-262.

Park, B. and Lu, R., 2015. Hyperspectral imaging technology in food and agriculture. New York, USA: Springer-Verlag,.

Piazzolla, F., Amodio, M.L. and Colelli, G., 2017. Spectra evolution over on-vine holding of Italia table grapes: prediction of maturity and discrimination for harvest times using a Vis-NIR hyperspectral device. Journal of Agricultural Engineering 48 109-116.

Rodríguez-Pulido, F.J., Gil-Vicente, M., Gordillo, B., Heredia, F.J. and González-Miret, M.L., 2017. Measurement of ripening of raspberries (Rubus idaeus L) by near infrared and colorimetric imaging techniques. Journal of Food Science and Technology 54 2797-2803.

Rungpichayapichet, P., Nagle, M., Yuwanbun, P., Khuwijitjaru, P., Mahayothee, B. and Müller, J., 2017. Prediction mapping of physicochemical properties in mango by hyperspectral imaging. Biosystems Engineering 159 109-120.

Sun, D.W., 2010. Hyperspectral imaging for food quality analysis and control. London, UK: Academic Press Inc.,.

Tian, X., Li, J., Wang, Q., Fan, S. and Huang, W., 2018. A bi-layer model for non-destructive prediction of soluble solids content in apple based on reflectance spectra and peel pigments. Food Chemistry 239 1055-1063.

Urraca, R., Sanz-Garcia, A., Tardaguila, J. and Diago, M.P., 2016. Estimation of total soluble solids in grape berries using a hand-held NIR spectrometer under field conditions. Journal of the Science of Food and Agriculture 96 3007-3016.

Wendel, A. and Underwood, J., 2017a. Illumination compensation in ground based hyperspectral imaging. ISPRS Journal of Photogrammetry and Remote Sensing 129 162-178.

Wendel, A. and Underwood, J., 2017b. Extrinsic parameter calibration for line scanning cameras on ground vehicles with navigation systems using a calibration pattern. Sensors 17 2491-2518.

Wendel, A., Underwood, J. and Walsh, K., 2018. Maturity estimation of mangoes using hyperspectral imaging from a ground based mobile platform. Computers and Electronics in Agriculture 155 298-313.

Zhang, N., Liu, X., Jin, X., Li, C., Wu, X., Yang, S., *et al.,* 2017. Determination of total iron-reactive phenolics, anthocyanins and tannins in wine grapes of skins and seeds based on near-infrared hyperspectral imaging. Food Chemistry 237 811-817.

Potential of optical sensors for nitrogen management in spring barley

R. Hackett
Teagasc, Crops Research Centre, Oak Park, Carlow, Ireland; richie.hackett@teagasc.ie

Abstract

Fertiliser nitrogen (N) decisions in Ireland for spring barley are made based on historical information (e.g. previous crop yield and previous crop grown) and no account is taken of the growth or N dynamics of the current crop. Spectral reflectance measurements of crops can indicate the N status of crops and may be useful for guiding N inputs to barley crops in a more site-specific way. This work examined the potential of an active reflectance sensor to indicate biomass and N status characteristics of spring barley crops during the growing season. Results indicate that reflectance measurements were well related to both N uptake and nitrogen nutrition index in spring barley and may be useful to guide N inputs to the crop.

Keywords: spring barley, nitrogen nutrition index, nitrogen fertiliser, reflectance

Introduction

Spring barley is currently the main arable crop in Ireland with the majority being grown for animal feed but an increasing proportion being grown for use in malting. Currently only historical data (past yields and cropping history) is used to determine fertiliser N inputs, which have a large influence on grain protein concentration and consequently suitability of the grain for malting in Ireland and no account is taken of the growth or N status of the current crop.

It is well established that the nitrogen concentration of the aboveground biomass of a crop decreases as the crop progresses through the vegetative growth stages principally due to the increasing proportion of structural tissues which have a low N concentration and which 'dilute' the high N concentration of leaf laminae (Lemaire and Gastal, 1997). Critical N dilution curves, which indicate the minimum N concentration of the aboveground biomass required to maximise biomass production, have been described for a number of crops, including barley (Greenwood *et al.*, 1991, Justes *et al.*, 1994, Zhao, 2014). The nitrogen nutrition index (NNI) which is calculated as the ratio between the measured N concentration and the critical N concentration derived from the N dilution curve, can be used to describe the relative N status of a crop canopy.

Generally good correlations have been found between spectral measurements and N status indicators such as NNI. Mistele and Schmidhalter (2008) reported canopy reflectance could give a good indication of NNI in wheat but did indicate that the relationship could vary between seasons. Zhao *et al.* (2018) reported good relationships between canopy reflectance and NNI for maize. However, relatively little work has been carried out examining the ability of reflectance measurements to indicate the N status of spring barley in high yielding temperate climates.

The objective of this study was to determine the relationship between reflectance measurements and biomass and N characteristics of spring barley at a range of growth stages between early stem extension and the grain filling stage of crop growth.

Materials and methods

An experiment was carried out at the Teagasc Crops Research Centre, Carlow, Ireland, (52.86° N, 6.92° W, 62 m a.m.s.l.) in each of three seasons 2014, 2015 and 2016. A different site, but with a

DOI 10.3920/978-90-8686-888-9_36, © Wageningen Academic Publishers 2019

similar soil type, was used in each year. The same cultivar of spring barley (cv. Tavern) was used in all experiments. Fertiliser N rates ranging from 0 to 240 kg N/ha, with 30 kg N/ha increments, were included in the trial where 30 kg N/ha of the total was applied at sowing and the remainder applied during the tillering stage. All fertiliser was applied as CAN (Calcium ammonium nitrate). Crops received all other inputs as per standard farm practice. The experiments also included a range of additional treatments with applications of fertilizer N at later growth stages as described in Hackett (2018) but data from these plots are not used here and therefore no further description of these treatments is given. Each experiment had four replications and plot size was 2.3×12 m. To facilitate repeated crop sampling, duplicate, adjacent, plots of each treatment of the N response were included. Crop reflectance, biomass accumulation and N concentration of the biomass were determined at four growth stages (GS 31, GS 39, GS 61 and GS 77 determined according to Zadocks *et al.* (1974)) in each experiment, with the exception that GS 77 measurements were not made in 2014. Canopy reflectance was measured using a hand held Crop Circle ACS 470 instrument (Holland Scientific Inc., Lincoln, Nebraska, USA). Reflectance was determined at 760, 730 and either 670 nm (2014 and 2015) or 550 nm (2016). At each measurement date, 25-35 readings were made along the length of each plot at a height of ~0.7 m above the crop canopy, avoiding plot edges. The normalised difference red edge vegetation index was calculated as:

$$NDRE = \frac{R_{760} - R_{730}}{R_{760} + R_{730}} \quad (1)$$

where R_{760} and R_{730} are reflectance at 760 and 730 nm respectively.
For biomass and N determination, plants in three adjacent 0.5 m row lengths at two locations in each plot were cut at ground level and dried in a forced air oven at 70 °C for 48 h. Dry weight of the samples was obtained and samples were subsequently milled (2 mm sieve) and their N concentration determined using the Dumas method (LECO FP327, Leco Corp, St. Joseph, MI, USA). Total N accumulation was calculated as dry matter yield multiplied by the N concentration. The nitrogen nutrition index (NNI) was calculated as:

$$NNI = \frac{N_{act}}{N_{crit}} \quad (2)$$

Where N_{act} is the measured N concentration of the crop biomass and N_{crit} is the critical N content for that crop given its biomass calculated according to Zhao (2014).
The relationship between NDRE and crop biomass, N concentration, crop N accumulation and NNI was examined using regression analysis. Linear and quadratic effects were sequentially included in the models and the quadratic effect was only retained when significant. In addition to analysing data for individual growth stages for each season, data for a given growth stage combined over seasons, data for individual seasons combined over growths stages and data combined over seasons and growth stages was also analysed. Statistical analysis was carried out using PROC GLM (SAS 9.4, SAS Institute Inc., Cary, NC, USA)

Results

A linear relationship between NDRE and biomass was found at GS 31 in all seasons and at GS 39 in the 2015 season (Table 1, Figure 1). Curvilinear relationships were found in 2014 and 2016 at GS 39 and for all seasons at GS 61 and GS 77. The range between the maximum and minimum biomass values increased to a much larger extent than the corresponding range in NDRE values giving rise to differences in the shape of the relationships. As a result, while high r^2 values were recorded for a given growth stage in a given season, when data for all seasons and growth stages were combined, there was no significant relationship between NDRE and biomass.

Table 1. Correlation coefficients (r^2) and root mean square error (RMSE) of models of the relationship between NDRE and crop biomass and N status indicators at four growth stages and three seasons. Quadratic models were fitted except where indicated with an asterisk.

Seaons	GS31		GS39		GS61		GS77		Combined	
	r^2	RMSE	r^2	RMSE	r^2	RMSE	r^2	RMSE	r^2	RMSE
Biomass (g DM/m^2)										
2014	0.64*	32.03	0.85	48.23	0.85	71.50	-	-	0.06	273.26
2015	0.68*	32.20	0.75*	60.15	0.89	62.77	0.86	91.11	0.14	376.54
2016	0.64*	31.62	0.79	41.30	0.92	51.06	0.94	72.05	0.15	261.45
Combined	0.62	33.07	0.57	80.57	0.76	93.56	0.72	167.80	0.10	316.71
Crop N concentration (g N/100 g DM)										
2014	0.85	0.44	0.87	0.31	0.83	0.22	-	-	0.60	0.81
2015	0.79	0.48	0.94	0.16	0.93	0.11	0.92	0.10	0.66	0.73
2016	0.90	0.37	0.95	0.21	0.81	0.26	0.96	0.07	0.54	0.70
Combined	0.84	0.46	0.83	0.36	0.82	0.22	0.90	0.10	0.61	0.76
Crop N accumulation (g N/m^2)										
2014	0.91	1.18	0.89	2.10	0.87	2.27	-	-	0.57	3.75
2015	0.90	1.15	0.93	1.20	0.93	1.27	0.91	1.45	0.52	3.82
2016	0.89	1.28	0.94	1.24	0.87	2.19	0.96	1.32	0.70	2.39
Combined	0.88	1.29	0.88	1.85	0.86	2.16	0.90	1.78	0.60	3.35
Nitrogen nutrition index										
2014	0.90	0.12	0.89	0.13	0.86	0.11	-	-	0.86	0.14
2015	0.87	0.13	0.95	0.06	0.94	0.06	0.93	0.06	0.89	0.12
2016	0.92	0.11	0.96	0.07	0.85	0.12	0.96	0.04	0.84	0.13
Combined	0.88	0.13	0.90	0.12	0.86	0.11	0.94	0.05	0.86	0.13

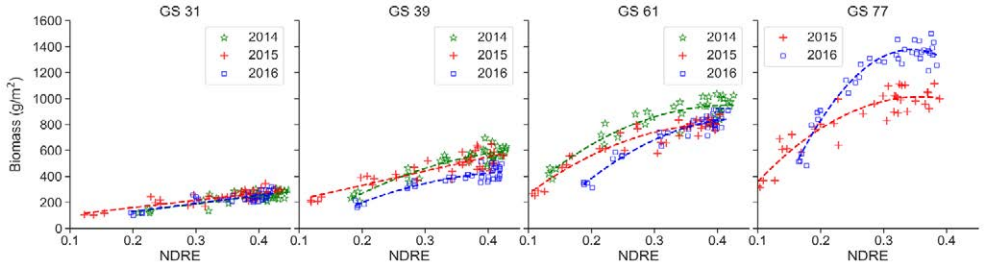


Figure 1. Relationship between NDRE and biomass of spring barley for four stages of growth over three seasons. Dashed lines represent regression between biomass and NDRE.

Significant quadratic relationships, with r^2 values mostly in excess of 0.80, were detected between N concentration of the biomass and NDRE in all seasons for all growth stages (Table 1, Figure 2). The shape of the curves was similar between seasons for a given growth stage and RMSE values where the data for a given growth stage were combined across seasons were similar to those obtained for the individual seasons. However the shape of the curves differed between growth stages with the slope of the curves decreasing as the crop progressed through its growth stages. As a result the r^2

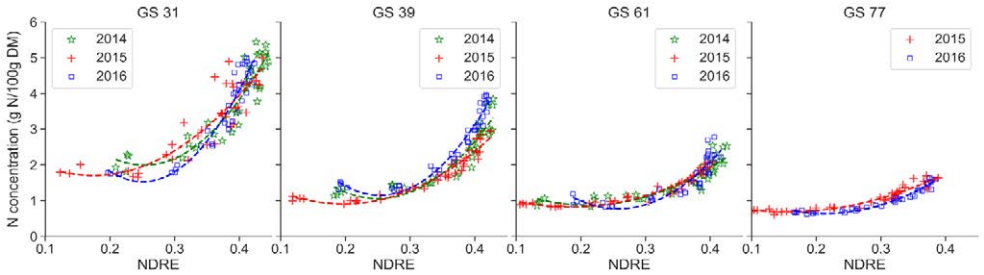


Figure 2. Relationship between NDRE and N concentration of the biomass of spring barley for four stages of growth over three seasons. Dashed lines represent regression between N concentration and NDRE.

values where data was combined over growth stages for a given season or over seasons were lower (0.54-0.66) and RMSE values were greater.

Significant quadratic relationships, with r^2 values mostly in excess of 0.85, were detected between N accumulation in the aboveground biomass and NDRE in all seasons for all growth stages (Table 1, Figure 3). When data for the three seasons for a given growth stage were combined, high r^2 values were also detected and RMSE values were similar to those observed for the individual sampling dates. When data were combined across growth stages r^2 values decreased (0.52-0.70) and RMSE values increased.

A strong relationship between NDRE and NNI was detected for all growth stages in all seasons (r^2=0.88-0.96). The shape of the curves for the three seasons for a given growth stage were similar such that when a common curve was fitted to the data for a given growth stage across seasons, high r^2 values (0.86-0.94) were obtained and RMSE values were similar to those obtained for individual seasons (Table 1, Figure 4). The range of values for NNI decreased as the crop progressed through growth stages but this was accompanied by a reduction in the range of values obtained for NDRE. As a result when data were combined across growth stages, either for an individual season or for the three seasons combined, r^2 values of 0.84-0.89 were obtained and RMSE values were of a similar magnitude to those obtained for individual season/growth stage combinations.

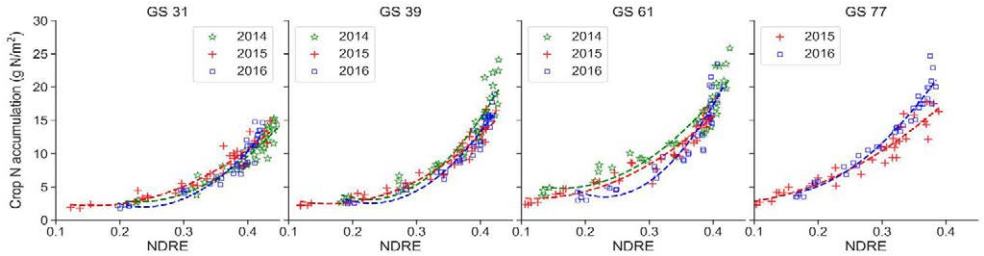


Figure 3. Relationship between NDRE and aboveground N accumulation of spring barley for four stages of growth over three seasons. Dashed lines represent regression between N accumulation and NDRE.

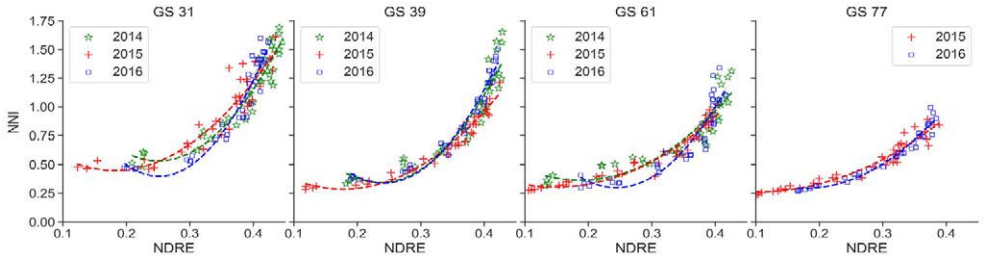


Figure 4. Relationship between NDRE and NNI of spring barley for four stages of growth over three seasons. Dashed lines represent quadratic regression between NNI and NDRE.

Discussion

Fertiliser N is typically applied to spring barley during the early growth stages and, in many cases, fertiliser N inputs will have been completed before GS31, the earliest growth stage considered in this work. However recent work has indicated that N inputs to spring barley can be made at later growth stages without negatively influencing grain yield or protein content provided a large deficiency has not occurred (Hackett, 2018; Beillouin *et al.,* 2018), suggesting that the first two sampling dates in this study are relevant in terms of guiding N inputs.

Critical N dilution curves were developed to describe the changes in N concentration that occurred up to anthesis, where the high N concentration of leaves was 'diluted' by the increasing amounts of low N concentration structural tissues. Therefore calculation of NNI for samples taken during grain fill may not be a valid measure of the N status of plants due to the accumulation of N in the grain. However, despite this, the relationship between NDRE and NNI at GS 77 was similar to that observed at earlier growth stages in this study. Knowledge of the N status of the crop at this stage may be particularly useful for indicating protein concentration of the mature grain.

While a good relationship between N accumulation and NDRE was obtained for a given growth stage, the poorer relationships obtained when data from different growth stages was combined suggests that growth stage-specific relationships between reflectance measurements and N accumulation would be needed if this information is to be used to reliably guide N inputs. In any case, N accumulation does not provide any direct indication of the N sufficiency or deficiency of the crop. NNI, on the other hand, indicates whether or not a canopy has a sufficient N concentration for maximum growth in relation to its actual biomass. Therefore it provides information on whether or not a crop requires more N, although it provides no information as to the quantity required. The close relationship between NDRE and NNI found in these experiments suggests that measurements with an active optical sensor can be useful for indicating the nitrogen status of spring barley. Of particular interest was the relatively constant relationship that existed between NDRE and NNI when data were combined over a wide range of growth stages from early stem extension to mid grain fill suggesting that growth stage specific relationships would not be required. Similar close relationships between NNI and reflectance measurements, including wavelengths in the red edge part of the spectrum, were obtained by Mistele and Schmidhalter (2008) and Erdle *et al.* (2011) for winter wheat.

While NNI can be used to indicate a requirement for additional N inputs, it gives no information as to the quantity of additional N required. An estimate of crop biomass could be used in conjunction with NNI to calculate the quantity of additional crop N required. In this study, there was a weaker relationship between NDRE and biomass than between NDRE and NNI for a given growth stage. In addition, the relationship was not stable across growth stages indicating that a growth stage-specific

relationship would need to be developed. Mistele *et al.* (2004) also reported a poorer relationship between reflectance measurements and biomass than between reflectance and N status.
A key issue with the relationships between NDRE and crop parameters described here is their curvilinear nature, indicating reduced sensitivity of the reflectance measurements to changes as the NNI or N accumulation increases. However detection of crops with sub-optimal N status, where reflectance measurements have higher sensitivity to change is more important, when determining the requirement for further fertiliser N inputs.

Conclusions

The work presented here indicates that reflectance measurements could be used to indicate the nitrogen status of spring barley at a range of growth stages. This information could be useful to guide N inputs to crops. However, the robustness of the relationships between reflectance measurements and N status over a greater number of sites needs to be investigated.

References

Beillouin, D., Trépos, R., Gauffreteau, A. and Jeuffroy, M.H., 2018. Delayed and reduced nitrogen fertilization strategies decrease nitrogen losses while still achieving high yields and high grain quality in malting barley. European Journal of Agronomy 101 174-182.

Erdle, K., Mistele, B. and Schmidhalter, U., 2011. Comparison of active and passive spectral sensors in discriminating biomass parameters and nitrogen status in wheat cultivars. Field Crops Research 124 (1) 74-84.

Greenwood, D.J., Gastal, F., Lemaire, G., Draycott, A., Millard, P. and Neeteson, J.J., 1991. Growth rate and % N of field grown crops: theory and experiments. Annals of Botany 67 (2) 181-190.

Hackett, R., 2018. Effect of delayed N application on spring barley productivity. In: Proceedings of the 20th Nitrogen workshop, Rennes, France: INRA pp. 274-275.

Justes, E., Mary, B., Meynard, J.M., Machet, J.M. and Thelier-Huché, L., 1994. Determination of a critical nitrogen dilution curve for winter wheat crops. Annals of Botany 74 397-407.

Lemaire, G. and Gastal, F., 1997. N uptake and distribution in plant canopies. In: G. Lemaire (Ed.) Diagnosis of the Nitrogen Status in Crops, Heidelberg, Germany: Springer, pp. 3-43

Mistele, B. and Schmidhalter, U., 2008. Estimating the nitrogen nutrition index using spectral canopy reflectance measurements. European Journal of Agronomy 29(4) 184-190.

Mistele, B., Gutser, R., Schmidhalter, U., 2004. Validation of field-scaled spectral measurements of nitrogen status in winter wheat. In: D. Mulla (ed.) Proceedings of the 7th International Conference on Precision Agriculture and other Precision Resources Management, Minneapolis, USA, pp. 1187-1195.

Zadoks, J.C., Chang, T.T., and Konzak, C.F., 1974. A decimal code for the growth stages of cereals. Weed Research 14 415-421.

Zhao, B., 2014. Determining of a critical dilution curve for plant nitrogen concentration in winter barley. Field Crops Research 160 64-72.

Zhao, B., Duan, A., Ata-Ul-Karim, S.T., Liu, Z., Chen, Z., Gong, Z. *et al.*, 2018. Exploring new spectral bands and vegetation indices for estimating nitrogen nutrition index of summer maize. European Journal of Agronomy 93 113-125.

Disentangling the sources of chlorophyll-content variability in banana fields

J. Lamour[1,2*], C. Leroux[2,3], G. Le Moguédec[4], O. Naud[2], M. Léchaudel[5,6] and B. Tisseyre[2]
[1]*Compagnie Fruitière, Marseille 13000, France; j.lamour@fruitiere.fr*
[2]*UMR ITAP, Montpellier SupAgro, Irstea, Univ. Montpellier, France*
[3]*SMAG, Montpellier 34960, France*
[4]*INRA, UMR AMAP, Montpellier 34000, France*
[5]*CIRAD, UMR QualiSud, 97130 Capesterre-Belle-Eau, Guadeloupe, France*
[6]*QualiSud, Univ Montpellier, CIRAD, Montpellier SupAgro, Univ d'Avignon, Univ de La Réunion, Montpellier, France*

Abstract

The objective of this study was to evaluate the effect of banana phenological stages on the chlorophyll content of its leaves in order to assess possible restrictions on the potential use of remote sensing for nutritional diagnoses when phenological stages cannot be discriminated on images. Chlorophyll measurements were made with a Dualex sensor at different positions on banana leaves with different ranks on the plant, on plants at different phenological stages and located in different fields of a commercial farm. The effect of these factors was studied using mixed models with nested factors and the results showed that they have a significant effect on the chlorophyll content. It is therefore suggested to take into account the diversity of phenological stages in banana fields before considering nutritional diagnoses services based on remote sensing.

Keywords: banana crop, chlorophyll content, mixed models, spatial analysis, nested factors

Introduction

Chlorophyll is the pigment responsible for the leaf light interception during photosynthesis. As such, there is an interest in measuring the chlorophyll-content to evaluate plant health status and productivity. Amongst all the leaf constituents, a major advantage of chlorophyll is that its concentration can be estimated through optical measurements. This property has led to the development of multiple proxy (SPAD®, Dualex®, N-sensor®) and remote (multi and hyper-spectral images) sensors. An empirical relation has also been shown between the chlorophyll-content and the leaf nitrogen content. Some studies have therefore tried to estimate nitrogen nutritional status from chlorophyll optical measurements to avoid time-consuming and costly nitrogen laboratory measurements. These studies were made by proxy detection (Debaeke *et al.,* 2006; Prost and Jeuffroy, 2007) or remote sensing studies (Haboudane *et al.,* 2002; Gitelson *et al.,* 2005). Operational services, like Farmstar® that rely on this principle, exist to manage crop fertilization in arable crops such as wheat or barley (Coquil and Bordes, 2005).

Applying such approaches on banana crops is not trivial. Contrary to most of the crops for which vegetation cover can be considered uniform (same phenological stage), banana fields contain banana plants at different phenological stages, e.g. before flowering and after flowering. The biological and cropping causes that lead to this phenomenon are detailed in Tixier *et al.* (2004). In simple words, it must be understood that each banana plant blooms only once and that each sucker (the new stem) produced by the banana corm ensures the following blooming (Figure 1). Suckers have different growth rates and flowering dates, leading to a diversity of phenological stages in the population (field, field group). This is called asynchronism and it makes it more difficult to evaluate the banana nutritional status given the nitrogen evolution dynamics within the plants. Indeed, the leaf nitrogen

DOI 10.3920/978-90-8686-888-9_37, © Wageningen Academic Publishers 2019

content varies with the phenological stage, especially because of dilution phenomena (the nitrogen content diminishes when the biomass increases, Lemaire, 1997) but also because of nitrogen reallocation phenomena during fruit filling.
Thus, in the case of a banana crop, it is assumed that the chlorophyll content in a population is not only due to the nutritional status of the banana plants but also to the different phenological stages caused by asynchronism. The main objective of this study was to quantify and compare the influence of these two factors on the banana leaf chlorophyll-content in a commercial plantation. The interest being, in the long-term scale, to evaluate the reliability of remote sensing to make spatialized nutritional diagnoses in banana plantations when information on phenological stages are missing.

Material and methods

Definitions and notations
Within a banana plantation, a field contains banana plants at different phenological stages (Figure 1). Each banana plant has different leaves that are characterized by their leaf rank on the pseudo-trunk. The youngest fully unfolded leaf is located at the top of the banana plant and is the 1-rank leaf. The leaf that appeared just before the 1-rank leaf and situated just below it is the 2-rank leaf, and so forth until the last leaf. The central rib splits the leaf's limb in two sides (Figure 1). The left side is the one on the left when one is looking at the upper face and when the peduncle is the nearest to the operator.

Study site and experiments
The two experiments were carried out on a 3,000 ha commercial banana plantation located near Njombe, Cameroun (WGS84: E:4.612, N: 9.639). Chlorophyll measurements were performed on the leaves of banana plants using the Dualex scientific sensor (FORCE-A, Orsay, France) that estimates it through light transmittance (μg/cm^2).

- Experiment 1: 15 out of the 1,500 fields of the plantation were selected according to the farmers' expertise to obtain fields that they considered representative of the diversity in growing conditions in the plantation, e.g. soil fertility, climate, water status. These fields are represented by a letter from A to O. Approximately 10 banana plants at flowering were chosen in the diagonals of these rectangular fields. For each of these banana plants, approximately 10 chlorophyll measurements were performed on one of the two sides of the 3-rank leaf that was previously cleaned. Measurements per leaf were made following a X shape (Figure 1B) on the leaf side of interest. The main objective of this preliminary experiment was two-fold: (1) make sure that

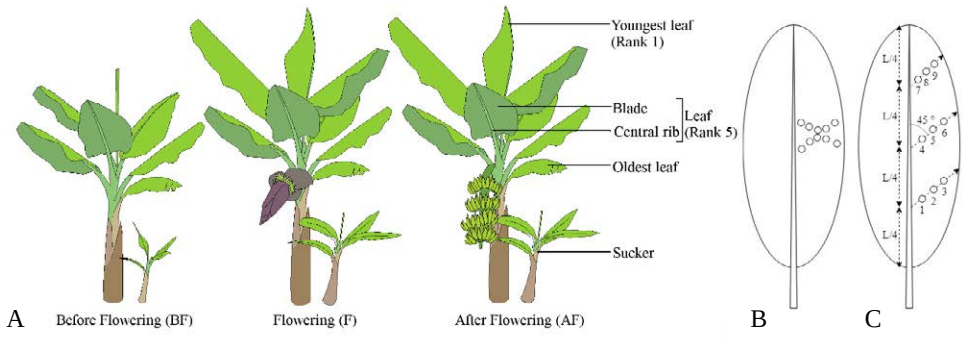


Figure 1. Banana plants at different phenological stages (A) Location of chlorophyll measurements on a leaf for the first (B) and second (C) experiments.

the growing conditions had an influence on the fields' chlorophyll content, and (2) compare the inter-field, inter-plant, and intra-leaf chlorophyll-content. This sample selection protocol is very similar to the one used for foliar analysis in banana crops (Lassoudière, 2007).

- Experiment 2: 8 out of these 15 fields (A, B, C, E, F, G, H, M) were used to evaluate the sources of chlorophyll-content variability due to banana plant asynchronism and architecture. Table 1 sums up the number of leaves that were studied for each leaf rank and phenological stage of interest. Five different phenological stages were considered: before flowering (BF), at flowering (F), and after flowering (AF2, AF6, AF10). For the last three stages (AF2, AF6, AF10), the number corresponds to the age of the bunches in weeks since flowering. More measurements were performed on the 3-rank leaf and for flowering banana plants because those are the rank and stage that are used by farmers to make leaf nutritional diagnoses (Table 1). In this experiment, uncommon banana plants or those for which some leaves were missing were avoided. For practical reasons and to reduce the within-plant variability, banana plants were selected within a restricted perimeter in the fields.

For all the leaves of this experiment 2, 9 measurements were made at different locations on each side of the leaves (18 measurements per leaf) following two axes x and y (Figure 1C). The y-axis corresponds to the main rib of the leaf. The x-axis is a fictive axis with a 45° angle with respect to the y-axis. The main objective of these measurements was to quantify the leaf chlorophyll-content heterogeneity. Working with a 45° angle was interesting to make sure that measurements were made in different secondary ribs to better estimate the within-leaf chlorophyll-content variability. One of the two sides of each leaf was cleaned to remove dust and phytosanitary treatment residues. This was done because deposits are likely to be more important on aged leaves, i.e. with advanced ranks and phenological stages. To avoid any confusion with a possible leaf lateralization effect (the left side of the leaf always unfolds first), the side of the leaf that was cleaned was not always the same. In total, four sources of chlorophyll-content variability were investigated, i.e. the variabilities due to: (1) the banana plants' growing conditions (that are assumed to be different amongst the investigated fields); (2) the banana plants' phenological stages; (3) the banana plants' architecture (location of the measurements on the leaf, side and rank of the leaf); and (4) the leaf cleanliness.

Data analysis

The influence of the sources of chlorophyll-content variability were characterized using a mixed model with nested factors. These models are an extension of the fixed models to which random effects are added to the models' residuals (Equation 1). One of the main limits of fixed models is that they assume that the response variable is independent amongst the model's individuals, e.g. between the chlorophyll-content of different leaves on the same banana plant in this case study. Often, this assumption of independence cannot be accepted, especially when measurements are made at different levels, on the same individual. The main interest of adding random effects, and therefore to work with mixed models, is to not only simply represent a population effect but also to

Table 1. Number of samples by phenological stage and leaf rank in experiment 2.

		Phenological stage				
		BF	F	AF2	AF6	AF10
Leaf Rank	1	9	23	9	9	9
	3	23	24	14	14	14
	5	9	23	9	10	4
	8	9	9	8	6	2

quantify the variability that is due to each level of variability. Modelling these sources of variability also enables to model the relationship between observations made at the same hierarchical level and to dispose of unbiased statistical tests on fixed effects.
For the first experiment, the inter-field, inter-plant, and intra-leaf variabilities were evaluated with a mixed linear model (Equation 1). The 'fields' effect was modelled by a fixed effect as fields were selected by the farmers. The 'banana plants' effect was modelled by a random effect nested to field as they were randomly selected in each field:

$$Chl_{b,i} = \mu + \mu_{field_b} + a_b + \varepsilon_{b,i} \quad (1)$$

Where $Chl_{b,i}$ is the i^{th} chlorophyll observation that was made on the banana plant b, μ is the population's reference, μ_{field_b} is the constant of the field in which b can be found, a_b is the random effect of the b that is the difference between the mean of observations on b and the mean of the observations in the field where b is located, $\varepsilon_{b,i}$ are the model's residuals.
It is assumed that the variables a_b are independent and identically distributed and that $a_b \sim N(0,\sigma_a^2)$. Similarly, $\varepsilon_{b,i} \sim N(0,\sigma^2)$ and it is assumed that a_b and $\varepsilon_{b,i}$ are independent.
For the second experiment, the effects of the banana phenological stage, leaf rank, within-leaf measurement location, leaf cleanliness and leaf side were modelled by fixed effects. The effects of the banana plant nested within the field and the leaf nested within the banana plant are represented by nested random effects.

$$\begin{aligned} Chl_{b,l,i} = {} & \mu + \mu_{field_b} + \alpha_{stage_b} + \beta_1 * rank_l + \gamma_1 * stage_b * rank_l + \beta_2 * rank_l^2 + \gamma_2 * stage_b * rank_l^2 \\ & + \mu_{side_l} + \mu_{clean_l} + \delta_1 * e^{x_i} + \delta_2 * e^{y_i} + \delta_3 * e^{x_i} * e^{y_i} \\ & + a_b + m_b * rank_l + c_l \\ & + \varepsilon_{b,l,i} \end{aligned} \quad (2)$$

Fixed effects
Random effects
Residuals

Where $Chl_{b,l,i}$ is the i^{th} chlorophyll measurement made on leaf l on banana plant b.

- *Fixed effects*: μ is the population's reference; μ_{field_b} is the constant of the field in which b can be found; α is the effect of the phenological stage of b; β_1 and β_2 represent the linear and quadratic coefficients between the chlorophyll-content and leaf rank; γ_1 and γ_2 represent the interaction terms between the phenological stage and leaf rank; μ_{clean_i} and μ_{side_i} are constants related to the cleanliness and side of leaf l; δ_1, δ_2 and δ_3 are the coefficients of the relationship between the chlorophyll and the location of the observations on the x and y axes of leaf l.
- *Random effects*: a_b and m_b represent the random intercept and slope associated with b, c_l is the random effect of l. a_b and m_b model the fact that the relationship between the leaf rank and leaf chlorophyll-content varies across banana plants around the mean relationship represented by the fixed effects. c_l models the fact that the mean chlorophyll content of l nested within b can deviate from the mean linear relationship between the chlorophyll content and the rank of the leaves of b.
- *Model's residuals*: $\varepsilon_{b,l,i}$ are the residuals of the model. It is assumed that $\begin{bmatrix} a_b \\ m_b \end{bmatrix} \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_m^2 \end{bmatrix})$, $c_l \sim N(0, \sigma_l^2)$, $\varepsilon_{bli} \sim N(0, \sigma^2)$, and that $\begin{bmatrix} a_b \\ m_b \end{bmatrix}$, c_l and ε_{bli} are mutually independent.

Note that the relationships between the chlorophyll-content and the leaf rank, observations locations on axes x/y and phenological stages are not linear (Equation 2). This choice was made because the assumptions of linearity were not met when looking at the structure of residuals. The effects of the fixed factors were tested with a top-down approach using the Bayesian information criterion (BIC). Once the structure of the fixed effects was determined, the estimation parameters was done using the restricted maximum likelihood (REML). Model were fitted with the *nlme* package (Pinheiro *et al.,* 2017) in R (R Core Team, 2018).

Results and discussion

Analysis of model 1

Figure 2 presents the chlorophyll-content variability of the 3-rank leaf for flowering banana plants on the 15 fields under study. Boxplots highlight variations in the leaf chlorophyll-content at multiple scales, i.e. between fields (different colors), within fields (same color), and in the 3-rank leaf of each banana plant (same boxplot). In model 1, the fixed effect 'field' is significant meaning that the growing conditions set out by farmers have an influence on the leaf chlorophyll-content. The variance estimations σ_a^2 and σ^2 are 18.41 $\mu g^2/cm^4$ and 12.43 $\mu g^2/cm^4$, respectively. The σ_a^2 estimation shows a banana plant effect and also corresponds to the covariance between observations made on the same banana plant and justifies the use of a mixed model. Finally, σ^2 shows the residual variability. It can be due to the leaf heterogeneity (leaf thickness, rib proportion) that can affect the Dualex measurements, to the chlorophyll gradients within the leaf, and to the acquisition condition (operator, temperature). It must be noted that the variability related to the repeatability of the measurements is weak if one refers to the standard deviation obtained by Cerovic *et al.* (2012) i.e. 0.13 $\mu g/cm^2$. These results demonstrate the impact of the banana growing conditions on the leaf chlorophyll-content in a commercial plantation and the possible interest of non-destructive measurements to better assess these sources of variability. However, this experiment was done with plants at the same phenological stage. The objective of the following experiment is to better characterize the sources of variability related to the banana architecture and phenological stages within a field.

Analysis of model 2

The analysis of the fixed effects of model 2 (Equation 3) is presented in Table 2 and Figure 3. Apart from μ_{clean_i}, the remaining fixed factors of model 2 had a significant influence on the leaf chlorophyll-content. This result shows that in banana plants, cleaning of the leaves is not mandatory in order to obtain reliable measurements. The importance of the 'field' effect that was already highlighted in the first experiment was here again confirmed (Table 2). However, the classification of fields according to their average chlorophyll content is different between Experiment 1 and Experiment 2 (μ_{field} in Table 2 and mean levels in Figure 2). The sampling design for Experiment 1 was built to ensure that the measurements are representative of the mean chlorophyll level of the fields. This was not the case for experiment 2, whose objective was to see the effect of phenological stages on banana plants located in the same environment. These differences in the sample design may explain the different classification between the two experiments.

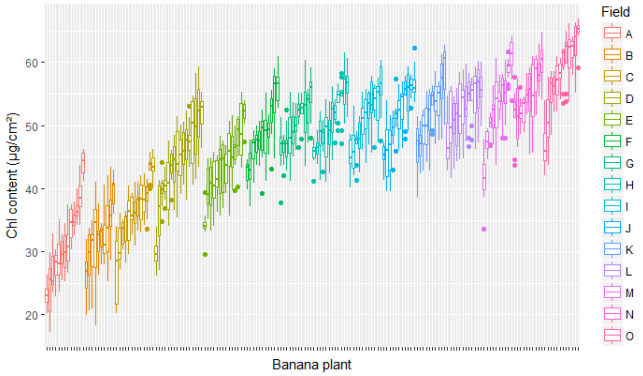


Figure 2. Chlorophyll-content variability in the 3-rank leaf for flowering banana plants in the 15 fields under study.

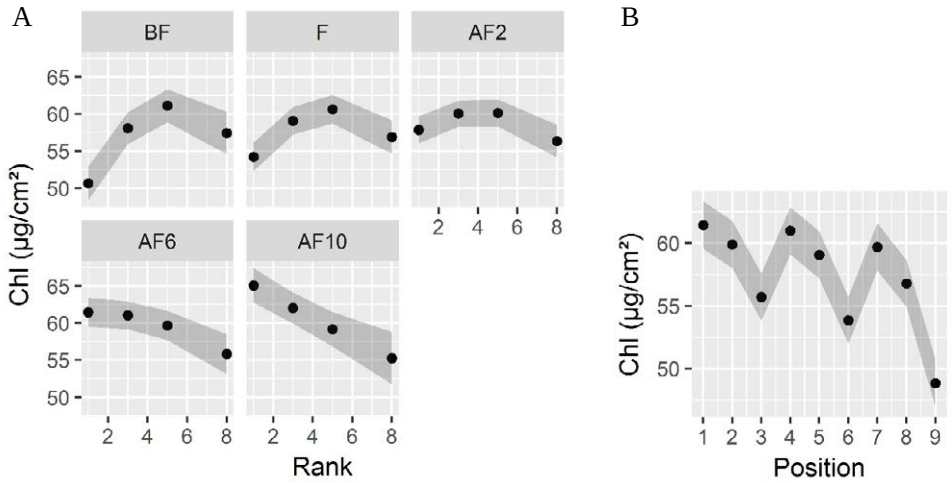


Figure 3. (A) Mean effect of the leaf rank and phenological stage on the chlorophyll content. (B) Mean effect of the within-leaf observation's location on the chlorophyll content. The hull represents the confidence interval at 95%.

Whatever the banana phenological stage, the chlorophyll-content varied with the leaf rank (Figure 3). The quadratic relationship between the chlorophyll-content and leaf rank was found to be dependent on the phenological stage. When looking at the 1-rank leaf, chlorophyll-contents were weak for stages BF, F and AF2, whereas they were much stronger for stages AF6 and AF10. The chlorophyll-content also varied spatially within a leaf. The chlorophyll-content decreased as measurements moved away from the central rib (Figure 3B). The left side of the leaf had a systematically higher chlorophyll-content than the right side. The chlorophyll-content also decreased along the y-axis, but more slightly. It must be noted that the mean difference between measurements made at positions 4 and 9 was 12 µg/cm^2. This is in the same range as the difference between the chlorophyll-contents of the 1-rank and 3-rank leaves at the BF stage (11 µg/cm^2). The effect of the phenological stage (α in Table 2) is not interpreted here. This is because banana plants at different phenological stages

Table 2. Parameter estimates of Model 2 through REML.

Parameter	Estimation	Standard error	Parameter	Estimation	Standard error
μ	29.14	1.85	γ_1	-1.88	0.22
μ_{fieldB}	-2.70	1.77	γ_2	0.14	0.02
μ_{fieldC}	4.00	1.90	$\mu_{side\ left}$	1.33	0.12
μ_{fieldE}	7.05	1.77	δ_1	-0.28	0.014
μ_{fieldF}	7.91	1.60	δ_2	-0.05	0.014
μ_{fieldG}	7.00	1.89	δ_3	-0.017	0.0011
μ_{fieldH}	4.89	1.77	σa^2	9.47	3.07
μ_{fieldM}	12.55	1.57	σm^2	0.20	0.44
α	5.37	0.48	σl^2	4.58	2.14
β_1	7.78	0.64	σ^2	15.18	3.90
β_2	-0.69	0.07			

appeared at different dates and certainly experienced different growing conditions. The experiment did not allow discrimination between the effect of the past growing conditions and the effect of the banana plant phenological stage. The variance associated with the random effects in Model 2 was not compared to the one in Model 1 because sampling was different in both experiments (within-leaf measurements' positions, choice of the banana plants within the fields).

Conclusions

This study highlighted that the growing conditions of the fields, the asychronism in the banana development cycles and the banana plant architecture had an effect on the chlorophyll content of the leaves. When measuring leaf-chlorophyll content on banana crops with a handheld chlorophyll meter, the phenological stage, rank and position can be known and controlled during the sampling. Using remote sensing, the interpretation of the maps for diagnoses purposes should be done with greater caution than for those obtained with ground-based measurements because of all the aforementioned sources of chlorophyll-content variability. Further research is needed before considering nutritional diagnoses services for banana crops based on remote sensing.

Acknowledgements

The experiments were made in Plantation du Haut Penja with the support of its agronomy department. We are grateful to all the plantation staff and particularly E. Andjengo, P. Dontsop, P. Esnault, R. Nkongo, A. Pugeaux, J. Simo and R. Solano for their help with the experiments.

References

Cerovic, Z. G., Masdoumier, G., Ghozlen, N. B., & Latouche, G., 2012. A new optical leaf-clip meter for simultaneous non-destructive assessment of leaf chlorophyll and epidermal flavonoids. Physiologia plantarum, 146, 251-260.

Coquil, B., Bordes, J. P., 2005. FARMSTAR: an efficient decision support tool for near real time crop management from satellite images. In J V Stafford (Ed.) Proceedings of the 5th European Conference on Precision Agriculture (pp. 873-880). Wageningen, the Netherlands: Wageningen Academic Publishers.

Debaeke, P., Rouet, P., & Justes, E., 2006. Relationship between the normalized SPAD index and the nitrogen nutrition index: application to durum wheat. Journal of plant nutrition, 29, 75-92.

Gitelson, A. A., Vina, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J., 2005. Remote estimation of canopy chlorophyll content in crops. Geophysical Research Letters, 32, Art. No. L08403

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., & Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote sensing of environment, 81, 416-426.

Lassoudière, A., 2007. Le bananier et sa culture. (The banana plant and its farming) Versailles, France: Editions Quae.

Lemaire, G. (Ed.)., 1997. Diagnosis of the nitrogen status in crops. Berlin, Germany: Springer-Verlag.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., *et al.,* 2017. Package 'nlme'. Linear and Nonlinear Mixed Effects Models, version, 3-1.

Prost, L., & Jeuffroy, M. H., 2007. Replacing the nitrogen nutrition index by the chlorophyll meter to assess wheat N status. Agronomy for Sustainable Development, 27(4), 321-330.

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org.

Tixier, P., Malezieux, E., & Dorel, M., 2004. SIMBA-POP: a cohort population model for long-term simulation of banana crop harvest. Ecological Modelling, 180(2-3), 407-417.

Sentinel-2 vegetation indices and apparent electrical conductivity to predict barley (*Hordeum vulgare* L.) yield

J.A. Martínez-Casasnovas*, A. Uribeetxebarría, A. Escolà and J. Arnó
Research Group in AgroICT & Precision Agriculture (GRAP), University of Lleida, Agrotecnio Center, Lleida, Catalonia, Spain; j.martinez@macs.udl.cat

Abstract

The launch of the satellites of the Sentinel-2 mission has opened a new era of possibilities in precision agriculture. The objective of the present work was to explore the use of Sentinel-2 visible, red-edge and near infrared bands to predict barley (*Hordeum vulgare* L.) yield in early spring phenological stages at field level. Spectral vegetation indices for different dates were analysed together with apparent electrical conductivity (ECa) to obtain the best linear model to predict yield in critical moments of the barley cycle, in which it is still possible to apply variable rates of sidedress fertilizer to improve yield. The research was carried out in 2017 in a farm located in NE Spain. ECa surveys were by means of a Veris 3100 ECa surveyor and yield data were acquired by means of a yield monitor. Images acquired the first week of April (barley Zadoks DC 30-32) were the ones presenting the best correlation with yield. Multiple regression models resulted in R^2 values of 0.75, confirming that vegetation indices from the first week of April produced lower prediction error than earlier dates.

Keywords: Sentinel-2, yield prediction, remote sensing, vegetation indices, apparent electrical conductivity

Introduction

Crop yield prediction is an important issue in agriculture at different scales. At farm level, farmers try to know how much yield they will get from their fields. At local, national or supra-national food systems, the benefits of yield prediction include improved inventory, revenue, cost and risk management along the entire value chain (Wall *et al.*, 2008; Lambert *et al.*, 2017). Two main modelling approaches are available for crop yield prediction: process-based models and statistical models (Lobell and Burke, 2010). The first usually require extensive input data about crop, soil, climate and management. The second are simpler alternatives, since they link the variable of interest (yield) to predictors known for the current and/or previous seasons. In this respect, remote sensing has been recognised as a source of data for characterizing spatial variability of crop grain yield early in the growing season, mainly using regression analysis techniques and multitemporal data sets (Wall *et al.*, 2008; Panda *et al.*, 2010; Bolton and Friedl, 2013).
From the early work of Tucker *et al.* (1980), who identified a relationship between NDVI and crop yield in experimental fields, a wide range of vegetation indices have been used in crop monitoring and grain yield prediction. Most of the research works using remote sensing imagery have been done with vegetation indices derived from the NOAA's Advanced Very High Resolution Radiometer (AVHRR) (Ferencz *et al.*, 2004, Wall *et al.,* 2008, and Labus *et al.,* 2010), or with the NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) (Bolton and Friedl, 2013, Qader *et al.*, 2018). More recently, the early prediction of crop yield has gained particular relevance (Panda *et al.*, 2010). Nevertheless, at the farm scale, where precision agriculture applications require high resolution data, the low resolution of MODIS (250 m bands 1-2) or AVHRR-NDVI (1 km) are not useful (Labus *et al.*, 2010). Other traditional satellite systems such as Landsat, with 30-m resolution, still may have too coarse resolution to fully understand the spatial dynamics of crop status (Scudiero *et al.*, 2016; Kumhálová and Matějková, 2017). Other researchers have used higher resolution satellite images

DOI 10.3920/978-90-8686-888-9_38, © Wageningen Academic Publishers 2019

(QuickBird-2, WorldView-2, SPOT-7), together with Landsat images to predict winter barley yield, with NDVI as predictor, or in combination with a handheld GreenSeeker crop sensor to estimate winter barley and winter wheat yield. Kumhálová and Matějková (2017), obtaining good results for images acquired in later phenological stages (e.g. BBCH 59 End of heading: inflorescence fully emerged) with R between 0.784 and 0.856.

The launch in 2015 and 2017 of the two satellites of the Sentinel-2 mission (an Earth observation mission developed by the European Space Agency as part of the Copernicus Programme of the European Commission, Drusch *et al.*, 2012), has opened a new era of possibilities in agriculture that overcomes some limitations of previous public and/or private missions. In particular, the 10 and 20 m resolution bands in the visible and infrared spectrum offer the possibility of calculating a wide range of spectral indices. Such capacities make Sentinel-2 satellites very appealing for crop monitoring in the framework of map-based precision agriculture (Escolà *et al.*, 2017). Many potential applications of Sentinel-2 imagery were published before the first satellite of the mission was launched but few real applications have been described, and very few use it to predict grain yield at farm level. An example is the work by Lambert *et al.* (2017), who tried to estimate yield at parcel level from Sentinel-2 time series in sub-Saharan smallholder farming systems (cotton, maize, millet, sorghum and peanut). The technological paradigm of precision agriculture also offers the possibility to consider information about the spatial variability of soils, usually ignored in yield prediction due to the difficulties of acquiring detailed soil data (Stadler *et al.*, 2015). One approach to know the spatial variability of soil integrated properties at field scale is the non-invasive measurement of the apparent electrical conductivity (ECa) (Uribeetxebarria *et al.*, 2018). Stadler *et al.* (2015) made a review of works in which ECa was correlated with yield at harvest in different crops. All of these studies only investigated the relationship between ECa and yield. Then, possible relationships between ECa and spectral vegetation indices with yield have not been sufficiently investigated yet. In this respect, the integration of multi-sensor data (ECa and spectral vegetation indices) might be most useful in yield prediction at field scale.

According to this background, the objective of the present work was to explore the possibilities offered by Sentinel-2 bands in the visible, red-edge and near infrared spectrum to predict barley (*Hordeum* vulgare L.) yield in early spring phenological stages at field level. Spectral vegetation indices for different dates and ECa were analysed to obtain the best linear models to predict yield before sidedress fertilizer application.

Material and methods

Study area

The study area comprised two plots of a farm: Plot 1 (17.2 ha) and Plot 2 (65.0 ha) located in NE Spain (Lat 41.740837°, Long 0.358143° WGS84, Figure 1). The plots were sown with barley (*Hordeum vulgare* L.) on December 10th and 12th, 2016, and were harvested between June 19th and 24th, 2017. The average yield per plot was 5,915 kg/ha (Plot 1) and 6,758 kg/ha (Plot 2). The variety cultivated was Gustav (Svalöf Weibull AB, Sweden), which is a two-row spring barley. The average slopes were 4.8% (Plot 1) and 2.4% (Plot 2). The plots were irrigated by means of sprinkler irrigation. Soils of the area are classified as Typic Xerorthents and Typic Xerofluvents (Soil Survey Staff, 2014).

ECa surveys

Two ECa surveys were conducted in June 23rd, 2016 (Plot 1) and December 20th, 2017 (Plot 2). They were carried out with a Veris 3100 sensor (Veris Technologies Inc. Salina, Kansas, USA), measuring ECa at 0-30 cm (shallow ECa) and 0-90 cm (deep ECa) soil depths simultaneously. Values above or below ±2.5 standard deviations (SD) were considered outliers and were removed from the original data file according to the criteria of Taylor *et al.* (2007). The final ECa data sets consisted of 3,270 points (Plot 1) and 12,198 points (Plot 2) with shallow and deep readings. These data were

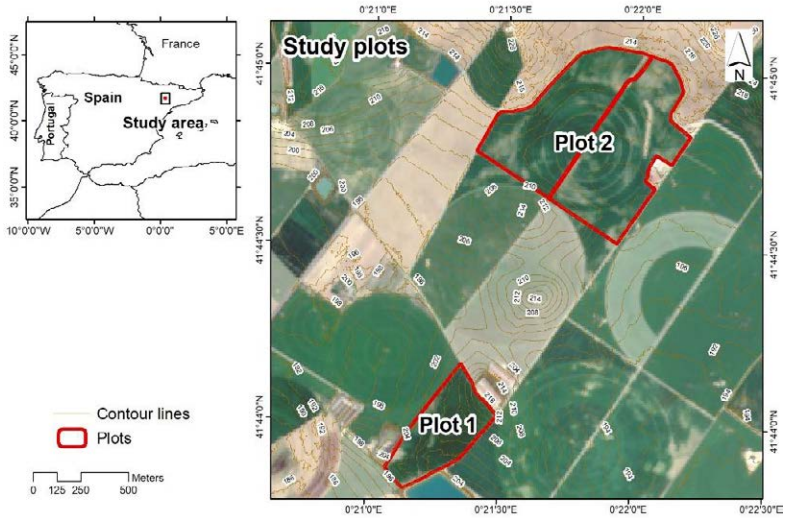


Figure 1. Location of the case study area.

interpolated on a 2-m grid by means of ordinary kriging using the exponential semivariogram model. For this purpose, ArcGIS Geostatistical Analyst 10.5 (ESRI, Redlands, California, USA) was used.

Sentinel-2 imagery: data acquisition and processing

Different Sentinel-2A images corresponding to barley's half tillering (Decimal Code 25-27 of the scale of Zadoks, Zadoks *et al.*, 1974) and early stem elongation (DC 30-32) were considered in the present study: 17/03/2017 (DC 23-25) and 06/04/2017 (DC 30-32). The selected crop stages mainly correspond to the slow expansion phase, in which barley's green area expands slowly and requires low nitrogen (N). After GS30 the barley starts the rapid expansion phase, with higher N requirements. Then, yield prediction in the studied stages (DC 25 to 30) can be useful to recommend N side-dressing applications.

In the present research, twenty-two vegetation indices (listed in https://www.indexdatabase.de for Sentinel-2) were selected according to their possible calculation from Sentinel-2 bands, using different combinations of the visible, red edge and near infrared bands (see indices below). For that, the Level-1C (Top-of-Atmosphere, ToA reflectance) images were pre-processed with the open source ESA Sentinel Application Platform (SNAP) and the Sen2Cor plugin (http://step.esa.int/main/third-party-plugins-2/sen2cor), in order to perform the physical atmospheric, terrain and cirrus correction of ToA Level-1C to Bottom-of-Atmosphere (BoA) Level-2A.

The vegetation indices initially considered were (see formulas and references in https://www.indexdatabase.de/): Chlorophyll Index Green (CIgreen), Chlorophyll Vegetation Index (CVI), Enhanced Vegetation Index (EVI), Enhanced Vegetation Index 2 (EVI2), Green Normalized Difference Vegetation Index (GNDVI), Green-Red Vegetation Index (GRVI), Normalised Difference Vegetation Index (NDVI), Normalized Difference Vegetation Index – Green, Modified Triangular Vegetation Index 1 (MTVI1), Modified Triangular Vegetation Index 2 (MTVI2), Optimized soil-adjusted vegetation index (OSAVI), Soil Adjusted Vegetation Index (SAVI), Wide Dynamic Range Vegetation Index (WDRVI), Chlorophyll Index Red Edge (CIred-edge), Inverted Red-Edge Chlorophyll Index (IRECI), Normalized Difference Index bands 4 and 5 (NDI45), Normalized Difference Red Edge Index 1 (NDRE1), Normalized Difference Red Edge Index 2 (NDRE2), Modified

Chlorophyll Absorption in Reflectance Index (MCARI), Red Edge NDVI (RENDVI), Transformed Chlorophyll Absorption Reflectance Index (TCARI) and Triangular Vegetation Index (TVI).

Yield data

Yield data were acquired by means of a Trimble Ag FMX yield monitor (Trimble Inc., Sunnyvale, California, USA) installed in a Fendt 5275 C combine. The monitor was calibrated with the real grain weight of the first load for each day of the harvest (from 19th to 24th June, 2017). Yield monitor data consisted of two sets of points (29,024 points for Plot 1 and 103,454 points for Plot 2). Similarly, as done with ECa data, values above or below ±2.5 standard deviations (SD) were considered outliers and were removed from the original data file. Then, raw yield data were converted to the reference grain moisture content of 13%. After this process, yield data consisted of 27,606 points (Plot 1) and 97,716 points (Plot 2), which were interpolated to a continuous grid. For that, ordinary kriging in 20 m blocks projecting interpolated data over a regular 5 m grid was done using VESPER (Minasny *et al.*, 2005).

Statistical analysis

To model yield using the ECa and the different vegetation indices as possible predictors, a linear model including several variables for each date and field was chosen so that the AIC (Akaike Information Criterion) was minimized (criterion-based procedure). To do this, a total of 28 points (locations) in Plot 1 and 102 points in Plot 2 randomly distributed according to 5 quantile yield classes were used as training samples to build and estimate the models. Then, testing sample of the same size but different observations was used to validate the models, being the root mean square error (RMSE) the adopted statistic to check how well does the model do in predicting the observations (actual yield) in the test sample. In all cases, model selection, detection of multi-collinearity (through the variance inflation factors, VIFs), and goodness of fit (R^2 and RMSE) were made by programming in R software (RStudio Version 1.1.456). The analysis were done on the basis of the coarser resolution of the input data (20×20 m, Sentinel-2 red-edge bands). For that, the input variables with finer resolution were aggregated to 20×20 m areas.

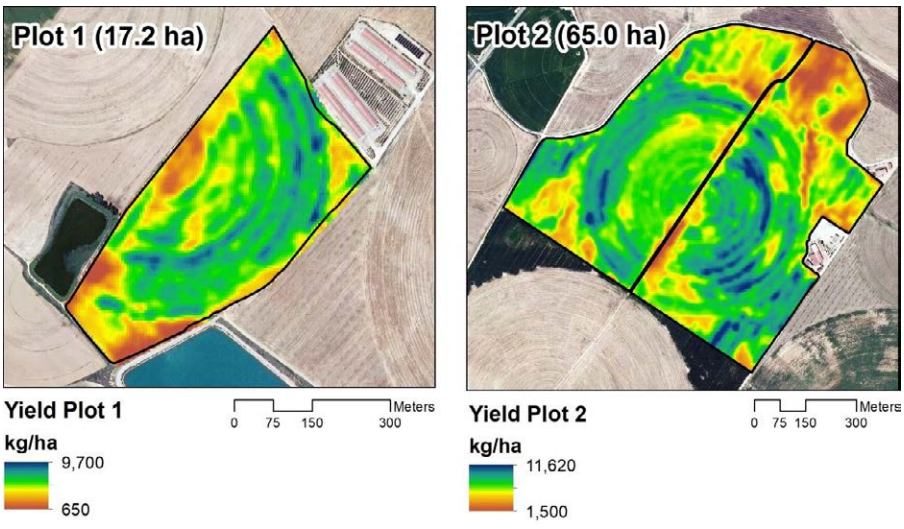


Figure 2. Barley yield maps of the study plots.

Results and discussion

Yield maps

Figure 2 shows the yield maps of the two barley study plots. The respective average yield values were 5,955±1,611 kg/ha of barley in Plot 1 and 6,832±1,804 kg/ha in Plot 2. The inclusion of the plot as a factor of two levels (Plot 1 and Plot 2) in a unique linear prediction model showed (data not included) the convenience of obtaining a separate model for each plot given the significance of this categorical variable.

Multiple linear regression prediction models of yield vs ECa and vegetation indices

Multiple linear regression models were obtained for each plot and date in order to achieve the prediction of yield in early spring barley stages. Table 1 shows the prediction models for Plot 1 and Table 2 for Plot 2. As mentioned above (methods), the initial set of vegetation indices were reduced using the Akaike Information Criterion (AIC). The selected models presented an average value of variance inflation factors (VIFs) less than 10, which was indicative of low correlation between the explanatory variables and, therefore, of more accurate estimates of regression coefficients. The results indicated that the models for both plots were different, with distinct explanatory variables. In addition, those variables were not the same in the two dates. For Plot 1, yield prediction models obtained for the second date (06/04/2017) produced less RMSE validation values than the earlier date (17/03/2017). Moreover, these RMSE values were similar to the RMSE ones calculated with the model data set. Regarding the explanatory variables, in the two dates, the CVI and the NDVI were selected in the simplified models, being on 06/04/2017 the only explanatory variables.

The results for Plot 2 are presented in Table 2. In this case, only one variable, ECa Deep, was selected in both types of models (complex and simplified) and for the two dates. This variable presented a negative coefficient, indicating an inverse relationship with yield. In this case, the prediction models of the first date were the ones explaining higher amount of yield variance (R^2 0.75) and with lower RMSE values (978 and 991 kg/ha). Nevertheless, the simplified model of 06/04/2017 had also similar results with only two explanatory variables (ECa Deep and NDRE2). In all the cases, the RMSE values for the model set and the validation set for Plot 2 presented lower differences than for the models of Plot 1.

The fact that the yield prediction models incorporated different explanatory variables in both plots could be due to the significantly different levels of soil apparent electrical conductivity (ECa), which

Table 1. Complex and simplified linear models for yield prediction for Plot 1.[1]

Date	Equation	Model		Validation[2]
		R^2	RMSE kg/ha	RMSE kg/ha
17/03/2017	-10596.89 – 6.87 ECa_Deep + 2639 CVI + 106564.72 NDVI_ Green + 11418.29 NDVI – 58947.15 RE_NDVI_03 – 9.27 TCARI	0.70	1,052	1,562
	-6115.48 + 1679.08 CVI + 14780.71 NDVI – 47765.74 RE_NDVI	0.67	1,108	1,497
06/04/2017	1990.95 – 8.58 ECa Deep – 309.50 CVI – 9811.42 NDVI Green + 2326.70 NDRE1 – 953.64 RE_NDVI – 810.73 TCARI – 66.12 TVI + 4996.66 NDVI	0.69	1,072	965
	1178.49 – 540.05 CVI + 5979.35 NDVI	0.69	1,123	1,031

[1] VIF: variance inflation factor <10. N Plot 1=28, N Plot 2=102.

[2] Validation points different than the model points.

Table 2. Complex and simplified linear models for yield prediction for Plot 2.[1]

Date	Equation	Model		Validation[2]
		R^2	RMSE kg/ha	RMSE kg/ha
17/03/2017	-3,220.59 – 12.80 ECa Deep + 810.61 CVI + 2,209.89 NDVI + 66,714.86 NDVI Green + 371.92 CI_RE – 3,726.92 NDI45 + 5,466.08 NDRE2 + 931.569 RE_NDVI + 2,466.89 TCARI	0.75	978	1,091
	-4,297.94 – 14.06 ECa Deep + 1,126.45 CVI + 84,219.78 NDVI Green + 6808.62 NDRE2	0.75	991	1,098
06/04/2017	375.48 – 14.90 ECa Deep – 79.40 CVI – 3,179.05 NDI45 + 8,109.97 NDRE2 + 4,044.71 RE_NDVI – 828.72 TCARI + 132.78 TVI + 782.22 NDVI	0.69	1,072	957
	1045.56 – 16.09 ECa Deep + 10,033.50 NDRE2	0.75	994	1,122

[1] VIF: variance inflation factor <10. N Plot 1=28, N Plot 2=102.
[2] Validation points different than the model points.

in the present study area are related to soil salinity (average ECa in Plot 1= 34.2±24.8 mS/m versus average ECa in Plot 2= 59.7±24.1 mS/m). Nevertheless, in Plot 2, the main soil salinity problems were concentrated in one specific zone of the field (northeaster part, Figure 2, with average yield between 1,500-2,000 kg/ha), being the rest of the zones of higher productive potential than Plot 1, which explains the higher average yield obtained in Plot 2 with respect to Plot 1. The fact that the ECa surveys were carried out in different dates in the two plots, could affect to the model coefficients, but not the importance of this variable in the prediction of grain yield in these soils with salinity problems.

Conclusions

The use of vegetation indices derived from Sentinel-2 images in early spring stages, and in combination of apparent soil electrical conductivity (ECa), allowed 75% of barley yield variance to be explained (predicted) with similar model and validation errors. The explanatory variables, however, were different in the two tested plots, although consistent between the two analysed dates in each respective plot. A reason could be the different soil properties reflected in ECa measures. Then, the consideration of ECa as possible explanatory variable of yield was relevant when combined with the vegetation indices derived from early spring stages to predict barley yield.

Acknowledgements

The present work was done within the framework of Contract C16022 between the University of Lleida and Ventafarinas, S.L. Funding from Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya under Grant 2017-SGR-646 is also acknowledged.

References

Bolton, D.K., Friedl, M.A., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agricultural and Forestry Meteorology 173, 74-84.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., *et al.,* 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. Remote Sensing of Environment 120, 25-36.

Escolà, A., Badia, N., Arnó, J., Martínez-Casasnovas, J.A., 2017. Using Sentinel-2 images to implement precision agriculture techniques in large arable fields: First results of a case study. In JA Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences 8(2), 377-382.

Ferencz, Cs., Bognár, P., Lichtenberger, J., Hamar, D., Tarcsai, Gy., Timár, G., *et al.,* 2004. Crop yield estimation by satellite remote sensing. International Journal of Remote Sensing 25(20), 4113-4149.

Kumhálová, J., Matějková, Š., 2017. Yield variability prediction by remote sensing sensors with different spatial resolution. International Agrophysics 31(2), 195-202.

Labus, M.P., Nielsen. G.A., Lawrence, R.L., Engel, R., Long, D.S., 2010. Wheat yield estimates using multi-temporal NDVI satellite imagery. International Journal of Remote Sensing 23(20), 4169-4180.

Lambert, M.J., Blaes, X., Traoré, P.S., Defourny, P., 2017. Estimate yield at parcel level from S2 time series in sub-Saharan smallholder farming systems. In Proceedings of the 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), IEEE, Brugge, Belgium, pp. 1-7.

Lobell, D.B., Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to climate change. Agricultural and Forest Meteorology 11, 1443-1452.

Minasny, B., McBratney, A.B., Whelan, B.M., 2005. VESPER version 1.62. Australian Centre for Precision Agriculture, University of Sydney, NSW 2006.

Panda, S.S., Ames, D.P., Panigrahi, S., 2010. Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques. Remote Sensing 2(3), 673-696.

Qader, S.H., Dash, J., Atkinson, P.M., 2018. Forecasting wheat and barley crop production in arid and semi-arid regions using remotely sensed primary productivity and crop phenology: A case study in Iraq. Science of the total Environment 613-614, 250-262.

Scudiero, E., Corwin, D.L., Wienhold, B.J., Bosley, B., Shanahan, J.F., Johnson, C.K., 2016. Downscaling Landsat 7 canopy reflectance employing a multi-soil sensor platform. Precision Agriculture 17, 53-73.

Soil Survey Staff, 2014. Keys to Soil Taxonomy, 12th ed. USDA-Natural Resources Conservation Service, Washington, DC, USA.

Stadler, A., Rudolph, S., Kupisch, M., Langensiepen, M., van der Kruk, J., Ewert, F., 2015. Quantifying the effects of soil variability on crop growth using apparent soil electrical conductivity measurements. European Journal of Agronomy 64, 8-20.

Taylor, J.A.; McBratney, A.B.; Whelan, B.M., 2007. Establishing management classes for broadacre agricultural production. Agronomy Journal 99, 1366-1376.

Tucker, C.J., Holben, B.N., Elgin, J.H., McMurtrey, J.E., 1980. Relationship of spectral data to grain yield variation. Photogrammetric Engineering and Remote Sensing 46, 657-666.

Uribeetxebarria, A., Daniele, E., Escolà, A., Arnó, J., Martínez-Casasnovas, J.A., 2018. Spatial variability in orchards after land transformation: Consequences for precision agriculture practices. Science of the Total Environment 635, 343-352.

Wall, L., Larocque, D., Léger, P.M., 2008. The early explanatory power of NDVI in crop yield modelling. International Journal of Remote Sensing 29(8), 2211-2225.

Zadoks, J.C., Chang, T.T., Konzak, C.F., 1974. A decimal code for the growth stages of cereals. Weed Research 14, 415-421.

Accessing the plant architecture in 3D for plant phenotyping – recent approaches and requirements

S. Paulus
Institut für Zuckerrübenforschung, Holtenser Landstr. 77, 37079 Göttingen, Germany; paulus@ifz-goettingen.de

Abstract

Three-dimensional measuring devices are able to provide highly accurate surface descriptions of crop plants in different scenarios. During recent decades, 3D devices have been used for measuring plant geometry at single plant and organ level. Various techniques such as laser triangulation, structured light scanning, time of flight and structure from motion approaches are used at different scales in laboratories, in greenhouses and in the field. This work aims to give an overview about the state-of-the-art 3D parameters that were extracted from the literature together with the focused plants and their biological links. Two main techniques, laser triangulation and structure-from-motion, are introduced and described regarding the extraction of simple and more complex plant traits in 3D to give an overview about state-of-the-art geometrical plant traits that are only measurable in three dimensional plant images. As a specific example, plant traits were described and introduced for a dataset showing the development of a maize plant over time from a greenhouse experiment.

Keywords: laser triangulation, structure-from-motion, plant geometry

Introduction

Remote sensing methods use spectral or geometrical properties of plants to get information about the plant status regarding biomass, stress reaction or growth response. As spectral information mostly focuses on information like water or chlorophyll content, geometrical information can be directly linked to harvest-relevant data such as biomass or the grow stage.
Spatial measurements have been processed to enable the derivation of geometrical parameters such as height and width, plant volume and leaf area by using mathematical algorithms. By tracking these parameters over time, growth stages can be described or development under stress can be investigated (Paulus, 2014b).
Access to the geometrical appearance can be provided by laser-based systems using a laser triangulation (LT) setup or structure from motion (SfM) approaches using standard RGB cameras (Rose, 2015). Both techniques can be used at the different scenarios of interest namely laboratory, greenhouse and field. As LT emits a laser line and uses a photo-active array (CCD) to measure distances it can be assumed to be almost independent of ambient light. It provides highly accurate point measurements within a defined space. By setting up a laser line illumination and a 2D photoactive grid (CCD array)it is possible to measure 2D profiles that can be extended by moving the sensor to a complete 3D point cloud (Dupuis *et al.,* 2015).
SfM approaches use multiple RGB images from standard cameras. These single images are stitched using overlapping parts within the images. Viewpoints are calculated and the 3D geometry of the focused object is reconstructed. Unlike the LT setup, that needs costly hardware but only limited software, SfM uses standard hardware and sophisticated software algorithms to reconstruct the objects and their 3D information (Rose *et al.,* 2015). Both techniques are optical measurement methods that differ in the way they measure as LT uses active illumination while SfM uses the ambient light (passive illumination). The result from both methods is a XYZ point cloud with a further parameter describing the backscattered laser intensity (LT) or the RGB color (SfM). This

DOI 10.3920/978-90-8686-888-9_39, © Wageningen Academic Publishers 2019

point cloud is commonly used to derive plant parameters as it is a highly accurate description of the plant geometry. The accuracy for both approaches regarding their use for plant parameterization has been shown by Paulus *et al.* (2014b) and Rose (2015).
Plant phenotyping is a very diverse scientific field. Surveying methods have been combined with algorithms from bioinformatics and have been linked to biological traits to fulfill the demands on future yield production. This results in a huge variety of measurements that have to be examined for the needs of plant phenotyping and for high throughput applications.
This paper gives an overview on two most used 3D measuring methods (LT and SfM). Showing new possibilities in 3D plant parameterization is one focus of this paper. Common (width and height) and quite new parameters (convex hull and biomass distribution) derived from XYZ point clouds are introduced and shown over time for a representative case study in maize. The applicability is shown for a two week measuring period of a maize plant. In addition a short literature overview is given for 3D plant parameters for the three scenarios of laboratory, greenhouse and field.

Materials and methods

For this case study, a single maize plant was cultivated in a greenhouse and measured by a laser scanner measuring arm – laser triangulation scanner combination (Romer Infinite 2.0 (2.8m Version), Hexagon Metrology Services Ltd., London UK and Perceptron Scan Works V5, Perceptron Inc., Plymouth, MI, USA). The laser scanner provided a point to point distance of 17 µm and an accuracy of about 50 µm within a scanning window of 110×105 mm. By using a measuring arm – laserscanner combination many different points-of-view were used and the point cloud was completed with every new position. This enabled a minimum of occlusion within the resulting point cloud. Measurements were performed every second day for a period of two weeks resulting in measurements on days 1, 3, 5, 7,9, 11 and 13. The laser-scanned point cloud was manually rastered (0.3 mm point to point distance). The parameters height and width, as well as convex hull area and convex hull volume were calculated using C++ and the Point Cloud Library (Rusu and Cousins, 2011). The statistical analysis, graphics and figure assembling was conducted using the statistic language R (R Core team, 2008) and the image processing tool Gimp (The GIMP team, GIMP 2.10, www.gimp.org, 1997-2018).
To link 3D point clouds to biological information it is essential to describe geometrical properties as parameters. Figure 1 shows the derivation of the point cloud parameter height as the distance between minimum and maximum z-coordinate, as well as the parameter width which is described by the maximum distance in x- or y-axis (Paulus *et al.,* 2014b). The convex hull describes the volume of the plant in terms of the volume used in 3D (Paulus *et al.,* 2014b). In 2D the convex hull describes

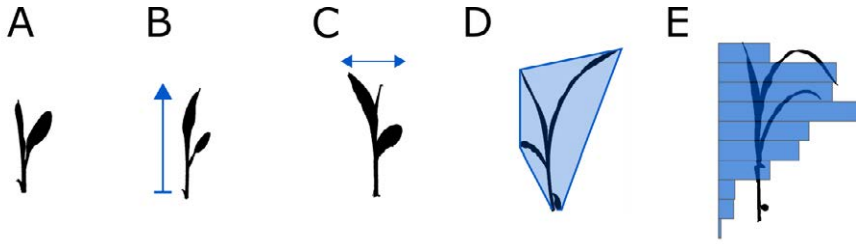


Figure 1. From a raw 3D point cloud, recorded by a LT system (A) different parameters can be derived. Height (B) and width (C) can be measured, as well as the convex hull volume (D) and the biomass distribution (E). From left to right the point clouds show the plant on the different measurement days 1, 3, 6, 9 and 13 (shown in a scaled geometry).

the inner area of a rubber band that is contracted to the maximum with the outside lying points as anchor points. Finally the biomass distribution gives a measurement of the distribution of biomass along the vertical plant axis. It shows the levels of biomass concentration. This is possible as the point cloud is a representation of the complete surface. Thus, the distribution of the surface points can be depicted as the plant biomass.

Results and discussion

When focusing on geometry describing parameters it is essential to define which parts of the point cloud and how these parts were parameterized. Based on the parameters introduced for different scans over time a growth graph can be generated. Width can be defined to be the longest distance (Euclidian) between two points perpendicular to the main axis (height) according to the main coordinate axis (x- and y-axis). Depending on the point-of-view, the width measurement regarding x- and y-axis do vary a lot. Figure 2 shows a measurement for the uncertainty when measuring that way. Figure 2A-D show different point-of-views for the maize plant, which results in different width measurements (Figure 2E). Through the growing stages a clear growth trend is visible. The grey

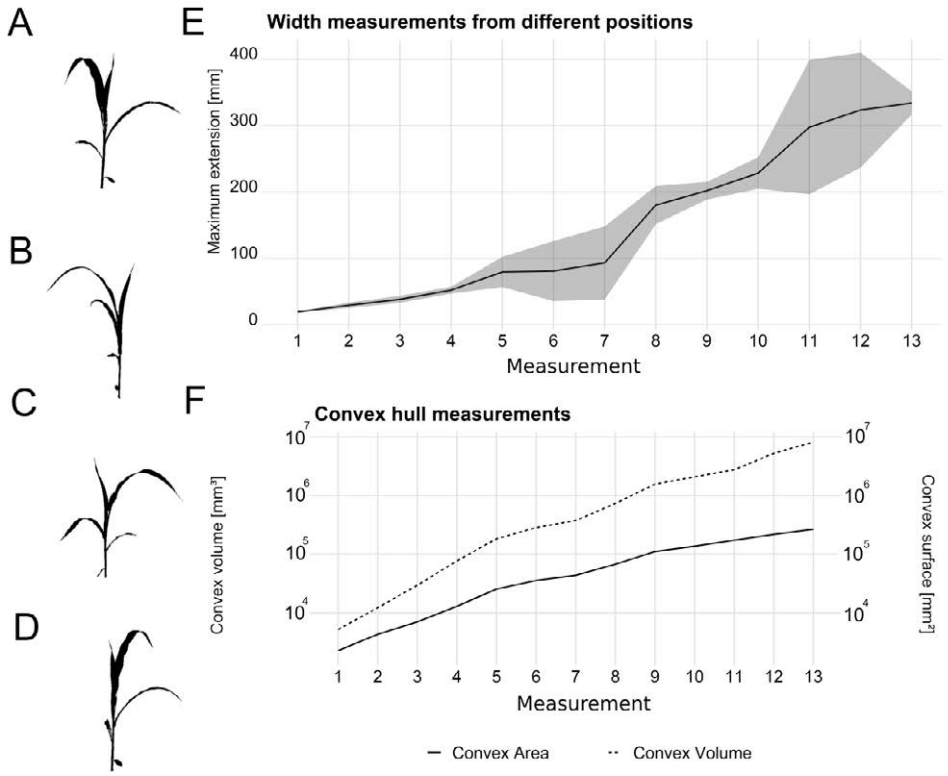


Figure 2. A maize plant from measurement 13 is shown in four different positions by rotating around the z-axis in 90° spin intervals (A, B, C, D). Graph E shows the mean width measurement (black line) together with the uncertainty using the minimum and maximum width value (gray area) from four different points-of-view. Width measurements result in a rather big inaccuracy depending on the point of view used. 3D measurement of the convex hull result in a point-of-view independent measurement for the volume / area of the convex hull (graph F). The area measurements have been scaled logarithmical to improve the comparability.

background shows the amount of variation between lowest and highest possible measurement for the width parameter. This error can cover more than 50% of the mean value (Figure 2E, measurement 6, 7 and 11).
In contrast to this, a 3D volume and area measurement as shown by the convex hull is well defined and its errors are exclusively coming from the sensor. Thus, a 3D measurement of the plant enables the derivation of point-of-view independent measurements that are well defined (Figure 2F). Both graphs show a clear growth curve with similar changes in slope.
By investigating the distribution of points along the x-axis the distribution of biomass can be analyzed. Therefore a prerequisite is high resolution point cloud with a very small amount of occlusions. Figure 3 shows the evolution of the point density / biomass for 13 points of measurement during two weeks of development (measurement 1-13), starting at growth stage 13 and ending at growth stage 15 (BBCH, Meier, 2018). To clearly connect point density to biomass, the point cloud was rastered to a 1 mm 3D grid. It shows clearly the levels of biomass concentration. Furthermore these measurements were only affected by measuring inaccuracies and not by different points-of-view. 3D point clouds enable a highly accurate measurement of well-defined and established parameters like height and leaf area, but also enable the derivation of new parameters like the convex hull as a measure for volume.
As the parameters introduced (Figure 1) do only depict a small overview of possible measurements, Table 1 gives a deeper understanding of which parameters were derived from LT or SfM approaches. These parameters were sorted regarding the applications for laboratory, greenhouse and field, respectively. As it is shown LT approaches are commonly used for laboratory whereas high resolution and high accuracy are needed, whereas structure from motion approaches are used in the greenhouse and in the field. This is due to the fast imaging needed in the field with a later processing, whereas LT needs time for accurate measurements which is often not applicable. Measurements in the greenhouse and in the field mostly concentrate on height measurements and parameters describing properties of a group of organs like leaf area index or projected canopy area. In contrast to this in laboratory parameters can be derived on plant organ level. The biological relevance of the traits shown is introduced. As 3D measuring devices work non-invasively these devices are well-suited to image growth processes over time as it has been shown in Cotton, Arabidopsis, Barley, Wheat and

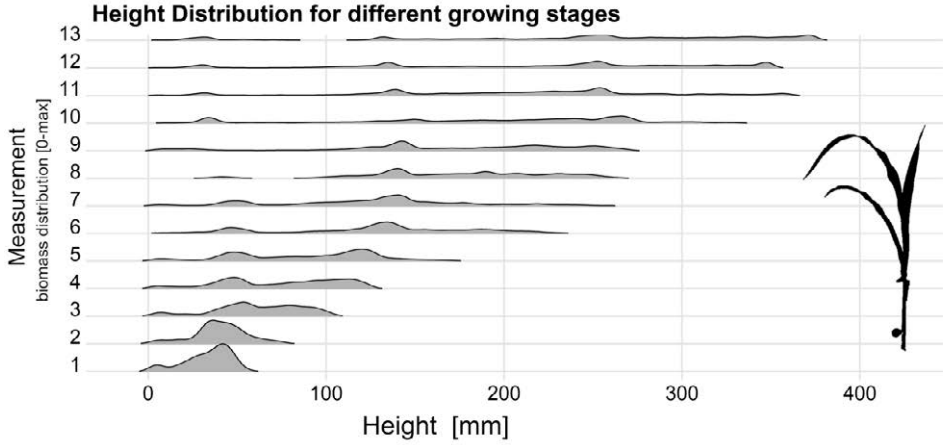


Figure 3. Distribution of plant points along the plants vertical axis is shown for the different growing stages between measurement 1 and 13. The x-axis shows the height level of the plant. The y-axis shows the different measurements and the gray values show the density of biomass at this level.

Pepper (Table 1). Traits have been extracted at different scales and can be used to be linked to QTL (quantitative trait loci, Van der Heijden *et al.,* 2012) for breeding, or to monitor drought effects as has been shown for sugar beet and soybean (Table 1).

Usually 3D reconstruction (SfM) or 3D measuring (LT) is accompanied by different problems like plant movement or environmental influences that can be minimized in the laboratory, but get stronger when the scale change to the greenhouse or to the field. In the laboratory as the number of photos that are taken can be optimized and the speed of the laser scanner can be slow, the resulting quality of the point cloud can be very high and complete. With increasing speed for high throughput in the greenhouse and with the use of less photos the result is a point cloud with a bigger amount of occlusion. Occlusion is low for small plants, but it increases with bigger plants and more complexity by more leaves. Self-occlusion by the plant leaf or stems can no more be eliminated by a larger number of viewpoints. Thus the quality decreases.

A further factor for greenhouse and field is wind. While in the greenhouse, air movement is mostly generated by fans or moving parts on the field there is usually a permanent wind. This produces motion within the plant. Hence, the different parts from single scans (LT and SfM) cannot be merged together. Doubled leaves or huge holes – namely parts without information – are generated.

Nevertheless the biggest challenge for 3D plant measuring is to link different scans over time. The issue with the previously described motion and occlusion are increased greatly by growth of the plant. Leaves change their position and the complete plant deforms. These problems can be approached by using highly accurate plant models such as L-Systems and Functional-Structural-Plant Models (FSPM), that allow tracking of single geometrical shapes (organs) over time (Vos *et al.,* 2009). To achieve this, the use of geometrical surface measurements is essential to generate models with high reliability.

As shown in Table 1, most research on 3D plant parameters on organ scale has been conducted during the last recent years. This is due to the increase of resolution and accuracy of algorithms for reconstruction and measuring systems and at the same time due to a decrease in price of these systems and the high availability of cameras on the market. Examples shown in the literature for use

Table 1. The derived plant parameters, the plant material used as well as the technology used and the biological connection are shown for the collected 3D plant parameters stated in literature at laboratory, greenhouse and field scales.

Scenario	Measured Trait	Plant	Used tech.	Biological connection	Literature
Laboratory	plant height & width	sugar beet	LT	trait analysis	Paulus *et al.,* 2014b
	root surface & volume	sugar beet	LT	trait analysis	Paulus *et al.,* 2014a
	leaf area	barley	LT	drought response	Paulus *et al.,* 2014b
	leaf length & width	cotton	Sfm	growth analysis	Paproki *et al.,* 2012
	leaf movement	arabidopsis	LT	growth analysis	Dornbusch *et al.,* 2012
	single leaf growth	barley	LT	growth analysis	Paulus *et al.,* 2014b
	ear volume	wheat	LT	yield estimation	Paulus *et al.,* 2013
Greenhouse	plant Height	pepper	SfM	QTL analysis	Van der Heijden *et al.,* 2012
	proj. canopy area	pepper	LT	growth analysis	Kjaer & Ottosen 2015
	leaf angle	soybean	SfM	drought response	Van der Heijden *et al.,* 2012
Field	plant height	wheat	SfM	growth analysis	Virlet *et al.,* 2017
	leaf area index	maize & sorghum	SfM	trait analysis	Shafiekhani *et al.,* 2017
	leaf area	grapevine	SfM	trait analysis	Klodt *et al.,* 2015

of 3D measuring systems indicate that the main application can be found in the field of trait and growth analysis, drought response and QTL analysis. The maize example can be seen as a prototypical application for specific traits and plants. In future, more and more plants will be measured and new traits will be parameterized. Studies will concentrate especially on the organ scale and for high throughput analysis routines in greenhouse and field will be the focus of future research.

Conclusions

Access to the geometrical information of plants is a key aspect to describe plant growth and to measure shape responses to environmental influences. The experiment described introduced the advantages of 3D measuring compared to established 2D plant images. Furthermore, two state-of-the-art techniques to measure 3D geometry of plants, i.e. active laser triangulation and structure from motion were introduced and application examples from the literature were shown and grouped into use at the different scales (laboratory, greenhouse and field) and linked to the measured plant type, the measuring method used and its biological relevance.
Three dimensional measurement is proven to be a well-suited tool for plant phenotyping with a high accuracy and the chance to measure completely new traits as a volume description or the biomass distribution.

References

Dornbusch, T., Lorrain, S., Kuznetsov, D., Fortier, A., Liechti, R., Xenarios, I., *et al.,* 2012. Measuring the diurnal pattern of leaf hyponasty and growth in Arabidopsis – a novel phenotyping approach using laser scanning. Functional Plant Biology, 39(11), 860. https://doi.org/10.1071/fp12018

Dupuis, J., Holst, C., Kuhlmann, H., 2015. Journal of Applied Geodesy, 10 (1) 37-44. https://doi.org/10.1515/jag-2015-0028.

Kjaer, K., & Ottosen, C.-O., 2015. 3D Laser Triangulation for Plant Phenotyping in Challenging Environments. Sensors, 15(6), 13533-13547. https://doi.org/10.3390/s150613533

Klodt, M., Herzog, K., Töpfer, R., & Cremers, D., 2015. Field phenotyping of grapevine growth using dense stereo reconstruction. BMC Bioinformatics, 16 (1) 143. https://doi.org/10.1186/s12859-015-0560-x

Meier, U., 2018. Entwicklungsstadien mono- und dikotyler Pflanzen. BBCH Monografie. Quedlinburg. Open Agrar Repositorium.

Paproki, A., Sirault, X., Berry, S., Furbank, R., & Fripp, J., 2012. A novel mesh processing based technique for 3D plant analysis. BMC Plant Biology, 12(1), 63. https://doi.org/10.1186/1471-2229-12-63

Paulus, S., Behmann, J., Mahlein, A.-K., Plümer, L., & Kuhlmann, H., 2014a. Low-Cost 3D Systems: Suitable Tools for Plant Phenotyping. Sensors, 14(2), 3001-3018. https://doi.org/10.3390/s140203001

Paulus, S., Dupuis, J., Mahlein, A.-K., & Kuhlmann, H., 2013. Surface feature based classification of plant organs from 3D laser scanned point clouds for plant phenotyping. BMC Bioinformatics, 14(1), 238. https://doi.org/10.1186/1471-2105-14-238

Paulus, S., Schumann, H., Kuhlmann, H., & Léon, J., 2014b. High-precision laser scanning system for capturing 3D plant architecture and analysing growth of cereal plants. Biosystems Engineering, 121, 1-11. https://doi.org/10.1016/j.biosystemseng.2014.01.010

R Development Core Team, 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rose, J., Paulus, S., & Kuhlmann, H., 2015. Accuracy Analysis of a Multi-View Stereo Approach for Phenotyping of Tomato Plants at the Organ Level. Sensors, 15(5), 9651-9665. https://doi.org/10.3390/s150509651

Rusu, R. B., & Cousins, S., 2011. 3D is here: Point Cloud Library (PCL). In IEEE International Conference on Robotics and Automation. IEEE. https://doi.org/10.1109/icra.2011.5980567

Shafiekhani, A., Kadam, S., Fritschi, F., & DeSouza, G., 2017. Vinobot and Vinoculer: Two Robotic Platforms for High-Throughput Field Phenotyping. Sensors, 17(12), 214. https://doi.org/10.3390/s17010214

Van der Heijden, G., Song, Y., Horgan, G., Polder, G., Dieleman, A., Bink, M., *et al.* (2012). SPICY: towards automated phenotyping of large pepper plants in the greenhouse. Functional Plant Biology, 39(11), 870. https://doi.org/10.1071/fp12019

Virlet, N., Sabermanesh, K., Sadeghi-Tehran, P., & Hawkesford, M. J., 2017. Field Scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring. Functional Plant Biology, 44(1), 143. https://doi.org/10.1071/fp16163

Vos, J., Evers, J. B., Buck-Sorlin, G. H., Andrieu, B., Chelle, M., & de Visser, P. H. B., 2009. Functional–structural plant modelling: a new versatile tool in crop science. Journal of Experimental Botany, 61(8), 2101-2115. https://doi.org/10.1093/jxb/erp345

Vegetation indices from remote sensing imagery as proxies for yield and grain N in wheat

M. Quemada[1*], J.L. Pancorbo[1], M. Alonso-Ayuso[1], J.L. Gabriel[2], J. López-Herrera[1] and E. Pérez-Martín[1]
[1]*Universidad Politécnica de Madrid. Avda. Complutense sn 28040 Madrid, Spain; miguel.quemada@upm.es*
[2]*INIA-INAGEA. Ctra. de la Coruña km 7,5, 28040 Madrid, Spain*

Abstract

The use of vegetation indices (VIs) might help to adjust fertilizer and irrigation, increase yield and reduce costs and nitrogen (N) losses. The objectives were to evaluate the use of VIs extracted from remote sensing imagery to estimate crop N status, yield and grain N content in wheat (*Triticum aestivum*, L.). A field experiment conducted with four N fertilizer levels randomly distributed in 32 plots (25×25 m^2), half irrigated and half rain-fed. At two sampling times, biomass and N uptake were determined, ground sensors measurements were taken, and multi-spectral imagery was acquired by an unmanned aerial vehicle. At harvest, yield was recorded with a combine and grain N determined. VIs obtained from ground measurements were highly correlated with those from the aerial platform. Using planar domain VIs, which relate N concentration and biomass, correlated well to crop N status and showed high potential for fertilizer recommendation, as well as yield and grain N prediction in wheat.

Keywords: nitrogen fertilization, crop sensors, unmanned aerial vehicle

Introduction

According to FAO, 15.9% of the total primary crops area harvested in the world was wheat (*Triticum aestivum*, L) in 2016. Adjusting fertilization to wheat requirements is a crucial strategy for increasing N use efficiency and reducing water and soil pollution (Arregui *et al.*, 2006). In recent years, different techniques have been applied and tested by researchers and farmers to determine the crop N status. Laboratory plant tissue analysis is a reliable method for calculating crop N status but it is expensive, slow and hard to apply to large fields (Haboudane *et al.*, 2002). In addition, in many cases, by the time the results are available, the physiological state of the crop has changed and it is of little support for decision making. The use of handheld chlorophyll meters (CM) has increased in recent years because it is a non-destructive and rapid method for measuring leaf chlorophyll and detecting N deficiencies (Tremblay *et al.*, 2009). Nevertheless, using CM to account for spatial variability in large fields could be very time-consuming. These limitations could be overcome by remote sensing, and its application to fertilizer management is providing valuable insights into improving decision making. Airborne sensors are used to measure canopy reflectance at different wavelengths, and the actual spectrum is affected by both vegetation structure and crop N concentration (Gabriel *et al.*, 2017). Based on this, many vegetation indices (VIs) have been developed to estimate crop N status (Chen *et al.*, 2010, Fitzgerald *et al.*, 2010). However, each crop requires a specific strategy for adjusting N fertilizer application depending on temporal evolution of crop N demand and quality criteria. Monitoring temporal variations in crop N content could allow adapting N application to crop N requirements over large areas of specific crops (Quemada *et al.*, 2014). In winter wheat, a common strategy is to split the synthetic fertilizer into two topdressings applications: one at the beginning of tillering and the rest during stem elongation for ensuring yield (Arregui *et al.*, 2006). Thereafter, foliar applications around flowering may be applied to increase grain protein content.

DOI 10.3920/978-90-8686-888-9_40, © Wageningen Academic Publishers 2019

The objectives of this study were to evaluate remotely sensed data to estimate biomass and crop N status at early stages and to predict yield and grain N concentration at harvest. Additionally, the use of planar domain VIs, that combine chlorophyll and structural VIs, was tested to see if they improved the estimation and prediction of crop variables.

Materials and methods

The study was carried out at La Chimenea field station, Central Spain (40°04'N, 03°32'W, 550 m a.s.l.) in a quarter of a 220 m radius pivot. The soil was mapped as *Haplic Calcisol*, had a pH ≈8.1, medium organic matter content, and a silty clay loam texture with low stone content throughout the soil profile. The study site was flat and the slope was <1%. The climate was classified as cold semi-arid (Bsk) according to Köppen classification. Normally, spring and summer are characterized by a large water deficit, but the spring 2018 was unusually wet. In the pivot, uniformly sown with winter wheat on November 2, 2017, were marked 32 plots (25×25 m).
A factorial experiment was established by randomly distributing the plots into four N (0, 50, 100 and 150 kg N/ha; N0, N1, N2 and N3, respectively) and two water levels, with four replications (Figure 1). Nitrogen fertilizer was hand broadcasted to plots in two stages: two thirds on January 25 (GS22) and one third on March 25 (GS35). Growth stages (GS) followed the decimal code (Lancashire *et al.*, 1991). Half of the plots received irrigation on May 8 to match crop water evapotranspiration with availability, whereas the other half remained rain-fed. All plots were harvested on July 20. Plots were georeferenced with RTK (Real Time Kinematic) through Global Navigation Satellite System (GNSS) technique with a HiPer® Pro receptor (Topcon Singapore Holdings Pte. Ltd, Singapore). The previous year, the field was sown with maize to ensure uniformity and low levels of soil inorganic N. Before sowing wheat, 50 kg P/ha and 70 kg K/ha were applied.

Crop analysis
Crop data were obtained at two different stages: at the mid-stem elongation (GS34, March 22) and at flowering (GS65, May 11). At each date, two biomass samples (0.50×0.50 m) were collected in each plot, dried at 65 °C and weighed to determine aerial biomass. A subsample was analysed for determining N concentration (% N) by the Dumas combustion method (LECO FP-428 analyser, St. Joseph, MI, USA). At harvest, the central stripe (1.5×20 m) of each plot was harvested with an experimental combine to measure yield, and a grain subsample of each plot was saved for analyzing grain N concentration by the Dumas method. For each plot, N content was calculated by multiplying N concentration times the aerial biomass. At harvest, the grain N content was calculated by

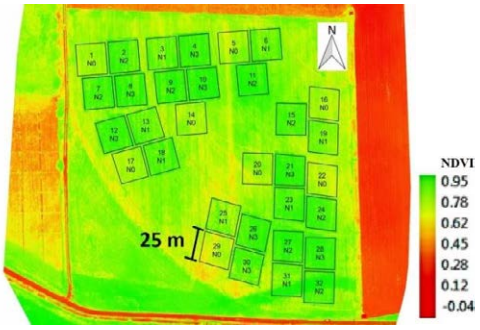


Figure 1. Normalized difference vegetation index (NDVI) from a field multi-spectral image with plot number and fertilizer level.

multiplying the grain N concentration times the yield. To measure the crop N status, the Nitrogen Nutrition Index (NNI) was calculated by referring the %N and biomass obtained at each sampling date (Mistele and Schmidhalter, 2008).

Ground-level optical measurements

Non-destructive measurements with two commercial optical sensors were taken to monitor the crop status at each sampling date. Dualex® Scientific (Force-A, Orsay, France), is a leaf clip sensor for testing crop N status and measures chlorophyll content (Chl) based on the ratio of the transmitted light at the red and infrared wavelengths (Tremblay *et al.*, 2009). This device also measures leaf flavonoid (FLAV) concentration by determining the optical absorption of the leaf epidermis under UV light. The Nitrogen Balance Index (NBI) is calculated as the ratio of Chl and FLAV. Greenseeker® (Handheld Crop Sensor Model HCS-100, Trimble, Sunnyvale, CA, USA) is an active light proximal sensor that measures NDVI. The sensor emits a burst of red and infrared light and measures the ratio of the light that is reflected. Each sampling date, the mean Chl, NBI, and Greenseeker were obtained as the average of 15 measurements per plot.

Remote sensing data acquisition and VI extraction

During stem elongation and flowering, two airborne campaigns were conducted as close as possible to ground-level measurements, ensuring a cloud-free sky day. The UAV images were acquired with a multi-spectral sensor (RedEdge-M, MicaSense, Inc., Seattle, USA) with an Instrument Landing System sensor and GNSS, 80 m over the experimental plots, on the solar plane at 11:40 GMT. Before and after the flights, images of the reflectance calibration panel for the camera were taken. The flight was conducted optimizing the path of the UAV through parallel tracks, with automatic pilot function according to the project plan, with a ground sampling distance of 6.81 cm. The UAV hexacopter used a system GPS A3 Pro of dual frequency, and a Ronin-MX stabilizer for the cameras with an inertial system for the positioning of the photo-centers. Seven control points were positioned on the ground with a GNSS in order to ensure an ortho-rectified image of the plots, providing a

Table 1. Vegetation indices used in this study.

Index	Equation	Reference
Structural indices		
Normalized Difference Vegetation Index	NDVI = (R800 – R670) / (R800 + R670)	Rouse *et al.*, 1974
Reformed Difference Vegetation Index	RDVI = (R800 – R670) / R800 + R670)^2	Roujean *et al.*, 1995
Chlorophyll indices		
Normalized Difference Red Edge	NDRE = (R790 – R720) / (R790 + R720)	Fitzgerald *et al.*, 2010
Transformed Chlorophyll Absorption in Reflectance Index	TCARI = 3[(R700 – R670) – 0.2 × (R700 – R550) × (R700 / R670)]	Haboudane *et al.*, 2002
Optimized soil-adjusted vegetation index	OSAVI = (1 + 0.16) × (R800 – R670) / (R800 + R670 + 0.16)	Rondeaux *et al.*, 1996
Combined	TCARI/OSAVI	Haboudane *et al.*, 2002
Double-Peak Canopy Nitrogen Index	DCNI = [R720 – R700) / (R700 – R670)] / (R720 – R670 + 0.03)	Chen *et al.*, 2010
NG	NG = NIR/Green	

5.4 cm/pixel resolution. This cinematic support was contrasted with the on-ground data, obtaining a precision of 29 mm in X, 33 mm in Y in planimetry and 21 mm in altimetry Z.
The camera had five independent image sensors that captured narrow wavelength bands centered at 475, 560 nm with a bandwidth of 20.0±2 nm, 668, 700 nm with a bandwidth of 10.0±2 nm, and 800 nm with a bandwidth of 40.0±2 nm. The sensors provided images made up of 1,280×960 pixels. The ortho-image was used to extract eight different VIs, classified as structural or chlorophyll indices, for each experimental plot using equations in Table 1.

Planar domain VI development
Planar domain indices were developed by representing in X axis each structural index and in Y axis each chlorophyll index from Table 1 (Fitzgerald *et al.*, 2010). These indices measure the relationship between the two VIs represented by comparing the distance of each point to the upper and bottom line that involve the cloud of points. Well fertilized plots will have a high value of both indices so will be close to the upper line, whereas N deficient plots will approach the bottom line. The canopy chlorophyll content index (CCCI) is a planar domain VI developed by Fitzgerald *et al.* (2010) designed to detect canopy N using NDVI as structural, and NDRE as chlorophyll indices. The indices obtained in this study will be compared with CCCI.

Statistical analysis
Multifactorial ANOVA analysis was conducted with the data obtained at harvest to determine the effect of N and water factors. A correlation analysis was performed to obtain the determination coefficient (R^2) between VIs and agronomic variables, and for measurements of different sensors. Statistical analyses were performed with R software (R Core team, 2018).

Results

Crop data
Yield was highly affected by N treatments ($P<0.001$), but not by the irrigation level ($P>0.1$), probably due to the unusually wet spring. Therefore, the water factor was not taken into account when studying the crop yield response. A linear-linear plateau model fit the yield response curve to N fertilizer application (Figure 2). The optimal N fertilizer rate was 111 kg N/ha and the yield plateau 5.8 Mg/

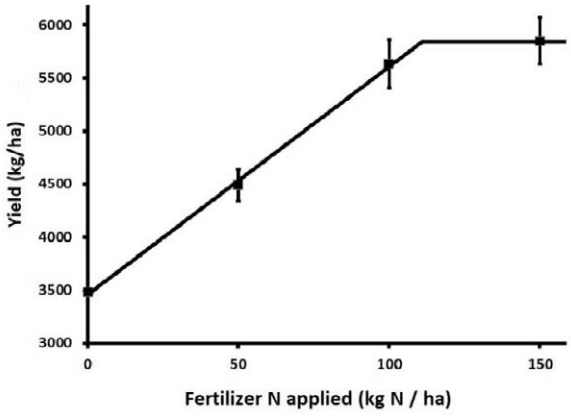


Figure 2. Yield response curve of wheat to N fertilizer rate. Squares are the mean with standard errors as bars, and the line is the adjusted linear-linear plateau model.

ha. The mean NNI of the three sampling dates was 0.8 for N0 treatment, 0.9 for N1, 1.1 for N2, and 1.4 for N3, highlighting that N0 and N1 treatments were N limited whereas N2 and N3 were not. The statistical analysis showed that the N treatments affected in a significant way grain N and grain N content ($P<0.01$). The irrigation levels affected significantly grain N ($P<0.01$) but not grain N content ($P>0.1$). Results indicated that with a higher irrigation amount, the grain N concentration decreased (Figure 3)

Crop N status and biomass

The relationship between ground level sensors and crop N status were better for the measurements taken at GS65 than at GS34 (Table 2). Determination coefficients were higher for N content, then for N concentration and NNI. Dualex measurements, both Chl and NBI, had an $R^2 \geq 0.6$ with N content at GS34 showing high potential for adjusting N fertilization in early stages. The relationship between Dualex and Chlorophyll indices from the multi-spectral image with crop biomass was significant and improved in GS65 compared to GS34.

The relationships between airborne structural indices (NDVI, RDVI) with biomass and crop N status improved when measurements were taken at GS65 (Table 2). The chlorophyll indices (TCARI, NDRE, DCNI, OSAVI, COMBINED, NG) were highly correlated to N content and biomass, but less to N concentration, especially at GS65. Among Chlorophyll indices, NG and OSAVI showed the best correlation with N concentration at both stages. The best correlation with most of the physiological variables was obtained with planar domain indices, particularly using NG as chlorophyll and NDVI as structural index (NG-NDVI). The main reason is that NG-NDVI did not saturate at high N fertilization rates.

The multi-spectral VIs that had the best relationship with Dualex Chl were $DCNI_{UAV}$, $NDRE_{UAV}$, and NG_{UAV}. Relationships between these airborne indices and Dualex Chl improved at a later GS. The relationship between NDVI measured with Greenseeker and extracted from the multi-spectral image was highly significant ($R^2>0.9$).

Crop yield

Indices obtained from sensor measurements, either at ground level or from UAV, had a highly significant relationship with wheat yield at harvest at both growth stages (Table 3). In similar studies conducted in maize, relationships between optical sensors in early growth stages and crop N status or yield were much lower than in wheat (Quemada *et al.*, 2014; Gabriel *et al.*, 2017). This could be explained by differences in canopy structure between both crops, while wheat reached a

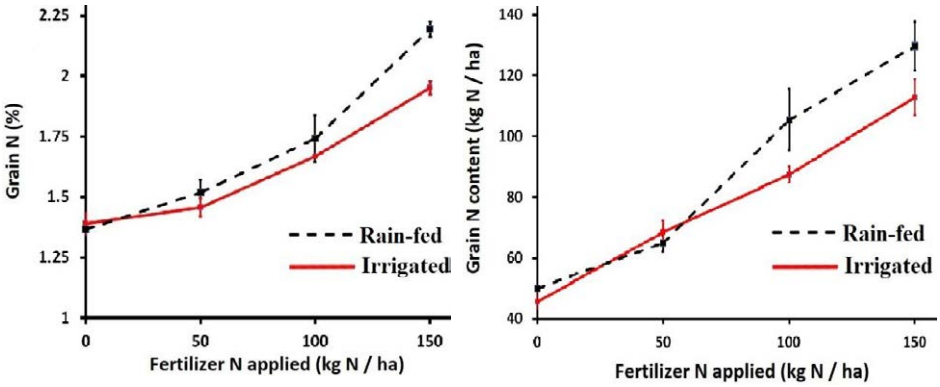


Figure 3. Grain N concentration (%) and grain N content (kg N/ha) at harvest for the various N levels and for the rain-fed and irrigated treatments. Squares are the mean with standard error as bars.

Table 2. Determination coefficient (R^2) between crop variables, ground level sensors (Dualex, Greenseeker) and indices extracted from the multi-spectral image acquired from a UAV at different growth stages (GS). Highly significant determination coefficients are marked in bold. [N] is biomass Nitrogen concentration.

	Stem elongation (GS34)				Flowering (GS65)			
	Biomass	[N]	N content	NNI	Biomass	[N]	N content	NNI
Dualex Chl	0.54	0.44	0.6	0.44	**0.76**	0.51	0.65	0.51
Dualex NBI	0.48	0.51	0.61	0.51	**0.79**	0.54	**0.7**	0.54
Greenseker NDVI	0.51	0.29	0.48	0.29	0.61	0.56	0.62	0.56
$NDVI_{UAV}$	0.55	0.29	0.51	0.29	0.65	0.56	0.62	0.56
$RDVI_{UAV}$	0.45	0.37	0.49	0.37	0.6	0.25	0.37	0.25
$TCARI_{UAV}$	0.3	0.22	0.34	0.22	0.65	0.54	0.62	0.54
$NDRE_{UAV}$	0.04	0.05	0.05	0.05	**0.78**	0.6	**0.71**	0.59
$DCNI_{UAV}$	0.17	0.07	0.14	0.07	**0.71**	0.35	0.52	0.35
$OSAVI_{UAV}$	0.51	0.35	0.52	0.35	**0.71**	0.61	0.69	0.61
$COMBINEDI_{UAV}$	0.32	0.23	0.36	0.23	0.68	0.56	0.64	0.56
NG_{UAV}	0.59	0.35	0.57	0.35	**0.75**	0.6	**0.71**	0.6
$NG\text{-}NDVI_{UAV}$	0.6	0.35	0.58	0.35	**0.76**	0.59	**0.71**	0.59
$NG\text{-}RDVI_{UAV}$	0.62	0.31	0.56	0.31	**0.73**	0.54	0.66	0.54
$CCCI_{UAV}$	0.03	0	0.02	0	**0.84**	0.57	**0.72**	0.56

near complete ground cover since early GS, maize only attains a high ground cover at advanced GS. These results highlight the potential of optical sensors for adjusting N fertilization to wheat demand.

Grain protein content

Dualex indices provided better estimation of grain N concentration and grain N content than the proximal sensor Greenseeker (Table 3). This is in agreement with results from the literature that showed the potential of leaf clip chlorophyll meters for estimating grain protein in wheat (Arregui *et al.*, 2006).

In a similar way, structural indices obtained from airborne imagery had a lower relationship with grain N and grain N content than planar domain indices. Particularly, the NG-NDVI planar domain index had an R^2>0.60 in GS37 and GS65 (Table 3), these are dates in which decision making for foliar N application to increase grain protein are usually taken by farmers.

Conclusions

These results highlight that a relevant application of remote sensing is the prediction of wheat yield and grain protein. Multi-spectral imagery was a useful tool for N fertilizer adjustment and, therefore, it could be used to reduce costs and increase the N use efficiency. Planar domain indices proposed in this study showed a better predictive capacity for wheat crop at stem elongation and flowering than other VIs, like NDVI. Combining structural and chlorophyll indices in planar domain indices improved the relationship with biomass and N content at both growth stages, GS34 and GS65.

Table 3. Determination coefficient (R^2) between estimations developed at different GS with yield, grain N concentration and grain N content at harvest. Highly significant determination coefficients are marked in bold.

	Stem elongation (GS34)			Flowering (GS65)		
	Yield	Grain N	Grain N content	Yield	Grain N	Grain N content
Dualex Chl	**0.7**	0.63	**0.72**	0.69	0.64	**0.74**
Dualex NBI	0.68	0.61	**0.71**	**0.7**	0.68	**0.76**
Greenseker NDVI	**0.72**	0.55	**0.7**	0.67	0.56	0.66
$NDVI_{UAV}$	**0.7**	0.48	0.63	**0.7**	0.52	0.66
$RDVI_{UAV}$	**0.7**	0.54	0.67	0.47	0.24	0.35
$TCARI_{UAV}$	0.39	0.28	0.37	0.67	0.49	0.63
$NDRE_{UAV}$	0.06	0.17	0.11	**0.75**	0.6	**0.74**
$DCNI_{UAV}$	0.23	0.04	0.13	0.43	0.55	0.55
$OSAVI_{UAV}$	**0.72**	0.53	0.67	**0.74**	0.59	**0.72**
$COMBINEDI_{UAV}$	0.42	0.29	0.39	**0.71**	0.51	0.66
NG_{UAV}	**0.74**	0.61	**0.74**	**0.74**	0.62	**0.75**
$NG\text{-}NDVI_{UAV}$	**0.74**	0.62	**0.74**	**0.73**	0.63	**0.76**
$NG\text{-}RDVI_{UAV}$	**0.71**	0.57	**0.7**	**0.7**	0.56	**0.7**
$CCCI_{UAV}$	0.03	0	0	**0.72**	0.62	**0.73**

Acknowledgements

Project funded by CICYT (AGL2017-83283-C2-1-R and CAM (AGRISOST-III).

References

Arregui, L.M., Lasa, B., Lafarga, A., Iraneta, I., Baroja, E. and Quemada, M., 2006. Evaluation of chlorophyll meters as tools for N fertilization in winter wheat under humid Mediterranean conditions. European Journal of Agronomy, 24, 140-148.

Chen, P., Haboudane, D., Tremblay, N., Wang, J., Vigneault, P. and Li, B., 2010. New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. Remote Sensing of Environment, 114(9), 1987-1997.

Fitzgerald, G., Rodriguez, D. and O'Leary, G., 2010. Measuring and predicting canopy nitrogen nutrition in wheat using a spectral index – The canopy chlorophyll content index (CCCI). Field Crops Research, 116(3), 318-324.

Gabriel, J.L., Zarco-Tejada, P.J., López-Herrera, P.J., Pérez-Martín, E., Alonso-Ayuso, M. and Quemada, M., 2017. Airborne and ground level sensors for monitoring nitrogen status in a maize crop. Biosystems Engineering, 160, 124-133

Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J. and Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote Sensing of Environment, 81, 416-426.

Lancashire, P.D., Bleiholder, H., Boom, T.V.D., Langelüddeke, P., Stauss, R., Weber, E. *et al.*, 1991. A uniform decimal code for growth stages of crops and weeds. Annals of applied Biology, 119(3), 561-601.

Mistele, B. and Schmidhalter, U., 2008. Estimating the nitrogen nutrition index using spectral canopy reflectance measurements. European Journal of Agronomy, 29(4), 184-190.

Quemada, M., Gabriel, J.L. and Zarco-Tejada, P., 2014. Airborne hyperspectral images and ground-level optical sensors as assessment tools for maize nitrogen fertilization. Remote Sensing, 6, 2940-2962.
R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rondeaux, G., Steven, M. and Baret, F., 1996. Optimization of soil-adjusted vegetation indices. Remote Sensing of Environment, 55, 95-107.
Roujean, J.L. and Breon, F.M., 1995. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. Remote Sensing of Environment, 51, 375-384.
Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W. and Harlan, J.C., 1974. Monitoring the vernal advancement of retrogradation of natural vegetation, NASA/GSFC, Type III, Final Report. Greenbelt, MD, USA, 371.
Tremblay, N., Wang, Z. and Belec, C., 2009. Performance of Dualex in spring wheat for crop nitrogen status assessment, yield prediction and estimation of soil nitrate content. Journal of Plant Nutrition, 33(1), 57-70.

Using terrestrial photogrammetry for leaf area estimation in maize under different plant growth stages

D. Reiser*, A. Kamman, M. Vázquez Arellano and H.W. Griepentrog
University of Hohenheim, Institute of Agricultural Engineering, Garbenstr. 9, 70599 Stuttgart, Germany; dreiser@uni-hohenheim.de

Abstract

A continuous determination of leaf area (LA) could help to adjust fertilization, pest management and harvesting. In this research, a non-destructive measuring method for LA estimation was correlated with results based on the analysis of dense point clouds, gained by photogrammetry. Maize rows were planted and were analysed weekly over one month, resulting in five different growth stages. For determining the LA out of the resulting point clouds, the plant points were separated from ground points. Afterwards, a point distance filter was applied and the resulting points were counted. The results showed a high correlation between LA and the filtered point cloud number. The filter distance could be set to a polynomial correlation dependent on the growth stage.

Keywords: photogrammetry, leaf area, autonomous robot, point cloud, spatial distance filter

Introduction

Farmers use their experience and visual feedback to assess plant production and plant health. However, as farmers have to deal with bigger areas today, than in previous times, it gets more difficult to detect changes at the right time. Modern technologies like smart sensors, drones and autonomous robots could help to close the gap if they are able to detect developments reliably and automatically. One option to estimate plant development is biomass change, which can be assessed by estimating the leaf area (LA) of plants (Francis *et al.,* 1969). In particular, the leaf area index is an important parameter for estimating plant heath, yield and development of plants (Kalisperakis *et al.,* 2015). Unfortunately, it is not easy to obtain reliable and accurate data. Direct manual and non-destructive measurements are time-consuming, labour intensive and not practical for modern optimized farming. If there is a way to automate the LA estimation of plants, it could be used to adjust treatments and could help to react to local heterogeneities. 3D imaging and reconstruction with photogrammetry could be a good way to automatically estimate the LA of plants. In combination with autonomous robots acquiring the data, the estimation of LA could be fully automated.
There have been many attempts to evaluate LA with different methods and sensors. Methods to evaluate LA have included satellite or airborne data with sensors like LiDAR (Li *et al.,* 2017), multispectral imaging (Liu *et al.,* 2016), depth sensors (Andújar *et al.,* 2016) or photogrammetry (Aguilar *et al.,* 2010). Photogrammetry uses the structure from motion technique for gaining 3D information out of distributed pictures (Vázquez-Arellano *et al.,* 2016). This technology creates 3D point clouds, which could be assessed in post-processing for the objects of interest. For terrestrial robots, this technology could be used by attaching a camera to the robot while driving through the crop rows. Terrestrial scanning for point cloud generation has already been done with laser scanners for individual trees, orchards and maize plants (Arnó *et al.,* 2013; Li *et al.,* 2017; Su *et al.,* 2018). Typical methods for estimating the LA out of point clouds were voxel grid filtering, 3D surface reconstruction and to cut the plants into slices to count the points at each area individually (Zheng and Moskal, 2009). The use of terrestrial photogrammetry for LA estimation of maize plants and the evaluation of the methodology over different growth stages has not been done until now.

DOI 10.3920/978-90-8686-888-9_41, © Wageningen Academic Publishers 2019

The objective of this research was to compare a manual measurement method for LA estimation with results based on the analysis of dense point clouds obtained by photogrammetry. The interest is to see if a point distance filter-based method could be used to robustly estimate the LA of maize plants, without taking shape analysis into account. To test the robustness, the filter should be applied under different growth stages, to see if the method is applicable to different plant shapes and sizes.

Material and methods

Experimental area and LA ground truth

Four 10 m long maize rows (*Zea mays* L.) were planted in a greenhouse. This area was used because it was independent of weather conditions for making measurements. The plant spacing was 0.13 m, with a Gaussian distribution as normal for maize, seeded in outdoor conditions. Each plant position was measured after emergence with a total station (SPS 930, Trimble, Sunnyvale, USA). The accuracy of the total station for static measurements is ±2 mm. For the ground truth of the LA of maize, the following formula was used (1):

$$LA = L \times W \times A \quad (1)$$

With *L*=leaf length, *W*= Leaf width and *A* as constant factor of 0.75. This formula was first used by Montgomery (1911) who evaluated an error of 2% for this method. In other research, factor *A* varied between 0.72 and 0.79 (Mokhtarpour *et al.,* 2010; Montgomery, 1911). For obtaining the ground truth of the LA, a measuring tape was used.

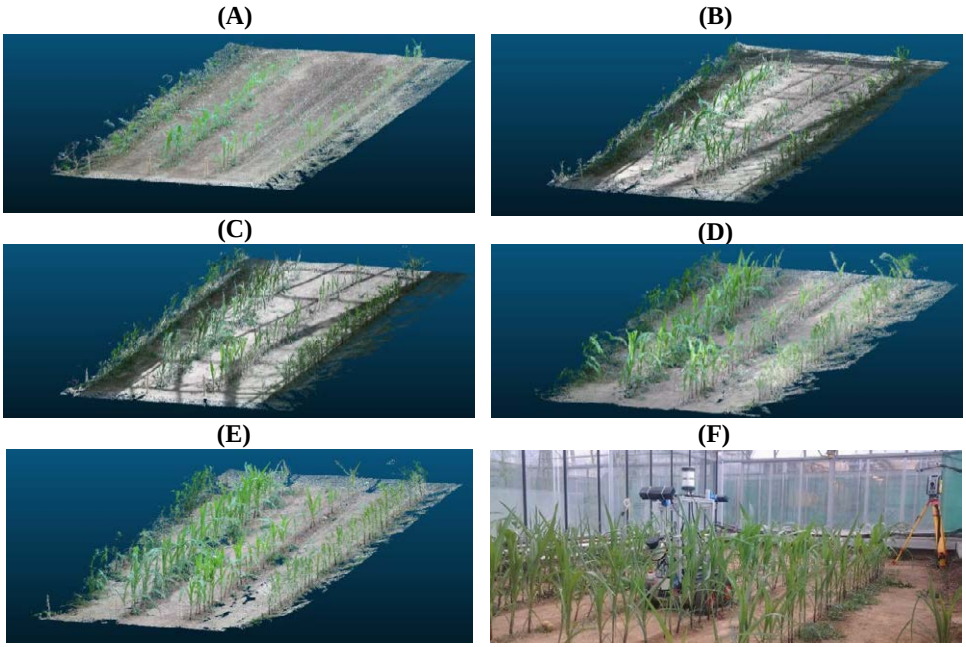


Figure 1. Visualisation of the point clouds generated by pix4D photogrammetry software 40 (A), 49 (B), 54 (C), 62 (D) and 69 (E) days after seeding. (F) The robot with camera mounting and total station 69 days after seeding.

Hardware and sensors

For collecting the images, the Kinect V2 camera was used (Microsoft, Redmond, WA, USA). Just the RGB camera information with a 1,920×1,080 (HD) resolution was processed in this experiment. The camera was mounted on as mobile robot with four wheels and size of 0.5×0.6×1.0 m. The robot was able to drive through the crop rows, adjusted by remote control to not harm the plants by a malfunction (Reiser *et al.,* 2016; Vázquez-Arellano *et al.,* 2018). The sensor was mounted at a height of 0.94 m above ground level and was turned downwards at 30 degrees (Figure 1F). It was taken care, that the camera was covering the whole area, recording images at a rate of 4 Hz. With the average driving speed of 0.02 m/s, around 50 images per path were collected. Both directions were covered, resulting in a total rate of roughly 300 images per test day.

Software and data processing

For creating the 3D point clouds for the analysis, the images were loaded into Pix4D software (Pix4D, Lausanne, Switzerland) to create a 3D environment of the greenhouse (Stretcha *et al.,* 2014). The software performed image stitching and computed the point clouds for the five measurements. The average computation time was around 30 minutes per point cloud (Toshiba i5-4200M – 2.50 GHz, 16 GB RAM, Internal intel graphic card). The point clouds were exported and analysed using the Cloud Compare software (CloudCompare, 2018). For referencing the point clouds to the same coordinate frame, ground truth of the plant positions obtained by the total station was used. The point clouds were transformed manually to the measured stem positions, using the cloud compare plugin for Iterative Closest Point (ICP) alignment of point clouds. The scaling of the point clouds was adjusted based on the total station points. These aligned point clouds were processed further to separate the plant points. First, the surroundings of the greenhouse were cropped and afterwards the soil was cut out, using the cloud compare tools. As the ground was even, this could be performed with simple distance filters. The resulting plant points were analysed with different filter settings.
To automate the filtering process and to compare different filter settings, an Excel macro was programmed. As the cloud compare software could be used as a command line tool, the filter settings could be adjusted by the Excel macro and the corresponding point cloud number could be directly registered in the Excel sheet. As the main tool for the filtering process, the cloud compare subsample procedure was used. The function picks out the single points of a point cloud so that the distance between the points do not exceed a specified value (CloudCompare, 2018). For calculating the filter distance for a specific point density, the following formula was used:

$$distance = \sqrt{\frac{A}{\pi}} * 2 \quad (2)$$

For example, when one point per cubic centimetre (A) remained, a distance value of 11.3 mm would be suitable for a rough estimation. The Excel macro was programmed to change the distance automatically and increase the distance with 0.1 mm steps. The number of resulting points for different filter distances was used and compared to the hand measurement ground truth. For each point cloud, 200 measurements were analysed. The best filter result was assumed, when the point cloud number fitted with the measured LA in square centimetres. For estimating the best fit, the Matlab R2017b Curve Fitting tool was used.

Results and discussion

Data acquisition and point cloud processing

The matching of the images for generating the point clouds was in all datasets successfull but did not manage to match all provided images, as the similarity of the features was sometimes too close. Anyhow, a point cloud could be generated for each day. The results of the five different growth stages can be seen in Figure 1.

The row to the far left was defined as row 1, counting up to the far right. The overall LA changed for the five test days from 1.1470 to 5.9219 m^2 based on the manual measurements. For further processing, each row was handled as an individual point cloud and was compared to the ground truth. The original point cloud number of the single rows and the corresponding ground truth is shown in Table 1. The LA increased for each row from test day to test day. In general, rows 2 and 4 developed slightly better than rows 1 and 3. Row 3 had the smallest development of leaf area. This difference could be detected at the increasing point cloud number of all rows. Just in row 1, the point cloud number decreased once, with increasing times after seeding. One reason for this could be the incomplete row reconstruction of row 1, caused by the adjacent wall of the greenhouse.

LA estimation and algorithm performance

After applying the distance filter to the point clouds, the remaining plant points were counted and compared to the ground truth value. As a first approximation, the point number was directly compared to the LA ground truth in m^2. In Table 2, the resulting best fit distance filter values for each row is shown.

The mean filter results for the rows were between 0.0137 and 0.0067 m. The mean values per test day varied between 0.0094 and 0.0135 m. This fits roughly with the expected filter value of 0.0113 m when assuming the area of one square centimetre, as shown in Equation 2. It could be also seen that the values from row 1 were much smaller than the other values of the rows. That could be caused by the incomplete reconstruction of the row, as most parts of the plants were not reconstructed and not part of the point clouds (Figure 1). Therefore, the results from row 1 were excluded from the following analysis. At each row, the value for the optimal filter value decreased with increasing

Table 1. Description of the original plant points per row and the ground truth.

Days after seeding	LA ground truth in [m^2]				Original point cloud number			
	row 1	row 2	row 3	row 4	row 1	row 2	row 3	row 4
40	0.2824	0.4191	0.1634	0.2819	8,342	65,048	21,575	18,636
49	0.6332	0.5722	0.2620	0.4174	14,893	90,389	53,239	21,717
54	1.0409	1.0729	0.5068	1.0398	22,702	92,477	50,970	33,182
62	1.5175	1.5402	0.6218	1.1695	22,936	194,458	92,494	47,100
69	1.7100	1.7763	0.7997	1.6357	19,180	205,525	122,369	43,543

Table 2. Best distance filter results.

Days after seeding	Filter values in [m]				
	row 1	row 2	row 3	row 4	mean value day
40	0.0090	0.0151	0.0136	0.0163	0.0135
49	0.0078	0.0141	0.0151	0.0135	0.0126
54	0.0078	0.0111	0.0125	0.0101	0.0104
62	0.0053	0.0113	0.0136	0.0101	0.0101
69	0.0036	0.0119	0.0139	0.0082	0.0094
Mean value row	0.0067	0.0127	0.0137	0.0116	

LA measured. This led to the assumption that there is a direct correlation between LA and filter values possible. When comparing the values of the ground truth together with the best-fit filter value obtained, it could be estimated, that bigger LA results in a lower value for the distance filter. Using a regression function could show a linear or polynomial relation between the best point cloud filter value and LA (Figure 2). When using a linear fit curve, an R-square of 0.62 was calculated. Using a polynomial fit curve with a second degree, the R-square was 0.66. The best regression with a polynomial fit curve with a second degree was:

$$f(x) = 0.00195x^2 - 0.00752x + 0.01666 \quad (3)$$

The Figure 2 shows the fit function of Equation 3 together with the filter values and the corresponding LA ground truth.
This fitting curve could be used to estimate better filter values based on the expected LA. When using the ground truth for estimating the filter value based on the fit curve, the error could be estimated from the original cloud. The Figure 3 shows estimates of two scenarios. The first takes a fixed value

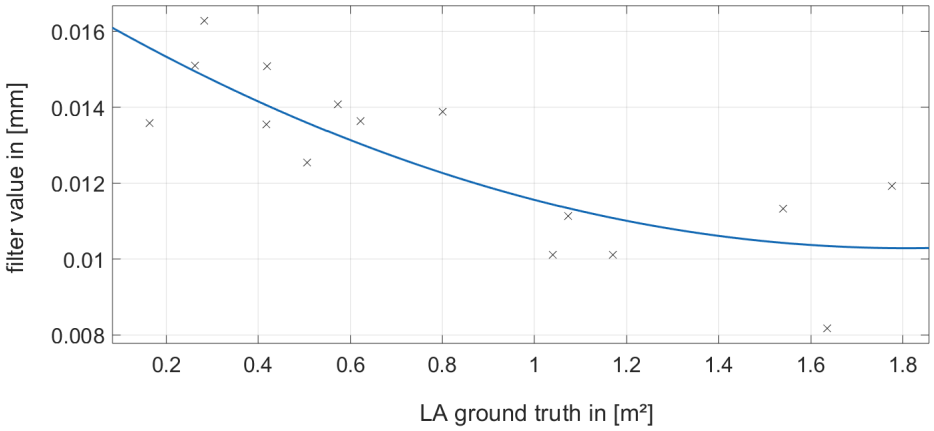


Figure 2. Behaviour between LA ground truth and the resulting filter value.

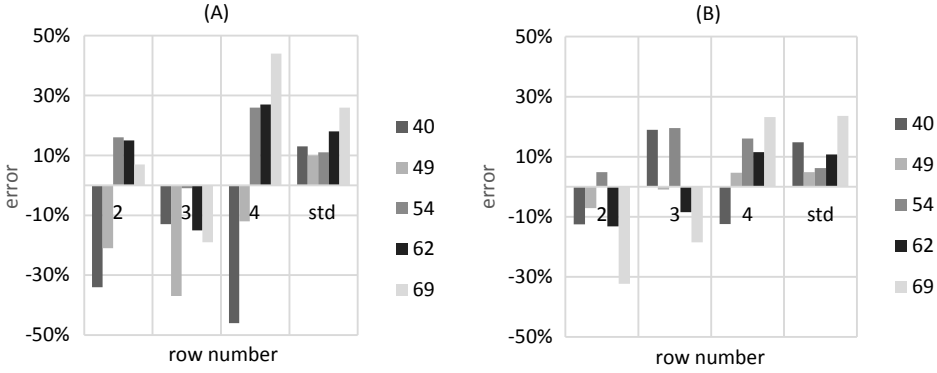


Figure 3. (A) constant value and (B) adaptive threshold error development. The legend describes the test days based on the days after seeding.

based on the mean value of all rows gained from row 2-4 of 0.0125 m (Figure 3A). The second scenario uses the regression curve to adaptively change the filter value based on the LA (Figure 3B). The results indicate that using the regression curve could decrease the error rate for the LA estimation. When comparing Figure 3A with Figure 3B, the error decreased from 46% to a value of less than 32%. The overall standard deviation for all three rows (std at Figure 3) decreased from 26 to 24% with the new estimation method. The method described is easy to carry out, to implement and to automate. Therefore, the error is acceptable for the given task, as five different growth stages were analysed. It should be mentioned, that one reason for errors could be the ground truth estimation process. As this time-consuming task needs the measuring person to work precisely for several hours, this could lead to some errors in calculating the leaves. A second reason for errors could be the inhomogeneity of the plant height for each row. As just one filter value per row was applied, this could have caused over and underestimations based on the LA. The estimated filtering method assumes, that the absolute value of the LA is known. This could be estimated out of growth models or using specific point cloud information like the plant height or volume. As the technique is independent on the measurement platform, it could be applied when integrated into a conventional implement or by an autonomous robot. The LA information could afterwards be used, to adjust fertilization, spraying or pest control, detection of water stress, or could help to assess new phenotypes.
Future work could address automatic filter value estimation based on point cloud information. The camera position and quality could be adjusted for better 3D reconstruction. Additionally, the colour information of the point cloud could be included, for better plant segmentation. To transfer the technique to the field, the photogrammetry must be calibrated with known objects or an exact geo-referencing of the image positions. Separating the areas into smaller areas inside the rows could help to improve the quality of the assessment.

Conclusions

A mobile robot was used to acquire camera data from maize rows for five different growth stages. This data was used to generate 3D point clouds and to estimate the LA for all growth stages. A distance filter was suitable to detect a high correlation between LA and point cloud number. When using an assumption of the real LA, a polynomial exponential function could be used to minimize the error from 46 to 32% and the standard deviation from 26 to 24% for all growth stages.

References

Aguilar, M. A., Pozo, J. L., Aguilar, F. J., García, A. M., Fernández, I., Negreiros, J., *et al.,* 2010. Application of close-range photogrammetry and digital photography analysis for the estimation of leaf area index in a greenhouse tomato culture. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives, 38(PART 5), 5-10. http://www.scopus.com/inward/record.url?eid=2-s2.0-84896969703&partnerID=tZOtx3y1

Andújar, D., Dorado, J., Fernández-Quintanilla, C., & Ribeiro, A., 2016. An approach to the use of depth cameras for weed volume estimation. Sensors 16(7), 1-11. doi:10.3390/s16070972

Arnó, J., Escolà, A., Vallès, J. M., Llorens, J., Sanz, R., Masip, J., *et al.,* 2013. Leaf area index estimation in vineyards using a ground-based LiDAR scanner. Precision Agriculture, 14(3), 290-306. doi:10.1007/s11119-012-9295-0

CloudCompare, 2018. https://www.danielgm.net/cc/. Accessed 22 November 2018

Francis, C. A., Rutger, J. N., & Palmer, A. F. E., 1969. A Rapid Method for Plant Leaf Area Estimation in Maize (Zea mays L.)1. Crop Science, 9(5), 537-539. doi:10.2135/cropsci1969.0011183X000900050005x

Kalisperakis, I., Stentoumis, C., Grammatikopoulos, L., & Karantzalos, K., 2015. Leaf area index estimation in vineyards from UAV hyperspectral data, 2D image mosaics and 3D canopy surface models. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives, 40(1W4), 299-303. doi:10.5194/isprsarchives-XL-1-W4-299-2015

Li, S., Dai, L., Wang, H., Wang, Y., He, Z., & Lin, S., 2017. Estimating Leaf Area Density of Individual Trees Using the Point Cloud Segmentation of Terrestrial LiDAR Data and a Voxel-Based Model. Remote Sensing, 9(11), 1-16. doi:10.3390/rs9111202

Liu, K., Zhou, Q., Wu, W., Xia, T., & Tang, H., 2016. Estimating the crop leaf area index using hyperspectral remote sensing. Journal of Integrative Agriculture, 15(2), 475-491. doi:10.1016/S2095-3119(15)61073-5

Mokhtarpour, H., Christopher, B. S., Saleh, G., Selamat, A. B., Asadi, M. E., & Kamkar, B., 2010. Non-destructive estimation of maize leaf area, fresh weight, and dry weight using leaf length and leaf width. Communications in Biometry and Crop Science, 5(1), 19-26.

Montgomery, E. G., 1911. Correlation studies in corn. Nebraska Agricultural Experiment Station Annual Report, 24, 108-159.

Reiser, D., Garrido, M., Arellano, M. V., Griepentrog, H. W., & Paraforos, D. S., 2016. Crop Row Detection in Maize for Developing Navigation Algorithms under Changing Plant Growth Stages. In Advances in Intelligent Systems and Computing (Vol. 417, pp. 371-382). Lisbon, Portugal: Springer. doi:10.1007/978-3-319-27146-0_29

Stretcha, C., Krull, M., Betschart, S., & Pix4D., 2014. The Chillon Project: Aerial/Terrestrial and Indoor Integration. Muldersdrift, South Africa: EE Publishers (Articles: PositionIT), 40-49. http://www.ee.co.za/article/chillon-project-aerialterrestrial-indoor-integration.html#

Su, W., Zhu, D., Huang, J., & Guo, H., 2018. Estimation of the vertical leaf area profile of corn (Zea mays) plants using terrestrial laser scanning (TLS). Computers and Electronics in Agriculture, 150, 5-13. doi:10.1016/j.compag.2018.03.037

Vázquez-Arellano, M., Griepentrog, H. W., Reiser, D., & Paraforos, D. S., 2016. 3-D Imaging Systems for Agricultural Applications – A Review. Sensors, 16(618), 24. doi:10.3390/s16050618

Vázquez-Arellano, M., Reiser, D., Paraforos, D. S., Garrido-Izard, M., Burce, M. E. C., & Griepentrog, H. W., 2018. 3-D reconstruction of maize plants using a time-of-flight camera. Computers and Electronics in Agriculture, 145. doi:10.1016/j.compag.2018.01.002

Zheng, G., & Moskal, L. M., 2009. Retrieving Leaf Area Index (LAI) Using Remote Sensing: Theories, Methods and Sensors. Sensors, 9(4), 2719-2745. doi:10.3390/s90402719

Early detection of *Fusarium* infection in corn using spectral analysis

T. Sandovsky[1], Y. Edan[1], S. Gad[1], A. Etzioni[2], T. Nacson[3] and V. Alchanatis[3*]
[1]*Department of Industrial Engineering & Management, Ben-Gurion University of the Negev, Beer Sheva 8410501, Israel*
[2]*Evogene, 13, Gad Feinstein, Rehovot 76120, Israel*
[3]*Institute of Agricultural Engineering, Agricultural Research Organization (ARO), Volcani Center, Rishon-LeZion, Israel; victor@volcani.agri.gov.il*

Abstract

This work presents a non-destructive methodology for early detection of *Fusarium* infection, by spectral analysis in the 350-2,500 nm range. Corn plants in greenhouse conditions were analysed using spectral analysis. The Lasso model was used to differentiate infected from non-infected plants based on the first derivative of leaf spectral reflectance. *Fusarium* infection was successfully recognized in plants at V2 growth stage with 74% success rate. This result enables infection detection at a stage which currently is not possible without destroying the plant, which can be further applied to map the disease in field scale.

Keywords: multispectral, spectral analysis, disease detection, *Fusarium*

Introduction

Plant diseases cause major production and economic losses in agriculture. It is estimated that the crop losses due to plant pathogens in the USA amount to about 33 billion dollars every year (Savary *et al.*, 2012). In recent years, *Fusarium* has been pronounced worldwide as a disease of economic importance (Windels, 2000). *Fusarium* is a phytopathogenic fungus with a global distribution, capable of infecting a wide range of crop plants, including cereals such as maize, wheat or barley (Jurado *et al.*, 2006). It may rapidly result in very high crop losses and quality reductions. Moreover, mycotoxins, potentially generated by these fungi, are poisonous and harmful for both human and animal nutrition (Browne *et al.,* 2005). These compounds may be present even after removal of mycelium and since most of them are resistant to physical and chemical treatments, they usually stay in the food during processing and storage (Falasconi *et al.*, 2005).
Early detection of the fungal species producing mycotoxins or of the mycotoxins themselves has become very important to prevent the human and animal risk deriving from entry of mycotoxins into the food chain (Falasconi *et al.*, 2005). Traditional methods used for plant assessment are still time-consuming, labour intensive and destructive in nature (Busemeyer *et al.*, 2013). Introduction of modern sensing technologies can improve crop yield, provide information to enable better in-field management decisions, reduce chemical and fertilizer costs through more efficient application, permit more accurate farm records, increase profit margin and reduce pollution (Li *et al.,* 2014).
The appearance of *Fusarium* in infected ears and plants, largely changes during the development of this disease, mostly due to degradation of chlorophyll contents and pronounced water losses. Hence, these changes lead to pronounced variations in spectral properties of infected grains and total ears. Both fungi and bacteria usually cause damage at molecular, cellular and/or tissue levels, which, in turn, can be detected as changes in the spectral signatures (West *et al.*, 2003).
Utilization of spectral analysis for detection of fungal and bacterial diseases is currently extensively under investigation with several working applications (Ray *et al.*, 2017). According to Bauriegel *et al.* (2011), in the spectral range of 400-1000 nm, there was a significant difference between the spectral patterns of healthy and diseased ears of wheat during BBCH-stage 71-85 (Lancashire *et al.,* 1991). There are no known works of early detection of *Fusarium* in corn.

DOI 10.3920/978-90-8686-888-9_42, © Wageningen Academic Publishers 2019

The objective of this study was to develop a method to detect infected and non- infected corn plants at early growth stages, before it can be visually identified on the leaves, based on leaf spectral reflectance.

Material and methods

Plant material

Corn seeds were infected and germinated in a greenhouse for a period of 3 weeks, at Evogene's research facilities in central Israel (31° 52' 55.80' N 34° 50' 30.77' E). Twelve bio pesticide microbial treatments were applied in two commercial corn varieties (P2088, 7210), on infected and non-infected plants in order to create variation in disease status of the various plants. Infection was created by drench with 10^4 spores of *Fusarium graminearum*. Four seeds were germinated in 380 ml pots. Five pots were repeated for each treatment. 140 pots were sampled, each pot contained 4 plants, but during the experiment only one plant (the same plant each time) was sampled.

The greenhouse temperature was maintained at 24±2 °C during the day and at 20±2 °C during the night. Drip irrigation with no fertilization (inhibits *Fusarium* infection) was applied, and irrigation frequency was defined according to plot weight at 50% water content.

Experimental and measuring procedure

The plants were sampled four times during a two-week period (from 11 days after seeding, stage V1, until 22 days after seeding, stage V2) between 09:00 to 12:30 local time (UTC-2). Developmental stages of growth were graded according to the BBCH scale (Lancashire *et al.,* 1991). At each sampling day, all plants were sampled with a field spectro-radiometer. Data were acquired from the last leaf that was fully developed. After the last sampling day, the roots of the plants were exposed and the disease infection level was manually examined by an agronomist expert. The disease was classified from level 0 to 5 where 0 is healthy, 5 is an infected plant, according to the state of the roots of the different plants (Figure 1). Due to the small number of samples, the number of infestation levels were aggregated from 6 to 2: levels 0 and 1 were aggregated to non-infected, and levels 2 to 5 were aggregated to infected.

Spectral measurement system

Spectral measurements were performed by point spectral measurements with a field spectro-radiometer (ASD FieldSpec 4 hi-res, ASD Inc. Malvern Panalytical, Boulder, Colorado, USA) with a 25° field of view in the range of 350-2,500 nm with 3 nm spectral resolution in the visible and near- infrared (VNIR) range and 8 nm in the short-wave infrared (SWIR) range. All measurements

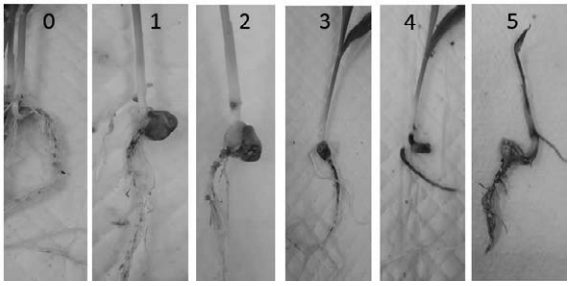


Figure 1. Labelling protocol of plants to different levels of disease according to plant status.

were conducted at fixed locations at approximately 10 mm distance from the leaf surface, resulting to an effective sampling area of about 60 mm^2. A 100 mm diameter spectralon plate was used as a white reference, to obtain reflectance curves.
Each recorded spectrum was the average of four sequential measurements at the same spot, where each measurement was the average of 30 full spectrum scans. Total acquisition time for each leaf was about 10 s.

Data analysis

Normalized difference indices
The normalized difference vegetation index (NDVI) (Rouse *et al.,* 1974) is the most widely used for retrieval of vegetation canopy biophysical properties (Jiang *et al.,* 2006). This index normalizes the differences between two selected bands. It has been related to crop variables such as biomass, leaf area, plant cover, leaf gap fraction, nitrogen and chlorophyll in cereals (Hansen and Schjoerring, 2003). NDVI values can range from (-1) to (1) while the normalization is effective in standardizing the spectral response to observed targets.
The generic form of normalized difference indices (NDI) was used in this work, where all combinations of wavelengths {i, j} in {350…2,500 nm} were substituted in Equation 1.

$$NDI = \frac{R_i - R_j}{R_i + R_j} \quad (1)$$

Partial least squares regression
Partial least squares (PLS) is a method that uses data compression to reduce the large number of collinear measured spectral variables, to a few orthogonal latent variables (LV's). It is employed in remote sensing for studying vegetation and soil characteristics (Atzberger *et al.*, 2010). The LVs represent the relevant structural information, which is present in the reflectance measurements to predict the dependent variable. In principle, PLS regression uses component projection successively to find latent structures. Visual inspection of score-plots and validation residual variance plots was used to find the optimal number of LVs, to prevent over-fitting. In most cases, this procedure can reduce the number of spectral variables to a few independent LVs. Validation of the models was carried out by the area under the curve (AUC) of receiver operating characteristic curve (ROC) – explained below.

Performance measures
A receiver operating characteristic curve (ROC) is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied. The ROC curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) for various threshold settings. The area under the curve (AUC) is the definite integral of ROC and is a known metric to evaluate the model quality. The performance of all the models was evaluated with the leave-one-out cross validation method (Tsenkova *et al.*, 2009).

Results and discussion

Data analyses were performed on the data from the last sampling day, 22 days after seeding, stage V2, closest to the day when the manual classification of disease level was performed. Different models were built for the P2088 variety.
It has been suggested that spectral derivatives have important advantages over spectral reflectance, such as their ability to reduce variability due to changes in illumination reflectance (Blackburn, 2007). In this work, the first derivative of the spectral reflectance was calculated, and simultaneous smoothing was applied (using the Savgol function implemented with R software; https://www.

rdocumentation.org/packages/pracma/versions/1.9.9/topics/savgol) (Savitzky and Marcel, 1964), in order to minimize noise from the environment.

The first derivative of all individual spectra was subjected to a partial least squares regression (PLS) to build a model that discriminates between non- infected and infected plants (model 1). In addition, the confidence intervals of the mean for the non- infected plants and the infected ones were calculated, to find wavelength ranges where there is a difference between the mean of the non- infected and the infected plants.

Figure 2 shows the spectral reflectance and the first derivative of the spectral reflectance of infected and non-infected plants in the entire spectral range. The graph shows that there are several wavelength ranges where there is a statistically significant difference between the mean of the non-infected and the infected plants.

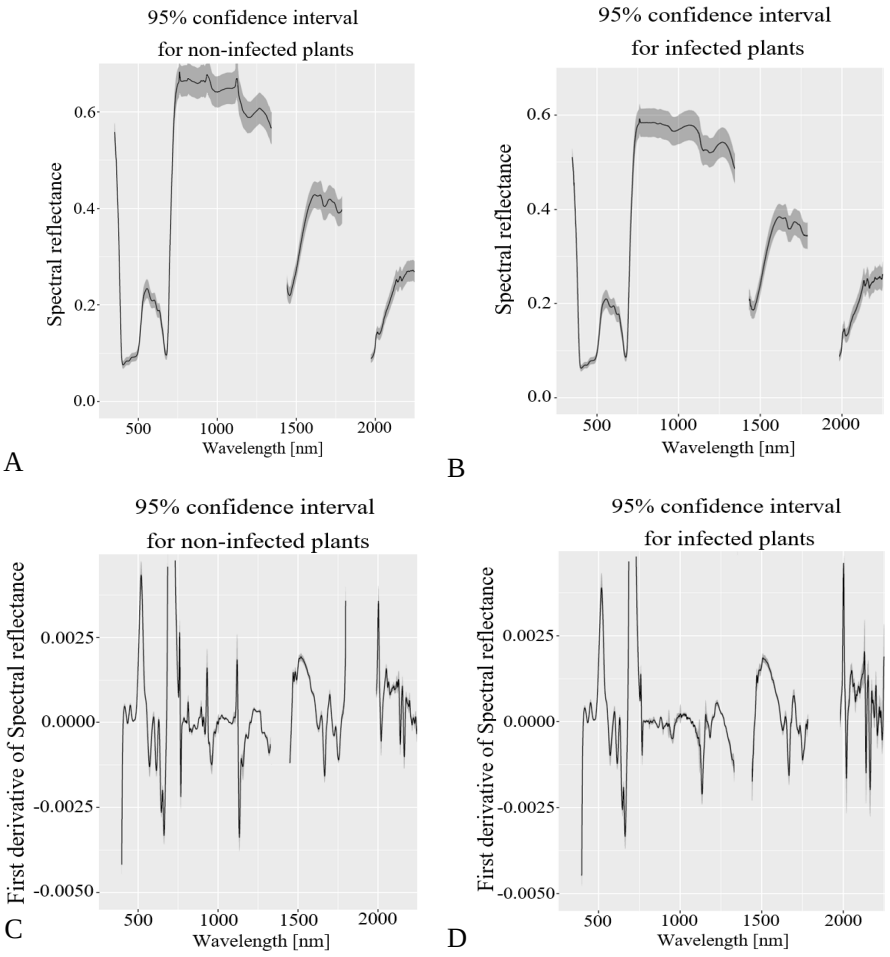


Figure 2. Mean and the 95% confidence interval of leaves' spectral reflectance of (A) non-infected plants and (B) infected plants. Mean and the 95% confidence interval of first derivative of leaf spectral reflectance of (C) non-infected and (D) of infected plants.

Although there was a significant difference in reflectance between infected and non-infected plants, this measure was not used despite its low computation complexity. Instead the derivative of reflectance was used since it is known to minimize noise and (Blackburn, 2007) and revealed more significant results.

Using the first derivative, the wavelengths with the most significant differences between non- infected and infected plants were derived using the statistical T-test. The most significant wavelengths were 715, 755, 920, 950, 975, 1,160, 1,170 and 1,280 nm. These wavelengths marginally include spectral regions sensitive to leaf [Chl] (715 nm) while all remaining wavelengths are near-infrared (NIR), sensitive to leaf parenchyma conditions.

Using the first derivative at these wavelengths, a classification model was built to separate non-infected from infected plants. The least absolute shrinkage and selection operator (Lasso) classification method was used (Tibshirani, 1996). The Lasso classifier model simultaneously conducts prediction and variable selection, selecting the most relevant wavelengths to discriminate between non- infected and infected plants (model 2). Classification results are depicted in Table 1, model 2.

An additional model was built, based on NDIs. Figure 3 shows the correlation between the normalized difference index (NDI) to the plant visually assessed level of the disease, for every two-band combination in a defined spectral range.

Table 1. Summary of classification results of the three models.

Model	Performance measure	
	AUC	Truth rate
Model 1 (PLS)	0.744	50 / 70 = 0.71
Model 2 (derivative of single wavelength)	0.740	52 / 70 = 0.74
Model 3 (NDI)	0.799	49 / 70 = 0.70

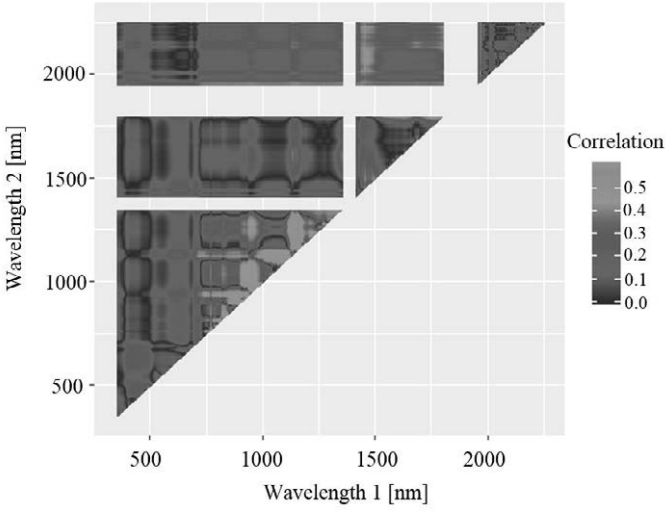


Figure 3. Correlation map of all possible two-band combinations by NDIs with the visually assessed level of the disease.

The areas with the higher correlation were for the following NDIs:

$$\frac{R_{890}-R_{900}}{R_{890}+R_{900}},\quad \frac{R_{950}-R_{1040}}{R_{950}+R_{1040}},\quad \frac{R_{1020}-R_{1120}}{R_{1020}+R_{1120}},\quad \frac{R_{950}-R_{1245}}{R_{950}+R_{1245}},\quad \frac{R_{1155}-R_{1220}}{R_{1155}+R_{1220}}.$$

Logistic regression models were examined for these indices. A logistic model is a widely used statistical model that uses a logistic function to model a binary dependent variable. The model which maximized the value of the AUC was based on the NDI with wavelengths 1,155, 1,220 nm. Results are depicted in Table 1, model 3.

Table 1 summarizes the classification results of the 3 models with AUC and accuracy depicted for model comparison. There is no significant difference between models 1, 2 and 3. The best model according to the truth rate values is model 2 followed by models 1 and 3. Model 3 yielded best results for AUC values. Furthermore, models 2 and 3 are simpler than model 1, since in model 1 all the wavelengths in the range 350-2,500 nm are included, whereas in models 2 and 3, only significant wavelengths are included.

Model 2 was able to predict the state for 74% of the plants. This result is very significant as identification was carried out at an early stage where it was not possible to classify between non-infected and infected plants without destroying the plant.

Figure 4 shows the density plots of the predicted probability for the non-infected and infected plants, for model 2. The three models show the same trend, while the prediction of the non- infected plants is consistently better than infected plant prediction. Another trend that appears in all the models are the 2 peaks in the distributions, on both sides of the threshold. This behaviour should be examined in future work.

For future work, more samples should be collected over the course of the growing stages to provide a better understanding of the spectral discrimination over time. Additionally, further research should focus on developing measures to identify at earlier stages. The point sampling technique that was used in this work cannot describe the spatial variability of the leaves spectral reflectance. Hyper-spectral images (HSI) can be useful to capture the complementary information from both domains – spatial domain and spectral domain. Utilizing a hyperspectral imaging system in order to classify vegetation can add an important data source in order to achieve these goals successfully.

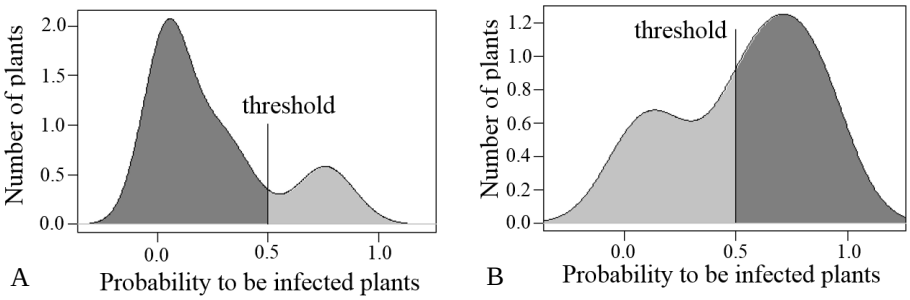


Figure 4. (A) Density plot distributions of the predicted values of the non- infected plants for model 2. The plants with predicted probability below 0.5 were correctly classified. (B) Density plot distributions of the predicted values of the infected plants for model 2.

Conclusions

To the best of the authors' knowledge, early non-destructive detection of *Fusarium* infection is still a big challenge that might have a large impact on the future agriculture economy. Therefore, achieving a classification accuracy of 74% of infected plants by applying spectral analysis is a step towards early detection of the disease.
The evaluated wavelength range 350-2,500 nm reliably detected corn plants infected with *Fusarium*. The Lasso classifier model was applied to identify distinct wavelengths that allow differentiation between spectra of infected and non-infected corn leaves: 715, 755, 920, 950, 975, 1,160, 1,170 and 1,280 nm.

Acknowledgements

This research was partially supported by the Israeli Innovation Authority, fund no. 580653293 and by W. Gunther Plaut Chair in Manufacturing Engineering, Ben-Gurion University of the Negev. We would like to thank the technical team from Evogene (Noa Guedj, Aviv Himmel, Tomer Mahler) for their help in running the experiments.

References

Atzberger, C., Guérif, M., Baret, F.& Werner, W., 2010. Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat. Computers and Electronics in Agriculture, 73 (2), 165-173.

Bauriegel, E., Giebel, A., Geyer, M., Schmidt, U., Herppich, W.B., 2011. Early Detection of *Fusarium* Infection in Wheat Using Hyper-Spectral Imaging. Computers and Electronics in Agriculture 75 (2): 304-12.

Blackburn, G. A., 2007. Hyperspectral remote sensing of plant pigments. Journal of Experimental Botany, 58 (4), 855-867.

Browne, R. A., Murphy, J. P., Cooke, B., Devaney, D., Walsh, E.J., Griffey, C.A., *et al.*, 2005. Evaluation of components of *Fusarium* head blight resistance in soft red winter wheat germ plasm using a detached leaf assay. Plant Disease, 89 (4), 404-411.

Busemeyer, L., Mentrup, D., Möller, K., Wunder, E., Alheit, K., Hahn, V., *et al.*, 2013. Breedvision – A multi-sensor platform for non-destructive field-based phenotyping in plant breeding. Sensors (Switzerland), 13 (3), 2830-2847.

Falasconi, M., Gobbi, E., Pardo, M., Della Torre, M., Bresciani, A., & Sberveglieri, G., 2005. Detection of toxigenic strains of *Fusarium* verticillioides in corn by electronic olfactory system. Sensors and Actuators, B: Chemical, 108, 250-257.

Hansen, P. M., & Schjoerring, J. K., 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sensing of Environment 86 (4): 542-53.

Jurado, M., Vázquez, C., Marín, S., Sanchis, V., & Teresa González-Jaén, M., 2006. PCR-based strategy to detect contamination with mycotoxigenic *Fusarium* species in maize. Systematic and Applied Microbiology 29 (8): 681-89.

Lancashire, P., Bleiholder, H., Van Den Boom, T., Langeluddeke, P., Stauss, R., Weber, E., *et al.*, 1991. A Uniform Decimal Code for Growth Stages of Crops and Weeds. Annals of Applied Biology 119 (3): 561-601.

Li, Lei, Zhang, Q. & Huang, D., 2014. A Review of Imaging Techniques for Plant Phenotyping. Sensors (Switzerland) 14 (11): 20078-111.

Ray, M., Ray, A., Dash, S., Mishra, A., Achary, K. G., Nayak, S., *et al.,* 2017. Fungal disease detection in plants: Traditional assays, novel diagnostic techniques and biosensors. Biosensors and Bioelectronics, 87(August 2016), 708-723.

Rouse, R. W. H., Haas, J. A. W., Schell, J. A., & Deering, D. W., 1974. Monitoring Vegetation Systems in the Great Plains With Erts. Proceedings Third ERTS-1 Symposium, 351.

Savary, S., Ficke, A., Aubertot, J. N., & Hollier, C., 2012. Crop losses due to diseases and their implications for global food production losses and food security. Food Security, 4 (4), 519-537.

Savitzky, A., & Marcel, J. E. G., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry, 36 (8), 1627-1639.

Tibshirani, R., 1996. Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58 (1), 267-288.

Tsenkova, R., Meilina, H., Kuroki, S., & Burns, D. H., 2009. Near infrared spectroscopy using short wavelengths and leave-one-cow-out cross-validation for quantification of somatic cells in milk. Journal of Near Infrared Spectroscopy, 17 (6), 345-351.

West, J. S., Bravo, C., Oberti, R., Lemaire, D., Moshou, D., & McCartney, H. A., 2003. The potential of optical canopy measurement for targeted control of field crop diseases. Annual Review of Phytopathology, 41 (1), 593-614.

Windels, C. E., 2000. Economic and Social Impacts of *Fusarium* Head Blight: Changing Farms and Rural Communities in the Northern Great Plains. Phytopathology, 90 (1), 17-21.

Zhangyan, J., Alfredo, R., Jin, C., Yunhao, C., Jing, L., Guangjian, Y., *et al.,* 2006. Analysis of NDVI and Scaled Difference Vegetation Index Retrievals of Vegetation Fraction. Remote Sensing of Environment 101 (3): 366-78.

Proximal versus remote sensing to monitor pasture quality in a Mediterranean Montado ecosystem

J. Serrano[1], S. Shahidian[1], F. Moral[2] and J. Marques da Silva[1,3*]
[1]*ICAAM, Departamento de Engenharia Rural, Universidade de Évora, P.O. Box 94, Évora 7002-554, Portugal; jmrs@uevora.pt*
[2]*Departamento de Expresión Gráfica, Escuela de Ingenierías Industriales, Universidad de Extremadura. Avenida de Elvas s/n, 06006 Badajoz, Spain*
[3]*Agroinsider Lda. (spin-off da Universidade de Évora), PITE, R. Circular Norte, NERE, Sala 18, 7005-841 Évora, Portugal*

Abstract

Montado is an agro-forestry system that covers substantial areas in countries of the Mediterranean region. In this system, the natural dryland pasture is the principal source of animal feed in extensive grazing. The climatic seasonality associated with the inter-annual irregularity of precipitation greatly determines the development of the pasture vegetative cycle. As the spring reaches its end, there is a notable reduction in the nutritive value of the plants, and a critical period begins. The objective of this work was to evaluate, through the relationship between pasture quality indices (pasture quality degradation index, PQDI and normalized difference vegetation index, NDVI), two technological approaches to monitor the evolution of the quality of a biodiverse pasture in the period of greatest vegetative development (between February and June): (1) through proximal sensing (PS), with the use of an active optical sensor; (2) through remote sensing (RS), using images captured by the Sentinel-2 satellite. The results of this study show strong and significant relationships between PQDI and NDVI (obtained by PS or RS). These two techniques (PS or RS) can, therefore, be used in a complementary way to identify and anticipate the animal supplementation needs and support the farmer's decision making process.

Keywords: quality degradation index, NDVI, Sentinel-2, optical active sensor, vegetative cycle

Introduction

Montado is the most extensive European agro-forestry system, as it covers 3.5-4.0 Mha in Portugal and Spain (Seddaiu *et al.*, 2013). The most important premise for increased livestock productivity in this ecosystem is the improvement of pastures in terms of quality (Aguilar *et al.*, 2018). These biodiverse pastures are characterized by marked seasonal variation in productivity and quality, which reflects the seasonality of the Mediterranean climate, caused by inter-annual changes in precipitation (Lumbierres *et al.*, 2017; Aguilar *et al.*, 2018). From autumn to spring, pasture is the main source of animal feed in extensive grazing. However, starting in late spring or early summer, low precipitation and high temperatures accelerate the final stages in the pasture vegetative cycle, clearly decreasing its nutritional value (Efe Serrano, 2006). In this critical period, farmers need essential data on the spatial and temporal variability of pasture quality in order to safeguard a minimum supply of high quality feed throughout the year (Pullanagari *et al.*, 2018) and evaluate possible feeding supplementation strategies for the animals, to at least cover their maintenance needs. Pasture quality is assessed by indicators, such as crude protein (CP) and neutral detergent fibre (NDF) concentration (Lugassi *et al.*, 2015). Throughout the vegetative cycle of the pasture, the CP contents tend to decrease while the NDF contents tend to increase. Serrano *et al.* (2018) presented an index designated for pasture quality degradation (PQDI), which uses the NDF/CP ratio to quantify the evolution of the natural degradation of pasture quality throughout its vegetative cycle.

DOI 10.3920/978-90-8686-888-9_43, © Wageningen Academic Publishers 2019

The most widely accepted and accurate method for assessing these indicators of pasture quality (CP and NDF) is based on in-situ chemical analysis. However, this conventional measurement involves intensive field work and destructive laboratory methods, which are relatively expensive and time-consuming, inapplicable at farm scale (Lugassi *et al.*, 2015; Louhaichi *et al.*, 2017; Lumbierres *et al.*, 2017; Pullanagari *et al.*, 2018).
The availability of efficient and accurate techniques and technologies that allow the mapping and understanding of this variability (Zhang *et al.*, 2016) is fundamental.
With recent advancements in precision agriculture (PA), information technologies, remote sensing (RS) and proximal sensing (PS) and geospatial analyses supported by global navigation satellite system (GNSS), fast and non-destructive methods are becoming available for mapping the spatio-temporal changes of vegetation characteristics with high spatial resolution (Moeckel *et al.*, 2017).
Normalized difference vegetation index (NDVI) is a vegetation index (Rouse *et al.*, 1973) derived from the normalized difference between near infrared and visible red reflectance, corresponding respectively to the strong chlorophyll absorption region and high reflectance plateau of vegetative canopies (Chai *et al.*, 2015). This index can be obtained by RS or PS. Types of RS technology are categorized by platform, including space-borne (satellites) or airborne (aircraft and unmanned aerial vehicles, UAV). 'OptRx' is an example of a commercial optical crop proximal sensor which allows the measurement of reflectance for calculation of NDVI (Serrano *et al.*, 2018).
Applications of satellite observations are, however, often limited due to the coarse spatial resolution (Hassan-Esfahani *et al.*, 2017) associated with poor quality images on cloudy days (Wang *et al.*, 2013). In the Montado ecosystems, satellite, aircraft or UAV images have the disadvantage of not being able to access the pasture under the trees, which highlights the importance of proximal sensors as a solution for monitoring pasture under these conditions (Handcock *et al.*, 2016; Louhaichi *et al.*, 2017).
Advances in non-invasive technologies, based on proximal sensors mounted on mobile platforms, allow the collection of samples at a high spatial resolution. This enables the exploration of the spatial variability of pastoral ecosystems at fine scale (Nawar *et al.*, 2017), which complements RS (Aguilar *et al.*, 2018) and, thus, makes it possible to evaluate, under and outside tree canopy the variables that determine the equilibrium and sustainability of this ecosystem (Nawar *et al.*, 2017), supporting grazing management decisions (Moeckel *et al.*, 2017).
The objective of this work was to evaluate, through the relationship between pasture quality indices (PQDI and NDVI), two technological approaches to monitor the evolution of the quality of a biodiverse pasture during the period of greatest vegetative development (between February and June 2017): (1) through PS, with the use of an active optical sensor (AOS, 'OptRx'); (2) through RS, using images captured by Sentinel-2 satellite system.

Materials and methods

Study area

The study was conducted on a paddock, with an area of 2.3 ha, located at the Mitra farm (co-ordinates 38°32.2' N; 8° 01.1' W), Évora University, Southern Portugal. The field of oak trees (Q. ilex ssp. rotundifolia Lam.), with a relatively reduced tree density (approximately 10 trees/ha), has had an understory of permanent biodiverse pasture (legumes, grasses and other species) grazed by sheep in a rotational system. The predominant soil of this field is classified as a Cambisol derived from granite (FAO, 2006).

Characterization of the climate

The Mediterranean climate can be considered as a transition between temperate and dry subtropical climates. It is characterized by summer drought, variable rainfall and mild or moderately cold winters. The monthly average temperature is between 8 °C and 26 °C; minimum temperatures are close to 0 °C between December and February. The annual rainfall in the region is between 400

and 600 mm; rainfall occurs mainly between October and March and is practically non-existent during the summer.

Pasture monitoring

The evolution of the pasture in the study area was recorded in 12 sampling points, areas without trees corresponding to 10×10 m pixel of Sentinel-2 imagery. The vertices of these regions of interest (ROI) were materialized on the ground with wooden stakes positioned with a real time kinematic (RTK) Global Positioning System (GPS) instrument (Trimble RTK/PP-4700 GPS, Trimble Navigation Limited, Sunnyvale, CA, USA).

Monitoring dates (of 2017) coinciding with the dates of the satellites passing and without cloud cover: at the end of winter (February 14 and March 16), and then monthly during spring (April 15, May 25, and June 14).

Three locations representative of pasture variability were identified previously in each ROI and in each monitoring date. In each of these locations and immediately before cutting each pasture sample, NDVI was measured with a commercial 'OptRx™' active optical proximal sensor (Ag Leader, Ames, Iowa 50010, USA). A GNSS receiver with sub-meter precision (Trimble GeoExplorer 6000 series model 88951, Raunheim, Germany) provided the coordinates of each measurement done by the sensor. The operator stood still holding the sensor at 0.75 m above ground and, in each location, performed tree measurements. Each complete measurement with the PS ran for about one minute, with one second recording interval. In total, nine complete measurements were carried out within each ROI. After each complete PS measurement, another operator placed a frame with a 0.25 m^2 area (dimensions 0.50×0.50 m) and pasture was harvested with portable electric grass shears at 10-20 mm above ground level and stored in marked plastic bags. The collected composite pasture samples (one for each ROI, resulting of nine sub-samples) were weighed to establish pasture productivity in terms of biomass (kg/ha). The dehydrated samples (72 h at 65 °C), were analysed in order to determine the key components of pasture quality (content of crude protein, CP, and neutral detergent fibre, NDF, both in % of dry matter) using standard methods.

Reflectance was also measured by remote sensing (Sentinel-2). $NDVI_{SAT}$ data were extracted in the 12 geo-referenced pixels (ROI) where the measurements with the active optical sensor were performed. To generate the NDVI values the surface reflectance bands 4 (B4; 665 nm) and 8 (B8; 842 nm) of Sentinel-2 spectral data were used. The 'Sen2Cor 2.3' processor (implemented on Sentinel Application Platform from European Space Agency) was used for atmospheric correction.

Statistical analysis of the data

Descriptive statistics analysis, including mean, standard deviation (SD), coefficient of variation (CV) and range, were determined for each pasture data set (CP, NDF) and sensor parameters (NDVI, measured by PS and by satellite).

The statistical treatment of the results was performed using 'MSTAT-C' software with a significance level of 95% ($P<0.05$) and consisted of analysis of regression between pasture parameters (CP, NDF) and sensor variables (NDVI, measured by PS and by satellite).

Results

Table 1 summarizes the results of descriptive statistics (mean, SD, CV and range) of pasture parameters, measured at the 12 sampling points of the field, for each monthly sampling date. The spatial variability of CP and NDF was particularly evident in February (CV of the order of 25%), and decreased throughout the vegetative cycle of the pasture, accompanying the decrease in pasture quality.

Figure 1-3 show the relationship between NDVI (measured both by PS and by satellite images) and CP, NDF and PQDI. All these regression analysis were significant.

Table 1. Descriptive statistics of pasture parameters, measured at the 12 sampling points of the field, for each monthly sampling date.[1]

Pasture parameters	Mean±SD	CV (%)	Range
CP, %			
February	14.2±3.5	24.6	[8.6-18.7]
March	14.8±3.2	21.6	[10.2-19.8]
April	7.4±1.3	17.8	[5.6-10.3]
May	7.7±1.2	16	[5.5-9.8]
June	5.4±0.8	14.5	[4.4-6.9]
NDF, %			
February	51.2±12.9	25.2	[35.0-77.5]
March	49.6±7.8	15.7	[36.9-64.6]
April	60.4±5.1	8.4	[54.0-70.2]
May	72.7±3.7	5.1	[65.6-79.6]
June	78.1±2.6	3.4	[74.2-83.9]
PQDI			
February	3.6±2.1	51.8	[2.0-9.0]
March	3.4±1.2	33.6	[2.3-6.4]
April	8.2±1.9	22.9	[5.4-12.5]
May	9.4±2.0	20.3	[6.9-14.4]
June	14.5±2.5	17.1	[11.1-18.8]
$NDVI_{PS}$			
February	0.748±0.109	14.5	[0.563-0.860]
March	0.813±0.053	6.5	[0.680-0.890]
April	0.554±0.065	11.7	[0.450-0.650]
May	0.314±0.077	24.5	[0.240-0.450]
June	0.173±0.011	6.1	[0.160-0.190]
$NDVI_{SAT}$			
February	0.576±0.059	10.2	[0.475-0.649]
March	0.640±0.039	6.1	[0.560-0.692]
April	0.596±0.051	8.5	[0.516-0.686]
May	0.340±0.052	15.4	[0.264-0.444]
June	0.261±0.048	18.3	[0.201-0.345]

[1] CP = crude protein; CV = coefficient of variation; NDF = neutral detergent fibre; $NDVI_{PS}$ = normalized difference vegetation index measured by optical active sensor; $NDVI_{SAT}$ = normalized difference vegetation index measured by satellite; PQDI = pasture quality degradation index; SD = standard deviation.

Discussion

The regression analysis between NDVI, measured by PS or RS, and pasture quality parameters (CP and NDF, Figure 1 and 2, respectively) showed significant coefficients of determination (R^2). The values of R^2 (between 0.50 and 0.75 for CP and between 0.76 and 0.83 for NDF) were of the order of magnitude of those obtained by Serrano *et al.* (2018) with the PS (0.69 and 0.78, respectively for CP and NDF) and of those obtained by Zhao *et al.* (2007) (0.61 and 0.58, respectively for CP and NDF) with reflectance data obtained by RS. Albayrak (2008) and Pullanagari *et al.* (2018) also

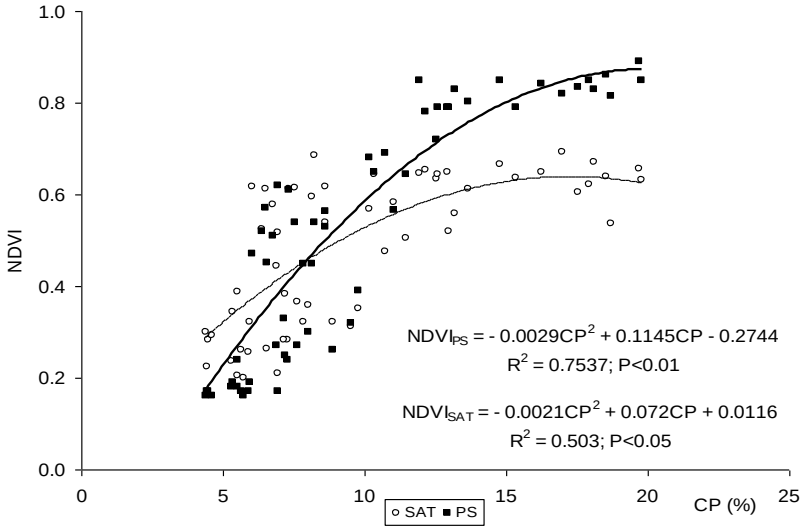


Figure 1. Relationship between NDVI (obtained by proximal sensor, $NDVI_{PS}$ and obtained by satellite, $NDVI_{SAT}$) and CP, in the set of monthly sampling dates.

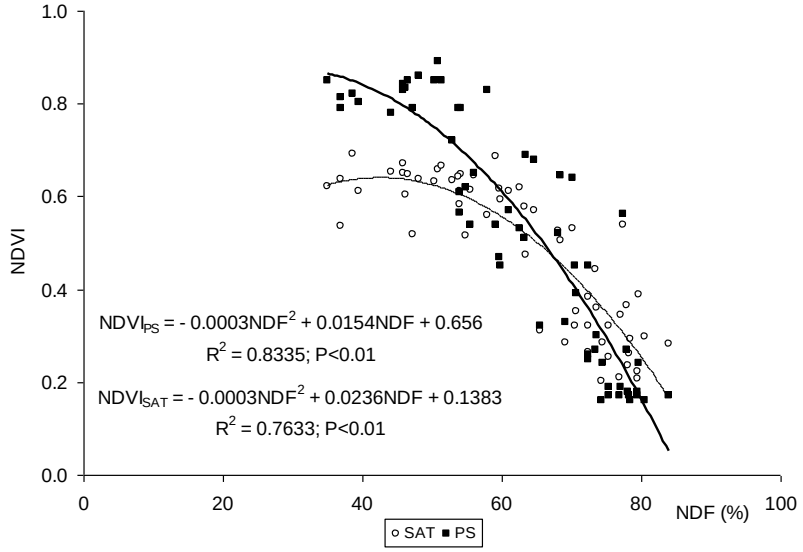


Figure 2. Relationship between NDVI (obtained by PS, $NDVI_{PS}$ and obtained by satellite, $NDVI_{SAT}$) and NDF, in the set of monthly sampling dates.

found satisfactory relationships between spectral measurements and pasture quality parameters such as CP content, which can be attributed to absorbance of visible radiance by chlorophyll that is abundant in green vegetation.

The regression analysis between NDVI, measured by PS or RS, and pasture quality degradation index (PQDI, Figure 3) showed also significant R^2.

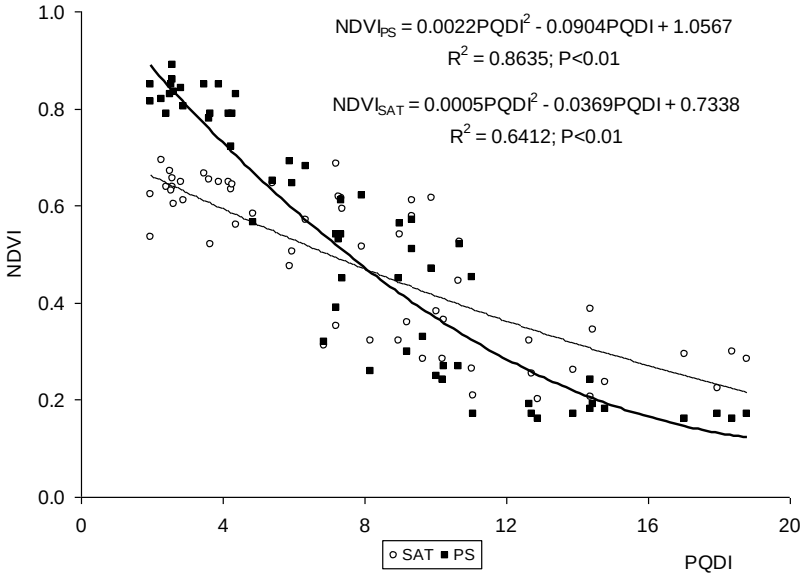


Figure 3. Relationship between NDVI values (obtained by proximal sensor, $NDVI_{PS}$ and obtained by satellite, $NDVI_{SAT}$) and pasture quality degradation index (PQDI), in the set of monthly sampling dates.

Figure 1, 2 and 3 show that the coefficients of determination between NDVI and pasture quality parameters considered (CP, NDF and PQDI) were higher for PS than RS (by satellite). These results may reflect the sampling process: while $NDVI_{SAT}$ involved a sampling area of 100 m^2 (pixel area), $NDVI_{PS}$ was obtained from PS measurements carried out directly on the pasture that was to be cut and subjected to analytical determinations (an area of 2.25 m^2), a factor particularly important in biodiverse pastures (Lugassi *et al.*, 2015). The high spatial and temporal heterogeneity of grassland and pastures, due to the diversity of species, natural environmental conditions and management practices (Lugassi *et al.*, 2015), suggest that future studies of calibration between PS and RS should consider more complete elementary sampling unit (ESU). More extensive pasture sampling of the ROI, in this case the Sentinel-2 pixel area, as suggested by Baret *et al.* (2003), will be the way to capture the spatial variability associated with this ecosystem.

The synoptic and repetitive coverage provided by RS images, along with ecological indicators and vegetation biophysical parameters, allows the application of integrated data analysis and the development of predictive models useful for decision makers (Aguilar *et al.*, 2018). The seasonality of pasture production and the typical inter-annual variability of the Mediterranean climate accentuate the importance of producing monitoring and management systems and even alarm systems for farmers based on the measurement of NDVI from satellite images (RS), with a very short turn-around time across large areas. Proximal sensing despite presenting the disadvantage of requiring lengthy in situ surveying, can be used in a complementary way, helping to overcome the main difficulties associated with satellite imagery, in particular those related to resolution, access under tree canopy in Montado ecosystem and, especially, with cloud cover.

Conclusions

The results of this exploratory study in pastures in Portugal show strong and significant coefficients of determination between PQDI and NDVI, obtained by PS or RS. Although RS, through satellite imagery, presents several operational advantages over PS, such as the scale of the response, speed of processing and low cost, frequent spring cloudiness may be an obstacle to regular determination of VI, which can lead to longer periods of time without information. This obstacle, together with the difficulty of evaluating VI under tree canopy in Montado ecosystem, justifies the interest in using these two techniques in a complementary way as information and alarm systems to support soil, pasture and grazing management.

Acknowledgements

This work was funded by National Funds through FCT (Foundation for Science and Technology) under the Project UID/AGR/00115/2019) and by the projects PDR2020-101-030693 and PDR2020-101-031244 ('Programa 1.0.1-Grupos Operacionais').

References

Aguilar, A. L., Flores, H., Crespo, G., Marín, M. I., Campos, I. and Calera, A., 2018. Performance assessment of MOD16 in Evapotranspiration evaluation in Northwestern Mexico. Water 10 901.

Albayrak, S., 2008. Use of reflectance measurements for the detection of N, P, K, ADF and NDF contents in Sainfoin pasture. Sensors 8 7275-7286.

Baret, F., Weiss, M., Leroy, M., Allard, D., Garrigues, S., Bohbot, H., *et al.,* 2003. VALERI: a network of sites and a methodology for the validation of land satellite products. Available online: http://w3.avignon.inra.fr/valeri/documents (accessed on 04 February 2019).

Chai, X., Zhang, T., Shao, Y., Gong, H., Liu, L. and Xie, K., 2015. Modeling and mapping soil moisture of plateau pasture using RADARSAT-2 imagery. Remote Sensing 7, 1279-1299.

Efe Serrano, J., 2006. Pastures in Alentejo: Technical Basis for Characterization, Grazing and Improvement. Universidade de Évora, ICAM, Évora, Portugal, 165-178.

FAO, 2006. World Reference Base for Soil Resources. Food and Agriculture Organization of the United Nations, World Soil Resources Reports no. 103. Rome, Italy: IUSS Working Group WRB.

Handcock, R. N., Gobbett, D. L., González, L. A., Bishop-Hurley, G. J. and McGavin, S.L., 2016. A pilot project combining multispectral proximal sensors and digital cameras for monitoring tropical pastures. Biogeosciences 13 4673-4695.

Hassan-Esfahani, L., Ebtehaj, A.M., Torres-Rua, A. and McKee, M., 2017. Spatial scale gap filling using an unmanned aerial system: A statistical downscaling method for applications in precision agriculture. Sensors 17 2106.

Louhaichi, M., Hassan, S., Clifton, K. and Johnson, D. E., 2017. A reliable and non-destructive method for estimating forage shrub cover and biomass in arid environments using digital vegetation charting technique. Agroforestry Systems 92 1341-1352.

Lugassi, R., Chudnovsky, A., Zaady, E., Dvash, L. and Goldshleger, N., 2015. Spectral slope as an indicator of pasture quality. Remote Sensing 7 256-274.

Lumbierres, M., Méndez, P. F., Bustamante, J., Soriguer, R. and Santamaria, L., 2017. Modeling biomass production in seasonal wetlands using Modis NDVI land surface phenology. Remote Sensing 9 392.

Moeckel, T., Safari, H., Reddersen, B., Fricke, T. and Wachendorf, M., 2017. Fusion of ultrasonic and spectral sensor data for improving the estimation of biomass in grassland with heterogeneous sward structure. Remote Sensing 9 98.

Nawar, S., Corstanje, R., Halcro, G., Mulla, D. and Mouazen, A. M., 2017. Delineation of soil management zones for variable-rate fertilization: A review. Advances in Agronomy 143 175-245.

Pullanagari, R.R., Kereszturi, G. and Yule, I. 2018. Integrating airborne hyperspectral, topographic, and soil data for estimating pasture quality using recursive feature elimination with random forest regression. Remote Sensing 10 1117.

Rouse, J.W., Haas, R.H., Schell, J.A. and Deering, D. W., 1973. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. Progress Report RSC 1978-1.

Seddaiu, G., Porcua, G., Ledda, L., Roggero, P., Agnelli, A. and Cortic, G., 2013. Soil organic matter content and composition as influenced by soil management in a semi-arid Mediterranean agro-silvo-pastoral system. Agriculture, Ecosystems and Environment 167 1-11.

Serrano, J., Shahidian, S., Marques da Silva, J. and Carvalho, M., 2018. A holistic approach to the evaluation of the montado ecosystem using proximal sensors. Sensors 18 570.

Wang, X., Linlin, G. and Li, X., 2013. Pasture monitoring using SAR with COSMO-SkyMed, ENVISAT ASAR, and ALOS PALSAR in Otway, Australia. Remote Sensing 5 3611-3636.

Zhang, Y., Xiao, Y., Zhuang, Z., Zhou, L., Liu, F. and He, Y., 2016. Development of a near ground remote sensing system. Sensors 16 648.

Zhao, D., Starks, P., Brown, M., Phillips, W. and Coleman, S., 2007. Assessment of forage biomass and quality parameters of bermudagrass using proximal sensing of pasture canopy reflectance. Grassland Science 53 39-49.

Section 6
Applications of unmanned aerial systems

Investigation of spraying efficiency of an aerial spraying system in a super-high density olive grove in Greece

F. Aru[1*], A. Gertsis[2], G. Vellidis[3] and F. Morari[1]
[1]*Department of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova, Via dell'Università 16, Legnaro, PD, Italy; federica.aru@studenti.unipd.it*
[2]*Department of Agro-Environmental Systems Management, Perrotis College/American Farm School, Precision Agriculture Laboratory, 54 Marinou Antypa St, Thessaloniki, Greece*
[3]*Department of Crop and Soil Sciences, University of Georgia, 2360 Rainwater Rd, Tifton, GA, USA*

Abstract

Non-optimal and untimely plant protection products (PPP) applications adversely affect the environment. Unmanned aerial vehicles (UAVs), recently introduced in super-high density (SHD) olive groves, increase sustainability of chemical treatments by reducing sprayed volume. With the general aim to use UAVs for precision pesticide application, spraying performances of a UAV were evaluated and compared with two traditional ground systems, a backpack sprayer and a mist-blower sprayer. Coverage and quality of spray applications were assessed by water sensitive paper (WSP), using a software (DepositScan) and a smartphone application (SnapCard). Performances were higher using the mist-blower sprayer, while UAV showed a more uniform distribution with a low coverage. Further improvements include manipulating UAV speed and using different type of nozzles. UAVs offer interesting opportunities to increase PPP application efficiency, as was found in this study.

Keywords: environment, plant protection products, precision pesticide application, super-high density olive groves, unmanned aerial vehicles

Introduction

The olive tree (*Olea europaea* L.) is an evergreen and long-lived species, cultivated worldwide over the centuries for its edible fruits (Russo *et al.*, 2016). Due to tree characteristics and depending on the cultivation method, agrochemicals are frequently applied uniformly over the whole orchard, without regard to the spatial variability of the target tree profile and using higher rates than needed. This phenomenon increases spray losses associated with deposition on the ground, off-target application and drift (Pérez-Ruiz *et al.*, 2011) and in turn decreases application efficiency. One of the most important off-target application phenomena is pesticide runoff which is caused either by overdosing, as a consequence of not having an appropriate dosing system or by performing low uniformity treatments as a result of the inadequate use and poor maintenance of application equipment (Miranda-Fuentes *et al.*, 2016). Olive farmers often apply products to the point of runoff to guarantee a high biological efficacy (Miranda-Fuentes *et al.*, 2015) but they contribute to environmental pollution. Main issues regarding plant protection quality in olives are the uncertainty about optimal application parameters and about the influence of operational parameters on the application homogeneity and efficacy (Miranda-Fuentes *et al.*, 2016). Moreover, olive grove characteristics and the improper spraying equipment, especially in new systems such as super-high density olive groves, can negatively influence the application efficiency (Miranda-Fuentes *et al.*, 2015). Unmanned aerial vehicles are precision agriculture solutions for developing spraying processes that are probably safer and more precise than the use of manned agricultural aircraft and traditional ground sprayers. Compared to ground technology, aerial systems allow faster spraying without the need for paths inside the crop field (Faiçal *et al.*, 2017), even in sloping groves. UAVs can be beneficial for operator's health by reducing the risk of human contact with chemicals; for

DOI 10.3920/978-90-8686-888-9_44, © Wageningen Academic Publishers 2019

improving spraying operation performance and avoiding the chemical drift outside designated areas (off-target application), which helps to preserve neighbouring fields; for protecting natural areas and water sources (Faiçal *et al.*, 2014).
In order to evaluate UAVs suitability for precision pesticide applications, it is important to measure spraying performance indices such as spraying range, deposition and droplets distribution. Moreover, height and angle of spraying, velocity of operation and plant shape should be considered (Zheng *et al.*, 2017). Evaluation of spraying quality and distribution indices by mean of portable devices can be cheaper and less time-consuming than conventional methods (Nansen *et al.*, 2015).
This study aimed to assess spraying efficiency and quality of an aerial spraying system (UAV) in a SHD olive production system (1,670 trees/ha) and to compare it with spraying technologies commonly used for plant protection purposes in olive groves. A novel smartphone application was tested as a simple and rapid method for spraying efficiency evaluation for future precision pesticide application tasks.

Materials and methods

Study site and experimental layout
The experimental area was a long-term super-high density olive grove adapted for mechanical harvesting, planted in late 2010, at Perrotis College/American Farm School, in Thessaloniki, Greece (Lat 40° 34' 13" N Long 22° 59' 12" E). The grove is composed of 12 rows, with between-rows and between-trees spacing equal to 4 and 1.5 m, respectively. The second out of four rows of the super-high density section (1,670 trees/ha) was selected to be totally sprayed under real field conditions by using two ground and one aerial plant protection products distribution systems. The SHD system was chosen because of its geometric regularity. This condition facilitates the study of spraying deposit homogeneity inside tree crowns, as shown by Miranda-Fuentes *et al.* (2015). From May to July, three trials using tap water were carried out, each one included four spraying events by two conventional ground systems and a UAV technology. All variables influencing the spraying (i.e. wind speed and direction, relative humidity, temperature, rain in the previous 24 h and leaf wetness) have been measured before and during each spraying event by using a weather station installed in the experimental field. The three trials were conducted under acceptable weather conditions, i.e. no extreme temperatures, low wind and medium-high relative humidity. Volume of water sprayed and experiment duration of every spraying event were measured.

Spraying techniques
Two ground and one aerial spraying systems were used in this study. Ground systems were a 16-l backpack sprayer and a mist-blower sprayer pulled by a tractor. The backpack had a double conical mist nozzle and it operated with a flow rate of 1.65 l/min and a pressure of 0.18 MPa. It was carried by a technician who walked down the trees line, on both sides, performing the spraying as conventionally and being careful in covering all canopies.
The mist-blower sprayer had a speed of 5.3 km/h, a turbine working at 1,500 rpm and 10 out of 14 – Albuz ATR 80 red – open nozzles. Working pressure, spraying angle and flow rate were equal to 0.7 MPa, 80° and 1.57 l/min, respectively. It was pulled by a tractor that proceeded both on the right and on the left sides, along the tree line.
A DJI AGRAS MG-1 octocopter spraying UAV – 10 l tank capacity – was used. Flight speed and height over the canopy were equal to 1 m/s and 1.5 m, respectively. Nozzles model was XR11001 orange, being flow rate, spraying angle and working pressure equal to 1.6 l/min, 110°, and 60 psi (0.4 MPa), respectively. It flew according to two trajectories: on both sides of the tree line (the same trajectory followed by technician and tractor) and above the line.

Droplets sampling

Four trees were randomly selected for coverage and spraying quality assessment by using Syngenta water sensitive paper (WSP) (26×76 mm) (Spraying Systems Inc., Wheaton, IL, USA), distributed by TeeJet® Technologies (Wheaton, IL 60187 USA). It consists of rigid, yellow paper, coated with a layer of bromoethyl blue that turns dark blue when aqueous drops contact it (Hoffman and Hewitt, 2005). To homogenize replicates, take in consideration differences among tree canopies and assign a specific number of WSPs to each tree, tree standardization was performed. Olive tree canopy volume was calculated by using the truncated inverted cone volume equation, being upper and lower diameters of the canopy and its height manually measured. It is critical to distribute WSPs evenly throughout each target tree canopy, being careful to place them at the top, centre and bottom of the canopy as well as laterally and internally from canopy outer edge. For that purpose, according to canopy volume, 24 WSPs were used for each spraying event in the first trial, while in second and third trials they were increased to 26 due to canopy growth. Branches, where WSPs were placed, were marked to ensure to put them in the same position at each spraying event. To avoid cross contamination between treatments, WSPs were placed just prior to being sprayed and removed after they dried. Dried WSPs were analysed by using DepositScan software and SnapCard smartphone application. DepositScan is a scanning program that can quickly evaluate spray deposit distribution on water sensitive paper and other kinds of collectors that can show visual differences between spray deposits and the background. The program consists of a set of custom plugins that are used by an image-processing program (ImageJ) to produce a number of measurements useful for expressing spray deposit distribution (Zhu *et al.*, 2011). DepositScan worked with an office multifunctional printer-scanner-copier, using a scanning resolution of 600 dpi as justified by Marçal and Cunha (2008). The program reports coverage percentage (SC (%)), deposits/cm^2 and volume diameters – $DV_{0.1}$, $DV_{0.5}$ and $DV_{0.9}$ – (μm) (Zhu *et al.*, 2011). SnapCard smartphone application is the outcome of a research collaboration between the entomology group at the Department of Agriculture and Food, Western Australia, and the University of Western Australia's (UWA) applied entomology program. The user takes a photo of each WSP by the same smartphone in which the application is installed and only the area of interest, i.e. WSP, is analysed. SnapCard calculates the percentage of blue out of the total number of pixels within the analysed area, giving as output the coverage percentage (SC (%)) (Nansen *et al.*, 2015).

Data analysis

Effects of spraying techniques on WSP were analysed by JMP12 software (SAS Institute, www.jmp.com) using one-way analysis of variance (ANOVA) with α=0.05. In addition, overspray index (I_o), coverage homogeneity index (HSC (%)), number of impacts homogeneity index (HNi (%)) and relative span factor (RSF) were calculated.

I_o quantifies the magnitude of overspray to compare spray deposition qualities of the spraying systems. Values range from 0 – no overspray (30% coverage) to 1 – saturated spray (100% coverage). It was determined for WSPs with coverage percentage greater than 30%, according to Chen *et al.* (2013). I_o calculation was performed by normalizing the spray coverage according to the following equation:

$$\mathrm{Io} = \frac{(SC-30)}{(100-30)} \tag{1}$$

where SC (%) is the spray coverage on WSP determined by DepositScan and SnapCard.

HSC (%) and HNi (%) range between 0 and 100, with higher values indicating higher mean homogeneity (Miranda-Fuentes *et al.*, 2017). Their calculation was performed as follows:

$$HSC = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{SCmax\ i - SCmin\ i}{SCmax\ i}\right) * 100 \tag{2}$$

$$HNi = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{HNimax\ i - HNimin\ i}{HNimax\ i}\right) * 100 \quad (3)$$

where N is the number of trees per treatment (equal to 4), SC_{max} and Ni_{max} are the maximum values of coverage and deposits/cm^2 per tree, SC_{min} and Ni_{min} are the minimum values of coverage and deposits/cm^2 per tree. HSC (%) was calculated by using DepositScan and SnapCard data, while HNi (%) by using DepositScan data. RSF is an index of uniformity of drop size distribution, ranging between 0 and 1 (the closer its value to 1 the more uniform is the spray) (Patel *et al.*, 2015). It was calculated for all spraying systems and trials according to Equation 4, where $DV_{0.9}$, $DV_{0.1}$ and $DV_{0.5}$ are volume diameters given by DepositScan.

$$RSF = \frac{(DV0.9 - DV0.1)}{DV0.5} \quad (4)$$

Volume diameters refer to droplet size, $DV_{0.9}$ value indicates that 90% of the volume of spray is in droplets smaller (or 10% larger) than this value. The same for $DV_{0.5}$ and $DV_{0.1}$ where volume is 50 and 10%, respectively.

Results and discussion

Coverage (SC (%)) and coverage homogeneity index (HSC (%))

The bivariate fit in Figure 1 shows a strong correlation (r=0.978) between coverage (SC (%)) determined by DepositScan (D) and SnapCard (S). In spite of the strong relationship, the latter clearly underestimates DepositScan SC (SC-D). The mist-blower sprayer showed a SC-D close to the optimal value of 30% (36.5%), while lower values were observed for backpack sprayer and UAV (Table 1). Conversely, homogeneity (HSC-D) was >90% for backpack sprayer and UAV flying over-the-line whereas smaller values were observed for the mist-blower sprayer, and UAV flying in parallel (Table 1). Lower HSC values were calculated when using SnapCard data.

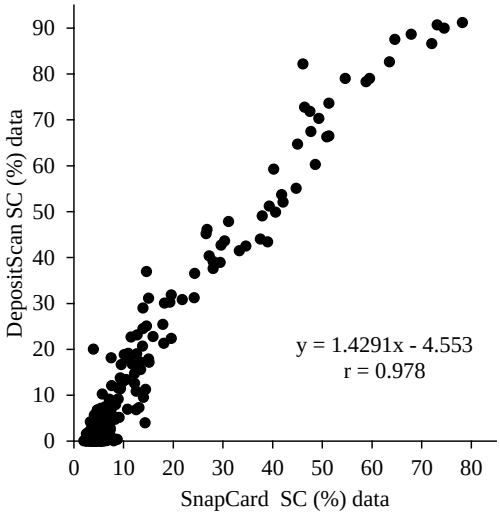


Figure 1. Comparison between coverage (%) data obtained by DepositScan and SnapCard. A strong correlation (r=0.978) exists between the two data sets.

Table 1. Coverage-SC (%), coverage homogeneity index-HSC (%) and overspray index-I_o calculated for DepositScan (D) and SnapCard (S) data. Letters indicate significant differences at α=0.05.

	SC (%) D	SC (%) S	HSC (%) D	HSC (%) S	I_o D	I_o S
Mist-blower sprayer	36.53 a	26.42 a	88.46 a	85.83 a	0.43 a	0.31 a
UAV over	0.81 c	4.71 c	91.58 a	53.89 c	/	/
Backpack	10.72 b	10.97 b	94.86 a	76.17 b	0.19 b	0.14 b
UAV parallel	0.32 c	4.39 c	84.38 a	53.93 c	/	/

Overspray index (I_o)

I_o was calculated for WSPs that showed a SC higher than 30%, emphasizing conditions of overspray (I_o>0). Therefore, only treatments with backpack and mist-blower sprayer were analysed, including 16 and 11% of DepositScan and SnapCard data, respectively. Worse performances were observed in mist-blower sprayer, both considering I_o-D (0.43) and I_o-S (0.31). Conversely, backpack sprayer index displayed values closer to the optimum, being in both cases <0.2 (Table 1).

Deposits/cm^2 and number of impacts homogeneity index (HNi (%))

Deposits were determined by DepositScan. Higher deposit amounts were detected when the mist-blower sprayer performance was assessed while low values were obtained for the spraying UAV, especially in parallel flight treatments (Table 2). Indeed, when SC was high even deposits were high. Average HNi was >90% for all spraying systems with best performances obtained by UAV in the over-the-line treatment (97.3%) (Table 2).

Experimental parameters

The mean water sprayed (Table 2) ranged from 78 l/ha for UAV flying over-the-line to 1,052 l/ha in the mist-blower sprayer. According to Matthews (2000), the average water sprayed by each system was classified as high, low and very low for mist-blower sprayer, backpack and UAV, respectively. Backpack spraying showed longer spraying event length with higher variability (327.3±24.6 s) while mist-blower sprayer and UAV had close means (<105 s) and standard deviations. Longer time observed in the former was due to human operations which yielded a larger imprecision in walking at a constant speed during the spraying and due to differences on forward speed, especially if it is compared with 5.6 km/h used with the mist-blower sprayer.

Table 2. Deposits/cm, number of impacts homogeneity index-HNi (%), spraying event length (s) and water sprayed (l/ha). Letters indicate significant differences at α=0.05.

	Deposits/cm	HNi (%)	Event length (s) mean ± s.d.	Water sprayed (l/ha)
Mist-blower sprayer	370.93 a	91.07 a	66.67±6.43	1,052 a
UAV over	40.56 c	97.31 a	50.00±5.29	78 c
Backpack	216.90 b	94.75 a	327.33±24.58	423 b
UAV parallel	19.22 c	95.05 a	104.67±5.86	137 c

Volume diameters and relative span factor (RSF)

Average volume diameters $DV_{0.1}$, $DV_{0.5}$ and $DV_{0.9}$ (μm) for each spraying system are reported in Table 3. According to the American Society of Agricultural and Biological Engineers (ASABE, former ASAE) which developed the standard ASAE 572.1 'Spray Nozzle Classification by Droplet Spectra' (ASAE, 2009), the mist-blower sprayer showed the highest $DV_{0.1}$ (higher the $DV_{0.1}$, lower the risk of drift), while all other systems show average $DV_{0.1}$ lower than 150 μm, (high risk of drift). The $DV_{0.5}$ measured for the mist-blower sprayer allows classification of its spray as ultra-coarse (>665 μm) while backpack and UAV values correspond to medium (236-340 μm) and fine (106-235 μm) ranges of spray quality, respectively. With respect to the $DV_{0.9,}$ high values were obtained for the mist-blower sprayer (higher the $DV_{0.9,}$ higher the risk of runoff). The other systems gave intermediate diameters. Moreover, the higher the $DV_{0.9}$, the lower is the droplets coverage ability. Average RSF values (Table 3), determined by using DepositScan volume diameters data, highlighted an optimal drop uniformity distribution in backpack spraying (RSF=0.99), while poorer performances were observed in the other treatments. In mist-blower sprayer (RSF=1.28), the size distribution was skewed toward bigger diameters whereas the opposite was observed in UAV flying over-the-line (0.85) and UAV flying in parallel (0.78).

Conclusions

UAV had a strong consistency, displaying a more uniform deposit distribution within each tree when compared to the mist-blower sprayer (which gave higher coverage percentages). In terms of sustainability of chemical treatments, UAV was demonstrated to be a viable system for reducing spraying volumes even though average volume diameters measured for this technology can be linked to the risk of spray drift. SnapCard application gave satisfactory performance in terms of coverage assessment. Therefore, it can represent a reliable decision-support tool for farmers and technicians. Many challenges are still open to assess feasibility of UAVs for precision pesticide application such as testing different flying paths, speed and type of nozzles to improve pesticide spray coverage.

Table 3. Drop size distribution – average volume diameters (μm) determined by using DepositScan for each spraying system and RSF calculated by using these diameters. Letters indicate significant differences at α=0.05.

	$DV_{0.1}$ (μm)	$DV_{0.5}$ (μm)	$DV_{0.9}$ (μm)	RSF
Mist-blower sprayer	434.09 a	1,282.95 a	2,149.34 a	1.28 a
UAV over	81.22 b	152.66 b	206.28 b	0.85 c
Backpack	138.46 b	298.93 b	468.79 b	0.99 b
UAV parallel	83.07 b	135.4 b	182.86 b	0.78 c

Acknowledgements

The authors express their sincere gratitude to the company 3D S.A. General Aviation Applications, for providing the UAV used in this study and a certified operator by the Hellenic Aviation Agency.

References

ASAE S572.1., 2009. Spray Nozzle Classification by Droplet Spectra. St. Joseph, MI., USA: American Society of Agricultural and Biological Engineers. 4 pp.

Chen, Y., Ozkan, H. E., Zhu, H., Derksen, R. C. and Krause, C. R., 2013. Spray deposition inside tree canopies from a newly developed variable-rate air-assisted sprayer. Transaction of the ASABE-American Society of Agricultural and Biological Engineers 56 1263-1272.

Faiçal, B. S., Freitas, H., Gomes, P. H., Mano, L. Y., Pessin, G., de Carvalho, A. C. P. L. F. *et al.,* 2017. An adaptive approach for UAV-based pesticide spraying in dynamic environments. Computers and Electronics in Agriculture 138 210-223.

Faiçal, B. S., Pessin, G., Filho, G. P. R., Carvalho, A. C. P. L. F., Furquim, G. and Ueyama, J., 2014. Fine-tuning of UAV control rules for spraying pesticides on crop fields. Institute of Electrical and Electronics Engineers-Computer Society 14 1082-3409.

Hoffman, W. C. and Hewitt, A. J., 2005. Comparison of three imaging systems for Water Sensitive Papers. Applied Engineering in Agriculture 21 961-964.

Lazzaro, L., Otto, S. and Zanin, G., 2008. Role of hedgerows in intercepting spray drift: evaluation and modelling of the effects. Agriculture, Ecosystems and Environment 123 317-327.

Marçal, A. R. S. and Cunha, M., 2008. Image processing of artificial targets for automatic evaluation of spray quality. Transactions of the ASABE-American Society of Agricultural and Biological Engineers 51 811-821.

Matthews, G. A., 2000. Pesticide application methods. Third Edition. Oxford, UK: Blackwell Science Ltd. 46 pp.

Miranda-Fuentes, A., Llorens, J., Rodríguez-Lizana, A., Cuenca, A., Gil, E., Blanco-Roldán, G. L. *et al.,* 2016. Assessing the optimal liquid volume to be sprayed on isolated olive trees according to their canopy volumes. Science of the Total Environment 568 296-305.

Miranda-Fuentes, A., Rodríguez-Lizana, A., Cuenca, A., Gonzàlez-Sànchez, E. J., Blanco-Roldàn, G. L., Gil-Ribes, J. A., 2017. Improving plant protection product applications in traditional and intensive olive orchards through the development of new prototype air-assisted sprayers. Crop Protection 94 44-58.

Miranda-Fuentes, A., Rodríguez-Lizana, A., Gil, E., Agüera-Vega, J. and Gil-Ribes, J. A., 2015. Influence of liquid-volume and airflow rates on spray application quality and homogeneity in super-intensive olive tree canopies. Science of the Total Environment 537 250-259.

Nansen, C., Ferguson, J.C., Moore, J., Groves, L., Emery, R., Garel, N. *et al.,* 2015. Optimizing pesticide spray coverage using a novel web and smartphone tool, SnapCard. Agronomy for Sustainable Development 35 1075-1085.

Patel, M. K., Sahoo, H. K., Nayak, M. K., Kumar, A., Ghanshyam, C. and Kumar, A., 2015. Electrostatic nozzle: new trends in agricultural pesticides spraying. Seventh Sense Research Group-International Journal of Electrical and Electronics Engineering 6-11.

Pérez-Ruiz, M., Agüera, J., Gil, J. A. and Slaughter, D. C., 2011. Optimization of agrochemical application in olive groves based on positioning sensor. Precision Agriculture 12 564-575.

Russo, C., Cappelletti, G. M., Nicoletti, G. M., Di Noia, A. F. and Michalopoulos, G., 2016. Comparison of European olive production systems. Sustainability 8 (825) 1-11.

Zheng, Y. J., Yang, S. H., Zhao, C. J., Chen, L. P., Lan, Y. B. and Tan, Y., 2017. Modelling operation parameters of UAV on spray effects at different growth stages of corns. International Journal of Agricultural and Biological Engineering 10 57-66.

Zhu, H., Salyani, M. and Fox, R. D., 2011. A portable scanning system for evaluation of spray deposit distribution. Computers and Electronics in Agriculture 76 38-43.

UAV-based hyperspectral imaging for weed discrimination in maize

R. Casa[1*], S. Pascucci[2], S. Pignatti[2], A. Palombo[2], U. Nanni[3,5], A. Harfouche[4], L. Laura[3,5], M. Di Rocco[3,5] and P. Fantozzi[3,5]
[1]*Department of Agricultural and Forestry scieNcEs (DAFNE), University of Tuscia, Via San Camillo de Lellis, 01100 Viterbo, Italy; rcasa@unitus.it*
[2]*Institute of Methodologies for Environmental Analysis (IMAA), Consiglio Nazionale delle Ricerche, Via del Fosso del Cavaliere 100, 00133 Roma, Italy*
[3]*Department of Computer, Control, and Management Engineering (DIAG), University of Rome 'La Sapienza', Via Ariosto 25, 00185 Rome, Italy*
[4]*Department for Innovation in Biological, Agro-food and Forest systems (DIBAF), University of Tuscia, Via San Camillo de Lellis, 01100 Viterbo, Italy*
[5]*Research Centre in Transport and Logistics (CTL), University of Rome 'La Sapienza', Via Eudossiana 18, 00185 Rome, Italy*

Abstract

Timely weed mapping in crop post-emergence situations is a challenging task required for developing precision weed management solutions. It is necessary to discriminate the crop from the weeds and, if possible, to distinguish different weed species. The ability to map weeds using hyperspectral images acquired from an unmanned airborne vehicle (UAV) over a maize field was evaluated by comparing different classification strategies. The results were mainly affected by the variability in crop and weed spectral signatures. The discrimination between maize and weeds allowed the quantification of their relative ground cover, showing moderate relationship with their relative leaf area index.

Keywords: weed maps, drones, imaging spectroscopy, image classification, machine learning

Introduction

For precision weed management, is necessary to discriminate the crop from the weeds and, if possible, to distinguish at least broad-leaved weed species (dicots) from grasses (monocots), since they are usually treated with different herbicides. There are several works showing successful weed detection on maize using UAVs equipped with multispectral sensors (e.g. Peña *et al.*, 2013; López-Granados *et al.*, 2016; Castaldi *et al.*, 2017; Pflanz *et al.,* 2018).
For this task, however, hyperspectral data are expected to have an advantage over multispectral data (Herrmann *et al.,* 2013). The former are able to capture, for example, differences in leaf pigment content, leaf mesophyll cell structure and leaf epidermis properties, practically invisible in coarse spectral resolution multispectral data, but potentially very useful for the discrimination of different plant species.
The advances in the miniaturization of imaging spectrometers (Aasen *et al.*, 2018), has led to their availability on-board unmanned airborne vehicles (UAV), paving the way for a potentially more rapid and efficient large-scale weed monitoring than from ground-based proximal platforms.
The cost of hyperspectral sensors is decreasing, so it is predictable that they will gain much wider diffusion in the coming years, allowing operational UAV-based hyperspectral crop monitoring and mapping.
However, it is important to investigate whether the discriminatory power of hyperspectral data at the leaf level, is still carried over and effective at the canopy scale, in a real field scenario, such as the one that may be encountered in an operational precision weed mapping context. In such a situation, there are many confounding factors that could interfere with the ability to discriminate crops and weeds: (1) the interaction of the geometry of illumination and observation with the canopy structure,

DOI 10.3920/978-90-8686-888-9_45, © Wageningen Academic Publishers 2019

mainly depending on the leaf angle distribution; (2) the variability in spectral signatures of plant species, which depend on their development stage, as well as on their polymorphism (phenotype) and growth conditions; (3) the actual spectral and spatial resolution and quality that is possible to achieve in operational contexts, given the specificities of UAV-based sensors and acquisition modes (Aasen *et al.*, 2018).
The objective of the present work was to assess the possibility of discriminating between maize and weeds in post-emergence conditions using UAV-based hyperspectral imaging, by exploring different image processing strategies.

Materials and methods

Hyperspectral images were acquired on June 14, 2016, over a maize field (Lat. 42.33° N, Lon. 12.99° E) located in Vetralla (Central Italy), from a height of about 30 m with an octocopter Multirotor G4 Skycrane UAV (Multitrotor, Brieselang, Germany) equipped with a pushbroom Nano-Hyperspec sensor (Headwall Photonics, Bolton, USA). This sensor records single hypercubes of 275 spectral bands in the visible and near-infrared (VNIR) range (400-1000 nm) with ~2.2 nm spectral interval and 5 nm spectral resolution (fwhm with 20 µm slit width). It is equipped with a calibrated f/1.8 17 mm lens, which results in 15.9° field-of-view over 640 pixels (spatial resolution). A 2-axis gimbal was used to stabilise the sensor in relation to the aircraft flight dynamics, and a global navigation satellite system (GNSS) receiver and inertial measurement unit (IMU, Xsens MTI-G-700 GPS/INS, Xsens Technologies, Enschede, the Netherlands) collected both geographical data and roll/pitch/yaw information to perform hypercube ortho-rectification. Since the flight was performed manually, directly over sampling frames placed on the ground, no mosaicing was carried out and only the images containing the quadrats were selected. Pre-processing of the hyperspectral data was performed using HyperSpecIII and SpectralView software (Headwall Photonics), which integrates the GNSS and IMU data to geometrically correct images and produce ortorectified hypercubes. The hypercubes were converted into calibrated radiance using the SpectralView software. A white Spectralon panel (Labsphere, North Sutton, USA) was placed in the field for the correction of raw data into spectral reflectance. Images were pre-processed into absolute reflectance by using the FLAASH tool based on MODTRAN4 radiative transfer code implemented in the ENVI 5.4.1 software (Harris Geospatial Solutions, Boulder, USA). The ground spatial resolution of the images obtained was of about 10 mm.
At the time of the flight, the maize crop, planted with a distance of 0.75 m between rows, was at the 6^{th} leaf stage (V6) and was naturally infested by weeds. The weed species found in the field were the dicots white goosefoot (*Chenopodium album* L.), redroot pigweed (*Amaranthus retroflexus* L.) and wormwood (*Artemisia absinthium* L.), and the monocots purple crabgrass (*Digitaria sanguinalis* L.) and sedge (*Cyperus rotundus* L.). In correspondence with the flights, 8 rectangular frames (hereafter quadrats) with a size of 0.75 by 1.5 m, were placed on the ground inside the field and georeferenced using a RTK Trimble Geo7x (Trimble, Sunnyvale, USA) with a precision of 15 mm. Reflectance spectra were collected, using an ASD FieldSpec FR Pro (Malvern Panalytical Ltd, Malvern, UK) with a 400-2,500 nm spectral range, acquiring 4 replicate spectra for each weed species found inside each quadrat, as well as for maize and bare soil. A white reference Spectralon panel spectrum was acquired every 4 samples to calibrate the data into absolute reflectance. Subsequent to the flights and spectra collection, the whole above-ground biomass inside the quadrats was harvested, placed inside plastic bags and taken to the lab, where plants from different weed species where separated and identified. The dry biomass and leaf area index (LAI), the latter obtained using a LI-3100C Area Meter (Li-cor, Lincoln, USA), was then measured, separately for maize and for each weed species. Three image classification strategies were then compared: (1) unsupervised K-means clustering and supervised Support Vector Machine (SVM); (2) a physiologically-based approach, exploiting the differences in chlorophyll/carotenoid pigments among plant species; (3) a two-phase machine

learning (ML) approach based on both supervised neural networks (NN) and unsupervised K-means, the latter using a subset of selected spectral bands to reduce the computational burden.

In the first approach, two commonly used classification algorithms were applied to the hyperspectral images acquired from the UAV. The unsupervised K-means algorithm was employed, by fixing *a priori* the number of clusters at six, encompassing all the broad classes occurring in the UAV imagery (i.e. soil, maize, weeds, shadow and plastic quadrat frames). The supervised SVM classifier was applied by defining regions of interest (ROI) of the training classes, manually on each image, by means of ground-truth data (photographs acquired on the ground above the quadrats). The SVM parameters used in the ENVI 5.4.1 software were (1) the radial basis function as kernel type, with a gamma in the kernel function of 1.0; (2) a penalty parameter of 100 that allows a certain degree of misclassification; and (3) a classification probability threshold of 0.30, which is used to set the probability that is required for the SVM classifier to classify a pixel, so that pixels where all rule probabilities are 0.30 were unclassified.

The second approach explored, was based on the retrieval of pigments, i.e. carotenoids (Car) and chlorophylls (Cab) and their ratio. Two methods were used for the estimation of the pigments: (1) the inversion of the weeds spectra acquired by the ASD and recorded by UAV VNIR data by using the radiative transfer model PROSAIL (Berger *et al.*, 2018); (2) the use of narrow-band spectral indices, sensitive to chlorophylls or carotenoids. The first method is based on the PROSAIL inversion process based on a least-squares merit function and a numerical optimization routine. The procedure was stopped at the 32nd iteration. Before applying the model inversion procedure to field collected spectra and UAV image data, it was tested on two experimental datasets, LOPEX and ANGERS (Féret *et al.,* 2011), in which both spectral and lab pigment measurements were available. The small root mean squared error (RMSE) values obtained for pigment retrieval in these datasets, 0.76 μg/cm^2 for chlorophyll and 2.21 μg/cm^2 for carotenoids, confirmed the correctness of the inversion process. The second method, based on the use of indices designed for use with imaging spectrometers and therefore use specific wavelengths sensitive to pigments (Table 1), was used to study the separability of crop and weed species in 2-D index space scatter plots. In particular, the GRBD index measures the band depth between 540 and 690 nm and can, therefore, be applied only to hyperspectral data. The use of pigment indices for species discrimination was first tested on ground collected spectra and subsequently applied to the UAV images.

In the third approach, a two-phase methodology was employed. Firstly, a supervised approach using a convolutional neural network was used to separate weeds from the other classes (maize, soil, shadow and plastic) in the UAV reflectance images with 276 bands. The training set was based on the visual selection of pixels in several portions of one of the images. In the second phase, an unsupervised approach (k-means algorithm), was used to separate the different weed species. In this case, 60 bands in the range 506-616 nm were selected, on the basis of the inspection of ground collected spectra: i.e. bands for which relevant differences in reflectance among weed species were chosen. Thus, a sub-classification of the 'weeds' pixels in two or three classes was obtained. The

Table 1. Spectral indices used for the discrimination of maize and weed species

Index name	Formulation	Source
CARI	719 / 721	Zhou *et al.,* 2017
Cnorm	(700 – 515) / (700 + 515)	this work
GRBD	band depth 540 – 690	this work
OSAVI	(800 – 670) / (800 + 680 + 0.16)	Rondeaux *et al.,* 1996
RE	linear extrapolation	Cho & Skidmore (2006), modified

classes were then labelled as the different weed species, by matching the percentages of classified pixels with the ranking of ground measured LAI of the weeds species in each quadrat.

Results

The hyperspectral images obtained from the UAV had a good spectral quality when compared to ground collected spectra, but required spectral smoothing (binning) to get rid of noise, especially in the NIR bands. The ortho-rectification of the hypercubes, built from the scan lines of the pushbroom sensor, showed some residual small distortions (Figure 1A). An example of the results obtained by applying the standard K-means clustering algorithm and the SVM classifier on the UAV reflectance images is shown in Figure 1. The K-means clustering, after labelling the classes according to image inspection, provided a better relationship than SVM between the pixel fraction of maize and the ground measurements of relative LAI (i.e. percentage of total LAI) obtained from the lab, with a determination coefficient R^2=0.28 and RMSE of 20%. On the other hand SVM provided better results for weed percent cover, with a R^2=0.29 and RMSE of 16%, as compared to K-means for which a R^2=0.21 was obtained.

By analysing the field spectral data set acquired with the ASD FieldSpec FR Pro, it was observed that some weeds could be discriminated from maize on the basis of their leaf chlorophyll/carotenoids pigment content ratio, as retrieved by PROSAIL inversion. In particular, *Amaranthus*, *Cyperus* and maize could be spectrally separated in the Cab/Car pigment space (Figure 2A), or even better in the Cnrom/GRBD index space (Figure 2B). The others weeds species did not show such a clear separability.

Figure 3 shows an example of the application of this separability strategy to UAV images, for the quick discrimination of maize, *Amarantus* and *Cyperus* in a sampling quadrat. The pixel density plot, in the Cnrom/GRBD index space (Figure 3A), allowed three clusters to be highlighted, roughly similar to those observed in Figure 2B. These clusters of pixels corresponded to areas of the UAV image (Figure 3B) assumed to be covered by maize, *Cyperus* and *Amaranthus* (Figure 3C). The fraction of pixels of the different classes retrieved in the quadrat, i.e. 49% maize, 40% *Cyperus* and 11% *Amaranthus*, were in the same order to the ground measurements of relative LAI, respectively 59% maize, 40% *Cyperus* and 1% *Amaranthus*. The difference in the maize LAI assessment by using physiologic traits can be related to the shadow component that was not masked off when applying the indices.

The two-step ML procedure allowed firstly the classification of UAV reflectance images into the same classes as the first approach (i.e. maize, weeds, shadows and plastic frame) using a supervised convolutional neural network (Figure 4). An RMSE of 16% and a R^2=0.1 was obtained for the relationship between pixel fraction and relative LAI of weeds measured on the ground.

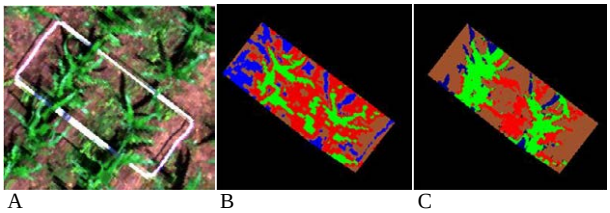


Figure 1. Example of classification of a hyperspectral UAV image: (A) surface reflectance image in natural colour (RGB), (B) K-means and (C) SVM classification maps, respectively. In the two classifications (B and C), green, red, brown and blue colors represent maize, weeds, soil and shadows classes, respectively.

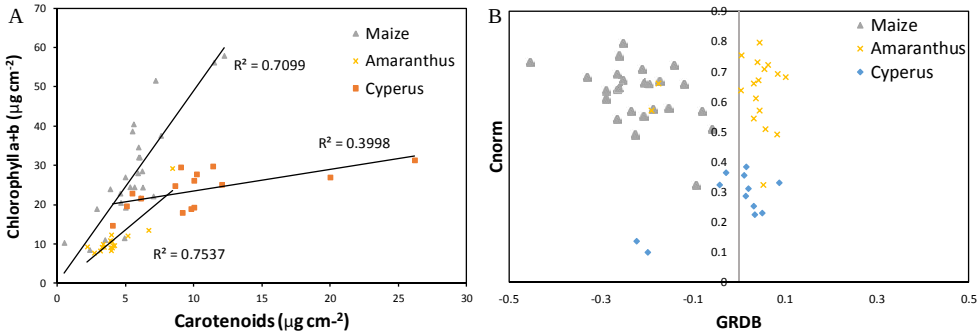


Figure 2. Separability among plant species in the 2-D space of (A) leaf chlorophylls and carotenoids retrieved by PROSAIL inversion and (B) spectral indices Cnorm and GRDB (Table 1), for spectral reflectance data collected on the ground.

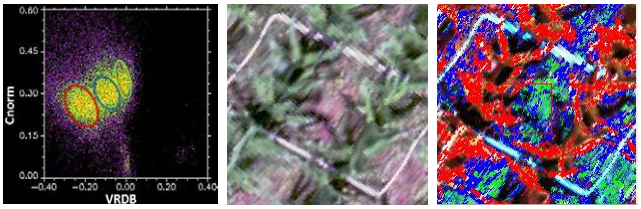


Figure 3. Application of the classification method based on chlorophyll and carotenoids spectral indices, to a UAV hyperspectral image: (A) pixel density plot in the Cnrom/GRBD index space with ellipses identifying three different clusters; (B) natural color UAV image of a ground sampling quadrat; (C) UAV image with highlighted pixels corresponding to clusters of Figure 3A: red for maize, green for *Amarantus* and blue for *Cyperus*.

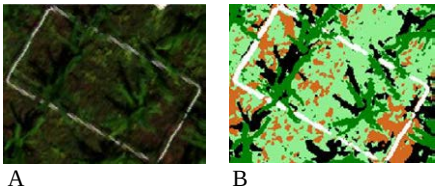
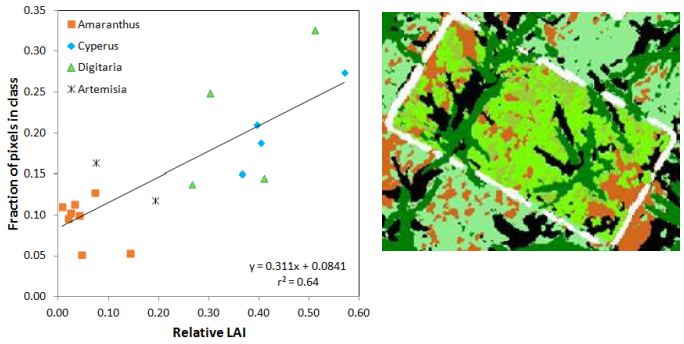


Figure 4. Example of classification of a UAV sampling frame quadrat (A) using a supervised convolutional neural network (B). In the classified image (B) dark green is maize (26.5%), light green is weeds (44.0%), brown is soil (21.2%), black is shadows (5.5%) and white is plastic (2.8%).

In the second step, using the unsupervised K-means approach, it was possible to discriminate among weed species (Figure 5), obtaining a good relationship between the fraction of pixels of the different species and their absolute LAI, measured on the ground, with an R^2=0.71 which reached R^2=0.85 after removal of one outlier. This outlier occurred in the quadrat where the area classified as shadow was maximum (13.33%). Furthermore, the error (differences between the measured LAI and the percentage of the classified pixels) was highly dependent on the weed species. The relationship of



A B

Figure 5. Results of the second step of ML classification approach, using K-means: (A) relationship between the relative LAI of weed species and their corresponding pixel fractions; (B) example of classified image. In the classified image (B) dark green is maize, brown is soil, white is plastic, bright green is *Digitaria* (32.6%) and olive green is *Amaranthus* (11.4%).

pixel fractions with the relative LAI, resulted in a RMSE of 0.15 and a R^2=0.64 for all weed species together (Figure 5A).
An underestimation of the relative LAI was nevertheless evident (Figure 5A) and it can be possibly attributed to the shadow areas and to a fraction of the LAI of the weeds that is masked from the aerial view by maize plants. Therefore there is evidence that correcting factors (e.g. depending on shadows and the weed species) might allow improvement of the accuracy of these estimates.

Conclusions

The hyperspectral images acquired with a UAV and corrected into ortho-rectified surface reflectance, had some residual distortion effects resulting from the pushbroom nature of the imager. It is clear that, with these data, the classification of weeds inside the maize crop is rather difficult, taking into account the typically small size of weed plants. Nevertheless, standard supervised and unsupervised classification methods, could allow discrimination between the crop and the weeds.
Furthermore, an original approach was developed, based on the concept of exploiting the differences chlorophyll/carotenoids content ratios, among weeds and maize, since these pigments can be retrieved from hyperspectral data. The under-estimation of maize relative LAI, in the physiological traits retrieval from UAV hyperspectral images, can be partially explained as an effect of the shadow areas produced by the maize foliage.
The most encouraging results, however, were obtained by a two-step machine learning approach, which allowed discrimination between different weed species and estimation of their relative LAI. However, an under-estimation of weed relative LAI occurred. It was attributed to the presence of unclassified shadow areas and to the masking of part of the leaf area of weeds by the maize plants or upper weed leaf layers.

Acknowledgements

This research was funded by MIPAAF within the APREINF project. The authors thank Federico Pelosi and Riccardo Salvati for data collection. The authors gratefully acknowledge MIUR (Minister for education, University and Research) for financial support (Law 232/216, Department of excellence) to DAFNE.

References

Aasen, H., Honkavaara, E., Lucieer, A. and Zarco-Tejada, P.J., 2018. Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: a review of sensor technology, measurement procedures, and data correction workflows. Remote Sensing 10 1091.

Berger, K., Atzberger, C., Danner, M., D'Urso, G., Mauser, W., Vuolo, F. *et al.*, 2018. Evaluation of the PROSAIL model capabilities for future hyperspectral model environments: a review study. Remote Sensing 10(1) 85.

Castaldi, F., Pelosi, F., Pascucci, S., Casa, R., 2017. Assessing the potential of images from unmanned aerial vehicles (UAV) to support herbicide patch spraying in maize. Precision Agriculture 18 76-94

Cho, M. A., and Skidmore, A. K. 2006. A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method. Remote Sensing of Environment 101 (2) 181-193.

Féret, J.-B., François, C., Gitelson, A., Asner, G.P., Barry, K.M., Panigada, C. *et al.*, 2011. Optimizing spectral indices and chemometric analysis of leaf chemical properties using radiative transfer modeling. Remote Sensing of Environment 115, 2742-2750.

Herrmann, I., Shapira, U., Kinast, S., Karnieli, A., Bonfil, D.J., 2013. Ground-level hyperspectral imagery for detecting weeds in wheat fields. Precision Agriculture 14(6), 637-659.

López-Granados, F.; Torres-Sánchez, J.; Castro, A.-I. D.; Serrano-Pérez, A.; Mesas-Carrascosa, F.-J.; Peña, J.-M., 2016. Object-based early monitoring of a grass weed in a grass crop using high resolution UAV imagery. Agronomy for Sustainable Development. 36, 67.

Peña, J. M.; Torres-Sánchez, J.; de Castro, A. I.; Kelly, M.; López-Granados, F., 2013. Weed Mapping in Early-Season Maize Fields Using Object-Based Analysis of Unmanned Aerial Vehicle (UAV) Images. PLoS ONE 8, e77151.

Pflanz, M., Nordmeyer, H., & Schirrmann, M., 2018. Weed Mapping with UAS Imagery and a Bag of Visual Words Based Image Classifier. Remote Sensing, 10 (10), 1530.

Rondeaux, G., Steven, M., and Baret, F., 1996. Optimization of soil-adjusted vegetation indices. Remote Sensing of Environment 55 95-107.

Zhou, X., Huang, W., Kong, W., Ye, H., Dong, Y., and Casa, R., 2017. Assessment of leaf carotenoids content with a new carotenoid index: development and validation on experimental and model data. International Journal of Applied Earth Observation and Geoinformation 57 24-35.

A precision viticulture UAV-based approach for early yield prediction in vineyard

S.F. Di Gennaro[1*], P. Toscano[1], P. Cinat[1], A. Berton[2] and A. Matese[1]
[1]*Institute of Biometeorology, National Research Council, 50145 Florence, Italy; f.digennaro@ibimet.cnr.it*
[2]*Institute of Clinical Physiology, National Research Council, 56124 Pisa, Italy*

Abstract

Early yield prediction is a key factor for vineyard management optimization, to obtain the desired grape production and quality. Ground observation and manual weighing are time consuming and frequently provide low representative data. The aim of the study was to develop an automated early yield estimation system (5 weeks before harvest) using high-resolution RGB images, acquired through an unmanned aerial vehicle (UAV) platform in a representative zone of vigour variability of the whole vineyard. An unsupervised recognition algorithm was applied to derive the number of clusters and size, which have been used to estimate production. This fast and accurate methodology, which operated with a low cost setup, has shown high accuracy in yield prediction, providing interesting potential to support grape production management both in vineyard and in cellar.

Keywords: UAV, image analysis, grape detection, early yield prediction, low-cost

Introduction

Knowledge of spatial variability across a field plays a key role in vineyard management, affecting a wide range of aspects related to crop development (Bonilla *et al.*, 2015). Among those, grape yield is an important component that drives the return of investment in wine production (Van Leeuwen and Darriet, 2016) and its estimation is a challenging process. Moreover, yield prediction can improve the efficiency both for harvest operations in the field and winemaking management (Cunha *et al.,* 2010; De la Fuente *et al.,* 2015). Forecast prediction is a key aspect for farmers but yield variability across years, makes a correct estimation between different seasons difficult (Sabbatini *et al.*, 2012). Traditional yield estimation is realized by taking measurements manually in the field and extrapolating to the entire vineyard – making it extremely time consuming, destructive and subjective (Nuske *et al.,* 2014). Furthermore, there is a high risk that a limited number of samples fails to properly characterize spatial variability within the vineyard, thus compromising the data (Tisseyre *et al.,* 2018). In this context, the emerging technologies applied in agriculture provide new solutions to overcome the weaknesses of traditional ground measurements (Dhondt *et al.,* 2013; Di Gennaro *et al.,* 2017; White *et al.,* 2012). Viticulturists have developed different systems for yield estimating. De La Fuente *et al.* (2015) provided a performance review of six yield prediction methods applied to different phenological stages, identifying fruit set and veraison as best solutions in terms of earliness of prediction in the growing season, accuracy and precision.
In the last decade, advances in sensor technologies and image computing software provide a valuable tool to overcome those issues. As summarized by Aquino *et al.* (2018), different approaches have been tested to realize a partial or total unsupervised cluster for berry detection by high resolution RGB images. The most widely used approaches to analyse pixel colours is the RGB colour space based on three colour components which are red (R), green (G), and blue (B). However, several colour spaces have been introduced to exploit their mathematical properties (Khattab *et al.,* 2014). Among different colour spaces, the L*a*b* and HSV are the most frequently used (Bora *et al.,* 2015). Some authors have reported studies of image processing techniques taken manually in a vineyard (Liu and Whittey, 2015; Aquino *et al.,* 2017; Mack *et al.,* 2017). Other recent studies have reported

DOI 10.3920/978-90-8686-888-9_46, © Wageningen Academic Publishers 2019

methodologies for large-scale yield monitoring in the field, by means of on-the-go image acquisition using modified terrestrial vehicles and 2D image analysis approaches (Nuske *et al.,* 2014; Font *et al.,* 2015; Aquino *et al.,* 2018). Those solutions have acquired images at night using artificial illumination on defoliated cluster zones, which provide accurate yield detection. However, they are time-consuming due to low forward speed especially in night condition. The aim of this study was to implement a fast and unsupervised methodology for yield prediction several weeks before harvest with reduced time and costs in terms of acquisition and processing of data. To obtain early automated yield estimation, high resolution RGB images were taken with a low cost UAV platform in a representative zone of vigour variability in a vineyard.

Materials and methods

Experimental site

The research was carried out in a 1.4 ha vineyard (43°25'45.30'N, 11°17'17.92'E) planted in 2008 and located on a slight southern slope at 355 m above sea level near Siena (Italy). Sangiovese cv. (*Vitis vinifera*) vines are trained to a vertical shoot-positioned trellis and spur-pruned on a single cordon with 2.2×0.75 m vine spacing and NW-SE rows.

Image acquisition

UAV technologies were chosen considering their potential in term of flight speed, low quote operative flight and custom angle of image acquisition. Remote aerial surveys were performed using an open-source UAV platform consisting in a modified multi-rotor MikrokopterOktoXL (HiSystems GmbH, Moomerland, Germany). The system was equipped with a multispectral camera ADC-Snap (Tetracam Inc., CA, USA) and a RGB camera Sony Cyber-shot DSC-QX100 (Sony Corporation, Tokyo, Japan). The UAV flight campaign was conducted in the experimental vineyards on 26 June and 07 August 2018. A preliminary survey on first date was performed equipping the multispectral camera to analyze vigour spatial variability, and to identify 2 representative zones, one in low vigour (LV) and one in high vigour (HV). A second multispectral survey was undertaken on 07 August to confirm that a vigour representative zone had been chosen, and produce the yield map. Multispectral image acquisition was at 50 m AGL, yielding a ground resolution of 0.025 m/pixel and 70% of overlap in both directions. During the second flight date, 8 consecutive vines within each representative zone were defoliated at fruit zone level. A RGB survey was carried out at 14.00 PM with 5 m AGL flight quote, with oblique camera (45°) along the row, providing a ground resolution of 0.002 m/pixel on the fruit zone. A series of RGB images of vines were acquired hovering about 20 s in front of each 8-vine plot. Sentinel-2 images were freely downloaded from the Copernicus Scientific Data Hub web site as a Level 2A product. This level is ortho-rectified and UTM geocoded, bottom-of-atmosphere multi-spectral reflectance.

Image processing

Multispectral images acquired by UAV were processed using Agisoft Photoscan Pro 1.1.6 (Agisoft LLC, St. Petersburg, Russia) to produce the ortho-mosaic and the digital elevation model (DEM). Matlab software (The MathWorks, Inc., Natick, USA) was employed for radiometric calibration based on reference panels and row filtering (Matese *et al.,* 2016). Next, NDVI data were extracted at single vine level with Quantum GIS (QGIS – http://www.qgis.org/), to correlate vigour and yield of each sample vine. NDVI from Sentinel 2 imagery was calculated using near-infrared B8 band (842 nm) and red B4 band (665 nm), producing a 10-m spatial resolution map.

The image processing workflow is shown in Figure 1. A first supervised step has been done to identify best image fitting on target vines for each vigour zone and to cut the sample vines region. Then, ImageJ software was used to perform an unsupervised processing of selected images. At this point, each original RGB image was converted to the three layers HSV color space: H

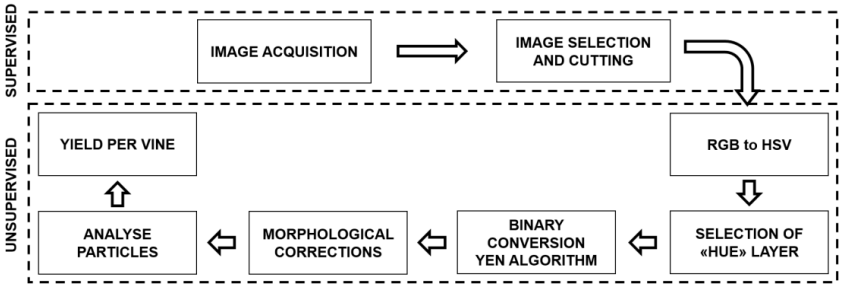


Figure 1. Image processing workflow.

(hue) distinguishes among the perceived colors, S (saturation) refers to how much of the light is concentrated at each specific wavelength of the hue; and V (value) represents the total brightness. The H layer was selected for image analysis, since it has the best performing layer for cluster segmentation. A threshold and two different filter methods were now applied to the H layer. (1) The Yen filter (Yen *et al.,* 1995) converted the layer to a 8 bit binary image, providing a black and white image with clusters segmentation from other objects (soil, leaf, trunk). (2) A median filter eliminated 'salt and pepper' noise, replacing each pixel with the median value in its 3×3 neighborhood. (3) A filling holes filter removed the small holes produced by the threshold based segmentation step. Finally, the command 'Analyze particles' was used to count and measure each cluster area detected in the binary 2D image.

Evaluation of yield prediction map

Yield estimation was carried out using the number and dimensions of detected clusters obtained from image analysis. First, an average yield per vine was calculated in each vigour zone from UAV images by converting total cluster area per vine (cm^2/vine) from UAV images in yield per vine (kg/vine), on the basis of calibration parameters obtained from previous work in the same vineyard (2017 season). The yield per vine within different vigour zones were then correlated with NDVI data extracted with a polygon centered on each sample vine from 07 August 2018 NDVI map. The linear estimation model was applied to produce a yield prediction map based on vigour spatial variability in the vineyard. The same approach was tested with the Sentinel 2 NDVI map acquired closest to cluster detection survey (18 July 2018).

Results and discussion

Following the workflow described above, the processing of each image required a maximum of one minute. Figure 2 shows graphical flow output applied in LV and HV zones, beginning with supervised selection and cutting of image of sample vines (Figure 2A,B), and then RGB to HSV transform (Figure 2C), and clusters data extraction from binary image corrected with morphological filters (Figure 2D).

Data analysis showed a strong linear correlation between estimated yield and observed yield in terms of ripe cluster per vine (R^2=0.84) (Figure 3A). The evaluation of the protocol response against ground traditional measurements showed a strong linear relationship during early survey at the beginning of August in mid veraison stage. However, unripe and partial overlapping conditions affect the absolute value detected, leading to an error of 17.8% in LV and 23.3% in HV in respect to ripe yield per vine, and higher concerning the total yield per vine (LV=30.5% and HV=27.8. Seasonally variability strongly affects cluster features, so high differences in terms of cluster morphology, berry ripeness within the cluster, cluster overlap level, canopy development and light condition could occur

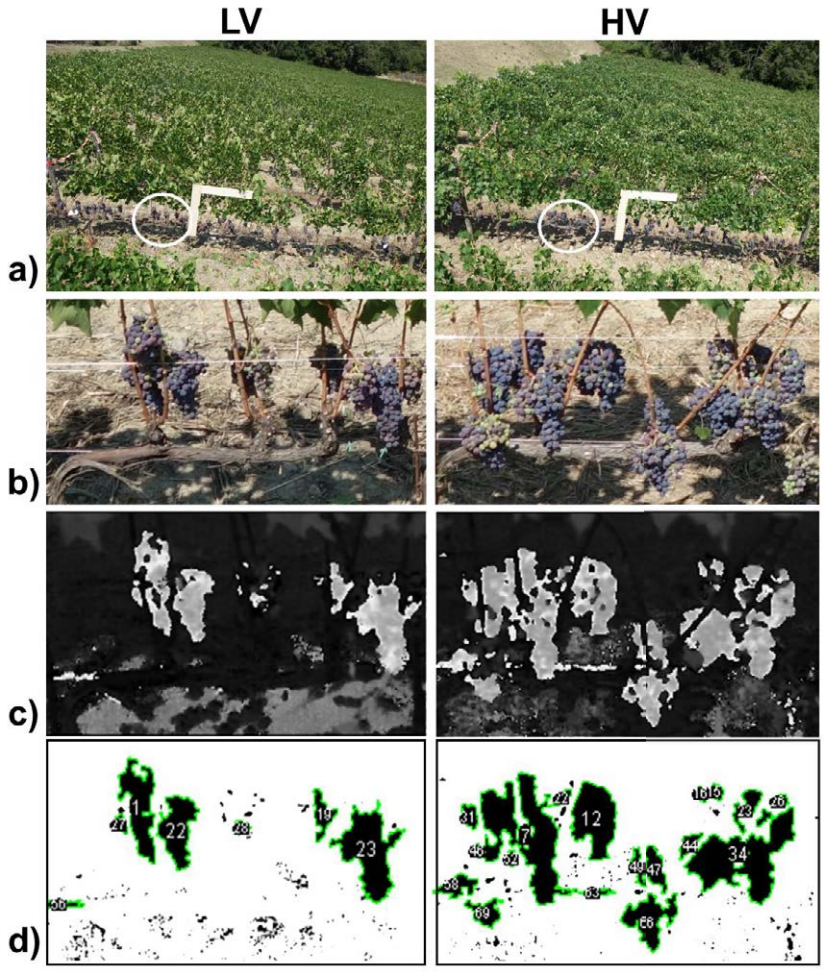


Figure 2. Example of image analysis chain from supervised (A, B) to unsupervised (C, D) step on low (LV) and high (HV) vigour sample vines: (A) original image, (B) sample vine detail; (C) hue layer of HSV color space; (D) binary image based on Yen threshold algorithm with morphological filter and clusters area extraction.

between different seasons. Consequently, it would be interesting to investigate the methodology performance across different seasons, and, finally potentially identify a seasonal coefficient to correct yield estimation.

However, errors in yield prediction are highlighted in other on-the-go research at harvest time with fully ripe grape: Font *et al.* (2015) achieved 16% prediction error while Nuske *et al.* (2014) reported yield prediction errors between 6.5-11.7%. Taking into account an early approach, Aquino *et al.* (2018) described the highest performance methodology reaching a 12.83% of error before veraison. As reported in other work on yield monitoring systems in vineyard (Taylor *et al.*, 2016), the methodology was confirmed to be effective for characterization of relative spatial yield patterns but, to improve absolute yield data, a seasonal adjustment is necessary against ground measurement on sample vines before it can be applied in large scale yield estimation. In regards to absolute mean

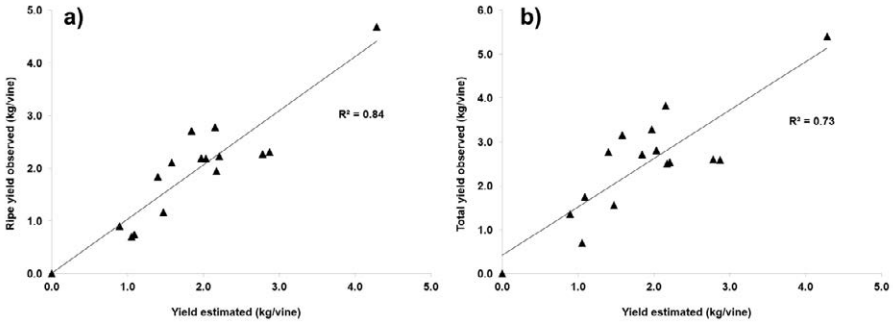


Figure 3. Estimated yield correlated with observed ripe cluster yield per vine (A), and observed total yield per vine (B).

values in different vigour zones, a discrete presence of partially overlapping between clusters in HV vines caused a mean yield under-estimation of the methodology (HV=2.33±0.97 kg, LV=1.78±0.69 kg) with respect to ground observation (HV=2.57±1.07 kg, LV=1.74±0.64 kg). Observing the performance in total yield estimation monitored at ground (HV=3.00±1.43 kg, LV=2.14±1.06 kg), the methodology loses all clusters that are partially or fully unripe, providing interesting linear correlation (R^2=0.73) but a higher under-estimation (Figure 3B). Undertaking a survey close to harvest, could reasonably achieve higher performance also in final yield estimation. Furthermore, an increase in berry weight from the beginning of August (UAV survey) could be another cause of the slight yield under-estimation compared to that measured at the harvest time 5 weeks later. Subsequently, NDVI data were extracted for sample vines in both vigour zones and correlated with the ripe yield per vine estimated by UAV approach. The linear regression between estimated yield of ripe grapes and NDVI extracted for each vine allowed a 'maturity yield map' generation based on NDVI spatial variability within the vineyard. A two-class yield map was generated based on the median distribution of NDVI values between high and low vigour (HV=2.84±1.20 kg/vine, LV=2.05±1.05 kg/vine) (Figure 4A). The same approach has been tested on a Sentinel 2 NDVI map

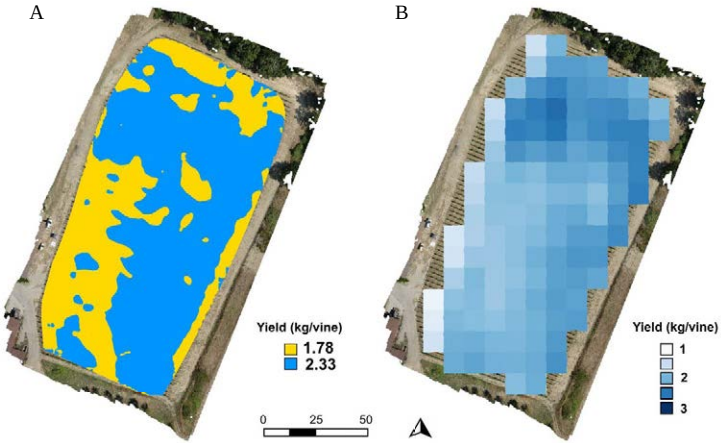


Figure 4. Maturity yield map based on vigour spatial distribution in UAV (A) and Sentinel 2 (B) NDVI map.

through correlation from NDVI pixel values containing 8 sample vines and mean estimated yield per vine providing similar results (Figure 4B). Those maps represent a snapshot of the ripening process within the vineyard to the farmer at the time of flight. Based on the knowledge of the vineyard potential yield, this tool could provide fast information for an accurate yield forecasting or eventually modify the cluster load with thinning practices to achieve yield and quality desired. Considering the short time needed to move within the vineyard, acquire images of fruiting zones and process images, the increase of the sampling size will add minimal impact in respect to a high improvement in accuracy. In fact, taking into consideration the possibility of monitoring also intermediate vigour zone, the methodology would be able to realize yield maps with greater prediction performance. However, given that the study vineyard is characterized by a clear distinction of vigour, the separation into 2 classes was sufficient.

Low-altitude inclined photogrammetry allows the acquisition of extreme detailed images of the fruiting zone. A primary requirement on the application of this method concerns the line of sight between sensor and cluster that has to be free to assure a good acquisition. However, the common practice of partial leaf removal a few weeks before harvest, to optimize ripening and preserve grape health, is a way to overcome this requirement. Furthermore, the automatic procedure based on color thresholds cannot discriminate green clusters from leaves, and therefore only partially and totally ripe clusters are identified. Except for these aspects, there are many advantages of the methodology. Firstly, the speed of acquisition and processing, which exploits single images taken in representative areas that are quickly reached by the UAV. This is a highly transferable approach, as it uses low-cost hardware and open source analysis software. Furthermore, the remote sampling approach based on spatial variability acquired from farmer knowledge and/or vigour map (high resolution UAV or free low resolution satellite) ensures representative data especially in large areas or highly fragmented and sloping vineyards. Although other on-the-go methods perform better in terms of accuracy in yield estimation, another strong point is the timeliness of acquiring images intra- and inter-fields, data processing and output report as simple spatial information or yield map. The proposed method also has high flexibility of application providing an automated way to acquire information such as leaf area or chromatic alterations caused by leaf symptomatology (biotic and abiotic stress).

The methodology is not intended to replace visual scouting on the ground, but as a tool to support scouting with potential for timely monitoring of large areas.

Conclusions

The innovative approach shows an interesting potential in practical application both for farmer support in grape production management both in vineyard and in cellar, and for service companies. The development of a yield estimation protocol based on low cost hardware (UAV and RGB camera) and software (open source image analysis tool), with high level of automation, provides a flexible and user-friendly solution for yield monitoring. The methodology developed has considerable potential in terms of recognizing the number of clusters and quickly estimating ripe yield at the time of flight; however, it suffers from an unavoidable physical limit represented by leaf cover.

The proposed approach requires further investigation to evaluate performance in different seasons, varieties and training systems, although an obvious limit remains the need to have a free line of sight between sensor and clusters.

Acknowledgements

The authors are grateful to Castello di Fonterutoli Marchesi Mazzei Spa (Siena, Italy) for the use of their vineyard to carry out this study.

References

Aquino, A., Diago, M. P., Millan, B., Tardaguila, J., 2017. A new methodology for estimating the grapevine-berry number per cluster using image analysis. Biosystems Engineering 156 80-95.

Aquino, A., Millan, B., Diago, M. P., Tardaguila, J., 2018. Automated early yield prediction in vineyards from on-the-go image acquisition. Computers and Electronics in Agriculture 144 26-36.

Bonilla, I., Martínez de Toda, F., Martínez-Casasnovas, J. A., 2015. Vine vigor, yield and grape quality assessment by airborne remote sensing over three years: Analysis of unexpected relationships in cv. Tempranillo. Spanish Journal of Agricultural Research 13 (2) e0903.

Bora, F. D., Bunea, C. I., Rusu, T., Pop, N., 2015. Vertical distribution and analysis of micro-, macroelements and heavy metals in the system soil-grapevine-wine in vineyard from North-West Romania. Chemistry Central Journal J 9 1-19.

Cunha, M., Maral, A. R. S. and Silva, L., 2010. Very early prediction of wine yield based on satellite data from vegetation. International Journal of Remote Sensing 31 3125-3142.

De La Fuente, M., Linares, R., Baeza, P., Miranda, C. and Lissarrague J. R., 2015. Comparison of different methods of grapevine yield prediction in the time window between fruitset and veraison. Journal International des Sciences de la Vigne et du Vin 49 27-35.

Dhondt, S., Wuyts, N. and Inzé, D., 2013. Cell to whole-plant phenotyping: the best is yet to come. Trends in Plant Science 18 (8) 428-439.

Di Gennaro, S. F., Rizza, F., Badeck, F. W., Berton, A., Delbono, S., Gioli, B., *et al.,* 2017. UAV-based high-throughput phenotyping to discriminate barley vigour with visible and near-infrared vegetation indices. International Journal of Remote Sensing 39 (15-16) 5330-5344.

Font, D., Tresanchez, M., Martínez, D., Moreno, J., Clotet, E., Palacín, J., 2015. Vineyard yield estimation based on the analysis of high resolution images obtained with artificial illumination at night. Sensors 15 8284-8301.

Khattab, D., Ebied H. M., Hussein A. S. and Tolba M. F. 2014. Color Image Segmentation Based on Different Color Space Models Using Automatic GrabCut. The Scientific World Journal 10 2014

Liu, S., Whitty, M., 2015. Automatic grape bunch detection in vineyards with an SVM classifier. Journal of Applied Logic 13, 643-653.

Mack, J., Lenz, C., Teutrine, J., Steinhage, V., 2017. High-precision 3D detection and reconstruction of grapes from laser range data for efficient phenotyping based on supervised learning. Computers and Electronics in Agriculture 135 300-311.

Matese, A., Di Gennaro, S. F. and Berton, A., 2016. Assessment of a canopy height model (CHM) in a vineyard using UAV-based multispectral imaging. International Journal of Remote Sensing 2016 1-11.

Nuske, S., Wilshusen, K., Achar, S., Yoder, L., Narasimhan, S., Singh, S., 2014. Automated visual yield estimation in vineyards. Journal of Field Robotics 3 837-860.

Sabbatini, P., Dami, I. and Howell, G. S., 2012. Predicting Harvest Yield in Juice and Wine Grape Vineyards. Extension Bulletin 3186, Michigan State University, USA.

Taylor, J.A., Sanchez, L., Sams, B., Haggerty, l., Jakubowski, R., Djafour S. *et al.*, 2016. Evaluation of a commercial grape yield monitor for use mid-season and at-harvest. Journal International des Sciences de la Vigne et du Vin 50 (2) 57-63

Tisseyre, B., Leroux, C., Pichon, L., Geraudie, V., Sari, T., 2018 How to define the optimal grid size to map high resolution spatial data? Precision Agriculture 19 (5) 957-971.

Van Leeuwen, C. and Darriet, P., 2016. The impact of climate change on viticulture and wine quality. Journal of Wine Economics 11 (1) 150-167.

White, J.W., Andrade-Sanchez, P., Gore, M.A., Bronson, K.F., Coffelt, T.A., Conley, M.M., *et al.,* 2012. Field-based phenomics for plant genetics research. Field Crops Research 133 101-112.

Yen, J. C., Chang, F. J. and Chang, S., 1995. A new criterion for automatic multilevel thresholding. IEEE Transaction on Image Processing IP-4 370-378.

Estimating melon yield for breeding processes by machine-vision processing of UAV images

A. Kalantar[1,2], A. Dashuta[1,3], Y. Edan[2], A. Dafna[4], A. Gur[4] and I. Klapp[1*]
[1]*Institute of Agricultural Engineering, Agricultural Research Organization (ARO), Volcani Center, Rishon-LeZion, Israel; iftach@volcani.agri.gov.il*
[2]*Department of Industrial Engineering & Management, Ben-Gurion University of the Negev, Beer Sheva 8410501, Israel*
[3]*School of Electrical Engineering, Tel Aviv University, Tel Aviv 69978, Israel*
[4]*Newe Ya'ar, ARO, Volcani Center, Ramat Ishay, Israel*

Abstract

Monitoring plants, for yield estimation in melon breeding, is a highly labor-intensive task. An algorithmic pipeline for detection and yield estimation of melons from top-view images of a melon's field is presented. The pipeline developed at the individual melon level includes three main stages: melon recognition, feature extraction, and yield estimation. For each region of interest classified as a melon, the melon features were extracted by fitting an ellipse to the melon contour. A regression model that ties the ellipse features to the melon's weight is presented. The modified R^2 value of the regression model was 0.94. Comparing yield estimation to ground truth, the average estimation error was 16%. The yield accuracy is highly dependent on the ellipse estimation accuracy, with promising results of only 4% error for the best ellipse-fitted melons.

Keywords: machine learning, CNN, active contour, melon, breeding, phenotyping, yield estimation

Introduction

Melon breeding requires a detailed account of accumulated yield and general yield distribution, in addition to melon size and location, which are essential for connecting yield to treatment/melon genetics. Manual plant monitoring for yield estimation is a highly labor-intensive task and thus expensive (Gongal *et al.*, 2015). Hence, automation of the process is beneficial. One proposed method for achieving an automated qualitative yield analysis is to apply computer-vision and machine-learning methods to color images of a top view of a melon field. Machine-learning is extensively used for fruit detection and recognition, using different recognition methods such as K nearest neighbors, K means with color information (Anisha *et al.*, 2013), (Qureshi *et al.*, 2017), Faster R-CNN (Bargoti and Underwood, 2017). Detection of almonds, mangoes and apples resulted in precision in the range of 0.7 to 0.9 (Bargoti and Underwood, 2017). Machine-learning techniques were also used to estimate yield loss such as a system that detected fruit on the orchard ground (Choi *et al.*, 2013). This work proposes a framework for detection and yield estimation of melons from color images, acquired from a digital camera mounted on a drone. Object detection under the highly variable outdoor conditions along with extraction of melon features related to yield, such as size, are detailed in the described algorithmic pipeline along with the results. The focus is on yield detection based on individual melons.

Materials and methods

Data acquisition
Data were acquired at an experimental agricultural research farm in Newe Ya'ar, Izrael valley, northern Israel. Latitude/longitude 32.718492/35.181951 respectively. Foliage coverage was reduced

DOI 10.3920/978-90-8686-888-9_47, © Wageningen Academic Publishers 2019

to improve image detection by stopping irrigation 1 week before the acquisition. The acquisition was performed at midday on 17 Aug 2017, at fruit ripening stage. RGB images were acquired from a Sony ILCE-5000 camera mounted on a drone (Quad-Copter) hovering about 15 m above the field, with the camera facing vertically downward. An area of 280×15 m was acquired. Before starting the image acquisition, 30 melons with different shapes, sizes and colors were randomly tagged in the field and used for ground-truth data. These melons were analyzed after image acquisition with a 'Tomato Analyzer' tool (Gonzalo *et al.*, 2009). The analysis provided a dataset of the ground-truth features of each melon, Additional similar data acquired in the following year (2018) were used to develop the statistical model to tie extracted section geometry to melon weight. The images were saved in .jpg format with a resolution of 3,064×5,456 pixels. Overall there were thousands of melons with high diversity. The variation in the dataset was intended to ensure the modeling of a robust and accurate detection algorithm that would generalize well under outdoor environmental conditions.

Data preparation

The algorithms included two machine-vision phases: a cascade object detector (COD) (Viola and Jones, 2001), and a neural network classifier (NNC). To train the COD, 3,031 negative training images (of background) and 1,933 positive training images (of melons) were created. NNC training was based on 1,383 positive (melon) and 1,349 negative (background) training images. All melons were randomly selected. Augmentation techniques were used to increase robustness, resulting in a training set with 44,256 positive images and 43,168 negative images. Augmentation was achieved by applying blurring, a rotation operation, and adding Gaussian noise, resulting in 31 additional new images for each original training image.

Algorithm

The proposed algorithm pipeline for automated yield prediction from RGB images of a melon field includes the following three sequential main stages: (1) melon recognition; (2) feature extraction; (3) yield estimation. Each stage uses results from the previous stage. The input for the system are RGB images of a melon field, and the output is a report that includes each melon's location and weight. The algorithm pipeline is illustrated in Figure 1.

Analysis

Detection performance was evaluated using recall and precision indicators. A fruit detection was considered to be a true positive if the predicted and ground-truth bounding box had an intersection over union (IoU) greater than a fixed threshold (Bargoti and Underwood, 2017b). Two different thresholds were examined. The first was 0.5, used in PASCAL Visual Object Classes (VOC) challenges (Everingham *et al.*, 2009), resulted in relatively small fruit size with respect to the image resolution, errors in detecting melons were received, causing them to be registered as false negative. To reduce the false negative misclassification, a smaller threshold of 0.2 was applied, which equates to a 58%

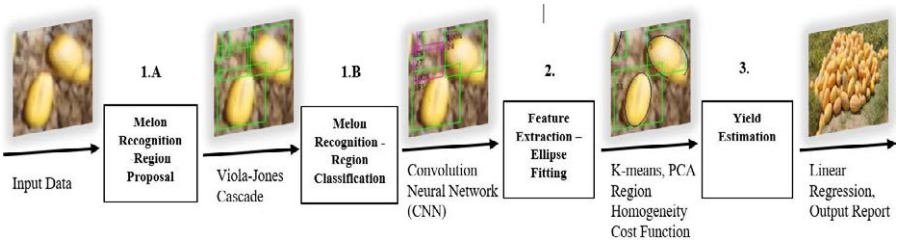


Figure 1. Algorithm pipeline for automated yield tracking.

overlap along each axis of the object; this was considered sufficient for the fruit-mapping application (Krizhevsky *et al.*, 2012).

Algorithm

Melon recognition implementation details

The melons make up only a small portion of typical melon field images. Therefore, to reduce computational effort, the recognition process was split into two sub stages: detection of proposed 'candidate' regions, followed by region classification. First, regions of interest (ROI) are derived, the proposal model runs in a sliding window over the entire input image and detects possible melon candidates. Since the detection of candidate regions requires an exhaustive search over the entire image, it is assigned to have low time complexity. It is also designed to have high recall, at the expense of low precision. Second, the candidates detected by the candidate region proposal model are classified by an algorithm, which is more computationally expensive, but compensates for the region proposal model's low precision (Dashuta and Klapp, 2018).

Candidate region proposal

The Viola-Jones face detector was chosen for the candidate region detection model: it proposes ROI suspected of containing a melon. The Viola-Jones detector, a COD, is an ensemble of a number of weak classifiers. A weak classifier is constructed by simply thresholding one co-ordinate of some feature vector of the concurrent window (Dalal and Triggs, 2005). The feature type selected for the detector is the histogram of oriented gradients (HOG) (Viola and Jones, 2001), which can capture the elliptical nature of the melon. The COD is arranged in stages with increasing complexity. The role of each stage is to decide whether the concurrent window is certainly NOT an object. If a stage decides that the concurrent window is not an object, the remaining stages are not evaluated. Hence, only true object windows trigger the entire cascade of stages. This mechanism ensures low time complexity of the model.

Region classification

A convolutional neural network (CNN) (Krizhevsky *et al.*, 2012) was used for the region classification model. The model was trained using the well-known 'transfer-learning' methodology (Li *et al.*, 2016). The learning process was initiated with a network that was pre-trained on the CIFAR-10 dataset (Krizhevsky *et al.*, 2012), a 10 category 32×32 pixels color image dataset. In this work, the final fully connected layer, which is essentially an image classifier based on a 64-dimensional feature vector produced by the hidden layers, was changed from a layer with output size of 10 labels (the original dataset has 10 classes) to one with an output size of 2 labels-melon and background. The training data were expanded using augmentation techniques (Wang *et al.*, 2014). Results revealed that the algorithm achieves 85.01% precision and 89.8% recall compared to the human identification used for the ground-truth measurement.

Feature extraction

This stage works on patches of the image, that have been detected as ROI and classified as melon. Each ROI was assumed to contain only one melon. For each melon, the exact location in the field and the features that were related to melon size were derived. These features helped to estimate the melon weight. The pre-knowledge that the shape of each individual melon can be well approximated by a spheroid (Heinzen *et al.*, 1998) implies that from a top view, the melon contour will be recognized as an ellipse. This enables translating the feature-extraction problem into an ellipse-fitting problem. Once the ellipse with the best fit to each detected melon is derived, the feature extraction is trivial. The ellipse partitions the patch into two regions: inner region (inside the ellipse) and outer region (in the proposed candidate region, outside the ellipse), using the following decision rule: pixels in

the inner region are classified as 'melon' and pixels in the outer region are classified as 'background'. Every possible ellipse can be parametrized by 5 parameters: semi-major axis (c), semi-minor axis (a), centroid x co-ordinate (x_0), centroid y co-ordinate (y_0) and angle of tilt (θ). Finally, the contour of every melon in the field can be approximated as an ellipse:

$$\frac{\left[(x-x_0)\cos(\theta)-(y-y_0)\sin(\theta)\right]^2}{a^2}+\frac{\left[(x-x_0)\sin(\theta)+(y-y_0)\cos(\theta)\right]^2}{c^2}=1 \quad (1)$$

Every point $x = (a,c,x_0,y_0,\theta)\in\gamma^5$ in the parameter space corresponds to a single ellipse (single solution) in the given ROI. To derive these parameters, first an initial ellipse that fits the contour is derived, then parametrization is conducted, followed by solution optimization using a cost function minimization problem (Dashuta and Klapp, 2018).

To estimate the yield, a model that ties the melon's geometry to its weight was determined. A spheroid model was applied using a 3D shape with same-sized width and depth axes (Figure 2).

Assuming that a typical melon has a spheroid shape, the parameters derived above from the fitted ellipse in the 2D image of the melon, in particular the sizes of the minor and major axes, should be correlated to the semi-height and semi-width of the melon. Using these two parameters, the melon weight is predicted. The regression model was built using information from 30 randomly selected individual melons which were both imaged by the UAV and measured for their weight and geometry in the laboratory. The best derived regression model was:

$$W = 0.1096653 + 0.003397929 \cdot c \cdot a^2 \quad (2)$$

Where W is the melon weight, a is the ellipse semi-height and c is the ellipse semi-width.

Following this model, fidelity to data points to the model was estimated using the adjusted R^2 figure of merit (Montgomery, 1997), resulting in a R^2_{Adj} of 0.94, and implying a strong fit for the specific dataset (for n=30, Figure 3). Another dataset of 135 melons obtained in the following year (July 2018) was used for cross-validation (Hawkins, 2004). The overall yield estimation from the ellipsoid parameter is presented in Figure 4.

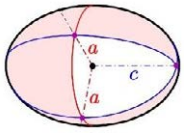


Figure 2. Ellipsoid model.

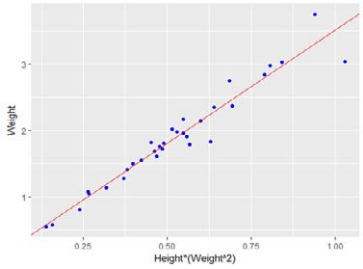


Figure 3. Weight and spheroid parameter regression results on training set analysis of regression.

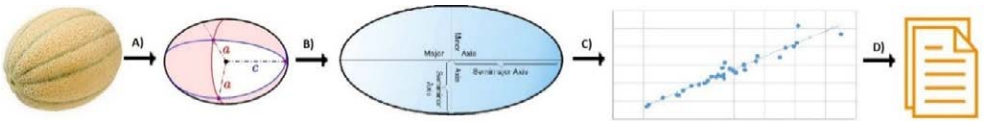


Figure 4. Yield estimation process. Stage (A) represents a melon by spheroid shape – use pre-knowledge of melon shape; stage (B) converts from 3D to 2D – the width and depth are minor axes and length is the major axis; stage (C) uses linear regression to predict yield of each melon; stage (D) generates a report for each melon location and yield estimation for the entire field.

The above method to fit an ellipse contour to the melon image was suggested. Feature extraction of such a contour is an ellipsoid section. Thus, assigning the ellipse parameters (Equation 1) in the proposed regression (Equation 2) results in an estimation of melon weight from post-processing of the UAV images.

The parameters in Equation 1 are given in pixels. Pixels were converted into centimeters by calculating the ground sample distance (GSD). The GSD was estimated from the images of a signboard that was placed next to the 30 randomly selected individual melons. A calibration ratio (1/GSD), was derived by dividing the number of pixels in the signboard image (along an edge), by the edge size in centimeters. This ratio was applied to the features of the melon marked in the image and located near the signboard. For all other melons, the averages of all of the ratios were calculated and applied.

Results

Precision vs recall of the experimental results is presented in Figure 5, revealing that the precision is retained through almost all of the recalled range. The detection algorithm was able to detect only 87% of the tagged fruit used for the calibration. The GSD metric was used to convert the pixel units into centimeters with a derived ratio of approximately 0.4 cm per pixel. Using this ratio, the actual yield was forecast. The result overestimated the actual yield by 16%. The error was due to the ellipse error and the error in the GSD estimation. The UAV's distance from the melon effect the image GSD. Wrong GSD resulted in estimation error in melon's size. By applying a more accurate GSD estimation based on the size of the local known targets for the 5 best recognized melons (accurate contour detection), the error was reduced to 4%.

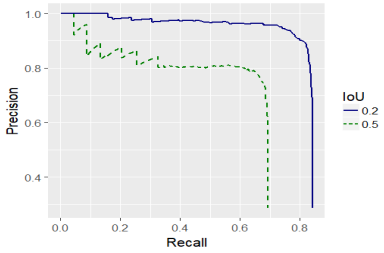


Figure 5. Experimental results precision vs recall curve, average precision for melon detection using different IoU thresholds.

Discussion

Yield accuracy strongly depends on the ellipse estimation accuracy. The weight estimation error for the best fitted melons was only 4%. Improvement of melon recognition and the fitting model can potentially result in an applicable method for estimation of melon yield from aerial surveys. Ongoing work is focused on improving recognition by using improved object-detection techniques, such as a faster R-CNN object detector. In addition, several models for improving the ellipse-fitting problem, such as the Chan-Vese model for active contours, are under examination. Since individual melons in the same field ripen at different rates, harvesting cannot be done in a single pass. Therefore, the field workers must sample every melon in the field to determine their ripeness. An automated yield analysis system could guide the field workers directly to the ripe melons and thus save time.

Conclusions

The above pipeline includes three main stages for yield detection based on calculation derived at the single melon level: melon recognition, feature extraction, and yield estimation. This scheme show promising capabilities. First, to recognize individual melon's couture. Second upon correct feature extraction predict melon weight with high accuracy. To improve the performances of the proposed pipe line the ongoing research focused on improving of learning process and possibly the active couture schema.

Acknowledgements

This work was partially supported by BARD Program number IS-4911-16 and by the Rabbi W. Gunther Plaut Chair in Manufacturing Engineering at Ben-Gurion University of the Negev.

References

Anisha. S, Divya. G, Shanu. S., 2013. A Survey of Computer Vision Methods for Counting Fruits and Yield Prediction, International Journal of Computer Science Engineering (IJCSE), 2, (6), 2319-7323

Bargoti S., and Underwood, J., 2017b. Image Segmentation for Fruit Detection and Yield Estimation in Apple Orchards, Journal of Field Robotics 34(6), 1039-1060.

Bargoti.S and Underwood, J., 2017a. Deep Fruit Detection in Orchards arXiv:1610.03677v2 [cs.RO] (last accessed February 2019).

Chan T.F. and Vese L.A., 2011. Active Contours without Edges, IEEE Transactions on Image Processing 10 (2), 266-277.

Choi.D, Suk-Lee.W, Ehsani. R., 2013. Detecting and counting citrus fruit on the ground using machine vision Paper No. 131591603, St Joseph, MI, USA: ASABE

Dalal N., and Triggs B., 2005. Histograms of Oriented Gradients for Human Detection, In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1 (Institute of Electrical and Electronics Engineers,), 886-893.

Dashuta A., and Klapp I., 2018. Melon Recognition in UAV Images to Estimate Yield of a Breeding Process, Light, Energy and the Environment (E2, FTS, HISE, SOLAR, SSL), OSA Technical Digest (Optical Society of America), paper ET4A.2. https://www.osapublishing.org/abstract.cfm?URI=EE-2018-ET4A.2 (last accessed February 2019).

Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A., 2010. The pascal visual object classes (voc) challenge. International Journal of Computer Vision, 88(2), 303-338.

Gongal A., Amatya S., Karkee M., Zhang Q., and Lewis K., 2015. Sensors and systems for fruit detection and localization: A review. In: Computers and Electronics in Agriculture 116, 8-19.

Gonzalo M.J., Brewer M., Anderson C., Sullivan D., Gray S., Knaap E., 2009. Tomato Fruit Shape Analysis Using Morphometric and Morphology Attributes Implemented in Tomato Analyzer Software Program, Journal of the American Society for Horticultural Science 134(1) 77-87.

Hawkins D.M., 2004. The Problem of Overfitting, Journal of Chemical Information and Modeling 44(1), 1-12.

Heinzen A., Shimmel C., Groppe R. and Davidson E.A., 1998. Apparatus for Removing Rind from Spheroidal Fruits and Vegetables, U.S. Patent 5,806,414, 15 Sep 1998.

Krizhevsky A., Sutskever I., and Hinton G.E., 2012. Imagenet Classification with Deep Convolutional Neural Networks, in Advances in Neural Information Processing Systems 25, (Lake Tahoe, CA, USA). pp 1097-1105.

Li Z., Song Y., Mcloughlin I., and Dai L., 2016. Compact Convolutional Neural Network Transfer Learning for Small-Scale Image Classification, in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp 2737-2741.

Montgomery D.C., 1997. Chapter 13 in Design and Analysis of Experiments (fourth ed.), Arizona, USA: John Wiley & Sons,

Qureshi. W. S, Payne. A, Walsh. K. B, Linker, R, Cohen, O, Dailey. M. N., 2016. Machine vision for counting fruit on mango tree canopies, Precision Agriculture 8, 224. https://doi.org/10.1007/s11119-016-9458-5

Viola P. and Jones M.J., 2001. Rapid object detection using a boosted cascade of simple features, In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1 (Institute of Electrical and Electronics Engineers), pp 511-518.

Wang J., Song Y., Leung T., Rosenberg C., Philbin J., Chen B., 2014. Learning Fine-grained Image Similarity with Deep Ranking, In the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 1386-1393.

Estimation of the leaf area index in maize based on UAV imagery using deep learning techniques

J. Martínez-Guanter*, G. Egea, M. Pérez-Ruiz and O.E. Apolo-Apolo
Aerospace Engineering and Fluids Mechanics Department, University of Seville, Ctra. Sevilla-Utrera km 1, 41013 Seville, Spain; martinezj@us.es

Abstract

Leaf area index (LAI) is a biophysical crop trait of great interest to maize breeders since it is directly related to crop productivity. Indirect methods based on gap fraction theory are frequently used for *in situ* LAI estimation. UAVs have allowed for crop information inference at greater spatial and temporal resolutions, while image analysis techniques generate increasingly accurate results as a useful output in agriculture. This paper describes a methodology to estimate LAI from RGB images taken with UAV in a maize breeding trial. A pre-trained artificial neural network has been used to automatically perform the LAI calculation. The output results were compared with ground direct LAI measurements. The first obtained results show that this approach can be a feasible and accurate methodology to estimate the LAI of maize breeding plots in a significantly shorter time and at a low cost.

Keywords: LAI estimation, machine learning, RGB images, phenotyping, image analysis

Introduction

Field phenotyping used to be a labour-intensive and time-consuming process with low efficiency, and subject to bias and inaccuracies (Araus and Cairns, 2014). A method to monitor and measure biophysical parameters of the crop in a fast, accurate, non-destructive and economical way is needed. In recent years, field phenotyping has been commonly understood to be the capacity for sensor-based data collection and processing, in a non-destructive and non-invasive way of plant properties (Andrade-Sanchez *et al.*, 2014; Araus and Cairns, 2014). Biophysical crop traits define variables of interest to quantify phenotypic factors such as biomass, leaf area index (LAI), the fractional vegetation cover (FVC) or the fraction of absorbed photosynthetically active radiation (FAPAR). Among these factors, the LAI has been widely used as an indicator of crop health status, its use of available nutrients, or the crop's agronomic management. Since LAI is defined as the leaf area of the canopy per unit area projected on the soil (Watson, 1947; Chen and Black., 1992), this index is a key biophysical variable required for modelling crop processes such as photosynthesis, radiation interception, energy balance or transpiration (Kalisperakis *et al.*, 2015; Bryson *et al.*, 1997). The two methods for LAI measurement, the direct measurement and the indirect LAI estimation, have been widely assessed and reviewed in the literature. Direct methods are the most accurate and helpful to validate indirect measurements, but they remain extremely time-consuming and prohibitively expensive when applied to large crop areas (Sonnentag *et al.*, 2007). Alternatively, in recent years, indirect assessment methods including plant canopy analyzers, hemispherical images and ceptometers have been developed (Jonckheere *et al.*, 2004). Chen and Black (1992) described the indirect estimation derived from optical sensors as the 'effective LAI', as these methods based on gap fraction are linking the canopy architecture with the LAI. Among these methods, the use of hemispherical digital RGB images has become popular since it has the practical advantages of its relatively low cost and low processing time. However, the repeatability in image capture still requires time and patience and the need to do it in-situ during the growing cycle can represent an obstacle that can be addressed in a complementary way with other methods.

DOI 10.3920/978-90-8686-888-9_48, © Wageningen Academic Publishers 2019

Recent advances in sensor equipment and platforms alongside developments in computer science offer novel prospects for a timely recognition and identification of crop traits (Moshou *et al.*, 2015). In recent years, UAVs (unmanned aerial vehicles), have become powerful sensor-bearing platforms for various agricultural applications. During the last decade, several studies have been carried out to obtain images and information mainly from multispectral sensors, in order to make indirect estimates of LAI from UAVs and satellites (Lee *et al.*, 2004; Yao *et al.*,2017). Recently, other authors have used these multispectral images with a multi-angle approach similar to the authors to estimate LAI and chlorophyll content in other crops such as potatoes (Roosjen *et al.*, 2018).
The non-independent relationship between the LAI and the FVC (Carlson and Ripley, 1997), enables the analysis of zenithal images obtained with a UAV-borne RGB sensor to lead to the calculation of the FVC and the subsequent estimation of the LAI. FVC, understood as the ratio of pixels corresponding to vegetation over the total number of pixels in an area (ranging from 0-1), requires effective segmentation of images using color-based discrimination in computer vision algorithms (Sadeghi-Tehran *et al.*, 2017). This segmentation and discrimination between plant and background soil depends to a great extent on natural illumination (Slaughter *et al.*, 2008) and on aspects such as the pre-processing of the image that have been studied in detail previously (Hamuda *et al.*, 2016). In the last few years, a set of automated analysis techniques have emerged largely due to increasing computational capacity. Among them, machine learning and deep learning methods have boosted the power of obtaining relevant insights from the precision agriculture domain (Kamilaris and Prenafeta-Boldú, 2018), from automated image analysis to applications in robotics and advanced statistical analysis. In this paper, the use of RGB images from UAVs to accurately estimate the LAI was assessed in the context of a maize breeding trial carried out near to Seville, Spain. To this end, a machine learning model based on deep learning techniques for gap fraction analysis and LAI derivation has been developed. The specific objectives, given this approach, were: (1) the development of a multi-angle top-down image analysis approach using deep learning techniques; (2) the automation of the FVC calculation process and generation of the relational FVC-LAI model; and (3) the comparison of the LAI obtained by the model with non-destructive allometric measurements taken in the field.

Materials and methods

Field test description

The trial was conducted in an experimental maize field (*Zea mays* L.) during the summer of 2018. The field is located 11 km from Seville, in southwestern Spain (37°27'38.61"N, 5°58'30.32"W). The crop had no irrigation limitations or stress, while the soil contained a topsoil with 1.3% organic matter. In the field, 18 small plots of two rows of maize each were delineated for testing (Figure 1A) with 2 varieties of maize, each with a different water stress tolerance range. Thus, plots 1, 6, 7, 9, 10, 11, 14 and 16 were more tolerant to stress, while plots 2, 3, 4, 5, 8, 12, 13, 15, 17 and 18 were more sensitive. This sensitivity to stress is relevant because it can significantly affect the LAI. The experiment was arranged with four direct measurement replications per plot. Each experimental plot took up an area of 4.50 m^2 (6×0.75 m), which accommodated 2 maize rows with 35 plants per row. Both the performance of the flights and the direct measurements of the LAI that are presented were carried out on DOY 193 (July 13th).

Direct LAI measurements

To estimate the leaf area of individual maize leaves in a non-destructive way, obtaining direct in-field measurements for calculating the allometric relationship is key. This allowed the validation of other non-destructive methods in trials where the plant material cannot be sampled, as is the case of a trial of varieties where the variation in the number of plants per sub-plot would alter the productivity results of the varieties.

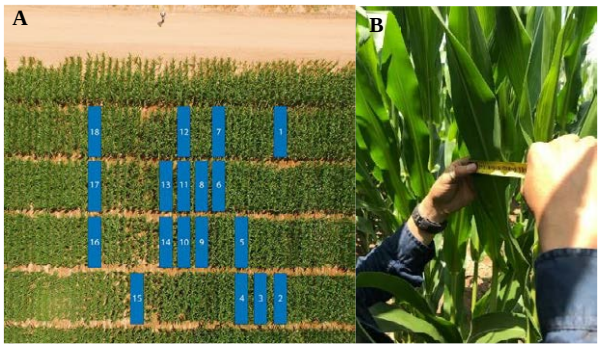


Figure 1. (A) Experimental maize field and the location of the assessed plots from a high-altitude UAV image. (B) direct measurements taken on the maize leaves (width and length parameters).

To measure leaf area (LA), four representative plants from each plot were selected to manually measure their leaves with a flexible tape graduated in mm (Figure 1B). For each leaf, the linear dimensions, length (L) and maximum width (W), were measured. A previous empirically obtained first-degree equation relation was used, (y = ax), where the independent variable was calculated as the product of W and L, being 0.7678 the slope of the linear relationship. These dimensions were used to calculate the leaf area of each leaf using the above mentioned allometric relationship. The total leaf area for each plant was calculated as the sum of unitary leaf areas for all plant leaves. The LAI for the plots was obtained according to the following expression:

$$LAI = \frac{PLA\ (m^2)}{PD\ (m^2)} \quad (1)$$

where PLA is the plant leaf area and PD is the plant density.

Image acquisition from UAV

Crop images were acquired by an RGB sensor embedded in the DJI Phantom 3 Pro commercial multirotor flight platform (DJI, Shenzen, China). The built in RGB sensor has a resolution of 4,000×3,000 pixels, with a focal length of 3.57 mm. The field of view (FOV) of the RGB sensor was 81° horizontal × 66° vertical. The flight was at a very low altitude over the crop (10 m), obtaining a ground sampling distance of 4.3 mm/pixel. The flight plan was designed in such a way that, for each plot, at least three images were obtained with three different angular perspectives, from the zenith, from the left side and from the right side (Figure 2A). In this way, the FVC varies in each of the images, while the LAI remains constant.

Image processing

Identification of row crops on images

For each of the 3 images with a different perspective corresponding to the same plot, the angular range in which the rows of the crop under study were found was calculated. For this, knowing the field of view (FOV) of the RGB sensor, the center of the image was determined as angle 0°. The distance in pixels from the center of the image to the center of the plot was quantified. The product of this distance in pixels by the Ground Sampling Distance (GSD) allowed to obtain the real distance in meters, and by the trigonometric relation between the height and the base of the triangle, the angle of capture with respect to the crop was obtained. This process was repeated for the 18 plots, determining in each of the images in which angular range the plot was. For this process, only the horizontal FOV of the sensor was used.

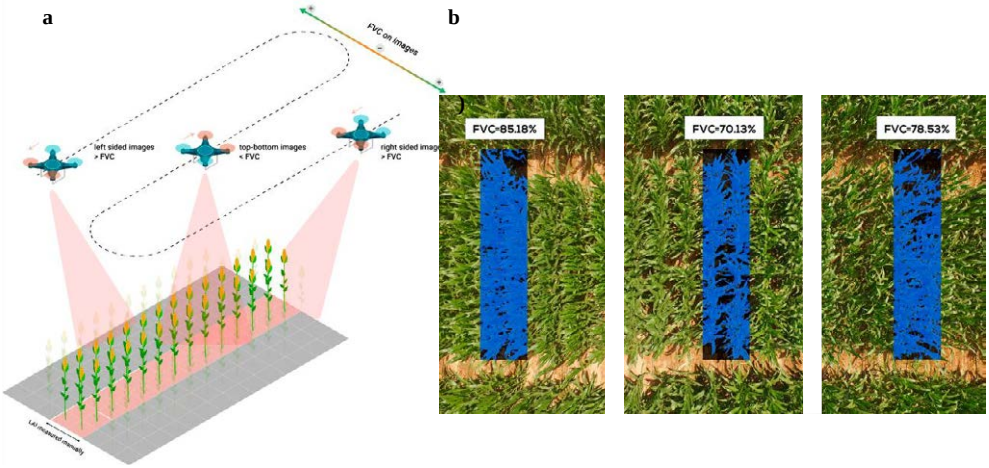


Figure 2. (A) Design scheme of the flight plan to obtain three images with different perspectives of the crop for the calculation of the FVC. (B) Example of the masking obtained through the DL model developed and variation of the FVC depending on the perspective in the same image.

FVC calculation based on Deep Learning model and LAI estimation

An instance segmentation model for masking based on pixel colour has been designed for the automatic calculation of the FVC in each of the images. This process of computer image analysis enables to identify each of the pixels in the image and associate them to a specific object (in this work, bare soil or vegetation), with the aim of segmenting and classifying it. The algorithm used to perform this instance segmentation was the Mask RCNN (He *et al.*, 2017) of the TensorFlow Object Detection API (Abadi *et al.*, 2015). This segmentation algorithm adds an additional binary mask to the Fast RCNN image object detection algorithm, creating an architecture that unifies it with a Fully Convolutional Network, widely used for semantic segmentation. The model was trained using an own dataset labelled by hand with 20 images and 2,000 labels (Figure 2B), in 5 steps with 500 epochs and a confidence of 0.9. Once trained, the model was used to automatically generate the masking of the images with the identified crop plots, separating between vegetation and soil. Subsequently, the FVC was obtained by dividing the number of pixels in the mask by the total number of pixels in the image. This processing was done on a computer with an NVIDIA GTX1080Ti CUDA-enabled GPU (NVIDIA Technologies, Santa Clara, CA, USA). Once the FVC was obtained using the ML technique described above, the LAI was estimated using a simple predictive model. To feed this model based on a decision tree classifier, LAI data obtained with CAN-EYE (INRA, Avignon, France) from the analysis of the same images obtained in the trial was used. Once this initial dataset with FVC:LAI structure was obtained, the prediction model was generated which received as input the FVC value for a given angular range.

Results

Direct LAI measurements

As stated in the last section, to obtain an allometric relationship of maize leaves is of great interest for validating other indirect measurement methods. In this paper, a simple linear regression equation with the capacity to explain 96% of the variability in observed single leaf area was developed by taking the linear dimensions, length (L) and maximum width (W) of a plant. The resulting relationship is

a first-degree equation, $y = a \times x$, where the independent variable is calculated as the product of W and L, being 0.77 the slope of the linear relationship. Allometric methods have been successfully used for many vegetation types and can be applied to larger areas. In the measurements made by hand on the total of the leaves of 4 corn plants of each of the 18 plots, the results were obtained as shown in Table 1 below.

FVC calculation results

Obtaining FVC from RGB images obtained from UAV has been carried out using Deep Learning techniques. The model used to generate the segmentation of the crop and background and then generate a mask on which to calculate the FVC was evaluated visually, by comparing the original and segmented images. First of all, the same plot has been detected manually in three different images. In each of them, the angular range from the camera has been determined from the FOV and the distance to the centre of the image. Once the mask has been generated using the automatic model Mask RCNN, the FVC was calculated by isolating the pixels corresponding to the maize crop. The accuracy of the segmentation was evaluated by comparing the values with those obtained from the manual segmentation using ImageJ software (Rueden *et al.*, 2017). An error function (Equation 2) was created in which the reference mask values used as reference were binaries (1 for vegetation or 0 for background) and compared pixel by pixel (i,j) with the mask generated by the developed model P_m, in a similar way to that developed by Sadeghi-Tehran *et al.* (2017), obtaining an outcome value of 0.93 out of 1 (in which 1 represents a perfect match between the reference and the masked image).

$$E_{seg} = \frac{\Sigma_{ij}(P_{m(i,j)} \cap P_{R(i,j)})}{\Sigma_{ij}(P_{m(i,j)} \cup P_{R(i,j)})} \quad (2)$$

Finally, the FVC values calculated by the model were used again as inputs in an associative function created from a predictive model to estimate the LAI. An angular variable corresponding to the calculated perspective of the plot on the image accompanied this FVC as input.

As seen in the first results obtained (Table 1), this methodology shows promise, with a good approximation to the LAI measured manually. In all the cases, the estimated LAI was higher than the true LAI measured directly, obtaining an average error of 0.366 in LAI estimation. This may be due to the estimation model itself, in which the weights of the variables may have been adjusted and the leaf occlusion variable is being over-compensated.

Conclusions

The LAI is a very relevant biophysical factor to define the state of the crop, although its precise estimation still presents some challenges in the field of precision agriculture. This work presents a methodology for the calculation of the fraction covered from RGB images captured with UAV, and later analysed by means of machine learning methods. Although the obtained results are still not as precise as the use of hemispheric images in reference software designed for this purpose like CAN-EYE, there is no doubt that the computing capacity and the refinement of the algorithms will give better results in future work. The authors consider that the use of this type of methodology for LAI estimation in a fast and cheap way can be a helpful and reliable tool in high-throughput phenotyping techniques, and they will continue to work along these research lines in the future.

Table 1. Summary of the values obtained by direct measurement of the LAI (True LAI) and the average values estimated from 3 zenithal images of the UAV.

		True LAI		Estimated LAI		Difference (Avg. Estimated – Avg. True)	Percentage Difference (Difference/True LAI) × 100
Water stress	Plot	Avg.	std	Avg.	std		
Tolerant	1	6.04	0.17	6.37	0.61	0.32	5.30
Sensitive	2	6.68	0.29	6.92	0.20	0.23	3.44
Sensitive	3	6.50	0.59	6.37	0.36	-0.13	2
Sensitive	4	6.18	0.45	6.57	0.71	0.39	6.31
Sensitive	5	6.84	0.59	6.90	0.12	0.06	0.88
Tolerant	6	6.29	0.36	6.53	0.31	0.24	3.82
Tolerant	7	6.36	0.73	6.76	0.90	0.40	6.29
Sensitive	8	7.54	0.33	7.63	0.42	0.08	1.06
Tolerant	9	7.53	0.33	7.70	0.31	0.17	2.26
Tolerant	10	5.39	0.89	5.86	0.75	0.47	8.72
Tolerant	11	6.41	1.48	6.65	1.64	0.25	3.90
Sensitive	12	5.49	1.06	6.01	1.30	0.53	9.65
Sensitive	13	5.90	0.55	6.23	0.80	0.34	5.76
Tolerant	14	6.18	0.40	6.45	0.45	0.72	11.65
Sensitive	15	6.09	1.39	6.82	1.10	0.73	11.99
Sensitive	16	4.83	0.70	5.73	0.54	0.90	18.63
Sensitive	17	5.79	0.67	6.21	0.64	0.42	7.25
Sensitive	18	5.84	0.42	6.31	0.40	0.47	8.05

Acknowledgements

This work was made possible through the support of the Spanish Ministry of Economic and Competence (PROJECT: GL2016-78964-R). The authors would also like to thank the Ministry of Science, Innovation and Universities for their support of the Torres-Quevedo grant program PTQ-17-09506.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., *et al.,* 2016. Tensorflow: A system for large-scale machine learning. In 12th {USENIX} Symposium on Operating Systems Design and Implementation 16) (pp. 265-283).

Andrade-Sanchez, P., Gore, M. A., Heun, J. T., Thorp, K. R., Carmo-Silva, A. E., French, A. N., *et al.,* 2014. Development and evaluation of a field-based high-throughput phenotyping platform. Functional Plant Biology, 41 (1), 68-79.

Araus, J.L., Cairns, J.E., 2014. Field high-throughput phenotyping: The new crop breeding frontier. Trends in Plant Science 19, 52-61.

Bryson, R. J., Paveley, N. D., Clark, W. S., Sylvester-Bradley, R., & Scott, R. K., 1997. Use of in-field measurements of green leaf area and incident radiation to estimate the effects of yellow rust epidemics on the yield of winter wheat. European Journal of Agronomy, 7 (1-3), 53-62.

Carlson, T. N., & Ripley, D. A., 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote sensing of Environment, 62 (3), 241-252.

Chen, J. M., Black, T. A., 1992. Defining leaf area index for non-flat leaves. Plant, Cell & Environment, 5 (4), 421-429.

Hamuda, E., Glavin, M., Jones, E., 2016. A survey of image processing techniques for plant extraction and segmentation in the field. Computers and Electronics in Agriculture 125, 184-199.

He, K., Gkioxari, G., Dollár, P., & Girshick, R., 2017. Mask r-cnn. In IEEE International Conference on Computer Vision (ICCV), (pp. 2980-2988). IEEE.

Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M., *et al.,* 2004. Review of methods for in situ leaf area index determination: Part I. Theories, sensors and hemispherical photography. Agricultural and forest meteorology, 121 (1-2), 19-35.

Kalisperakis, I., Stentoumis, C., Grammatikopoulos, L., & Karantzalos, K., 2015. Leaf area index estimation in vineyards from UAV hyperspectral data, 2D image mosaics and 3D canopy surface models. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 40(vol. 1), page 299.

Kamilaris, A., & Prenafeta-Boldú, F. X., 2018. Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70-90.

Lee, K. S., Cohen, W. B., Kennedy, R. E., Maiersperger, T. K., & Gower, S. T., 2004. Hyperspectral versus multispectral data for estimating leaf area index in four different biomes. Remote Sensing of Environment, 91 (3-4), 508-520.

Moshou, D., Pantazi, X.E., Oberti, R., Bravo C., West, J., Ramon, H., *et al.*; 2015. Crop health condition monitoring based on the identification of biotic and abiotic stresses by using hierarchical self-organizing classifiers. In J V Stafford (Ed.) Proceedings of the 10th European Conference on Precision Agriculture, Precision Agriculture '15. Wageningen, the Netherlands: Wageningen Academic Publishers. pp 619-626

Roosjen, P. P., Brede, B., Suomalainen, J. M., Bartholomeus, H. M., Kooistra, L., & Clevers, J.G., 2018. Improved estimation of leaf area index and leaf chlorophyll content of a potato crop using multi-angle spectral data–potential of unmanned aerial vehicle imagery. International Journal of Applied Earth Observation and Geoinformation, 66, 14-26.

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T. *et al.,* 2017. ImageJ2: ImageJ for the next generation of scientific image data, BMC Bioinformatics 18 (1) 529, doi:10.1186/s12859-017-1934-z.

Sadeghi-Tehran, P., Virlet, N., Sabermanesh, K., & Hawkesford, M. J., 2017. Multi-feature machine learning model for automatic segmentation of green fractional vegetation cover for high-throughput field phenotyping. Plant methods, 13 (1) 103.

Slaughter, D.C., Giles, D.K., Downey, D., 2008. Autonomous robotic weed control systems: A review. Computers and Electronics in Agriculture. 61, 63-78.

Sonnentag, O., Talbot, J., Chen, J. M., & Roulet, N. T., 2007. Using direct and indirect measurements of leaf area index to characterize the shrub canopy in an ombrotrophic peatland. Agricultural and Forest Meteorology, 144 (3-4), 200-212.

Watson, D. J., 1947. Comparative Physiological Studies on the Growth of Field Crops: I. Variation in Net Assimilation Rate and Leaf Area between Species and Varieties, and within and between Years. Annals of Botany XI (vol 41), 41-76.

Yao, X., Wang, N., Liu, Y., Cheng, T., Tian, Y., Chen, Q., *et al.,* 2017. Estimation of wheat LAI at middle to high levels using unmanned aerial vehicle narrowband multispectral imagery. Remote Sensing, 9 (12) 1304.

A UAV-based system for monitoring crop growth in wheat, barley and triticale phenotyping field trials

J.M. Peña[1*], F.J. Ostos-Garrido[2], J. Torres-Sánchez[2], F. Pistón[2] and A.I. de Castro[2]
[1]*Institute of Agricultural Sciences, Spanish National Research Council (CSIC), C/ Serrano 115-B, 28006 Madrid, Spain; jmpena@ica.csic.es*
[2]*Institute for Sustainable Agriculture, Spanish National Research Council (CSIC), Avda. Menéndez Pidal s/n, 14004 Córdoba, Spain*

Abstract

This research evaluated the capability of an unmanned aerial vehicle (UAV) –based system to monitor the plant heights of several wheat, barley and triticale field trials over time, by applying a customised object-based image analysis algorithm to the digital surface models created with the structure-from-motion technique. Comparisons between UAV-based and ground-based measurements over time reported an overall coefficient of determination (R^2) of 0.917 and an average root mean square (RMS) error of 0.10 m. With a target of field phenotyping, a ranking of genotypes in term of plant heights and crop growth was accurately provided, which can facilitate the decision making process in cereal breeding programs.

Keywords: 3D digital surface model (DSM), plant height, remote sensing, object-based image analysis, cereal breeding program

Introduction

Research on crop production is based to a certain extent on the analysis of the genotype-phenotype relationships. While genotyping tools have made great progress in the last 20 years and have led to a better understanding of the plant genome, the phenotyping tools have not been able to meet the needs of high-performance, low-cost and non-destructive plant data collection and analysis (Araus and Cairns, 2014). However, in recent years, new high-throughput phenotyping platforms are undergoing a rapid evolution that could significantly improve understanding between the association between genes and phenotype in the actual field conditions, including both on-ground platforms (Busemeyer *et al.*, 2013) and aerial vehicles (Shi *et al.*, 2016). Of all systems, unmanned aerial vehicles (UAVs) are taking relevance as a tool for phenotyping because of some advantages compared to other platforms (Yang *et al.*, 2017). UAVs are a low-cost and reliable platform for taking frequent remote images that allow observation of the variation of plant trials over crop development (Torres-Sánchez *et al.*, 2014). Furthermore, the capacities of the UAV to use a wide range of sensors, to operate at low flight altitude and to capture many consecutive images with high overlaps provide high-resolution spatial, spectral and three-dimension (3D) information of the studied plants. 3D data is obtained from the digital surface model (DSM) created by applying the structure-from-motion technique (Nex and Remondino, 2014) and, as a result of this process, the physical and morphological characteristics of the observed plants or trees can be potentially quantified (Jiménez-Brenes *et al.*, 2017; Moeckel *et al.*, 2018). In the specific case of winter cereals, this technology has been applied for assessing plant height in barley (Bendig *et al.*, 2014) and sorghum (Watanabe *et al.*, 2017) and growth status in winter wheat (Du and Noguchi, 2017), but to authors´ knowledge, there are no applications that assess its performance to monitor an experiment of multiple varieties of cereals with diverse phenotypic characteristics. In this work, the field trial of a breeding program designed to determine lignocellulosic bioethanol potential in cereals, in which it is essential to determine plant growth, was used to assess the capability of a UAV-based system to determine the height of the plants over time.

DOI 10.3920/978-90-8686-888-9_49, © Wageningen Academic Publishers 2019

Materials and methods

A field trial with sixty-six accessions belonging to the species *Hordeum vulgare* (barley, 21 accessions), *Triticum aestivum* (bread wheat, 24 accessions), *Triticum durum* (durum wheat, 11 accessions) and *x Triticosecale* (triticale, a hybrid of wheat and rye, 10 accessions) was established at the experimental station of the Institute for Sustainable Agriculture Center in Cordoba, Spain (co-ordinates 37.859 N, 4.796 W, datum WGS84). The experiment was sown on 15th November 2013 following a completely randomized block design with three replications. Each block counted in 78 plots distributed in ten rows, with inter-plot distance of 0.30 m and inter-row distance of 0.50 m. Each plot included four plants at approximately 0.15 m apart. A quadrocopter UAV model md4-1000 (microdrones GmbH, Siegen, Germany) was used to collect the multi-temporal set of aerial images by using a still point-and-shoot visible-range camera, model Olympus PEN E-PM1 (Olympus Corporation, Tokyo, Japan) (Figure 1). The camera acquired 12-megapixel images in red, green and blue (RGB) bands at 14 mm focal length. The UAV route was configured for flight at 3 m/s and at 10 m flight altitude over the ground, and for taking down-facing photos at an interval of 1 s to achieve a side overlap of 60% and a forward overlap of 90%. At this flight altitude, the spatial resolution of the RGB images was 3.81 mm/pixel of ground sampling distance. The UAV system collected the remote images on seven different dates: (1) 17 February (94 days after sowing, DAS), (2) 27 February (104 DAS), (3) 14 March (119 DAS), (4) 25 March (130 DAS), (5) 7 April (143 DAS), (6) 25 April (161 DAS), and (7) 9 May (175 DAS). The set of UAV images were processed to generate the ortho-mosaics of the experimental field at each study date (Figure 2). Additionally, the remote images were also used to build the digital surface model (DSM) of the field by applying the structure-from-motion technique (Jay *et al.*, 2015). The DSM is a 3D polygon mesh that reflects the irregular geometry of the ground and contains information of the plant heights (Figure 3). Agisoft PhotoScan Professional software (Agisoft LLC, St. Petersburg, Russia) was used to generate these geo-spatial products.

The ortho-mosaicked images enabled visual identification of each one of the 234 trial plots, which were manually defined over the image and saved as a vector file. Then, a customized algorithm was created with the eCognition Developer software (Trimble Geospatial, Munich, Germany) to analyze the images and DSMs of each studied date by using an object-based approach after image



Figure 1. The UAV with the RGB camera flying over the experimental field at 10 m altitude in the date 1 (94 DAS).

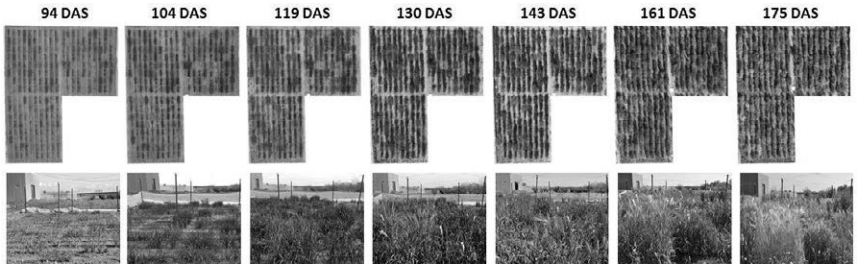


Figure 2. Aerial ortho-mosaicked images (in the top) and on-ground pictures of the experiment (in the bottom) at each study date. DAS: days after sowing.

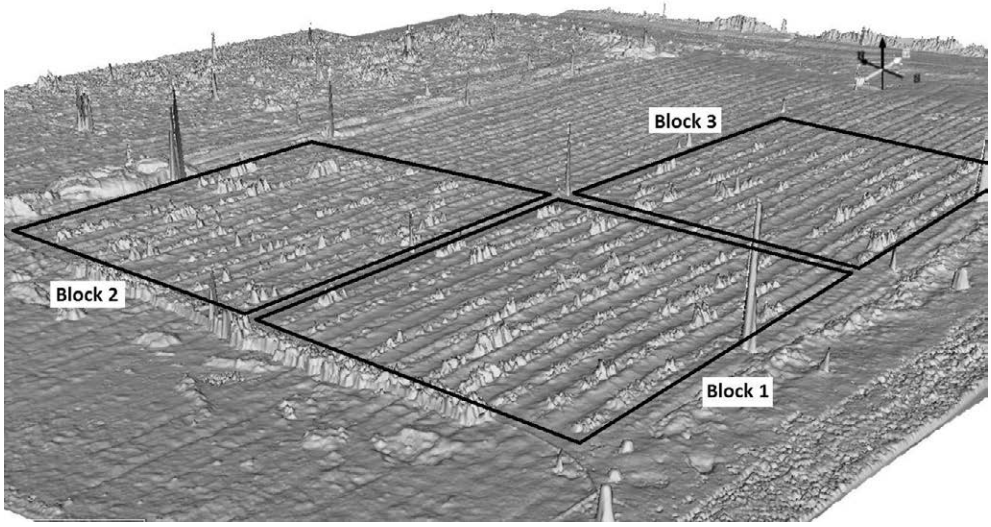


Figure 3. The Digital Surface Model (DSM) of the experimental field with 3D information of the overflown area in the date 1 (17th February).

segmentation (Blaschke *et al.*, 2014). The algorithm was specifically programmed to run in a fully automatic manner without the need for user intervention, and with the ability to sequentially discriminate the vegetation fraction of every trial plot over time by applying the Otsu thresholding method described in Torres-Sánchez *et al.* (2015a). Once the crop objects were classified in each plot, the algorithm computed the central co-ordinates, the crop height and the relative position of every plot within the experiment design. The height of every crop object was obtained by comparing its maximum DSM value to the average DSM values of the surrounding bare soil objects. Finally, the customized algorithm automatically exported all the trial plot data as a table file (e.g. CSV or ASCII format) for further descriptive and statistical analysis. Data analysis was conducted with the statistical software JMP version 10 software (SAS Institute Inc., Cary, NC, USA). Performance of the UAV-based phenotyping system was evaluated by comparing UAV-based and manual on-ground measurements over time, and thus determining the accuracy of the plant height estimations. The actual plant heights were manually measured on-ground with a tape measure over an aluminum bar at the same dates of the UAV flights. Plant height values were averaged for each trial plot.

Results and discussion

Given the multitude of screened genotypes, many different phenotypes were observed in the study trial plots in terms of plant heights. The averaged plant heights collected on ground through the crop development ranked from minimum values of 0.38-0.40 m in the cases of the *H. vulgare* CP20 (Franka), the *T. durum* TP23 (Avalon) and the *T. aestivum* TP27 (Renan) genotypes, to maximum values of 0.92-0.99 m in the cases of the *x Triticosecale* TS58 (Wapiti) and TS78 (Juanillo 95) and the *T. durum* TP20 (Rugby) genotypes. At the level of plant species, averaged heights of *T. aestivum* and *x Triticosecale* (0.62 m and 0.81 m, respectively) were significantly smaller and higher, respectively, than the other species, while *H. vulgare* and *T. durum* did not show significant differences in averaged plant heights (0.66 m and 0.70 m, respectively).

The UAV-based phenotyping system computed the plant heights of every trial plot over time (Figure 4A). The height data were automatically retrieved from the DSMs created with the RGB images collected at each one of the seven UAV flights and applying the customized OBIA procedure developed in this investigation. The system measured a range of values between 0.04 m in the first date to 1.63 m in the last date. In addition, three different growth patterns were also detected (Figure 4B). The plants were identified as prostrate, semi-erect and erect during the tillering stage, and as semi-erect and erect at maturity. The accuracy of the estimates was evaluated by direct comparison with on-ground manual measurements (Figure 5). The UAV-based system under-estimated the ground-true measurements in most cases, computing values between 27% (date 1, 94 DAS) and 6% (date 7, 175 DAS) lower than the real plant heights on average. Despite these under-estimates, high correlations were determined between UAV-based and ground-based measurements over time, reporting a global R^2 of 0.917 and an average RMS of 0.10 m. Specifically for each studied date, the best estimates were obtained on the date 3 (119 DAS), with a R^2 of 0.894 and RMS of 0.05 m, and the lowest determinations corresponded to the date 6 (161 DAS), with a R^2 of 0.615 and RMS of 0.13 m. The UAV-based measurements were smaller than the ones manually taken on the ground. The UAV-based system computed the plant heights from the DSM created after combining numerous overlapped remote images, which usually lacks the precision to detect higher parts of the plants such as the spikes, which are observed during manual on-ground measurements. In addition, on-ground measurements are usually an average value of the maximum height of the plants, which is a subjective measure and depends on the experimenter. The under-estimates reported in the investigation were

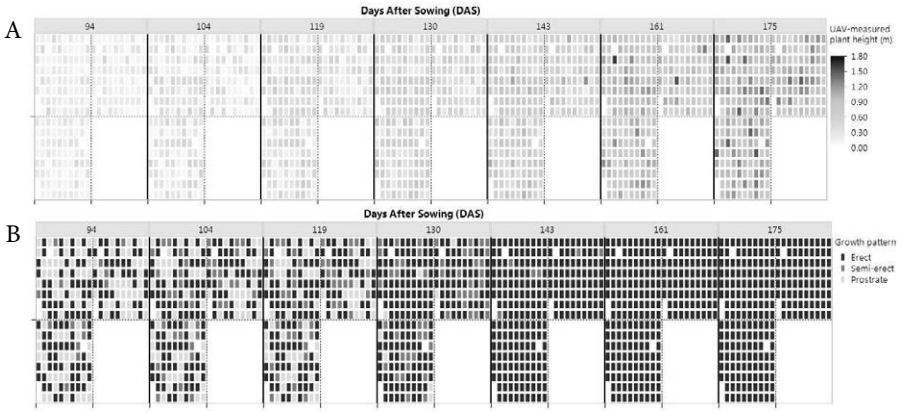


Figure 4. Plant variability measured by the UAV-based phenotyping system over time: (A) plant heights, and (B) growth pattern.

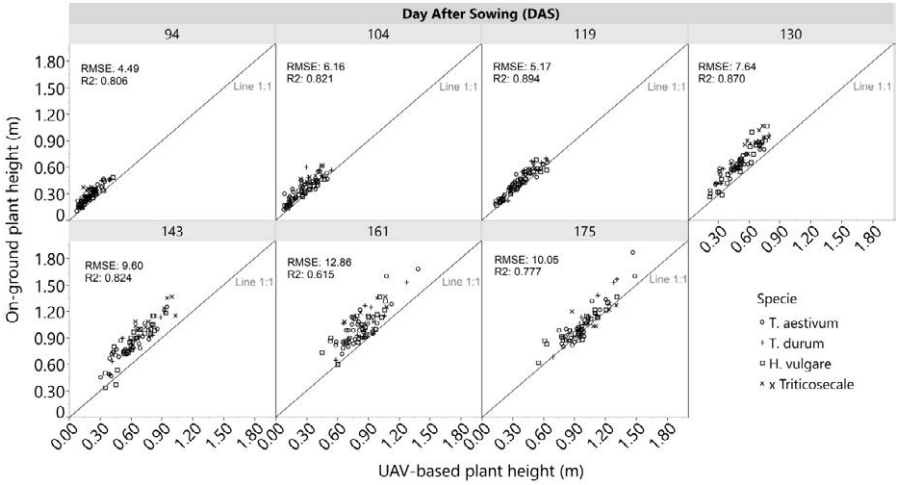


Figure 5. UAV-based estimated height vs observed height as affected by study date.

in line with the ones obtained in Bendig *et al.* (2014), who obtained a coefficient of determination of 0.92 in barley fields by using a UAV-based RGB imaging platform, and slightly lower than in Ehlert *et al.* (2008), who obtained coefficients of determination of 0.93 to 0.99 in winter rape, rye and wheat by using a ground based non-destructive laser rangefinder. In addition, Busemeyer *et al.* (2013) reported coefficients of determination of 0.97 and 0.92 in triticale by using a ground-based multi-sensor platform and a UAV-based system, respectively. This degree of under-estimation was also reported in previous case studies in several tree crops with larger dimensions such as vineyard (de Castro *et al.*, 2018), olive trees (Torres-Sánchez *et al.*, 2015b), poplar (Peña *et al.*, 2018) and almond trees (Torres-Sánchez *et al.*, 2018).

Conclusions

The system computed the plant heights of 234 trial plots at seven different dates over crop development with high accuracy, which can drastically reduce the time needed for manual on-ground measurements. Consequently, differences of plant heights among numerous genotypes of triticale, wheat and barley species were reported, which provides useful non-destructive information to predict crop growth and crop production as a valuable resource for investigations in cereal breeding programs.

Acknowledgements

This research was partly financed by the AGL2017-83325-C4 and the AGL2011-22596 Projects (Spanish Ministry of Economy, Industry and Competitiveness – MINECO, and FEDER funds). Research of Dr. Peña, Mr. Ostos-Garrido and Dr. de Castro was supported by the Ramon y Cajal (RYC-2013-14874), FPI (grant BES-2012-457 052455) and the Juan de la Cierva (IJCI-2015-24811) Programs of the Spanish MINECO funds, respectively. We thank Tadeo Bellot for his help in field and laboratory work.

References

Araus, J.L., Cairns, J.E., 2014. Field high-throughput phenotyping: the new crop breeding frontier. Trends in Plant Science, 19, 52-61. https://doi.org/10.1016/j.tplants.2013.09.008

Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S., Bareth, G., 2014. Estimating Biomass of Barley Using Crop Surface Models (CSMs) Derived from UAV-Based RGB Imaging. Remote Sensing, 6, 10395-10412. https://doi.org/10.3390/rs61110395

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., *et al.,* 2014. Geographic Object-Based Image Analysis – Towards a new paradigm. ISPRS Journal of Photogrammetry and Remote Sensing, 87, 180-191. https://doi.org/10.1016/j.isprsjprs.2013.09.014

Busemeyer, L., Mentrup, D., Möller, K., Wunder, E., Alheit, K., Hahn, V., *et al.,* 2013. BreedVision – A Multi-Sensor Platform for Non-Destructive Field-Based Phenotyping in Plant Breeding. Sensors, 13, 2830-2847. https://doi.org/10.3390/s130302830

De Castro, A.I., Jiménez-Brenes, F.M., Torres-Sánchez, J., Peña, J.M., Borra-Serrano, I., López-Granados, F., 2018. 3-D Characterization of Vineyards Using a Novel UAV Imagery-Based OBIA Procedure for Precision Viticulture Applications. Remote Sensing, 10(4), 584. https://doi.org/10.3390/rs10040584

Du, M., Noguchi, N., 2017. Monitoring of Wheat Growth Status and Mapping of Wheat Yield's within-Field Spatial Variations Using Color Images Acquired from UAV-camera System. Remote Sensing, 9(3), 289. https://doi.org/10.3390/rs9030289

Ehlert, D., Horn, H.-J., Adamek, R., 2008. Measuring crop biomass density by laser triangulation. Computers and Electronics in Agriculture, 61, 117-125. https://doi.org/10.1016/j.compag.2007.09.013

Jay, S., Rabatel, G., Hadoux, X., Moura, D., Gorretta, N., 2015. In-field crop row phenotyping from 3D modeling performed using Structure from Motion. Computers and Electronics in Agriculture, 110, 70-77. https://doi.org/10.1016/j.compag.2014.09.021

Jiménez-Brenes, F.M., López-Granados, F., de Castro, A.I., Torres-Sánchez, J., Serrano, N., Peña, J.M., 2017. Quantifying pruning impacts on olive tree architecture and annual canopy growth by using UAV-based 3D modelling. Plant Methods, 13, 55. https://doi.org/10.1186/s13007-017-0205-3

Moeckel, T., Dayananda, S., Nidamanuri, R.R., Nautiyal, S., Hanumaiah, N., Buerkert, A., *et al.,* 2018. Estimation of Vegetable Crop Parameter by Multi-temporal UAV-Borne Images. Remote Sensing, 10(5), 805. https://doi.org/10.3390/rs10050805

Nex, F., Remondino, F., 2014. UAV for 3D mapping applications: a review. Applied Geomatics, 6, 1-15. https://doi.org/10.1007/s12518-013-0120-x

Peña, J.M., de Castro, A.I., Torres-Sánchez, J., Andújar, D., Martín, C.S., Dorado, J., *et al.,* 2018. Estimating tree height and biomass of a poplar plantation with image-based UAV technology. AIMS Agriculture and Food, 3, 313-326. https://doi.org/10.3934/agrfood.2018.3.313

Shi, Y., Thomasson, J.A., Murray, S.C., Pugh, N.A., Rooney, W.L., Shafian, S. *et al.,* 2016. Unmanned Aerial Vehicles for High-Throughput Phenotyping and Agronomic Research. PLOS ONE, 11, e0159781. https://doi.org/10.1371/journal.pone.0159781

Torres-Sánchez, J., de Castro, A.I., Peña, J.M., Jiménez-Brenes, F.M., Arquero, O., Lovera, M., *et al.,* 2018. Mapping the 3D structure of almond trees using UAV acquired photogrammetric point clouds and object-based image analysis. Biosystems Engineering, 176, 172-184. https://doi.org/10.1016/j.biosystemseng.2018.10.018

Torres-Sánchez, J., López-Granados, F., Peña, J.M., 2015a. An automatic object-based method for optimal thresholding in UAV images: Application for vegetation detection in herbaceous crops. Computers and Electronics in Agriculture, 114, 43-52. https://doi.org/10.1016/j.compag.2015.03.019

Torres-Sánchez, J., López-Granados, F., Serrano, N., Arquero, O., Peña, J.M., 2015b. High-Throughput 3-D Monitoring of Agricultural-Tree Plantations with Unmanned Aerial Vehicle (UAV) Technology. PLOS ONE, 10, e0130479. https://doi.org/10.1371/journal.pone.0130479

Torres-Sánchez, J., Peña, J.M., de Castro, A.I., López-Granados, F., 2014. Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. Computers and Electronics in Agriculture, 103, 104-113. https://doi.org/10.1016/j.compag.2014.02.009

Watanabe, K., Guo, W., Arai, K., Takanashi, H., Kajiya-Kanegae, H., Kobayashi, M., *et al.,* 2017. High-Throughput Phenotyping of Sorghum Plant Height Using an Unmanned Aerial Vehicle and Its Application to Genomic Prediction Modeling. Frontiers in Plant Science, 8: 421. https://doi.org/10.3389/fpls.2017.00421

Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., *et al.,* 2017. Unmanned Aerial Vehicle Remote Sensing for Field-Based Crop Phenotyping: Current Status and Perspectives. Frontiers in Plant Science, 8: 1111. https://doi.org/10.3389/fpls.2017.01111

The application of structure from motion techniques in late-season corn damage

J.M. Prince Czarnecki*, S. Samiappan and L.A. Hathcock
Geosystems Research Institute, Mississippi State University, Mississippi State, Mississippi, 39762, USA; joby.czarnecki@msstate.edu

Abstract

'Green snap' occurs when corn stalks are broken in high wind. Late in the season, green snap produces yield losses that are immediate and irreversible. Losses are proportional to the percentage of plants lost, thus estimation of damage extent is important for calculating economic loss. Unmanned aerial vehicles are capable of creating a 3D-model, called a digital surface model, of the crop canopy using structure from motion techniques. The objective of this study was to compare image- versus structure-based calculations of damaged area. The structure-based calculations were more reflective of the field assessment. While the latter approach showed promise, more development of protocols is needed for reliability and operational success.

Keywords: unmanned aerial systems, crop damage, green snap, wind damage

Introduction

As the popularity of unmanned aerial vehicles (UAV) increases, one obvious application of the technology for agriculture is to perform rapid assessment of crop damage. Compared to a human, UAVs have the potential to cover land more quickly, and gather more comprehensive information. This is particularly true where large areas of damage are present, or where crops are sufficiently mature as to make visual inspection of the entire field difficult. As early as the mid-70s, Towery *et al.* (1975) identified the potential for the use of aerial photography in delineating areas of crop damage. Subsequent experiments have further demonstrated the validity of airborne imagery to detect damage and stand loss in corn resulting from naturally-occurring (Peters *et al.*, 2000) and simulated (Erickson *et al.*, 2004) damage in corn.

Despite improvements in genetics, lodging remains an issue in corn production systems (Duvick, 2005). Wind is one natural causal agent for lodging of corn that has the potential to create unexpected economic hardship for growers. Growth stage at the time of a high wind event dictates the ability of the crop to recover, and more importantly, the impact on final yields. Lodging from wind in late-vegetative to early-reproductive stages reduced yields by up to 30% in controlled field studies with simulated wind damage (Carter and Hudelson, 1988). The impact on yield is two-fold as wind damage often changes the morphological characteristics of the plant, reducing harvest efficiency for stalks that are still standing (Cleugh *et al.*, 1998; Gardiner *et al.*, 2016). When corn stalks are broken in high wind, the phenomenon is referred to as 'green snap' (Figure 1). Late in the season, green snap produces yield losses that are immediate and irreversible. These yield losses have been shown to be directly proportional to stalk breakage (Elmore and Ferguson, 1999). For this reason, assessment of damage extent can be critically important for calculating the expected economic impact of green snap.

Vescovo *et al.* (2016) reported success in detecting hail-damage corn with airborne LiDAR, but noted that UAVs represented a future opportunity for innovation in this area. With sufficient overlap between successive image frames, UAVs are capable of creating a 3D model of the crop canopy via structure from motion (SfM). SfM is a technique which relies on the concept of parallax to estimate 3D surfaces from 2D images. SfM is a low-cost alternative to other methods which estimate surface

DOI 10.3920/978-90-8686-888-9_50, © Wageningen Academic Publishers 2019



Figure 1. Left: wind damage to mature corn stalks, referred to as 'green snap,' which results in significant yield losses because the corn plant is unable to recover. Right: corn plants which are laid over by wind have potential to right themselves if the plant is young enough and the stalk is not broken. Timing of data collection is critical to accurately predict if corn stalks will recover.

topography (e.g. LiDAR, terrestrial laser scanning) and is the basis for phenotyping, which is also increasingly performed with UAVs. Chu *et al.* (2017) successfully demonstrated the potential for the use of UAVs and SfM techniques to assess lodging severity in small-scale corn plots. Therefore it seems reasonable to expect that UAVs and SfM techniques might be useful in delineating areas within production corn fields where the crop has been laid over by green snap. However, additional questions about operational efficiency, as well as the relative benefits of 3D structure over more traditional 2D methods remain open and unanswered. Thus the objective of this study was to compare image- versus structure-based calculations of damaged area.

Materials and methods

The field site used for this study was a production corn field of *c.* 200 ha, located in Tallahatchie County, Mississippi (USA), near Minter City. UAV missions were conducted in late July 2016. Flights were conducted with a fixed wing aircraft (Nova 7200, Altavian Inc., Gainsville, Florida, USA) equipped with a colour infrared camera sold by the aircraft manufacturer. Imagery was captured at a spatial resolution of *c.* 650 mm^2. Wind damage was naturally occurring. Damage occurred in mid-July, when corn was already at physical maturity (i.e. R6 growth stage per Abendroth *et al.* (2011)). Collected imagery was processed to produce an ortho-mosaic and digital surface model (DSM) using PhotoScan image processing software (AgiSoft LLC, St. Petersburg, Russia). Digital surface models were corrected using 1.5-m digital elevation models for the area of interest sourced from the Mississippi Automated Resource Information Service. Supervised image classification methods were applied to the generated ortho-mosaic images with colour and texture as distinguishing feature

information, with a support vector machines classifier, following methods from Samiappan *et al.* (2017). Image classification was conducted with eCognition (Trimble Inc., Westminster, Colorado, USA). Classification analyses produced vector layers of damaged and non-damaged areas based on simulated normalized difference vegetation index calculated from colour infrared reflectance. These layers were generated as both binary classes (damage vs non-damage) and binned classes of severity (low, med, high).

Georeferenced field scouting was conducted by the producer's contracted consulting company. The field scout was provided with a 1-m resolution image product derived from UAV imagery before visiting the field to guide ground-truth collection. The consulting company then provided georeferenced field data which identified verified areas of damage and also severity of damage, in addition to field photographs. The consulting company also provided spatially-explicit yield data from the field site. Yield data were cleaned to remove outliers, with anything more than twice the average expected corn yield for the geographic region of the study removed (13,800 kg/ha), as these values were likely erroneous. Following this quality control, yield data were prepared in multiple representations for analyses. Yield from individual rows was summarized to obtain a row average as it became quickly apparent that some smoothing was necessary (see Results section for explanation). Row averages were calculated in ArcGIS (ESRI, Redlands, California, USA) by creating polygons for each row and spatially joining the polygon layer to the yield points, summarizing the underlying points to obtain an average yield for the row.

A confusion matrix was used to calculate the kappa statistic (Viera and Garrett, 2005) for both image and the DSM approaches against the ground-truthed damage assessment from the field scout. The value of the kappa statistic was used as the basis for conclusions presented regarding the comparison between image- versus structure-based calculations of damaged area.

Results

A visual inspection of image data products confirmed expectations regarding damage impacts on yield. Areas showing significant damage aligned well with low yielding areas (Figure 2). After preliminary investigation with yield on an individual point basis, it became apparent that the realities of the yield monitor data made it impractical for the intended analyses. Particularly in damaged areas, incoming flow to the combine was irregular, resulting in spikes in biomass, rather than continuous variable intake.

Thus yield at any given point was not considered truly representative of the point, especially in the areas of light damage and on the borders of areas with heavy damage where corn was matted together. The inaccuracies introduced by this effect rendered the initial correlation analyses meaningless. As a next step, tiles with dimensions matching the swath width were created using the yield point as a centroid. This did not alleviate all issues associated with irregular yield, but did improve the representative nature of the data. Ultimately, row averages for yield were selected to perform analyses as this incorporated the biomass spikes and the gaps surrounding the spike in the calculation, and seemed most representative of the actual yield obtained. Thus, all results presented are based on row averages for yield.

Analysis of corn damage at any stage is likely to be confounded by the single-point-in-time nature of the image collection. In growth stages up to silking (i.e. R1 growth stage), previous researchers noted that the upper portions of lodged corn stalks righted themselves two days post event (Carter and Hudelson, 1988). It was also noted by Chu *et al.* (2017) that recovery of lodged corn plants made them more difficult to detect in earlier stages; detection ability improved as the crop matured. In the current study, the field scout noted that many stalks that were broken continued to fall after the image data were collected. This means that these damaged portions of the field were potentially not accounted for in the image analysis, but would have affected the final yield output. Additionally, in these instances, the severity would have been rated by the scout as less severe than areas where

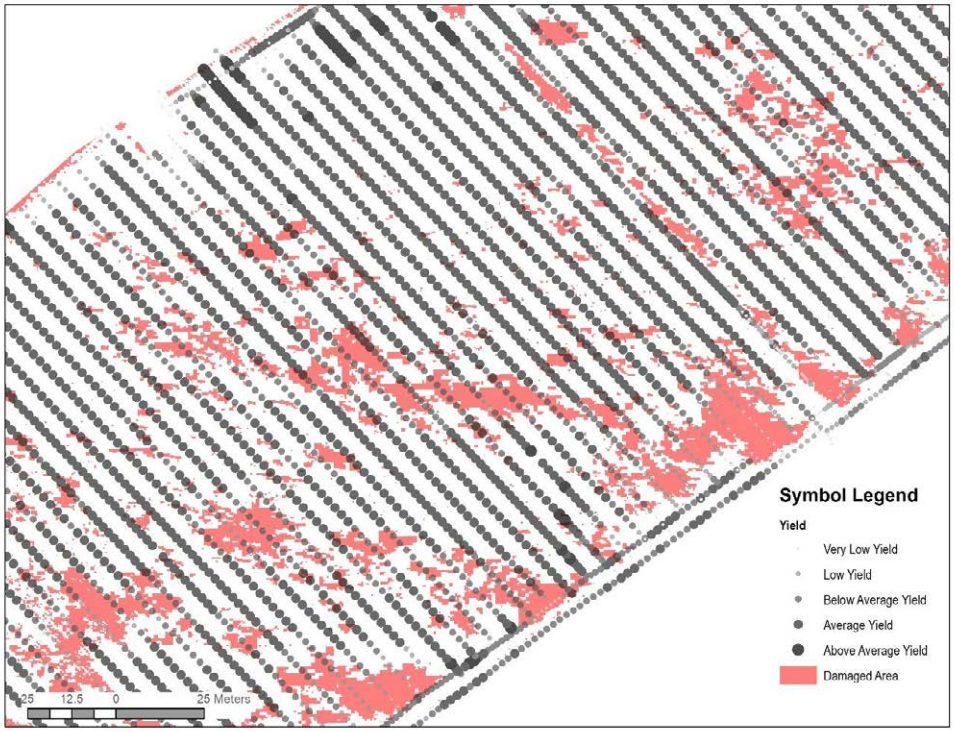


Figure 2. Vector layers of damaged and non-damaged areas of the corn field were generated from unmanned aerial vehicle imagery. Areas of damage were compared against yield monitor data. Areas classified as 'damaged' aligned well with areas of below average to very low yield, although spikes in biomass were apparent in the data as the harvester moved between areas of damaged corn and undamaged corn.

corn stalks were laying on the ground; however, had the field scouting been performed several days later, the outcome may have been markedly different for the severity rating. Upon reflection of these issues, the basis for the number of classes and where to place class breaks seemed more subjective than objective. Thus the binary damage layer was selected for use in analyses. Under the binary system, corn stalks that were bent, broken, or flattened were all considered damaged. Thus this system is more likely to align with final yield values.

Both comparisons produced kappa values interpreted as moderate agreement. Although the image data alone provided highly accurate classification of damaged areas, it was poorer at detecting non-damaged areas. High accuracies for both are expected as damage was clearly visible in the high-resolution imagery (Figure 3). The improvement in the false negative rate improved the overall accuracy, and accordingly the kappa value. The classification informed by both the image data and the DSM yielded a kappa of 0.49, which was superior to the 0.41 yielded by the image data alone (Table 1). This indicates that the 3D information does improve field assessment of damage, although both methods were able to detect damage reliably.



Figure 3. Gray-scaled colour infrared imagery taken with an unmanned aerial vehicle of a wind-damage corn field. Areas of damage are apparent within the image.

Table 1. Classification accuracy results for image analysis of unmanned aerial vehicle data from a corn field with green snap damage.

	Image data alone	Image data with digital surface model data
Damaged class accuracy	99%	96%
Non-damaged class accuracy	48%	60%
Overall accuracy	68%	73%
Kappa statistic	0.41	0.49

Discussion

Working with the farmer and his consultant provided a number of insights into the chances for operational success of this technique in production agriculture. Typically for the consultant, when large-scale damage is suspected, the consulting company will seek some form of aerial imagery; generally from an airplane as UAVs have not been available until more recently. Particularly in mature corn, it is inefficient to scout fields to search for all areas of damage. For this study, the flight time was approximately 1 hr. Collection of a significant amount of images is a by-product of the high overlap needed to produce reliable DSM using SfM. The flight collected over 3,000 images, which contributed greatly to lengthy processing times. The total time to generate a DSM was 8 hrs for the

computer used (Intel Xeon processor, 3.70 GHz with 32 GB RAM), which was approximately twice the time necessary to generate an ortho-mosaic only (it should be noted that Photoscan does produce a low-quality DSM as a pre-cursor to generating an ortho-mosaic, but this is not the same as the SfM-derived DSM used for analysis). At the time of this study, the average price of commissioning a single mission from an airplane-based aerial imaging service was approximately USD$ 2.50/ha, meaning the cost for this field would have been USD$ 500.

In practice, the high resolution of the UAV imagery proved impractical in terms of file size for the consultant, which should be a consideration for future efforts. Previous research on detection of hail damage in corn reported acceptable results from resolutions as high as 5 m (Peters *et al.*, 2000). To obtain a workable option for the scout, it became necessary to downgrade the resolution of the UAV imagery from 650 mm^2 to 1 m^2. The file also had to be transformed to a more mobile-friendly file type, in this instance a *.kmz file, for use in-field. The temptation with UAVs is to collect at ultra-high resolutions because it is feasible, even though it is often not necessary. This temptation is exacerbated in many controlled research projects by the fact that small plots are reasonably quick to image and low altitude flight is obtainable without considerations for battery life. On the other hand, at least in the USA, flight altitude limitations (generally 120 m AGL) combined with the high resolution of commercial sensors may mean ultra-high resolutions at even the highest altitudes allowed by the Federal Aviation Administration for regular commercial operations.

Intuitively, the use of a sensor capable of providing near infrared information should be useful in identifying those corn plants with broken, but still standing stalks, as the reflectance should be decreasing. In fact, this phenomenon has been observed in previous remote sensing studies on corn unrelated to crop damage where lodging occurred in research fields (e.g. GopalaPillai and Tian, 1999). In the later stages of corn growth, it can be challenging to separate decreased reflectance due to damage from that due to crop maturity. Future efforts should include detailed information about stalk condition and general crop health in field reports. Given what has already been said about the snapshot nature of the image data, the detailed field notes may be extremely valuable for teasing out subtle differences between damaged and merely-senescing corn plants. Where possible, the recommendation would be to conduct more than one mission to monitor continued loss within the canopy. This should be undertaken, however, with the understanding that crop maturity will also likely be changing between missions, and this change must be accounted for. Although less problematic in late stages of maturity such as this, previous researchers have commented that changes in corn maturity in earlier stages have limited the ability to conduct pre/post analysis in simulated damage trials (Peters *et al.*, 2000).

From an equipment standpoint, a number of low-cost UAVs are commercially available and can be auto-piloted by even low-skilled users with mobile applications that conduct flight plans sufficient for DSM generation. This makes the technology broadly accessible. However, to obtain accurate results, a certain amount of effort to collect ground control points is required that is more rigorous than normal ortho-mosaic needs. Additionally, the images must be processed into a data product that can be analysed. This represents a secondary cost for either a cloud-based service or a desktop software license. The desktop license requires a trained operator to manipulate the data to generate the 3D output. Both options require a trained analyst to interpret results. Reflectance in the near infrared theoretically adds more capability to detect damage. However, a sensor capable of providing this information will carry a higher price tag than the RGB cameras integrated into low-cost UAV platforms. Given the accuracy of the image data alone in detecting damaged areas, the additional processing time and effort to generate a DSM may not be warranted. However, if an accurate account of non-damaged areas is also needed, the DSM is recommended to the extent it remains cost competitive with status quo methods.

Conclusions

While the approach showed promise, more development of protocols is needed for reliability and operational success. A targeted collection strategy should reduce the time and effort necessary to collect and process data. However, the value proposition for producers and the likelihood of adoption bears more examination.

Acknowledgements

Much appreciation is extended to Southern Ag Consulting, LLC for collaborating with the research team to gather data and evaluate the utility of the proposed approach, and the producer who allowed access. This work was partially supported by funding from the Mississippi Agricultural and Forestry Experiment Station and US Department of Agriculture National Institute of Food and Agriculture Hatch funds.

References

Abendroth, L. J., Elmore, R. W., Boyer, M. J., Marley, S. K., 2011. Corn Growth and Development (No. PMR-1009). Iowa State University Extension Service, Ames, IA, USA. 50 pp.

Carter, P., Hudelson, K., 1988. Influence of simulated wind lodging on corn growth and grain yield. Journal of Production Agriculture 1(4), 295-299.

Chu, T., Starek, M.J., Brewer, M.J., Murray, S.C., Pruter, L.S., 2017. Assessing lodging severity over an experimental maize (*Zea mays* L.) field using UAS images. Remote Sensing 9(9), 923. Cleugh, H., Miller, J., Böhm, M., 1998. Direct mechanical effects of wind on crops. Agroforestry Systems 41(1), 85-112.

Duvick, D., 2005. Genetic progress in yield of United States maize (*Zea mays* L.). Maydica 50 (3/4), 193-202.

Elmore, R.W., Ferguson, R.B., 1999. Mid-season stalk breakage in corn: hybrid and environmental factors. Journal of Production Agriculture 12(2), 293-299.

Erickson, B.J., Johannsen, C.J., Vorst, J.J., Biehl, L.L., 2004. Using remote sensing to assess stand loss and defoliation in maize. Photogrammetric Engineering & Remote Sensing 70(6), 717-722.

Gardiner, B., Berry, P., Moulia, B., 2016. Wind impacts on plant growth, mechanics and damage. Plant Science 245, 94-118.

GopalaPillai, S., Tian, L., 1999. In-field variability detection and spatial yield modeling for corn using digital aerial imaging. Transactions of the ASAE 42, 1911-1920.

Peters, A.J., Griffin, S.C., Viña, A., Ji, L., 2000. Use of remotely sensed data for assessing crop hail damage. Photogrammetric Engineering & Remote Sensing 66(11), 1349-1355.

Samiappan, S., Turnage, G., Hathcock, L., Casagrande, L., Stinson, P., Moorhead, R., 2017. Using unmanned aerial vehicles for high-resolution remote sensing to map invasive *Phragmites australis* in coastal wetlands. International Journal of Remote Sensing 38(8-10), 2199-2217.

Towery, N.G., Eyton, J.R., Changnon Jr, S.A., Dailey, C.L., 1975. Remote sensing of crop hail damage. Illinois State Water Survey, Urbana, IL, USA. 29 pp.

Vescovo, L., Gianelle, D., Dalponte, M., Miglietta, F., Carotenuto, F., Torresan, C., 2016. Hail defoliation assessment in corn (*Zea mays* L.) using airborne LiDAR. Field Crops Research 196, 426-437.

Viera, A.J., Garrett, J.M., 2005. Understanding interobserver agreement: The kappa statistic. Journal of Family Medicine 37(5), 360-363.

A multispectral processing chain for chlorophyll content assessment in banana fields by UAV imagery

G. Rabatel[1*], J. Lamour[2], D. Moura[1] and O. Naud[1]
[1]*IRSTEA, UMR ITAP, 361 rue Jean-François Breton, 34196 Montpellier Cedex 05, Fance; gilles.rabatel@irstea.fr*
[2]*Compagnie Fruitière, 33 Bd Ferdinand de Lesseps, 13014 Marseille, France*

Abstract

A specific processing chain has been designed to build leaf chlorophyll content (LCC) prediction maps in banana fields using multispectral UAV imagery. A supervised regression model was first calibrated out-of-field using multispectral images of collected leaves laid on the ground, and then applied to build a LCC ortho-mosaic for a complete field using the same multispectral sensor. Though the quality of the resulting LCC map was limited mainly due to light scattering inside the canopy, a significant correlation with within-field vigor heterogeneity has been found, making this approach promising for further banana production management.

Keywords: multispectral sensor, banana, UAV imagery

Introduction

Banana production for export is an input intensive agricultural system organized in large area plantations (several hundred hectares). To optimize the use of agricultural inputs with a precision farming approach, characterization of the spatial variability induced by growing conditions is necessary. This is presently done by means of in-field data measurements, representing a very labour intensive task. Moreover, due to natural asynchronism of banana plants, various phenological stages usually occur simultaneously in the same field, inducing specific sampling difficulties.
On the other hand, unmanned aerial vehicle (UAV) imagery is more and more used for the spatial characterization of various crops, thanks to its low cost and its ability to measure vegetation characteristics over large fields. Such advantages would obviously be of great interest for banana production but feasibility studies are still required.
The objective of the present study was to evaluate the feasibility of using an UAV equipped with a multispectral imaging sensor as an alternative to ground measurements in banana fields. As a first step, the leaf chlorophyll content (LCC) has been chosen as variable of interest, due to its strong relationship with the plant vigor and nitrogen status and its ability to be accessed through optical measurements (Cerovic *et al.*, 2012).

Material and methods

In order to build LCC maps from multispectral data, a supervised prediction model based on partial least square regression (PLS-R) has been chosen (Wold *et al.,* 2001). In contrast to spectral based vegetation indices like MCARI (Modified Chlorophyll Absorption in Reflectance Index) (Daughtry *et al.*, 2000), supervised models require a learning dataset. In counterpart, they are more likely to be adapted to specific crops and available spectral bands of a given image sensor. In the present case, the learning dataset must combine multispectral imager data with reference LCC values. Due to the practical difficulty in collecting a sufficient number of learning samples in banana fields, a specific procedure has been designed: the PLS-R model was built out-of-field by cutting and collecting individual banana leaves in various growing conditions.

DOI 10.3920/978-90-8686-888-9_51, © Wageningen Academic Publishers 2019

Such a learning procedure presents several advantages: (1) reference LCC measurements can be made in better conditions, allowing measurement all over the leaf area; (2) multispectral image data are not biased by leaf inclination and multiple reflection of light inside the canopy (Al Makdessi *et al.*, 2017); (3) correspondence between imaged leaves and reference measurements is straight forward (instead of very tedious or uncertain in canopy images). In counterpart, it requires a specific data processing chain to overcome technical issues like band co-registration and proper application of the resulting out-of-field model on canopy images acquired by UAV.

Calibration of a LCC prediction model

Leaf collection and LCC reference measurements

The experiments were carried out in commercial fields of the Plantation du Haut Penja (PHP), located near Njombe in Cameroon (WGS84: E: 4.612, N: 9.639). Eight fields were selected to provide a set of leaves representative of the diversity that can be found in a banana plantation: three fields with a low average reference chlorophyll content (around 35 µg/cm^2), two fields with a medium content (around 45 µg/cm^2) and three fields with a high content (around 50 µg/cm^2). In addition, the leaves were collected at various leaf order numbers on the banana plant (1 to 8, mostly 1, 3 and 5) and at various phenological stages: before flowering, at flowering and 2, 6 or 10 weeks after flowering. In total, 234 leaves were collected and transported to PHP facilities where chlorophyll measurements were made with a Dualex Scientific+® (Force-A, Orsay, France). Average LCC_D values per leaf (D stands for Dualex) were estimated from distributed measurements all over the leaves (9 on each leaf side).

Multispectral image acquisition of laid leaves

After Dualex measurements, leaves were laid outdoors on a gravel floor made of black pouzollane, and imaged using a multi-spectral sensor Sequoia® (Parrot, Paris, France) mounted on a quadrotor UAV (Quad Mikrokopter®, Moormerland, Germany). The elevation of the UAV was 10 m above the soil, providing a sufficient resolution (1 cm/pixel) to limit pixel mixing at the strip edges of banana leaves (Figure 1).

Radiometric correction of multispectral bands

The Sequoia sensor includes a standard RGB camera (image example Figure 1 right) as well as four monochromatic cameras providing 1,280×960 images in four spectral bands. These bands, named GRE, RED, REG and NIR, have central wavelengths and FWHM (full width at half maximum) of

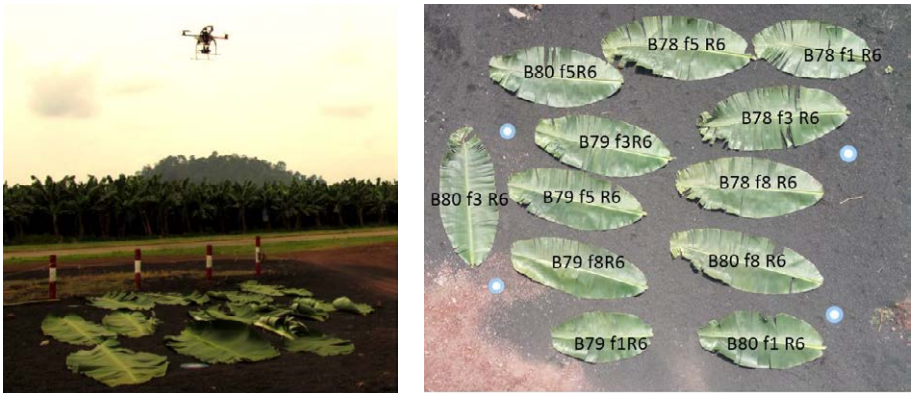


Figure 1. Image acquisition of laid leaves left: image acquisition (Sequoia sensor shown top-left) right: example of Sequoia image (RGB channel) with labelled leaves.

respectively 550 and 40 nm (green), 660 and 40 nm (red), 735 and 10 nm (vegetation red-edge), 790 and 40 nm (near infrared).
These band images require radiometric correction in order to get a reflectance signal independent of the lighting conditions. For this purpose, the Sequoia sensor comes with a sunshine sensor that has to be mounted on the top of the UAV to measure ambient light during acquisition. This sunshine sensor includes four photoelectric cells with respectively the same spectral sensitivities as the GRE, RED, REG and NIR cameras. Its data are recorded and stored in the Exif header (Exchangeable image file format) of each image file. Camera-specific parameters are also available for raw signal to radiance conversion and image vignetting and distortion corrections, as detailed in Sequoia application notes (2017). All these corrections have been systematically applied in order to obtain a reflectance image for every band. Simultaneously, pixels with saturated raw signals were marked to be eliminated in further processing.

Co-registration of multispectral bands
The GRE, RED, REG and NIR camera lenses of the Sequoia® are set at the corners of a 15×15 mm square (Figure 1 left). As a consequence, at 10 m elevation, parallax effects create a significant spatial shift between multispectral components (Figure 2 left). The co-registration method proposed by (Rabatel and Labbé, 2016), with the GRE band as master band, was used to correct this shift and build a usable multispectral image (Figure 2 right).

PLS-R model calibration
PLS-R is usually devoted to spectrometric data involving hundreds of wavelengths, due to its ability to optimize the reduction of the variable space dimensionality according to the output variable of interest. However, it can be used with multispectral data as well, with no dimensionality reduction. For every collected leaf, its position in the corresponding multispectral image was manually recovered and the average value of (GRE, RED, REG, NIR) pixel components computed. The resulting 234×4 spectral values were associated with corresponding LCC_D leaf average values, and then randomly split into a learning set and a test set of respectively 161 and 73 samples (~69 and 31%).
A PLS-R model was calibrated using the learning dataset to predict LCC_S (S stands for Sequoia) from the spectral values. The prediction quality was evaluated by calculating the coefficient of determination (R^2) and the root mean square error (RMSE).

In-field assessment of the model

Field image acquisition by UAV
A particular field among the eight fields where leaves were collected was chosen for UAV image acquisition. This field was chosen because of the heterogeneity of its growing conditions, highlighted

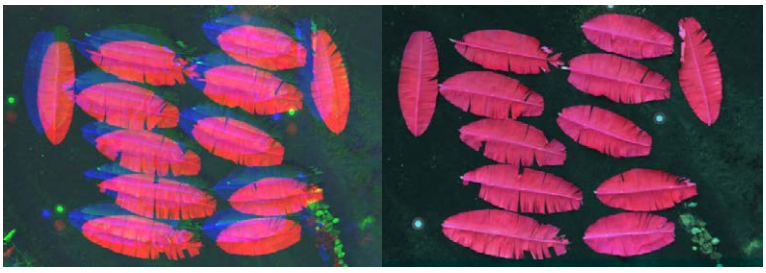


Figure 2. Multispectral image before (left) and after (right) co-registration (images shown in false colors with R=NIR, G=RED, B=GRE).

by a previous mapping of plant's circumferences. Details on this circumference mapping are given in (Leroux *et al.*, 2018).
The UAV with the Sequoia sensor was flown over the whole field at an elevation of 30 m above the banana plants. The flight was designed to have an image overlapping of 80% in the direction of travel and 60% between travels, giving a total of 230 RGB and 4 band images all over the field (after removing take-off and landing phases).

Field UAV image processing
UAV image processing usually relies on the use of specialized mosaicking software to produce a geo-referenced ortho-mosaic from the set of acquired images. The bulk of such software consists in searching correspondence points between overlapping images, and then determining camera 3D-positions and a digital surface model (DSM) of the covered scene. Finally, an ortho-mosaic is built from these data.
In the present study, a slightly different approach has been used to ensure the correct implementation of the PLS-R model designed above:

- Radiometric and band co-registration procedures (see above) were applied to every original set of four band images, and the PLS-R model was applied on the resulting multispectral images, leading to a set of 230 LCC_S images.
- RGB Sequoia images were mosaicked using the commercial software package Photoscan (Agisoft, St. Petersburg, Russia).
- Finally, LCC_S images were assembled to build a LCC map by re-using camera 3D positions and DSM issued from the Photoscan RGB mosaicking process.

LCCS quality assessment on selected leaves
Before UAV acquisition, 25 leaves were chosen in the field at various phenological stages, and colored papers were stapled on them in order to be visible from the sky. After the UAV acquisition, the 25 marked leaves were collected, brought to the PHP facilities and included in the PLS-R calibration dataset previously described.
Thanks to the colored papers, those leaves were retrieved in the RGB ortho-mosaic, and their average LCC values in the LCC_S map were calculated and compared with their LCC_D reference values by means of a linear regression.

Filtering and relevance assessment of the LCCS map
The LCC_S map resolution was much higher than the plant circumference map resolution (resp. 3.3 cm and 1 m per pixel). In addition, the LCC_S map showed short-scale variations within leaves or within plants. In order to filter out these short scale variations and only highlight inter-plant variations, which were assumed to reflect different growing conditions, the LCCs map was averaged on a grid of blocks of 4×4 m. The variogram was then computed, modelled and used for interpolating values at the same locations as the circumference map by ordinary kriging (see Oliver and Webster, 2014). This final LCC_S map was then compared to the circumference map.

Results and discussion

PLS-R model for LCC estimation
The range of LCC_D values for all collected leaves was about 35-65 µg/cm^2. The PLS-R model calibrated on the learning set gave the following prediction equation:

$$LCC_S = -389.33 \times GRE + 377.11 \times RED - 98.12 \times REG + 100.33 \times NIR + 61.86 \quad (1)$$

Its application on learning and test sets gave respectively R^2 coefficients 0.617 and 0.646, with a RMSE of resp. 3.85 and 3.62, i.e. ~12% of the total range (Figure 3).

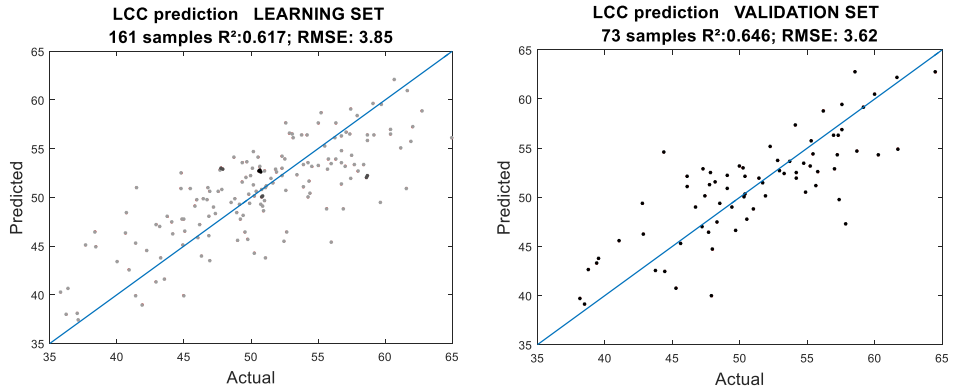


Figure 3. LCC_S estimation by PLS-R on learning and test sets.

The interpretation of Equation 1 is not straight forward. While optical prediction of LCC in chlorophyll-meters usually relies only on red-edge and near infrared bands (Cerovic *et al.*, 2012), Equation 1 also involves GRE and RED spectral components, in an unexpected way (the more the RED component, the more the predicted LCC_S).
Indeed, measurement conditions are slightly different: image reflectance is more sensitive to the shininess of leaf cuticles than transmittance signal of the Dualex. Moreover, the numerous veins of the banana leaves are taken into account, but were not included in Dualex measurements. The PLS-R model probably compensates for such variability factors by combining all available spectral bands.

In-field LCC estimation
Figure 4 shows an example of marked leaf retrieved in the RGB and LCC_S maps, and the linear regression between LCC_D and mapped LCC_S values for all marked leaves. The R^2 coefficient has been drastically reduced compared to the one obtained when leaves were laid on the soil (R^2=0.35 versus more than 0.6), and LCCS values are significantly over-estimated.
The low R^2 can be explained by the difficulty in correctly estimating the average LCC_S for marked leaves in the map due to their positions and orientations in the scene, and by multiple light reflections

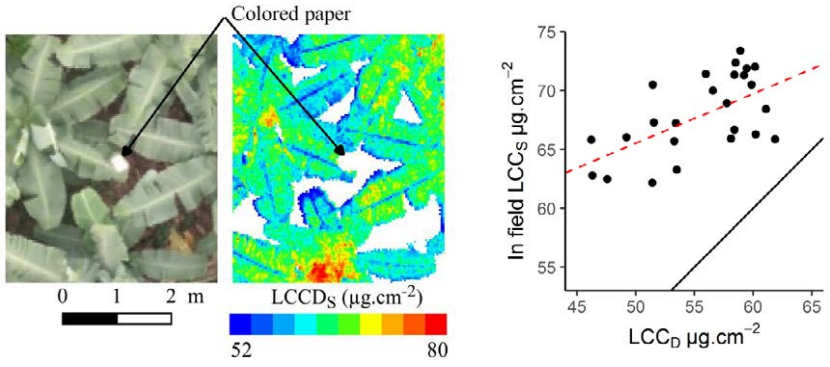


Figure 4. Leaf LCC estimation on selected leaves in canopy (left: RGB and LCCS map detail. right: LCC_S-LCC_D relationship).

in canopy, which modify the leaf spectra. This last cause is also likely to explain the over-estimation of LCC_S (Al Makdessi *et al.*, 2017). Unfortunately, the correction method proposed by the same author (Al Makdessi *et al.*, in press) cannot be used due to an insufficient number of spectral bands.

Assessment of the LCCs map to identify agronomic zones

The previous results were acquired at the leaf level, but the agronomic diagnosis of fields is generally carried out at a larger scale in order to identify spatial trends. The map obtained after filtering and kriging is presented in Figure 5 (right). Indeed, different LCC_S zones are visible within the field. Pseudo-trunk circumference maps for adult plants and for all plants are also given in Figure 5 (left and middle). Zones of small circumferences are very similar to those of low LCC_S (R^2=0.50), showing that valuable agronomic information can be obtained from LCC_S map at this scale.

This result must be moderated. In this field, pseudo-trunk circumference maps obtained with adult plants and with all plants are very similar, indicating a uniform mixing of the phenological stages, due to plant asynchronism effect over a long period. In younger fields with more organized phenological stages, biases could be expected.

Conclusions

In this study, an original approach for LCC monitoring in banana fields by UAV has been developed, using a multispectral sensor. It relies on a supervised PLS-R model. The learning phase has been made possible by combining reference measurements and image acquisition in optimal out-of-field conditions. As a counterpart, specific image processing developments were required to properly manage multispectral image radiometry both in model calibration and model application to aerial canopy images.

In terms of results, LCC prediction performance was limited when applied to field images, probably due to canopy scattering effects. More spectral information (e.g. hyperspectral imagery) would be required both to improve the prediction model and investigate canopy effect corrections. However, this UAV approach remains a promising tool for within-field heterogeneity monitoring at large spatial scale, at least for fields where plant phenological stages are uniformly distributed.

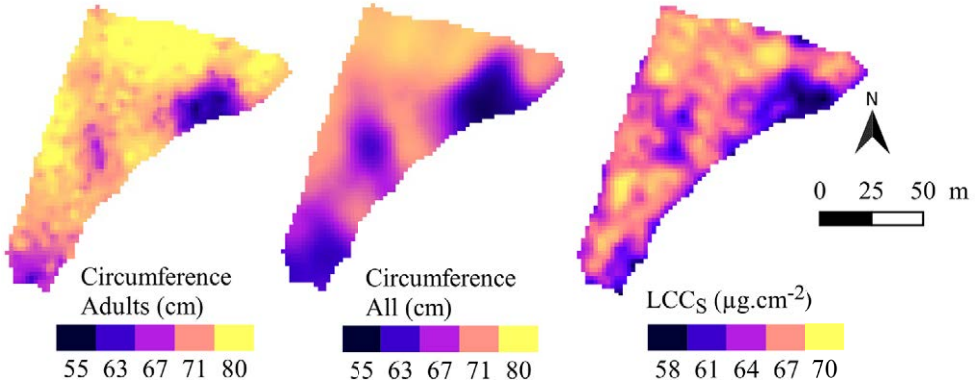


Figure 5. Comparison between LCC_S and vigor maps (left and middle: circumference maps; right: LCC_S map after kriging).

Acknowledgements

The experiences were made in Plantation du Haut Penja with the support of its agronomy department. We are grateful to all the staff of the plantation.

References

Al Makdessi, N, Ecarnot, M., Roumet, P., Rabatel, G., In press. A spectral correction method for multi-scattering effects in close range hyperspectral imagery of vegetation scenes: application to nitrogen content assessment in wheat. Precision Agriculture. https://doi.org/10.1007/s11119-018-9613-2

Al Makdessi, N., Jean, P.-A., Ecarnot, M., Gorretta, N., Rabatel, G., Roumet, P., 2017. How plant structure impacts the biochemical leaf traits assessment from in-field hyperspectral images: A simulation study based on light propagation modeling in 3D virtual wheat scenes. Field Crops Research, 205: 95-105.

Cerovic, Z. G., Masdoumier, G., Ghozlen, N. B. and Latouche, G., 2012. A new optical leaf-clip meter for simultaneous non-destructive assessment of leaf chlorophyll and epidermal flavonoids. Physiologia Plantarum, 146: 251-260.

Daughtry, C.S., Walthall, C., Kim, M., de Colstoun, E.B., McMurtrey, J., 2000. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. Remote Sensing of Environment. 74, 229-239. doi:10.1016/S0034-4257(00)00113-9

Leroux, C., Jones, H., Pichon, L., Guillaume, S., Lamour, J., Taylor, J. *et al.*, 2018. GeoFIS: An Open Source, Decision-Support Tool for Precision Agriculture Data. Agriculture, 8(6), 1-21.

Oliver, M. A., & Webster, R., 2014. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. Catena, 113, 56-69.

Rabatel, G, Labbé, S., 2016. Registration of visible and near infrared unmanned aerial vehicle images based on Fourier-Mellin transform. Precision Agriculture 17:564-587

Sequoia application notes, 2017. https://forum.developer.parrot.com/t/parrot-announcement-release-of-application-notes/5455 (last accessed December 2018).

Wold, S., Sjöström, M., & Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems, 58, 109-130

Developing crop canopy model for irrigation of high-density olive groves by using UAV imagery

E. Zancanaro[1*], A. Gertsis[2], G. Vellidis[3], F. Marinello[4] and F. Morari[1]
[1]*Department of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova, Viale dell'Università 16, Legnaro, PD, Italy; ester.zancanaro@studenti.unipd.it*
[2]*Department of Agro-Environmental Systems Management, Precision Agriculture Laboratory, Perrotis College/American Farm School, 54 Marinou Antypa St, Thessaloniki, Greece*
[3]*Department of Crop and Soil Sciences, University of Georgia, 2360 Rainwater Rd, Tifton, GA, USA*
[4]*Department of Land, Environment, Agriculture and Forestry, University of Padova, Viale dell'Università 16, Legnaro, PD, Italy*

Abstract

Plant architecture is a key issue in understanding olive tree water needs. The aim of this study is to evaluate water requirements of a high-density olive grove by coupling an agro-hydrological model with UAV-based, high-resolution 2D imagery and 3D structure from motion (SfM) reconstruction. Three planting densities (1670, 1000, 500 trees/ha) were drip irrigated according to two irrigation regimes. Evapotranspiration was estimated using FAO56 and FAO66-modified approaches. FAO66 input data (canopy volume and coverage) were quantified using manual ground-based surveys, 2D imagery and 3D SfM reconstruction. 2D imagery showed better relationships to ground-based area and volume estimates than the 3D SfM approach. Moreover, coupling the ET model with 2D data accurately predicted crop water needs.

Keywords: evapotranspiration, olive orchard, high-density olive, unmanned aerial vehicle, irrigation

Introduction

Olive trees (*Olea europea* L.) have been sparsely planted for centuries in the Mediterranean basin. In addition, the high drought tolerance of olive trees was conducive to rainfed production in most cases, and irrigation scheduling never considered a valuable tool to increase water use efficiency (WUE; Gucci and Fereres, 2012). With the adoption of super-high intensity (SHD) systems (1,500-2,500 trees/ha in rows 3-4 m apart; Connor *et al.*, 2014) dense canopies and low soil volume available for each tree result in high water demands. For this reason, irrigation scheduling is essential to prevent drought stress and apply water when needed and with the proper amount (Gucci and Fereres, 2012). Crop evapotranspiration (ETc) is a fundamental factor for estimating crop water demand. Usually, ETc is calculated as the product of reference evapotranspiration (ETo) and a crop coefficient (Kc), where ETo is a measure of the evaporative demand of a given uniform grass sward and Kc quantifies the effect of crop characteristics on ETo. Allen *et al.* (1998) provided a list of Kc values for each crop and developmental stage. For olive, Kc is set as 0.65 (Kc_{ini}) for the initial stage, 0.7 for mid (Kc_{mid}) and late stages (Kc_{end}; Allen *et al.*, 1998).
In tree crops, Kc varies seasonally and is additionally affected by other factors such as canopy and roots architecture, tree density, pruning, crop load, irrigation method and soil surface management (Fereres and Goldhamer, 1990). Moreover, even in mature, intensive groves, full cover is never reached, and soil wetness could impact Kc (Orgaz *et al.*, 2006).
Plant architecture is a key issue in understanding olive water needs as it determines the amount of solar radiation that is intercepted by the canopy (Orgaz *et al.*, 2006). Alternatives to the traditional ground measurement of 3D morphological parameters of tree canopies have recently emerged. Most of these alternatives are based on the use of remote sensing tools. Photo reconstruction enables the

DOI 10.3920/978-90-8686-888-9_52, © Wageningen Academic Publishers 2019

generation of high-resolution digital surface models (DSM) and orthomosaics with high spatial and temporal resolution (Fritz *et al.*, 2013; Gini *et al.*, 2014; Turner *et al.*, 2014).
In this study, all aforementioned factors have been considered in developing a crop-canopy model for the estimation of water requirements of an olive orchard. The final aim was to develop a tool that increased the effectiveness of irrigation, addressing the orchard variability and allowing site-specific irrigation management.

Materials and methods

Study site and experimental design

To investigate potential applications of a canopy-based ET model for precision irrigation management, a long term experimental, 0.24 ha olive orchard was established at the Perrotis College/American Farm School, Thessaloniki, Greece (40° 34' 13' N, 22° 59' 12' E). Olive trees from the Arbequina and Koroneiki varieties were planted in late 2010, at three planting densities (1,670, 1000, 500 trees/ha). The experimental area was composed of 12 tree rows, 50 m long and spaced 4 m apart. A grass cover was maintained between tree lines. The climatic regime is Mediterranean with an annual rainfall of 425 mm and an average daily temperature of 15.7 °C. Measurements were collected during an irrigation period from April to September 2018.
The orchard soil type is sandy clay loam (50% sand, 24% silt, 26% clay) with a bulk density of 1.47×10^3 kg/m^3, pH 7.85, electrical conductivity (EC) of 0.164 S/m and soil organic matter of 1.6%. Total available water content (TAW) in the 500 mm deep root zone is equal to 115 mm.

Irrigation scheduling and evapotranspiration model

The field was irrigated with a drip irrigation system consisting of one 3.8 l/h self-compensating dripper per tree. The amount of water varied between each planting density. Starting from June, each planting density was irrigated according to two irrigation treatments: regular (R) and deficit corresponding to 40 to 60% of the water given to the regular one (D). ETc (mm/day) was calculated using two approaches: FAO56 tabulated Kc values (Allen *et al.*, 1998) and a modified FAO66 model (Fereres *et al.*, 2012) with the addition of the plant direct evaporation component, according to Orgaz *et al.* (2006). The two models are described in Equation 1 and 2, respectively.

$$ETc = (Kc\ Ks)ETo \tag{1}$$

$$ETc = Tr\ (Ks) + Epd + Ewz + Edz \tag{2}$$

where *Tr* is tree transpiration (Equation 3), *Ewz* is surface evaporation from the soil wetted by the emitters (Equation 4), *Edz* is surface evaporation from dry soil (Equation 5) and *Epd* is direct evaporation of rainy water from tree foliage (Equation 6). *Ks* (Equation 7) is the water stress coefficient calculated according to Allen *et al.* (1998).

$$Tr = ETo \times (Qd\ F1\ F2) \tag{3}$$

$$Ewz = 0.6\ \times ETo \tag{4}$$

$$Edz = ETo\ [(0.28 - 0.18G - 0.03ETo + \tfrac{3.8F(1-F)}{ETo})(1 - wz)] \tag{5}$$

$$Epd = ETo\ \left(2.7\ \tfrac{g\ F}{ETo}\right) \tag{6}$$

$$Ks = \tfrac{TAW-Dr}{TAW-(1-p)TAW} \tag{7}$$

ETo (mm/day) is the reference crop evapotranspiration calculated using the FAO Penman-Monteith combination method (Allen *et al.*, 1998); *Qd* is the fraction of solar radiation intercepted by the tree; *F1* and *F2* are two factors related to canopy conductance (Orgaz *et al.*, 2006); *g* is the ground cover fraction of tree canopies; *F* is the monthly frequency of rainy days; *wz* is the fraction of soil surface wetted by the emitters (set at 0.2); *TAW* is the total available water content; *p* is the fraction of *TAW* that can be depleted before moisture stress; *Dr* is the root zone depletion according to the FAO56 daily time step water balance model (Allen *et al.*, 1998).
FAO56 and FAO66-modified models were run for each planting density (d1=1,670 trees/ha, d2=1000 trees/ha, d3=500 trees/ha) and for each irrigation treatment (R = regular and D = deficit irrigation). In addition, the FAO66 model was run using canopy shape data obtained from (1) ground-based surveys (G) performed by measuring canopy height and diameters of each tree by means of a retractable measuring tape or (2) orthomosaic unmanned aerial vehicle (UAV) imagery. Shape data obtained from 3D SfM reconstructions were not used as model input due to the low resolution of the achieved orchard 3D models.

Determination of canopy parameters
Tree area and volume were measured according to the following three approaches: (1) ground-based surveys (from ground measurements, 10 mm resolution); (2) orthomosaic reconstructions (from UAV images, 10×10 mm^2 pixel resolution); (3) 3D area reconstruction (from Structure from Motion (SfM) reconstruction, voxel resolution >1×10^{-4} m^3).
Manual ground-based surveys were performed three times (21 April, 6 June, 7 July): canopy shape was assumed a truncated inverted cone (Figure 1; Equation 8) and volumes (V) calculated for each tree, by manually measuring top (r_1) and bottom (r_2) radii and r_1-r_2 height (h). Canopy area was estimated measuring the ground shade at noon. In d2 and d3, tree shade was assumed to be elliptical and the area was calculated by measuring major and minor axes. In d1, trees created a continuous shade on the ground, thus the shade of each line was assumed to be a rectangle 50 m long (tree line length) and width obtained averaging the widths next to each tree.

$$V_T = \frac{1}{3}\pi h\,(r_1^2 + r_2^2 + r_1 r_2) \tag{8}$$

$$V_S = \frac{4}{3}\pi r_S^3 \tag{9}$$

Airborne imagery data were acquired by collecting images of the experimental area three times (2 April, 26 June, 17 September) during the growing season with a high resolution RGB camera (Canon Ixus 125 HS, Canon Inc, Ōta, Tokyo, Japan) mounted on a quadricopter (GEP 4/2, Graupner/SJ GmbH, Kirchleim unter Teck, Germany). The UAV flew 50 m above the ground, image overlap was 65% (33 images on the field test) during the first flight, 85% (44 images) during the second and 40%

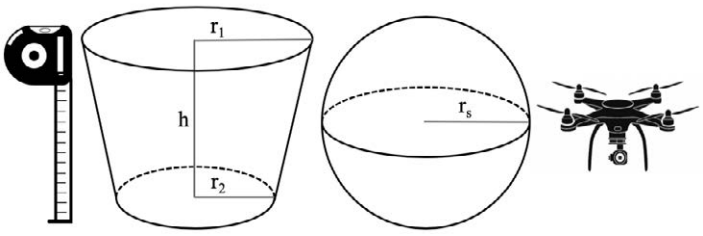


Figure 1. Canopy shape is approximated to a truncated inverted cone in case of ground-based surveys and to a sphere in case of airborne imaging. r_1, r_2, h and r_S (as depicted) are the relative estimated parameters.

(20 images) during the third one. Images from each UAV flight were processed in order to obtain three digital elevation models (DEMs).
Orchard orthomosaics relative to each UAV flight were generated using Agisoft Photoscan Professional (v. 1.0.4, Agisoft LLC, St. Petersburg, Russia). Tree segmentation was manually carried out by outlining individual tree borders in order to overcome software limitations. Indeed, the available software tools could hardly distinguish tree canopies from grass cover. Areas were computed by means of SPIP™ software (v. 4.6.1, Scanning Probe Image Processor, Image Metrology A/S, Hørsholm, Denmark) applying a specific routine (Particle & Pore Analysis module) for quantification of segmented images. Volume was eventually calculated by extracting average radii (r_S) per line and approximating canopy shape to a sphere (Figure 1; Equation 9).
With regard to 3D SfM reconstructions, images were processed taking advantage of Agisoft Photoscan software. Point cloud segmentation was made by applying a SPIP™ software routine (Particle & Pore Analysis module) that automatically placed tree thresholds at a fixed height from the ground (0.5 m). 3D maps were processed for residual local slope correction in order to avoid thresholding errors. Areas and volumes of each tree line were extracted and applied for further analyses.
Canopy area and volume estimates for the three different sampling dates and methods (G, 2D and 3D) were compared in all the area and volume combinations to investigate possible relationships between UAV-based estimates and ground-based measurements. Regression analysis was performed by comparing FAO66-modified daily ETc values calculated by using ground (G) and 2D data (Figure 1) and by comparing FAO56 and FAO66-modified daily ETc data.

Results

Weather data, canopy volume and area
During the irrigation period, from April to September, rainfall was 320 mm, average air temperature was 23 °C and average relative humidity (RH) was 72%.
The 3D models for April and September flights showed poor relationships with ground measurements for both areas and volumes apparently due to the low image overlap (%) and the total number of images. R^2 values were 0.16 (April) and 0.01 (September) for volume comparison, 0.18 (April) and 0.01 (September) for area comparison (Table 1). The correlation between ground and 2D data was stronger because image numbers were suitable to obtain high resolution orthomosaics (Table 1).

Crop evapotranspiration (ETc)
Total ETc (mm) calculated using FAO56 and FAO66-modified models and FAO66 ETc components (Ep, Epd, Eso, Esw) are shown in Table 2. When ETc was calculated by FAO66-modified model, water stress (Equation 7) occurred in the deficit irrigation treatment, during July, August and September for d1 (average Ks 0.96, 0.8, 0.83) and during September for d2 (average Ks of 0.84). For the FAO56

Table 1. R^2 values from regression analyses performed between 2D, 3D and ground areas and volumes on the three sampling dates (April, June and September). Bold indicates significant values ($P<0.05$).

	April	June	September
Area 2D vs G	**0.66**	**0.92**	**0.84**
Area 3D vs G	0.18	**0.8**	0.01
Volume 2D vs G	**0.56**	**0.84**	**0.73**
Volume 3D vs G	0.16	**0.75**	0.01

Table 2. Total ETc calculated by FAO56 and FAO66-modified models and FAO66 ETc components (Ep, Epd, Eso, Esw). d1, d2 and d3 are the three planting densities (1,670, 1000, 500 trees/ha), R and D indicate regular and deficit irrigation treatments, 2D and G are 2D and ground-based canopy data; CC is cover crop; * means that water stress occurred.

	Treatment	Etc (mm)	Ep (mm)	Epd (mm)	Eso (mm)	Esw (mm)
FAO66	d1-R-G	685	106	13	47	6
	d1-R-2D	541	72	11	48	6
	d1-D-G	642	98*	13	47	4
	d1-D-2D	534	72	11	48	4
	d2-R-G	529	71*	9	51	4
	d2-R-2D	457	54	9	50	4
	d2-D-G	519	69*	9	51	3
	d2-D-2D	453	54	9	50	3
	d3-R-G	346	28	4	56	2
	d3-R-2D	311	20	4	55	2
	d3-D-G	344	28	4	56	1
	d3-D-2D	309	20	4	55	1
FAO56	CC	191	-	-	-	-
	d1-R	508	-	-	-	-
	d1-D	508	-	-	-	-
	d2-R	508	-	-	-	-
	d2-D	508	-	-	-	-
	d3-R	508	-	-	-	-
	d3-D	480*	-	-	-	-

model-based ETc, water stress occurred in d3 planting density, deficit irrigation treatment (D), during September (average Ks of 0.56).

The regression analysis between FAO66 daily ETc calculated using ground and 2D data showed a strong correlation, particularly in d1 and d2 planting densities, for both regular (R^2 values of 0.93, 0.88 and 0.79 for d1, d2 and d3, respectively) and deficit irrigation (R^2 values of 0.85, 0.87 and 0.79 for d1, d2 and d3, respectively) treatments (Figure 2).

The regression analysis between FAO56 and FAO66-modified (ground measurements as input) resulted in significant R^2 values ($P<0.05$; Table 3). The correlation was stronger in d2 planting density, both regular (R) and deficit (D) irrigation treatments, when compared to d1 and d3 (R^2 values equal to 0.73 and 0.72). Moreover, daily ETc values of d1 and d2 planting densities closely agree to a 1:1 relationship (Table 3).

Discussion

Canopy volume and area

2D reconstruction showed better relationships to ground-based area and volume estimates than the 3D SfM approach. The number of acquired images, particularly on the April and September flights, was not adequate for the software to reconstruct accurate orchard 3D models while it was acceptable to build orchard orthomosaics. The number of images strongly affected area and volume estimates with higher R^2 values when the overlap was set at 85% (0.92 for area and 0.84 for volume estimates; Table 1). The presence of a grass cover with colour appearance not easily distinguishable from tree

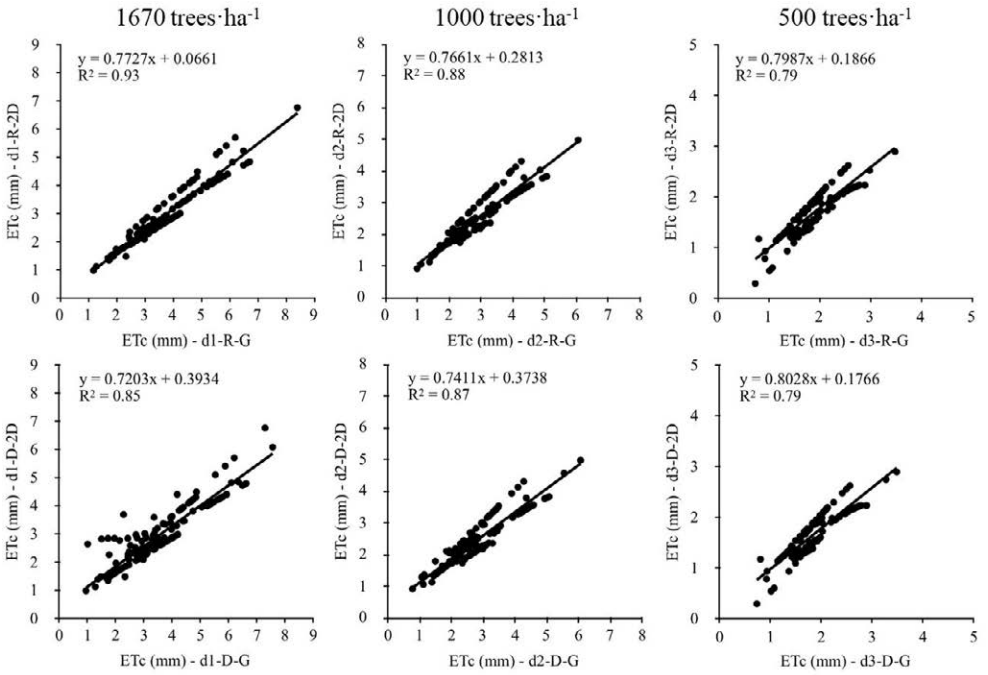


Figure 2. Regression analyses between FAO66-modified daily ETc with ground (G) and 2D data as input. Graphs in the first line refer to regular irrigation treatment (R), while graphs in the second line refer to deficit irrigation treatment (D).

Table 3. Results from the regression analyses between FAO56 and FAO66-modified daily ETc data with ground measurements as input.

	R^2	Line equation
FAO56 vs FAO66 – d1-R	0.6	y = 1.09x – 0.89
FAO56 vs FAO66 – d1-D	0.61	y = 1.29x – 0.09
FAO56 vs FAO66 – d2-R	0.73	y = 0.99x – 0.14
FAO56 vs FAO66 – d2-D	0.72	y = x – 0.04
FAO56 vs FAO66 – d3-R	0.64	y = 0.49x – 0.53
FAO56 vs FAO66 – d3-D	0.59	y = 0.4x – 0.82

canopies did not allow implementation of automatic thresholding routines. As a consequence, segmentation was manually performed in order to increase data accuracy, although this resulted in a time-consuming and poorly repeatable approach.

Crop evapotranspiration (ETc)

Fao66-modified model was applied using ground data and 2D data as model input. The correlation between them was stronger for d1 and d2 than d3 planting density. In the latter, the low canopy volume and poor resolution of orthomosaics did not allow to accurately reconstruct small canopies.

Therefore, it was difficult to distinguish tree canopies from grass cover and, in turn, perform the manual segmentation. The strong correlation between ETc estimated by using ground and 2D data showed that UAV could replace time-consuming and unprecise ground-based analyses.
Comparisons between ET models highlighted that FAO56 was able to describe lesser than 73% of the FAO66 variability. Standard Kc values used in FAO56 approach did not account for all the components that spatially contributed to ETc losses from the orchard such as planting density, pruning practices, etc. Conversely, the proposed FAO66 model combined with remote sensing technologies allowed to capture canopy architecture variability. Coupling this approach with variable rate irrigation techniques could optimize water use by delineating dynamic management zones (MZ) according to single tree requirements. To verify the accuracy of model estimates, future and in-field ET measurements should be needed.

Conclusions

Olive orchard ETc is affected by canopy shape parameters but ground-based surveys techniques for their estimation are time-consuming and not repeatable. This study demonstrated that consistent canopy area and volume estimation using sub-optimal overlapping images may be obtained with orthomosaics, while it is less reliable in 3D reconstruction. Because of the presence of a green grass cover between tree rows, manual thresholding was performed and this made the 2D approach time-consuming and not repeatable. By improving 3D SfM reconstruction, the process could become automatic, repeatable and more accurate as the canopy thresholding would be based on tree height estimates and not on the tree colour. Furthermore, the ET model fed with 2D-estimated plant architecture showed results comparable with ground-surveys techniques. The developed model can increase the precision and therefore the efficiency of irrigation in high density olive groves and can become transferable in similar linear production systems. Further validation of this model is needed.

References

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M., 1998. Crop evapotranspiration – Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56. FAO: Rome, Italy 300 pp.

Connor, D.J., Gomez-de-Campo, M., Rousseaux, M.C. and Searles, P.S., 2014. Structure, management and productivity of hedgerow olive orchards: a review. Scientia Horticulturae 169 71-93.

Fereres, E. and Goldhamer D.A., 1990. Deciduous fruit and nut trees. In: B.A. Stewart and D.R. Nielsen (Eds.) Irrigation of Agricultural Crops, Agronomy monograph no. 30, ASA: Madison, Wisconsin, USA pp. 987-1017.

Fereres, E., Goldhamer, D.A. and Sadras, V.O., 2012. Yield response to water of fruit trees and vines: guidelines. In: P. Steduto, T.C. Hsiao, E. Fereres, D. Raes (Eds.) Crop yield response to water, FAO Irrigation and Drainage Paper 66, FAO: Rome, Italy pp. 246-295.

Fritz, A., Kattenborn, T. and Koch, B., 2013. UAV-based photogrammetric point clouds-Tree stem mapping in open stands in comparison to terrestrial laser scanner point clouds. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Rostock, Germany XL-1/W2 141-146.

Gini, R., Passoni, D., Pinto, L. and Sona, G., 2014. Use of unmanned aerial systems for multispectral survey and tree classification: A test in a park area of northern Italy. European Journal of Remote Sensing 47 251-269.

Gucci, R. and Fereres, E., 2012. Fruit trees and vines: olive. In: P. Steduto, T.C. Hsiao, E. Fereres (Eds.) Crop yield response to water, FAO Irrigation and Drainage Paper 66, FAO: Rome, Italy pp. 297-313.

Orgaz, F., Testi, L., Villalobos, K.J. and Fereres E., 2006. Water requirements of olive orchards-II: determination of crop coefficients for irrigation scheduling. Irrigation Science 24 77-84.

Turner, D., Lucieer, A. and Watson, C., 2014. An Automated Technique for Generating Georectified Mosaics from Ultra-High Resolution Unmanned Aerial Vehicle (UAV) Imagery, Based on Structure from Motion (SfM) Point Clouds. Remote Sensing 4 1392-1410.

Section 7
Satellite-based applications for precision agriculture

Assessing spatial and temporal variability in evapotranspiration for olive orchards in Tunisia using satellite remote sensing

A. Bchir[1], D.J. Mulla[2*], A. Ben Dhiab[1], F. Ben Meriem[1], W. Bousetta[1] and M. Braham[1]
[1]*Institut de l'Olivier, Unité spécialisée de Sousse, Rue Ibn Khaldoun, B.P. 40, Tunisia*
[2]*Dept. Soil, Water & Climate, Univ. Minnesota, 1991 Upper Buford Circle, St. Paul, MN 55108 USA; mulla003@umn.edu*

Abstract

The objective of this study was to evaluate satellite remote sensing to estimate evapotranspiration (ET) for olive orchards in Tunisia. Crop water stress index (CWSI) and reference or actual ET were estimated for treatments having deficit drip irrigation at 60, 50 and 40% of crop ET. Results showed significant differences in CWSI or ET between treatments. ET estimates from satellite imagery were strongly correlated with ground truth measurements, with r^2 values of 0.86 and 0.62 for potential ET and actual ET, respectively, showing that satellite imagery can accurately quantify spatial and temporal patterns in ET in Tunisian olive orchards.

Keywords: olive orchard, precision irrigation, evapotranspiration, METRIC, EEFlux, Tunisia

Introduction

Water is a scarce resource in Tunisia. While annual precipitation in Tunisia averages 300 mm/yr, it ranges from a high of over 1,500 mm/yr in the northeast to less than 100 mm/yr in the southwest. Crop production uses 31% of the nation's surface and groundwater resources. By far the largest consumer of water is olives at 7,570 Mm^3/yr, of which 270 MCM/yr is from surface and groundwater irrigation (Chouchane *et al.*, 2013).

Better water management is challenging as a result of spatial and temporal variability in climate, coupled with seasonal changes in crop growth and development. There is increasing awareness for the need to use water saving strategies to adapt to climate change, which result (IPCC, 2013) in a decrease in number of days with precipitation as well as of amount of precipitation.

The olive tree (*Olea europaea* L.) is well adapted to the semi-arid and arid environment of Tunisia, and is usually grown in rainfed regions with limited water resources. Whereas olive trees can grow and produce adequately under naturally low annual water supply, irrigation is essential to achieve high yields. Therefore, there is an increasing agronomic and economic interest in optimizing water use in olive production in arid and semi-arid lands through precision applied deficit irrigation.

Precision irrigation water management is challenging as a result of spatial and temporal variability in climate and spatial variability in soil, coupled with seasonal changes in crop growth and development. With the advent of Earth observations from satellite-based sensors, it is possible to use remote sensing data to assess spatial and temporal variability in crop water stress and evapotranspiration. Remotely sensed information can help in monitoring olive tree water stress. Evapotranspiration (ET) is one of the most important components of the water balance in semi-arid areas; it is a key factor for optimizing irrigation water management (Poussin *et al.*, 2008). With advances in Earth observations from satellite-based sensors, numerous recent studies have used remote sensing data to assess drought impacts (Ghulam *et al.*, 2007; Peters *et al.*, 2002; Tadesse *et al.*, 2005; Wan *et al.*, 2004). Allen *et al.* (2007) pioneered the use of a satellite based energy balance method for mapping evapotranspiration with internalized calibration (METRIC). This approach has been operationalized using Landsat 8 satellite imagery available through Google Earth's EEFlux engine (Allen *et al.*, 2015), but has never been evaluated in Tunisia.

DOI 10.3920/978-90-8686-888-9_53, © Wageningen Academic Publishers 2019

The hypothesis of this study is that using ground and satellite based approaches will lead to development of precision crop water irrigation strategies that can conserve water and improve productivity of olive orchards in Tunisia. As a first step, the general objective of this project is to examine the feasibility of using satellite remote sensing to estimate spatial and temporal irrigation requirements in Tunisian olive orchards.

Materials and methods

The experimental study (Figure 1) was conducted on irrigated olive trees (cv. Koroneiki) in a semi-arid zone near Menzel Mhiri, Tunisia (35°25'N, 9°48'E). Three water stress treatments were imposed using drip irrigation at 1 to 4 day intervals, namely:

1. Sustained deficit irrigation applying 60% of ETc;
2. Sustained deficit irrigation applying 50% ETc;
3. Sustained deficit irrigation applying 40% ETc.

In each treatment, Crop Water Stress Index (CWSI) was estimated using temperature of sunlit leaves measured by an FLIR IR Thermal camera. Canopy and soil temperatures were also measured using the IR Thermal camera.

The crop water stress index *CWSI* was calculated according to Jones (1992) and Alchanatis *et al.* (2010) as:

$$CWSI = \frac{T_{canopy} - T_{wet}}{T_{dry} - T_{wet}} \quad (1)$$

where T_{canopy} is the actual canopy temperature obtained from the thermal image and T_{wet} and T_{dry} are the lower and upper boundary temperatures, corresponding to a fully transpiring leaf with open stomata and a non-transpiring leaf with closed stomata, respectively (Alchanatis *et al.*, 2010).

A meteorological station installed on the site measured net solar radiation (Rn), wind speed (U), relative humidity and air temperature (T) at 15 minute intervals at a height of 2 m above ground. These data were aggregated from Jan 30, 2018 to August 26, 2018 into daily averages. A rain gauge was used to measure daily precipitation over the same time period.

These data were used to estimate daily reference evapotranspiration (ETo) using the Penman-Monteith formula (Equation 2):

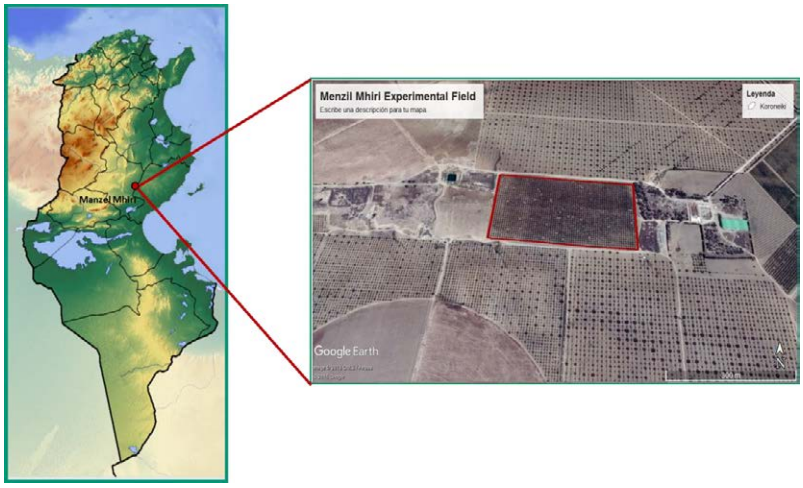


Figure 1. Location of experimental olive orchard at Menzil Mhiri, Tunisia.

$$ET_0 = \frac{[0.408\,\Delta\,(Rn-G)] + \left[900\frac{\gamma}{T+273}\left(e^0 - e\right)U_2\right]}{\Delta + [\gamma\,(1+0.34\,U_2]} \quad (2)$$

Where G is soil heat flux, γ is the apparent psychrometer constant, e^0 and e are the saturated and ambient vapor pressures, and the subscript 2 indicates measurement at a height of 2 m. Time varying crop coefficients (Kc) were multiplied by ETo to obtain crop evapotranspiration (ETc).
Independently, Landsat 8 satellite remote sensing data were used to estimate ETo as well as actual crop evapotranspiration ETa based on the mapping evapotranspiration at high resolution using internal calibration (METRIC) method available through Google Earth Engine evapotranspiration flux (EEFlux, Allan *et al.*, 2015). Calculations were based on cloud free images obtained on seventeen dates from Jan. 30, 2018 to August 26, 2018, with one image per month from January to April, two images in May, three images in June, and four images each in July and August.

Results

The cumulative seasonal amount of irrigation applied was 242.5 mm in treatment 1, 187.1 mm in treatment 2 and 134.1 mm in treatment 3 (Figure 2). From the middle of January to the end of June, less than 20 mm of precipitation fell. Between June 23 and August 2 there was no precipitation, but a large rainstorm totalling 16.2 mm occurred on August 3, causing olive trees to recover from the extended drought. This period produced water stress in treatments 2 and 3.
During the extended drought from June to August, canopy temperatures in the olive orchard were lowest in irrigation treatment 1 and increased in treatments 2 and 3 by about 1 and 2 degrees Celsius, respectively (Table 1). CWSI values were lowest during this period in treatment 1, and increased in treatments 2 and 3 by about 0.15 to 0.19, respectively. Olive tree transpiration decreased from 4.05 to 2.35 to 1.65 mmol/m^2/s in treatments 1, 2 and 3, respectively.
After the large rainstorm on August 3, there was quick recovery from drought, with little evidence for water stress. Canopy temperatures decreased by 3 degrees Celsius in treatment 1, by 4 degrees Celsius in treatment 2 and by 5 degrees Celsius in treatment 3. CWSI dropped precipitously to

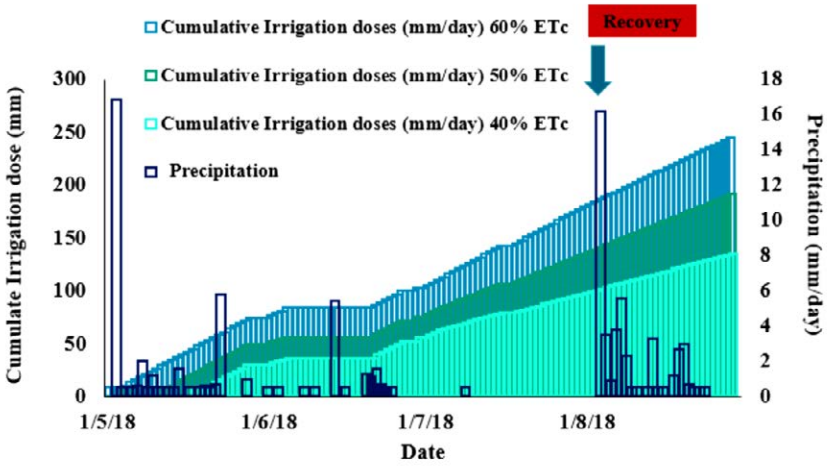


Figure 2. Precipitation at the Menzel Mhiri experimental site from January through September of 2018. Irrigation for treatments 1, 2 and 3 are shown on the same graph.

values in the range of 0.1 to 0.06 in all three treatments (Table 1). Transpiration rates increased to between 5.63 and 5.71 mmol/m^2/s in all three treatments.

Estimates of potential evapotranspiration (ETo) at the experimental olive orchard using the Penman-Monteith equation (Equation 2) were strongly correlated (r^2=0.86) with METRIC-based estimates of potential evapotranspiration from Landsat 8 through the Google Earth Engine EEFlux site (Figure 3A). Satellite based estimates, however, tended to under-estimate ground based estimates by about 15%. Regardless, temporal trends in potential evapotranspiration estimated by satellite accurately tracked temporal trends estimated by the Penman-Monteith method.

Estimates of actual crop evapotranspiration (ETc) using the Penman-Monteith method were strongly correlated (r^2=0.62) with METRIC-based estimates (Figure 3B) when averaged across measurement dates for each of three locations within each of the three irrigation treatments. Satellite based estimates under-estimated ground based ETc measurements by about 30%, but accurately reflected the differing levels of water stress imposed by the three irrigation treatments.

Discussion

The comparison of METRIC ETo and calculated FAO-ETo values indicate the potential of METRIC for providing high resolution ETo estimation (r^2=0.86). According to Numata *et al.* (2017), ET maps developed using METRIC provide the means to quantify ET in terms of both the rate and spatial distribution, and at high spatial resolution.

The underestimation of ETo and ETc using METRIC data can be explained by the high daily variation in daily ET rates due to changes in atmospheric evaporative demand (Numata *et al.*, 2017). For areas where availability of both satellite and meteorological data are adequate, well-characterized ET dynamics in space and time at fine spatial scales, derived from METRIC, will provide a more accurate representation of spatial variation in ET associated with agricultural production compared to the existing coarse resolution ET products, such as MODIS 16 ET (Mu *et al.*, 2007).

The results of the present work show that remote sensing based on satellite images may be used, with high accuracy and spatial representation to calculate crop water and irrigation requirements. These results are in agreement with results found by Saadi (2018). According to Numata *et al.* (2017), there are uses and applications for both coarser scale and finer scale ET maps. For regional or global applications, coarser scale ET estimates are typically adequate. A fine spatial resolution ET

Table 1. Measurements of canopy and soil and crop water stress index (CWSI) and actual crop evapotranspiration in an experimental olive orchard at Menzel Mhiri for three irrigation treatments during and after recovery from an extended drought during July and August of 2018.

Date	Water treatment	Canopy temperature (°C)	Soil temperature (°C)	Crop water stress index (CWSI)	Transpiration (mmol/m^2s)
20/07/2018 during water stress	60% ETc	33.3±0.1	44.7±1.3	0.56±0.01	4.05±0.45
	50% ETc	34.6±0.5	45.7±1	0.71±0.05	2.35±0.21
	40% ETc	35.0±0.2	46.7±1.7	0.75±0.05	1.65±0.64
28/08/2018 after recovery	60% ETc	30.6±0.3	43.6±0.4	0.10±0.03	5.66±0.81
	50% ETc	30.3±0.1	43.9±0.1	0.06±0.01	5.71±0.57
	40% ETc	30.4±0.1	43.8±0.1	0.08±0.01	5.63±1.69

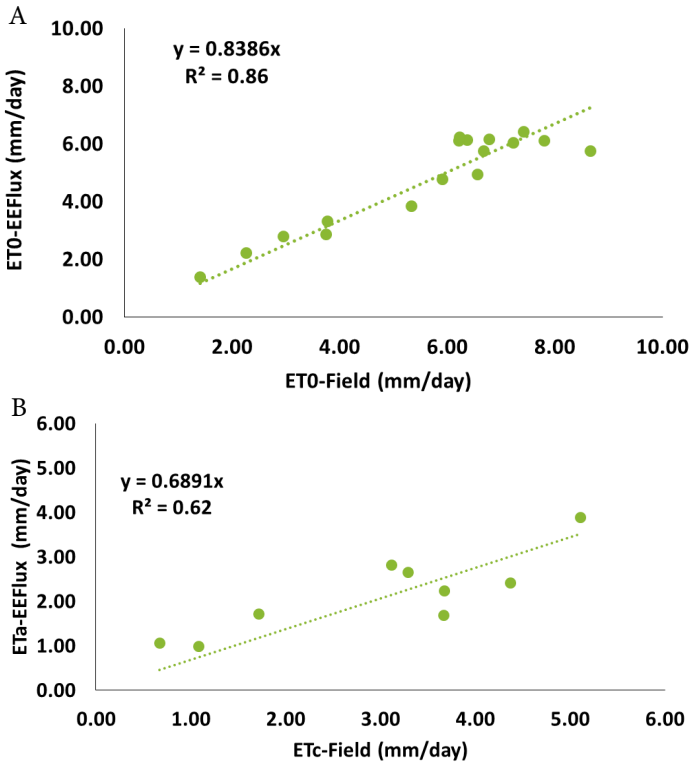


Figure 3. Relationship between field measured potential (A) or actual (B) evapotranspiration and satellite estimated evapotranspiration from January through August of 2018 at Menzel Mhiri experimental olive orchard in Tunisia.

map derived from Landsat 8 imagery (via EEFLux) will allow agricultural producers in developing countries such as Tunisia to better estimate olive tree water requirements for precision agriculture.

Conclusions

This study showed the feasibility of using Landsat 8-based estimates for actual crop evapotranspiration in an experimental olive orchard located in Tunisia. Three deficit irrigation treatments were imposed in the study, at rates of 60, 50 and 40% of ground based estimates for ETc. During a drought that occurred from late June to early August, ground based measurements of olive orchard canopy temperature, CWSI and ETc showed severe water stress in the 50 and 40% irrigation treatments. Drought was ended by a large rainfall event in early August, and measurements of canopy temperature, CWSI and ETc all improved considerably. Satellite-based estimates of potential (ETo) and actual (ETa) crop evapotranspiration in the three deficit irrigation treatments were strongly correlated with ground-based estimates of ETo and ETc from January to August, 2018. Satellite-based estimates were responsive to both irrigation treatment and temporal variability in climate.

Acknowledgements

This project was funded by the International Center for Biosaline Agriculture (ICBA) and the Civilian Research Development Foundation (CRDF Global) (JRCAFS-23425).

References

Alchanatis V., Cohen Y., Cohen S., Moller M., Sprinstin M., Meron M. *et al.,* 2010. Evaluation of different approaches for estimating and mapping crop water status in cotton with thermal imaging. Precision Agriculture 11(1), 27-41.

Allen, R.G., Morton, C., Kamble, B., Kilic, Huntington, J., Thau, D. *et al.,* 2015. EEFlux: A Landsat-based evapotranspiration mapping tool on the Google Earth Engine. In: 2015 ASABE/IA Irrigation Symposium: Emerging Technologies for Sustainable Irrigation-A Tribute to the Career of Terry Howell, Sr. Conference Proceedings (pp. 1-11). St Joseph, MI, USA: American Society of Agricultural and Biological Engineers.

Allen, R.G., Tasumi, M. and Trezza, R., 2007. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC) – Model. Journal of Irrigation Drainage Engineering 133, 380-394.

Chouchane, H., Hoekstra, A. Y., Krol, M. S., and Mekonnen, M. M., 2013. Water footprint of Tunisia from an economic perspective. Value of Water Research Report Series No. 61, Delft, the Netherlands: UNESCO-IHE.

Ghulam A., Li, Z. L., Qin, Q. and Tong, Q., 2007. Exploration of the spectral space based on vegetation index and albedo for surface drought estimation. Remote Sensing, 1: 013529-013512.

IPCC, 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J. *et al.* (Eds.)., Cambridge, UK and New York, NY, USA: Cambridge University Press, 1535 pp.

Jones, H. G., 1992. Plants and Microclimate. 2nd ed., pp. 231-236. Cambridge, UK: Cambridge University Press.

Mu Q., Heinsch F.A., Zhao M., Running S.W., 2007. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. Remote Sensing Environment 111, 519-536.

Numata I., Khand K., Kjaersgaard J., Cochrane M.A., Silva S.S., 2017. Evaluation of Landsat-Based METRIC Modeling to Provide High-Spatial Resolution Evapotranspiration Estimates for Amazonian Forests. Remote Sensing 9, 46.

Peters A.J., Walter-Shea, E. A., Ji, L., Vina, A., Hayes, M., and Svoboda, M. D., 2002. Drought monitoring with NDVI-based standardized vegetation index. Photogrammetric Engineering Remote Sensing 68, 71-75.

Poussin J., Imache, A., Béji, R., Le Grusse, P. and Ben Mihoub, A., 2008. Exploring regional irrigation waterdemand using typologies of farms production units: An example fromTunisia. Agricultural Water Management 95, 973-983.

Saadi S., 2018. Spatial estimation of actual evapotranspiration and irrigation volumes using water and energy balance models forced by optical remote sensing data (VIS / NIR/ TIR). PhD Thesis, Université de Toulouse, France. 294 pp.

Tadesse T., Brown, J. F., and Hayes, M. J., 2005. A new approach for predicting drought related vegetation stress: integrating satellite, climate, and biophysical data over the US central plains. ISPRS Journal of Photogrammetry and Remote Sensing 59, 244-253.

Wan Z., Wang, P. and Li, X., 2004. Using MODIS land surface temperature and normalized difference vegetation index products for monitoring drought in the southern Great Plains, USA. Remote Sensing 25, 61-72.

Accuracy of crop coefficient estimation methods based on satellite imagery

O. Beeri[1*], R. Pelta[1], T. Shilo[1], S. Mey-tal[1] and J. Tanny[2,3]
[1]*Manna-Irrigation, Gvat, 3657900, Israel; ofer.beeri@manna-irrigation.com*
[2]*Institute of Soil, Water and Environmental Sciences, Agricultural research Organization, POB 15159, Rishon-Leziyon 7528809, Israel*
[3]*HIT – Holon Institute of Technology, POB 305, Holon 5810200, Israel*

Abstract

The crop coefficient (K_c) is a crucial factor in irrigation decision making. The ability to estimate it directly from satellite imagery can immensely assist growers around the globe in their crop monitoring. Direct estimations of K_c from diverse spectral indices, enable the use of various multi-spectral sensors and increase the ability to map relatively small plots (<1.0 ha). The goal of this study was to examine the accuracy of K_c estimations using different spectral indices with different satellite sensors. To achieve this goal, a database was created. Datasets included 126 satellite images and corresponding K_c from five flux tower sites over four irrigated crops (citrus, corn, alfalfa, and soybean). Kc was estimated remotely from five sensors, by applying previously published indices equations. The results show that the normalized difference vegetation index (NDVI) achieved the highest accuracy (RMSE<0.1 and nRMSE of 10%) among the indices, characterized by only minor accuracy differences between the sensors. The lowest accuracy among all crops was observed in citrus.

Keywords: crop coefficient, Landsat, Sentinel-2, Linear Imaging Self Scanning, flux towers

Introduction

Crop irrigation is based on the evapotranspiration from the crop (ET_c), which can be estimated remotely from thermal imagery (Allen *et al.*, 2007; Anderson *et al.*, 2011) or by multiplying the reference evapotranspiration (ET_0) by the crop coefficient (K_c). K_c is a crop-specific coefficient as calculated from the ratio between the ET_c and ET_0, which represents the atmospheric evaporative demand together with the crop water requirements. The former approach of estimating K_c has the disadvantage of relatively coarse thermal spatial resolution, which makes it useless for small plots (<1.0 ha), whereas the latter requires standard climatic data to calculate ET_0 and K_c from published protocols (Allen *et al.*, 1998). Another approach is to estimate K_c from satellite imagery, allowing not only mapping the K_c for small plots but also to utilize different satellite sensors and by that to increase the number of cloud-free images.
Two main methods have been suggested in the literature to estimate K_c from satellite sensors: (1) direct relationships between spectral vegetation indices (SVI) and K_c (Toureiro *et al.*, 2017) and; (2) to estimate vegetation fraction (V_f) from SVI and then to re-estimate the K_c from Vf (Allen and Pereira, 2009). The validation of such methods can be carried out with soil sensors or lysimeters (Hunsaker *et al.*, 2003; Johnson and Trout, 2012), but for satellite imagery, the common validation approach is to utilize flux tower datasets (Wang *et al.*, 2007; Nagler *et al.*, 2009). The goal of this study was to estimate K_c in both methods mentioned above (i.e. SVI to K_c and V_f to K_c), by calculating different SVI from different satellite sensors namely, Landsat-5/7/8, Sentinel-2 and the Linear Imaging Self Scanning (LISS-3). To achieve this goal, three research questions were defined:

1. Does the accuracy vary between different equations (relating SVI to K_c)?
2. Does the accuracy vary between sensors?
3. Does the accuracy vary between crops?

DOI 10.3920/978-90-8686-888-9_54, © Wageningen Academic Publishers 2019

Materials and methods

Five flux towers sites (Table 1) inside five fields of irrigated corn, alfalfa, soybean and citrus were selected for this project. The data for each tower was downloaded from the AmeriFlux site (http://ameriflux.lbl.gov; last accessed 5-Dec-18), including flux and crop measurements.

Flux tower measurements

The main flux measurements used in this study were utilized in four steps. Step one: Evaluation of the flux data quality (energy balance closure). The half-hourly or hourly data of net-radiation (Rn), soil heat flux (G), sensible heat flux (H) and latent heat flux (LE) were utilized to calculate statistical regressions between the energy consumed by the plants (H+LE) and the available energy (Rn-G) as described in Wilson *et al.* (2002). The resulting values in Table 1 show good closure slope statistics for most datasets, except for the Lin tower dataset, which is in the lower edge (0.60) of accepted range (see Table 2 in Wilson *et al.*, 2002). As flux measurements in irrigated orchards are not common (https://fluxnet.fluxdata.org; last accessed 5-Dec-18), it was decided to proceed with the analysis of this dataset.
Step two: ET_c calculations. Half-hourly or hourly values of LE, obtained from the towers' data (units of w/m^2) were converted to ET_c (units of mm/day) according to Allen *et al.* (1998) and summed up to daily values.
Step three: ET_0 calculations. As weather stations were not nearby all sites, ET_0 was calculated from the flux tower meteorological data (Kamble *et al.*, 2013; Ramoelo *et al.*, 2014). The Rn and G, along with air pressure, air temperature (T_a), wind speed and vapor-pressure-deficit were utilized to calculate the Penman-Monteith reference evapotranspiration (Allen *et al.*, 1998). To test the accuracy of this method, Tw3 dataset was compared to a nearby (3.2 km) weather station from the California Irrigation Management Information System (https://cimis.water.ca.gov/WSNReportCriteria.aspx; last accessed 5-Dec-18). The regression between both ET_0 datasets showed good agreement, with a slope, intercept and R^2 of 0.87, 0.01 and 0.96, respectively. ET_0 was also summed up for each day to get a daily value. A daily K_c was calculated by dividing the daily ET_c over the daily ET_0.
Step four: Tower footprint. The footprint for every half-hourly or hourly time step was calculated according to Hsieh *et al.* (2000). The required footprint model parameters were the wind friction velocity (expressed by u_*, ustar, and calculated from the data of the ultrasonic anemometer of the flux tower), H and T_a from the flux tower database, as well as the height of the instrument and the crop (http://ameriflux.lbl.gov/data/measurement-height; last accessed 5-Dec-18). Crop height was interpolated from field measurements for each site, except for the Lin tower, where crop measurements were not available. For this site, the citrus tree height was determined according to Table 12 in Allen *et al.* (1998). These half-hourly, or hourly, footprints were used to map the daily footprint, which was the basis for

Table 1. Flux towers database.

Name	Location in USA	Time intervals of data	Crop	Selected growing seasons	Energy balance closure		
					Slope	Intercepts	R^2
Bi2	California	half-hour	corn	2017	0.73	10.33	0.97
Tw3	California	half-hour	alfalfa	2016-2018	0.79	8.25	0.97
Ne1	Nebraska	hour	corn	2002, 2004, 2006	0.85	3.36	0.96
Ne2	Nebraska	hour	soybean	2002, 2004, 2006	0.83	-0.14	0.97
Lin	California	half-hour	citrus	2010	0.60	28.51	0.85

the satellite imagery data collection. The average footprint area was 11.5 ha, with the lowest average for the alfalfa, 8.3 ha, and the highest for soybean, 16.5 ha. Tower-Kc was not calculated for days that had either missing values of LE or one of the footprint parameters.

Satellite imagery

The initial imagery database included more than 130 images from Landsat-5/7/8, Sentinel-2 and LISS-3. Landsat and LISS-3 images were downloaded from the United States Geological Survey website (https://earthexplorer.usgs.gov; last accessed 5-Dec-18) and Sentinel-2 images from the European Space Agency website (https://scihub.copernicus.eu/dhus/#/home; last accessed 5-Dec-18). Landsat images were downloaded in surface reflectance units, while Sentinel-2 images were converted to surface reflectance via the Sentinel Application Platform (SNAP). LISS-3 images were transformed to surface reflectance by applying the empirical-line method (Moran *et al.*, 2001) with Landsat-8 surface reflectance data as the reference values. Several images were removed due to clouds/cloud shadows above the flux tower sites or due to missing flux data, resulting in a total of 126 images (Table 2). For each image, 4 SVIs were calculated per field (Table 3): NDVI (Rouse *et al.*, 1974), soil adjusted vegetation index (SAVI, Huete, 1988), enhanced vegetation index (EVI2, Jiang *et al.*, 2008) and visible atmospherically resistant index (VARI, Gitelson *et al.*, 2002). These SVIs were spectrally stretched to minimize soil and saturation effects (Gillies and Carlson, 1995). Estimations of K_c were calculated according to two methods (Table 4): (1) direct transformation of SVI to K_c and (2) transformation of the SVI to vegetation fraction (Vf) and then to K_c, together with a local climatology adjustment of the maximum K_c (Allen and Pereira, 2009).

Statistical accuracy assessment

The satellite K_c estimations were compared to the flux tower K_c for each image and summarized for each crop, each sensor, and overall score for the entire monitoring system. This comparison was quantified by using bias (K_c units), root mean square error (RMSE, K_c units), normalized RMSE (nRMSE, as percentage) and coefficient of determination (R^2):

$$\text{Bias} = \sum (Si - Fi) / n \quad (1)$$

$$\text{RMSE} = [(1 / n) \times \sum (Si - Fi)^2]^{1/2} \quad (2)$$

$$\text{nRMSE} = 100 \times \text{RMSE} / (\text{maximum (Fi)} - \text{minimum (Fi)}) \quad (3)$$

where Si, Fi and n are the satellite estimation, flux measurement and the number of observations, respectively.

Table 2. Imagery database. Number of images that were utilized for each tower site.

Tower	Lansat-5	Landsat-7	Landsat-8	Sentinel-2	LISS-3
Bi2		5	7	7	3
Tw3		13	16	19	5
Ne1	10	6			
Ne2	11	3			
Lin	10	11			

Table 3. Spectral vegetation indices formulas.

Maximum	Minimum	Formula	Index
1.0	0.2	(NIR – Red) / (NIR + Red)	NDVI
0.9	0.1	2.5 × (NIR * Red) / (NIR + 2.4 × Red + 1)	EVI2
1.5	0.3	1.5 × (NIR * Red) / (NIR + Red + 0.5)	SAVI
1.0	0.1	(NIR – Red) / (NIR + Red – Blue)	VARI

Table 4. K_c calculation equations.

K_c code	Formula[1]	Source
K_c1_NDVI	1.18 × NDVI + 0.04	Tasumi *et al.*, 2006
K_c2_EVI2	(1.65 × (1 * exp(-2.25 × EVI2))) * 0.169	Nagler *et al.*, 2013
K_c3_SAVI	1.416 × SAVI + 0.017	Bausch, 1993
K_c4_ndvi-Vf	a + ((NDVI × 107.4 + 1.0) / 100) × (Kc_max * a)	The conversion of SVI to Vf is based on un-published data
K_c5_vari-Vf	a + ((VARI × 122.3 * 22.3) / 100) × (Kc_max – a)	
K_c6_evi2-Vf	a + ((EVI2 × 98.9 + 11.3) / 100) × (Kc_max * a)	

[1] Where a equals to 0.15.

Results and discussion

Accuracy of K_c estimation from SVI for the entire dataset

All six formulas of K_c estimation (Table 4) yielded an RMSE below 0.17 and nRMSE below 18%, which is compatible with an RMSE of 0.16-0.19 as described by Kamble *et al.* (2013). The K_c1, estimated solely from NDVI, reached the highest accuracy in all examined statistical parameters (Table 5).

Accuracy of K_c estimation per sensor

In general, K_c estimated directly from NDVI achieved the best results according to the RMSE, for all sensors. Among the sensors, the accuracy metrics of LISS-3 were the highest, which may be related to its small dataset (eight images) and the manual atmospheric corrections, while Landsat-5 and 7 had the lowest accuracy metrics. Nonetheless, the RMSE values of K_c1 for these sensors are below 0.11 (Table 6).

Table 5. Performance of K_c estimations for the entire dataset.

Metrics	K_c1 NDVI	K_c2 EVI2	K_c3 SAVI	K_c4 NDVI	K_c5 VARI	K_c6 EVI2
Bias	0.003	0.099	0.059	0.042	0.047	0.038
RMSE	0.097	0.169	0.136	0.132	0.140	0.144
nRMSE	10	18	14	14	15	15
R^2	0.87	0.80	0.86	0.76	0.75	0.62

Table 6. Performance of K_c estimations for each sensor

Sensor	Metrics	K_c1 NDVI	K_c2 EVI2	K_c3 SAVI	K_c4 NDVI	K_c5 VARI	K_c6 EVI2
Landsat-5	Bias	-0.007	0.054	0.041	-0.013	-0.002	-0.011
31 images	RMSE	0.107	0.172	0.146	0.128	0.139	0.149
	nRMSE	13	20	17	15	16	18
	R^2	0.87	0.87	0.87	0.73	0.72	0.56
Landsat-7	Bias	-0.002	0.095	0.061	0.002	0.013	-0.022
38 images	RMSE	0.105	0.167	0.143	0.144	0.149	0.148
	nRMSE	12	19	17	17	17	17
	R^2	0.80	0.74	0.80	0.69	0.69	0.56
Landsat-8	Bias	0.042	0.155	0.103	0.127	0.128	0.113
23 images	RMSE	0.088	0.198	0.139	0.147	0.149	0.149
	nRMSE	11	24	17	18	18	18
	R^2	0.90	0.77	0.90	0.90	0.89	0.88
Sentinel-2	Bias	-0.007	0.110	0.046	0.081	0.085	0.104
26 images	RMSE	0.089	0.150	0.125	0.112	0.116	0.134
	nRMSE	10	17	14	13	13	15
	R^2	0.90	0.87	0.90	0.91	0.91	0.89
LISS-3	Bias	-0.020	0.107	0.029	0.069		0.075
8 images	RMSE	0.057	0.124	0.073	0.089		0.124
	nRMSE	8	18	11	13		18
	R^2	0.95	0.93	0.95	0.94		0.90

The power of utilizing different sensors to estimate K_c for a specific location is presented in the 2017 dataset of the alfalfa field in Tw3 site (Figure 1). The importance of employing imagery from different sensors in one monitoring system is highlighted in the short growing cycles of this crop by capturing the short phenological stages.

Accuracy of K_c estimation per crop

The performance of K_c estimations from satellite imagery for the examined field crops (Table 7) revealed high accuracy for corn, alfalfa and soybean.

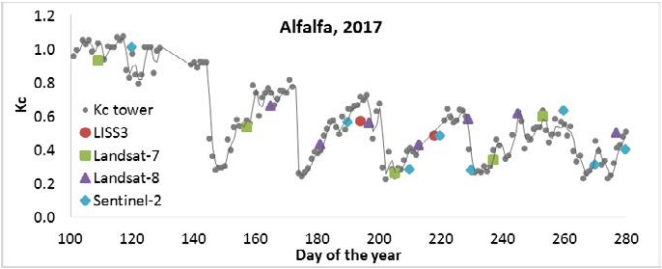


Figure 1. Daily estimated K_c1 values of the alfalfa field (Tw3 site) during 2017 season. K_c-tower (gray dots) was regressed with a moving average model.

Table 7. Performance of K_c estimations for each crop.

Crop	Metrics	K_c1 NDVI	K_c2 EVI2	K_c3 SAVI	K_c4 NDVI	K_c5 VARI	K_c6 EVI2
Corn	Bias	0.031	0.039	0.091	0.124	0.134	0.079
38 images	RMSE	0.091	0.155	0.157	0.142	0.155	0.120
	nRMSE	10	17	17	15	17	13
	R^2	0.86	0.93	0.96	0.97	0.97	0.91
Alfalfa	Bias	-0.021	0.139	0.032	0.065	0.073	0.088
58 images	RMSE	0.085	0.173	0.099	0.102	0.109	0.137
	nRMSE	11	22	13	13	14	17
	R^2	0.85	0.76	0.85	0.86	0.86	0.80
Soybean	Bias	-0.009	0.065	0.035	0.059	0.075	0.076
14 images	RMSE	0.088	0.158	0.129	0.086	0.107	0.125
	nRMSE	13	24	20	13	16	19
	R^2	0.94	0.92	0.94	0.94	0.90	0.79
Citrus	Bias	0.020	0.132	0.082	-0.178	-0.173	-0.192
21 images	RMSE	0.134	0.186	0.174	0.191	0.188	0.203
	nRMSE	38	53	50	54	53	58
	R^2	0.22	0.25	0.21	0.31	0.30	0.31

The range of R^2 for K_c1 is 0.85-0.94, which is comparable to R^2 values of 0.91 and 0.90 for corn and soybean, respectively, as presented by Kamble *et al.* (2013). However, the accuracy of K_c in citrus was very low as presented by R^2 (0.22 for K_c1) and the nRMSE (38%). This anomaly can be related to the small dataset for this crop (one plot, one tower, one growing season), the small seasonal changes in K_c values, and to the lowest quality of the flux data as previously explained. For all crops, K_c1 method achieved the highest accuracy, with random errors throughout the growing season (Figure 2).

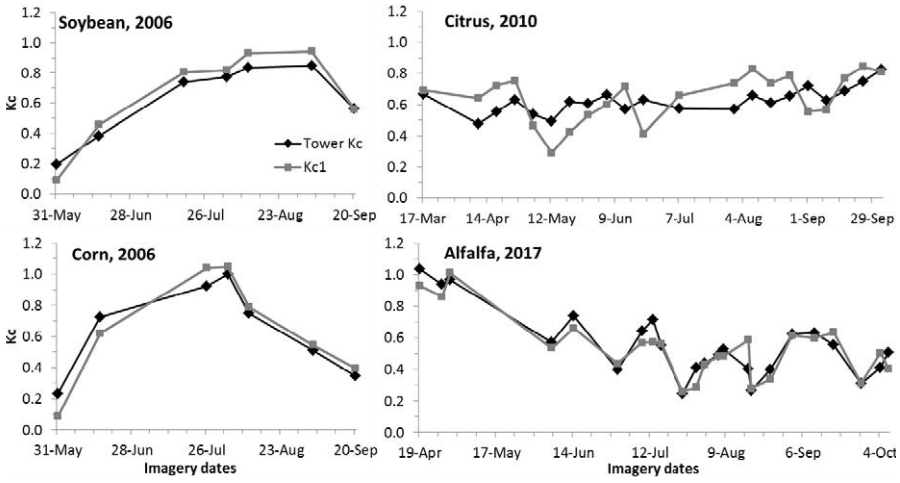


Figure 2. Daily estimated K_c1 values (grey squares) and flux tower K_c (black diamonds) for each crop.

Conclusions

The goal of this study was to assess the accuracy of K_c estimation using in-orbit multi-spectral satellite imagery. To this end, a large dataset was established over the course of several growing seasons. The dataset consisted of measured data from five flux towers, four crop types and 126 satellite images from five different sensors. The results show that K_c estimations derived from NDVI had the best accuracy, with an RMSE below 0.1 and nRMSE of 10%, while K_c estimations for citrus showed lower accuracy. Therefore, this study concludes that a good estimation of K_c can be achieved by using satellite imagery for intelligent irrigation, even when the irrigation scheduling system combines different sensors.

Acknowledgements

The authors wish to thank Prof. Dennis Baldocchi (University of California-Berkeley) for his continuous help.

References

Allen, R. G., Pereira, L. S., 2009. Estimating crop coefficients from fraction of ground cover and height. Irrigation Science. 28, 17-34.

Allen, R. G., Pereira, L. S., Raes, D., Smith, M., 1998. FAO Irrigation and drainage paper No 56: Crop Evapotranspiration. Rome, Italy: FAO.

Allen, R. G., Tasumi, M., Morse, A., Trezza, R., Wright, J. L., Bastiaanssen, W., *et al.*, 2007. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC) – applications. Journal of Irrigation and Drainage Engineering. 133, 395-406.

Anderson, M. C., Kustas, W. P., Norman, J. M., Hain, C. R., Mecikalski, J. R., Schultz, L., *et al.*, 2011. Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. Hydrology and Earth System Sciences. 15, 223-239.

Bausch, W. C., 1993. Soil background effects on reflectance-based crop coefficients for corn. Remote Sensing of Environment. 46, 213-222.

Gillies, R. R., Carlson, T. N., 1995. Thermal remote sensing of surface soil water content with partial vegetation cover for incorporation into climate models. Journal of Applied Meteorology. 34, 745-756.

Gitelson, A. A., Kaufman, Y. J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetation fraction. Remote Sensing of Environment. 80, 76-87.

Hsieh, C., Katul, G., Chi, T., 2000. An approximate analytical model for footprint estimation of scalar fluxes in thermally stratified atmospheric flows. Advance in Water Resources. 23, 765-772.

Huete, A., 1988. A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment. 25, 295-309.

Hunsaker, D. J., Pinter, P. J., Barnes, E. M., Kimball, B. A., 2003. Estimating cotton evapotranspiration crop coefficients with a multispectral vegetation index. Irrigation Science. 22, 95-104.

Jiang, Z., Huete, A. R., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. Remote Sensing of Environment. 112, 3833-3845.

Johnson, L. F., Trout, T. J., 2012. Satellite NDVI assisted monitoring of vegetable crop evapotranspiration in california's san Joaquin Valley. Remote Sensing. 4, 439-455.

Kamble, B., Kilic, A., Hubbard, K., 2013. Estimating crop coefficients using remote sensing-based vegetation index. Remote Sensing 5, 1588-1602.

Moran, M. S., Bryant, R., Thome, K., Ni, W., Nouvellon, Y., Gonzalez-Dugo, M. P., *et al.*, 2001. A refined empirical line approach for reflectance factor retrieval from Landsat-5 TM and Landsat-7 ETM+. Remote Sensing of Environment. 78, 71-82.

Nagler, P. L., Glenn, E. P., Nguyen, U., Scott, R. L., Doody, T., 2013. Estimating riparian and agricultural actual evapotranspiration by reference evapotranspiration and MODIS enhanced vegetation index. Remote Sensing. 5, 3849-3871.

Nagler, P. L., Morino, K., Murray, R. S., Osterberg, J., Glenn, E. P., 2009. An empirical algorithm for estimating agricultural and riparian evapotranspiration using MODIS enhanced vegetation index and ground measurements of ET. I. Description of method. Remote Sensing. 1, 1273-1297.

Ramoelo, A., Majozi, N., Mathieu, R., Jovanovic, N., Nickless, A., Dzikiti, S., 2014. Validation of global evapotranspiration product (MOD16) using flux tower data in the African savanna, South Africa. Remote Sensing. 6, 7406-7423.

Rouse, J. W., Haas, R. H., Schell, J. A., Deering. D. W., 1974. Monitoring vegetation systems in the Great Plains with ERTS, in: Third Earth Resources Technology Satellite-1 Symposium. Washington, USA:National Aeronautics and Space Administration, pp. 309-316.

Tasumi, M. Allen, R. G., Trezza R., 2006. Calibrating satellite-based vegetation indices to estimate evapotranspiration and crop coefficients. In: Ground Water and Surface Water Under Stress: Competition, Interaction, Solutions. Denver, USA: U.S. Committee on Irrigation and Drainage, 103-112.

Toureiro, C., Serralheiro, R., Shahidian, S., Sousa, A., 2017. Irrigation management with remote sensing: Evaluating irrigation requirement for maize under Mediterranean climate condition. Agriculture Water Managment. 184, 211-220.

Wang, J., Sammis, T. W., Andales, A. A., Simmons, L. J., Gutschick, V. P., Miller, D. R., 2007. Crop coefficients of open-canopy pecan orchards. Agriculture Water Managment. 88, 253-262

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., *et al.*, 2002. Energy balance at FLUXNET sites. Agriculture Forest Meteorology. 113, 223-243.

Can time series of multispectral satellite images be used to estimate stem water potential in vineyards?

Y. Cohen[1*], P. Gogumalla[1], I. Bahat[1,3], Y. Netzer[4,5], A. Ben-Gal[6], I. Lenski[2], Y. Michael[2] and D. Helman[1,2]
[1]*Institute of Agricultural Engineering, Agricultural Research Organization (Volcani Center), P.O. Box 15159 Rishon LeZion 7505101, Israel; yafitush@volcani.agri.gov.il*
[2]*Department of Geography and the Environment, Bar-Ilan University, Ramat Gan, Israel*
[3]*The Robert H Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot 76100, Israel*
[4]*Department of Agriculture and Oenology, Eastern R&D Center, Israel*
[5]*Department of Chemistry & Biotech Engineering, Ariel University, Ariel 40700, Israel*
[6]*Department of Soil, Water and Environmental Sciences, Agricultural Research Organization Gilat Research Center, Mobile post Negev, Israel*

Abstract

Vegetation indices (VI) derived from a new generation of high-spatial resolution satellites, namely Planet and Sentinel-2, were tested as proxies for stem water potential (Ψ-stem) in commercial vineyards. Multivariable linear regression models were developed from the Planet and Sentinel-2 data and in-situ Ψ-stem measurements in 82 vineyards in Israel, providing Ψ-stem estimates every 1-2 weeks. With multivariable regression VIs in the VIS-NIR region (Planet) and moisture VIs in the NIR-SWIR region (Sentinel-2) were correlated with in-situ Ψ-stem measurements. In general, Sentinel-based models performed better (higher correlations and higher ability to capture temporal and spatial variability) than Planet-based models. The trade-off between spectral, spatial and temporal resolutions of these two satellites are further discussed.

Keywords: Planet-labs, Sentinel-2, stem water potential, Google earth engine

Introduction

Viticulture has been the pride of Israel since biblical times. Due to wine vineyards being conservative water users and very profitable, the area under grape cultivation nationally is presently increasing. Two aspects primarily influence the profits made by commercial vineyards. The first is the level of fruit production, and the second is berry quality. Irrigation management influences both quantity and quality, especially of red wine grapes. Deficit irrigation is typically practiced in Israeli wine grape vineyards. Vineyard-scale deficit water management, designed to optimize yields and quality, depends on continuous water status monitoring. In order to achieve an accurate picture of the water status variability throughout the entire vineyard area, continuous stem water potential (Ψ-stem) measurements are required from a large number of vines spread within the vineyard. This task is time and labor consuming, requiring substantial human resources. Remote sensing can be an effective tool to address this need for monitoring. Thermal cameras mounted on airplanes or drones can provide information on plant water status through the use of the crop water stress index (CWSI), which has been shown to correlate well with Ψ-stem in vineyards (e.g Möller *et al.*, 2007; Baluja *et al.*, 2012). Acquiring thermal aerial images from drones is expensive due to relatively low area covered per unit time. To address this issue this study focused on two objectives. The first objective was to examine whether the VIs retrieved from a time series of images from two new satellite platforms, namely Planet-labs and Sentinel-2 can be used to estimate Ψ-stem in vineyards through

DOI 10.3920/978-90-8686-888-9_55, © Wageningen Academic Publishers 2019

the irrigation season. The second objective was to compare between the two satellite platforms as they are different in their spectral, spatial and temporal resolution.

Materials and methods

Study region

The study area included 82 commercial vineyards located in the Golan Heights, Upper Galilee and Judean hills. As part of a larger study in precision irrigation, in one of the vineyards located in the Judean Hills, Mevo Beitar (35° 6' 10' N, 31° 43' 46' E), in-field water status monitoring was conducted in 20 cells of 20×20 m. The dominant climate in the study area is Mediterranean with mean annual rainfall varying from approximately 500 mm/yr in the Judean hills to 800 mm/yr in the Golan Heights. The 81 commercial vineyards were irrigated by surface drip irrigation. Irrigation amounts were decided based on the daily computed ETo (reference evapotranspiration) provided by the Israel Meteorological Service (http://www.ims.gov.il/IMSEng). Weekly or bi-weekly midday Ψ-stem measurements from all 81 vineyards collected by managing viticulturists in the 2017 season were available for the study. Irrigation in the Mevo-Beitar vineyard started on day of year (DOY) 131 and ended at DOY 249. Ψ-stem was measured prior to irrigation at solar noon (from 12:00 to 14:30), using a portable pressure chamber in every cell (model Arimad 2, Kfar Charuv, Israel and Arimad 3000, MRC, Hulon Israel) according to the procedure described in Boyer (1995). Deficit irrigation applications ranged between 14 and 25% of calculated full vine evapotranspiration (ETc) throughout the season using corrections to the crop coefficient based on Ψ-stem and leaf area index measurements (Munitz *et al.*, 2017).

Satellite data

Planet-labs (Houborg and McCabe, 2016) is a commercial company of a constellation of small nano-satellites also called 'Doves'. Planet satellites provide high spatial resolution data (3 m) and daily global revisit time. Planet images has 4 spectral bands (3 in the visible (VIS) region and one in the near-infrared (NIR) region). Using these bands, commonly-used vegetation indices, i.e. NDVI, SAVI, EVI and GNDVI, were derived from Planet data for the 2017 growing season (Helman *et al.*, 2018; Table 1). The Sentinel-2 mission carries a multi-spectral instrument which measures reflected solar radiances in 13 spectral bands ranging from the VIS to the shortwave infrared (SWIR) bands with a 5-days revisit time and a spatial resolution of 10-60 m over land and coastal areas (Drusch *et al.*, 2012). In addition to the vegetation indices mentioned above, moisture based indices (MSI, NDMI, NMSI1, NMSI2) were derived, which specifically use spectral bands in the NIR and SWIR region (Table 1).

Data analysis

Time series of VIs from Planet and Sentinel-2 were generated using Java Script API on Google Earth Engine platform (GEE) (Gorelick *et al.*, 2017), which is a freely available cloud-based platform for processing geospatial datasets. Formulation of the different VIs used in this study are presented in Table.1. An empirical relationship was developed between in-situ Ψ-stem and the VIs (Table 1) in 81 vineyards using linear regression models. In addition, in order to represent the soil contribution to the VI, ΔVI was calculated representing the difference in fractional vegetation cover between peak of the growing season (vines with a fully developed canopy) and prior to growing season (minimum value representing soil and understory herbaceous vegetation cover) (Helman *et al.*, 2018). The weekly maximum and minimum VI values were extracted (VI max and VI min, respectively) from the daily Planet-derived VIs, which were added to the multivariable linear regression models. Because of the less frequent revisit time of Sentinel-2 (every 5 days), a single VI value was taken in each week and used with the ΔVI in the multivariable linear regression model (i.e. no VI max and VI min were included). The day of year (DOY) was used in both Planet and Sentinel-based models. Two multi linear regression model types were developed:

Table 1. List of vegetation indices derived from Planet and Sentinel-2 data and their formula.

	Index name	Formula	Reference
Vegetation indices (VIs)	EVI (enhanced vegetation index)	$2.5\times(\frac{NIR-Red}{NIR+6*Red-7.5*Blue+1})$	(Huete *et al.*, 2002)
	G/NDVI (green normalized difference vegetation index)	$\frac{NIR-Green}{NIR+Green}$ *or* $\frac{NIR-Red}{NIR+Red}$	(Gitelson & Merzlyak, 1997) or (Tucker, 1979)
	SAVI (soil adjusted vegetation index)	$\left(\frac{NIR-Red}{NIR+Red+L}\right)*(1+L)$	(Huete, 1988)
Moisture based indices (MVIs)	MSI (moisture stress index)	$\frac{NIR}{SWIR1}$	(Gao, 1996)
	NDMI (normalized drought monitoring index)	$\frac{NIR-(SWIR1-SWIR2)}{NIR+(SWIR1-SWIR2)}$	(Wang & Qu, 2007)
	NMSI1 (normalized moisture stress index1/2)	$\frac{NIR-SWIR1}{NIR+SWIR1}$ *or* $\frac{NIR-SWIR2}{NIR+SWIR2}$	(Gao, 1996) or (Fensholt & Sandholt, 2003)

1. Multivariable Linear Model for each VI (named after the index):
   Variables for Planet model – DOY, VI avg, VI max, VI min and ΔVI Variables for Sentinel-2 model – DOY, VI, and ΔVI
2. Multi-Index Multivariable Linear Model:
   Variables for Planet model – DOY, All VIs and their derivatives:, max, and min.
   Variables for Sentinel-2 model – DOY, All VIs and MVIs and their ΔVI.

The multi-index multivariable linear models are hereafter named MLM-Planet and MLM-Sentinel-2. It should be noted that, because of the higher temporal resolution of Planet, it was available to larger number of in-season Ψ-stem measurements than Sentinel-2 (n=951 vs n=794). Coefficient of determination (R^2) and RMSE were calculated to assess the robustness of the models using the various features of the VIs and MSIs. Best performing models were selected for further analysis. These models were validated using the data collected from the 20 cells in Mevo-Beitar Vineyard. The standard deviation (STD) was calculated and F-test was performed to the measured and estimated seasonal mean of Ψ-stem of individual vineyards (n=81) to test the capability of the models to capture the inter-vineyards variability. In addition to the whole-season evaluation of the models, weekly percent deviations were calculated between estimated and measured Ψ-stem for the above mentioned linear models. Furthermore, residuals (seasonal mean of estimated minus measured Ψ-stem) of the models from the two satellites were mapped for Mevo-Beitar cells to compare their capability to depict the in-field variability.

Results

Seasonal behavior of VIs and Ψ-stem

Changes in VIs and MVIs derived from Planet and Sentinel-2 along with the measured Ψ-stem over the course of the growing season in the 82 vineyards. Time series patterns of the VIs derived from

Planet and some of the MVIs derived from Sentinel-2 were similar to those of Ψ-stem measurements in the 82 vineyards (Figure 1A, 1C). The VIs derived from Sentinel-2 data, however, were less indicative of the Ψ-stem change during the season (Figure 1B).

Multivariable linear model

The best single VI models from each of the Planet (SAVI) and Sentinel-2 (MSI) data and the MLM models from both satellite are compared in Table 2. In general, the two Sentinel-2 models performed better than Planet-based models, having higher R^2 and lower RMSE and F-statistics (Table 2 and 3). Yet, the variances over the 81 vineyards of the estimated Ψ-stem values based on each of the 4 models were significantly different from the variance of the measured Ψ-stem (Table 3). This implies that the models did not fully capture the inter-vineyards water status variability. When the intercept of the regression line was forced to 0 to check for 1:1 relationships, the slopes of all models except for MSI-Sentinel-2 were not different from 1 but the Pearson correlations of the Planet models were 0.64 and 0.56 for the SAVI and MLM, respectively, and 0.78 and 0.73 for the MSI and MLM, respectively,

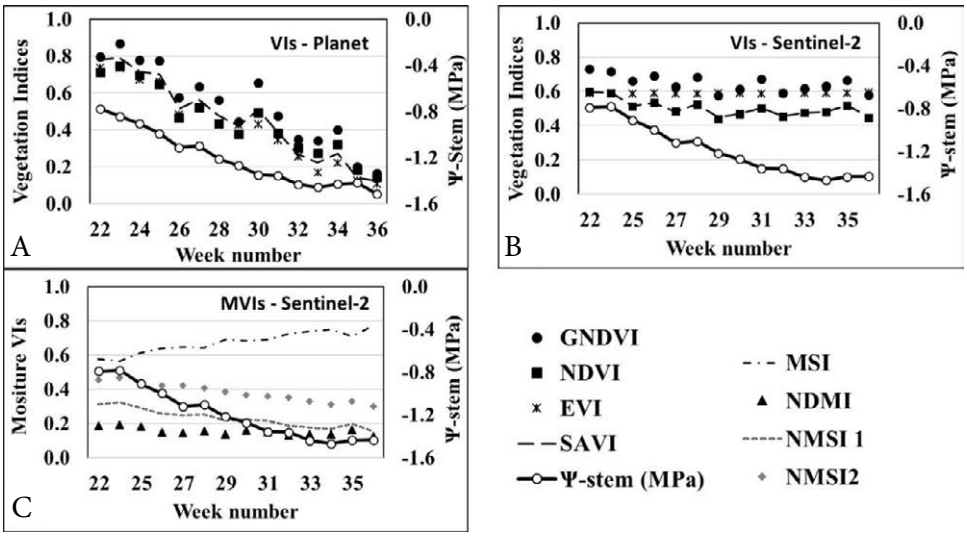


Figure 1. Weekly time series of the VIs and MVIs and measured stem water potential (Ψ-stem) in the 82 vineyards (including Mevo-Beitar) for (A) Planet satellite data; (B) and (C) Sentinel-2 satellite data during the 2017 growing season.

Table 2. R^2 and RMSE for calibration (81 vineyards) and validation (20 cells in Mevo-Beitar vineyard) data sets based on best performing models.

Model	N	Variables	Calibration		Validation	
			Adj. R2	RMSE	Adj. R2	RMSE
SAVI-Planet	951	DOY/SAVI/SAVI max,min/ΔSAVI	0.57	0.22	0.63	0.18
MLM-Planet	951	DOY/ΔNDVI/SAVI/NDVI max, min	0.58	0.22	0.63	0.18
MSI-Sentinel-2	794	DOY/MSI/ΔMSI	0.69	0.18	0.67	0.23
MLM-Sentinel-2	794	DOY/ΔEVI//MSI/NMSI1/ ΔNMSI2	0.65	0.19	0.63	0.18

of the Sentinel-2 models (data not shown). Further comparison was made only between MLM-Planet and MLM-Sentinel-2. The weekly measured Ψ-stem in the 81 vineyards and the relative deviations of the estimated from the measured Ψ-stem based on the MLM models are presented in Table 4. It shows that at a weekly basis, MLM-Sentinel-2 outperforms MLM-Planet with lower deviations than MLM-Planet in 12 out of 14 weeks. Generally, high relative deviations were obtained for the beginning and the end of the irrigation season. It should be also noted that the range of relative deviations (less than 15%) may be acceptable for in-season irrigation management.
The residuals between the seasonal mean of measured and estimated Ψ-stem of individual cells in the Mevo-Beitar vineyard is plotted in Figure 2. The difference ranged between -0.14 and 0.21 MPa (under and over- estimation of Ψ-stem, respectively) for MLM Sentinel-2 and -0.16 and 0.18 MPa for MLM-Planet. Most of the cells were mapped in a range of ±1.5 Mpa, with MLM-Sentinel-2 having more cells with small residuals (~ zero) than MLM-Planet.

Table 3. The standard deviation (STD) of the average seasonal estimated and measured stem water potential of individual vineyards and F-test results for equal variances.

Linear models	STD-estimated	STD-measured	F-statistic	*P*-value
SAVI-Planet	0.07	0.17	6.85	<0.0001
MLM-Planet	0.07	0.17	5.9	<0.0001
MSI-Sentinel-2	0.13	0.18	2	0.003
MLM-Sentinel-2	0.14	0.18	1.66	0.03

Table 4. Weekly measured Ψ-Stem and the percent deviation of the estimated Ψ-Stem.[1]

		% Deviation from measured Ψ-Stem	
Week	Measured Ψ-Stem	MLM-Planet	MLM-Sentinel-2
22	-0.74	12.4	**-2.3**
23	-0.76	9.0	**5.2**
25	-0.89	8.4	**1.3**
26	-0.96	**2.2**	6.1
27	-1.08	-2.0	**-1.3**
28	-1.09	**1.4**	**1.4**
29	-1.18	**-1.3**	2.3
30	-1.27	-4.0	**-3.5**
31	-1.32	**-5.5**	**-5.5**
32	-1.37	-3.3	**-0.6**
33	-1.41	0.9	**-0.3**
34	-1.45	2.1	**0.3**
35	-1.44	4.8	**1.4**
36	-1.43	11.0	**9.2**

[1] Bold values indicates best estimation.

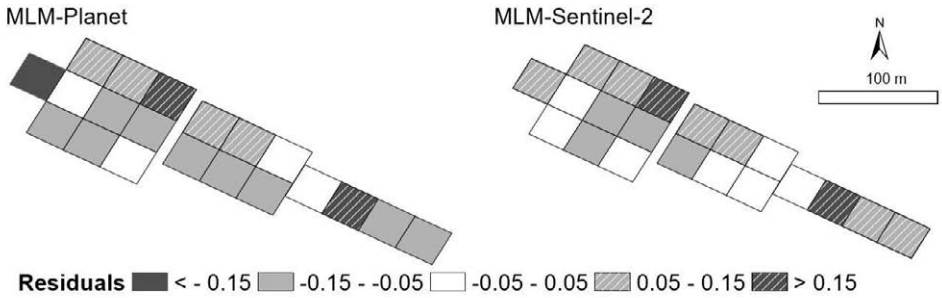


Figure 2. Residuals map between seasonal mean of measured and estimated Ψ-stem for Mevo-Beitar derived using MLM-Planet (left) and MLM-Sentinel-2 (right).

Discussion

The new generation of satellites opens a new era in the use of remote-sensing in precision agriculture. Two satellite constellations, namely, Planet Labs and Sentinel-2, with different spatial, temporal and spectral resolutions were evaluated for their capability to estimate water status in vineyards. Despite their differences, both of them demonstrated moderate success in water status estimation in 82 Israeli vineyards. This is probably due to their interchangeable advantages. Different features of VIs in VIS-NIR regions extracted from Planet data were correlated with water status while similar VIs extracted from Sentinel-2 were not (Figure 1). This difference may be attributed to the higher temporal and spatial resolutions of the Planet images. Moisture vegetation indices that utilize the SWIR bands of the Sentinel-2 were correlated with water status of the vineyards. Despite the capability of the frequent-and-high-spatial-resolution of the Planet data to capture some of the water status variability simply using VIs in the VIS-NIR region, the MSIs from the Sentinel-2 performed better in this matter. The Sentinel-2 models obtained better correlations for the 81 vineyards (calibration) as well as for the 20 cells of Mevo-Beitar vineyard (validation). They were more successful in capturing the inter-vineyards variability of the water status as reflected by their higher STD and lower F-statistic (Table 3). The Sentinel-2 models also significantly better estimated weekly Ψ-stem (Table 4). This means, that in a certain constellation, it might be used for in-season irrigation management (Cohen *et al.*, 2017; Cohen and Alchanatis, 2019). Finally, the Sentinel-2 models better captured the in-field water status variability in Mevo-Beitar (Figure 2) as reflected by lower residuals for the individual cells. Improved weekly and in-field variability water status estimations were obtained by the Sentinel-2 models, despite its lower temporal and spatial resolutions, compared to the Planet Labs. The higher spectral resolution with specific bands near those of water absorption may therefore be preferred over high spatio-temporal resolution. Continuing work utilizes a similar database from the 2018 season to enable more robust conclusions.

Conclusions

A large database of weekly Ψ-stem measurements from 82 vineyards in Israel was used to examine the capabilities of Planet and Sentinel-2 images to estimate water status in veinyards. The high spatio-temporal resolution of Planet enabled correlation of VIs in the VIS-NIR region with water status. Yet, the use of SWIR bands unique to Sentinel-2, allowed better estimation of water status in various spatial and temporal scales. In the future it is suggested to use the Planet time-series data to estimate, map and monitor vine vegetative growth and to combine this with Sentinel-2 time-series data to estimate water status.

Acknowledgements

This research was supported by the Fund of the Chief Scientist of the Ministry of Agriculture in Israel (Grant No. 20-12-0030). We thank the Israeli wineries Barken, Golan Heights and Carmel for sharing their data with us.

References

Baluja, J., Diago, M.P., Balda, P., Zorer, R., Meggio, F., Morales, F., *et al.,* 2012. Assessment of vineyard water status variability by thermal and multispectral imagery using an unmanned aerial vehicle (UAV). Irrigation Science 30 511-522.

Boyer, J.S., 1995. Measuring the water status of plants and soils. Academic Press, Inc.

Cohen Y., Agam, N., Klapp, I., Karnieli A., Beeri, O., Alchanatis V., *et al.,* 2017. Future approaches to facilitate large-scale adoption of thermal based images as key input in the production of dynamic irrigation management zones. Advances in Animal Biosciences 8 546-550.

Cohen, Y., and Alchanatis, V., 2019. Spectral and spatial methods for hyperspectral and thermal image analysis to estimate biophysical and biochemical properties of agricultural crops, in: P. S. Thenkabail, Lyon, G.J., Huete, A. (Ed.), Hyperspectral Remote Sensing of Vegetation (Second Edition, Four-Volume-Set), CRC Press-Taylor and Francis group, Boca Raton, London, New York.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, *et al.,* 2012. Sentinel-

ESA's optical high-resolution mission for GMES operational services. Remote Sensing of Environment 120 25-36.

Fensholt, R. and Sandholt, I., 2003. Derivation of a shortwave infrared water stress index from MODIS near- and shortwave infrared data in a semiarid environment. Remote Sensing of Environment 87 111-121.

Gao, B.-C., 1996. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment 58 257-266.

Gitelson, A.A. and Merzlyak, M.N., 1997. Remote estimation of chlorophyll content in higher plant leaves. International Journal of Remote Sensing 18 2691-2697.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote sensing of Environment 202 18-27.

Helman, D., Bahat, I., Netzer, Y., Ben-Gal, A., Alchanatis, V., Peeters, A. *et al.,* 2018. Using Time Series of High-Resolution Planet Satellite Images to Monitor Grapevine Stem Water Potential in Commercial Vineyards. Remote Sensing 10 1615-1637.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X. and Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote sensing of Environment 83 195-213.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote sensing of Environment 25 295-309.

Houborg, R. and McCabe, M.F., 2016. High-resolution NDVI from Planet's constellation of earth observing nano-satellites: a new data source for precision agriculture. Remote Sensing 8 768-787.

Möller M., Alchanatis V., Cohen Y., Meron M., Tsipris J., Naor A., *et al.,* 2007. Use of thermal and visible imagery for estimating crop water status of irrigated grapevine. Journal of Experimental Botany 58 827-838.

Munitz, S., Netzer, Y. and Schwartz, A., 2017. Sustained and regulated deficit irrigation of field-grown Merlot grapevines. Australian Journal of Grape and Wine Research 23 87-94.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote sensing of Environment 8 127-150.

Wang, L. and Qu, J.J., 2007. NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. Geophysical Research Letters 34 1-5.

Interaction between soil variability and maize nitrogen status assessment from Sentinel-2

A. Crema[1,3*], G. Vandini[2], M. Boschetti[3], F. Nutini[3], D. Cillis[4] and R. Casa[1]
[1]*Department of Agricultural and Forestry scieNcEs (DAFNE), University of Tuscia, Via San Camillo de Lellis, 01100 Viterbo, Italy*
[2]*University of Bologna, Via Zamboni 33, Bologna, Italy*
[3]*Institute for Electromagnetic Sensing of the Environment (IREA), Consiglio Nazionale delle Ricerche (CNR), Via Bassini 15, Milano, Italy; crema.a@irea.cnr.it*
[4]*IBF Servizi S.p.A. Via Cavicchini, 2, 44037 Jolanda di Savoia (Fe), Italy*

Abstract

Crop nitrogen (N) status assessment from remote sensing can be based on the estimation of the nitrogen nutrition index (NNI). It is possible to estimate NNI directly from vegetation indices (VI) or indirectly through the estimation of biomass or leaf area index (LAI) and canopy nitrogen content. In this study, different approaches for the estimation of NNI in maize were assessed. LAI retrieved from Sentinel-2 over-estimated measured LAI introducing excessive uncertainties in the estimation of critical N from allometric relationships between LAI and biomass. Vegetation indices including the red edge bands were more promising in the direct estimation of NNI.

Keywords: nitrogen nutrition index, NNI, maize, biophysical variables, vegetation index

Introduction

Soil fertility and primary nutrient elements availability are essential for yield and quality production of cereal crops. Nitrogen is often the most important element and limiting factor for plant growth. Crop nitrogen (N) status assessment relies on the estimation of the nitrogen concentration during the season and can be described through the use of the nitrogen nutrition index (NNI), defined as the ratio of actual to critical N content in the above-ground biomass (Lemaire *et al.*, 2008). It is possible to estimate NNI from remote sensing, either directly from vegetation indices (Zhao *et al.,* 2018), or indirectly through the estimation of crop biophysical variables such as biomass or LAI and actual nitrogen content (Cammarano *et al.*, 2014). Estimation methods using remote sensing are extremely useful, since they can provide spatialized data, suitable for the development of precision fertilisation applications (Nutini *et al.*, 2018).
The objectives of this study were to investigate the influence of nitrogen fertilization and soil variability in maize growth, on the estimation of NNI from remote sensing. Direct measurements of plant N concentration, for different N fertilisation treatments, were conducted to assess the plant N status of maize using the theory of critical N (Nc) dilution (Plénet and Lemaire, 2000). Satellite Sentinel-2 acquisitions provided the information to analyse spatial and spectral variability in crop status. Satellite images were exploited to estimate directly and indirectly NNI in order to find the relationship with the vegetation variables, in order to compare different estimation approaches.

Materials and methods

Study area and experimental design
During the 2018 maize growth season, a field experiment was conducted at the Bonifiche Ferraresi farm, located in Jolanda di Savoia (lat. 44°52'26' N, long. 11°56'24' E, elev. -3 m) in the North-East of Italy. This area is characterized by soils originating from reclaimed land after drainage during

DOI 10.3920/978-90-8686-888-9_56, © Wageningen Academic Publishers 2019

the early 20th century. Large and sudden changes in soil properties, occurring in the area, are one of the most important issues that farmers have to manage for the implementation of precision farming techniques. In 2015, the farm conducted a survey for the determination of soil properties using the Automatic Resistivity Profiler (ARP) geophysical sensor (Priori *et al.*, 2013) as well as soil sampling. From these data, a soil map of the whole farm was derived by the surveying company (SOING, Livorno, Italy). This information was used to identify areas with contrasting soil properties potentially related to different productivity. An experimental field of 16 ha was selected, in which four different classes (CDS2, JOL3, JOL1, FOR1) of the soil map had been identified.
Soil samples were collected at 36 points in the field from 0 to 0.30 m soil layer, before sowing the maize crop on 17th April. The sampling sites were equally placed into the 4 classes found in the field according to the soil map. Samples were used to measure soil texture, soil organic matter (SOM), total nitrogen, total organic carbon (TOC), pH, C/N. The results of soil laboratory analyses were assessed using one way-analysis of variance (ANOVA) and Tukey post hoc test ($P<0.05$) to assess differences in soil properties among soil classes. All statistical tests were conducted using Rstudio 1.1.383 software (Rstudio Team, Boston, USA)
The size of the plots, 36 m width by 50 m length, was dimensioned based on the width of a variable rate fertilizer spreader Sulky model x36 (Sulky-Burel, Chateaubourg, France) and sufficiently large to contain at least 10 pixels of a 10 m resolution Sentinel-2 image, considering an internal buffer area. Prescriptions maps were used for applying the prescribed N rates to the different plots, using a John Deere StarFire receiver (John Deere, Moline, USA) for GNSS signal, JD GreenStar 2630 display and a Sulky x36 spreader to apply the fertilizer.
Five different N fertilizer treatments were applied to the plots with 4 replicates, one for each identified soil class. The N treatments (kg N/ha) were: N1(10), N2 (90), N3 (180), N4 (300) and N5 (350). The total fertilisation was performed in 5 applications. The N treatments were: 10, 90, 180, 300 and 350 kg N/ha. The total fertilisation was performed in 5 applications. A uniform basal dressing was applied before sowing, with 200 kg/ha Pheoscor and 125 kg/ha of 8-18 (10 kg N/ha) starter fertilizer were uniformly placed at planting. During the crop development phases, three top dressing applications with urea were executed at phenological stages of 2nd leaf, 6th leaf and of beginning of flowering. The last N application at flowering was carried out with a uniform distribution of nitrogen due to the absence of VRT machinery for liquid fertilizer distribution.

Field data collection

In situ measurements were performed at 4 different dates (9th of May, 29th of May, 18th of June, 3rd of July). The first campaign was carried out before the first top dressing at 2nd leaf phenological phase (BBCH 12), while the other 3 campaigns were carried out right after the fertilizations (BBCH 14, BBCH 16, BBCH 63). A Trimble GNSS receiver was used to place each sampling area, represented by an ESU (Elementary Sample Unit) of 20 m^2, at the centre of each plot. Above ground biomass was measured by harvesting 8 plants for each ESU, which were then dried in oven at 80 °C until constant weight. The whole plant samples were then finely ground and used for N content determination using the Kjeldhal method. Leaf chlorophyll and nitrogen balance index (NBI) readings for 20 representative plants were acquired with the Dualex (Force-A, Orsay, France) for each ESU at each sampling date. A Canon digital camera with a fish-eye lens was used to acquire 16 hemispherical photographs in each ESU at each date, subsequently processed with the Can-Eye V6.494 software (https://www6.paca.inra.fr/can-eye/) to estimate crop LAI. The maize critical N dilution curve was calculated as the ratio between actual and critical N content, adopting the maize specific coefficients proposed by Plénet and Lemaire (2000):

$$Nc = 3.4W^{-0.37} \quad (1)$$

where Nc is the critical N content (%) and W is maize above-ground biomass (t/ha) above 1 t/ha (Nc is 3.4% for W<1 t/ha).
The NNI of maize was calculated using Equation 2 as the ratio between the actual crop N uptake (Na) and the critical N uptake, Nc.

$$NNI = Na/Nc \quad (2)$$

and was used to defined conditions of N deficiency (NNI<1), N surplus (NNI>1) and optimal status (NNI=1).

Satellite images acquisitions
Four Sentinel-2 SAFE level-2A cloud free images, the closest in time to field sampling dates, were downloaded and pre-processed using sen2r R toolbox (Ranghetti and Busetto, 2018). The image dates were: 19th of May (10 days later than the 1st ground sampling), 3rd of June (5 days after the 2nd sampling date), 23rd of June (5 days after the 3rd sampling date), 8th of July (5 days after the 4th sampling date). These data were processed to obtain the vegetation indices (VI) listed in Table 1 and to retrieve LAI and CC (canopy chlorophyll) biophysical variables using the ESA SNAP biophysical processor (Weiss and Baret, 2016). For each plot, 10 pixels, centred on the ESU, were selected to extract VI values and biophysical variables for each N treatment plot. Pearson correlation analyses across VIs, Sentinel-2 biophysical variables and field measured vegetation variables and NNI were carried out to investigate their relationships.

Results and discussion

The soil sampling revealed that the study field is characterised by a clay soil with generally good fertility and high SOM and total N content and good C/N. The comparison among soil classes showed a statistically significant difference (P<0.05) in pH and in total N % for the FOR1 class only, for which pH value was sub-acidic compared to the sub-alkaline pH of the other classes (Table 2). Crop response to N fertilization was strongly influenced by soil properties as shown in Figure 1. Whereas, for most soils, a rate of 90 or 180 kg N/ha allowed to reach the maximum biomass, with values generally comparable and to those at higher N rates, the plots with soil type FOR1 showed a remarkably lower response to N, probably due to other nutritional limitations caused by the low pH. As an example of such a soil pattern, it can be seen from Figure 1 that the JOL1 soil class was characterized by increasing biomass values according to N application on both dates, 29th of May

Table 1. List of vegetation indices used in the present study.

Index	Name	Reference
NDVI	normalized difference vegetation index	Rouse *et al.*, 1974
NDRE	normalized difference red edge index	Barnes *et al.*, 2000
MSAVI	modified soil-adjusted vegetation index	Qi *et al.*, 1994
GNDVI	green normalized difference vegetation index	Gitelson *et al.*, 1996
MCARI	modified chlorophyll absorption in reflectance index	Daughtry *et al.*, 2000
MCARI2	modified chlorophyll absorption in reflectance index 2	Haboudane *et al.*, 2004
MTVI2	modified triangular vegetation index 2	Haboudane *et al.*, 2004
OSAVI	optimized soil adjusted vegetation index	Rondeaux *et al.*, 1996
TCI	triangular chlorophyll index	Hunt *et al.*, 2012

Table 2. Results of the soil sampling carried out in the experimental field.[1]

Soil Class	N total (%)	TOC (%)	C / N	SOM (%)	Clay (%)	Silt (%)	Sand (%)	pH
CDS2	0.33 a	3.44	10.39	5.93	63.72	24.01	12.28	7.56 a
JOL3	0.33 a	3.31	10.07	5.70	63.84	22.78	13.39	7.24 a
JOL1	0.32 a	3.40	10.68	5.87	62.28	25.66	12.06	7.63 a
FOR1	0.40 b	4.00	9.86	6.89	62.24	26.85	10.91	5.48 b
Mean	0.34*	3.54	10.25	6.09	63.02	24.82	12.16	6.98*

[1] Soil types with the same letter, or no letter, are not significantly different. * significant differences among means ($P<0.05$).

(Figure 1A) and 08th of July (Figure 1B), whereas the FOR1 soil class remained stable with a lower biomass accumulation.

The relationship between biomass and actual N content, as compared to the critical N from the dilution curves proposed by Plénet and Lemaire (2000) or Li *et al.* (2012) reflected these differences in the crop response to N (Figure 2). Plots located on soil type FOR1 showed N deficiency even at high N rates (e.g. 350 kg N/ha), according to the NNI concept, being below the critical N content. Changes in NNI values under varied N applications during the season did not show the expectations of a clear gradient according to N amount as demonstrated by Zhao *et al.*, 2018. NNI did not always increase with the increasing N application rate at each growth stage showing a comparable N status at BBCH 63 between 180, 300 and 350 kg N/ha rates. The substantial differences of NNI at BBCH 63 were observed under different soil classes, where classes FOR1 and JOL1 markedly differed in NNI values (FOR1=1.68, JOL1=2.34). FOR1 had a constant limiting factor that influenced expected N uptake regardless of N applications. FOR1 N intra-class variability (st.dev 0.08) was very low and the lowest with respect to the other classes (JOL1 st.dev=0.28, CDS2 st.dev=0.36, JOL3 st.dev=0.36), as a consequence of the fact that no differences were obtained with variable rate fertilization as in the other zones of the field.

Estimation of NNI from Sentinel-2 data

The indirect estimation of NNI from remote sensing relies on the estimation of the actual N content and the biomass. Although there was an under-estimation of LAI from Sentinel-2 as compared to ground measurements (Figure 3A), a good relationship was found between measured and estimated CC (Figure 3B), i.e. the product between LAI and chlorophyll. The CC estimated from Sentinel-2 also had a good relationship with the maize above-ground biomass (r=0.92).

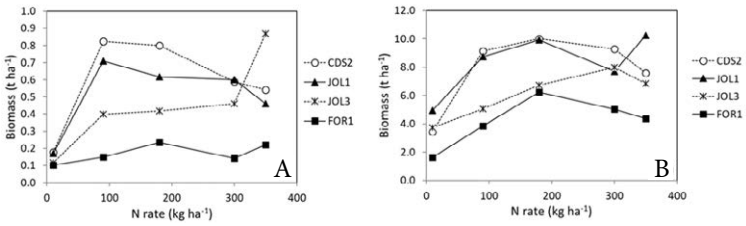


Figure 1. Response of above ground maize biomass to N fertilizer rates at (A) BBCH13 on 29th May and (B) at BBCH 63 on 3rd of July, for different soil types (Table 2).

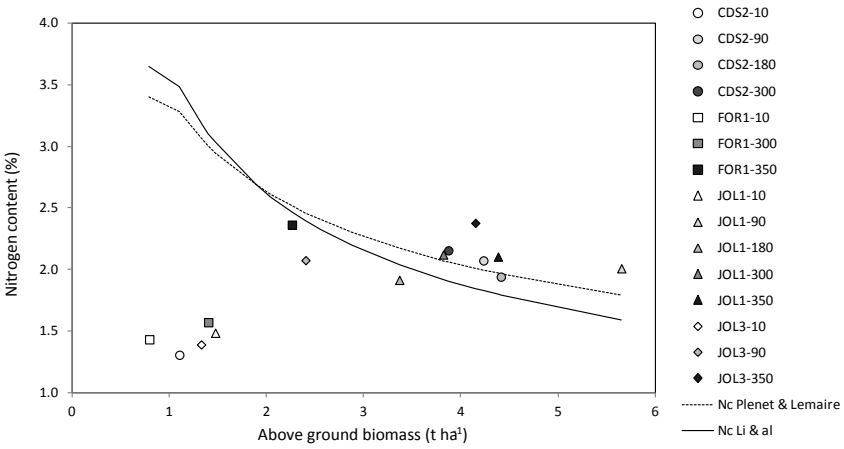


Figure 2. Relationships between maize biomass and actual N content for the different soil types (Table 2) at different N rates at the BBCH 16 maize growth stage. The critical N curves (Nc) proposed by Plénet and Lemaire (2000) and Li *et al.* (2012) are also plotted.

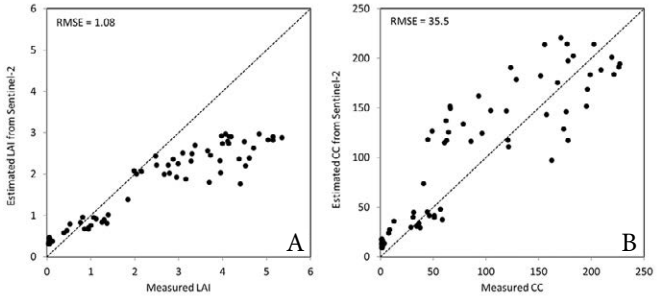


Figure 3. Relationship between (A) measured LAI and estimated LAI from Sentinel-2 biophysical processor; (B) measured and estimated canopy chlorophyll.

However, the estimation of actual crop N content was not very good: the best relationship was found between crop N content and TCI (r=-0.45) or MCARI (r=-0.44). A somewhat better relationship was found between the actual crop N uptake, i.e. the product of N content and biomass, and canopy chlorophyll CC (r=0.89) or leaf chlorophyll Cab (Figure 4A) (r=0.89) estimated from Sentinel-2. The latter is the ratio of the two variables output from the SNAP Sentinel-2 biophysical processor, i.e. CC and LAI. Therefore, it would be possible to use the approach of estimating NNI as the ratio of actual to critical N uptake. Alternatively, NNI could be directly estimated using Cab from Sentinel-2 (r=0.78) (Figure 4B) or TCARI (r=-0.76).

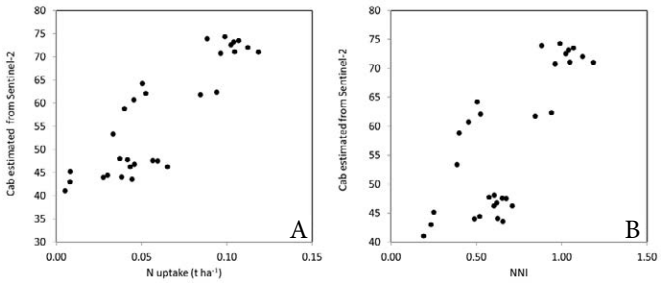


Figure 4. Relationship between (A) actual crop N uptake and Cab estimated from S2; (B) nitrogen nutrition index and Cab estimated from Sentinel-2.

Conclusions

In this study, maize nutritional status was analysed through consolidated techniques such as the use of the NNI directly and indirectly derived, providing a key to understanding the influence of important factors such as soil variability. Soil mapping with geoelectric instruments provided spatialized information layers that can be used with satellite-derived maps to provide guidance on how to interpret the nutritional status of the crop by considering not only nitrogen as a limiting factor. Different soil classes have an effect on growth independently of N rates and an impact on dry matter production. The estimation of NNI from Sentinel-2 has been tested, finding good relationship between canopy chlorophyll (CC) and the maize above-ground biomass, though the estimation of actual crop N content was not very good. More promising relationships were found with the approach of estimating NNI directly from vegetation indices or biophysical variables. In this case, the best relationships with NNI were found with leaf chlorophyll (Cab) and the TCARI index.

Acknowledgements

This research was funded by the company IBF Servizi SpA in the framework of a scientific collaboration agreement with the University of Tuscia. A.C. PhD is funded by CNR-IREA. The authors would like to thank the staff of Bonifiche Ferraresi SpA and IBF Servizi SpA for their technical assistance and for their support during the field experiment. The authors would also like to thank Prof. Pietro Marino, Dr Paolo Villa and Dr Giampiero Ubertini for help with samples analysis. The authors gratefully acknowledge MIUR (Minister for education, University and Research) for financial support (Law 232/216, Department of Excellence) to DAFNE.

References

Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., *et al.,* 2000. Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data. In: Robert, P. C.; Rust, R. H.; Larson, W. E. (Eds.) Proceedings of the 5th International Conference on Precision Agriculture, Madison, WI, USA: American Society of Agronomy. pp. 1-15.

Blondlot, A., Gate, P., & Poilvé, H., 2005. Providing operational nitrogen recommendations to farmers using satellite imagery. In J V Stafford (Ed.) Precision Agriculture '05. Proceedings of the 5th European Conference on Precision Agriculture, the Netherlands: Wageningen Academic Publishers pp 345-352.

Cammarano, D., Fitzgerald, G. J., Casa, R., & Basso, B., 2014. Assessing the robustness of vegetation indices to estimate wheat N in mediterranean environments. Remote Sensing, 6 2827-2844.

Daughtry, C. S. T., Walthall, C. L., Kim, M. S., De Colstoun, E. B., & McMurtrey, J. E., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sensing of Environment 74: 229-239.

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N., 1996. Use of a green channel in remote sensing of global vegetation from EOS- MODIS. Remote Sensing of Environment 58: 289-298.

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B., 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. Remote Sensing of Environment 90: 337-352.

Hunt, E. R., Doraiswamy, P. C., McMurtrey, J. E., Daughtry, C. S. T., Perry, E. M., & Akhmedov, B., 2012. A visible band index for remote sensing leaf chlorophyll content at the Canopy scale. International Journal of Applied Earth Observation and Geoinformation 21 103-112.

Lancashire, P. D., Bleiholder, H., Boom, T. Van Den, Langelüddeke, P., Stauss, R., *et al.,* 1991. A uniform decimal code for growth stages of crops and weeds. Annals of Applied Biology 119 561-601.

Li, W., He, P., & Jin, J., 2012. Critical Nitrogen Curve and Nitrogen Nutrition Index for Spring Maize in North-East China. Journal of Plant Nutrition 35 1747-1761.

Nutini, F., Confalonieri, R., Crema, A., Movedi, E., Paleari, L., Stavrakoudis, D., *et al.,* 2018. An operational workflow to assess rice nutritional status based on satellite imagery and smartphone apps. Computers and Electronics in Agriculture 154 80-92.

Plénet, D., & Lemaire, G., 2000. Relationships between dynamics of nitrogen uptake and dry matter accumulation in maize crops. Determination of critical N concentration. Plant and Soil 216 65-82.

Priori, S., Fantappiè, M., Magini, S., & Costantini, E. A. C., 2013. Using the ARP-03 for high-resolution mapping of calcic horizons. International Agrophysics 27 313-321.

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S., 1994. A modified soil adjusted vegetation index. Remote Sensing of Environment 48 119-126.

Rondeaux, G., Steven, M., & Baret, F., 1996. Optimization of soil-adjusted vegetation indices. Remote Sensing of Environment 55 95-107.

Rouse, J. W., R. H. Haas, J. A. Schell, and D. W. D., 1973. Monitoring vegetation systems in the Great Plains with ERTS, Third ERTS Symposium. NASA SP-351 I, 309-317.

Weiss, M., & Baret, F., 2016. S2ToolBox Level 2 products: LAI, FAPAR, FCOVER, 53.

Zhao, B., Duan, A., Ata-Ul-Karim, S. T., Liu, Z., Chen, Z., Gong, Z., *et al.,* 2018. Exploring new spectral bands and vegetation indices for estimating nitrogen nutrition index of summer maize. European Journal of Agronomy 93 113-125.

Adopting precision agriculture to improve the cultivation of old wheat varieties in Tuscany (Italy)

C. Fabbri, M. Napoli*, M. Mancini, G. Brandani, R. Vivoli and S. Orlandini
Department of Agrifood Production and Environmental Sciences (DISPAA), University of Florence, Piazzale delle Cascine, 18, 50144 Florence, Italy; marco.napoli@unifi.it

Abstract

A field experiment (September 2016 – July 2018) was carried out in Monteroni D'Arbia, Tuscany, Italy to compare three old genotypes (Andriolo, Sieve and Verna) of bread wheat to the modern cultivar Bologna. Sentinel-2A images were collected and analysed using QGIS program to calculate five vegetation indices (NDVI, NDRE, GNDVI, CCCI, SAVI). The results of the field campaign showed that the vegetation indices calculated from Sentinel-2A bands showed satisfactory correlations with respect to protein content and grain dry biomass. Sentinel-2A vegetation indices were shown to be suitable for supporting precision agriculture in old bread wheat varieties fields.

Keywords: Sentinel-2A, old wheat varieties, yield, remote sensing, protein content

Introduction

Wheat is one of the most consumed foods worldwide for humans. This trend has meant that, over the years more productive and resistant (modern) wheat species have been put on the market. This has had a positive effect on productivity, but negative aspects on the other side: increase in the use of fertilizers, chemical effects on human health. Currently, it is important to cultivate varieties that are well suited for food health and to take advantage of resilient farming techniques (Arzani and Ashraf, 2017). Among these, the most studied are old wheat varieties. In the past few years, old wheat varieties have been valued to be rich sources of bioactive components, classified as 'healthy food' (Lachman *et al.*, 2013). Added to these benefits are the lower nutritional requirements of these varieties to obtain maximum yield and the higher nitrogen uptake efficiency compared to modern varieties, which allows greater environmental protection (Pourazari *et al.*, 2015). The re-cultivation of old wheat varieties could be recommended in modern agriculture, where precision farming techniques make it possible to manage crops in an increasingly sustainable way. One of the major problems encountered is undoubtedly the management of nitrogen for wheat nutrition, because there is a poor relationship between applied nitrogen fertilizer and crop nitrogen demand. It leads to low nitrogen use efficiency in current crop production systems (Diacono *et al.*, 2013). However, considering that the nitrogen aerial concentration and the nitrogen absorption is different between the vegetative and reproductive stages, the characteristics of biomass are important in determining the relations between spectral indices and crop nitrogen status (Li *et al.*, 2014) and, accordingly, determining the final yield. The assessment of canopy nitrogen status from vegetation indices evaluation is also difficult, due to the different crop structure and canopy cover level (Cammarano *et al.*, 2014). The old wheat genotypes show a crop growth behavior and vegetative development different from modern ones. Therefore, it is important to adopt remote sensing techniques to identify the differences that characterise these varieties. This challenge is essential both for yield estimation and for optimal nitrogen management. In the past few years, attention has been paid to systems that use satellite images to detect the vegetative state of crops. The Sentinel-2A satellite is currently available freely; it is part of the Sentinel 2 mission project of the European program Copernicus and the European Space Agency (ESA). Sentinel-2A carries a multispectral instrument (MSI) and it has 13 spectral bands with a spatial resolution from 10 to 60 m.

DOI 10.3920/978-90-8686-888-9_57, © Wageningen Academic Publishers 2019

This work presents a result of remote sensing analysis using satellite data Sentinel-2A to study the productive behavior of modern and old wheat varieties in comparison. The aim is to investigate the capability of vegetation indices to detect final yield and production quality (protein content). Thus, it will be possible to understand the development of different wheat varieties which is important for optimising nitrogen management practices.

Materials and methods

Study site

This study was carried out in a farm located in Monteroni D'Arbia, 13 km from Siena, Tuscany, Italy, (43°13'49'08 N, 11°25'18'84 E; 161 m asl) (Figure 1) during two growing season (2015/2016 and 2017/2018).

The experimental field was cultivated with four bread wheat varieties: three old (Verna, Sieve and Andriolo) and one modern (Bologna). Data were collected during the main phenological wheat stages: tillering (March), stem elongation (mid-April); flowering (mid-May); grain filling (June); harvesting (early-July). The fields were sown in late November and harvested in early July. The experimental design consisted of 24 plots (one each treatment): four wheat varieties, three different nitrogen treatment (N1:35, N2:85 and N3:135 kg N/ha) and two different densities (D1:90 and D2:180 kg of seed/ha). During the growing season, phenological stage evaluations, using the BBCH scale (Hack *et al.*, 1992), were performed monthly in triplicate for each treatment. Then, the destructive samples were collected for determination of aboveground biomass and nitrogen content at flowering and at harvesting. Three sub-samples (0.5×0.5 m) of aboveground biomass were collected manually from each plot. Samples were oven-dried (48 hours at 105 °C) and then weighed and analysed for the determination of the nitrogen and protein as a percentage of the dry matter.

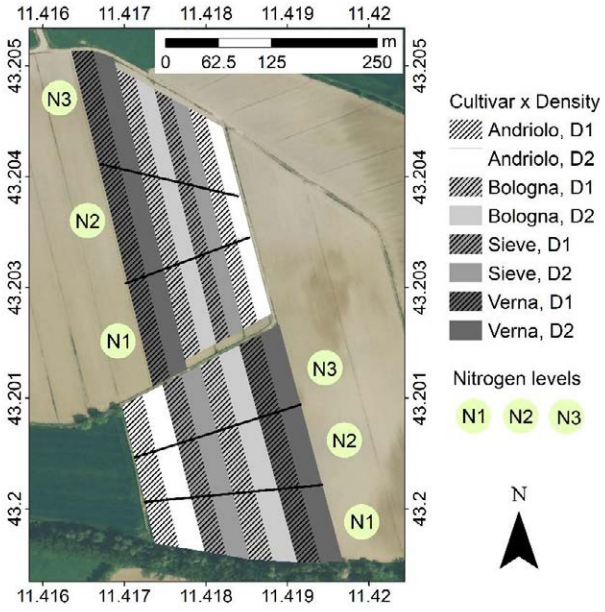


Figure 1. Map of the study field with the scheme of the fertilisation experiment. Four wheat varieties (namely Andriolo, Bologna, Sieve, and Verna), three different nitrogen treatment (N1:35, N2:85 and N3:135 kg N/ha) and two different densities (D1:90 and D2:180 kg of seed/ha).

Multispectral images

Datasets produced by Sentinel-2A satellite were downloaded from Europe's Copernicus repository (https://scihub.copernicus.eu/dhus). All available Sentinel-2 images from January to harvesting were collected for both growing seasons (i.e. 2017 and 2018). The 20 m spatial resolution was used for each satellite image obtained; Blue (B2, 490 nm), Green (B3, 560 nm), Red (B4, 665 nm) and NIR (B8, 842 nm). For this research, four Sentinel-2A images of the available collection were selected to be correlated with yield data. Images about 1, 2, 3, and 4 months prior to harvest were chosen to correlate with yield data.

The remote sensing images were uploaded on opensource QGIS (v2.18; QGIS Development Team, 2016) and combined with the georeferenced-sampled point locations, collected using a GPS receiver. A Python script was compiled to process the downloaded bands and to compute the vegetative index reported in Table 1. The two red based vegetation indices (NDVI and SAVI) and the three green (NDVIRE, GNDVI and CCCI) represent commonly used vegetation indices for estimation of nitrogen content and biomass determination on crops.

An example of maps to show the NDVI evolution through the 2018 growing season of the experimental field is shown in Figure 2.

Table 1. Definitions of the vegetative indices applied.

Index	Acronym	Equation	Reference
Normalized difference vegetation index	NDVI	(NIR – Red) / NIR + Red)	Rouse *et al.*, 1973
Red-edge NDVI	NDVIRE	(NIR – Red_{edge}) / (NIR + Red_{edge})	Gitelson and Merzlyak, 1994
Green NDVI	GNDVI	(NIR – Green) / (NIR + Green)	Gitelson *et al.*, 1996
Canopy chlorophyll content index	CCCI	((NIR – Red_{edge}) / (NIR + Red_{edge})) × ((NIR – Red) / (NIR + Red))$^{-1}$	Barnes *et al.*, 2000
Soil adjusted vegetation index	SAVI	(NIR – Red) / (NIR + Red + 0.428) × (1 + 0.428)	Huete, 1988

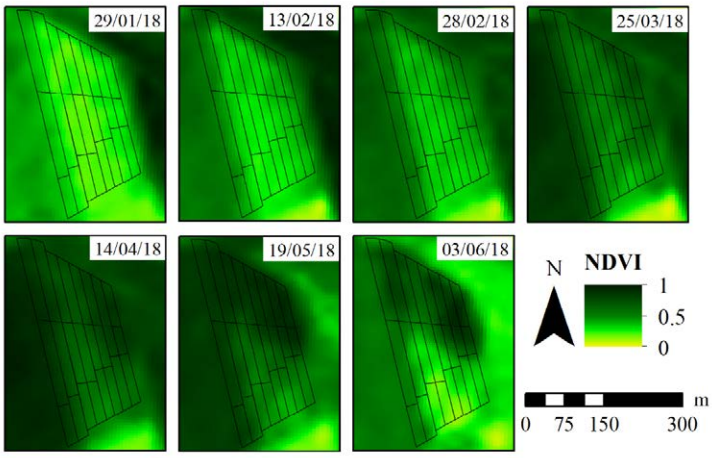


Figure 2. NDVI evolution along the 2018 growing season using all available and useful Sentinel-2 products.

Results and discussion

Vegetation indices and growth stage comparison

The old varieties were the latest in reaching flowering stage (BBCH 61), while Bologna was the earliest. In comparison to Bologna, the Verna and Sieve cultivars had longer growing season of about 4 days, while Andriolo of about 16 days. The effect of sowing density, nitrogen levels and cultivars on the radiometric response gave the following results: the modern cultivar showed NDVI and SAVI values significantly lower than those of the ancient varieties until booting. However, in the flowering stage, differences were no longer statistically significant. No significant differences in NDVI and SAVI values were found among the ancient varieties between the booting and the flowering stage. Significant differences in NDVI and SAVI values were observed, at flowering, for N3 with respect to N2 and N1. Nitrogen doses above or equal to 80 kg N/ha did not show significant differences, regardless of the seeding density and varieties. No significant differences were observed between the two densities for the old varieties.

Vegetation indices and yield comparisons

Table 2 shows the correlations between grain yield from the field sampling and the vegetation indices calculated from the Sentinel-2A images. According to statistical results, significant correlation between vegetation indices and grain biomass at harvesting were found from April to June (from stem elongation to grain filling) for both years for Bologna, Sieve and Verna. In these cases, positive correlation coefficients were found, thus indicating a positive relationship between plant vigour and the harvested yield. Bologna and Sieve showed a high correlation from April to June (from stem elongation to grain filling) for both years. Verna showed high correlation from April to May (from stem elongation to flowering), except for April 2018, where no indices explain the final yield. In contrast, Andriolo, which seems to have a completely different behavior compared to other varieties, showed a negative low correlation throughout the considered period for the first growing seasons and a significant positive correlation throughout the second growing season. The best correlations were obtained with the GNDVI index. The NDVI and SAVI showed an equal performance. Finally, the CCCI showed a poor uniformity of results and the lowest correlation values for every variety (the best results were related to Sieve). This case study highlighted the possibility to understand and quantify the within-field spatial variability of wheat vigour and yield along the whole season. This information could be considered to improve the fertilization and crop protection actions.

Vegetation indices and grain protein percentage

In contrast to what was observed for grain yield, the correlation coefficients were negative which means that as the indices increased, the grains protein as a percentage of the dry matter content decreased (Table 3). These results were consistent with the well-known negative relationship between grain yield and grains protein as a percentage of the dry matter content (Simmonds, 1995). In this analysis, the results seemed to be very dependent on the weather, as the results were different from 2017 to 2018. The old bread wheat varieties showed higher correlations in 2018 than in 2017, while the modern Bologna, showed an opposite trend. Probably, the trend is explained by the fact that 2017 was a particularly dry year and the modern variety has suffered less from heat stress compared to the old ones. In general, May and June seem to be always the most suitable months to predict protein content of grain from vegetation indices for the considered wheat bread variety. Andriolo showed very different correlation trends compared to other old varieties, while the similarity between Bologna and Sieve correlations trend was repeated. Sieve showed the highest correlation values compared to the other varieties and more constant over time.

Table 2. Correlation coefficients between yield and vegetation indices for each bread wheat variety (n=18). The analysis was performed for images acquired in March, April, May, and June of 2017 and 2018. * and ** indicates significant ($P<0.05$ and $P<0.01$) correlations.

Variety	Index	2017				2018			
		March	April	May	June	March	April	May	June
Andriolo	NDVI	0.081	-0.237	-0.272	0.107	0.548*	0.518*	0.506*	0.492*
	SAVI	0.081	-0.237	-0.271	0.107	0.548*	0.518*	0.506*	0.492*
	NDVIRE	0.123	-0.307	-0.014	0.109	0.489*	0.515*	0.488*	0.45
	GNDVI	0.46	-0.068	-0.028	0.103	0.531*	0.522*	0.502*	0.498*
	CCCI	0.013	0.003	0.295	0.255	-0.014	0.461	0.416	0.273
Bologna	NDVI	0.441	0.813**	0.59*	0.617**	0.545*	0.668**	0.703**	0.692**
	SAVI	0.441	0.813**	0.59*	0.617**	0.545*	0.668**	0.703**	0.692**
	NDVIRE	0.142	0.643**	0.603**	0.763**	0.596**	0.633**	0.721**	0.684**
	GNDVI	0.449	0.831**	0.621**	0.721**	0.487*	0.669**	0.712**	0.695**
	CCCI	-0.549*	0.1	0.599**	0.442	0.398	0.362	0.709**	0.336
Sieve	NDVI	0.035	0.527*	0.799**	0.628**	0.663**	0.814**	0.791**	0.787**
	SAVI	0.035	0.527*	0.799**	0.628**	0.663**	0.814**	0.791**	0.787**
	NDVIRE	0.12	0.587*	0.802**	0.793**	0.72**	0.82**	0.77**	0.805**
	GNDVI	0.309	0.709**	0.81**	0.82**	0.63**	0.808**	0.784**	0.799**
	CCCI	0.001	0.131	0.724**	0.616**	0.516*	0.62**	0.634**	0.642**
Verna	NDVI	-0.083	0.673**	0.913**	0.398	0.01	0.325	0.597**	0.622**
	SAVI	-0.083	0.673**	0.913**	0.398	0.01	0.325	0.597**	0.622**
	NDVIRE	0.05	0.729**	0.923**	0.916**	0.1	0.34	0.655**	0.657**
	GNDVI	0.202	0.81**	0.925**	0.91**	-0.01	0.396	0.643**	0.615**
	CCCI	0.324	0.354	0.888**	0.113	0.453	0.338	0.668**	0.565*

Conclusions

The aim of this research was to find derived Sentinel 2-A vegetation indices able to assess wheat grain production and quality. The study was carried out on different wheat varieties (ancient ones and a modern one) cultivated in Val D'Orcia area; the aim was also to detect if differences were visible based on the cultivar behaviour. Five vegetation indices most used were utilized for the purpose and statistically analysed. A strong correlation between vegetation indices and yield data was found for all vegetation indices. High correlation values were also found between selected vegetation indices and protein content of grains. The relationship varied in the selected crop varieties. Andriolo showed a trend that is very different from the other varieties. Bologna and Sieve had a strong similarity in correlation results for both yield and protein content results. The results for protein content of grains in a modern variety were influenced by weather.

The results showed the high capability of Sentinel–2A to detect both quality and quantity grain characteristics independently of wheat variety, even if some differences are evident. Further analysis is necessary to validate the observations.

Table 3. Correlation coefficients between grain protein content and vegetation indices for each bread wheat varieties (n=18). The analysis was performed for images acquired in March, April, May and June of 2017 and 2018. * and ** indicates significant (P<0.05 and P<0.01) correlations.

Variety	Index	2017				2018			
		March	April	May	June	March	April	May	June
Andriolo	NDVI	0.093	-0.333	-0.388	-0.348	-0.535*	-0.682**	-0.588*	-0.606**
	SAVI	0.093	-0.333	-0.388	-0.348	-0.535*	-0.682**	-0.588*	-0.606**
	NDRE	0.178	-0.3	-0.279	-0.346	-0.52*	-0.669**	-0.528*	-0.555*
	GNDVI	0.282	-0.278	-0.335	-0.402	-0.543*	-0.671**	-0.526*	-0.606**
	CCCI	0.054	0.142	-0.055	-0.317	-0.101	-0.596**	-0.377	-0.345
Bologna	NDVI	0.183	0.324	-0.556*	-0.754**	-0.239	-0.299	-0.472*	-0.485*
	SAVI	0.183	0.324	-0.556*	-0.754**	-0.239	-0.299	-0.472*	-0.5170*
	NDRE	0.174	0.097	-0.541*	-0.61**	-0.383	-0.323	-0.515*	-0.482*
	GNDVI	0.125	-0.017	-0.553*	-0.534*	-0.157	-0.292	-0.487*	-0.488*
	CCCI	-0.187	-0.174	-0.456	-0.803**	-0.493*	-0.322	-0.625**	-0.232
Sieve	NDVI	0.515*	0.012	-0.476*	-0.621**	-0.41	-0.519*	-0.602**	-0.579*
	SAVI	0.515*	0.012	-0.476*	-0.621**	-0.41	-0.519*	-0.602**	-0.579*
	NDRE	0.41	-0.376	-0.502*	-0.489*	-0.54*	-0.603**	-0.646**	-0.601**
	GNDVI	0.362	-0.155	-0.468*	-0.414	-0.387	-0.574*	-0.612**	-0.575*
	CCCI	-0.408	-0.511*	-0.494*	-0.555*	-0.627**	-0.691**	-0.745**	-0.509*
Verna	NDVI	0.18	-0.264	-0.49*	0.265	-0.441	-0.561*	-0.574*	-0.539*
	SAVI	0.18	-0.264	-0.49*	0.265	-0.441	-0.561*	-0.574*	-0.539*
	NDRE	0.07	-0.269	-0.485*	-0.423	-0.5*	-0.606**	-0.565*	-0.504*
	GNDVI	-0.031	-0.385	-0.469*	-0.49*	-0.405	-0.534*	-0.553*	-0.561*
	CCCI	-0.321	-0.05	-0.534*	0.27	-0.238	-0.658**	-0.528*	-0.101

Acknowledgements

This activity was part of the project 'GRANT GRani Antichi Nuove Tecniche di coltivazione' founded by PSR2014/2020 Regione Toscana.

References

Arzani, A. and Ashraf, M., 2017. Cultivated Ancient Wheats (Triticum spp.): A Potential Source of Health-Beneficial Food Products. Comprehensive Reviews in Food Science and Food Safety 16 477-488.

Barnes, E.M., Clarke, T.R. and Richards S.E., 2000. Coincident detection of crop water stress, nitrogen status and canopy density using ground based Multispectral data. In P. C. Robert, R. H. Rust and W. E. Larson (Eds.) Proceedings of the 5th International Conference on Precision Agriculture, Madison, WI, USA: ASA-CSSA-SSSA.

Cammarano, D., Fitzgerald, G., Casa, R., and Basso, B., 2014. Assessing the Robustness of Vegetation Indices to Estimate Wheat N in Mediterranean Environments. Remote Sensing 6(4) 2827-2844.

Diacono, M., Rubino, P., Montemurro, F., 2013 Agronomy for Sustainable Development 33: 1 219-242.

Gitelson, A.A., and Merzlyak, M.N., 1994. Spectral reflectance changes associated with autumn senescence of *Aesculus Hippocastanum* L. and *Acer Platanoides* L. leaves: spectral features and relation to chlorophyll estimation. Journal of Plant Physiology 143 286-292.

Gitelson, A.A., Kaufman, Y.J., and Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote sensing of environment 58 289-298.

Hack, H., Bleiholder, H., Buhr, L., Meier, U., Schnock Fricke, U., *et al.*, 1992. Einheitliche Codierung der phänologischen Entwicklungsstadien mono- und dikotyler Pflanzen – Erweiterte BBCH-Skala Allgemein Nachrichtenblatt. (Uniform coding of phenological developmental stages of monocotyledonous and dicotyledonous plants – Extended BBCH scale). Deut Pflanzenschutzd 44 265-270.

Huete, A.R., 1988. A soil adjusted vegetation index (SAVI). Remote Sensing of Environment 25 295-309.

Lachman, J., Hejtmánková, K., and Kotíková, Z., 2013. Tocols and carotenoids of einkorn, emmer and spring wheat varieties: Selection for breeding and production, Journal of Cereal Science, 57(2) 207-214 doi:10.1016/j.jcs.2012.05.011.

Li, F., Mistele, B., Hu, Y., Chen, X., and Schmidhalter, U., 2014. Reflectance estimation of canopy nitrogen content in winter wheat using optimized hyperspectral spectral indices and partial least squares regression. European Journal of Agronomy 52 198-209.

Pourazari, F., Vico G., Ehsanzadeh P., and Weih M., 2015. Contrasting growth pattern and nitrogen economy in ancient and modern wheat varieties. Canadian Journal of Plant Science 95 851-860.

QGIS Development Team, 2016. QGIS Geographic Information System. Open Source Geospatial Foundation. http://qgis.org

Rouse, J.W., Haas, R.H., Schell, J.A., and Deering D.W., 1973. Monitoring vegetation systems in the great plains with ERTS. Third ERTS Symposium, NASA SP351 I: 309-317.

Simmonds, N.W., 1995. The relation between yield and protein in cereal grain. Journal of the Science of Food and Agriculture 67(3), 309-315.

A cotton yield estimation model based on agrometeorological and high resolution remote sensing data

A. Falagas and K. Karantzalos*
Remote Sensing Laboratory, National Technical University of Athens, Athens, Greece; karank@central.ntua.gr

Abstract

In this paper, a cotton yield forecasting model based on agrometeorological and high resolution Sentinel-2 (10 m pixel size) satellite data is proposed. In particular, the WOFOST crop growth and yield simulation model was employed only during the training process. Daily LAI and yield simulations along with multi-temporal satellite observations were integrated to calibrate and constrain the predictions. The developed model was validated against the reference actual yield. The experimental results were quite promising. The model managed to rapidly estimate cotton yield with an overall accuracy of approximately 5 kg per 100 m^2. The calculated correlation between the actual and predicted yields was above 82% indicating the great potential of the proposed approach.

Keywords: earth observation, satellite, multispectral, LAI, crop model, WOFOST

Introduction

Remote sensing and space technology provides operationally frequent spatial information for understanding the heterogeneity of the environment and, therefore, is of significant importance in modeling the spatial dynamics of any land system (Karantzalos *et al.*, 2015; Karmas *et al.*, 2016). In particular, remote sensing has a key role and great potential in agricultural crop monitoring by delivering high temporal and spatial earth observation (EO) data for describing the spatio-temporal variability of crop biophysical variables (Leon *et al.*, 2003; Karantzalos *et al.*, 2017; Jin *et al.*, 2017), including leaf area index (LAI), biomass, canopy cover, the fraction of absorbed photosynthetically active radiation (fAPAR), leaf nitrogen accumulation, evapotranspiration and soil properties.
Without remote sensing observations, these canopy state variables were usually simulated through process-based models (e.g. CERES-Maize, Sirius, SUCROS2, STICS, SUNFLO) by combining information for the main biophysical processes and simple empirical parameterization (e.g. AqYield, PolyCrop, WOFOST). However, due to modeling errors (e.g. generalization of parameters and of spatial, soil, weather and crop data, merging varieties and regional management parameters, effect of plant diseases and insect pests), EO data have a critical role in assimilating the information of canopy state variables into crop models towards improving crop management (Jin *et al.*, 2018; Guo *et al.,* 2018).
In particular, during the last decade, a significant research effort has been concentrated in coupling remote sensing and crop growth models towards efficient and accurate crop yield estimation (Ma *et al.*, 2013). Most methods are coupling a radiation transfer model with a specific crop model and/or employing medium resolution (e.g. MODIS, MERIS) satellite data in order to reinitialize crop model parameters by minimizing the differences between predictions and remote sensing observations. To this end, in this paper, similar to recent research efforts (Lobell *et al.*, 2015) which exploit high resolution Landsat satellite data (30 m spatial resolution), the goal is to integrate frequent, high resolution Sentinel-2 data (10 m spatial resolution) as well as to employ the estimated daily canopy state variables from the WOFOST (De Wit *et al.*, 2019) model in order to train and develop a novel forecasting model for estimating efficiently LAI and cotton yield.

DOI 10.3920/978-90-8686-888-9_58, © Wageningen Academic Publishers 2019

In the following sections, the employed data and processing methods are described along with the proposed approach for training and developing the new model. Then, experimental results and validation are presented along with the overall conclusions.

Materials and methods

Study area

The study was conducted in Palamas (39.4690°N, 22.0823°E), Karditsa, Greece. The region of interest, covers 200 km^2 in total (Figure 1) and the topography is characterized by a smooth terrain and mainly plains. The climate of the region is continental with cold winters and hot summers. The average annual rainfall is 410 mm and the average annual temperature is 15.6 °C. Cotton which is the dominant crop in the region, is generally planted in the beginning of May and harvested in mid-October.

In-situ *data*

Several *in-situ* campaigns were conducted in order to gather the required agronomic cotton parameters during the growing season of 2017 and 2018. In particular, measurements of plant height, leaf area and leaf count number were conducted in order to calculate the LAI of cotton plants in specific sites. Moreover, other agronomic data were collected, including cotton variety, irrigation and nutrition management and the actual yield per parcel from the local farmers.

Earth Observation data

Sentinel 2A-2B level L1C satellite data were collected in the study area. Atmospheric corrections were conducted using Sen2Cor 2.5.5 algorithm to convert L1C-TOA data to L2A-BOA. From the L2A data, 12 vegetation indices were calculated including NDVI, NDRE1, NDRE2, SAVI, MCARI,

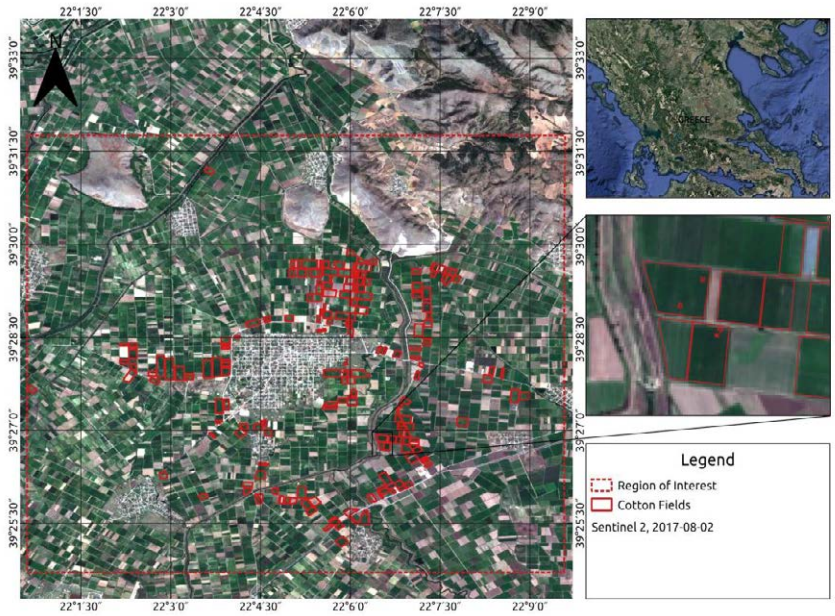


Figure 1. The study area in Palamas Greece, where field campaigns were conducted during the years of 2017 and 2018.

TCARI, Chlorophyll Red-Edge, Chlorophyll Green Index, MTCI, EVI, EVI2, PSRI-NIR in order to train a statistical model for LAI estimation. A number of vegetation indices have been considered and computed in order to exploit the high resolution multispectral Sentinel 2 satellite data.

Satellite-based LAI estimation

LAI is defined as the leaf area per unit area of horizontal land below (Watson, 1947), while it is a proxy of the gas exchange potential. The field LAI measurements (ground truth LAI) was employed and combined with the corresponding satellite Sentinel 2 data and calculated vegetation indices. In particular, after several experiments, the EVI and PSRI-NIR indices showed the highest correlations with the ground truth LAI data (r^2=0.89 for both indices). These multi-temporal satellite-based LAI observations were used to feed and calibrate the WOFOST model.

WOFOST calibration and daily LAI and yield predictions

Since statistical (or any other machine learning, deep learning) models require optimally a huge amount of training data, the goal here is to calibrate the WOFOST model with the appropriate parameters and actual agrometeorological data for the considered study area and produce dense (daily) temporal predictions (LAI and cotton yield) for the following step.

In particular, WOFOST can simulate crop growth taking into account photosynthesis, respiration and the manner that these processes are influenced by environmental conditions. Since, WOFOST has several parameters and simulation values, in order to evaluate and sort them by their sensitivity the standard local sensitivity analysis technique was employed to evaluate the sensitivity of its parameters. To do so, the effective range of the parameters was determined and these cotton parameters were set to follow uniform distribution in their individual range of value. Based on a random sampling of 3,000 values and Monte Carlo propagation the sensitivity rate was estimated. The most sensitive parameters were the specific leaf area (SLATB), conversion efficiency of assimilates into storage organ (CVO), the conversion efficiency of assimilates into leaf (CVL) and the maximum leaf CO_2 assimilation rate (AMAXTB) parameters (Figure 2).

In order to calibrate the WOFOST model against the actual in-situ measurements, the considered assimilation procedure minimized the cost function of the satellite-based LAI and simulated LAI along with the actual and simulated yields.

The employed shuffled complex evolution (SCE) optimization algorithm (Duan *et al.*, 1993) performed an iterative WOFOST re-initialization until a criterion threshold was satisfied in order to acquire the optimal sets of parameters. SCE follows a global optimization strategy based on a constrained random search and has been demonstrated that it can be successfully exploited for calibrating crop models. In this case, the considered cost function was the following one:

$$J_{LAI,YIELD} = \sum_{i=t_i}^{t_n}\left(LAI_{,t_i} - LAI_{EO,t_i}\right)^2 + \left|\left[\left(Yield_{,t_n} - Yield_{actual}\right)/Yield_{actual}\right] * 100\right| \quad (1)$$

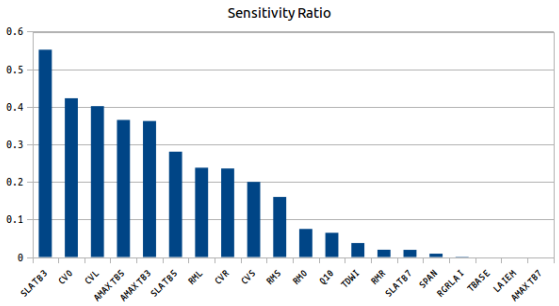


Figure 2. The sensitivity analysis results for WOFOST cotton parameters.

where J is the value of the cost function, LAI_{EO} is the value of satellite-based LAI, LAI_{SIM} is the value of simulated LAI, $Yield_{SIM}$ is the simulated yield in maturity date and $Yield_{actual}$ is the actual yield per parcel.

New cotton yield prediction model

Once calibrated, the WOFOST model was used to produce daily simulations for each monitored parcel in the study area for a two year period i.e. 2017 and 2018. Each individual simulation refers to the estimated from the calibrated WOFOST LAI and yield per cotton parcel. Then based on the extracted simulated LAI outputs (per parcel), numerous (pseudo) satellite EVI indices were computed from the corresponding Sentinel 2 images.

Given the daily simulated yields, the (pseudo) satellite EVI indices and meteorological weather data (i.e. mean temperature, mean solar radiation and mean precipitation per month), numerous training data were derived in order to train the new cotton yield statistical model. In particular, a generic multiple linear regression formulation was employed and was defined as follows:

$$Yield_d = \beta_{0,d} + \beta_{1,d} * RI_d + \beta_{2,d} * W_d * RI_d \quad (2)$$

where W is a vector of meteorological data, RI_d corresponds to remote sensing index and d corresponds to the considered date.

Experimental results and evaluation

The cotton yield prediction model developed was applied in the study area for two years based on Sentinel-2 multispectral data and ingested meteorological ones. Four parcels for the year of 2017 and eleven parcels for 2018 were used for the validation of the cotton yield maps produced, covering approximately 60 ha in total.

The quantitative assessment indicated relatively high correlations between the predicted and actual cotton yields. In particular, the comparisons between the actual yields and the predicted ones resulted in an r^2 of 0.97 and 0.82 for the years 2017 and 2018, respectively.

Since the model developed is able to provide yield predictions per pixel (corresponding to a 10 m by 10 m square on the ground) the per-pixel difference was assessed between the actual and predicted

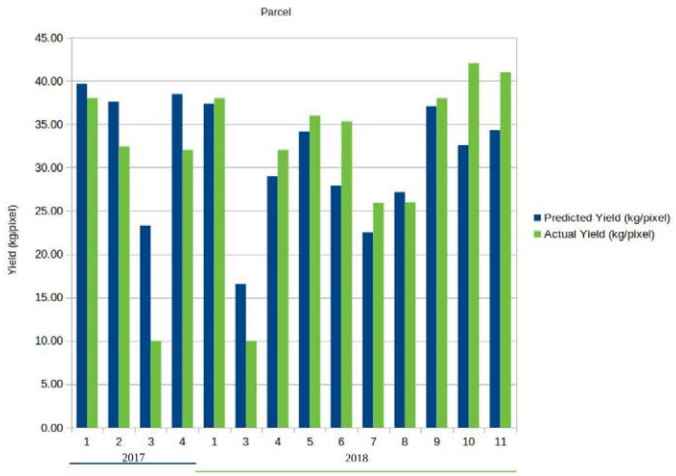


Figure 3. The per-parcel predicted vs actual yields for the years 2017 and 2018.

yield. Indeed, the average per-pixel yield difference between the predictions and the actual yields was calculated at 5 kg/pixel for both years (Figure 3).
Generally speaking, the relatively lower correlation rate for the 2018 parcels was due to the fact that the additional parcels covered a larger variability regarding the agricultural management, the soil parameters and maturity levels. Moreover, for both years, it should be noted that the meteorological data employed included the monthly average temperatures, precipitations and solar radiation, while daily information would have addressed more effectively any weather extremes or absolute correspondences between EO observations during the training process of model development.
For 2017, the cotton yield prediction model developed, although performing significantly accurately, it systematically slightly overestimated the corresponding yields (approximately 4-5 kg per pixel). This was mainly due to the fact that during the creation of the pseudo satellite-based observations for the optimization SCE procedure, the simulated LAI was lower than the LAI calculated from the Sentinel-2 data.

Qualitative assessment

In order to qualitatively examine the performance of the model, in Figure 4 and 5, the yield maps produced per day-of-year (DOY) are presented for 2017 and 2018, respectively. After a closer look, one can observe that, in all cases, the model can capture and exploit the differences of both maturity level as well as canopy state variables between August and September.
In Figure 4 and 5, each pixel at a given DOY predicts the yield based on the satellite-based spectral index following the formulation of the $Yield_d$ calculation. This per pixel prediction actually corresponds to the cotton yield for an area of 100 m^2.

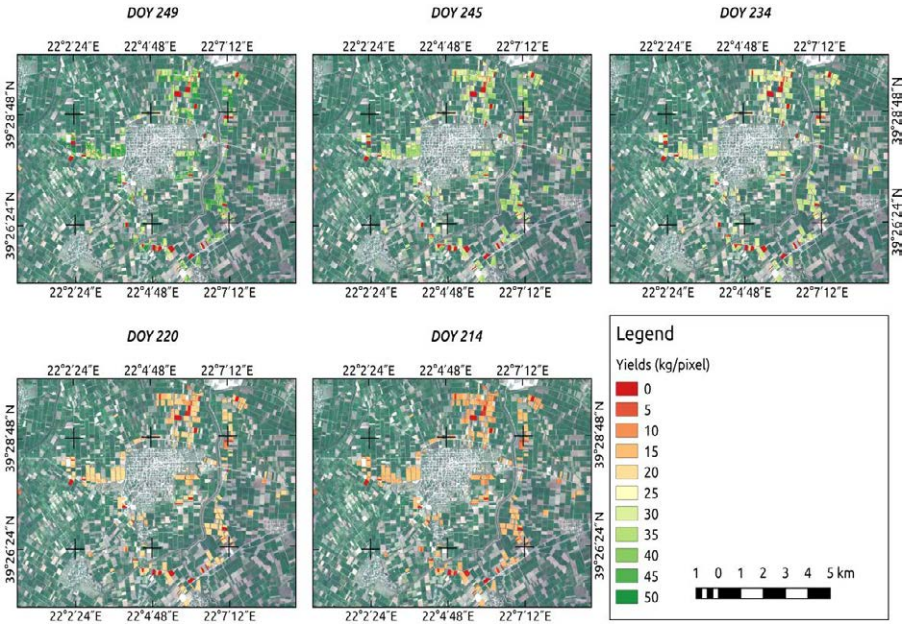


Figure 4. Yield (kg/pixel) maps for different DOYs 2017.

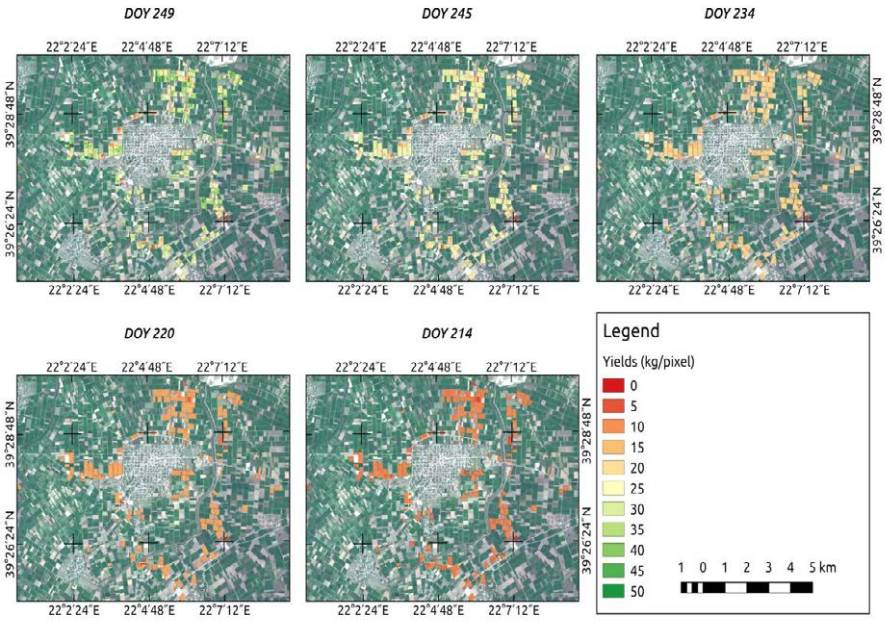


Figure 5. Yield (kg/pixel) maps for different DOYs of 2018.

Therefore, at the DOY #214 of 2018 the average amount of cotton in each 100 m^2 region (pixel) started from around 8.3 kg, while one month later i.e. at the DOY #249 of 2018 it reached an average of 27 kg, which was the final estimation just before the harvesting period.
In a similar manner, at the DOY #201 of the year 2017, the average amount of cotton in a 100 m^2 region was 12.7 kg and reached an average of 30.8 kg at the DOY #249 (2017) just before the harvesting period which was also significantly close to the actual yield.

Conclusions

The paper presents an approach to map accurately cotton yields based on remote sensing and agrometeorological data. The proposed approach exploits the power of crop models to capture the physiological knowledge of how crops respond to weather, management and soil parameters, as well as the power of high resolution remote sensing data to observe the spatial variability of the observed canopy state variables. In particular, the WOFOST model was employed and calibrated in order to produce accurate daily simulated LAI and yield predictions against the available reference data. Then, based on these predictions, a new statistical model was developed able to map cotton yield from satellite multispectral images and meteorological data. The cotton forecasting model developed was validated against the actual yield in the study area. The developed method was proved to be reliable with a per pixel difference between the predicted and actual yield of approximately 5 kg. The calculated correlation between the actual and predicted yields were above 82% indicating the quite promising potentials of the approach.

References

De Wit A., Boogaard H., Fumagalli D., Janssen S., Knapen R., Kraalingen D. *et al.*, 2019. 25 years of the WOFOST cropping systems model, Agricultural Systems, 168, 154-167

Duan Q.Y., Gupta V.K., Sorooshian S., 1993, Shuffled complex evolution approach for effective and efficient global minimization, Journal of Optimization Theory and Application. Plenum Publishing Corporation, 76, (3), 501-521.

Guo, C., Zhang, L., Zhou, X. *et al.,* 2018. Integrating remote sensing information with crop model to monitor wheat growth and yield based on simulation zone partitioning, Precision Agriculture, 19(1) 55-78.

Jin X., Kumar L., Li Z., Feng H., Xu X., Yang G., *et al.,* 2018. A review of data assimilation of remote sensing and crop models. European Journal of Agronomy, 92, 141-152.

Jin Z., Prasad, R., Shriver J., Zhuang, Q., 2017. Crop model- and satellite imagery-based recommendation tool for variable rate N fertilizer application for the US Corn system, Precision Agriculture, 18(5), 779-800.

Karantzalos K., Karmas A., Tzotsos A., 2017. Monitoring Crop Growth and Key Agronomic Parameters through Multitemporal Observations and Time Series Analysis from Remote Sensing Big Data, In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture, Advances in Animal Biosciences, 8(2), 394-399.

Karantzalos K., Karmas A., Tzotsos, A., 2015. RemoteAgri: Processing Online Big Earth Observation Data for Precision Agriculture, In J V Stafford (Ed.) Proceedings of the 10th European Conference on Precision Agriculture. Wageningen, the Netherlands: Wageningen Academic Publishers. pp 421-428, doi: http://dx.doi.org/10.3920/978-90-8686-814-8_52.

Karmas A., Tzotsos A., Karantzalos K., 2016. Big Geospatial Data for Environmental and Agricultural Applications, Yu and Guo (Eds.), Big Data Concepts, Theories and Applications, Switzerland: Springer International Publishing, DOI 10.1007/978-3-319-27763-9.

Leon C.T., Shaw D.R., Cox M.S., Abshire M.J., Ward B., Wardlaw M.C. III *et al.*, 2003. Utility of Remote Sensing in Predicting Crop and Soil Characteristics, Precision Agriculture, 4: 359.

Lobell D., Thau D., Seifert C., Engle E., Little B., 2015. A scalable satellite-based crop yield mapper, Remote Sensing of Environment, 164, 324-333.

Ma G., Huang J., Wu W., Fan J., Zou J., Wu S., 2013. Assimilation of MODIS-LAI into the WOFOST model for forecasting regional winter wheat yield, Mathematical and Computer Modelling, 58, 634-643.

Watson, D.J., 1947. Comparative physiological studies on the growth of field crops: I. Variation in net assimilation rate and leaf area between species and varieties, and with and between years, Annals of Botany, 11, 41-76.

Can satellite-derived vigour maps be used to delineate homogeneous zones in hedgerow olive orchards?

F.J. Moral[1*], F.J. Rebollo[2], S. Millán[3], H. Prieto[3], J.M. Pérez[3] and C. Campillo[3]
[1]*Departamento de Expresión Gráfica, Escuela de Ingenierías Industriales, Universidad de Extremadura. Avenida de Elvas s/n, 06006 Badajoz, Spain; fjmoral@unex.es*
[2]*Departamento de Expresión Gráfica, Escuela de Ingenierías Agrarias, Universidad de Extremadura. Avenida de Adolfo Suárez s/n, 06007 Badajoz, Spain*
[3]*Centro de Investigaciones Científicas y Tecnológicas de Extremadura (CICYTEX), Finca La Orden, Junta de Extremadura. 06187 Guadajira (Badajoz), Spain*

Abstract

Some high resolution satellite images were downloaded in different phenological stages of the olive tree cycle during two years and the normalised difference vegetation index (NDVI) was computed in two olive orchards. Using map algebra techniques, all NDVI maps were integrated and zones with different plant vigour were delineated. Results showed that significant differences in some soil properties and yield exist between the three zones determined. Consequently, the proposed methodology, using satellite images at some important selected dates and computing a global NDVI map, can be a quick technique to define homogeneous zones in the more intensively cultivated olive groves.

Keywords: satellite images, Sentinel 2, intensive olive grove, normalised difference vegetation index (NDVI)

Introduction

Spain is the country with the highest number of olive trees (around 300 million), and is currently the world's leading olive and olive oil producer and exporter. Of the 2.5 Mha of olive groves, 92% are dedicated to olive oil production with an average annual production between 600,000 and 1,000,000 t. About 60% of the crop is concentrated in Andalusia, the biggest olive growing region on the planet, and Extremadura is the third, with around 10% of the area. Due to the suitable environmental conditions (Mediterranean climate and appropriate soil types), olive production has increased considerably during recent decades, particularly because of the adoption of high plantation densities and drip irrigation (e.g. Gómez *et al.*, 2014).

Although conventional farming is predominant in Extremaduran olive orchards, the existence of spatial variability indicates that site-specific management may be the best option. The spatial variation in crops is the result of a complex interaction of biological, edaphic, anthropogenic, topographic and climatic factors (Moral *et al.*, 2010).

Agriculture is oriented towards production systems that are more respectful of the environment, based on the quantitative knowledge of the agrarian system and its inter-relations with the environment. Site-specific crop management (SSCM) is an application of precision agriculture to crop production, identifying zones (management zones) with similar characteristics, such as texture, topography and nutrient levels (e.g. Schenatto *et al.*, 2017; Gavioli *et al.*, 2016), with the purpose of improving crop productivity and environmental quality. When a farmer decides to consider management zones, different tools are available: conventional soil sampling, yield monitoring or its estimation by intensive sampling, soil apparent electrical conductivity by contact or non-contact sensors, or vegetation indices derived from aerial or satellite multispectral imagery. However, the use of some of the aforementioned tools has important disadvantages. Conventional soil sampling

DOI 10.3920/978-90-8686-888-9_59, © Wageningen Academic Publishers 2019

is time consuming, expensive and provides only a limited spatial coverage. Similar problems can be also found when yield is considered. The alternative approach based on the analysis of surrogate data such as soil electrical properties and multispectral images, easily available at high resolution, can be the best option to delineate different management zones. For instance, Pedrera-Parrilla *et al.* (2014) and Moral *et al.* (2010) used soil electrical data for zoning purposes, and Hall and Wilson (2013) and Martínez-Casanovas *et al.* (2012) utilised vegetation indices computed from multispectral images. The use of satellite data to evaluate changes in vegetation properties aimed at applications in precision agriculture has been investigated elsewhere (e.g. Sibanda *et al.*, 2015; Thenkabail, 2003). Coarse resolution satellite images are necessary to describe phenology of vegetation (Testa *et al.*, 2014). The monitoring of crop offers direct information to analyse the spatial variability of the crop area. This information helps to justify management actions to improve production practices. The information can be used to adjust and direct fertilisation, to determine the development of the crop, to adapt irrigation to the needs of the crop and to schedule the harvest. One of the measures for determining crop development is the reflectance measurement, which is different depending on the type of surface and, in the case of plants, the species, the cultivar and the state of the plants (Plant, 2001). From this information, the normalised difference vegetation index (NDVI) can be computed (ratio of the difference of the values of reflectance in the near infrared and red bands and their sum). It is known that NDVI is a good indicator of the state of development of vegetation and, in consequence, can be used to detect canopy differences within a field (e.g. Gómez *et al.*, 2011).
In the light of previous research, the goal of this study was to test the potential of high spatial resolution data provided by multispectral satellite imagery to delineate homogeneous zones at a within-field scale. As far as the authors know, there is no other study in which vegetation indices have been used to define homogeneous zones in hedgerow olive orchards and, what is more important, an approach is proposed which may be used by farmers in a short timescale. Data utilised in the present study fits with practical constraints in terms of resolution and cost.

Materials and methods

Site description

The study was carried out in a farm located in the municipality of Alvarado, about 16 km East of Badajoz, in Southwestern Spain. The farm, called El Carrascal, covers an area of approximately 16 ha (38° 49' 27.15' N, 6° 46' 18.39' W, datum WGS84) and the field is divided into two separate hedgerow olive orchards: a rectangular-shaped field and a larger square-shaped field (Figure 1). Each orchard is planted with the variety Arbequina at a high density (1.35×4 m). The predominant soil in the field has been classified as a Haplic Fluvisol (FAO, 2006).

Data collection and image processing

Soil samples were collected in November 2016. Their co-ordinates were determined with sub-meter precision using a Mesa-Geode positioning system (Juniper Systems, Utah, USA). The sampling points were selected using a stratified random sampling scheme, according to a previous characterisation of the fields using a soil apparent electrical conductivity survey (not shown). Soil samples were taken at 0.30 m depth from 40 sampling sites (Figure 1). All soil samples were transported back to the lab in plastic bags and were air-dried, ground and passed through a 2-mm sieve. Texture, organic matter (OM) content, macronutrients (N, P, and K), and cation exchange capacity (CEC) were measured on each soil sample using conventional methods (Serrano *et al.*, 2014).
Additionally, 23 sampling points were selected to determine the yield. Each of these locations had 4 olive trees that were hand-harvested. The yield of each tree was weighed separately, and sub-samples of 100 olive fruits of the total harvest in each tree were collected to determine the average weight of the fruit and the number of fruits per tree. All subsamples were weighed fresh. Olive fruits were harvested in late October and early November 2016 and 2017.

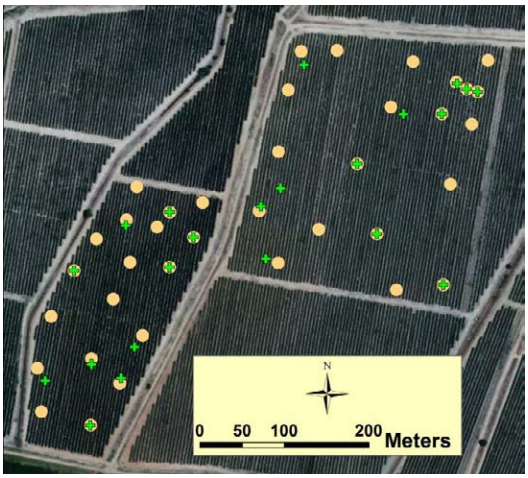


Figure 1. Aerial photography of the experimental fields. Locations of soil samples (circles) and olive trees for fruit yield sampling (crosses) are shown.

During 2016 and 2017, four images from the Sentinel satellites (2A and 2B), with a spatial resolution of 10 m, were selected and downloaded in different phenological stages of the olive tree cycle: phase I (vegetative development) in June, phase II (fruit development and stone hardening) in July and August, and phase III (ripening) in September. Satellite images were processed and analysed, using the SCP plugin in the geographical information system QGIS 2.18 (Congedo, 2018), performing the atmospheric correction and computing the NDVI values of each pixel. Spatial and temporal resolutions and the availability of images free of charge make Sentinel 2 one of the best options when satellite imagery is required for crop monitoring. Due to the high density of both olive groves, the effect of canopy cover discontinuity is of minor importance. NDVI values from plants are significantly higher than from bare soil. Consequently, within-field variation of NDVI due to soil variability was insignificant compared to variation of NDVI values caused by plant variability. Therefore, with a 10 m resolution pixel, NDVI values were predominantly from the tree canopy.

NDVI mapping for each acquisition date was performed using ArcGIS (version 10.3, ESRI Inc, Redlands, California, USA). NDVI values were reclassified into three types (low, medium, and high) and, later, all NDVI maps were combined with the raster calculator in ArcGIS using map algebra to compute the sum of the values associated to each cell in the overlays.

Finally, a statistical analysis was undertaken to check significant differences between zones. Since normality assumptions were not met with the data corresponding to some soil properties, the comparison of their mean values between NDVI zones was carried out using the Kruskal-Wallis non-parametric test and the Dunn test as a post hoc analysis. Differences of mean yield of some olive trees between NDVI zones were also checked.

Results and discussion

Firstly, the statistical parameters of the NDVI data corresponding to each selected date were computed (Table 1). The amplitudes of the NDVI values were important in all cases. Although mean NDVI values during the selected dates in 2016 were higher than in the selected dates in 2017, their coefficients of variation were lower than in 2017. However, in all cases, these values may suggest the existence of significant variability of the state of the olive canopy and, in consequence, the convenience of defining different homogeneous zones.

Table 1. Descriptive statistics for NDVI at each acquisition date of the satellite images.

Acquisition date	Minimum	Maximum	Mean	Coefficient of variation (%)
21 June 2016	0.45	0.56	0.51	3.92
24 July 2016	0.45	0.57	0.50	4.00
30 August 2016	0.45	0.58	0.52	3.85
19 September 2016	0.50	0.67	0.59	3.39
29 June 2017	0.31	0.45	0.39	5.13
27 July 2017	0.39	0.49	0.44	4.55
25 August 2017	0.39	0.50	0.43	4.65
17 September 2017	0.35	0.48	0.40	5.00

Figure 2 shows the four maps corresponding to 2016. In all cases, NDVI was mapped in the same quantiles, removing absolute differences between dates due to different conditions of image, climate or other factors. A temporal stability of NDVI was apparent throughout the selected dates in 2016 since the correlation coefficients between NDVI maps were higher than 0.92. Although NDVI maps for 2017 are not shown, the same temporal stability throughout the selected dates occurred (correlation coefficients between NDVI maps were higher than 0.94), which means that low zones remain low and high zones remain high over the year, but not that the same NDVI values are constant throughout the year.

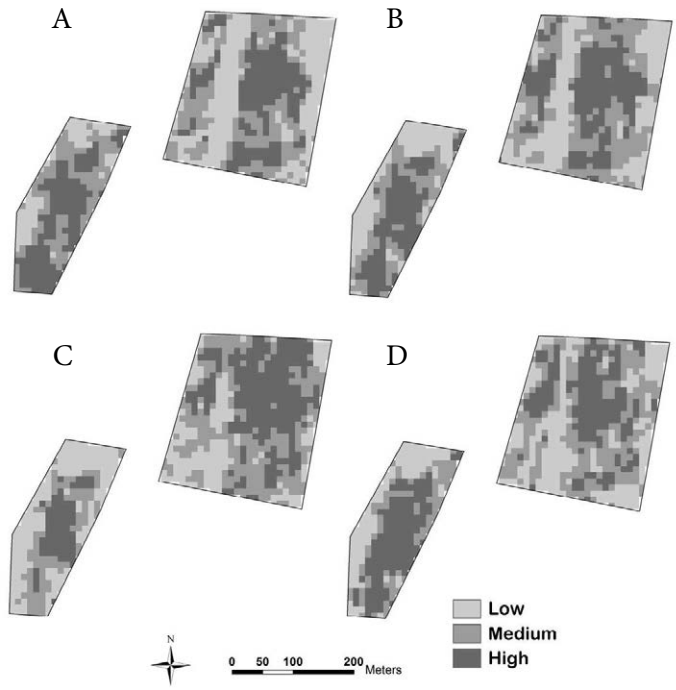


Figure 2. NDVI maps at different dates in 2016: (A) June, (B) July, (C) August, and (D) September. Quantile classification into three classes was used.

Figure 3 shows the global map after combining all reclassified NDVI maps, from 2016 and 2017. It was obtaining after considering the three categories for each individual map and assigning a value for each of them (low = 1, medium = 2, high =3); later, using the map algebra tool (raster calculator) in ArcGIS, the eight images were summed and a final classification was considered: low for cell values lower than 12, medium for values between 12 and 20, and high for values higher than 20. Finally, it was obtained that 26.69% of the area is in the 'high' class, so higher tree vigour and better water status can be expected and, consequently, higher olive fruit yield. The rest of the surface area corresponds to the 'medium' class, 47.52%, and 'low' class, 25.79%.

The significance of each homogeneous zone was determined after performing the non-parametric test. Table 2 shows that zoning by NDVI was able to explain variations in some soil properties. No significant differences in texture were observed among the three zones. However, differences were significant for other important soil properties, such as OM, pH, CEC, TN and K_2O. But, as may be unexpected, zones in which global NDVI were low or high had similar mean values for all soil properties. Consequently, if only the soil properties were taken into account, both zones would be merged. Nevertheless, when the mean yield of the olive trees was considered, differences were significant among the three zones. The yield in the zone with low global NDVI was 14.85 t/ha, in the zone with medium global NDVI was 14.06 t/ha, and in the zone with high global NDVI was 16.41 t/ha. Although there is not a direct relationship between the development of the vegetal cover, which is estimated by the NDVI data, and the tree yield, in this case study a significant relationship was evident between high NDVI values and high olive fruit yield.

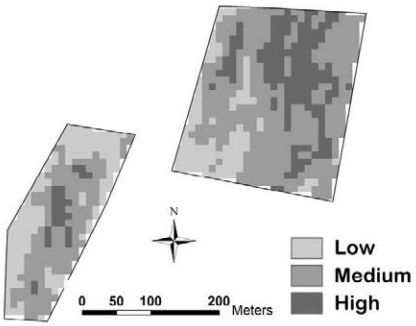


Figure 3. Global NDVI map obtained with the combination of the NDVI maps corresponding to the selected dates in 2016 and 2017 (June, July, August and September).

Table 2. Mean values of some soil properties of each homogeneous zone as defined by NDVI. Within each column, different letters indicate significant differences ($P<0.05$) according to the Dunn test.[1]

Zone	Sand %	Silt %	Clay %	OM %	pH	CEC cmol/kg	TN g/kg	P_2O_5 g/kg	K_2O g/kg
Low	69.44[a]	20.98[a]	9.57[a]	0.72[a]	8.20[b]	5.59[a]	0.07[a]	81.32[a]	0.30[a]
Medium	60.51[a]	22.22[a]	17.27[a]	0.94[b]	7.41[a]	10.12[b]	0.10[b]	59.49[a]	0.49[b]
High	72.62[a]	19.34[a]	8.03[a]	0.52[a]	6.94[a]	4.44[a]	0.07[a]	75.65[a]	0.29[a]

[1] CEC = cation exchange capacity, OM = organic matter; TN = total nitrogen.

The three homogeneous zones are based on tree vigour. While the high vigour zone corresponds with high tree yield, as was previously indicated, there are no differences in yield between the other two zones. Moreover, the low and high vigour zones show similarities in many soil properties compared to the medium vigour zone. In consequence, the cause for the differences in tree yield cannot be explained either by the NDVI or by the soil properties. Thus, the proposed zoning can lead to the delineation of homogeneous vigour zones but not to the determination of management zones.
Results have demonstrated that multitemporal NDVI images are insufficient for the delineation of management zones in hedgerow olive orchards. Despite a temporal stability of NDVI being apparent throughout the selected dates and during the two years considered in this study, only using homogeneous vigour zones to define management zones for implementing variable rate management is inadequate because the causes for the differences are unclear. It means that the spatial variation of those soil properties influencing the yield of the olive trees have to be considered to properly delineate management zones.

Conclusions

This study has shown that high spatial resolution data provided by multispectral satellite imagery is relevant to delineate homogeneous vigour zones in hedgerow olive orchards, which can be used to direct efficient soil or plant sampling and can be incorporated with other spatial layers to delimit management zones.
Since multispectral images, such as those provided by the Sentinel satellites, can be downloaded without any cost, the proposed approach can be applied widely, in particular when there are not many sampling data or in large-scale fields. However, if coherent management zones have to be delineated to implement SSCM, these homogeneous zones based on satellite images cannot be used solely. NDVI-based layer is very important but other layers related to some essential soil properties, such as soil apparent electrical conductivity, soil depth or texture, should be considered and integrated in the process to delineate management zones.

Acknowledgements

This research was co-financed by the Junta de Extremadura and the European Regional Development Fund (ERDF) through the projects GR (Research Groups TIC008 and AGA001) and regional strategy project AGROS. It was also funded by the INIA through the projects RTA2012-00018-C02 and RTA2013-00045-C04.
The authors are also very grateful to two anonymous reviewers for providing constructive comments which contributed to improve the final version of this paper.

References

Congedo, L., 2018. Semi-Automatic Classification Plugin Documentation. https://www.researchgate.net/publication/307593091_Semi-Automatic_Classification_Plugin_Documentation_Release_6011. (last accessed 28/09/18).

FAO, 2006. World reference base for soil resources. World Soil Resources Reports N. 103. Food and Agriculture Organization of the United Nations, Rome, Italy.

Gavioli, A., de Souza, E. G., Bazzi, C. L., Guedes, L. P. C. and Schenatto, K., 2016. Optimization of management zone delineation by using spatial principal components. Computers and Electronics in Agriculture 127 302-310.

Gómez, J. A., Infante-Amate, J., De Molina, M. G., Vanwalleghem, T., Taguas, E. V. and Lorite, I., 2014. Olive cultivation, its impact on soil erosion and its progression into yield impacts in Southern Spain in the past as a key to a future of increasing climate uncertainty. Agriculture 4 170-198.

Gómez, J. A., Zarco-Tejada, P. J., García-Morillo, J., Gama, J. and Soriano, M. A., 2011. Determining Biophysical Parameters for Olive Trees Using CASI-Airborne and Quickbird-Satellite Imagery. Agronomy Journal 103 644-654.

Hall, A., Wilson, M. A., 2013. Object-based analysis of grapevine canopy relationships with winegrape composition and yield in two contrasting vineyards using multitemporal high spatial resolution optical remote sensing. International Journal of Remote Sensing 34 1772-1797.

Martínez-Casasnovas, J., Agelet-Fernández, J., Arno, J. and Ramos, M., 2012. Analysis of vineyard differential management zones in relation to vine development, grape maturity and quality. Spanish Journal of Agricultural Research 10 326-337.

Moral, F. J., Terrón, J. M. and Da Silva, J. R., 2010. Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. Soil & Tillage Research 106 335-343.

Pedrera-Parrilla, A., Martínez, G., Espejo-Pérez, A. J., Gómez, J. A., Giráldez, J. V. and Vanderlinden, K., 2014. Mapping impaired olive tree development using electromagnetic induction surveys. Plant and Soil 384 381-400.

Plant, R.E., 2001. Site-specific management: the application of information technology to crop production. Computers and Electronics in Agriculture 30(1-3) 9-29.

Schenatto, K., de Souza, E. G., Bazzi, C. L., Gavioli, A., Betzek, N. M. and Beneduzzi, H. M., 2017. Normalization of data for delineating management zones. Computers and Electronics in Agriculture 143 238-248.

Serrano, J., Shahidian, S. and Marques da Silva, J., 2014. Spatial and temporal patterns of apparent electrical conductivity: DUALEM versus Veris sensors for monitoring soil properties. Sensors 14 10024-10041.

Sibanda, M., Mutanga, O., and Rouget, M., 2015. Examining the potential of Sentinel-2 MSI spectral resolution in quantifying above ground biomass across different fertilizer treatments. ISPRS Journal of Photogrammetry and Remote Sensing 110 55-65.

Testa, S., Mondino, E. C. B., and Pedroli, C., 2014. Correcting MODIS 16-day composite NDVI time-series with actual acquisition dates. European Journal of Remote Sensing 47 285-305.

Thenkabail, P. S., 2003. Biophysical and yield information for precision farming from near-real-time and historical Landsat TM images. International Journal of Remote Sensing 24(14) 2879-2904.

Estimating cotton water requirements using Sentinel-2: model development and validation

O. Rozenstein[1*], N. Haymann[1], G. Kaplan[1], and J. Tanny[1,2]
[1]*Institute of Soil, Water and Environmental Sciences, Agricultural Research Organization, Volcani Center, HaMaccabim Road 68, P.O.B 15159 Rishon LeZion 7528809, Israel; oferr@volcani.agri.gov.il*
[2]*HIT – Holon Institute of Technology, POB 305, Holon 58102, Israel*

Abstract

Crop coefficient (K_c)-based estimation of crop water consumption is one of the most commonly used methods for irrigation management. In this study, cotton evapotranspiration was measured in the field during two seasons using the eddy covariance method. K_c was estimated as the ratio between reference evapotranspiration and the measured cotton evapotranspiration. In addition, a time series of Sentinel-2 imagery was processed to produce 21 vegetation indices (VIs) based on the sensor's unique spectral bands. The results suggest that most VIs that are based on Sentinel-2 bands are suitable predictors for cotton K_c, and that those based on the red and red-edge spectral bands are the best ones. Consequently, this work sets the scene for near-real-time irrigation decision support systems.

Keywords: eddy covariance, remote sensing, irrigation, evapotranspiration, spectral modelling

Introduction

In addition to dwindling land resources, water scarcity is a growing concern. Water supply estimates for 2050 are such that close to half of the planet's population, mostly in developing nations, will suffer an acute water shortage (World Water Assessment Programme (WWAP), 2014; Mayor *et al.*, 2016). This water shortage is expected to affect food supply significantly. Consequently, food security starts with adequate and sufficient water for agriculture and with efficient use of water for irrigation. Crop coeffcient (K_c)-based estimation of crop water consumption is one of the most commonly used methods for irrigation management (Allen *et al.*, 1998). This approach for estimating the crop evapotranspiration (ET_c) is based on the estimation of reference evapotranspiration from a hypothetical crop (ET_0) such that: $ET_c = K_c \times ET_0$. While ET_0 is commonly derived using the Penman-Monteith Method, or based on pan evaporation measurements, K_c is empirically determined in water consumption experiments for specific crops in specific environments to isolate the atmospheric evaporative demand from the plant reaction. However, the use of standard K_c tables often results in erroneous ET_c estimations because of deviations from standard growing practices, extreme weather conditions, pests, crop diseases and other stress factors that inhibit regular crop development. Moreover, standard K_c tables that assume uniform field conditions, cannot handle within-field variations in ET_c. In previous studies, these issues were addressed by using vegetation indices derived from satellite imagery as surrogates of K_c (Duchemin *et al.*, 2006; Kamble *et al.*, 2013). Since some vegetation indices closely follow the K_c temporal trajectory, linear regression models between satellite derived vegetation indices and K_c can be used to create K_c maps of the field for satellite images throughout the season. These maps capture the actual crop development including the variability within the field, thus enabling more accurate estimation of ET_c and even precision irrigation. Up until recently, the application of this method was hindered by the lack of public domain imagery at a high revisit time with sufficient spatial resolution.
In a previous study (Rozenstein *et al.*, 2018), K_c estimation models were developed for cotton based on Sentinel-2 imagery that is available at a revisit frequency of 5 days at 10-20 m spatial resolution.

DOI 10.3920/978-90-8686-888-9_60, © Wageningen Academic Publishers 2019

This enabled, for the first time, to estimate K_c, at a frequency high enough to support irrigation decisions, and at a fine spatial resolution that well captures within-field variability. These results suggested that most vegetation indices that are based on Sentinel-2 bands are suitable predictors for cotton K_c. Moreover, vegetation indices that are based on the red and red-edge spectral bands performed better than indices developed for satellite sensors of previous generations (e.g. Landsat, SPOT, Modis) that do not use these bands. This progress in estimating cotton water consumption using satellite imagery that are available at no cost was a step forward towards the development of crop water requirements models. Following these results, the study was extended for another season of measurements to test and validate the remote sensing-based K_c estimation models for cotton. Accordingly, the specific objectives of this study were: (1) to measure the actual daily ETc empirically in a cotton field; (2) to validate and assess the performance of the Rozenstein *et al.* (2018) models in a cotton field different from the one in which they had been developed; (3) to combine the measurements from both seasons to improve the generality of the Sentinel-2 based ET_c prediction models.

Materials and methods

Study site and local climate

The measurements took place during the summer of 2017 in a cotton field near Negba, in the Shfela region in Israel (Figure 1). This location is about 20 km away from Tel-Nof, where field measurements were conducted in the previous year (Rozenstein *et al.*, 2018). This field was chosen because it was relatively flat, and large enough to contain most of the flux footprint measured by the eddy-covariance system (Burba, 2013). The field had a rectangle shape, with a mean length of 770 m and mean width of 500 m (the long dimension was oriented nearly north-south, see Figure 1. Cotton, cv. *Pima*, was sown on March 27, 2017, and drip-fertigated throughout the season. The irrigation was stopped on August 24, 2017, and the cotton was harvested on September 25, 2017. Field measurements of evapotranspiration were conducted between 25 July and 11 September 2017. The local summer climate is rainless. According to data supplied by the Israel Meteorological Service from the Negba meteorological station, about 2.5 km south-west of the experiment location, between 2015-2017 the

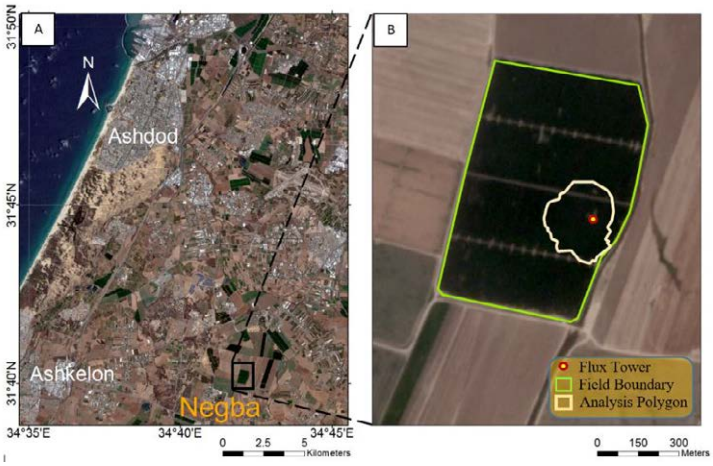


Figure 1. (A) Sentinel-2A RGB image acquired on 30 June 2017. The black square represents the footprint of the image on the right, (B) showing a blow-up of the area around the study site using a PlanetScope RGB image acquired on 29 June 2017.

annual average number of rainy days, i.e. rainfall ≥1 mm, was 59, with an annual average rainfall of 409 mm. There was no significant rainfall during the summer months – June through August (total summer rainfall for 3 years only 4.7 mm). Between 2015 and 2017, the average annual and growing season average (March-September) meteorological data were: relative humidity 68 and 68%, respectively; relative humidity at 12:00 – 52 and 50%, respectively; mean daily air temperature – 20.4 and 23.2 °C, respectively; mean daily minimum air temperature – 15.6 and 18.1 °C, respectively; and corresponding mean daily maximum temperature – 25.6 and 28.5 °C, respectively.

Agro-meteorological measurements

These measurements were conducted to determine ET_0 and ET_c from which the crop coefficient K_c was calculated as ET_c / ET_0. ET_0 was calculated based on meteorological data and ET_c was measured with an eddy covariance system. The Agro-Meteorology unit of the Israel Ministery of Agriculture and Rural Development regularly estimates the reference evapotranspiration (ET_0) at selected locations across Israel. In this study, ET_0 was calculated according to the Penman-Monteith method (Allen *et al.*, 1998) based on meteorological measurements of temperature, relative humidity, wind speed and solar irradiance at the Negba meteorological station located about 2.5 km south-west of the eddy-flux tower.

Direct measurements of ET_c were performed by an eddy covariance system that measured latent (LE) and sensible (H) heat fluxes in a manner similar to that described in Rozenstein *et al.* (2018). An eddy covariance (EC) system was installed on a tower at 3.7 m above the ground at a position that allowed a fetch of about 400 m for the prevailing northwesterly wind (Figure 1B). The system consisted of a three-axis ultrasonic anemometer (Model CSAT3; Campbell Scientific, Logan, UT, USA) that measured the wind speed vector and sonic temperature, and an Infra-Red Gas Analyzer, IRGA (Model LI-7500; LI-COR, Lincoln, NE, USA) that measured water-vapor concentration. To minimize wind distortion effects, the CSAT3 anemometer head was oriented towards the approximate direction of the predominant wind. Raw signals from the EC system were sampled at 20 Hz. Signals were recorded on a CR3000 data logger (Campbell Scientific, Logan, UT, USA) and later processed by the EddyPro software (LI-COR, Lincoln, NE, USA) to generate the flux data. The EddyPro software estimated the distance of the 90% flux footprint of each 30-min data point, later used to demark the area that contributed to the flux, thus defining an analysis mask (see the analysis polygon in Figure 1B).

Additional measurements facilitated energy-balance closure analysis. Net radiation (Rn) was measured with a net radiometer (Q*7.1; REBS, Seattle, WA, USA) installed at 3.65 m height, on the same pole of the EC system. Air temperature and relative humidity were measured by a sensor (HMP45, Campbell Sci., Logan, UT, USA) positioned at a height of 3.5 m within a ventilated radiation shield. Soil heat flux (G) was measured with four soil heat flux plates (HFT-3.1; REBS, Seattle, WA, USA) installed at a depth of 0.08 m within the soil and two thermocouples type T (copper-constantan) that were installed in the soil layer above each plate at depths of 0.02 and 0.06 m (total of 8 thermocouples). The calculation of soil heat flux was done similarly to the procedure described by Rosa *et al.* (2013) and Tanny *et al.* (2006). Thirty-minute averages of measured variables (net radiation, air temperature and relative humidity, soil heat flux and soil temperatures) were recorded on a CR23X data logger (Campbell Scientific, Logan, UT, USA). All the equipment was powered by car batteries that were charged during the day by solar panels.

Satellite image processing

A total of seven relatively cloud-free Sentinel-2A images acquired during the cotton-growing season were analyzed in this study. The images were downloaded as Level-2A atmospherically corrected products from the Sentinels Scientific Data Hub and processed using SeNtinel Application Platform (SNAP) (Gascon and Ramoino, 2017). In order to perform the analysis only for pixels that contributed to the measured water vapor flux, an analysis mask was defined using a similar

protocol to Rozenstein *et al.* (2018). In short, the pixels selected for the analysis were within the 90% flux footprint of eddy-covariance measurements taken between 7 am and 6 pm, cloud-free throughout the imagery time-series, and contained within the field boundary. The analysis mask was produced by thresholding the half-hourly flux point density histogram at 1%. This final analysis mask corresponds to 393 Sentinel-2 pixels at 10 m^2 each.
Twenty-one vegetation indices included in SNAP were derived for all of the images throughout the growing season. Linear regression models were developed for the time series of field measured K_c and the time series of each spectral index. R^2 and root mean square error (RMSE) values were calculated for every model. The K_c –NDVI model developed in this study was compared to other similar models from the literature that were developed for cotton (Montgomery *et al.*, 2015) and with a general field crop model (Kamble *et al.*, 2013). In addition, the field measured K_c values, and the estimates of K_c -NDVI models were compared with the recommended K_c for cotton in this region that is provided by the Israeli Extension Service.

Results

The energy balance closure was used to assess the validity of the EC flux measurements (Wilson *et al.*, 2002). In this analysis, half-hourly values of consumed energy of latent and sensible heat, LE+H, are regressed against corresponding values of the available energy, Rn-G, where Rn is net radiation and G is soil heat flux. The 0.84 regression slope of the LE+H against Rn-G indicates a good energy balance closure, and good reliability of the flux measurements achieved using the eddy-covariance system (albeit, this was not as good as the Tel-Nof energy balance closure). Most of the regression models between each of the 21 vegetation indices and K_c were found to be good at the Tel-Nof experiment but the models performance in Negba were not as good. For example, the NDVI model did not perform satisfactorily, however, when compared against other NDVI models (Kamble *et al.*, 2013; Montgomery *et al.*, 2015), it gave a better prediction for K_c (Table 1). Still, MTCI performed best, while REIP and S2REP that were found to perform well in Tel-Nof also worked well in Negba. The Israeli Extension Service K_c table values were usually higher than those measured and predicted for mid-season, and lower at the end of the season (Figure 2).

Discussion

The results suggest that most VIs that are based on Sentinel-2 bands are suitable predictors for cotton K_c, and that those based on the red and red-edge spectral bands (MTCI, REP, and S2REP) are the best ones. Other indices (i.e. EVI and TSAVI) show acceptable performance, but not as good as red-edge based MTCI, which demonstrated the best K_c prediction ability of all the indices tested in this study. Therefore, these results demonstrate the potential of developing K_c prediction for field crops using Sentinel-2 that would perform better compared to older public domain satellite sensors that did not cover the red-edge region, and did not provide as short a revisit time as the Sentinel-2 pair of sensors (i.e. Landsat, SPOT).
In this study, the field measurements were carried out during the mid-late stages of the growing season. Thus, the models developed in this paper still require further testing and adaptations to extend their application to the entire growing season. In future studies, irrigation trials would be conducted to test the MTCI based model against the common irrigation practice in Israel to see if it can increase the water use efficiency without significant yield reduction.

Table 1. R^2 and RMSE of K_c – Vegetation Index regression models ranked according to the RMSE of the combined models: each model was tested against the data from the location for which it was developed and the combined models were tested for the combined dataset of both seasons.

Index	Tel Nof models		Negba models		Combined models	
	R^2	RMSE	R^2	RMSE	R^2	RMSE
MTCI	0.9915	0.0079	0.9111	0.0511	0.8892	0.0565
TSAVI	0.7673	0.0413	0.8171	0.0751	0.8608	0.0639
EVI	0.9206	0.0241	0.7780	0.0808	0.8596	0.0645
GEMI	0.9942	0.0125	0.8065	0.0755	0.8282	0.0714
WDVI	0.9942	0.0134	0.8065	0.0787	0.8097	0.0752
S2REP	0.9372	0.0214	0.7228	0.0903	0.8043	0.0790
PVI	0.8178	0.0365	0.7466	0.0864	0.7759	0.0816
DVI	0.8959	0.0276	0.7029	0.0935	0.7759	0.0816
REIP	0.896	0.0276	0.7031	0.0935	0.8042	0.0845
MSAVI	0.9492	0.0193	0.6330	0.1079	0.7078	0.0930
SAVI	0.9408	0.0208	0.6582	0.1003	0.7052	0.0934
IRECI	0.9576	0.0176	0.6295	0.1044	0.7044	0.0937
ARVI	0.9582	0.0175	0.5953	0.1091	0.6912	0.0950
MSAVI2	0.952	0.0188	0.5871	0.1102	0.6567	0.1005
TNDVI	0.949	0.0193	0.4645	0.1255	0.5189	0.1183
IPVI	0.9493	0.0193	0.4581	0.1263	0.5125	0.1191
NDVI	0.9491	0.0193	0.4581	0.1263	0.5124	0.1191
GNDVI	0.9484	0.0194	0.3714	0.1360	0.4630	0.1254
PSSRa	0.8902	0.0284	0.3237	0.1411	0.2918	0.1407
RVI	0.8843	0.0291	0.3155	0.1419	0.2786	0.1418
NDI45	0.91	0.0257	0.1446	0.1587	0.1897	0.1510
NDVI (Kamble *et al.*)		0.1969		0.3470		0.3057
NDVI (Montgomery *et al.*)		0.2118		0.3623		0.3210

Conclusions

Sentinel-2 is superior to older generations of public domain satellite data in terms of spatial, temporal and spectral resolutions. This allows for the first time to estimate K_c, an important parameter for irrigation management, at a high frequency that can support irrigation decisions, at a fine spatial resolution of 10 m that well captures within-field variability, and at higher accuracy than before, owing to the sensor's unique spectral bands that cover the red-edge region.

Acknowledgements

This study was partially supported by an ARO Start-Up Grant held by Offer Rozenstein. Gregoriy Kaplan was supported by an absorption grant for new immigrant scientists provided by the Israeli Ministry of Immigrant Absorption. We thank Mr Victor Lukyanov for technical support with the field measurements. The field measurements were funded by the Chief Scientist of the Ministry of Agriculture, Israel, under grant number 304-0505.

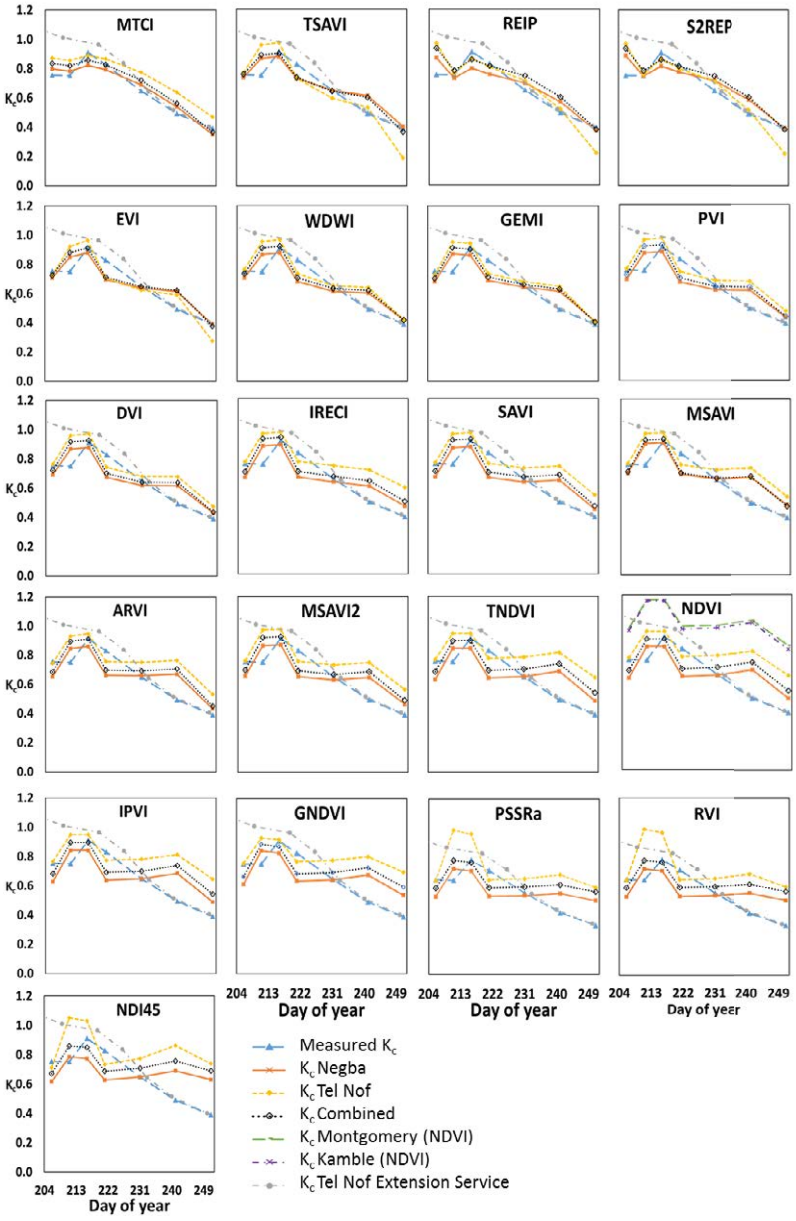


Figure 2. The temporal trajectory of K_c based on Eddy Covariance (EC), estimated by different vegetation indices from seven Sentinel-2 images, the two different models developed in Tel-Nof and Negba, a combined model for both sites, models from the literature (Montgomery *et al.*, 2015; Kamble *et al.*, 2013), and the recommendation of the Israeli Extension Service.

References

Allen, R.G., Pereira, L.S., Raes, D., & Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Rome,Italy: FAO 300*, D05109.

Burba, G., 2013. Eddy covariance method for scientific, industrial, agricultural and regulatory applications: A field book on measuring ecosystem gas exchange and areal emission rates. Lincoln, NE, USA: LI-Cor Biosciences.

Cohen, J.E., 2005. Human population grows up. *Scientific American, 293*, 48-55.

Duchemin, B., Hadria, R., Erraki, S., Boulet, G., Maisongrande, P., Chehbouni, A., *et al.,* 2006. Monitoring wheat phenology and irrigation in Central Morocco: On the use of relationships between evapotranspiration, crops coefficients, leaf area index and remotely-sensed vegetation indices. *Agricultural Water Management, 79*, 1-27.

Gascon, F., & Ramoino, F., 2017. Sentinel-2 data exploitation with ESA's Sentinel-2 Toolbox. European Geosciences Union *(EGU) General Assembly Conference Abstracts, 19*, 19548.

Kamble, B., Kilic, A., & Hubbard, K., 2013. Estimating crop coefficients using remote sensing-based vegetation index. *Remote Sensing, 5*, 1588-1602.

Mayor, B., Casado, R.R., Landeta, J., López-Gunn, E., & Villarroya, F., 2016. An expert outlook on water security and water for energy trends to 2030-2050. *Water Policy, 18*, 1-18.

Montgomery, J., Hornbuckle, J., Hume, I., & Vleeshouwer, J., 2015. IrriSAT – Weather based scheduling and benchmarking technology. In T. Acuña, C. Moeller, D. Parsons and M. Harrison (Eds.). *Building Productive, Diverse and Sustainable Landscapes*'. *Proceedings of the 17th Australian Agronomy Conference*. http://agronomyaustraliaproceedings.org/images/sampledata/2015_Conference/pdf/agronomy2015final00449.pdf (last accessed Feb 2019)

Rosa, R., Dicken, U., & Tanny, J., 2013. Estimating evapotranspiration from processing tomato using the surface renewal technique. *Biosystems Engineering, 114*, 406-413.

Rozenstein, O., Haymann, N., Kaplan, G., & Tanny, J., 2018. Estimating cotton water consumption using a time series of Sentinel-2 imagery. *Agricultural Water Management, 207*, 44-52.

Schmitz, C., van Meijl, H., Kyle, P., Nelson, G.C., Fujimori, S., Gurgel, A., *et al.,* 2014. Land-use change trajectories up to 2050: insights from a global agro-economic model comparison. *Agricultural Economics, 45*, 69-84.

Tanny, J., Haijun, L., & Cohen, S., 2006. Airflow characteristics, energy balance and eddy covariance measurements in a banana screenhouse. *Agricultural and Forest Meteorology, 139*, 105-118.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., *et al.,* 2002. Energy balance closure at FLUXNET sites. *Agricultural and Forest Meteorology, 113*, 223-243.

World Water Assessment Programme (WWAP), 2014. The United Nations World Water Development Report 2014: Water and Energy. *Paris, France: UNESCO.*

Satellite-based detection of sowing dates in field crops

Y. Sadeh[1*], X. Zhu[1], K. Chenu[2] and D. Dunkerley[1]
[1]*School of Earth, Atmosphere and Environment, Monash University, Clayton, Victoria 3800, Australia; yuval.sadeh@monash.edu*
[2]*The University of Queensland, Queensland Alliance for Agriculture and Food Innovation (QAAFI), 203 Tor Street, Toowoomba, QLD 4350, Australia*

Abstract

Sowing dates affect the environmental conditions that plants will experience during their lifetime and thus greatly influence yields. This paper presents a newly developed innovative, semi-automated method to detect sowing of crops on no-till fields using high spatio-temporal resolution CubeSat satellites. This method succeeded in identifying the actual sowing dates of individual fields with an RMSE of 0.9 and 1.9 days in a set of 16 national trial fields and in 50 fields of a commercial farm, respectively. The method is robust and can be applied over a wide range of soil types, atmospheric conditions, crop types and sensors, to detect sowing dates even at sub-field scales. It opens avenue to improve satellite-based yield prediction.

Keywords: sowing dates, CubeSats, satellite images, farming management, planting dates

Introduction

Sowing dates have a large impact on crop development and growth and ultimately crop yields as they affect what environmental conditions the plants are likely to encounter (Coventry *et al.,* 1993; Flohr *et al.,* 2017). Sowing dates are thus important to organizations and individuals who are interested in production, processing, marketing and trade of food and its products (Guo, 2013). They are also major inputs of crop models for assessing the expected yield effects of different management practices (Chenu *et al.,* 2017; Holzworth *et al.,* 2014). However, obtaining accurate sowing dates from ground reports is difficult and time consuming (Marinho *et al.,* 2014; Sacks *et al.,* 2010). Only a few census datasets containing this information are made publicly available (Urban *et al.,* 2018). The advantage of remote sensing for collecting data on crop management practices has long been recognised. However, a satellite-based methodology for directly detecting sowing dates has not yet been developed. That is mainly because there is a lag-time between the sowing date and the plant emergence. Most existing studies have attempted first to detect the crop 'green-up', i.e. the earliest reliable evidence of vegetation that can be sensed on satellite images (Lobell *et al.,* 2013). Next, they backcast emergence and sowing dates by assuming the lag time between the sowing and the green-up detection to be constant (e.g. Gao *et al.,* 2017; Lobell *et al.,* 2013; Manfron *et al.,* 2017). In recent decades, the number of farmers around the world who adopted no-tillage sowing increased dramatically (Hobbs *et al.,* 2008). When no-tillage farming is implemented, seeds are directly sown into the previous crop residue and stubble. From the remote sensing perspective, no-tillage creates a distinct contrast between the sown and un-sown soil. To accurately detect sowing dates at a field or sub-field scale from space, very high spatial and temporal resolution satellite remote sensing systems, such as constellations of micro or nano-satellites (known as CubeSats), is needed. In contrast to satellites such as Landsat or Sentinel-2, which traditionally developed and launched by space agencies, these new CubeSat satellites, such as PlanetScope (Planet Labs Inc., San Francisco, CA, USA), are relatively inexpensive to build and allow collection of both high spatial (<5 m) and temporal resolution (<1 week) imagery at lower cost (Dash *et al.,* 2016; Jain *et al.,* 2016). However, CubeSat constellations can suffer from cross-sensor inconsistency (Houborg *et al.,* 2016), which may limit the accuracy of spectral measurements and analysis such as vegetation indices, but also

DOI 10.3920/978-90-8686-888-9_61, © Wageningen Academic Publishers 2019

could hinder the use of this platform to monitor changes on the Earth surface. In this paper, a robust method to use PlanetScope satellite images to detect sowing dates at the field or sub-field scales via detection of the no-tillage sowing operation is proposed. It has been applied to detect sowing dates with high accuracy on experimental wheat (*Triticum aestivum*) fields across Australia and a typical commercial farm in Victoria having a wide range of soil, climate conditions and crop species.

Materials and methods

In this paper, sowing dates are detected via image processing that detects changes between consecutive images using PlanetScope data. The approach assumes that sowing corresponds to the first detectable field-scale change in the surface after the harvest of the previous crop, as expected for the widely-used no-tillage farming practices (Sadeh *et al.,* 2019).

Study fields

The approach was first tested on 16 testing fields operated by the Australian National Variety Trials (NVT). The NVT fields (sized 70×40 m on average) were located at eight sites in different regions across Australia, which have diverse management practices, soil types and climatic conditions. In a second stage, the method was expanded to include 50 fields of a commercial farm located near Birchip, Victoria. This 6,400 ha farm (the average field size is 116 ha) grows a large variety of crops including wheat, barley (*Hordeum vulgare*), canola (*Brassica napus*), lentils (*Lens culinaris*), oats (*Avena* spp.), vetch (*Vicia sativa*) and peas (*Pisum sativum*). The studied sowing period in this farm spanned from late-March to mid-June. All fields, from both NVT and farm, were sown with no-tillage, which now is the most common practice used in Australia.

Imagery

Images were sourced from the satellite constellation PlanetScope, which comprises multiple launches of groups of individual CubeSat, each 10 cm by 10 cm by 30 cm in size (Planet Team, 2018). A complete PlanetScope constellation consists of about 120 satellites, which have the capability to imaging the entire earth's land surface on a daily basis. The PlanetScope satellites have four spectral bands: Blue, Green, Red and Near Infrared (NIR) with a spatial resolution of 3-4 m at nadir and positional accuracy of <10 m RMSE (Planet Team, 2018). This study used a total of 205 (78 for the NVT and 127 for the farm) PlanetScope scenes acquired by 78 different satellites (61 for the NVT and 30 for the farm). For each NVT location, PlanetScope images covered a period of ten days before and after reported sowing. The NVT fields had a median cover of 9.5 images over the whole period, with a 2-d median gap between the images. For the farm, 31 images could be used (from 17/04/2017 to 20/06/2017), with a 5-d median gap between cloud-free images.

Semi-automated sowing detection

The semi-automated sowing detection approach is based on the assumption that the first change at the field-scale on the surface after harvesting is related to the no-tillage sowing practice. It involves detecting sowing dates based on spectral changes between successive images. The method is semi-automated, requiring three inputs: (1) two satellite images acquired on different dates ($Image_{t1}$ and $Image_{t2}$); (2) the field boundaries; and (3) a threshold value to detect changes between images (Figure 1). This threshold corresponds to a pixel value in the processed image histogram that indicates a change (Figure 2C). First, the images from the same day (not necessarily acquired by the same satellite) mosaicked and clipped using the fields' boundaries. Then, for each image, the principal component analysis (PCA) method was applied, to transform the images to four principal components (PC_1, PC_2, PC_3 and PC_4) (Figure 1). PC_1 systematically found to be the best PC to represent the variance in the soils, while PC_2 tended to represent variations in the vegetation, PC_3

and PC_4 mainly contained noise. Afterwards, PC_1s from two consecutive images ($Image_{t_1}^{PC_1}$ – the early image and $Image_{t_2}^{PC_1}$ – the later image) were used to create a ratio-image.
The ratio image (Figure 1) represents the magnitude of change between the two images at the pixel level (Figure 2B). A pixel value closer to 1 indicates no change. The histogram of the ratio image generally exhibits a normal distribution (bell-shaped) around the value of 1 if there is no change across the image. When changes occurred, the histogram would show a multimodal distribution, with several distributions combined, i.e. would have multiple peaks and valleys. By examining the distributions rather than the one representing 'no change', a threshold pixel (ratio) value to identify changes was determined. For example, in Figure 2C, there are two distributions of the pixel values: one is a normal distribution around 1 and the other is a normal distribution around about 2.8. The second distribution represents the changes. The left end of the second distribution is set as the

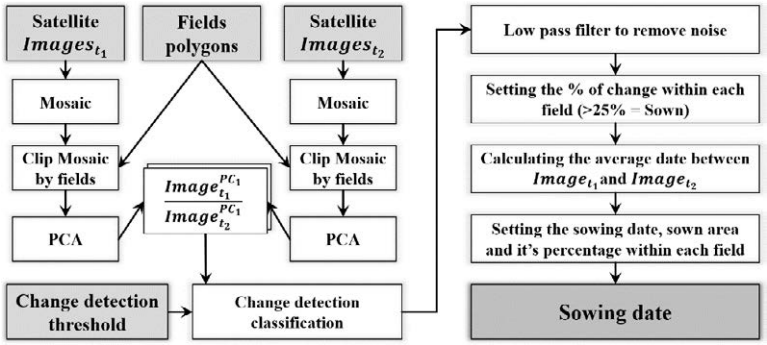


Figure 1. Sowing-detection workflow.

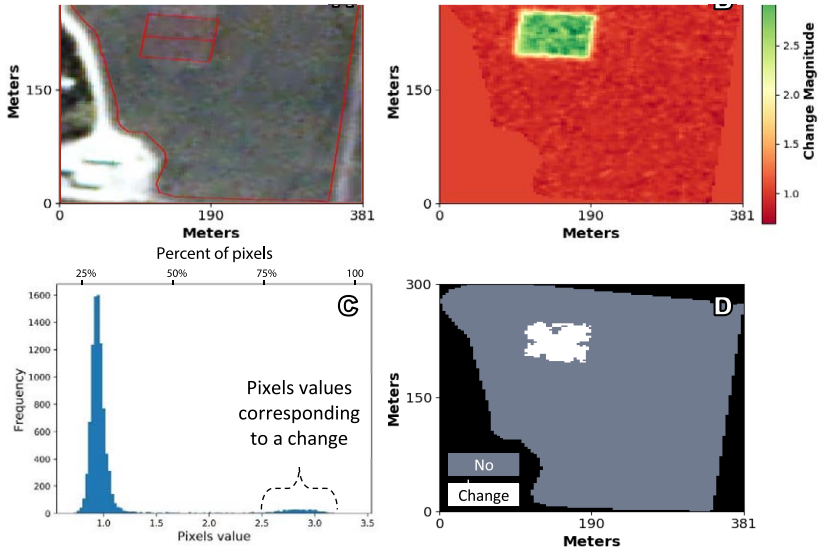


Figure 2. Sowing detection of a newly sown experimental field. (A) Image of a NVT wheat fields after sowing (red rectangles). (B) Image resulting from dividing $Image_{t_1}^{PC_1}$ by $Image_{t_2}^{PC_1}$ (C) Histogram of the image in B. (D) Results of an image classification based on the threshold set to 2.5.

threshold value. A pixel with a value equal to or greater than this threshold is considered as 'change'. The threshold values may change from scene to scene due to cross-sensor inconsistencies among the different satellites in the constellation. In this case study, threshold values ranged between 1.2 and 3.

Results

A very high correlation (R^2= 0.99) was found between reported and detected sowing dates in both the NVT and commercial farm fields (Figure 3). This method detected 100% (RMSE=0.9 days, mean error=-0.6 day gap) of the sowings in NVT fields and 80% in the farm fields (RMSE=1.9 days, mean error=-0.1 day gap) (Table 1). During the analysis, a few false detections were found, where a field was incorrectly classified as sown. In the farm analysis, 5% of unsown fields were falsely detected as sown, but 95% of the sown fields were correctly identified. In some cases the presence of clouds in the images was translated as a change that covered most of the field and therefore was determined as sown. Detection failures, i.e. no detection of sowing, which occurred in 20% of the farm's sown fields, were also partly due to cloud cover (50% of cases) and the large time gap (more than 4 days) between $Image_{t1}$ and $Image_{t2}$ (33% of cases).
Figure 4 shows that most of the successful detections occurred when the gap between the sowing and the satellite image date was under four days, and the probability for a successful detection decreased as the gap in days between the sowing and image capture increased.

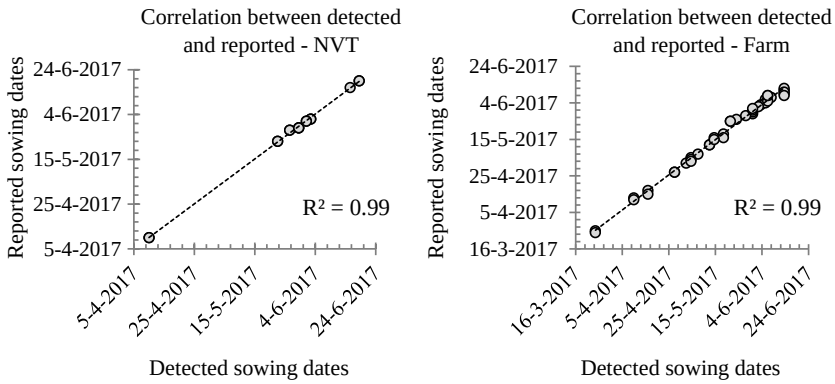


Figure 3. Correlation between the reported and detected sowing dates of the 16 NVT fields (on the left) and the 50 farm fields (on the right).

Table 1. The accuracy of sowing detection.

	NVT	Birchip farm
Sowing detected (No. of fields)	16/16 (100%)	40/50 (80%)
Failed to detect (No. of fields)	0 (0%)	10 (20%)
RMSE (days)	0.9	1.9
Mean error (days)	-0.6	-0.1
Median error (days)	0.0	0.0
Standard deviation (days)	0.7	1.9

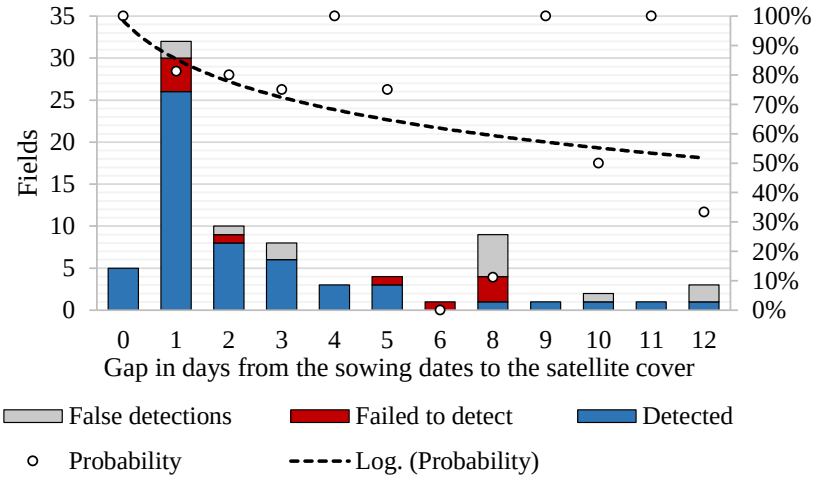


Figure 4. Distribution of the 56 successful (in blue), 10 unsuccessful (in red) and 10 false (in grey) detections of the sowing dates (left Y-axis). The dashed line shows the probability of achieving a successful detection (black points) relatively to the gap in days from the sowing date to the satellite cover (right Y-axis).

Although the method was designed to determine the sowing dates at a field-scale, it could be applied to identify sub-field sowing. For instance, sowing of a small proportion of the field was detected in the Birchip farm (Figure 5), as the grower did not have time to complete sowing of the entire field, which is relatively common in this region.

Discussion

Remote sensing has the potential for mapping sowing dates. To the authors' knowledge, no methodology has been published on how to detect sowing dates directly from satellite images. Until

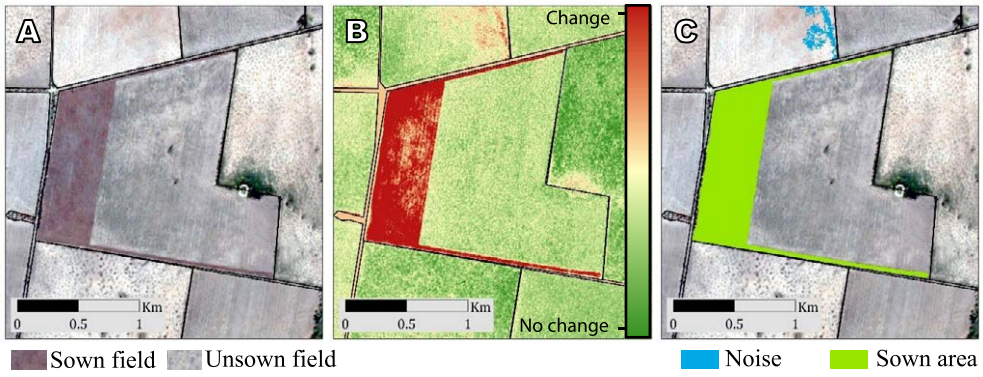


Figure 5. Sub-field sowing detection. (A) the RGB satellite image of a partly sown field (the area in brown). (B) the ratio image of $\text{Image}_{t_1}^{PC_1}$ and $\text{Image}_{t_2}^{PC_1}$. A change between the images resulted in high values (red). (C) the subfield detected as sown (in green).

recently, such detection was not possible due to technological limitations (low spatial resolution and long revisit time). Very high spatial and temporal resolution CubeSats images may overcome them, particularly for detecting no-tillage sowing, which is becoming the main sowing practice in major cropping regions worldwide (Kassam *et al.,* 2015). CubeSats such as PlanetScope suffer from cross-sensor inconsistencies (as the CubeSats satellites are not identical) and from relatively low signal-to-noise ratio (McCabe *et al.,* 2017). Therefore, the main challenge when conducting change detection using this type of satellite is to separate the noise from the actual change that happened on the earth surface between the earlier and later images. To overcome these limitations, PCA was used as it is useful for detecting decorrelations between images (Du *et al.,* 2007) and also for separating an underlying systematic data structure from noise (Wold *et al.,* 1987). It was found that PCA was very effective in extracting the dominant patterns in the images and detecting changes between the main patterns on each pair of images (Figure 2). This method allowed minimising the cross-sensor inconsistencies between the images, even when images acquired by 78 different PlanetScope satellites with different atmospheric conditions were used.

The results of the semi-automated sowing detection method have better performance and accuracy than previous methods. These studies were able to estimate sowing dates with a RMSE from 6.68-12.41 days, with R^2 ranging between 0.15 and 0.90 between their estimations and official regional statistics (e.g. Jain *et al.,* 2016; Lobell *et al.,* 2013). In contrast to the Green-up approach, the sowing detection method approach does not depend on local conditions (soil or crop types, rainfall), does not require the identification of a fixed number of days to estimate emergence and is therefore more suitable for diverse environments, as long as no-tillage is practiced. The results show that most successful detections took place when the gap between the sowing and the satellite image dates was <4 days (Figure 4). In these conditions, the successful detection rate was 91%. However, when this gap was larger than 5 days, the percentage of successful detections declined to 62% (Figure 4). Therefore, using satellites with a temporal resolution of >4 days such as Sentinel-2 or Landsat, will be less likely to achieve the high accuracy of detection, as did the method reported here.

Studies showed that high-resolution information of sowing dates could be used to reduce the uncertainty of yield prediction using crop simulations (e.g. Mathison *et al.,* 2017). However, the existence of such datasets is currently scarce and, when released, these data are often shared at aggregated regional scales (Urban *et al.,* 2018). Establishing a large-scale database of field/farm management practices is time consuming, expensive, relies heavily on manual labour and requires some training (Jain *et al.,* 2016). The methodology presented in this study, can be used to produce low-cost field level statistics, which can be updated on a near-daily basis. In addition, statistics at the field-level may improve the understanding of the reasons for yield gaps, and could assist in identifying low-performing fields that need to be treated in order to increase their productivity (Jain *et al.,* 2016).

In spite of its very high accuracy, this sowing detection methodology has limitations, such as sensitivity to the presence of clouds in the image and manual identification of the change detection threshold. In order to fully automate the process, it is recommend that future studies: (1) evaluate the suitability of existing cloud and cloud shadow detection methods to mask clouds from these new satellite images; (2) explore machine-learning clustering techniques to identify the pixel values that represent change in the histogram; and (3) implement classification and segmentation techniques to extract field boundaries. Despite the inconsistency of spectral data, CubeSats are only starting to unleash their potential. As more advanced CubeSat constellations deployed into space, the number of CubeSats-based applications at the farm, field and sub-field level will increase and their efficacy for monitoring farm managements will improve.

Conclusions

This study investigated the potential of using high spatio-temporal resolution CubeSats images to detect sowing dates at the field scale. An innovative semi-automated sowing detection methodology was developed. It uses PlanetScope data to detect changes on the field surface caused by no-tillage sowing. The method demonstrated capabilities to achieve successful detections over a wide range of soil types, atmospheric conditions, crop types and sensors. Overall, the method detected 85% of the fields with a very high correlation (R^2=0.99) between actual and estimated dates. Time of sowing was detected with a median gap of 0 day. It achieved an RMSE of 0.9 and 1.9 days in a national dataset and in a representative commercial farm, respectively. The approach presented in this study may be used regionally or nationally to produce and update near-daily field level statistics in a spatial coverage that can identify crop sowing even at sub-field scale. As sowing dates influence grain yield, information about sowing dates may be used to improve yield prediction, e.g. with crop model simulations. In the future, the accuracy of this method is expected to increase as more CubeSats constellations are being built and deployed.

Acknowledgements

The authors would like to thank Tim McClelland (Partner A.W. McClelland & Co, Birchip VIC 3483) for sharing his farm sowing dates, constructive feedback and time. We also thank the National Variety Trials (NVT) program for providing field trials dataset, and Planet Labs, Inc. for providing PlanetScope imageries used in this study.

References

Chenu, K., Porter, J.R., Martre, P., Basso, B., Chapman, S.C., Ewert, F., *et al.,* 2017. Contribution of Crop Models to Adaptation in Wheat. Trends in Plant Science, 22(6), 472-490.

Coventry, D., Reeves, T., Brooke, H., & Cann, D., 1993. Influence of genotype, sowing date, and seeding rate on wheat development and yield. Australian Journal of Experimental Agriculture, 33(6), 751-757.

Dash, J., & Ogutu, B.O., 2016. Recent advances in space-borne optical remote sensing systems for monitoring global terrestrial ecosystems. Progress in Physical Geography, 40(2), 322-351.

Du, Q., & Fowler, J.E., 2007. Hyperspectral image compression using JPEG2000 and principal component analysis. IEEE Geoscience and Remote Sensing Letters, 4(2), 201-205.

Flohr, B.M., Hunt, J.R., Kirkegaard, J.A., & Evans, J.R., 2017. Water and temperature stress define the optimal flowering period for wheat in south-eastern Australia. Field Crops Research, 209, 108-119.

Gao, F., Anderson, M.C., Zhang, X., Yang, Z., Alfieri, J.G., Kustas, W.P., *et al.,* 2017. Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. Remote Sensing of Environment, 188, 9-25.

Guo, Z., 2013. Mapping the planting dates: An effort to retrive crop phenology information from MODIS NDVI time series in Africa. In, International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 3281-3284)

Hobbs, P.R., Sayre, K., & Gupta, R., 2008. The role of conservation agriculture in sustainable agriculture. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1491), 543-555.

Holzworth, D.P., Huth, N.I., deVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., *et al.*, 2014. APSIM – Evolution towards a new generation of agricultural systems simulation. Environmental Modelling & Software, 62, 327-350.

Houborg, R., & McCabe, M., 2016. High-Resolution NDVI from Planet's Constellation of Earth Observing Nano-Satellites: A New Data Source for Precision Agriculture. Remote Sensing, 8(9)

Jain, M., Srivastava, A., Balwinder, S., Joon, R., McDonald, A., Royal, K., *et al.,* 2016. Mapping Smallholder Wheat Yields and Sowing Dates Using Micro-Satellite Data. Remote Sensing, 8(10), 860-878.

Kassam, A., Friedrich, T., Derpsch, R., & Kienzle, J., 2015. Overview of the worldwide spread of conservation agriculture. Field Actions Science Report, 8

Lobell, D.B., Ortiz-Monasterio, J.I., Sibley, A.M., & Sohu, V.S., 2013. Satellite detection of earlier wheat sowing in India and implications for yield trends. Agricultural Systems, 115, 137-143.

Manfron, G., Delmotte, S., Busetto, L., Hossard, L., Ranghetti, L., Brivio, P.A., *et al.,* 2017. Estimating inter-annual variability in winter wheat sowing dates from satellite time series in Camargue, France. International Journal of Applied Earth Observation and Geoinformation, 57, 190-201.

Marinho, E., Vancutsem, C., Fasbender, D., Kayitakire, F., Pini, G., & Pekel, J.F., 2014. From remotely sensed vegetation onset to sowing dates: Aggregating pixel-level detections into village-level sowing probabilities. Remote Sensing, 6(11), 10947-10965.

Mathison, C., Deva, C., Falloon, P., & Challinor, A.J., 2017. Defining sowing and harvest dates based on the Asian Summer Monsoon. Earth System Dynmics Discussions, 1-39.

McCabe, M.F., Aragon, B., Houborg, R., & Mascaro, J., 2017. CubeSats in Hydrology: Ultrahigh-Resolution Insights Into Vegetation Dynamics and Terrestrial Evaporation. Water Resources Research, 53(12), 10017-10024.

Planet Team, 2018. 'Planet imagery product specifications.' Planet Labs Inc, Accessed 12/04/2018. https://www.planet.com/products/satellite-imagery/files/Planet_Combined_Imagery_Product_Specs_December2017.pdf.

Sacks, W.J., Deryng, D., Foley, J.A., & Ramankutty, N., 2010. Crop planting dates: An analysis of global patterns. Global Ecology and Biogeography, 19(5), 607-620.

Sadeh, Y., Zhu, X., Chenu, K., & Dunkerley, D., 2019. Sowing date detection at the field scale using CubeSats remote sensing. Computers and Electronics in Agriculture, 157, 568-580.

Urban, D., Guan, K., & Jain, M., 2018. Estimating sowing dates from satellite data over the US Midwest: A comparison of multiple sensors and metrics. Remote Sensing of Environment, 211, 400-412.

Wold, S., Esbensen, K., & Geladi, P., 1987. Principal component analysis. Chemometrics and Intelligent Laboratory Systems, 2(1-3), 37-52.

Cost-effectiveness and performance of optical satellites constellation for Precision Agriculture

M. Sozzi[1*], A. Kayad[1], D. Giora[1], L. Sartori[1,2] and F. Marinello[1,2]
[1]*Department of Land, Environment, Agriculture and Forestry, University of Padova, Viale dell'Università 16, Legnaro (PD), Italy; marco.sozzi@phd.unipd.it*
[2]*NEOS srl, Spin-off of the University of Padova, Piazzetta Modin, 12, Padova (PD), Italy*

Abstract

A prerequisite of precision agriculture is monitoring of field variability, which is usually detected by means of proximal or remote sensors. Satellite images can easily provide information about such variability in a relatively short time. The present work aims to evaluate the main satellite constellations according to precision agriculture requirements, and to estimate their cost effectiveness in relation to expected precision agriculture benefits. Performance of satellite constellations were evaluated in terms of spatial resolution, spectral resolution, radiometric resolution, bandwidth and revisit time. Then, cost effectiveness of satellite constellations was retrieved for each constellation, calculating the break-even point where all costs are completely recovered by benefits. Preliminary results show how satellite image price is mainly correlated with spatial resolution and minimum order area. Satellite constellations with the lowest break-even point are Sentinel and Landsat, even though they represent a compromise of resolution features.

Keywords: satellite image price, satellite performance, within-field variability

Introduction

Within-field variability knowledge is a fundamental concept of precision agriculture methodologies, where soil, crops and environmental variabilities are clustered in management zones according to time and space (Cillis *et al.,* 2018). Over the years, several sensors and respective analysis methodologies have been proposed to detect sub-field variability, mainly using proximal and remote sensors (Dubbini *et al.,* 2017; Anastasiou *et al.,* 2018). Proximal sensor surveys, usually characterised by high measurement accuracy, are time-consuming tasks, even if they can be mounted on tractors, to combine surveys with farming operations. On the other hand, remote sensors have a relatively low accuracy, but they provide a massive amount of information in a short time. The latter characteristic has led to a rapid increase of interest from the precision agriculture sector on satellite remote sensing, especially over the last ten years (Atzberger and Clement, 2013; Georgi *et al.,* 2018). A quantitative analysis of scientific papers in Scopus database for 'precision agriculture' and 'remote sensing' keywords showed how the number of scientific papers rose from 30 in 2008 up to more than 70 in 2018.
In remote sensing, the main features for satellite classification are spatial, spectral radiometric and temporal resolution. The spatial resolution identifies smallest detectable detail in an image. It is affected by ground sample point, and it is also defined as pixel size. Spatial resolution is the primary limitation to apply satellite remote sensing in agriculture. Indeed, several earth observation satellites provide daily images with good spectral resolution, but the spatial resolution is incompatible with the within-field variability detection. Spectral resolution is another important feature of satellite image, especially for agriculture use. According to the healthy vegetation spectral signature, a peak of reflectance is detectable between 750 and 950 nm. This peak and its difference with reflectance in red wavelength have been studied over the years, and several vegetation indices have been proposed (Xue and Su, 2017). Among them, the normalized difference vegetation index (NDVI) is

DOI 10.3920/978-90-8686-888-9_62, © Wageningen Academic Publishers 2019

the most widely used. Thanks to the relationship between healthy vegetation and vegetation indices, numerous applications on these indices have been proposed in agriculture (Thomason *et al.,* 2011). Radiometric resolution affects the dynamic range of satellite images, and it proportionally affects the level of variability detectable from the sensors. Finally, temporal resolution, known as revisit time, affects the number of images available during the season. Precision agriculture needs frequent monitoring of the study area, and it cannot be prejudiced by cloud cover which is a problem for remote sensing of fields.
Several satellite image providers offer satellite images from different constellations and different image resolution. Moreover, collections of historical archives of satellite images are available, which can be used by farmers to delineate temporally stable zones (Crestey *et al.,* 2017). Precision agriculture has specific needs of resolutions which were analysed in order to define a ranking of satellite images ready for agriculture. The present study analysed, in terms of resolutions and costs, performance and affordability of the most common satellite constellations for precision agriculture use.

Materials and methods

Performance and cost analysis was based on a dataset of pricelist and quotations provided by 9 different satellite image companies. From each pricelist were selected all multispectral satellites (spectral resolution from 600 to 800 nm) with a revisit time lower than 20 days. A threshold in pixel size to use satellite images for precision agriculture was set at less than 30×30 m. These three parameters were used as a dataset selection criterion, extracting 60 prices for 16 constellations of satellites.
The selected constellations were:

- *Deimos-2:* A full private satellite developed by Elecnor company (Spain) (www.elecnor-deimos.com);
- *Dove (Planet):* A constellation of nanosatellites provided by Planet Labs Inc (US) (www.planet.com);
- *Formosat-2:* Earth observation satellite developed by the National Space Organization of Republic of China (www.nspo.narl.org);
- *GaoFen-2:* High resolution satellite developed by China National Space Administration (www.cnsa.gov.cn);
- *Kompsat-2/3/3A:* Three satellites constellation launched by the Korea Aerospace Research Institute. All the three satellites have different resolution features (www.kari.re.kr);
- *Landsat-7/8:* Mid resolution satellites developed as a collaboration between NASA and U.S. Geological Survey (landsat.gsfc.nasa.gov);
- *Pleiades-1A/1B:* Constellation of two satellites built by Airbus Defence and Space (www.intelligence-airbusds.com);
- *Rapideye:* A constellation of 5 satellites provided by Planet Labs Inc (US) (www.planet.com);
- *Sentinel-2:* European constellation consists of two optical satellites (sentinel.esa.int);
- *Spot-6/7:* Two identical satellites constellation built by Airbus Defence and Space (www.intelligence-airbusds.com);
- *Superview-1:* Commercial satellites constellation developed by Space View Tech (www.spaceview.com);
- *TripleSat:* A constellation of three satellites built by Surrey Satellite Tech. Ltd (www.sstl.co.uk);
- *WorldView-2/3/4:* Satellites developed by Digital Globe (US) (www.digitalglobe.com);

Several earth observation satellites provide images with high spatial resolution in visible band (panchromatic) and medium resolution in multispectral (even near-infrared) bands. The analysis took into account the panchromatic resolution, considering the possibility to apply the pan-sharpening process. Although the dataset was composed of satellites with a spectral resolution focused on red and near-infrared wavelengths, some satellites provide larger bands from ultraviolet

to thermal infrared. Spectral resolution was evaluated even in terms of bandwidth. A high bandwidth increases the probability to detect a peak in reflectance while, on the other hand, it increases noise. Bandwidth was evaluated as a performance index in terms of the ratio between red and near-infrared bandwidths.

An important parameter of sensors performance is Signal-to-Noise-Ratio (SNR). This value defines the ratio between signal and noise power for a single band. In vegetation index calculation, band math operations have an additive effect on single band SNR. Error propagation in the NDVI formula (Rouse *et al.,* 1973) was calculated using SNR retrieved from satellite manufacturers' websites.

In Table 1, name and resolution features of each constellation are summarized.

After defining the dataset, cost analysis was performed. In satellite images, the main cost items are the price per unit area and download cost. Furthermore, precision agriculture applications based on satellite images are time-consuming tasks, and they require a relatively high computing power. Thus, an additional cost item related to data processing was added to the price analysis.

Price per unit area was determined considering the minimum price between recent acquisition and fresh archive not older than 90 days. Indeed, precision agriculture requires fresh images in order to continuously analyse the crop through the season. In addition, all satellite image providers set a minimum order area (25-100 km^2) which strongly affect the price especially to the final user requiring small areas.

Data processing time for satellite images for precision agriculture was specifically evaluated. According to the Sentinel Data Access Annual Report (2017), 50% of Copernicus users have bandwidth available for the download between 10 and 50 megabits per second (European Space Agency, 2018), which means a download time for a Sentinel-2 product (mid resolution) between 13 and 3 minutes. Then, considering download, georeferencing, atmospheric correction, index calculation, masking and legend-making, the whole analysis takes about 30 minutes in case of coding script analysis and about 60 minutes in case of step-by-step analysis (based on a survey between professional users of remote sensing and geographical information systems). In any case, vegetation indices calculation requires only a band math operation and its processing time is proportional to

Table 1. Resolution features of the selected satellites.

Name	Launched (year)	Spatial resolution (m)	Spectral resolution (nm RED/NIR)	Radiometric resolution (bits/pixel)	Revisit time (dd)
Deimos-2	2014	0.75	640-697/770-892	10	2
Dove	2013	3	590-670/780-860	16	1
Formosat-2	2004	2	630-690/760-900	8	1
GaoFen-2	2014	0.8	630-690/770-890	14	5
GeoEye-1	2008	0.5	655-690/780-920	11	3
Kompsat-2	2006	1	630-690/760-900	14	6
Kompsat-3	2012	0.7	630-690/760-900	14	3
Kompsat-3A	2015	0.55	630-690/760-900	14	3
Landsat-7/8	2013	15	630-680/850-880	12	8
Pleiades-1A/1B	2011	0.5	600-720/750-950	12	1
Rapideye	2008	5	630-685/760-850	12	5.5
Sentinel-2	2015	10	645-685/785-900	16	5
Spot-6/7	2012	1.5	625-695/760-890	12	1
Superview-1	2016	0.5	630-690/770-890	11	4
TripleSat	2015	0.8	600-670/760-910	10	1
WorldView-2/3/4	2009	0.3	655-690/780-920	11	1

the amount of area analysed. Computational demand was estimated with the number of kilobytes per hectare starting from radiometric and spatial resolutions.According to the labour cost per hour, data processing cost was fixed at 100 $ per image (without image interpretation and advising). Download cost per hectare was calculated using the computational demand and the first half 2018 report of Rewheel/research (Rewheel, 2018). According to this report, in OECD countries the median fully allocated gigabyte price was 1.3 $/GB (April 2018). The Rewheel report analysed 4G mobile broadband or stationary (fixed-wireless) plans of 41 countries all over the world.
To calculate the cost-effectiveness of satellite images for precision agriculture, the study took advantage of the metanalysis proposed by Colaço and Bramley (2018) related to the advantages of variable rate nitrogen application in grain crops. In the 80% of papers analysed by Colaço and Bramley (48), the economic advantages due to precision farming ranged from -30 to 70 $/ha. The present study considered the weighted average profit reported in the metanalysis, which is 30 $/ha. Profitability point for each satellite constellation were compared with utilised agricultural area of farms in EU.

Results and discussion

Although the spectral resolution of all selected satellites was focused on red and near-infrared, a huge difference was found in bandwidth. Pleiades constellation is characterized by 120 nm and 200 nm bandwidth in red and near-infrared bands respectively. On the other hand, Landsat has 50 nm bandwidth in red and 30 nm in near-infrared. A wide band increase the sensitivity to reflectance peaks, but it reduces the possibility to discriminate features according to reflectance. Except for Dove and Landsat 7/8, all satellites are characterized by a wider bandwidth in near-infrared than in red. In Dove satellites, this is due to a wide red bandwidth with a lower limit close to the yellow wavelength (590 nm). Differently, in Landsat 7/8, the high bandwidth ratio is due to a narrow near-infrared band. Bandwidth in red and near-infrared is summarized in the bandwidth ratio in Table

Table 2. Price and performance parameters.

	Bandwidth ratio (Red/Nir)	NDVI errors propagation (based on SNR)	Computational demand (kB/ha)	Minimum order area (ha)	Price per hectare ($/ha)
Deimos-2	0.47	1.05%	25.3	2,500	0.060
Dove	1.00	1.83%	8.89	10,000	0.012
Formosat-2	0.43	0.38%	3.13	14,400	0.031
GaoFen-2	0.50	0.30%	34.2	2,500	0.045
GeoEye-1	0.25	0.47%	102	10,000	0.275
Kompsat-2	0.43	0.96%	21.9	2,500	0.055
Kompsat-3	0.43	0.96%	44.6	2,500	0.110
Kompsat-3A	0.43	0.92%	72.3	2,500	0.160
Landsat-7/8	1.67	0.41%	0.203	3,700,000	0
Pleiades-1A/1B	0.60	0.56%	75.0	10,000	0.213
Rapideye	0.59	1.02%	3.00	10,000	0.012
Sentinel-2	0.34	0.59%	0.838	1,200,000	0
Spot-6/7	0.53	0.32%	8.33	10,000	0.045
Superview-1	0.50	N/A	68.8	10,000	0.250
TripleSat	0.47	0.95%	24.4	2,500	0.080
WorldView-2/3/4	0.24	0.74%	179	10,000	0.275

2. Even SNR showed a high variability between satellites. The highest error was found in Dove, Deimos-2 and Rapideye satellites. On the other hand, GaoFen-2, Spot-6/7 and Formosat-2 showed a low error value. In Table 2 is reported the propagation of SNR in NDVI formula. Computational demand showed a similar trend to pixel size, due to the quadratic relation with spatial resolution.
According to the pricelists considered, all satellites require a minimum order area which strongly affects the price of the final product. Minimum order area ranges from 2,500 ha of Kompsat satellites up to 14,400 ha of Formosat-2. In Sentinel-2 and Landsat 7/8, tile extension is considered as minimum order area. In Table 2, the price per hectare is also reported, which, multiplied by minimum order area, returns the price of the minimum purchasable area. Worldview 2-3-4, Geoeye-1, and Pleiades 1A/B provide images with a sub-meter resolution and have a high cost of minimum area. Sensors with a spatial resolution of 10-15 m, like Sentinel-2 and Landsat 7/8 have the lowest price for minimum area, which is influenced only by download cost. Radiometric resolution showed a relatively low variability in the satellite's dataset. Values range from 8 bits/pixel in case of Formosat-2, up to 16 bits/pixel in case of Dove and Sentinel-2. The dynamic range of a satellite's sensors does not affect the overall image quality, but it may be considered to detect variability with a value of NDVI close to the saturation (Gu *et al.*, 2013).
According to the minimum order area, price per hectare, download cost, and price analysis, the price curve for each satellite was calculated. The minimum number of images needed to perform a good crop monitoring was set as 5 images per season.
As shown in Figure 1, there is no slope variation in curves up to 2,500 ha, due to the minimum order area. Then, the breakeven point for each constellation was calculated, taking advantage from

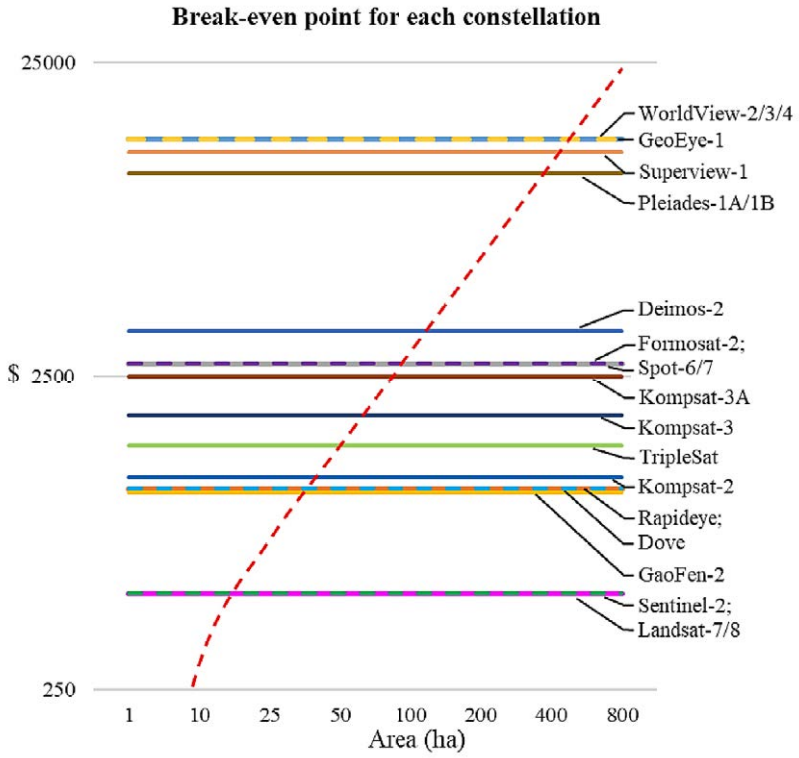


Figure 1. Break even point of satellite images, in ha.

the estimated precision agriculture advantages (30 $/ha). Break-even point represents the point after which the profitability of an investment exceeds the cost.
The lowest profitability point is approximately 17 ha, and it represents free of charge satellites (Landsat 7/8 and Sentinel-2). Starting from 40 ha even Dove, Rapideye, GaoFen-2 and Kompsat-2 become profitable. In this case, the other features should be considered for the choice: GaoFen-2 has a sub-meter resolution with five days of revisit time, Dove provides daily 3 m resolution images, Rapideye implements an additional red-edge band. TripleSat and Kompsat-3 have the break-even point close to 50 and 65 ha respectively. Beak-even point of Kompsat-3A, Formosat-2, Deimos-2 and Spot-6/7 ranges from 85 up to 120 ha and they provide about 1 m resolution images with a short revisit time (1-3 days). Daily sub-meter resolution images are provided by WorldView-2/3/4, GeoEye-1 and Pleiades-1A/B, with a break-even point of about 400-470 ha. Similarly, the break-even point of Superview-1 is approximately 400 ha even if revisit time is 4 days. Despite limited cost of images per hectare, the minimum purchasable area predominantly affects the final cost of satellite images. For agriculture use, the threshold of minimum area is excessively high (the lowest threshold is 2,500 ha), and the proportional price can be reached only by a few farms or associated companies. This is the reason why Kompsat 2, Kompsat-3 and Kompsat-3A, with a required area of 2,500 ha, have a break-even point of 40, 65 and 85 ha respectively.
With a focus on the European situation, the average utilised agricultural area per holdings is 16.1 ha (Eurostat, 2013). In this situation, satellite image use should mainly consider free of charge satellites.

Conclusions

This study analysed the most common satellite constellations used for precision agriculture. Performance and economic benefits resulting from each constellation were evaluated. Results show how each constellation of satellites requires a different area to be economically viable. Analysis of resolutions, revisit time, and minimum cost showed that profitability of satellites images for precision agriculture start at 17 ha, with satellites characterized by a spatial resolution of 10-15 m (Sentinel 2 and Landsat 7/8). With this kind of spatial resolution, field shape should be considered, because the border effect noise could affect data analysis. Farms with an area of 40 ha can profitably exploit Dove, RapidEye or GaoFen-2 images. Sub-meter resolution images are economically viable with more than 40 ha with 5 days of revisit time (GaoFen-2) or with 50 ha with daily revisit time (TripleSat). The possibility to use high-resolution images is important, because a high spatial resolution better detects within-field variability.
According to the analysis, satellites equipped with sub-meter spatial resolution seem to be less exploitable for agricultural activity and this is related to the fact that farms with this order of area are not so common in Europe. Anyway, it is also important to notice that data with a high spatial resolution requires a higher level of technology in field operations rather than data with lower spatial resolution and this is due to the capability of agricultural machinery to vary the rate of distribution of inputs in a small area. Since in Europe the average utilised agricultural area per holdings is 16.1 ha, Landsat-7/8 and Sentinel-2 satellite constellation seem to be a good solution due to their low break-even points.

Acknowledgements

This study was undertaken thanks to data provided by e-GEOS, LAND INFO Worldwide Mapping, Geocento, Satellite Imaging Corporation, Sysdeco Italia s.r.l., Harris geospatial, Apollo Mapping, Satellite Imaging Corporation and Planet Labs Inc (Planet Team (2017). Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. https://api.planet.com).

References

Anastasiou, E., Balafoutis, A., Darra, N., Psiroukis, V., Biniari, A., Xanthopoulos, G., *et al.,* 2018. Satellite and Proximal Sensing to Estimate the Yield and Quality of Table Grapes. Agriculture 8(7) 94.

Atzberger, C. and Clement, 2013. Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. Remote Sensing 5(2) 949-981.

Cillis, D., Pezzuolo, A., Marinello, F. and Sartori, L., 2018. Field-scale electrical resistivity profiling mapping for delineating soil condition in a nitrate vulnerable zone. Applied Soil Ecology 123 780-786.

Colaço, A. F. and Bramley, R. G. V., 2018. Do crop sensors promote improved nitrogen management in grain crops? Field Crops Research 218 126-140.

Crestey, T., Pichon, L. and Tisseyre, B., 2017. Potential of freely available remote sensing visible images to support growers in delineating within field zones. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences 8(2) 372-376.

Dubbini, M., Pezzuolo, A., De Giglio, M., Gattelli, M., Curzio, L., Covi, D., *et al.,* 2017 Last generation instrument for agriculture multispectral data collection. Agricultural Engineering International: CIGR Journal 19(1) 87-93.

European Space Agency, 2018. Sentinel Data Access Annual Report 2017 (1) 96.

Eurostat, 2013. Agriculture, forestry and fishery statistics 2013(1) 24-42.

Georgi, C., Spengler, D., Itzerott, S. and Kleinschmit, B., 2018. Automatic delineation algorithm for site-specific management zones based on satellite remote sensing data. Precision Agriculture 19(4) 684-707.

Gu, Y., Wylie, B. K., Howard, D. M., Phuyal, K. P. and Ji, L., 2013. NDVI saturation adjustment: A new approach for improving cropland performance estimates in the Greater Platte River Basin, USA. Ecological Indicators 30 1-6.

Rewheel, 2018. The state of 4G pricing, April 2018 – Digital Fuel Monitor 9 th release 1H2018.

Rouse, J. W., Hass, R. H., Schell, J. A. and Deering, D. W., 1973. Monitoring vegetation systems in the great plains with ERTS. Third Earth Resources Technology Satellite (ERTS) symposium 1(1) 309-317.

Thomason, W. E., Phillips, S. B., Davis, P. H., Warren, J. G., Alley, M. M. and Reiter, M. S., 2011. Variable nitrogen rate determination from plant spectral reflectance in soft red winter wheat. Precision Agriculture 12(5) 666-681.

Xue, J. and Su, B., 2017. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. Journal of Sensors 1 1-17.

Section 8
Site-specific nutrient, lime and seed management

Crop imaging and soil adjusted variable rate nitrogen application in winter wheat

F. Argento[1,2*], T. Anken[2], F. Liebisch[1] and A. Walter[1]
[1]*Institute of Agricultural Sciences, Crop Science Group, ETH Zürich, Universitätstrasse 2, 8092 Zürich, Switzerland; francesco.argento@usys.ethz.ch*
[2]*Digital Production Group, Agroscope, Tänikon 1, 8356 Ettenhausen, Switzerland*

Abstract

Site-specific nitrogen (N) management in precision agriculture is used to improve nitrogen use efficiency (NUE) at field scale. The aim of this study has been to contribute to improvements in decision support for application of N fertilizer. Specifically, image-based methods were applied and improved to better target the application of N fertilizer via variable rate application (VRA) for small fields, as typical in Swiss agriculture. Multi-spectral images acquired with an unmanned aerial vehicle (UAV) platform and soil N data at reference locations were used to explore the in-field variability of crop in a plot-based field experiment that compared uniform standard (ST) rate to variable rate (VR) application. The VR final yield resulted in the same range as the ST but showed an average 10% reduction in fertilizer use.

Keywords:, nitrogen management, winter wheat, UAV, variable rate application

Introduction

The spatial and temporal distribution of N available for agricultural crops in the field varies greatly according to a multitude of factors including soil properties, climate and management (Kindred *et al.*, 2015). In Swiss agriculture, farm size is considered medium to small scale as average farm extension is around 20 ha (Swiss FSO, 2018). It follows that the monitoring and management of in-field variability presents additional obstacles compared to large-scale precision farming operations (e.g. Farmstar-conseil, France). Nonetheless, N fertilizers are widely used and their application represents both an environmental issue and a cost factor that are being actively tackled. The implementation of a precision farming framework may be beneficial to further improve NUE at medium to small scale and an example for other countries with similar issues.

Aim and hypothesis

The study presented in this paper is part of a project that aims to develop methods for the implementation of site-specific N fertilization in Switzerland. The focus is on main crops such as winter wheat (*T. aestivum*) and maize (*Z. mays*). With the aid of low-altitude remote sensing technology and based on monitoring of various soil parameters (e.g. soil type and nitrate content), the project objectives are to map heterogeneity in fields and to provide data-based recommendations for variable rate N fertilization in fields of the size of a few ha. At the same time the economic and environmental costs and benefits are evaluated. This first experiment was designed to test the functionality of the sensors and materials involved, to validate principles from in the literature (Gabriel *et al.*, 2017; Heege, 2013; Nawar *et al.*, 2017) and to get first insights in to the practical implementation. The main hypothesis for this preliminary study was that the implementation of site-specific N management via VRA would reduce average N application compared to the standard fertilization strategy without affecting yield and thus increasing NUE.

DOI 10.3920/978-90-8686-888-9_63, © Wageningen Academic Publishers 2019

Material and methods

Experimental site and design

The experimental site is located at the Agroscope research centre in Tänikon, Canton Thurgau, Switzerland (47.4790021°N, 8.9059287°E) and managed by the Swiss Future Farm. The climate in this region is characterized by 1,170 mm annual rainfall and an average annual temperature of 8.6 °C recorded by a local weather station (data range 1970-2017) from the federal office of meteorology (MeteoSwiss). The field covered an area of 2.5 ha on which winter wheat (cultivar Arnold) was sown on the 19th of October 2017. The soil in this field is characterized as a Gleysol based on a soil survey map provided by the national soil monitoring network (NABO). The soil type was not homogenous across the field, in fact the central part was characterized by a more stagnating Gleysol.

A randomized block design with three treatments replicated in six blocks (n=6) covering 1.35 ha of the field was chosen to study in-field variability. The dimension of a single plot was 15×50 m, to match the operational range of the pneumatic fertilizer spreader (Rauch, Sinzheim, Germany) used. N fertilizer was applied in the form of mineral ammonium nitrate (24% N content), divided in three split applications (Table 1). The fertilization strategy is based on the mineral N (N_{min}) method provided in the 'Principles of fertilization of agricultural crops in Switzerland' (PRIF) (Sinaj and Richner, 2017). The treatments consisted of a standard (ST) and of a variable rate (VR) treatment, for which each plot was managed individually by evaluating the initial N_{min} (e.g. Plot 1 with N_{min}=30 kg N/ha received 80-30 = 50 kg N/ha fertilizer in the first application) and the in-season spectral indices (e.g. Plot 10 with mean index value of 0.25, 25% higher than the field average received 25% less fertilizer i.e. 40 kg N/ha in the second application). Finally, a non-fertilized (NF) treatment was used as a measure of soil N supply.

Soil and plant analysis

To obtain a measure of N_{min} available in the soil at the end of the winter season, mixed samples of 6-8 soil cores from each plot were collected on 27th February 2018 at three different depths 0-30 cm, 30-60 cm and 60-90 cm and the concentration of NH_4 and NO_3 was measured. The mean 0-90 cm N_{min} per plot were used as reference for the calculation of the first split application to apply as shown in Table 1.

Biomass samples were collected at BBCH 32 and few days before harvest (8th July 2018) in two locations per plot. The BBCH 32 samples were dried at 60 °C for 48 h to quantify dry matter and afterwards N concentration was measured in the laboratory. The samples at harvest were air dried for two weeks and threshed to separate the grain from the straw which together make up the plant dry weight. N concentration was measured in both the milled grains and straw material and then used to calculate N uptake (dry matter × N concentration).

Table 1. Fertilization strategy adapted from Swiss reference (Sinaj and Richner, 2017) and Levy and Brabant (2016). N_{min} and rates of N fertilizer per treatment in kg N/ha are reported per growth stage in BBCH decimal units (Meier *et al.*, 2009). For VR, the range is shown.

	N_{min} BBCH 20	Split 1 BBCH 23	Split 2 BBCH 32	Split 3 BBCH 45	Total
Reference	–	80 – N_{min}	60	20	160 – N_{min}
ST	44	36	60	20	116
VR	30 – 85	0 – 50	40 – 70	0 – 20	50 – 132
NF	43	–	–	–	0

Acquisition of aerial data

The low-altitude remote sensing UAV platform was composed of a multispectral camera (Parrot Sequoia, Paris, France) mounted on a Quadcopter (DJI P4P, Shenzhen, China) and flown over the field on a weekly to bi-weekly basis. Spectral information of the growing crop was recorded in four bands of the light spectrum, namely Green (G) centred at 550 nm, Red (R) at 660 nm, Red Edge (RE) at 735 nm and Near-Infrared (NIR) at 790 nm. The bandwidth was 40 nm for G, R and NIR and 10 nm for the RE. The output of the camera consisted of one separate image for each band. The images were captured while the camera was being flown over the field in a single grid automatic flight plan (Pix4D Capture, Lausanne, Switzerland) by the drone at 5 m/s speed and 1.9 s time interval at 50 m height. The ground sampling distance (GSD) of the single band images was 4.5 cm/pixel. The raw images were processed in the photogrammetry software Pix4D Mapper: first a radiometric correction was applied by using a Airinov reflectance target, then they were stitched together to create a reflectance map and finally they were georeferenced with ground control points (GCP) before generating final index maps of the field. The spectral index selected for this study, the Normalized Difference Red Edge index (NDRE = (NIR – RE) / (NIR + RE)), is known for its sensitivity to chlorophyll and N content of the crop (Basso *et al.*, 2016). The data were extracted per each plot with the zonal statistics package in the GIS software QGIS (2.18) and further analysed with RStudio (R version 3.4.1).

Results

Spectral measurements

There was a high correlation (R^2=0.53) between N uptake, analysed from mid-season destructive samples and NDRE measurements (Figure 1). This result was a good indication of the possibility of using the selected spectral index to make a qualitative assessment of the crop's N status. The changes in NDRE through the growing season showed an interesting temporal pattern (Figure 2) and increases in the index curve corresponded with the time points of the three split fertilizer applications. Differences between the NF treatment and the VR/ST became significant after the second split N-application, which also coincided with the end of the prolonged drought period in this region in 2018. After flowering and beginning of senescence, the NDRE reflectance signal changed and started to decrease.

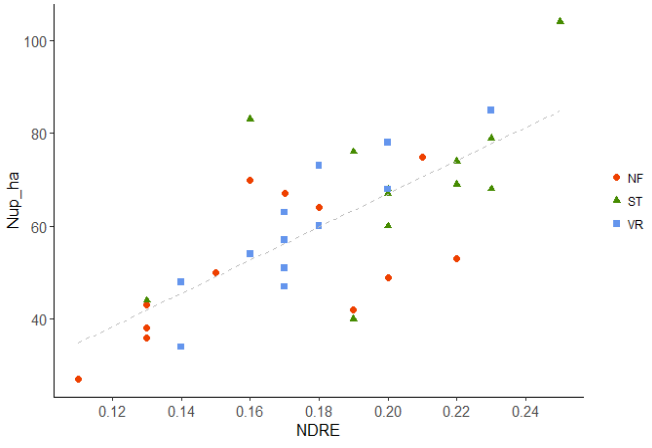


Figure 1. Linear regression between the NDRE and the N uptake at BBCH 32. Samples were collected from the same area of 1 m^2, twice per each plot.

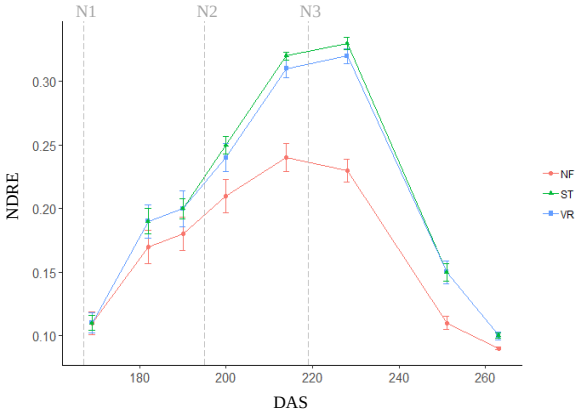


Figure 2. Seasonal patterns of NDRE measurements in 2018. Each measurement point represents average values per treatment with standard error (n=6). Time is represented on the x axis as days after sowing (DAS). N1, N2 and N3 denotes the first, second and third split application, respectively.

Grain yield

The grain yield obtained via manual sampling (Figure 3) showed no significant difference between the VR treatment and the ST treatment (6.34±0.27 t/ha and 6.23±0.23 t/ha, respectively) while both were above the yield achieved by non-fertilized plots (3.81±0.21 t/ha). Thus, the changes in management did not affect the yield of the VR treatment, in which the N fertilizer was redistributed and on average was reduced by 11 kg N/ha (Table 2) compared to the ST treatment.

Nitrogen use efficiency

The apparent fertilizer recovery (AFR) (Kindred *et al.*, 2015) as reported in Equation 1 was used as a measure of the quantity of N fertilizer recovered by the crop. In this equation, the value of N uptake from the NF plot (as a quantification of the soil N supply) was subtracted from the N uptake (where x = ST or VR) of a treatment and divided for the N fertilizer applied.

$$\text{AFR} = \frac{\text{N uptake (x) - N uptake NF}}{\text{N applied}} \quad (1)$$

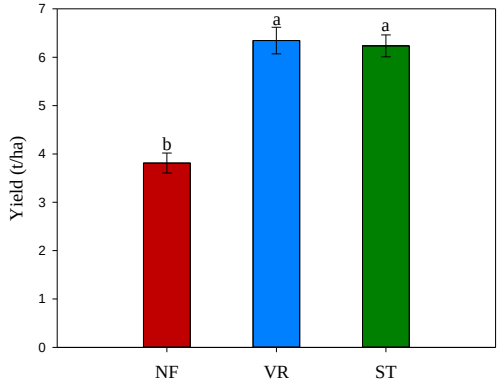


Figure 3. Grain yield of winter wheat at harvest in t/ha. Bars represent the average values per treatment with standard error (n=6). Letters denote the level of significance.

Table 2. N balance including N applied with the fertilizer and N uptake from the crop at harvest. Letters denote the level of significance.

	N application (kg N/ha)	N uptake grain (kg N/ha)	N uptake straw (kg N/ha)	N uptake total (kg N/ha)	AFRg %	AFR %
NF	-	68b	15b	84b	-	-
VR	105a	144a	35a	179a	72.3a	93.8a
ST	116a	146a	32a	179a	67.2a	83.1a

Total N uptake in both the VR and ST treatments resulted in an uptake of 179 kg N/ha (Table 2). The reduction in average applied N seemed to indicate a total higher use efficiency (AFR) in the VR (93.8%) compared to ST (83.1%). The difference was however not significant (P=0.26). The same pattern was true for the grain only (AFRg) with 72.3% for VR and 67.2% for ST. The average soil N supply estimated from the non-fertilized plot was 84 kg N/ha.

Discussion

Spectral indices can be reliable to follow vegetation development but a robust quantification needs further calibration data or even reference plots such as N rich strips and/or a non-fertilized plot (Raun *et al.*, 2005; Samborski *et al.*, 2009). Following the temporal evolution of NDRE differences across the six VR plots was the base of the rationale to adjust the fertilization strategy in-season in the different plots which seemed to result in a higher NUE due to a reduction in average N applied. This is in line with several experiments testing remote sensing based variable rate fertilization in wheat as reviewed by Diacono *et al.* (2013). Compared to other studies, in this experiment, tools such as management zones and an N rich strip were not directly delineated and used for the decision because, as mentioned in the beginning, the main aim of this preliminary study was to test the drone camera system and to gain first-hand field experience. The advantage of such tools is that in theory they provide a more objective means of comparison for the remote sensing technology and they facilitate the direct conversion to N recommendations via algorithms (Franzen *et al.*, 2016; Holland and Schepers, 2010). Furthermore, in this study, the fertilizer was reduced (up to 50%) in single plots showing better growth during the season and increased (up to 15%) in the plots that showed reduced growth. The benefits of this approach are that fertilizer is redistributed across the field in greater quantity to the areas in which plant growth stimulation is wanted and in reduced quantities to areas showing a prominent SNS. A possible negative consequence could be that in areas in which plant density or biomass is lower due to the combination of factors other than N (e.g. water limitation or compaction), the supplementary N applied could not be properly used (Ebertseder *et al.*, 2003) and could increase the environmental risk. Excessive N could also result in lodging. Another common approach is therefore to adapt fertilizer quantity to the yield potential of a certain plot or management zone (Bushong *et al.*, 2016). In this case, the benefit comes from targeting areas of high potential where more fertilizer will be provided to allow the crop to reach its full potential. The last winter wheat growing season (2017/2018) was characterized by an intense and prolonged drought, especially in the month of April and beginning of May with only 14 ml summed precipitation during this time compared to the average 90 ml (source: MeteoSwiss). Normally, yield expectation for winter wheat in this region is around 7 t/ha of grains. As expected, the drought affected the final yield and probably also the N supply in this period. Despite this drought, the yield achieved was still in a good range and was comparable to that of other winter wheat fields in the same area. The tendency of the Gleysol to retain more water and the soil heterogeneity in the field played a distinct

role in this year. The central area of the field characterized by a more organic soil was visible from early in the season and in the UAV images (Figure 4) where the vegetation developed better than in the surrounding plots.

Conclusions

The experiment described in this paper was a first experience into the implementation of site-specific N management in Switzerland supported by low-altitude remote sensing and information about soil heterogeneity. The knowledge gained was important to better understand the relation between the digital data and the actual processes in the field through the growing season. It provided positive indication that by managing in-field variability, even on a small scale, it is possible to influence the efficiency of N uptake from the crop, thus reducing the residual N that can accumulate after the season and the risk of loss by emissions. The environmental conditions, especially the soil and the dry weather played a crucial role for N management. Deciding on the most appropriate site-specific fertilisation strategy is crucial to improve the efficiency of nitrogen uptake in wheat. In the future studies of this project, the delineation of management zones and the estimation of zone-specific yield potential will be taken into account as factors to develop the basic strategy. UAV images provided a base for supporting the decision regarding the amount and spatial distribution of the second and third application of N fertilizer. This can be further improved by the use of data driven algorithms but it requires ultimately an expert evaluation before the information is transferred to the variable rate machinery. Potential improvements for the in-season management may be made through the integration of data from the continuous monitoring of the soil-nitrate content.

Acknowledgements

The authors would like to thank the Swiss federal office of agriculture (FOAG) for funding the project, the technical staff at Agroscope Tänikon and ETH Zürich, the Swiss Future Farm for the support with the field trial management as well as the members of the ETH Phenofly team for helping with the set-up of the UAV system.



Figure 4. NDRE map produced from images taken on 6th April 2018 (BBCH 24) with the UAV system. The variability is visible across the trail design.

References

Basso, B., Fiorentino, C., Cammarano, D. and Schulthess, U., 2016. Variable rate nitrogen fertilizer response in wheat using remote sensing. Precision Agriculture, 17: 168-182.

Bushong, J. T., Mullock, J. L., Miller, E. C., Raun, W. R., Klatt, A. R. and Arnall, D. B., 2016. Development of an in-season estimate of yield potential utilizing optical crop sensors and soil moisture data for winter wheat. Precision Agriculture, 17: 451-469.

Diacono, M., Rubino, P., & Montemurro, F., 2013. Precision nitrogen management of wheat. A review. Agronomy for Sustainable Development, 33: 219-241.

Ebertseder, T., Gutser, R., Hege, U., Brandhuber, R., & Schmidhalter, U., 2003. Strategies for site-specific nitrogen fertilization with respect to long-term environmental demands. In: J V Stafford, A Werner (Eds.) Precision Agriculture '03. Proceedings of the 4th European Conference on Precision Agriculture. Wageningen, the Netherlands: Wageningen Academic Publishers. pp. 193-198.

Franzen, D., Kitchen, N., Holland, K., Schepers, J., & Raun, W., 2016. Algorithms for in-season nutrient management in cereals. Agronomy Journal, 108: 1775-1781.

Gabriel, J. L., Zarco-Tejada, P. J., López-Herrera, P. J., Pérez-Martín, E., Alonso-Ayuso, M., & Quemada, M., 2017. Airborne and ground level sensors for monitoring nitrogen status in a maize crop. Biosystems Engineering, 160: 124-133.

Heege, H. J., 2013. Site-Specific Fertilizing. In: Precision in Crop Farming. Dordrecht, the Netherlands: Springer. pp. 193-271.

Holland, K. H., & Schepers, J. S., 2010. Derivation of a variable rate nitrogen application model for in-season fertilization of corn. Agronomy Journal, 102: 1415-1424.

https://www.bfs.admin.ch/bfs/en/home/statistics/agriculture-forestry/farming.assetdetail.5287765.html (last accessed 22/01/2019)

Kindred, D. R., Milne, A. E., Webster, R., Marchant, B. P., & Sylvester-Bradley, R., 2015. Exploring the spatial variation in the fertilizer-nitrogen requirement of wheat within fields. Journal of Agricultural Science, 153: 25-41.

Levy, L., & Brabant, C., 2016. Die Kunst, den Stickstoffdünger für einen optimalen Ertrag und Proteingehalt von Weizen aufzuteilen (The art of splitting nitrogen applications to optimise wheat yield and protein content). Agrarforschung Schweiz, 7: 80-87.

Meier, U., Bleiholder, H., Buhr, L., Feller, C., Hack, H., Heβ, M., *et al.,* 2009. The BBCH system to coding the phenological growth stages of plants-history and publications. Journal Für Kulturpflanzen, 61: 41-52.

Nawar, S., Corstanje, R., Halcro, G., Mulla, D., & Mouazen, A. M., 2017. Delineation of Soil Management Zones for Variable-Rate Fertilization: A Review. Advances in Agronomy, 143: 175-245.

Raun, W. R., Solie, J. B., Stone, M. L., Martin, K. L., Freeman, K. W., Mullen, R. W., *et al.,* 2005. Optical sensor-based algorithm for crop nitrogen fertilization. Communications in Soil Science and Plant Analysis, 36: 2759-2781.

Samborski, S. M., Tremblay, N., & Fallon, E., 2009. Strategies to make use of plant sensors-based diagnostic information for nitrogen recommendations. Agronomy Journal, 101: 800-816.

Sinaj S. and Richner W., 2017. Principles of fertilization of agricultural crops in Switzerland (PRIF 2017). Agrarforschung Schweiz, 8 (6), Special publication

Swiss Federal Statistical Office FSO, 2018. Agriculture and Food: pocket statistics 2018.

Using yield monitor data to guide precision nitrogen application

B.W. Brorsen[1*], E. Park[2] and X. Li[2]
[1]*Dept. of Agricultural Economics, Oklahoma State University, Stillwater, OK, USA; wade.brorsen@okstate.edu*
[2]*Dept. of Agricultural Economics, Mississippi State University, Starkville, MS, USA*

Abstract

The question addressed is how to best use yield monitor data to make variable rate nitrogen recommendations. A spatial plateau model is estimated using Bayesian Kriging and yield monitor data, which gives an estimate of the yield if nitrogen were not limiting. The model is then used to make nitrogen recommendations. On the one field considered, the variation in nitrogen recommendations was modest and so the economic advantage relative to uniform rate application is a few dollars per hectare, but the method has the potential for the costs to be small since it uses data that is already being collected.

Keywords: Bayesian Kriging, plateau functions, variable rate nitrogen, yield monitor data

Introduction

Agricultural producers and crop consultants are being inundated with data such as yield maps based on yield monitor data. Advanced precision application technology is available, but there is still no clear consensus on how to decide how much nitrogen to apply in each location. This article seeks to use the available yield monitor data to make variable rate nitrogen recommendations.
First, a stochastic plateau model (Tembo *et al.,* 2008) is estimated for each grid in the field. Only the plateau and its variance can be estimated. The other parameters have to come from prior information. Usually only a few years of yield monitor data are available and so estimates would be imprecise. To correct for this imprecision, Bayesian Kriging is used to smooth the parameter estimates across space. Previous research on precision nitrogen application has focused primarily on plant sensing with both hand-held sensors and satellite images (Raun *et al.,* 2002; Muñoz-Huerta, 2013). The approach developed here would have the potential to be combined with information from plant sensing.
Perhaps the most important place yield monitor data could be used is to make precision nitrogen fertilizer recommendations. Variable rates can be applied at low cost, but the problem of how best to determine the rate applied has not been solved. A new method of using yield monitor data to estimate the key parameters of the production function is proposed.

Materials and methods

The historical yield monitor data will typically be limited to a few years. Bayesian Kriging gives a way to essentially get more observations by using a weighting function of nearby observations. The method used by Park *et al.* (2018) is modified to estimate a stochastic plateau function where the parameters vary across space. Currently, the spatial correlation only considers distance, but it could be modified to add soil type and slope.
Past research has considered assessing growing plants (Franzen *et al.,* 2016), nitrogen soil sampling (Odberg, 2016), and management zones (Peralta *et al.,* 2015) as a means of determining nitrogen rates. The Pacific Northwest area considered by Peralta *et al.,* shows enough spatial variability that variable rate application looks promising. But, the economics of variable rate application in areas with less spatial variability like the field considered here have not shown as much economic potential for variable rate application (Bullock *et al.,* 2002; Boyer *et al.,* 2011).

DOI 10.3920/978-90-8686-888-9_64, © Wageningen Academic Publishers 2019

Bayesian Kriging differs from the spatial error model used in studies such as Anselin *et al.* (2004), Lambert *et al.* (2004), Hurley *et al.* (2004), Liu *et al.* (2006), and Lambert *et al.* (2006). These models assume the spatial correlation is in the error term and capture spatial heterogeneity in yield response using sub-block dummy variables. In contrast, in the Bayesian Kriging model the parameters are stochastic and the spatial autocorrelation is reflected in the stochastic parameters.

The stochastic plateau function of Tembo *et al.* (2008) which is used in a variety of research contexts (Harmon *et al.,* 2016, 2017; Ouedrago and Brorsen, 2018) is modified to be site-specific:

$$y_{it} = \min(\alpha + \beta N_{it}, P_i + \varepsilon_i) + \varepsilon_{it} \tag{1}$$

where x_{it} and y_{it} are the level of the nitrogen input and response yield on th plot at year t,i = 1,…,n and t = 1,…,T, the plateaus P_i are assumed to be spatially correlated by distance and multivariate normally distributed, $P \sim MVN(\bar{P}, \Sigma_p)$, where $P = [P_1,\ldots, P_N]'$, $\Sigma_P = \rho_p e^{-D_{ij}/\theta_P}$, ε_i is a spatially correlated plateau error (plateau variance), $\varepsilon \sim MVN(0, \Sigma_\varepsilon)$, where $\varepsilon = [\varepsilon_1,\ldots,\varepsilon_N]'$, $\Sigma_\varepsilon = \rho_\varepsilon e^{-D_{ij}/\theta_\varepsilon}$ and ε_{it} is an independently identically distributed error term $\varepsilon_t \sim MVN(0, \sigma^2 1_{NT})$, where $\varepsilon_t = [\varepsilon_{1t},\ldots,\varepsilon_{Nt}]$. Since the whole field receives the same amount of nitrogen, the only parameters that can be estimated are plateau on each site (P), the variance of the plateau on each site, and the spatial covariance parameters. The parameters α and β cannot be estimated. Therefore, for estimation the model is parametrized to consider the producer's target yield $\mu = \alpha + \beta_{Nit}$. In general, farmers in Mississippi set a target level 120% higher than the historical average they observed. Therefore, the deterministic target yield was set at 120% of historical average yield 16.3 Mg/ha (13.6 Mg/ha × 120%).

For a detailed explanation of Bayesian Kriging see Park *et al.* (2019). With Bayesian Kriging, parameters are stochastic and spatially correlated. The procedure results in estimating a different plateau and plateau variance for each grid with the estimates being smoothed across space. It is a hierarchical Bayesian procedure. The first layer is the normal likelihood as expressed in (1). The second layer is the Gaussian spatial process for the plateau and plateau variance. The final layer is the prior distributions. Since the data contribute nothing toward estimating the parameters α and β, their posterior distribution is the same as the prior distribution.

The goal is to eventually have a fully Bayesian solution for the optimal level of nitrogen as in McFadden, *et al.* (2018), but currently the plug-in solution of Tembo *et al.* (2008) is used where optimal N (N_i^*) is:

$$\boldsymbol{N_i^*} = \frac{1}{\beta}\left(P_i + \rho_{\epsilon i}\Phi^{-1}\left(1 - \frac{r}{p*\beta}\right) - \alpha\right) \tag{2}$$

where $\Phi^{-1} = (1-(r/p\beta))$ is the inverse of the standard normal cumulative distribution function (assuming $\beta \geq r/p$ since otherwise zero N would be optimal), r is the price of nitrogen, p is the output price, $\sigma_{\varepsilon i}$ is the plateau variance for the ith plot. A 2016 average nitrogen price and a 2016 average corn futures price were used as r and p. As mentioned, the parameters α and β cannot be estimated. Therefore, β was calculated from Mississippi corn nitrogen recommendation from the Mississippi State University Extension Service (0.041 Mg of corn/kg of nitrogen). The intercept α cannot be directly estimated, so it is calculated using the target yield μ (16.3 Mg/ha), and slope parameter β (0.041 Mg of corn/kg of nitrogen), and the actual target nitrogen rate N (242.38 kg/ha) in the dataset. The obtained α is 6.26 Mg/ha.

The empirical application uses data collected from corn production fields of about 160 ha in Mississippi, U.S.A. (90.26718 W 33.17676 N). Four years of production history (2012, 2014, 2015, and 2016) are covered. Consecutive corn was planted in those fields from 2011 to 2016. However, only four years have complete input data records. The major soil types are Dubbs silt loam and Dundee silt loam, with around 0 to 2% slopes. A small portion of the land area is covered by Dundee silty clay loam. Corn was harvested in early August of each year. The yield monitor was installed on a

John Deere corn harvester with six sensors in a row mounted on the header of the harvester. The distance between those sensors was about 2 m. The sensors recorded the instantaneous yield and moisture data at 1 second intervals during harvesting. The length of the intervals depends on the speed of the driving. At an average speed of about 7 km/h, the length is about 2 m. Therefore the original yield data are geo-referenced points about 2 m apart.

Yield data were obtained from a high-resolution yield monitor and were aggregated to 100×100 m grids. The relatively large grids are used to (1) reduce the computational burden for estimation and (2) make it feasible to use a bulk applicator. The data were cleaned before aggregating as in Li and Coble (2017). The value for each grid cell is the mean of the original point readings falling within the 100×100 m grid cell. The border areas of the field usually have some abnormal values in the original point readings. Those are likely due to the turning and speed changing of farming machines during operation. Since it is very cumbersome to correct for those data errors, all points within 20 m of the border line were discarded. Certain portions of the field had missing data in some years. Those portions were discarded as well. In the end, a balanced panel was obtained (see Figure 1 for the locations of the final grids).

The priors used are proper priors. To get the proper priors, geographically weighted regression was used to get an approximate estimate of the parameters. Metropolis-Hastings (MH) steps within a Gibbs sampler is used to update the posteriors of the model as in Park *et al.* (2019). In the process, Markov Chain Monte Carlo (MCMC) chains were run for 50,000 iterations and burn-in for 10,000 observations. Every grid has its own parameters, and the parameters are spatially smoothed. The estimation is written in R (R Core Team, 2018).

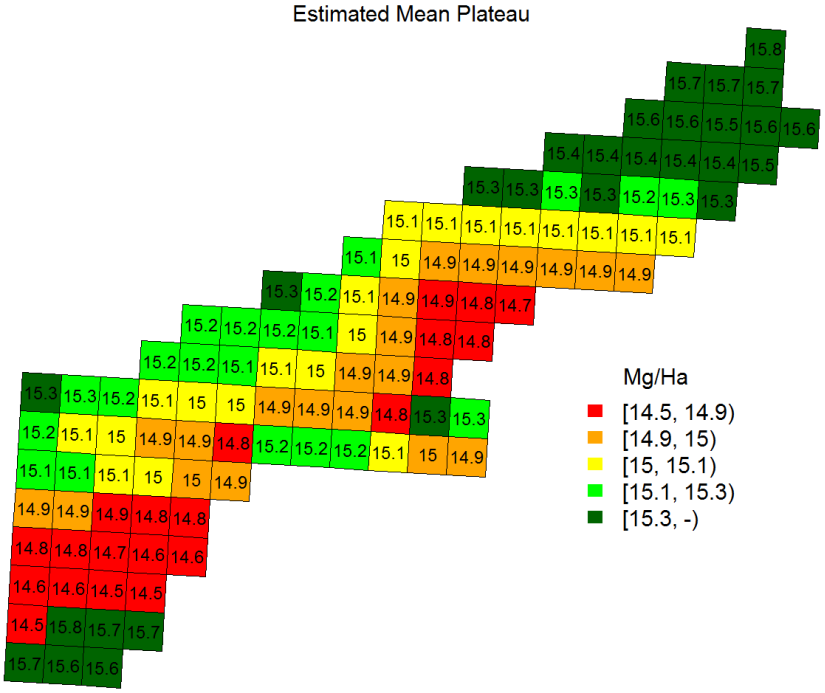


Figure 1. Estimated plateau corn yields (Mg/ha) at Mississippi farm from Bayesian Kriging with 100×100 m grids.

Results and discussion

Table 1 presents a summary of the estimated coefficients. All of the estimated plateaus are below the producer's target yield (μ). Therefore, there is limited ability to increase yield in this example, but there is potential for some small savings on nitrogen. The economic advantage would be higher if there had been more difference between the highest and lowest yield plateaus. Figure 1 shows these plateau levels across the field. There are clear areas of the field that are higher yielding than others. Figure 2 shows the optimal levels of N fertilizer across the field. All of the estimated optimal N recommendations are lower than the producer's target N rate of 242.38 kg/ha. More nitrogen is applied to areas of the field that have more yield potential. The economic advantage of the variable rate nitrogen application is a few dollars per hectare. Is this enough to justify adoption? The procedure could potentially be automated. Yield monitor data are already being collected. The current process of cleaning the yield monitor data is tedious and the cleaning of yield monitor data will need to be automated. The technology to apply variable rate nitrogen already exists. In order to be adopted, the recommendations may need to be made more precise by using information about soil type or weather. Going to smaller grids might also increase the return. The appeal of the approach is that it is possible that costs could be driven low enough to make this feasible even with small values of information.

The technology to use yield monitor data to guide variable rate technology already exists. The engineering technology is ahead of the agronomic and economic research (Bullock *et al.,* 2009; Bullock and Mineo, 2017). This paper is an effort to help catch up. A commercial firm might want to be willing to accept less costly, but less precise, estimation methods. Hamiltonian Monte Carlo methods such as those used in the software package RStan (Stan Development Team, 2018) might prove faster than the Markov Chain Monte Carlo methods used here. Parallel processing and writing the estimation in C++ are also potential ways to speed up the estimation. Commercial firms are seeking ways to exploit the technology that is available. The approach developed here is an important step in that direction.

Conclusions

The general approach used here to guide site-specific nitrogen application does work. Bayesian Kriging works even when the number of observations of each grid is only four years. The approach assumes a linear stochastic plateau production function. The yield monitor data are used to estimate the plateau for each point in the field. Since most historical data uses a uniform rate, the other

Table 1. Summary of coefficients of spatial plateau model for corn yield (Mg/ha) in Mississippi, USA.[1]

	Coefficients[1]			
	μ	*P*	σ^2_ξ	σ^2
Average	16.3	15.1	1.1	0.1
Max	16.3	15.8	1.4	0.1
Min	16.3	14.5	1.0	0.1
Site specific	No	Yes	Yes	No

[1] The coefficients are producer's target yield (μ), the average plateau yield (P), the plateau variance (σ^2_ξ), and the error variance (σ^2).

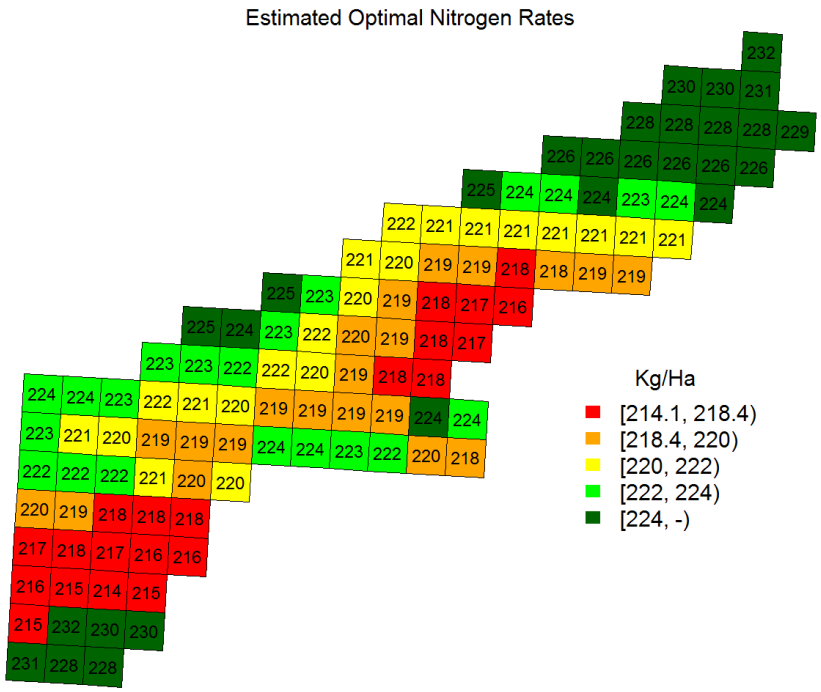


Figure 2. Spatial nitrogen recommendations based on stochastic plateau estimated with Bayesian Kriging.

parameters of the production function must come from prior information. Bayesian Kriging is used to smooth the estimates across space (future versions can also consider similarity in soil type). Recommendations can then be made using Bayesian decision theory, although the current paper uses a plug-in method. The Bayesian approach could be used to combine the information with other types of information such as plant sensing data. It is currently a computationally intensive method and so further research is needed to remove this limitation. The results look quite promising, but there is still much more to be done.

Acknowledgements

The research was funded by the A.J. & Susan Jacques Chair, the Mississippi Agricultural Experiment Station, and the Oklahoma Agricultural Experiment Station and USDA National Institute of Food and Agriculture, Hatch Project number OKL02939.

References

Anselin, L., Bongiovanni, R., and Lowenberg-DeBoer, J., 2004. A spatial econometric approach to the economics of site-specific nitrogen management in corn production. American Journal of Agricultural Economics 86(3) 671-683.

Boyer, C.N., Brorsen, B.W., and Raun, W.R., 2011. Profitability of variable rate nitrogen application in wheat production. Precision Agriculture 12(4) 473-487.

Bullock, D. and Mineo, T., 2017. An assessment of the value of information from on-farm field trials. University of Illinois Urbana-Champaign. https://agecon.unl.edu/seminars/bullock-mieno-on-farm-trials.pdf. (last accessed 1/10/2019).

Bullock, D.S., Lowenberg-DeBoer, and Swinton, S.M., 2002. Adding value to spatially managed inputs by understanding site-specific yield response. Agricultural Economics 27(3) 233-245.

Bullock, D.S., Ruffo, M.L., Bullock, D.G., and Bollero, G.A., 2009. The value of variable rate technology: An information-theoretic approach. American Journal of Agricultural Economics 91:209-223.

Franzen, D., Kitchen, N., Holland, K., Schepers, J., and Raun, W., 2016. Algorithms for in-season nutrient management in cereals. Agronomy Journal 108(5) 1775-1781.

Harmon, X., Boyer, C.N., Lambert, D.M., Larson, J.A., and Gwathmey, C.O., 2016. Comparing the value of soil test information using deterministic and stochastic yield response plateau functions. Journal of Agricultural and Resource Economics. 41(2) 307-323.

Harmon, X., Boyer, C.N., Lambert, D.M., and Larson, J.A., 2017. Temporal frequency of soil test information effects on returns to potassium fertilization in cotton production. Journal of Agricultural and Applied Economics 49(2) 251-272.

Hurley, T.M., G.L. Malzer, and B. Killian, 2004. Estimating site-specific nitrogen crop response functions. Agronomy Journal 96(5) 1331-1343.

Lambert, D.M., Lowenberg-DeBoer, J., and Bongiovanni, R., 2004. Comparison of four spatial regression models for yield monitor data: a case study from Argentina. Precision Agriculture 5 579-600.

Lambert, D.M., Lowenberg-DeBoer, J., and Malzer, G.L., 2006. Economic analysis of spatial-temporal patterns in corn and soybean response to nitrogen and phosphorus. Agronomy Journal 98(1) 43-54.

Li, X., and Coble, K.H., 2017. A spatial econometric analysis of within-field crop yield variability using yield monitor data from a Mississippi farm. Southern Agricultural Economics Association. http://ageconsearch.umn.edu/record/252824/files/yield%20variability_SAEA.pdf (last accessed 11-26-2018).

Liu, Y., Swinton, S.M., and Miller, N.R., 2006. Is Site-Specific Yield Response Consistent over Time? Does It Pay? American Journal of Agricultural Economics 88(2) 471-483.

McFadden, B.R., Brorsen, B.W., and Raun, W.R., 2018. Nitrogen fertilizer recommendations based on plant sensing and Bayesian updating. Precision Agriculture 19(1) 79-92.

Muñoz-Huerta, R.F., Guevara-Gonzales, R.G., Contreras-Medina, L.M., Torres-Pacheco, I., Prado-Olivarez, J., and R.V. Ocampo-Velazquez., 2013. A review of methods for sensing the nitrogen status in plants: Advantages, disadvantages and recent advances. Sensors 13(8) 10823-10843.

Odberg, E., 2016. Variable rate nitrogen. Pacific Northwest Extension Series No. 691.

Oldham, L., 2015. Nutrient Management Guidelines for Agronomic Crops Grown in Mississippi. Mississippi State University Extension Service Publication 2647.

Ouedraogo, F., and Brorsen, B.W., 2018. Hierarchical Bayesian estimation of a stochastic plateau response function: Determining optimal levels of nitrogen fertilization. Canadian Journal of Agricultural Economics 66(1) 87-102.

Park, E., Brorsen, B.W., and Harri, A., 2019. Using Bayesian Kriging for spatial smoothing in crop insurance rating. American Journal of Agricultural Economics 101(1) 330-351.

Peralta, N.R., Costa, J.L., Balzani, M., Franco, M.C., Córdoba, M., and D. Bullock, 2015. Delineation of management zones to improve nitrogen management of wheat. Computers and Electronics in Agriculture 110 103-113.

Raun, W.R., Solie, J.B, Johnson, G.V., Stone,M.L., Mullen, R.W., Freeman, K.W., *et al.,* 2002. Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. Agronomy Journal 944(4) 815-820.

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org.

Stan Development Team, 2018. RStan: the R interface to Stan. R package version 2.18.2. http://mc-stan.org.

Tembo, G., Brorsen, B.W., Epplin, F.M., and Tostão, E., 2008. Crop input response functions with stochastic plateaus. American Journal of Agricultural Economics 90(2) 424-434.

Integrated approach for site-specific nitrogen management in North Dakota, USA

D. Franzen
North Dakota State University, Dept. 7186, P.O. Box 6050, Fargo, 58108 ND, USA;
david.franzen@ndsu.edu

Abstract

Nitrogen (N) management is important for farmers in North Dakota to maximize economic return and environmental stewardship. Fall soil sampling for soil nitrate is important to N management. Zone soil sampling is highly recommended over grid soil sampling. Zone development tools include remote imagery, soil electrical and magnetic sensors, multi-year yield maps and topography. Nutrient rate research to support site-specific fertilizer application shows that N rate and yield are not related between fields. Therefore, N rate is not yield-goal based. Zone sampling does not aid in-season N needs. Algorithms are published for use of active-optical (AO) sensors for corn, and work continues in sugarbeet, confection sunflower, and spring wheat. Use of zone soil sampling and the use of active-optical sensors can provide improved site-specific N management.

Keywords: zone sampling, grid sampling, nitrogen, active-optical sensors

Introduction

North Dakota agriculture includes the cultivation of about 6 Mha of nitrogen (N)-requiring crops. The most important of these by area are spring wheat (*Triticum aestivum*, L.), durum wheat (*Triticum turgidum* ssp. Durum), corn (*Zea mays*, L.), and sunflower (*Helianthus annuus*, L.). The foundation of N management in this region is fall soil sampling for residual nitrate. Harvest is usually complete by the end of October, and the soil freezes shortly thereafter. For example, in 2018 at Grand Forks, ND, the bare soil temperature first measured 0 °C November 17. (NDAWN, 2018) Residual soil nitrate is trapped in the frozen soil until the spring thaw, which is closely followed by spring planting often while frost can still be found in the soil. In 2018 for example, bare soil temperature at Grand Forks first rose above 0 °C April 15. Planting began in the region within 2 weeks.
Starting in 1995, site-specific N application became of interest to farmers and their suppliers as fertilizer applicators capable of varying the rate of fertilizer became available and profitability of crop production became more difficult to attain. Grid soil sampling was first explored, followed by zone soil sampling. Fertilizer recommendations were revised to better serve site-specific field nutrient applications. Today, site-specific nutrient management should also include strategies to account for in-season nutrient requirements. This paper will describe the transition from whole-field nutrient management to site-specific nutrient management in North Dakota, and the integrated approach to nutrient management now recommended to farmers to maximize their profitability and environmental stewardship. This paper may serve to provide an incentive to researchers around the world to utilize some of these ideas to provide improved, practical site-specific nutrient strategies for their farmers to improve the efficiency of nutrient application and profitability of crop production.

Soil sampling

Soil sampling for site-specific nutrient application is not a new concept. The first recommendations to farmers for a site-specific soil sampling strategy was published by Linsley and Bauer (1929) as a guide for site-specific lime application that was labor-saving for the hand-labor used at the time.

DOI 10.3920/978-90-8686-888-9_65, © Wageningen Academic Publishers 2019

With the availability of equipment that could more easily apply fertilizers and lime with far less labor, the interest in site-specific application faded for almost 50 years. Interest in within-field nutrient management was again sparked by the development of fertilizer applicators with the ability to apply varying rates of fertilizer within a field, combined with the availability of global navigation satellite systems (GNSS) and the computer capacity to collect GNSS data and analyze it by companies and individuals practically and affordably (Franzen and Mulla, 2016).

A series of soil sampling studies were conducted in several states to determine what density of soil sampling would practically and meaningfully result in the determination of fertility patterns in fields that could be better managed through their site-specific nutrient application. The results from Wisconsin (Wollenhaupt *et al.*, 1994), Illinois (Franzen and Peck, 1995) and Nebraska (Gotway *et al.*, 1996) indicated that a sampling density of about 1 sample per 0.4 ha was required to direct meaningful site-specific fertilizer application. In these states, soil sampling is usually directed towards phosphorus (P), potassium (K) and soil pH, with a sampling depth of about 150 mm. Each sample that would be sent to a laboratory for analysis would consist of 3-5 soil cores in a composite.

In North Dakota, the sampling depth is most typically 600 mm due to the importance of measuring soil nitrate-N for use in N recommendations for small grains, canola (*Brassica napus*), sunflower, corn and other N-requiring crops. The recommended depth of soil cores for sugar beet (*Beta vulgaris* subsp. *vulgaris*) was 1,200 mm until recently. The deeper soil samples still required 3-5 cores within a composite. One sample per 0.4 ha would have been impractical for farmers or their consultants from the aspects of time and expense. A field near Valley City, ND was sampled in a 0.1 ha grid. Patterns of nitrate-N, P and other nutrients were mapped. The field was sampled at the same sample density the next year, and patterns mapped in year two were similar to those found in year one (Figure 1). The reason for the similarity in patterns was topography. Water movement and crop growth was influenced by topography (Franzen *et al.*, 1998), and patterns of nitrate and other nutrients were similar between years. The sampling pattern facilitated investigating different sampling patterns on the correlation of nitrate-N values using different soil sample density. The topography-based zone sampling, using a 5-sample composite for each zone was similar in significance or greater than the 0.4 ha grid and greater than grid density of 1 ha, or 2 ha per sample (Table 1). The use of zone sampling greatly reduced the time and expense of soil sampling for meaningful site-specific nutrient application while equaling or exceeding the predictive ability of a 0.4 ha grid. Zone sampling resulted in the practical application of site-specific nutrient application, particularly N, in North Dakota.

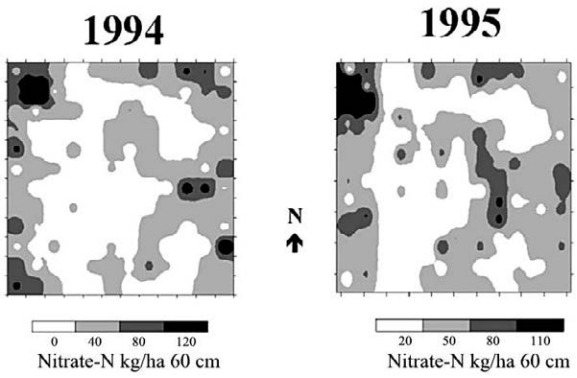


Figure 1. Residual soil nitrate-N mapping of a square 12.5 ha field near Valley City, ND in 1994 (after spring wheat) and in 1995 (after sunflower), sampled in a 0.1 ha grid. The patterns indicate stability of relative residual nitrate values between years.

Table 1. Regression coefficients (r^2) for comparison of zone-based and less dense grid estimates with the 0.1 ha sampling grid, Valley City, ND.[1]

Year	r^2 for estimates with 0.1 ha grid values			
	Topography	0.4 ha	1 ha	2 ha
1994	0.29**	0.18	0.07	0.07
1995	0.38**	0.50**	0.21*	0.21*
1996	0.49**	0.34**	0.41**	0.01

[1] * indicates significance of regression coefficient at $P<0.05$. ** indicates significance of regression coefficient at $P<0.01$.

Continued research in the region, amounting to over 50 site-years of whole-field data with experimental base grid sampling of 0.1 ha to 0.2 ha densities, showed that zone sampling was appropriate from eastern Montana through western Minnesota. Tools to develop effective zones were aerial and satellite imagery, soil electrical conductivity sensing, soil magnetic conductance sensing, topography (not elevation, but landscape structure), and multi-year yield maps (with each year standardized for relative yield) (Franzen *et al.*, 2011). It was important that at least two tools were utilized to produce the nutrient management zones. Tools that researchers have used to delineate zones include aerial imagery, satellite imagery, soil EC sensor data, soil EM sensor data, multi-year yield maps, farmer experiences and observations, and topography (Hornung, 2006; Franzen *et al.*, 2011). In a given year, use of only one tool has a chance of not representing nutrient patterns, while two tools has a much greater ability to represent patterns even if the other tool is not as useful.

A zone sampling approach for crop nutrients has also been explored by other researchers in the USA. Mallorino and Wittry (2004) found that zone sampling in Iowa was superior to grid sampling for soil pH, organic matter and potassium. Grid sampling was superior to zone sampling for phosphorus. Khosla *et al.* (2002) showed that zone-based N management improves N management in Colorado. Adoption of zone sampling is high among North Dakota farmers and their consultants. In 2016, use of variable-rate N application using zone sampling was used on nearly 54% of sugar beet area, with profitability of its use being about $120/ha/yr averaged over 10 years (American Crystal Sugar Co., Moorhead, MN, personal communication).

North Dakota N recommendations and recommendations for other plant nutrients of interest were formally yield-based. Yield-based recommendations tended to be embraced by farmers utilizing site-specific nutrient application because it was easy to link productivity to a nutrient rate. However, N rate trials on spring wheat (Franzen, 2018), corn (Franzen, 2017), and sunflower (Schultz *et al.*, 2018) showed that N rate and yield between fields were not related. Raun *et al.* (2011) first recognized the lack of yield and N relationship between fields, and several US Corn Belt states do not recognize yield as a factor in their economic production function N rate models (Sawyer, 2018). The lack of yield-goal based N recommendations places greater emphasis on attention to other tests, such as soil nitrate-N sampling and analysis. From this type of test, regional soil rate responses (as detailed in the literature) produce N fertilizer recommendations such as N-calculators (Sawyer and Nafziger, 2005; Franzen, 2009; Franzen, 2015; Franzen, 2016; Sawyer, 2018). According to the largest soil testing laboratory in the region, 40% of samples received from North Dakota are intended for direction of site-specific nutrient application, with zone samples by far the largest proportion of site-specific nutrient soil samples (Agvise, 2018).

The regression analysis of zone sampling with the experimental 0.1 ha base grid values used as a standard for comparison of less dense grids (i.e. 1 sample/ha) for soil nitrate-N suggests that soil

analysis only represents 50% of soil nitrate-N variability (Franzen *et al.*, 1998). Active-optical (AO) sensors are able to recognize small-scale variability in N availability and were developed as a soil fertility management tool to serve that purpose (Raun *et al.*, 2002).

Development of active-optical sensor algorithms for in-season N fertilization

In a series of N-rate experiments in corn, the AO sensors at growth stage V5 and V10 were used to relate the red NDVI (RNDVI) and red edge NDVI (RENDVI) (when possible with the instrument) to yield (Sharma *et al.*, 2015). The AO sensors used were the GreenSeeker® (Trimble, Inc., Sunnyvale, California, USA), and the Holland Scientific® Crop Circle A470 (Holland Scientific, Lincoln, NE, USA) sensor. The GreenSeeker (GS) only had a RNDVI light source/detector, while the Crop Circle (CC) generated RNDVI and RENDVI data. The results indicated a strong relationship between yield and both RNDVI and RENDVI at V5, but only RENDVI at V10.

The results of the study were further condensed and incorporated into a series of algorithms for use by farmers for site-specific on-the-go N application (Franzen *et al.*, 2014). These algorithms would enable the farmer to apply a base preplant N rate across the field of about 50-75% of total recommended seasonal N indicated by zone sampling and the N recommendation. At about V5 (Ransom, 2013), the sensor would be used to deliver any remaining N to the field in a small-scale site-specific manner on-the-go using the algorithm developed for their region and soil. The algorithms also contain a minimum value for sensor readings whereby the likely cause of deficiency is something other than N. Readings below the minimum are probably the result of a poor plant stand, high salts or another non-N factor. This prevents the grower from wasting fertilizer N on a field area where the N will not be helpful for yield improvement. The standard against which the algorithm rate is based comes from an N-non-limiting area established at the time of preplant N application. The readings from the N non-limiting area would represent the greatest yield possible with any rate of N applied. The difference in yield prediction between the N non-limiting area and any other area of the field would be used by the algorithm to calculate the N required to approach the higher yield. A similar concept is also used by others in the USA to direct in-season N application (Franzen *et al.*, 2016a). The regression equations from these studies that relate AO sensor readings to crop yield identify an additional 25% of field variability. The combination of zone soil sampling with the use of an AO sensor would identify about 75% of total N variability within fields.

Algorithm development for spring wheat, sugar beet and sunflower is currently on-going. The algorithms for spring wheat will include early-season N algorithms for yield increase, and anthesis-timing algorithms for protein enhancement N applications. The sugar beet algorithms will be needed for a series of harvest dates based on growing degree days, since sugar beet continues to develop and increase in tonne yield and sucrose concentration from the beginning of sugar beet harvest in the region until the end; about a 6 week period. The algorithm will therefore have to be growing degree day by yield by N rate, or a 3-dimensional model (Bu *et al.*, 2017).

Sunflower algorithms for confection sunflower are likely, though not yet developed, since there is a strong relationship between confection sunflower RNDVI, RENDVI and yield at V6 (Schultz *et al.*, 2018). Oil-seed sunflower yield was not related to AO sensor readings (Franzen *et al.*, 2019). However, they are related to crop height, which was determined using an acoustic height sensor (Sharma *et al.*, 2015). Corn yield is also related to crop height, but the RNDVI and RENDVI relationships with corn yield were strong enough that crop height was not included in corn AO sensor algorithms. The algorithms developed for corn may be modified by farmers. To do so, they need to turn off the in-season N application in a pass through the field, but continue to collect NDVI readings. At harvest, the relationship between the NDVI readings and yield will result in a new relationship equation. This can be combined with a weighted equation based on the number of sites used to develop the algorithm, and a new algorithm is produced which, over time, will become the farmer's farm algorithm. Since there will be year to year variation in growing season moisture

and temperature conditions, it would probably take multiple years to produce an algorithm that would provide farmer confidence.

Currently, few farmers are using AO sensors for variable-rate in-season N application. Feedback from farmers indicates that they do not trust the technology to automatically apply different N rates to their fields. However, an increasing number of farmers use the N non-limiting concept to help them determine whether or not they require in-season supplemental N applications. An N non-limiting area will also serve to indicate a S deficiency in corn, based on a recent study (Franzen *et al.*, 2016b). When N limits yield, S deficiency is not as severe. When N is non-limiting, S deficiency intensifies.

Conclusions

The goal of a site-specific nutrient management extension education program should be to move the farmer to a more integrated site-specific approach, with early crop health and in-season adaptable nutrient management due to early-season rainfall. In N research on high clay soils, those with clay content greater than 350 g/kg soil dry matter, excessive early-season rainfall can result in denitrification with losses in excess of 100 kg/ha. In addition, coarse textured soils are susceptible to high N leaching losses. Therefore, in these soils, preplant N application alone may result in N loss and lower yield in wet years. By using zone soil sampling as a basis for about half the required N applied as a preplant rate, followed by AO guided in-season N application, the efficiency of the total N applied increases, and small-scale in-season N variability is addressed. By using a zone sampling strategy for soil nitrate-N, farmers are able to address about 50% of the N variability within their fields. If farmers additionally utilize AO sensors to direct in-season N management, a total of 75% of field N variability would be addressed according to this research.

Acknowledgements

Portions of this research were funded by the US-National Science Foundation, grant number PFI-1114363, the International Plant Nutrition Institute, the North Dakota Corn Council, the National Sunflower Association, Pioneer Hibred, Int. Inc., USDA-CSREES-IFAFS award #00-52103- 9652, and the US EPA.

References

Agvise, 2018. Agvise soil test summary 2018. https://www.agvise.com/wp-content/uploads/2018/11/AGVISE-Soil-Test-Summary-2018.pdf (last accessed 11/2018).

Bu, H., Sharma, L.K. Denton, A. and Franzen D.W., 2017. Comparison of satellite imagery and ground-based active optical sensors as yield predictors in sugar beet, spring wheat, corn and sunflower. Agronomy Journal 109 1-10. doi:10.2134/agronj2016.03.0150

Franzen, D., 2009. North Dakota Wheat Nitrogen Calculator. https://www.ndsu.edu/pubweb/soils/wheat (last accessed 11/2018).

Franzen, D., 2015. North Dakota Corn Nitrogen Calculator. https://www.ndsu.edu/pubweb/soils/corn (last accessed 11/2018).

Franzen, D., 2016. North Dakota Sunflower Nitrogen Calculator. https://www.ndsu.edu/pubweb/soils/sunflower (last accessed 11/2018).

Franzen, D.W., 2017. Soil fertility recommendations for corn. NDSU Ext. Circ. SF722 (revised).
https://www.ag.ndsu.edu/publications/crops/soil-fertility-recommendations-for-corn/sf722.pdf (last accessed 11/2018).

Franzen, D.W., 2018. Fertilizing hard red spring wheat and durum. NDSU Ext. Circ. SF712 (revised). https://www.ag.ndsu.edu/publications/crops/fertilizing-hard-red-spring-wheat-and-durum-1/sf712.pdf (last accessed 11/2018).

Franzen, D.W., and Peck T.R., 2015. Field soil sampling density for variable rate fertilization. Journal of Production Agriculture 8 568-574. doi:10.2134/jpa1995.0568

Franzen, D.W., and Mulla D., 2016. A history of precision agriculture. p. 1-20. In-Precision Agriculture Technology for Crop Farming. Q.Zhang (ed.). New York, USA: CRC Press,.

Franzen, D.W., Cihacek, L.J. Hofman, V.L. and Swenson L.J., 1998. Topography-based sampling compared with grid sampling in the northern Great Plains. Journal of Production Agriculture 11 364-370. doi:10.2134/jpa1998.0364

Franzen, D., Long, D. Sims, A. Lamb, J. Casey, F. Staricka, J. *et al.,* 2011. Evaluation of methods to determine residual soil nitrate zones across the northern Great Plains of the USA. Precision Agriculture 12 594-606. doi:10.1007/s11119-010-9207-0

Franzen, D., Sharma, L.K. and Bu H., 2014. Active optical sensor algorithms for corn yield prediction and a corn side-dress nitrogen rate aid. NDSU Extension Circular SF1176-5. https://www.ndsu.edu/fileadmin/soils/pdfs/sf1176-5.pdf (last accessed 11/2018).

Franzen, D., Kitchen, N. Holland, K. Schepers, J. and Raun, W., 2016a. Algorithms for in-season nutrient management in cereals. Agronomy Journal 108 1-7. doi:10.2134/agronj2016.01.0041

Franzen, D.W., Schultz, E.C. DeSutter, T.M. Sharma, L.K. Ashley, R. and Bu H., 2019. Sunflower type influences yield prediction using active optical sensors. Agronomy Journal 111 (2) https://tinyurl.com/y6sjpsgn

Franzen, D.W., Sharma, L.K. Bu, H. and Denton. A., 2016b. Evidence for the ability of active-optical sensors to detect sulfur deficiency in corn. Agronomy Journal 108 2158-2162. doi:10.2134/agronj2016.05.0287

Gotway, C.A., Ferguson, R.B. Hergert, G.W. and Peterson T.A., 1996. Comparison of kriging and inverse-distance methods for mapping soil parameters. Soil Science Society of America Journal 60 1237-1247. doi:10.2136/sssaj1996.03615995006000040040x

Hornung, A., Khosla, R. Reich, R. Inman, D. and Westfall D.G., 2006. Comparison of site-specific management zones: soil-color-based and yield-based. Agronomy Journal 98:407-415.

Khosla, R., Fleming, K. Delgado, J.A. Shaver, T.M. and Westfall D.G., 2002. Use of site specific management zones to improve nitrogen management for precision agriculture. Journal of Soil and Water Conservation. 57 513-518.

Linsley, C.M. and Bauer F.C., 1929. Test your soils for acidity. University of Illinois Circular 346, Urbana, IL, USA.

Mallarino, A.P. and Wittry D.J., 2004. Efficacy of grid and zone soil sampling approaches for site-specific assessment of phosphorus, potassium, pH, and organic matter. Precision Agriculture 5 131-144.

NDAWN. North Dakota Agricultural Weather Network, 2018. https://ndawn.ndsu.nodak.edu

Ransom, J., 2013. Corn growth and management quick guide. NDSU Ext. Circ. A1173. N. Dak. St. Univ. Ext Serv., Fargo, ND https://www.ag.ndsu.edu/pubs/plantsci/crops/a1173.pdf (Accessed February, 2019).

Raun, W.R., Solie, J.B. Johnson, G.V. Stone, M.L. Mullen, R.W. Freeman, K.W. *et al.,* 2002. Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. Agronomy Journal 94 815-820. doi:10.2134/agronj2002.8150

Raun, W.R., Solie, J.B. and Stone M.L., 2011. Independence of yield potential and crop nitrogen response. Precision Agriculture 12 508-518. doi:10.1007/s11119-010-9196-z

Sawyer, J.E. and Nafziger E.D., 2005. Regional approach to making nitrogen fertilizer rate decisions for corn. In Proceedings of the 2005 North Central Extension-Industry Soil Fertility Conference. Vol. 21, pp 16-24Brookings, SD, USA: Phosphate & Potash Institute.

Sawyer, J., 2018. Corn Nitrogen Rate Calculator. http://cnrc.agron.iastate.edu (last accessed 11/2018).

Schultz, E., DeSutter, T. Sharma, L. Endres, G. Ashley, R. Bu, H. *et al.,* 2018. Response of sunflower to nitrogen and phosphorus in North Dakota. Agronomy Journal 110 1-11. doi:10.2134/agronj2017.04.0222

Sharma, L.K., Bu, H. Denton, A. and Franzen D.W., 2015. Active-optical sensors using red NDVI compared to red edge NDVI for prediction of corn grain yield in North Dakota, U.S.A. Sensors 15 27832-27853. doi:3390/s151127832

Wollenhaupt, N.C., Wolkowski, R.P. and Clayton M.K., 1994. Mapping soil test phosphorus and potassium for variable-rate fertilizer application. Journal of Production Agriculture 7 441-448.

Dividing the risk – theoretical exploration of increasing N management temporal granularity in maize

A.G. Hunt[1*], J.M. Sharp[2], P.R. Johnstone[1] and B.P. Searle[1]
[1]*New Zealand Institute for Plant and Food Research, Hawke's Bay, New Zealand; adrian.hunt@plantandfood.co.nz*
[2]*New Zealand Institute for Plant and Food Research, Lincoln, New Zealand*

Abstract

Crop demand and soil supply of nitrogen (N) varies significantly within and between seasons. This temporal variability poses an important challenge to growers looking to maximise yields and profitability whilst minimising nutrient losses. It is hypothesised that increasing the temporal granularity (number of occasions within a crop cycle) at which N fertiliser decisions are made and applied to a crop will increase N use efficiency, without reducing yield. To test this, the systems modelling platform APSIM Next Generation was used to simulate the growth of a maize grain crop over 35 different growing seasons and assessed N management scenarios ranging from one N decision per crop to daily N decisions. N management was informed by soil mineral N status and predicted N soil supply and crop demand before the next N management action. Results indicated that increasing the temporal granularity of N management substantially improves the alignment between N supply and demand without limiting yield.

Keywords: variability, nitrogen use efficiency, APSIM, maize, *Zea mays*

Introduction

Nitrogen (N) is a key macronutrient for plant growth. The majority of the world's population is dependent on the application of mineral N fertilisers to supplement soil supply, for energy and protein (Dawson and Hilton, 2011). However there are potential risks associated with N fertiliser applications. Nitrogen can be lost from production systems through leaching or gaseous losses, thus limiting plant productivity, as well as impacting negatively on the environment through eutrophication of waterways (Carpenter *et al.*, 1998) and greenhouse gas effects (Snyder *et al.*, 2009). Combined with other sources of N pollution, the cost to the EU alone has been estimated to be at least 70 billion euros per year (Sutton *et al.*, 2011). Hence there is a need to apply N fertilisers prudently to optimise productivity and environmental outcomes.

Shanahan *et al.* (2008) identified that mismatches between soil N supply and crop N demand, and failure to account for temporal variability in crop N demand, are potential causes of low N use efficiency (NUE) and high N losses from agricultural production systems. The challenge of supplying N to maximise productivity while minimising negative environmental impacts is made difficult by variability in factors which affect soil N availability, supply and crop N demand. These factors include crop species and cultivar, planting date, soil type, soil moisture, soil nutrient status, precipitation, radiation, temperature, weeds, pests and diseases. Of these factors, temperature, radiation and precipitation (i.e. climate) can vary substantially between and within seasons, and it has been predicted that their variability will increase as a result of global warming (Ray *et al.*, 2015). The consequence of this variability is that any N management strategy that is not responsive to changes in the environment risks under or over supply of N.

Current N management practice is dominated by upfront and/or split applications of N. These methods do not account for inter-annual or within-season variability. At best, these practices account for soil N supply based on soil mineral-N testing and N mineralisation values. Together

DOI 10.3920/978-90-8686-888-9_66, © Wageningen Academic Publishers 2019

with a yield potential estimated from experience or a long-term average climate set, the timing and rate of N applications can be determined using an N balance approach. These scheduling decisions are generally made prior to planting and not updated within the season, despite within-season variation in the drivers of soil N supply and crop N demand. Under this approach, to avoid a risk of yield loss due to insufficient N supply, scheduling typically assumes a good growing season and targets a high potential yield, while accounting for potentially low soil organic mineralisation and hence soil supply. While this approach rarely limits productivity, it will supply excess N to the system when environmental conditions do not result in the high potential yield.

Given the temporal variability of within-season climatic conditions that affect soil N supply and crop N demand, there may be an opportunity to improve crop NUE and reduce environmental losses by implementing an N management approach that responds to this variability, i.e. tailoring N supply to within-season climatic conditions. It is therefore hypothesised that increasing the temporal granularity (i.e. number of occasions within a crop cycle) at which N fertiliser decisions are made and applied to a crop will improve NUE and mitigate environmental losses. To explore the potential impact of this management paradigm, a modelling approach, using an agricultural systems model, was used.

Materials and methods

The mechanistic, processed-based agricultural systems model APSIM next generation, as described in Holzworth *et al.*, 2014, was used to explore how changes in the temporal granularity of N applications in response to within-season climatic conditions affected NUE and N losses. The system selected was irrigated maize grown for grain, modelled using the APSIM maize module, whose performance has been validated across New Zealand (Brown *et al.*, 2015), including the region used here. The location selected was a 21-ha field near Otane (S 39°52¢, E 176° 40¢) in the Hawke's Bay region of New Zealand and managed using typical management practices for the region, and for which the site and model parameterisation was previously reported by Sharp *et al.* (2017). In brief, measured soil characteristics, including soil mineral N, mineralisable N, water holding capacity, and 35 individual years (1980-2015) of weather data were used in the model. The site had a mixed rotation history, and is a Waimakariri loam over sandy loam with 252 mm of plant-available water in the top 2 m. Extensive field sampling and laboratory analyses indicated that soil chemical properties were unlikely to limit the growth of maize, other than mineral N.

To explore the impact of changing the temporal granularity of N applications, treatments were selected with different periods of time between N fertiliser input decisions. Specifically, these were: upfront (a single application of all fertiliser applied at planting), a starter and side dressing (a starter amount at planting and a side-dressings 60 days after planting), every 30 days, every 7 days, every 3 days and every day. The N management process for all treatments followed a yield risk-averse approach by assuming the maximum potential yield, and therefore crop N uptake (Figure 1A), and the lowest soil mineral N supply (Figure 1B). These values were determined from model simulations run for the 35 years of weather data under irrigated and unconstrained N conditions. N fertiliser applications were then calculated as the difference between the potential N uptake and the potential soil supply before the next fertiliser application or harvest of the crop (Figure 2). Scheduling decisions were also designed to maintain a minimum level of available soil N that did not limit crop N uptake through physical/chemical constraints. This margin was adjusted dependent on the potential daily N uptake and ranged from soil N of 15 kg/ha early and late in the season to 30 kg/ha during the peak in plant uptake.

Outputs from the simulations were collated in a database. Statistics and presentation of results were conducted in R (www.r-project.org), using the packages dplyr, mgcv and ggplot2 (package descriptions available at cran.r-project.org).

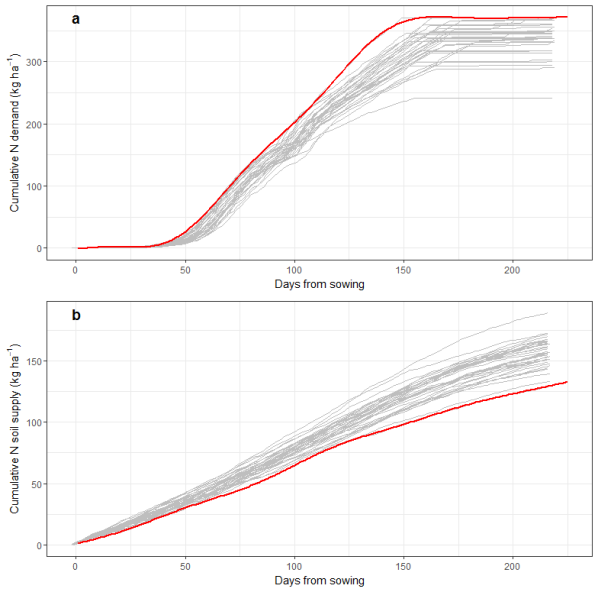


Figure 1. (A) Cumulative N demand (kg/ha) of the modelled crop in each of the 35 years under no N supply constraint, bold/red line indicates maximum uptake. (B) Cumulative soil mineral N supply (kg/ha) by mineralisation, bold/red line indicates minimum supply.

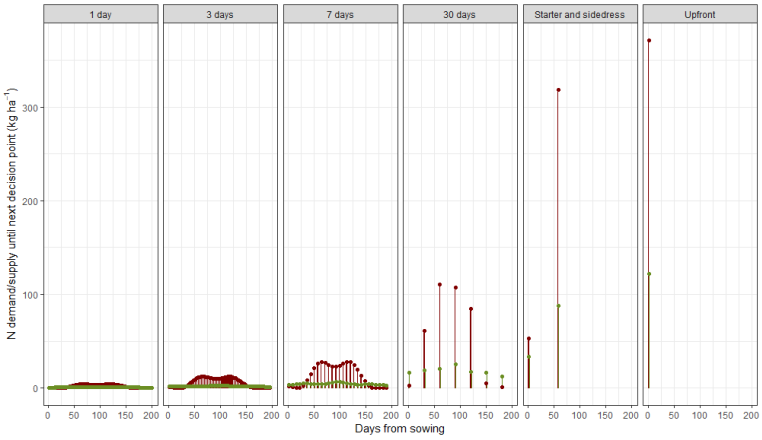


Figure 2. Predicted crop N demand prior to the next N application (or harvest) (dark/red dots), and predicted soil supply of plant available N (light/green dots), for each of the temporal N management treatments.

Results

Increasing the temporal granularity of N management reduced the concentration and residence time of NO_3 present in the soil. For instance, the single fertiliser application in the upfront treatment

resulted in soil NO_3 concentrations exceeding 280 kg/ha, in comparison with daily N management treatment in which most years the concentration did not exceed 50 kg/ha (Figure 3A). In addition, the granularity of N management affected the proportion of time during the season where the soil had an elevated soil NO_3 content, increasing the risk of N losses through leaching events. For instance, with daily N applications, soil NO_3 concentration was above 50 kg/ha in autumn for only 20% of all years, and in those years for only 2-10% of all days, compared with 100% of years and over 90% of days for the upfront treatment. This leaching risk can be represented simplistically as the sum of the daily soil NO_3 concentration accumulated through the cropping cycle (Figure 3B). Upfront N management resulted in a mean cumulative soil NO_3 concentration of 35,673 kg/ha days at the end of the season, with a standard deviation (sd) of 2,525 kg/ha days, while increasing the temporal granularity in N management to daily management resulted in a mean of 6,078 kg/ha and sd of 507 kg/ha, an approximate 6- and 5-fold decrease, respectively (Figure 3B).

In the upfront treatment, the rate of N fertiliser applied was consistent in all years at 279 kg/ha. This treatment responded only to the simulation starting conditions, which were the same for all simulations, and forecasting fertiliser N application from the maximum crop demand and the minimum soil supply over the whole season. As the level of temporal granularity increased, the mean seasonal N fertiliser application decreased. However, with that, an increase in among-year variability in seasonal N fertiliser application was observed. This is due to N management responding to within-season variation in environmental conditions that influence crop growth, and to a lesser extent soil N supply (Figure 4A). No other treatment applied as much fertiliser as the up-front treatment and, in some years, the total applied N was less than 200 kg/ha for the 7 days or less treatments. Despite these reductions in N fertiliser application through increased temporal granularity, grain yields were identical across all treatments, indicating that N was not limiting within each year (Figure 4B). Grain yields did however vary between 9.7 and 16.8 t/ha among simulation years. As all simulations

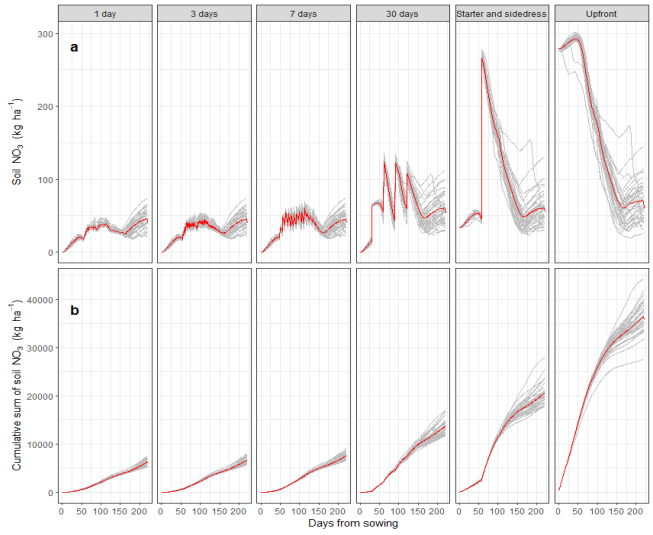


Figure 3. (A) Soil NO_3 content over time each year for the temporal N management treatments over 35 different growing seasons, bold line indicates mean. (B) Cumulative sum of soil NO_3 content accumulated through the cropping cycle each year for the temporal N management treatments, bold line indicates mean.

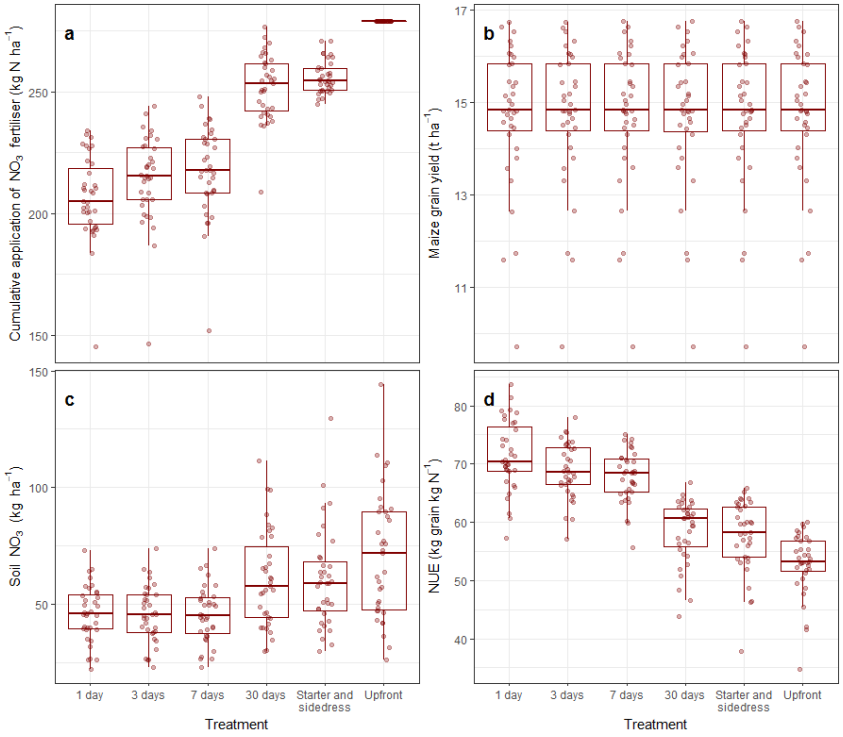


Figure 4. (A) Rate of fertiliser applied in each year kg N/ha by treatment. (B) Yield of grain Maize t/ha by treatment. (C) Soil NO_3 content kg/ha to a depth of 2 m. (D) Nitrogen Use Efficiency (NUE) kg grain per kg N applied as fertiliser.

received optimal irrigation, the primary driver of this yield variability would have been inter-annual differences in temperature and radiation.

Temporal granularity of N management substantially influenced the soil NO_3 content at the end of each season. Mean (sd) residual soil N decreased from 71 (27) kg N/ha with upfront treatment to 46 (12) kg N/ha with daily fertiliser treatment (Figure 4C). This represents a 35% decrease in the mean value and 56% decrease in the variability around that mean. The highest NUE was achieved in the most granular N management treatment (daily) with a mean of 71 kg grain per kg of nitrogen fertiliser applied. Decreasing granularity resulted in a reduction in NUE, with the most coarse N management treatment (upfront) resulting in the lowest NUE of 53 kg grain per kg of N applied. Despite a 26% decrease in mean NUE, the variability around the mean (sd) showed only an 18% decrease.

Discussion

Increasing temporal granularity of N management decreased the likelihood that crop demand for or soil supply of N would vary substantially from their predicted values before the next N management step. By breaking down the N management decisions into smaller steps, the risk of an N constrained yield did not change, but the opportunity to better match N supply to variation in N demand, resulting from within-season climatic conditions, increased. While the system simulated was irrigated, in an unirrigated growing system where insufficient soil water may further constrain yield and result in

greater within- and between-season variation in crop N demand, the opportunity to more closely match N supply to crop N demand through increased temporal granularity of N management may be greater. This may be especially pronounced in regions where there is substantial variability in seasonal rainfall and limited access to irrigation.

In this study, increased temporal granularity of N management decreased the amount of fertiliser applied but maintained yield, and therefore increased NUE. This is in contrast to previous studies where, although NUE values calculated using applied fertiliser ranges were greater with lower applications of N fertiliser, there was a significant decrease in yield (Maltese *et al.*, 2019, Haegele *et al.*, 2013, Moll *et al.*, 1982). In these studies N was applied as a starter with two or three side-dressings following a predetermined program, which did not allow for adjustment for seasonal variation.

In the study reported here, N management was scheduled on a calendar basis which was chosen to best reflect current grower norms. However, further improvements may be possible by moving to a scale using accumulated thermal time, which better matches the progression of crop development. Under such a system, N applications would become more frequent under optimal growth conditions. Using a thermal time-based approach may make it more difficult to plan farm operations, as it is more difficult to predict when an operation will need to be implemented. However integrating local environmental data and forecasting could enable a shift of this nature. At a coarse level, this is already the case where observation of key crop development stages is used to schedule N application and crop sensing is utilised to check for plant N deficiency (Scharf *et al.*, 2002).

Advances in fertiliser delivery technology such as fertigation are facilitating opportunities for increased temporal granularity of nutrient supply. These advances will only lead to beneficial outcomes in productivity and the environment if the delivery of nutrients is sufficient to keep up with plant demand without leaving excess N available for loss to the environment. Maize was chosen in this study due to the strong model available for simulations (Brown *et al.*, 2015). However greater benefits from increased temporal granularity may be achieved in other, 'higher risk' crops such as onions (Wright, 1993, Halvorson *et al.*, 2002) and potatoes (Errebhi *et al.*, 1998), which have shallower rooting depths and excess N supply can reduce quality. Improvements in the temporal management of N supply in these or any other crops will require a strong foundational knowledge of soil N supply and crop N uptake. This knowledge could be integrated using models, but it also seems likely that real-time measures or proxies of N supply and demand would be required to support the further refinement of N management and to build a robust practical system. Faced with an increasingly variable climate, greater food demand and focus on reducing the environmental impact of food production, building systems and tools to support a more reactive approach to temporal N management appears to be a wise investment.

Conclusions

There is evidence of potential benefits by moving from an upfront N management system to one which takes into account variability in within- and among-season climatic conditions that affect soil N supply and crop N demand. This approach offers an opportunity to substantially increase NUE and reduce the environmental impact of intensive cropping systems. However advances in real-time evaluation of crop N demand and soil N supply may be required to achieve practical implementation.

Acknowledgements

This work was funded by the New Zealand Ministry of Business, Innovation & Employment through the Strategic Science Investment Fund. (SSIF) programme Sustainable Agro-Ecosystems.

References

Brown, H. E., Teixeira, E. I., Huth, N. I. and Holzworth, D. P., 2015 The APSIM Maize Model. APSIM initiative. http://www.apsim.info/ApsimxFiles/Maize1757.pdf (last accessed December 2018).

Carpenter, S.R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N. and Smith, V. H., 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications 8 (3) 559-568.

Dawson, C. J. and Hilton, J., 2011. Fertiliser availability in a resource-limited world: Production and recycling of nitrogen and phosphorus. Food Policy 36 S14-S22.

Errebhi, M., Rosen, C. J., Gupta, S. C. and Birong, D. E., 1998. Potato yield response and nitrate leaching as influenced by nitrogen management. Agronomy Journal 90 10-15.

Haegele, J. W., Cook, K. A., Nichols, D. M. and Below, F. E., 2013. Changes in nitrogen use traits associated with genetic improvement for grain yield of maize hybrids released in different decades. Crop Science 53 (4) 1256-1268.

Halvorson, A. D., Follett, R. F., Bartolo, M. E. and Schweissing, F. C., 2002. Nitrogen fertilizer use efficiency of furrow-irrigated onion and corn. Agronomy Journal 94 442-449.

Holzworth, D. P., Huth, N. I., deVoil, P. D., Zurcher, E. J., Herrmann, N. I., McLean, G., *et al.,* 2014. APSIM – Evolution towards a new generation of agricultural systems simulation. Environmental Modelling & Software 62 327-350.

Maltese, N. E., Melchiori, R. J. M., Maddonni, G. A., Ferreyra, J. M. and Caviglia, O. P., 2019. Nitrogen economy of early and late-sown maize crops. Field Crops Research 231 40-50.

Moll, R. H., Kamprath, E. J. and Jackson, W. A., 1982. Analysis and interpretation of factors which contribute to efficiency of nitrogen utilization 1. Agronomy Journal 74 (3) 562-564.

Ray, D. K., Gerber, J. S., MacDonald, G. K. and West, P. C., 2015. Climate variation explains a third of global crop yield variability. Nature Communications 6 5989.

Scharf, P. C., Wiebold, W. J. and Lory, J. A., 2002. Corn yield response to nitrogen fertilizer timing and deficiency level. Agronomy Journal 94 435-441.

Shanahan, J. F., Kitchen, N. R., Raun, W. R. and Schepers, J. S., 2008. Responsive in-season nitrogen management for cereals. Computers and Electronics in Agriculture 61 (1) 51-62.

Sharp, J., Johnstone, P., Searle, B., Hunt, A., Qiu, W. and Wallace, D., 2017. The impact of spatial variability in soil nitrogen and the value in its management: A case study. The 7th Asian-Australasian Conference on Precision Agriculture, Hamilton, New Zealand http://doi.org/10.5281/zenodo.895491

Snyder, C. S., Bruulsema, T. W., Jensen, T. L. and Fixen, P. E., 2009. Review of greenhouse gas emissions from crop production systems and fertilizer management effects. Agriculture, Ecosystems & Environment 133 (3-4) 247-266.

Sutton, M. A., Oenema, O., Erisman, J. W., Leip, A., van Grinsven, H. and Winiwarter, W., 2011. Too much of a good thing. Nature 472 159.

Wright, P.J., 1993. Effects of nitrogen fertiliser, plant maturity at lifting, and water during field-curing on the incidence of bacterial soft rot of onions in store. New Zealand Journal of Crop and Horticultural Science 21 (4) 377-381.

Creating soil texture maps for precision liming using electrical resistivity and gamma ray mapping

S. Meyer[1*], C. Kling[2], S. Vogel[4], I. Schröter[3], A. Nagel[3], E. Kramer[3], R. Gebbers[4], G. Philipp[5], K. Lück[6], F. Gerlach[6], D. Scheibe[7] and J. Ruehlmann[1]
[1]*Leibniz-Institute of Vegetable and Ornamental Crops, 14947 Grossbeeren, Germany; meyer.swen@igzev.de*
[2]*Gut Wilmersdorf GbR, 16278 Angermünde OT Wilmersdorf, Germany*
[3]*Eberswalde University for Sustainable Development, Landscape Management and Nature Conservation, 16225 Eberswalde, Germany*
[4]*Leibniz Institute for Agricultural Engineering and Bioeconomy (ATB), Engineering for Plant Production, 14469 Potsdam, Germany*
[5]*Landwirtschaft Petra Philipp, 15234 Frankfurt/O. OT Booßen, Germany*
[6]*Land- und Forstwirtschaft Komturei Lietzen GmbH & Co KG, 15306 Lietzen, Germany*
[7]*LAB Landwirtschaftliche Beratung der Agrarverbände Brandenburg GmbH, 15374 Müncheberg, Germany*

Abstract

This paper shows the first results of the EU-funded research project 'pH-BB: Precision liming in Brandenburg'. The project aims to develop innovative solutions to adapt the current practice of pH management to the demands of a modern, resource-efficient and yield-optimized agricultural production. In the project, proximal soil sensors are used to rapidly generate high-resolution and cost-effective spatial datasets of soil parameters (i.e. soil pH, soil texture, soil organic matter) relevant for agricultural liming. Sensors include: (1) the Geophilus proximal soil sensing system to map apparent electrical resistivity (ERa) and natural soil-born gamma (γ) emissions; (2) the Soil pH Manager (Veris Technologies) to measure the active soil acidity. The derivation of soil texture maps based on ERa and γ data and first results that are important for a precise liming application are presented in this paper.

Keywords: soil maps, pH, proximal sensing, precision liming

Introduction

Liming is necessary for good nutrient availability and crop growth in acidic soils (Tunney *et al.,* 2010). Acidification of soils is caused by natural processes (carbon, nitrogen and sulphur – C, N, S cycles) and human activities (Holland *et al.,* 2018) and leads to negative effects on yield, as for example shown in Goulding (2016).
A study by Zimmer and Ellmer (2012) showed that only 26% of 56,320 soil samples taken in the state of Brandenburg between 2006 and 2009 have an optimal pH value, while 38% were too low (acidic) and 36% were too high (alkaline). These numbers indicate that, at present, the lime management on farms in Brandenburg is not sufficient. The problems farmers are facing include high soil pH variability based on a high soil heterogeneity within field sites, expenses and time required for soil sampling, uncertainties concerning the interpretation of soil information and fertilization decision making as well as problems related to the availability and handling of appropriate fertilization technology.
Based on the official guidelines of the Association of German Agricultural Investigation and Research Institutions (VDLUFA) (von Wulffen *et al.,* 2008), liming requirements are usually derived from 1 mixed soil sample per management unit of 3 ha. The mixed soil samples are tested for the current

DOI 10.3920/978-90-8686-888-9_67, © Wageningen Academic Publishers 2019

pH value and soil texture is derived by experts using a quick field texturing method. Based on this information, the liming requirement is defined by a look up table system for each management unit of 3 to 5 ha. However, this method implies several problems:

1. The average soil pH and soil texture as representation for the management unit neglect the variation of these parameters within the management units.
2. The field texturing method is less accurate than a lab analysis on soil texture, which can lead to uncertainties in defining the target pH value.
3. The soil texture classes in the VDLUFA system are rather coarse, distinguishing only 5 mineral soil texture classes and one class for peat soil, which can lead to unrealistic and severe differences in lime application recommendations of neighboring management units.

With respect to this and with regard to precision farming, farmers are confronted with the lack of high-resolution field maps of soil physical parameters as needed for decision support in acidity management. Publicly available soil information still does not reflect the small spatial variability of the soil properties present within the fields.

The European Innovation Partnership (EIP) AGRI funded project 'pH-BB' aims to enhance soil acidity management on farms in Brandenburg. In the pH-BB project, mobile soil sensors are applied to enable fast and cost-effective assessments of soil parameters like texture (sand, silt, clay content), soil organic carbon (SOC), and current pH value at high spatial resolution.

Therefore, two mobile multi-sensor platforms were used:

1. Soil texture information were derived based on data of the Geophilus system which measures the electrical resistivity (ERa) at 6 depths up to 1.5 m and the total counts of the natural gamma (γ) activity in the soil.
2. The mobile sensor platform (MSP) manufactured by Veris technologies was used to spatially derive the current soil pH, with the soil pH Manager.

The aim of the study was to develop a method that allowed an easy and automated generation of top soil texture maps based on mobile soil sensing data, which can be used for a lime management that respects the natural soil variability in a high level of detail and improves the currently available best practice as described above.

Materials and methods

In the project pH-BB, the mobile sensor based soil mapping and precision liming approaches were tested in collaboration with three agricultural farms in Brandenburg. In 2017 and 2018, several fields have been mapped. For this paper, field 6 of Komturei Lietzen has been chosen as a representative example of the pH-BB procedure.

Study area

Field 6 of the Komturei Lietzen (E 52.451551, N 14.135066) is located about 50 km east of Berlin, Brandenburg and has an area of 74 ha. Climatically, Brandenburg is located in the transition zone of the humid oceanic and the dry continental climates. The annual mean temperature is 9 °C; minimum temperatures are in January and maximum temperatures in July. The mean annual precipitation sum is 550 mm. Soil genesis is strongly affected by different ice ages. The test site is located in the young moraine landscape of the Weichsel glaciation and soils in this area are commonly characterized by well-established sand layers over loam (clay leaching) over marl (MUGV 2011).

Applied sensors

Most commonly used soil sensors map the bulk electrical conductivity (ECa), its reciprocal the electrical resistivity (ERa) or sample the soil's natural variation in γ activity. The soils ECa or ERa are affected by texture, water content, mineralogy, porosity, salinity, temperature, organic matter and bulk density (Corwin and Lesch, 2005). The influence of the soils' current moisture content,

which has on the one hand a strong influence on the soils' conductivity and resistivity, has on the other a very limited influence on the γ activity. Thus, a multiple sensor approach of ECa/ERa and γ measurements can improve the discrimination of different soil properties (Castrignanò *et al.,* 2012) and may provide significant advantages when compared to other proximal soil sensing methods (Mahmood *et al.,* 2013).

The applied experimental Geophilus-System (Lueck and Ruehlmann, 2013; Figure 1A) is a multiple sensor system that maps simultaneously the electrical resistivity (ERa) and the γ activity. It consists of seven rolling electrode pairs (one sending and six receiving) and measures the electrical resistivity in six layers from the soil surface up to a depth of 1.5 m. The attached γ probe measures soils γ activity (γ natural total counts) approximately in the upper 0.3 m layer. The system logs the sensor values each second at a preferred speed of 10 km/h. A dGNSS is used to geo-reference (incl. elevation) each logging.

The current pH values were mapped using Soil pH Manager (MSP) Veris technologies, (Salina, KS, USA). The pH values are derived from electrical voltage measurements of two ion selective antimony electrodes. A dGNSS signal is logged to determine the measurement locations. Further information can be found in Schirrmann *et al.* (2011).

Reference soil sampling

At 33 reference locations, soil samples (0-0.15 m depth) were taken and lab-analyzed for pH values, soil texture and soil organic carbon (SOC). Half of the sampling locations were chosen based on the pH mapping results and the other half based on the Geophilus mapping results. Reference samples were used to calibrate the MSP soil pH measurements and to regionalize soil textures using the Geophilus mapping results as covariables in multi-linear regression models.

Interpolation of ERa, γ and elevation data

The Geophilus system generates about 200 point measurements per hectare, which makes the interpolation with geostatistical methods like Kriging computationally very demanding and this method needs expert knowledge in the data preparation that farmers usually do nott have. Instead an inverse distance weighting (IDW) interpolation model (Shepard, 1968) was applied to the mapped ERa, γ and elevation point data to generate raster datasets. This technique has been widely used in the past and is probably one of the oldest spatial prediction methods (Hengl, 2009), it is computational less demanding and needs little expert knowledge. Raster datasets were created with a spatial resolution of 2 m for each parameter. Finally the dimensionless γ/ERa index (γEI) was calculated:

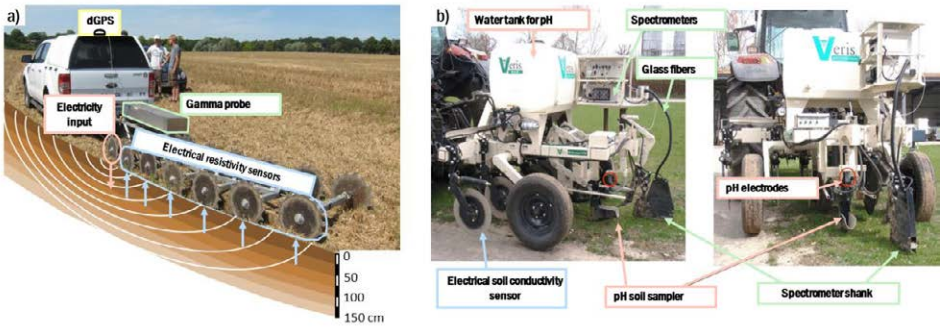


Figure 1. (A) The Geophilus system with 7 rolling electrode pairs and γ probe; (B) Mobile Sensor Platform (MSP) by Veris technologies with pH-manager, VIS-Nir spectrometer and OpticMapper (not shown).

$$\gamma EI = \gamma / ERa(_{Rho1}) \times 1000 \quad (1)$$

where γ is the γ raster dataset and $ERa(_{Rho1})$ the electrical resistivity raster of the smallest dipol array with an exploration depth of about 0-0.25 m.
The interpolated raster datasets were afterwards used as covariables for the regionalization of the soil textures.

Regionalization of soil textures
Raster values of the interpolated ERa, γ, γEI and elevation dataset were extracted at the reference locations of the 33 lab-analyzed soil samples, and values were added to the soil texture table. Based on this table, multi-linear regression (MLR) models were created, in a backward stepwise procedure. The prediction performance of applied MLR models was tested against the sampled soil physical properties. The model performance was assessed using a leave-one-out cross-validation (Isaaks and Srivastava, 1989) and the accuracy of each model was determined by the root mean square error (RMSE).
The German KA5 soil texture classification system (Eckelmann *et al.,* 2005) was applied to the raster datasets of clay, silt and sand content to create classified soil texture maps.

Regionalization of pH values
The mapped and calibrated soil pH values were regionalized using the Thiessen polygon method also known as the Voronoi diagram method. This regionalization method was applied and preferred over others like IDW or Kriging since extreme values fully remain within the regionalized result raster dataset. The extreme pH values are especially important for a precise acidity management that tries to improve the pH value at every location in the field towards its pH optimum.

Calculation of calcium oxide (CaO) amounts
The official German liming guideline of the VDLUFA (von Wulffen *et al.,* 2008), is a look up table system that allows farmers to determine an optimum pH based on the individual soil texture and organic matter (OM) classes present in the management unit. In comparison to the current pH status of the management unit, farmers can determine the application amount of CaO needed to change the pH-value towards the optimum.
In the pH-BB project, this lookup table system was used in a slightly modified approach with an automated R script to calculate the CaO application amount for every pixel of the resulting raster data sets (Figure 4 and 5). The modification implied the replacement of the soil texture and OM classes with clay and OM content.

Results

The field survey of the Geophilus-System and the Veris MSP was conducted at Field 6 on August 7th 2017. The interpolated mapping results are shown in Figure 2A-D and Figure 5A. The range in elevation is 18 m between the lowest point (66 m) in the southeastern part and highest point in the southwestern part (84 m) of the field (Figure 2D).
The yellow (bright) colored areas of the resistivity map >300 Ω/m (Figure 2A) represent very dry and/or very sandy areas while blue colors <50 Ω/m indicate areas with a high soil moisture content and/or high clay content. More or less corresponding to low-resistivity areas are the areas with a low γ activity (Figure 2B). Differences between the patterns of the two maps can be explained by the different soil moisture sensitivity of the two sensors. As a result, the γ/ERa index map (Figure 2C) may be interpreted as a soil moisture index map with yellow (bright) colored dry areas, blue (dark) colored moist areas and intermediate moisture in orange colored areas. The results of the 33 lab-analyzed soil samples (Figure 3) show that sand is the dominating fraction with a mean of

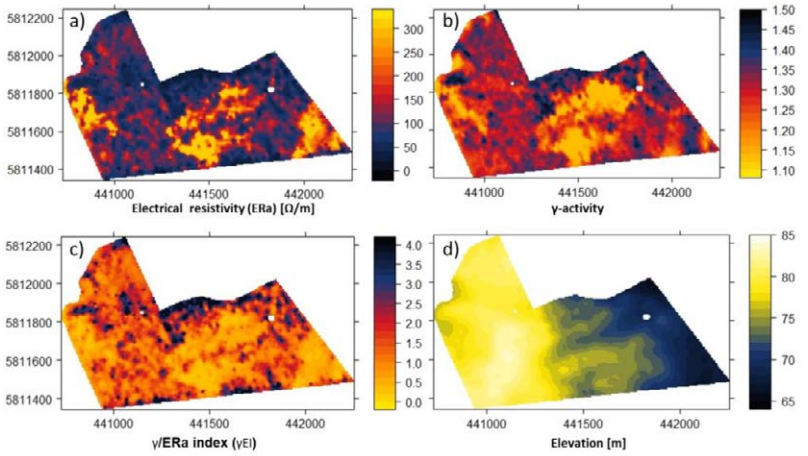


Figure 2. Mapping results of the Geophilus-System (A), electrical resistivity of the first channel (Rho1; 0-0.25 m) [Ω/m], (B) γ activity, (C) γ/Era index [dimensionless], (D) digital elevation model [m] recorded with DGPS.

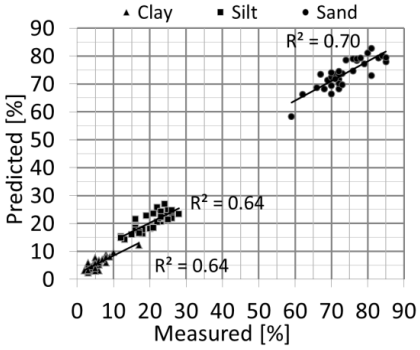


Figure 3. Scatterplots of the lab-analyzed reference soil texture samples and the predicted soil texture fractions clay, silt and sand.

74% and a range between 59 and 84%, while the clay content shows the lowest values with a mean of 5% and a range between 2 and 17%. The silt fraction reaches values between 12 and 28% and a mean of 21%. Field 6 is characterized by a low SOC content with a mean of 0.82% and a range between 0.5 and 1.3%.

The derived multi-linear regression models in Equations 2-4 show a good relationship between covariables and the target variables with a good adjusted R^2 between 0.64-0.70.

Multi-linear regression equations:

$$\text{Clay}[\%] = -5.84 + (253.4 \times \gamma EI) + (0.11 \times \text{Elevation}) \quad (2)$$

adj. R^2=0.64; α<0.001; RMSE=1.94

$$\text{Silt}\ [\%] = -49.05 + (33.76 \times \gamma + 0.035 \times \text{Elevation}) \quad (3)$$

adj. R^2=0.64; α<0.001; RMSE=2.92

Sand [%] = 116.00 + (-555.53 × γEI) + -0.46 × Elevation) (4)
adj. R^2=0.70; α<0.001, RMSE=3.7

where α is the significance level.
The RMSE is low for all fractions with values of 1.9 (clay), 2.9 (silt) and 3.7 (sand) and indicate a good prediction performance of the models. Derived regression models were afterwards applied on the interpolated Geophilus raster datasets and soil texture fraction were predicted for the entire field (Figure 4).
The predicted spatially distributed soil texture maps and the location of the 33 reference points are shown in Figure 4A-C with the same color range. The points are in good correspondence with the predicted soil texture maps. At the moment, the number of reference samples taken for the soil texture prediction is very high.
The project aims to reduce this number to a maximum of 5 reference samples, while keeping reliable prediction accuracy. The classified soil texture map in Figure 4D shows that the classes according to German KA5 system slightly silty sand (Su2) and slightly loamy sand (Sl2) cover the widest area of the field. The distribution of the class medium loamy sand (Sl3) is patchy but especially in the northern center and in the northwestern part bigger areas of this class can be identified. The classes of highly loamy sand (Sl4) and medium silty sand (Su3) are only visible in tiny patches of <0.03 ha.
The calibrated and interpolated pH values of the Veris MSP mapping are shown in Figure 5A. The results of the soil texture and pH mapping now allow different approaches to calculate liming recommendations.
Figure 5B shows the final CaO application amounts calculated based on the soil texture in Figure 4 and the current soil pH value in Figure 5A (at OM <4% assumed). CaO data was aggregated on a 12×40 m polygon raster in management direction (Figure 5B). Based on the CaO application map (Figure 5B) precise liming was applied with Granukal® lime using an AMAZONE ZA-TS with Amatron terminal on April 9th 2018.

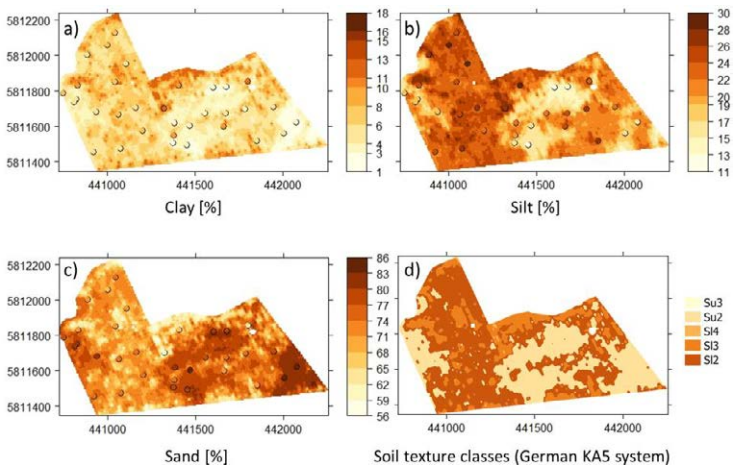


Figure 4. Predicted soil texture with reference soil sampling points, (A) clay content [%], (B) silt content [%], (C) sand silt content [%], (D) soil texture classes after the German KA5 system.

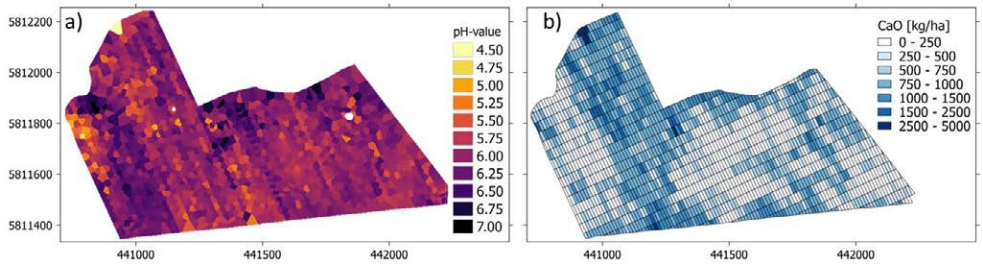


Figure 5. (A) MSP mapped pH values and (B) the CaO application amounts calculated for field 6 of Komturei Lietzen based on soil texture (Figure 4) and pH (Figure 5A) derived from proximal sensor data and aggregated to 12×40 m management units.

Conclusions

This paper shows how mobile sensor data of the Geophilus system together with mobile soil pH measurements of the Veris MSP system can be used to acquire spatially distributed precise soil texture and soil acidity maps. The regression models developed allow a soil texture prediction based on ERa and γ data with a high accuracy (RMSE<3.7) for all soil texture fractions. Therefore, these data can be used to calculate liming requirements that respect small scale soil and pH variabilities. The success of the precise liming will be controlled by repeated MSP pH mapping in the ongoing project.

References

Castrignano, A., Wong, M. T. F., Stelluti, M., De Benedetto, D., & Sollitto, D., 2012. Use of EMI, gamma-ray emission and GPS height as multi-sensor data for soil characterisation. Geoderma, 175, 78-89.

Corwin, D. L., & Lesch, S. M., 2005. Apparent soil electrical conductivity measurements in agriculture. Computers and Electronics in Agriculture, 46(1-3), 11-43

Eckelmann, W., Red.: Sponagel, H., Grottenthaler, W., Hartmann, K.-J, Hartwich, R. *et al.,* 2005. Bodenkundliche Kartieranleitung. (Soil scientific mapping manual) Ad-Hoc-AG-Boden, Bundesanstalt für Geowissenschaften und Rohstoffe (Hrsg.) 5, Hannover, Germany.

Goulding, K. W. T., 2016. Soil acidification and the importance of liming agricultural soils with particular reference to the United Kingdom. Soil Use and Management, 32(3), 390-399.

Hengl, T., 2009. A practical guide to geostatistical mapping (Vol. 52). Hengl. Science commons, Amsterdam. URL: http://spatial-analyst.net/book (last accessed February 2019).

Holland, J. E., Bennett, A. E., Newton, A. C., White, P. J., McKenzie, B. M., George, T. *et al.,* 2018. Liming impacts on soils, crops and biodiversity in the UK: a review. Science of the Total Environment, 610, 316-332.

Isaaks, E. & Srivastava, R., 1989. Applied Geostatistics. New York, USA: Oxford University Press.

Lueck, E., & Ruehlmann, J., 2013. Resistivity mapping with GEOPHILUS ELECTRICUS – Information about lateral and vertical soil heterogeneity. Geoderma, 199, 2-11.

Mahmood, H. S., Hoogmoed, W. B., & van Henten, E. J., 2013. Proximal gamma-ray spectroscopy to predict soil properties using windows and full-spectrum analysis methods. Sensors, 13(12), 16263-16280.

Ministerium für Umwelt, Gesundheit und Verbraucherschutz des Landes Brandenburg (MUGV) 2011. Steckbriefe Brandenburger Böden. (Fact sheets soils of Brandenburg). Potsdam, Germany

Schirrmann, M., Gebbers, R., Kramer, E., Seidel, J., 2011. Soil pH Mapping with an On-The-Go Sensor. Sensors 11, 573-598.

Shepard, D., 1968, January. A two-dimensional interpolation function for irregularly-spaced data. IN: Proceedings of the 1968 23rd ACM national conference, pp. 517-524. New York, USA: ACM

Tunney, H., Sikora, F. J., Kissel, D., Wolf, A., Sonon, L., & Goulding, K., 2010. A comparison of lime requirements by five methods on grassland mineral soils in Ireland. Soil Use and Management, 26(2), 126-132.

von Wulffen, U., Roschke, M., & Kape, H. E., 2008. Richtwerte für die Untersuchung und Beratung sowie fachlichen Umsetzung der Düngeverordnung (DüV): gemeinsame Hinweise der Länder Brandenburg, Mecklenburg-Vorpommern und Sachsen-Anhalt. (Guideline values for the investigation and advice as well as technical implementation of the Fertilizer Ordinance (DüV): common references of the federal states of Brandenburg, Mecklenburg-Western Pomerania and Saxony-Anhalt.) Bernberg, Germany: Landesanstalt für Landwirtschaft, Forsten und Gartenbau des Landes.

Zimmer, J., & Ellmer, F., 2012. Nährstoffversorgung ackerbaulich genutzter Böden im Land Brandenburg. (Nutrient supply of arable land in the state of Brandenburg). IN: Kommissionen IV und VI DBGGP, V., 92.

Effect of variable rate phosphorus and nitrogen fertilizing on winter wheat *(Triticum aestivum* L.*)* in Mezőföld, Hungary

G. Milics[1*], A. Vér[1], L. Szekeres[2] and J. Kauser[3]
[1]*Szechenyi Istvan University, Faculty of Agricultural and Food Sciences, Biosystems and Food Engineering, 2 Var, Mosonmagyarovar 9200, Hungary; milics.gabor@sze.hu*
[2]*Szekeres Ltd., 31. Batthyany st., Polgardi 8154, Hungary*
[3]*K-Prec Ltd., 98. Bela kiraly st., Piliscsaba 2081, Hungary*

Abstract

Variable rate technology (VRT) in nutrient management has been developed in order to apply crop inputs according to the required amount of fertilizers. There is an ongoing debate among experts on how to define management zones as well as how to define the required amount of fertilizers for phosphorus and nitrogen replenishment for winter wheat (*Triticum aestivum* L.) production. The objective of the study reported in this paper was to investigate the effect of variable rate phosphorus and nitrogen fertilizer application in winter wheat in Mezőföld, Hungary. Winter wheat production based on variable rate nutrient treatment resulted in 1.19 t/ha more yield than the farm average while applying an average 108 kg/ha less nitrogen and 37 kg/ha more phosphorus fertilizer.

Keywords: winter wheat, VRA phosphorus and nitrogen fertilizing

Introduction

Fields in Mezőföld region in Hungary are not uniformly structured in terms of soil types and fertility. Variable rate technology (VRT) in nutrient management allows application of a variable amount of fertilizer on each plot of land, depending on the initial content of soil nutrient. There is an ongoing debate among experts and practicing farmers on how to define management zones, as well as how to define the required amount of fertilizers for phosphorus and nitrogen replenishment for winter wheat (*Triticum aestivum* L.) production.
In order to manage within-field differences, management zones have to be defined. For this reason, management zone delineation techniques are used (Peralta *et al.*, 2015).
The basis of all management zone delineation (MZD) techniques is the grouping of data values into similar classes so that the members of the class are closer in relation than to members of other classes (Lark, 1998). A management zone, therefore, is a portion of a field that is more homogenous than the overall field based on a certain measurable characteristic (Zhang *et al.*, 2002). Various authors have proposed criteria for the delineation of management zones (Mulla, 1991; Basso *et al.*, 2001; Fleming *et al.*, 2001; Franzen *et al.*, 2002). Depending on the available input data, management zones can be decided in practice with relatively high accuracy.
Variable rate application of nitrogen and phosphorus increases the efficiency of winter wheat production (Raun *et al.*, 2001; Raun *et al.*, 2002). There is a general perception that P fertilizer use is very inefficient because P recovery by the crop rarely exceeds 25% in the year of application and more often is only 10-15% of the P applied. The P not taken up by the crop (i.e. the residual P) has been believed to be irreversibly fixed in the soil in unavailable forms, but those views are changing as agronomic field experiments have shown that that is not the case (Johnston *et al.*, 2014).
One approach of nitrogen fertilizing is to split applications between fall and spring. In Hungary, nitrogen fertilizer for winter wheat is generally applied when the soil is still cold in early spring (usually early March), and top-dressing is carried out in April in order to ensure a sufficient amount of nutrients for the plants throughout the growing season. Crop sensing technology is able to 'fine-

DOI 10.3920/978-90-8686-888-9_68, © Wageningen Academic Publishers 2019

tune' the amount of variable rate nitrogen fertilizer (Schwalbert *et al.*, 2019), however, it requires sensors and machinery for precise application.
Applying a spatially variable rate of nitrogen is challenging because it deals with the adoption of site-specific practices that aim at maximizing crop N uptake, minimizing N losses, and optimizing the indigenous soil's N supply (Basso *et al.*, 2011).
The economic optimum rate of nitrogen fertilizer varies for each geographical location. However, most studies point out that nitrogen application rates are much higher than they should be. For example, Stamatiadis *et al.* (2017) reported that in their high spatial resolution study, VRA delivered an economic optimum N rate using 72% less in-season N or 38% less total N (131 kg N/ha) than that applied by the farmer (212 kg N/ha) in practice.
The objective of the work reported here was to investigate the effect of variable rate phosphorus and nitrogen fertilizer application in winter wheat (*T. aestivum* L.) in Mezőföld, Hungary.

Materials and methods

Location
The investigated field was a 40 ha plot of land planted with winter wheat (*T. aestivum* L.) near Polgárdi, Mezőföld, Hungary (N 47°03'11', E 18°17'17') (Figure 1).
The previous crop was autumn rape in 2017 (Figure 2A). Within the last 5 years, winter wheat had been grown in 2013 (Figure 2B) and 2016. Maize yield data were collected from the field in 2013, 2014. Yield map for 2016 is not available due to sensor failure.
Management zone delineation was based on remotely sensed data segmentation, several year's yield map analysis and farmer's experience. Based on the data, seven management zones were determined (Figure 3). Areas without assigning a zone to them were treated according to the long-term, conventional management practice of the farm.

Instrumentation
Potassium fertilizing was carried out by a Bogballe M3W (Bøgballe A/S, Bøgballe, Uldum, Denmark) fertilizer spreader. Cultivation was carried out by a Bednar Terraland 4000 (Bednar FMT s.r.o., Rychnov nad Kněžnou, Czech Republic) deep cultivator. Seeding was carried out by a Case IH 240

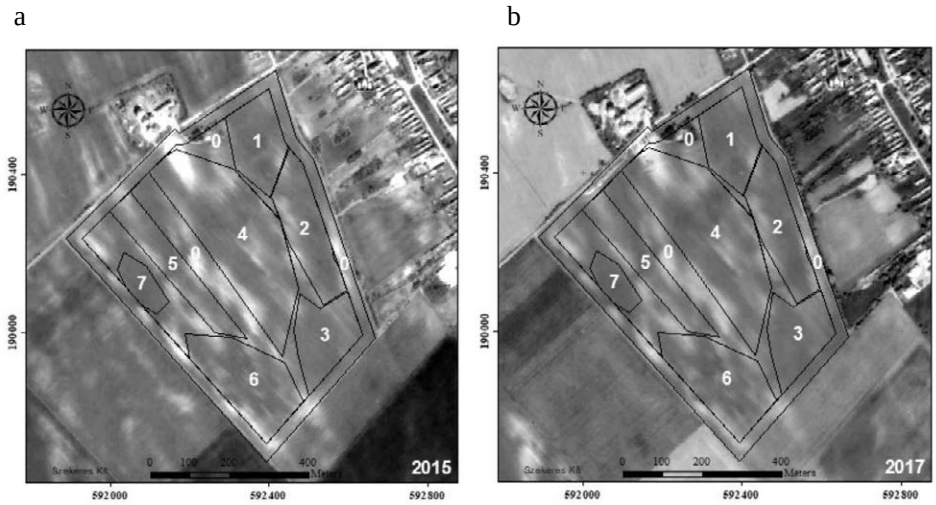


Figure 1. Satellite image of the investigated field in 2015 (A) and 2017 (B) (Source: Google Earth).

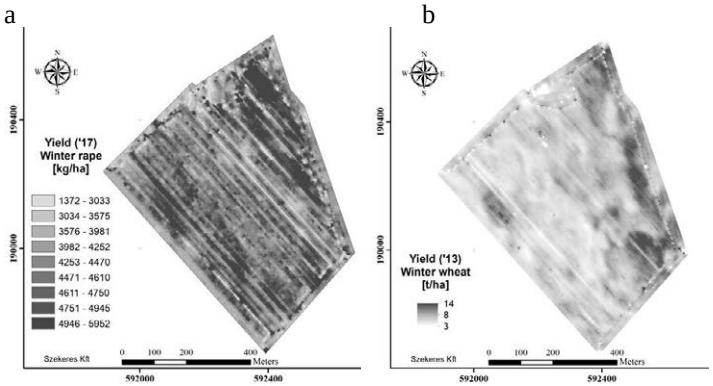


Figure 2. Yield map of winter wheat (2013) (A) and previous crop winter rape (2017) (B).

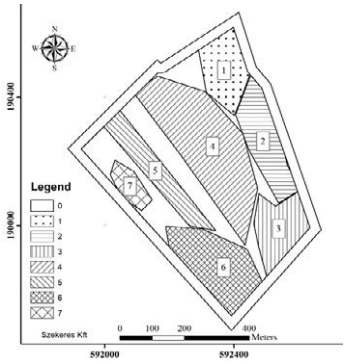


Figure 3. Management zones in the investigated field.

CVX (CNH Industrial Austria GmbH, St Valentine, Austria) tractor and a Horsch Pronto 6DC (Horsch Maschinen GmbH, Schwandorf, Germany) precision seeding machine on 25 September, 2017, the earliest possible date. Weed control was carried out in autumn by a Dammann DT2000 (Herbert DAMMANN GmbH, Buxtehude-Hedendorf, Germany) fertilizer. Variable rate nitrogen fertilizer replenishment was carried out by a Bogballe (Bøgballe A/S, Bøgballe, Uldum, Denmark) fertilizer spreader. For top-dressing a Dammann fertilizer was used.
Yield estimation was carried out based on the data collected by a Topcon CropSpec (Topcon Positioning Systems, Inc., Livermore, USA) instrument using a Fendt 716 (AGCO GmbH, Marktoberdorf, Germany) tractor. For harvesting, a Case IH 8120 (CNH Industrial America, LLC., USA) combine harvester was used. For control and data collection, an AFS700 (CNH Industrial America, LLC., USA) monitor was installed in the machines.

Soil sampling and experiment
The soil type is calcareous chernozem on a loess base. Soil samples were collected from each management zone and sent for laboratory analysis. Potassium (60%K_2O) was applied in a uniform amount of 60 kg/ha prior to seeding. Phosphorus and nitrogen rates for each management zone were determined by K-Prec Ltd. (Piliscsaba, Hungary) advisory system. Phosphorus was applied in the form of Monoammonium phosphate (MAP 12N:52P) during seeding in late September, the

prescribed rates were between 136 and 350 kg/ha. Winter wheat is treated as a highly intensive crop on the experimental field, therefore, splitting the amount of nitrogen fertilizer is practiced. Variable rate nitrogen was applied in the form of Calcium ammonium nitrate (CAN 27%N) in early March, the prescribed rates were between 240 and 300 kg/ha (Table 1, Figure 4). Top-dressing was applied in early April in the form of Nitrosol (UAN 30%N) at a uniform 180 kg/ha.
During the vegetation period of 2018, yield estimation was carried out before harvest with the help of a Topcon CropSpec (Topcon Positioning Systems, Inc., Livermore, USA) canopy greenness sensor aiming to monitor in-season differences within the investigated field. Differences between the management zones were visible on the map created by the sensor (Figure 5).
Estimation was carried out on 10 May for each management zone by counting the number of harvestable heads and kernels on 1 m^2.

Results

The planned yield was between 10.5 and 12 t/ha. Due to the drought in April and May, the yield was predicted at a much lower level during crop canopy measurements and counting for each

Table 1. Variable rate phosphorus and nitrogen application rates (kg/ha) for the management zones.

Management zone	MAP (12N:52P) [kg]	CAN (27%N) [kg]	Planned yield [t/ha]
MZ 0	200	370	9
MZ 1	136	240	12
MZ 2	215	248	12
MZ 3	313	249	12
MZ 4	197	255	11
MZ 5	350	269	10.75
MZ 6	270	274	10.5
MZ 7	179	300	12
Average	237	262	11.5
Control	200	370	9

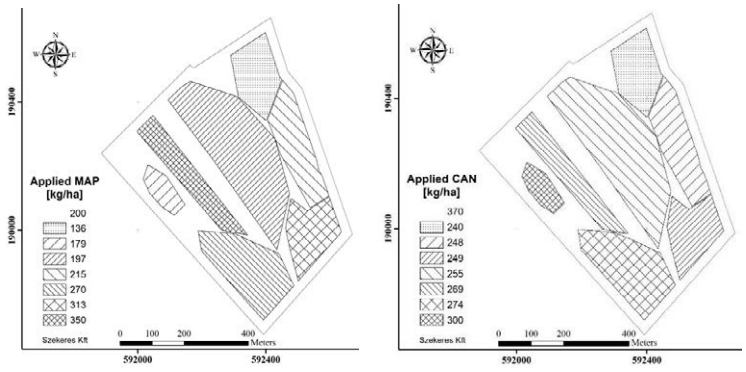


Figure 4. The applied amount of variable rate phosphorus (MAP) and nitrogen (CAN).

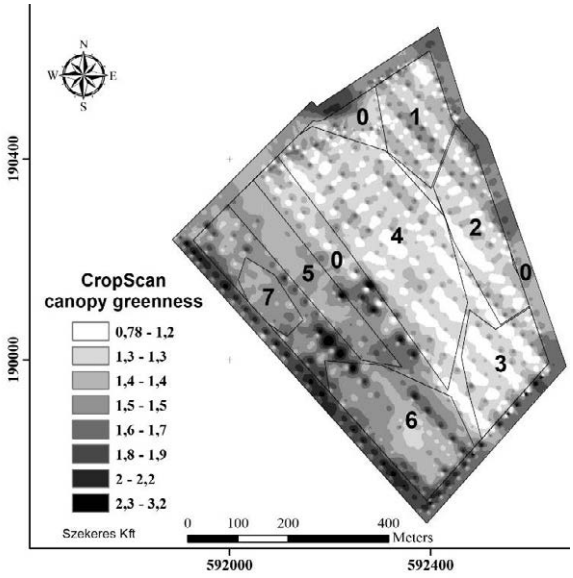


Figure 5. Topcon CropSpec canopy greenness map.

management zone (8.9-10.3 t/ha). Measured yield data showed even lower values (8.13-9.31 t/ha) (Table 2, Figure 6). The yield in the commercially treated control area (MZ 0) was 7.76 t/ha, whereas the lowest and highest yield in the site-specifically treated management zones were 0.2-1.38 t/ha above the control. The farm average for winter wheat production was 7.76 t/ha, which is 1.19 t/ha lower than the average 8.95 t/ha in the experimental field.

Discussion

Winter wheat production in Mezőföld, Hungary, currently has great potential. One important goal of the farmers in the region is a simultaneous increase of the yield and a decrease in the environmental effects by decreasing the amount of nitrogen applied. Variable rate application of

Table 2. Planned, estimated and measured yield in the investigated field.

Management zone	Planned yield [t/ha]	Estimated yield [t/ha]	Measured yield [t/ha]
MZ 0	9	8.5	7.76
MZ 1	12	10.3	9.46
MZ 2	12	10.1	9.37
MZ 3	12	10.2	9.58
MZ 4	11	9.35	8.81
MZ 5	10.75	9.1	8.18
MZ 6	10.5	8.92	8.25
MZ 7	12	10.2	8.97
Average	11.5	9.7	8.95
Control	9	8.5	7.76

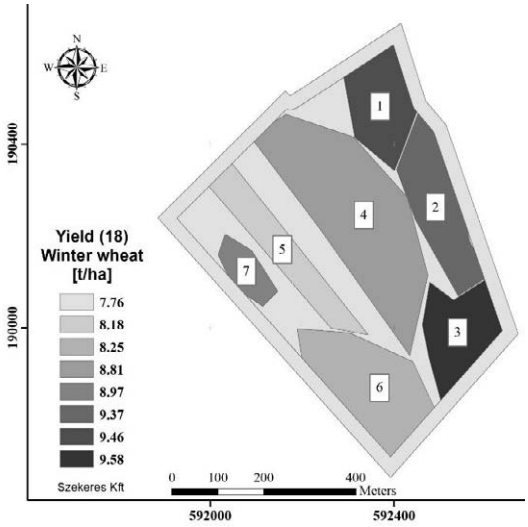


Figure 6. Yield map of winter wheat (2018) related to the management zone.

phosphorus and nitrogen is a key element for applying only the required amount of fertilizers. The experimental field was divided into seven management zones. Delineation was based on remotely sensed data segmentation, several year's yield map analysis and farmer's experience. Differences in the topographic location and soil properties required variable rate fertilizer application. In the experimental field, winter wheat production was on average 8.95 t/ha for the 40 ha field compared to the farm average of 7.76 t/ha. It has to be emphasized that this yield was achieved by reducing the amount of N fertilizer applied by slightly more than 100 kg/ha for the experimental field (Figure 7).

Conclusions

Variable rate phosphorus and nitrogen fertilizing in Mezőföld, Hungary, positively affected the yield in winter wheat production. Management zone delineation was carried out correctly on the investigated field. Each management zone produced a higher yield than the farm average despite the unfavourable conditions compared to the other fields of the farm. Variable rate phosphorus and

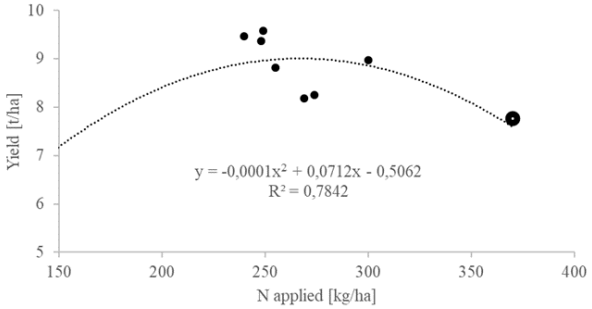


Figure 7. Effect of variable rate N fertilizing in the experimental field (The highlighted dot shows conventional (370 kg/ha N) farm practice).

nitrogen fertilizer application had a positive effect on the winter wheat yield. Despite the shortage of rainfall in the vegetation period, each management zone produced a higher yield than the farm average.

Acknowledgements

The authors would like to thank Szekeres Ltd. and K-Prec Ltd. for providing the data and equipment for the research. The authors carried out this research as part of the Project Networking European Farms to Enhance Cross Fertilisation and Innovation Uptake Through Demonstration (NEFERTITI) 772705 Hungarian HUB. This project was partly supported by the EFOP-3.6.3-VEKOP-16-2017-00008 project. The project is co-financed by the European Union and the European Social Fund.

References

Basso B., Ritchie J.T., Cammarano D., Luigi Sartori L., 2011. A strategic and tactical management approach to select optimal N fertilizer rates for wheat in a spatially variable field. European Journal of Agronomy 35 215- 222.

Basso B., Ritchie J.T., Pierce F.J., Braga R.P., Jones J.W., 2001. Spatial validation of crop models for precision agriculture. Agricultural Systems 68, 97-112.

Fleming K.L., Westfall D.G., Wiens D.W., Brodahl M.C., 2001. Evaluating farmer defined management zones for variable rate fertilizer application. Precision Agriculture 2, 201-215.

Franzen D.W., Hopkins D.H., Sweeney M.D., Ulmer M.K., Halvorson A.D., 2002. Evaluation of soil survey scale for zone development of site specific nitrogen management. Agronomy Journal 94 381-389.
https://doi.org/10.1007/s11119-017-9540-7

Johnston A.E., Poulton P.R., Fixen P.E., Curtin D., 2014. Phosphorus: its efficient use in agriculture. Advances in Agronomy, 123 177-228.

Lark R. M., 1998. Forming Spatially Coherent Regions by Classification of MultiVariate Data: An Example from the Analysis of Maps of Crop Yield. International Journal of Geographical Information Science, 12 (1), 83-98. doi:10.1080/136588198242021

Mulla D.J., 1991. Using geostatistics and GIS to manage spatial patterns in soil fertility. In: Kranzler, G. (Ed.), Proceedings of the Automated Agriculture for the 21st Century. St. Joseph, MI, USA: ASAE,

Peralta N.R., Costa J.L., Balzarini M., Franco M.C., Córdoba M. and Bullock D., 2015. Delineation of management zones to improve nitrogen management of wheat. Computers and Electronics in Agriculture 110, 103-113.

Raun, W.R., Solie J.B., Johnson G.V., Stone M.L., Lukina E.V., Thomason W.E., 2001. In-season prediction of potential grain yield in winter wheat using canopy reflectance. Agronomy Journal 93, 131-138.

Raun, W.R., Solie J.B., Johnson G.V., Stone M.L., Mullen R.W., Freeman K.W., 2002. Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. Agronomy Journal 94, 815-820.

Schwalbert, R.A., Amado, T.J.C., Reimche, G.B., 2019. Precision Agriculture 20 (1) 56-77 https://doi.org/10.1007/s11119-018-9581-6

Stamatiadis S., Schepers J.S., Evangelou E., Tsadilas C., Glampedakis A., Glampedakis M., *et al.,* 2018. Variable-rate nitrogen fertilization of winter wheat under high spatial resolution. Precision Agriculture 19 (3) 570-587

Zhang N., Wang M., Wang N., 2002. Precision agriculture – a worldwide overview. Computers and Electronics in Agriculture. 36 (2-3) 113-132.

Precision nitrogen and water management for maize production in the western great plains of the US

E. Phillippi[1], R. Khosla[1*], A. Andales[1] and L. Longchamps[2]
[1]*Colorado State University, Fort Collins, CO 80523, USA; raj.khosla@colostate.edu*
[2]*Agriculture and Agriculture Food Canada, St-Jean-sur-Richelieu, Quebec, Canada*

Abstract

Nitrogen and water continue to be the most limiting factors for profitable maize production in the western Great Plains. The objective of this research was to determine the most productive and efficient nitrogen and water management strategies for irrigated maize. This study was conducted in 2016 and 2017 in Colorado, USA. Treatments included, six N fertilizer rates of 0, 56, 112, 168, 224, 280 kg/ha by the V8 growth stage. Four rates of irrigation were applied at 60, 80, 100 and 120% of evapo-transpiration throughout the growing season. A significant (P<0.1) and positive grain yield response to increasing levels of N was observed only in 2016. A significant effect of irrigation (P<0.1) was observed across the field and within zones in both years.

Keywords: nitrogen, water, management zones

Introduction

Nitrogen (N) is among the most limiting nutrients in maize production and must be applied in adequate amounts to achieve maximum yield potential. Research has shown that, site-specific management zones (SSMZ) for variable rate application of N are a simple and effective way to increase nitrogen use efficiency and mitigate N pollution of the environment (Inman *et al.*, 2005; Khosla *et al.*, 2002; Delgado *et al.*, 2005). Water continues to be the most limiting factor for profitable maize production in the western Great Plains. Irrigated crop growers continually face many challenges related to water management notably: competition for water from increasing municipal demand and threats to the stability of water supply due to increasing severity and likelihood of drought (Tilman *et al.*, 2011). Precision irrigation management systems have the potential to significantly reduce water consumption by crops, by applying the right amount of water, at the right place, and at the right time without impairing grain yields. It is necessary to develop precision irrigation techniques to determine and apply optimal rates of water at every location in a crop field to maximize yield, water use efficiency and potentially conserve water. Additionally, the interaction between N rates evapotranspiration is crucial in achieving high grain yields. As high rates of applied N may result in increased water uptake at full irrigation when compared to low rates of N at limited irrigation. The overall objective of this research was to determine the most productive and efficient nutrient and water management strategy for irrigated maize. The specific objectives were:

1. To understand crop response to variable-rate nitrogen across the field and within management zones.
2. To understand variable rate irrigation water management across the field and within management zones.

Materials and methods

This study was conducted in 2016 and 2017 in Colorado, USA at the Agricultural Research, Development and Education Center near Fort Collins (40°40' 38.24" N, 104° 58' 44.76" W). The site has been under continuous maize cropping system with a center pivot irrigation system equipped

DOI 10.3920/978-90-8686-888-9_69, © Wageningen Academic Publishers 2019

with GNSS, telemetry, variable drive motor, and drop nozzles for variable rate application of water. Soils at the site are classified as a fine-loamy, mixed, super active, mesic, Aridic Haplustalf (Soil Survey Staff, 1980). The two soil series at the study site are a Kim loam and Nunn clay loam. Site specific management zones were previously delineated on the study field using the method described by Khosla *et al.*, 2002, and classified into low, medium, and high zones, or zones 1, 2, and 3 respectively. The experimental design consisted of a completely randomized block design with multiple treatment strips per block. Where, treatment strips were planted across the entire length of the field in the East to West direction such that the treatment strips traversed over each management zone, and such that, all treatment combinations were accomplished. Maize hybrid, DKC 46-20 (Dekalb) was planted with a six row John Deere precision vacuum planter (Deere & Company, Moline, IL, USA) at 76 cm row spacing in both years. In 2016 five N rates of 0, 56, 112, 168, 224 kg/ha using 32% Urea Ammonium Nitrate (UAN) were side-dressed by dribbling fertilizer in between the maize rows at the V8 crop growth stage (Ritchie, *et al.*, 1993). In 2017, an additional N rate of 280 kg/ha was also applied. A high clearance Lee Spider tractor equipped with DGPS unit was used to side dress UAN fertilizer (Lee Agra Inc., TX).

Based on the input received from the collaborating farmers, three rates of irrigation, 80, 100, and 120% of evapo-transpiration (ET) were applied in 2016. The irrigation rates were revised to 60, 80, and 100% of ET for 2017. Experimental layout was designed such that some or all of the levels of irrigation were randomly applied to clusters of experimental plots within each strip to achieve combinations of every factor level (i.e. N rate, management zone, and irrigation rate). The online tool, Water Irrigation Scheduling for Efficient Application (WISE) was used to schedule and determine the quantity of water applied for irrigation events (Andales *et al.*, 2014). The W.I.S.E tool uses a soil water balance approach by employing weather and soil data that is available online, to estimate daily crop ET using the ASCE Penman Monteith equation and plant available water in the soil throughout the growing season. Grain was harvested on October 21st and November 1st in 2016 and 2017 respectively, with a yield monitor equipped combine. The yield data was corrected for errors as described by Khosla and Flynn (2008), for subsequent statistical analysis.

Statistical analysis

Least squares regression analysis was performed to study the main effects of water, nitrogen, zone, and their interaction. Due to the complexity of this field experiment it was determined that a *P*-value of 0.1 was deemed appropriate. As has been the case in other studies at the discretion of the researchers (Cherry, 1998; Karlan *et al.*, 2006; Miller *et al.*, 2008). The factors and their levels were defined as irrigation rate (60, 80, 100, 120% of evapotranspiration estimated for the study site), N Rate (0, 56, 112, 168, 224 and 280 kg/ha), and SSMZs (1, 2, and 3). The response variable was the harvested maize grain yield (Mg/ha). A linear mixed model was fit to grain yield data where the study factors were assigned as fixed main and two-way interaction effects in the model. Random effects included were the experimental strip; and experimental strip by SSMZ combinations to account for the experimental design structure. F-tests were generated for all fixed effects. Post-hoc analysis included estimation of marginalized means, contrasts, and simple effect comparisons. Data analysis was done using R for Windows software, Version 3.3

Results and discussion

Weather

Northeastern Colorado is a semi-arid environment and the study site is located in the rain shadow of the Front Range of northeastern Colorado. As presented in Figure 1, the 2016 crop growing season experienced unusually hot and dry weather. The temperatures were 4.5% higher (1.5 °C) than the historical average during July and August, critical time period for the late vegetative and reproductive growth stages. Also, slightly more than 50 mm of precipitation is received on average historically

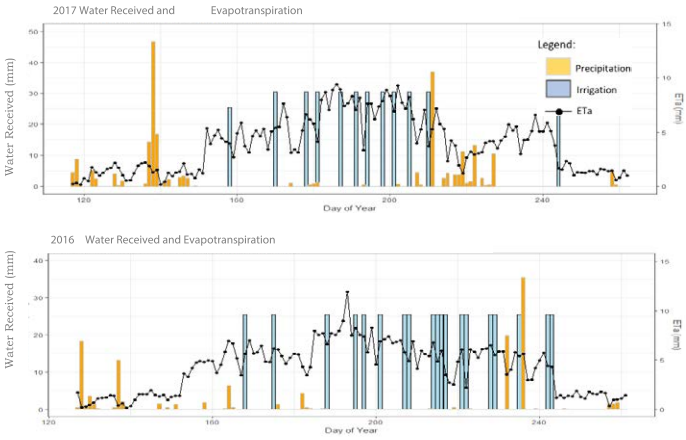


Figure 1. Bar and dot and line plots of water received at Field 3,100 in 2016 and 2017. Different colored bars indicate water received as well as type (irrigation or precipitation). Dot and lines indicate estimated daily actual evapotranspiration.

at the study site in the month of July, however, essentially 0 mm of precipitation occurred in 2016. This is in contrast with 2017, when temperatures were similar to the historical average but with large amounts of precipitation occurring early and later in the growing season. Mid-May 2017 saw blizzard conditions with nearly 100 mm of snow falling and a short period of below freezing temperatures. In late July to mid-August 2017, 46% more precipitation occurred relative to the historical average. As is logically expected, unusual weather events during the 2016 and 2017 growing season likely played a significant role in grain yield response to fertilizer and water.
The short-season maize hybrid (DKC46-20) accrued 2503 growing degree units from planting to harvest in 2016 and 2,483 growing degree units in 2017 (Bauder *et al.*, 2003).

Nitrogen
The findings of the N rate study treatments were both expected and unexpected. As expected, there was a significant (P<0.1) and positive grain yield response to increasing levels of N across all site-specific management zones in 2016 (Table. 1). Where, mean grain yield varied from a low, 7.1, to a high, 11.4 Mg/ha across all N treatments. Unexpectedly, no significant grain yield response was observed across or within site-specific management zones with increasing levels of N applied in 2017 (Table 1). Hence, N results for 2017 will not be discussed further.
In 2016, statistical comparison using least squares means pursuant to objective (i) for grain yield across all levels of N treatments, show significant differences between 0 and all higher rates of N. These differences range from 1.9 Mg/ha between 0 and 56 kg N/ha N rate treatment to 4.3 Mg/ha between 0 and 224 kg/ha N rate treatment. Likewise, a significant mean grain yield increase of 2.0 Mg/ha or more was observed for N rates of 168 and 224 kg/ha when compared to the 56 kg/ha N treatment. No significant differences were detected in grain yield among N rates of 112, 168, and, 224 kg/ha. These findings are similar to a previous study conducted in the eastern plains of Colorado (Al-Kaisi and Yin, 2003). It was reported that no significant grain yield advantage occurred at N rates above 140 kg/ha where variable rate irrigation was also employed.
Analyses performed on grain yield for N treatments within each management zone also display a positive yield response to increasing levels of N. For Zone 1 mean grain yield varied from a low, 7.0 to a high, 11.6 Mg/ha; for Zone 2 mean grain yield varied from a low, 7.2 to a high, 11.5 Mg/ha;

Table 1. Mean grain yield for corresponding rates of nitrogen applied averaged over irrigation and SSMZ treatments in this study in 2016 and 2017.[1]

Year	N-rate (kg/ha)	Mean grain yield (Mg/ha)
2016	0	7.1[a]
	56	9.0[b]
	112	9.9[bc]
	168	11.0[c]
	224	11.4[c]
2017	0	10.7[a]
	56	11.1[a]
	112	11.8[a]
	168	11.1[a]
	224	12.1[a]
	280	12.1[a]

[1] Different letters (a, b and c) indicate significant differences at alpha=0.90 using Tukey's HSD test.

and for Zone 3 mean grain yield varied from a low, 7.1 to a high, 11.1 Mg/ha. However, zone effect was not significant (P>0.1). As logically expected, the lowest mean grain yield occurred at 0 kg/ha N treatment. Maximum mean grain yield in zones 1 and 3 occurred at the 224 kg/ha N treatment. While zone 2 attained highest yield at the 168 kg/ha N treatment. However, no significant interactions (P>0.1) between zone and N treatments were observed.

The lack of interaction between N application rates and SSMZs appears to be anomalous, as long-term research has documented that the variable rate N fertilizer optimizes and/or maximizes yield across management zones (Inman *et al.*, 2005; Khosla *et al.*, 2002; Hornung *et al.*, 2006). However, observations in this study are not unique. A similar study conducted in Nebraska that employed site-specific SSMZ's and variable rate N with irrigated maize also reported no significant differences in grain yield between a uniform application of N and SSMZ approach (Ferguson *et al.*, 2002). There are multiple reasons that could be attributed to such an observation. The year 2016 was an unusually dry year, where only 16% of the crop water needs was met by precipitation. Field observations made during the season indicate periods of crop water stress in between irrigation cycles. It is worth noting that the management zones currently employed at this study site were delineated prior to 2012 under furrow irrigation system, where irrigation water moved from West to East direction of the field. This was also reflected in the productivity potential of management zones delineated on the field as they followed a West to East gradient, where East part of the field was the most productive or designated as high management zone. In 2012 this field underwent significant change when it was precision leveled from a previous field-gradient of 3% to a gentle slope of <1% to accommodate a transition from furrow irrigated system to a precision irrigated center pivot system. It is postulated that the management zones may have shifted both, spatially and temporally due to the change in irrigation methodology. To date, minimal research has been conducted to determine how and if zones shift spatially overtime and if there exists a necessity to reclassify fields as these shifts may occur. More research is needed to reclassify this field and similar other fields for most efficient maize production systems. In addition, studies (Seifert *et al.*, 2017) have reported yield reductions associated with continuous maize cropping system which may limit the efficacy of employing SSMZ's. Interestingly, a review of literature indicated no previous studies evaluating interaction between cropping system and SSMZ's.

Water

There was a significant main effect of irrigation treatment ($P<0.1$) in both 2016 and 2017. The irrigation rates of 80, 100, and 120% of ET resulted in mean grain yield of 7.6, 9.6, and 10.0 Mg/ha respectively in 2016 (Table 2). While in 2017, mean grain yield increased from a low of 11.1 Mg/ha to a high of 12.0 Mg/ha at rates of 60 and 80% ET respectively. There was no significant difference in grain yield at the 80 and 100% of the ET irrigation rate in 2017. Similar observation has been previously reported in another study by Al-Kaisi and Yin in 2003.

Additionally, zone and irrigation interaction was found to be significant ($P<0.1$) in both years as presented in Figure 2. In 2016, At the 100% of ET rate, a grain yield advantage of up to 1.5 Mg/ha was observed when comparing zones 1 and 2 to zone 3. When irrigation application rate was increased to 120% of ET, zones 2 and 3 outperformed zone 1 by up to 1 Mg/ha. Similarly, in 2017, the 80% ET rate resulted in the highest and significantly different mean grain yield response in both zones 1 and 2. While zone 3 achieved maximum yield at the highest rate of 100% ET.

These results seem to confirm that grain yield can be maximized, or significant losses avoided when a slight deficit irrigation approach is undertaken (Rathore *et al.*, 2017; Stone *et al.*, 2010; Zhang, 2003). As indicated previously, the 2016 growing season was hot and dry. Cumulative irrigation water applied at the 100% of ET treatment was 533 mm. It has been documented that the ASCE

Table 2. Mean grain for corresponding three rates of irrigation applied averaged over Nitrogen and SSMZ treatments in this study in 2016 and 2017.[1]

Year	Irrigation rate (%ET)	Mean grain yield (Mg/ha)
2016	80	7.6[a]
	100	9.6[b]
	120	10.0[b]
2017	60	11.1[a]
	80	12.0[b]
	100	11.7[b]

[1] Different letters (a and b) indicate significant differences at alpha level of 0.90 using Tukey's HSD test.

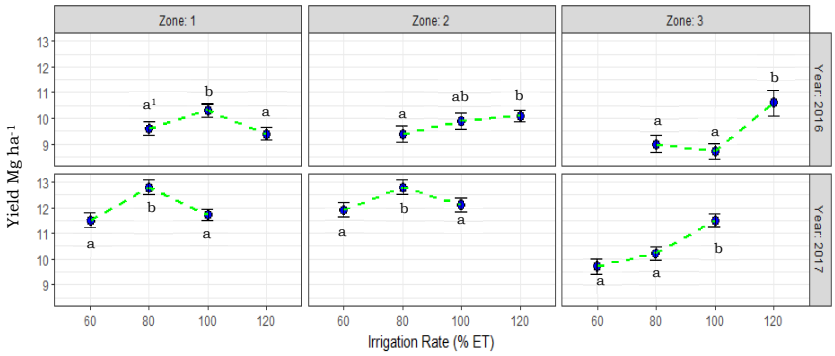


Figure 2. Irrigation rates for each management zone, and corresponding mean grain yield along with standard error of the means observed in this study. Different letters (a and b) indicate significant differences at alpha level of 0.90 within management zones across irrigation rates using Tukey's HD test.

Penman Monteith equation for calculating reference ET, used by WISE to estimate crop ET, can result in underestimations of ET in hot advective environments (Gowda *et al.*, 2008). Thus, the irrigation applied could have been misappropriated such that the 120% ET rate was in effect the 100% of the ET rate. However, in both years, zone 3 mean grain yield was most sensitive to the highest irrigation rate. Where in 2016 there was up to a 22% mean grain yield advantage when irrigation was increased to 120% ET. In 2017, a mean grain yield advantage of up to 19% was observed when irrigation was increased to 100% ET. Clearly more research is needed to evaluate such aspects of irrigation when employing SSMZ's.

Conclusions

The research objectives for this study were to examine the most productive and efficient nutrient and water management strategies coupled with site-specific management zones. Overall, the findings of this study reaffirm that macro-scale spatial variability can be managed by employing SSMZs along with variable rate fertilizer and irrigation. The extreme weather events that occurred during both growing seasons coupled with a shift in zones likely accounted for the lack of zone and SSMZ interaction in both years. More study is needed to examine the causes and timing of how a shift in zones may occur and the potential impact on variable rate application of N and water. Encouragingly, a moderate deficit irrigation did not significantly decrease grain yield relative to the largest level of irrigation water applied both within zones and across the entire field. Moreover, the study has demonstrated the challenges associated with extreme weather and zone shifts as it relates to employing variable rate technologies to maintain or increase grain yield.

References

Al-Kaisi, M., and Yin, X., 2003. Effects of nitrogen rate, irrigation rate, and plant population on corn yield and water use efficiency. Agronomy Journal, *95*(6), 1475-1482.

Andales A., Bauder T., and Arabi M., 2014 A mobile irrigation water management system using a collaborative GIS and weather station networks. In L.R. Ahuja, L. Ma, R. Lascano (Eds.), Practical Applications of Agricultural System Models to Optimize the Use of Limited Water, (pp. 53-84). Madison, WI, USA: ASA, CSSA, SSSA.

Bauder T., Waskom R., Schneekloth J., and Alldredge J., 2003. Best Management Practices for Colorado Corn. Colorado State University Cooperative Extension; Colorado Corn Growers Association-Colorado Corn Administrative Committee

Cherry, S., 1998. Statistical tests in publications of The Wildlife Society. Wildlife Society Bulletin, 26(4), 947-953.

Delgado, J., Khosla, R., Bausch, C., Westfall, D., and Inman, D., 2005. Nitrogen Fertilizer Management Zones Reduces Potential for Nitrate Leaching. Journal of Soil and Water Conservation 60: 402-410.

Ferguson R., Hergert G., Schepers J., Gotway S., and Cahoon, T., 2002. Site-specific nitrogen management of irrigated maize: yield and soil residuals nitrate effects Soil Science Society of America Journal, 66, 544-553.

Gowda, P., Chavez, H., Colaizzi, J., Evett, L., Howell, P., and Tolk, D., 2008. ET mapping for agricultural water management: Present status and challenges. Irrigation Science, 26(3), 223-237.

Hornung, A., Khosla R., Reich R., Inman D. and Westfall D.G., 2006. Comparison of Site-Specific Management Zones. Agronomy Journal 98 (2) 405-417.

Inman, D., Khosla R., and Westfall D., 2005. Nitrogen uptake across site-specific management zones in irrigated corn production systems. Agronomy Journal 97. 169-176.

Karlen, D.L., Hurley, E.G., Andrews, S.S., Cambardella, C.A., Meek, D.W., Duffy, M.D. *et al.*, 2006. Crop rotation effects on soil quality at three northern corn/soybean belt locations. Agronomy journal, 98 (3), 484-495

Khosla, R., and Flynn, B., 2008. Understanding and Cleaning Yield Monitor Data. In S Logsdon, D Clay, D Moore and T Tsegaye (Eds.) Soil Science: A Step-by-step Analysis:. Madison. WI. USA: ASA.

Khosla, R., Fleming K., Delgado J., Shaver T., and Westfall D., 2002. Use of site-specific management zones to improve nitrogen management for precision agriculture. Journal of Soil and Water Conservation 57 (6) 513-518.

Miller, P.R., Buschena, D.E., Jones, C.A. and Holmes, J.A., 2008. Transition from intensive tillage to no-tillage and organic diversified annual cropping systems. Agronomy Journal, 100 (3), 591-599

R Development Core Team, 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria

Rathore V., Nathawat N., Bhardwaj S., Sasidharan R., Yadav B., Kumar M., *et al.,* 2017 Yield, Water and Nitrogen Use Efficiencies of Sprinkler Irrigated Wheat Grown under Different Irrigation and Nitrogen Levels in an Arid Region. Agricultural Water Management, 187, 232-245.

Ritchie S., Hanway J., and Benson G., 1993. How a Corn Plant Develops? Special Report Number 48. Iowa State University of Science and Technology, Ames, IA, USA

Seifert C., Roberts M., and Lobell D., 2017. Continuous Corn and Soybean Yield Penalties across Hundreds of Thousands of Fields. Agronomy Journal. 109 541-548.

Soil Survey Staff, 1980. Classification of soil series of the United States, Puerto Rico, and the Virgin Islands. Soil Conservation Service, U.S. Dept. of Agriculture, Washington, DC, USA. 330 pp.

Stone K., Camp C., Sadler E., and Evans D., 2010. Corn Yield Response to Nitrogen Fertilizer and Irrigation in the Southeastern Coastal Plain. Applied Engineering in Agriculture. 26, 429-438.

Tilman D., Balzer C., and Hill B., 2011 Global food demand and the sustainable intensification of agriculture. Proceedings of the National Academy of Sciences USA., 108, 20260-2026.

Zhang H., Kijne J., Barker R., and Molden D., 2003. Improving water productivity through deficit irrigation: examples from Syria, the North China Plain and Oregon, USA. Water Productivity in Agriculture: Limits and Opportunities for Improvement., International Water Management Institute, Colombo Sri Lanka, pp. 301-309.

Statistical model to overcome rice variety effect in precision fertilisation

D. Sacco[1*], E. Cordero[1], B. Moretti[1], E.F. Miniotti[2], D. Tenni[2], G. Beltarre[2], C. Grignani[1] and M. Romani[2]
[1]*Dept. of Agricultural, Forest and Food Sciences, University of Turin, Grugliasco (TO), Italy; dario.sacco@unito.it*
[2]*Rice Research Centre, Ente Nazionale Risi, Castello d'Agogna (PV), Italy*

Abstract

Different rice varieties have different response to N fertilisation, depending on their agronomic traits. Actually, approximately 150 varieties are used in the Italian rice sector so the definition of correct N management strategies represents a large effort. The present study aimed to define strategies for rice N fertilisation based on yield components and substitution value of early and late supply. The statistical analysis conducted here suggests which parameters must be analysed in order to derive helpful rules able to drive best N fertiliser application and to suggest substitution values of late N application in order to compensate for insufficient early development.

Keywords: N fertilisation, optimal N splitting, yield components, PCA, topdressing

Introduction

Rice (*Oryza sativa* L.) is widely cultivated in north-west Italy, in flooded paddy fields. Nitrogen (N) plays a fundamental role in rice production, determining both grain yield and quality (Zhu *et al.,* 2016). Unbalanced N fertilisation hampers rice growth and development (Shaiful Islam *et al.*, 2009) and contributes to environmental pollution, particularly to water eutrophication and greenhouse effect, through N leaching and both NH_3 and N_2O emissions, respectively (Nguyen *et al.,* 2008; Yang *et al.,* 2013). Synchronising N supply to rice requirements both in time and space is needed, to enhance the overall sustainability of rice cropping systems. Rice N uptake varies over the growing season. Hashim *et al.* (2015) have pointed out that N uptake increases during vegetative and reproductive stage while, during the ripening stage, rice utilised N stored in plant tissues, by re-translocation from the vegetative parts to the developing grains. Rice grain yield is defined as the product of yield sink capacity and filling efficiency (Zhang *et al.,* 2013). Consequently, grain yield can be considered as the product of panicle density, spikelets number per panicle, 1000-grain weight and percentage of filled grains.
The right splitting of N supply between pre-plant and topdressing applications optimises rice response to applied fertilisers. Nitrogen supply at pre-sowing and early growth stages promotes rice growth and tillering, while topdressing applications increase the number of filled spikelets (Hirzel *et al.,* 2011; Zhang *et al.,* 2013). Several studies put forth that topdressing N application at panicle initiation stage (PI) can profitably improve yield attributes (Bah *et al.*, 2009; Lee *et al.*, 2009). Vegetative indices, based on spectral characteristics of the canopy, provide an objective tool to quantify amount of N to be supplied with the last N application. However, different rice varieties have different response to N fertilisation, depending on their agronomic traits (Hirzel *et al.,* 2011). Indeed, the optimal time and amount of N application varies with the rice cultivar (Bah *et al.*, 2009). Although total amount to be supplied is related to total N uptake, splitting between early and late applications and the substitution value of late applications supplied to compensate lacks in early application is based on rice variety knowledge. Actually, approximately 150 varieties are used in

DOI 10.3920/978-90-8686-888-9_70, © Wageningen Academic Publishers 2019

Italian rice sector so the definition of a correct N management and definition of the substitution value for each represents a large effort.
The statistical procedure to obtain functions linking N supply as sum of pre-sowing and tillering stage application and N supplied at PI was previously described by Cordero *et al.* (2018). However, the relationship is variety-specific, requiring several trials for adaptation.
The present study aimed to define through a statistical procedure, the yield components that mainly drive the final yields of the three analysed varieties, in order to derive general rules leading to definition of strategies for rice N fertilisation. Substitution value of late fertilisation to compensate lack of local early growth or development was also determined.

Materials and methods

Agronomic management
A multiple-year experiment was carried out in Castello d'Agogna (PV, Italy) (8° 41' 52" E; 45° 14' 48" N) between 2011 and 2017. The experimental area is characterised by a temperate climate, with cold winters and warm summers, and rainfall occurring mainly in spring and autumn. The soil of the experimental site has silty-loam texture, low organic matter content and well-balanced C/N ratio. The typical agronomic practices of the region were adopted throughout the growing seasons. Conventional tillage practices involved spring ploughing, laser-levelling and then harrowing. Phosphorus (P) and potassium (K) were uniformly distributed in all plots before harrowing. Rice was water seeded through broadcasting. Water management involved pin-point flooded method, maintaining a permanent ponding water depth of 5-20 cm until the final draining, with the exception of brief drainage periods to promote root extension, as well as fertiliser and herbicide applications. Three different rice varieties (Gladio, Centauro, and Carnaroli) belonging to three different grain types (long B, round and long A, respectively) have been selected, considering their broad cultivation in the area and their different response to N fertilisation. The experiment was conducted during two growing seasons for each variety. The experimental setup was prepared to diversify crop vigour before the last topdressing fertilisation, by supplying different N rates as sum of pre-sowing and tillering stage applications ($N_{PRE+TILL}$), then analysing the effect of increasing N supply at PI (N_{PI}). Treatments were organised in a split plot design, with $N_{PRE+TILL}$ in the main plots (4.5×26.4 m) and N_{PI} in the subplots (4.5×6.6 m), establishing four replicates for each treatment. Nitrogen rates have been adapted to the specific N requirements of the different rice varieties, as reported in Table 1. Dry granular urea was used as N fertiliser.

Data acquisition and analysis
At the end of the growing season, three areas of 0.25 m^2 were hand-harvested in each plot. Then, grain yield normalised to a moisture content of 14% was determined. Moreover, yield components

Table 1. Nitrogen rates used in each year of the experiment for each rice variety.

Rice variety	Grain type	Year	$N_{PRE+TILL}$ (kg/ha)	N_{PI} (kg/ha)
Gladio	Long B	2011	60-100-140-180	0-30-60-100
		2013	0-60-100-140	0-30-60-100
Centauro	Round	2014	0-60-100-140-200	0-30-60-100
		2015	0-60-100-140-200	0-30-60-100
Carnaroli	Long A	2016	0-30-60-90	0-30-60-90
		2017	0-30-60-75-90	0-30-60-90

(i.e. number of panicles per square meter, number of spikelets per panicle, 1000-grain weight and sterility percentage) were calculated.
Statistical analysis was performed in the R environment for statistical computing, using R version 3.5.1 (R Core Team, 2018). A General Linear Model (GLM) was applied according to Cordero *et al.* (2018), rice grain yield can be considered as the sum of the contribution due to early growth stage and topdressing N application at PI, as well as their interaction. The statistical methodology proposed by Cordero *et al.* (2018) was applied to each rice variety, with the aim of determining the optimal fertilisation strategy accounting for their specific agronomic traits. Results analysis showed huge differences among the rice varieties, pointing out the need of adapting N management to their agronomic traits. Therefore, further improvements of the statistical procedure were studied, to overcome rice variety effect for obtaining best N fertilisation strategy, then extending its application in rice cropping systems. Hence, the present research investigated a statistical procedure that can be conveniently applied to group rice varieties, determining their important traits to explain their different N response. Rice grain yield can be estimated as the product of panicle density, number of spikelets per panicle, and 1000-grain weight, reduced by the sterility percentage (Equation 1).

$$GY = D \times Ns \times P \times (1 - S) \quad (1)$$

where GY is rice grain yield, D is panicle density, Ns is spikelets number per panicle and S is sterility percentage.
Then, with the aim of building an additive model, natural logarithm can be applied to both members of Equation 1, as follows:

$$\ln GY = \ln D + \ln Ns + \ln P + \ln(1 - S) \quad (2)$$

Principal component analysis (PCA) was applied to the logarithmic values of grain yield and grain yield components, to define the most important explanatory variables. Before applying PCA, data were standardised by subtracting the average value and dividing the difference by the standard deviation. Commonalities, representing the proportion of variable variance explained by the principal component (PC), were used to identify the variables that better explained the variability among the rice varieties.
Last, a correlation analysis was applied separately for each rice variety, to response of grain yield components to N fertilisation.
Finally, the GLM reported by Cordero *et al.* (2018) was applied to the logarithmic values of the grain yield components, with the aim of evaluating the influence of N splitting on the final crop production.

Results

The statistical procedure proposed by Cordero *et al.* (2018) suggested the optimal fertilisation strategy for the different rice varieties, which allows obtaining maximum grain yield (Table 2). The model provides maximum reachable yield also in case $N_{PRE+TILL}$ is not optimal due to different N availability induced by local soil effects, suggesting corrected dose to supply in N_{PI}. Consequently, the application of the GLM provides the coefficients that can be used to determine N_{PI} that maximise grain yield corresponding to each $N_{PRE+TILL}$. Coefficients allowed the determination of the slope of the function representing the replacement value of N_{PI} with respect to $N_{PRE+TILL}$. Any reduction of $N_{PRE+TILL}$.equal to the coefficient, require one more unit of N_{PI}.
Clearly, rice varieties require different N management, both in terms of total N supply and splitting during the growing season. Gladio and Centauro need approximately the same total N amount, while Carnaroli requires 40% less N. Both Carnaroli and Centauro require less N at early growth

Table 2. Optimal fertilisation strategy for the different rice varieties, which allows obtaining maximum grain yield and replacement value of N_{PI} to compensate early distribution.

Variety	Total N fertilisation (kg/ha)	$N_{PRE+TILL}$ (%)	N_{PI} (%)	Replacement value
Gladio	200	50	50	-0.313
Centauro	190	20	80	-0.652
Carnaroli	120	33	66	-0.273

stage, and more at PI. Conversely, Gladio needs equal splitting of N application between early growth and topdressing applications.
However, lack of efficiency of early distribution in Centauro can be more easily compensated than in the other two varieties, with Carnaroli being the worst with low level of compensation.
Multivariate analysis applied to logarithmic values of grain yield and grain yield components is able to explain 84.3% of the total variance with the extraction of the two main PC (Figure 1).
In particular, PC1 and PC2 contributed to explain 63.4 and 20.9% of the total variance. First PC has the larger positive association with panicle density and grain yield, and a large negative association with 1000-grain weight, contributing 24.6%, 19.6%, and 24% respectively. Conversely, PC2 has a large positive association with spikelets number per panicle, crop density, which contributes 37.7%, 36.7%, and 20.4% respectively. Moreover, the graph clearly showed the presence of three distinct groups in the data that well reflect the different rice varieties.
Table 3 shows the loadings and the commonalities for each variable. The highest communality value was recorded by spikelets number per panicle, while the lower value was recorded by sterility.
Table 4 shows the effect of N supply on grain yield components that mostly influenced grain yield in the different rice varieties (i.e. panicle density for Gladio, spikelets number per panicle for Centauro and 1000-grain weight for Carnaroli).

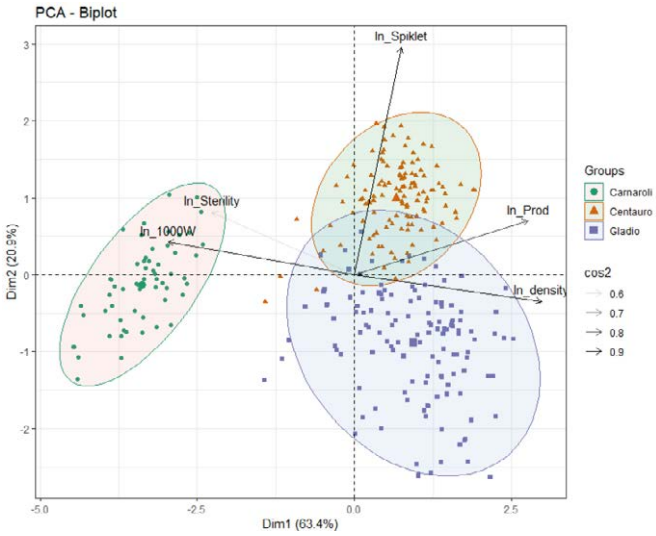


Figure 1. Biplot of Principal Component Analysis applied to log-transformed data.

Table 3. Loadings and commonalities for each variable. Log-transformed data have been used.[1]

Variable	PC1	PC2	Commonalities
Panicle density	**0.536**	-0.108	0.299
Spikelets number per panicle	0.133	**0.926**	0.875
1000-grain weight	**-0.531**	0.133	0.300
Sterility	**-0.406**	0.255	0.230
Grain yield	**0.497**	0.219	0.251

[1] Values in bold show in which factor the variable has the highest loading.

Table 4. GLM applied to main yield component of the varieties, based on N fertilisation quantity and splitting.

Covariates	Gladio Panicle density	Centauro Spikelets number per panicle	Carnaroli 1000-grain weight
$N_{PRE+TILL}$	0.007	n.s.	n.s.
$N_{PRE+TILL}^2$	0.000	n.s.	n.s.
N_{PI}	n.s.	0.000	0.000
N_{PI}^2	n.s.	0.031	n.s.
$N_{PRE+TILL*} N_{PI}$	0.000	0.000	0.021

All grain yield components were affected by the interaction between $N_{PRE+TILL}$ and N_{PI}. Moreover, in Gladio variety panicle density was also affected by $N_{PRE+TILL}$, while spikelets number per panicle and 1000-grain weight were influenced by N supply at PI.

Discussion

The slope of the function linking N_{PI} with $N_{PRE+TILL}$ was negative for all rice varieties, as expected. Then, N_{PI} has to be reduced corresponding to higher $N_{PRE+TILL}$. However, replacement value of N_{PI} with respect to $N_{PRE+TILL}$ was the highest for Centauro variety, followed by Gladio and Carnaroli. This finding demonstrated that, in Centauro variety, the interaction between early growth stage and topdressing N applications is more effective in determining rice grain yield. Consequently, in Centauro variety, N_{PI} decreased more at increasing $N_{PRE+TILL}$.

Application of PCA showed that the differences among the rice varieties can be connected to their different agronomic traits. Carnaroli variety was clearly separated from the others, mainly referring to PC1. Indeed, Carnaroli variety is characterised by a lower panicle density and grain yield, and a high 1000-grain weight. Conversely, Centauro and Gladio appeared as less distinct groups, distinguished mainly on the basis of PC2. In particular, Centauro variety showed a higher spikelets number per panicle. Moreover, maximum grain yield was slightly lower in Gladio variety. Consequently, to achieve maximum grain yield, N fertilisation should be focused to promote 1000-grain weight, spikelets number per panicle and panicle density in Carnaroli, Centauro and Gladio, respectively.

Spikelets number per panicle was assumed as the variable more representative of the total variability. Conversely, sterility was found to be insufficiently representative of the total characteristics. These

findings are consistent with previous studies by Ortega Blu (2007) and De Souza *et al.* (2017), stating that sterility is mainly associated with low temperature during pollen formation and flowering.
Overall, deficient N supply at early growth stage reduced panicle density, that was compensated mainly through higher 1000-grain weight.
The optimal splitting during the growing season can improve crop response to N application, as the interaction between early growth stage and PI application was significant for all grain yield components. Overall, early growth stage N application promoted rice growth and tillering, then positively affecting also re-translocation to the developing grains during the ripening stage, confirming previous results by Hirzel *et al.* (2011). Nitrogen supplied at PI was not effective on panicle density, as tillering process is already ended, but influenced spikelets number per panicle and 1000-grain weight, as also asserted by Zhang *et al.* (2013).
Consequently, for Gladio variety a consistent N amount has to be supplied at early growth stage. Conversely, in Centauro and Carnaroli, a lower N amount has to be supplied as sum of pre-sowing and tillering stage application, the focusing more on PI. Moreover, N_{PI} is higher in Centauro variety than for Gladio, stating the higher replacement value.

Conclusions

A statistical model can be used to optimise rice N fertilisation. The present study found that rice varieties showed wide differences in growth patterns, and consequently different response to N fertilisation. The slope of the function linking early growth stage and topdressing N application provides the replacement value of N_{PI} with respect to $N_{PRE+TILL}$. Logarithmic values of grain yield and grain yield components can be conveniently used in PCA analysis, allowing evaluation of variables that are more effective in determining the differences among the rice varieties. Moreover, PCA enabled evaluation of the specific contribution of each variable in the determination of grain yield. In addition, correlation analysis let to better understanding of the response of grain yield components to N fertilisation. This information, coupled with the investigation of how N supply affected the parameter more related to grain yield for each rice variety, provided useful information for optimising N fertilisation. Consequently, this procedure can be applied to improve N fertilisation in rice. Based on this finding, results of other experiments, not directly addressed to evaluate N splitting, but able to provide information on contribution of yield components to final yield, could be probably used in order to define strategies for improved fertilisation.

Acknowledgements

Authors want to thank Ente Nazionale Risi for having financed the project.

References

Bah, A., Syed Omar, S.R., Anuar, A.R., Husni, M.H.A., 2009. Critical time of nitrogen application during panicle initiation on the yield of two Malaysian rice cultivars (Oryza sativa L.). Pertanika Journal of Tropical Agricultural Science. 32(2): 317-322.

Cordero, E., Moretti, B., Miniotti, E. F., Tenni, D., Beltarre, G., Romani, M., *et al.*, 2018. Fertilisation strategy and ground sensor measurements to optimise rice yield. European Journal of Agronomy 99: 177-185.

De Souza, N. M., Marschalek, R., Sangoi, L., Weber, F. S., 2017. Spikelet sterility in rice genotypes affected by temperature at microsporogenesis. Revista Brasileira de Engenharia Agrícola e Ambiental 21(12):817-821.

Hashim, M. M., Yusop, M. K., Othman, R., Wahid. S. A., 2015. Characterization of nitrogen uptake pattern in malaysian rice MR219 at different growth stages using ^{15}N isotope. Rice Science, 22(5): 250–254.

Hirzel, J., Pedreros, A., Cordero, K., 2011. Effect of nitrogen rates and split nitrogen fertilization on grain yield and its components in flooded rice. Chilean Journal of Agricultural Research 71(3): 437-444.

Lee, J.H., Kang, C.H., Roh, A.S., Park, K.Y., Lee, H.J., 2009. Assessment of N topdressing rate at panicle initiation stage with chlorophyll meter-based diagnosis in rice. Journal of Crop Science and Biotechnology 12(4), 195-200.

Nguyen, H.T., Lee, K.J., Lee, B.W., 2008. Recommendation of nitrogen topdressing rates at panicle initiation stage of rice using canopy reflectance. Journal of Crop Science and Biotechnology 11(2), 141-150.

Ortega Blu, R., 2007. Analysis of factors affecting spikelet sterility in flooded rice under field conditions in Chile. Archives of Agronomy and Soil Science 53(2):183-192.

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org.

Shaiful Islam, M., Hasanuzzaman, M., Rokonuzzaman, M., Nahar, K., 2009. Effect of split application of nitrogen fertilizer on morpho-physiological parameters of rice genotypes. International Journal of Plant Production 3(1): 51-62.

Yang, S., Peng, S., Xu, J., Hou, H., Gao, X., 2013. Nitrogen loss from paddy field with different water and nitrogen managements in Taihu Lake region of China. Communications in Soil Science and Plant Analysis, 44: 2393-2407.

Zhang, Z., Chu, G., Liu, L., Wang, Z., Wang, X., Zhang, H. *et al.,* 2013. Mid-season nitrogen application strategies for rice varieties differing in panicle size. Field Crops Research 150: 9-18.

Zhu, D., Hang, H., Guo, B., Xu, K., Dai, Q., Wei, C. *et al.,* 2016. Effect of nitrogen management on the structure and physicochemical properties of rice starch. Journal of Agricultural and Food Chemistry., 64: 8019–8025.

Assessing sampling strategies and soil sensors performance in the detection of field scale variability of plant-available nitrogen

E. Wallor[1*], M. Bourouah[2], K.C. Kersebaum[1] and R. Gebbers[3]
[1]*Leibniz Centre for Agricultural Landscape Research (ZALF), Institute of Landscape System Analysis, Muencheberg, Germany; evelyn.wallor@zalf.de*
[2]*Hahn-Schickard-Society for Applied Research, Villingen-Schwenningen, Germany*
[3]*Leibniz Institute for Agricultural Engineering and Bioeconomy, Department of Engineering for Crop Production, Potsdam, Germany*

Abstract

Agriculture contributes the strongest to the nitrogen (N) enrichment of ecosystems. It is therefore important to know the spatial distribution of factors influencing N losses or to know the distribution of soil nitrate (NO_3^-) itself. One resource-extensive method to capture NO_3^- variability is the use of sensors. The present study examines the within-field variability of soil mineral N resulting from three sampling strategies and compares outputs of an ion-selective electrode (ISE) and a colorimeter to the results of the N_{min}-method. With regard to sampling, the differences in the measured soil states are more pronounced when grouped by soil texture classes. The sensory measured NO_3^- concentration normally exceeds that of the N_{min}-method. Furthermore, type and quality of the relation between sensory and conventional NO_3^- results depend on the preparation of soil samples and their organic carbon content.

Keywords: soil mineral nitrogen, soil variability, ion selective electrode, colorimeter

Introduction

The variability of plant-available or soil mineral nitrogen (N) at field scale is mainly influenced by the soil texture and its vertical distribution in the soil profile. This is due to the impact of soil texture on soil moisture dynamics which affects denitrification (Giles *et al.*, 2012). Both terms, plant-available and mineral N, consider nitrate (NO_3^-) and ammonium (NH_4^+), with NO_3^- accounting for much more of the soil. Risks to the environment, such as NO_3^- leaching, are caused inter alia by nutrient management that is not adapted to the specific soil inventory and the corresponding hydraulic properties of the field (Galloway *et al.*, 2003). In order to capture the temporal and spatial variability of plant-available N and to ensure site-adapted fertilization, at least two measurements per growing season in a corresponding sampling design are necessary (Nanni *et al.*, 2011). With conventional sampling and soil monitoring at a high spatio-temporal resolution is not economical. Until now, the lack of affordable methods for mapping relevant soil nutrients is a fundamental problem in precision agriculture (Gebbers *et al.*, 2016). To advance research in this area, the present study (1) compares the results of three different sampling strategies with respect to the variability of soil mineral N; (2) examines the importance of soil texture on the distribution of soil states; and (3) investigates the traceability of plant-available NO_3^- in soils by two sensor systems in parallel.

Materials and methods

Study sites and data
The data used for evaluating sampling strategies for assessing the variability of soil mineral N at field scale were collected on a 20-hectare agricultural site near Beckum, Germany (51.750096°N, 7.993412°E), which is characterised by seven soil texture classes (Figure 1). A total of 60 soil profiles

DOI 10.3920/978-90-8686-888-9_71, © Wageningen Academic Publishers 2019

of 0.9 m depth were sampled at three soil layers (0 to 0.3 m, 0.3 to 0.6 m, 0.6 to 0.9 m), at six different dates (3 times in spring, 3 times in autumn) over a period of three years (2000 to 2002). Each sample was analysed for N_{min} and WC (ISO 14256, ISO 11461). Grid points on nested grids (Figure 1) were additionally sampled to a depth of 3 m in order to investigate the vertical distribution of soil mineral N (results not shown). Detailed information about the site and data were provided by Wallor *et al.* (2018). Soil samples used for the comparison of conventionally and sensory examined soil NO_3^- originate from two experimental fields: (1) a long-term field experiment (LTFE) at Muencheberg, Germany (52.503748°N, 14.140469°E) and (2) a N fertilisation trial at Marquardt, Germany (52.457757°N, 12.966905°E). At Muencheberg, mixed samples were collected from the upper soil layer of 48 plots reflecting six fertilisation levels, each with eight repetitions. Sampling date was in October 2016 after harvest. The field trial at Marquardt consists of altogether 16 plots resulting from four fertilisation levels and four repetitions.
At Marquardt, soil samples were collected from three soil layers (0 to 0.3 m, 0.3 to 0.6 m, 0.6 to 0.9 m) in February 2018. At both experimental fields soil texture is dominated by sand (67 to 90%).

Sampling strategies
The present study compares different soil sampling strategies with respect to the spatial and temporal variability of soil N available to plants using the Beckum data introduced above. The sampling strategies considered are according to Schirrmann *et al.* (2011): (1) grid-based sampling practice in precision agriculture with one sampling point per hectare; (2) sampling practice based on soil inventory (zone-based) with one sampling point per soil unit; and (3) intensive sampling (Table 1). The grid-based sampling practice was imitated by the randomised selection of 20 regular grid points from the entire pool. Regarding the zone-based sampling practice the regular grid points were assigned to the corresponding soil texture class and results of soil analysis were averaged accordingly (cf. Figure 1).

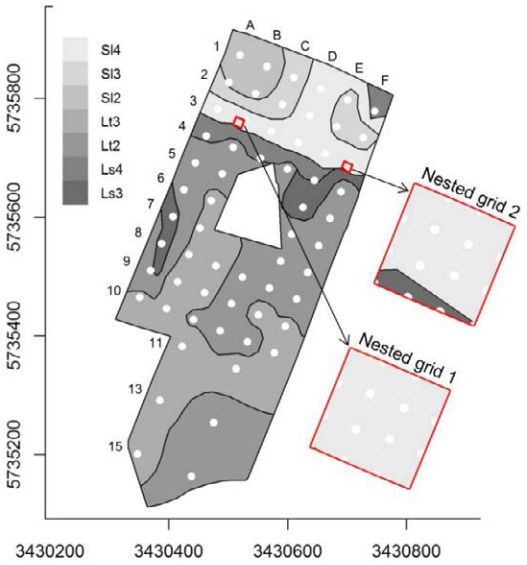


Figure 1. Beckum study site with location of grid points (regular grid: A1 to F10), nested grids and distribution of soil texture classes according to the German Soil Taxonomy (Ad-hoc AG Boden, 2005), aggregated over 0 to 0.9 m soil depth; clay content increases from Sl2 (5-8%) over Ls3 (17-25%) to Lt3 (35-45%).

Table 1. Regular grid points considered in the analysis of each described sampling strategy.

Sampling strategies	Regular grid points included	Number of sample locations per ha (on average)
grid-based	A4, A5, A7, A9, B1, B3, C8, C9, C11, C13, C15, D1, D4, E3, E6, E8, E13, F1, F4, F10	1.00
zone-based	points per soil texture class (cf. Figure 1)	0.35[1]
intensive	all	3.00

[1] Referring to the mean value of all sampling points in one soil texture class.

Sensor measurements

Ion selective electrodes (ISE) and a colorimetric sensor were tested with samples taken from Muencheberg and Marquardt. An ISE is a transducer that converts the concentration of a specific ion dissolved in a solution into an electrical potential. The ion-selective membrane is the key component of ISEs (Bomar *et al.*, 2017). An ISE generates a difference in electrical potential between itself and a reference electrode. This potential is proportional to the concentration of the target ion (here: NO_3^-) in the solution (Rundle, 2011). Colorimetry is an optical method which uses reagents that change their colour in the presence of the analyte (Gautheyrou *et al.*, 2001). The colour intensity, which is proportional to the concentration of the analyte, is measured by a photo diode at a defined wavelength. Soil samples were prepared according to Jones (2001) and afterwards the absorption factor was measured. The specific NO_3^- concentration was estimated using a pre-defined calibration equation and the Lambert-Beer Law (Mayerhöfer *et al.*, 2016). The amount of absorbed light by the reacted sample is directly related to the amount of NO_3^- concentration. For comparing sensory and conventionally measured NO_3^- concentrations, results of the conventionally applied N_{min}-method ($N_{min} = NO_3^-N + NH_4^+$-N, ISO 14256) were multiplied by 44.268 (= 10 × 4.4268) to consider the difference in the dimensions (g/100 g → ppm) and to convert them into NO_3^- concentration (NO_3^--N → NO_3^-).

Tools and statistics

All statistical analyses have been conducted using the R software (Core Team R, 2018). The stratified sampling method from the *spsample* function in R (*gstat* package) (Bivand *et al.*, 2013) was used to select grid points by geometry from the entire pool of the Beckum data to establish the grid-based sample (cf. Table 1). The *spsample* function marks a previously defined number of points in a field. In the case of the field examined here, the grid points closest to the *spsample* marks were selected. Results of soil mineral N as detected by the introduced sampling strategies where compared by using box and whisker plots (McGill *et al.*, 1978). Differences between the distributions of measured soil states according to the soil texture zones were determined by the Mann-Whitney U-test (Mann and Whitney, 1947) due to deviations from the normal distribution of some data. For the comparison of the vertical distribution of soil mineral N within soil profiles averaged over all grid points and over grid points belonging to different soil texture classes, the package *SoilDB* was used (Beaudette *et al.*, 2018). Conventionally and sensory measured NO_3^- contents in the soil samples were compared in a regression analysis. The quality of the corresponding regression models is described by Akaike's An Information Criterion (AIC) (Sakamoto *et al.*, 1986).

Results

Sampling strategies and field-scale variability of soil states

When comparing the distributions of soil mineral N according to the introduced sampling strategies no significant differences could be detected with respect to its within-field variability. This applies to the vast number of sampling dates and soil depths considered in this study (Figure 2). Only the value range was considerably reduced by grouping and averaging measured soil states according to the soil inventory.

In the grid-based sampling strategy the scattering of values declined due to a lower sample size. It also shows that some extremes are not captured by the grid-based sampling. As expected, the amount of soil mineral N is lower in spring than in autumn after harvest when looking at the same year. In addition, soil mineral N decreases with soil depth for each sampling date. In order to identify differences between the vertically aggregated soil texture classes N_{min} and WC results were divided by sampling dates (autumn, spring) and assigned to the corresponding soil texture class. Results are presented in Table 2. Obviously, differences between the units of the soil inventory are more pronounced for WC than for N_{min} (Figure 3) and, the zone-based variation of WC occurs more regularly over the year compared to N_{min}.

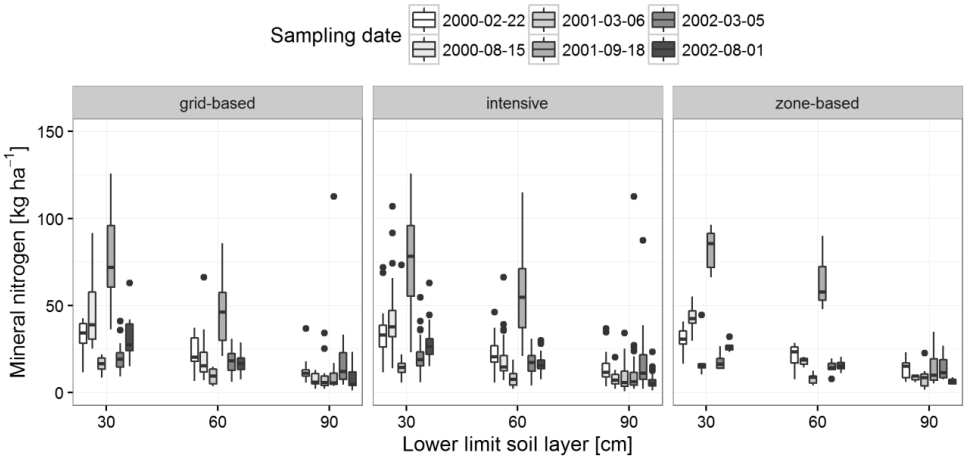


Figure 2. Field-scale variability of soil mineral N as detected by different sampling strategies at six sampling dates and for each soil layer; boxes refer to the lower and upper quartile (25^{th} and 75^{th} percentile), whiskers refer to the upper and lower extreme, and horizontals refer to the median.

The nitrogen state differs mainly in spring between soil texture classes. In autumn, variation in soil mineral N is present only in the soil layer of 0.6 to 0.9 m. However, compared to the overall sample the zone-based vertical distribution of mineral N shows slight differences, which is exemplarily shown for one sampling date in autumn (Figure 3). Furthermore, the range of measured N_{min} is the lowest in the upper layer of the sandy texture class Sl2.

Sensory detection soil NO_3^-

The quality and type of relation between sensory and conventionally measured soil NO_3^- differs according to sampling depth and preparation of the sample. Figure 4 presents the results of NO_3^- measurements from the soil sampling at Marquardt, where three soil layers were considered and

Table 2. Results of the Mann-Whitney U-test (Mann and Whitney, 1947); soil mineral N and WC values were grouped by spring and autumn sampling dates, soil layer and soil texture class; groups sharing the same letter do not show significant differences in the distribution of soil mineral N and WC, respectively.

Season	Soil state	Layer [m]	Soil texture class					
			Sl3	Sl4	Ls3	Ls4	Lt2	Lt3
Spring	Nmin	0-0.3	a	abc	b	abc	c	c
		0.3-0.6	a	abc	b	bc	c	c
		0.6-0.9	abc	ab	c	ac	b	abc
	WC	0-0.3	a	bc	b	c	d	d
		0.3-0.6	a	bc	b	c	d	d
		0.6-0.9	a	b	c	d	d	d
Autumn	Nmin	0-0.3	a	a	a	a	a	a
		0.3-0.6	a	a	a	a	a	a
		0.6-0.9	a	b	abcd	abc	d	c
	WC	0-0.3	a	b	b	b	c	d
		0.3-0.6	a	b	c	bc	d	e
		0.6-0.9	a	b	b	bc	c	c

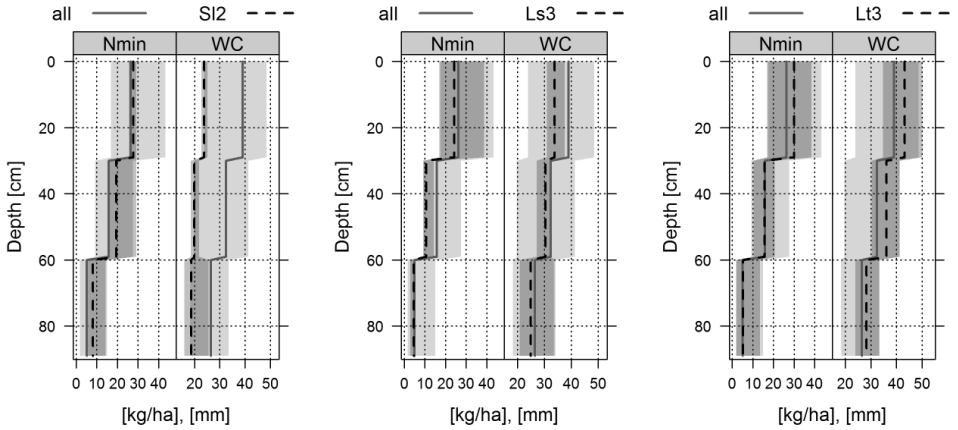


Figure 3. Comparison of soil mineral N and WC in the soil profile averaged over all regular grid points (grey) and over grid points belonging to different soil texture classes (black; Sl2, Ls3, and Lt3); sampling date 01.08.2002; plots show the median bounded by 5th and 95th percentiles.

samples were analysed field-fresh. It becomes obvious that samples from the lowest soil layer show no relation at all between sensory and conventionally measured soil NO_3^-. This is true for both sensors the ISE and the colorimeter. In total, ISE identifies higher NO_3^- contents than the colorimeter. In Figure 5 results from the comparative measurements of the Muencheberg and Marquardt sampling are depicted exclusively for the upper soil layer. In contrast to the Marquardt samples, the Muencheberg samples were dried and sieved before the sensory detection. Again, ISE

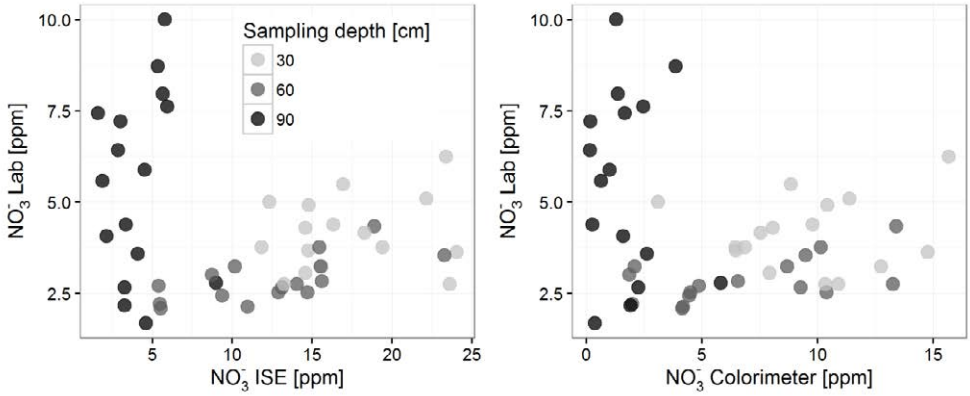


Figure 4. Relation between NO_3^- [ppm] measured conventionally in the laboratory and sensory by ISE and colorimeter; different point colours refer to individual sampling depths.

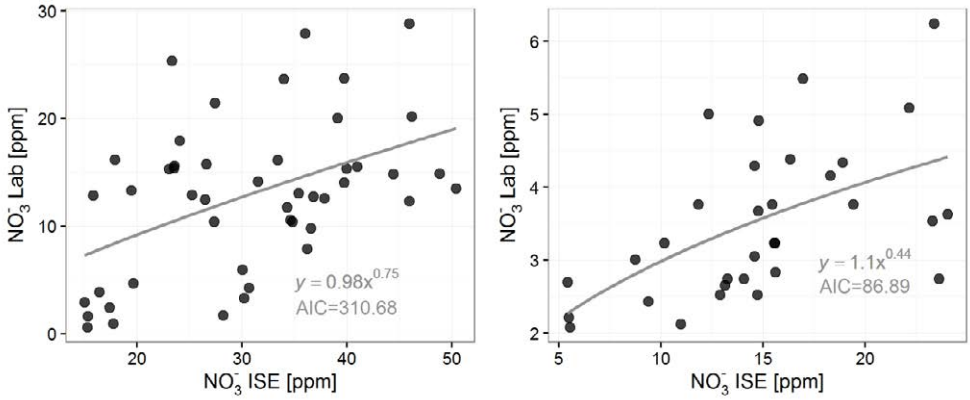


Figure 5. Relation between NO_3^- [ppm] measured conventionally in the laboratory and by ISE; left plot: samples from LTFE Muencheberg, samples were dried and sieved before ISE measurements; right plot: samples from the field trial at Marquardt, samples were measured field-fresh.

values exceed the conventionally measured ones. Although, the AIC of the corresponding regression models is not comparable due to different samples' origin and range of NO_3^- values, results suggest an improved relation when samples are field-fresh. In that case, the regression is rather non-linear compared to the regression of prepared samples.

Discussion

Due to the highly mutable character of soil NO_3^- it has to be stated that a certain number of impact factors is not covered by the present study (pH-value, N_{total}, C_{total}, microbiology). However, it is common practice to determine the required nitrogen fertilisation rate on the base of one soil sampling autumn and/or spring before crop growth is accelerating. The extent of within-field variability is covered especially by the intensive and grid-based sampling strategy. Calculating the average soil mineral N based on all available grid points per each soil unit in the zone-based strategy

causes a decline in the spans of the values for each sampling date. But from a practical point of view, the zone-based sampling strategy allows for the identification of quasi-homogeneous units in the field. Soil-related variations of mineral N were detectable especially for the spring sampling. However, the large observed within-field variability of soil mineral N underpins the need for its efficient detection in the field. Preliminary results of sensory measurements suggest an influence of soil organic carbon (SOC) on the detectability of NO_3^- by ISE and colorimeter. Referring to Figure 4 and the Marquardt sample the correlation between sensory and conventionally examined NO_3^- decreased considerably below 0.6 m sampling depth. In the corresponding soil layer SOC is the lowest and amounts to 0.18% per mass compared to 0.36% per mass and 0.77% per mass in the second and upper layer, respectively. Since the measuring principles of the two sensors used are quite different, there is also the possibility of a systematic error here. The correlation coefficient of the ISE and colorimeter outputs is 0.74 considering all Marquardt samples. The influence of sample preparation on the quality and type of regression and the generally higher levels of sensory detected soil NO_3^- need to be investigated more closely.

Conclusions

The comparison of sampling strategies has shown that intensive sampling produces the widest range of values as well as most outliers. This can lead to misinterpretations in the derivation of management zones. The texture-based sampling strategy is preferable to this, as there are clear differences between the soil types, especially during spring sampling, which is crucial for the planning of site-specific fertilization. The further development of reliable correlation functions for recording soil nitrate from sensor readings requires further reference analyses on different soil types from different soil depths.

References

Beaudette, D.E., Skovlin, J. and Roecker, S.M., 2018. soilDB: Soil Database Interface. R package version 2.0-1. http://CRAN.R-project.org/package=soilDB. (last accessed March 2019).

Bivand, R.S., Pebesma, E. and Gomez-Rubio, V., 2013. Applied spatial data analysis with R, Second edition. NY, USA: Springer. http://www.asdar-book.org

Bomar, E.M., Owens, G.S. and Murray, G.M., 2017. Nitrate Ion Selective Electrode Based on Ion Imprinted Poly(*N*-methylpyrrole). Chemosensors, 5 (2), 1-8.

Core Team R, 2018. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Austria, Vienna. http://www.R-project.org (accessed July 2018).

Galloway, J.N., Aber, J.D., Erisman, J.W., Seitzinger, S.P., Howarth, R.W., Cowling, E.B. *et al.,* 2003. The nitrogen cascade. BioScience 53 (4), 341-356.

Gautheyrou, J., Loyer, J.Y. and Pansu, M., 2001. Soil Analysis. A.A. Rotterdam, the Netherlands: Balkema Publishers. ISBN 90-5410-716-2.

Gebbers, R., Dworak, V., Mahns, B., Weltzien, C., Büchele, D., Gornushkin, I., *et al.,* 2016. Integrated approach to site-specific soil fertility management. In Proceedings of the 13th international conference on precision agriculture. www.ispag.org/proceedings/?action=abstract&id=2084&search=authors (accessed February 2019).

Giles, M., Morley, N., Baggs, E.M. and Daniell, T.J., 2012. Soil nitrate reducing processes – drivers, mechanisms for spatial variation, and significance for nitrous oxide production. Frontiers in Microbiology, 3 (407), 1-16.

Jones, J.B., 2001. Laboratory Guide for Conducting Soil Tests and Plant Analysis. Boca Raton, FL, USA: CRC Press, ISBN 9780849302060.

Mann, H. and Whitney, D., 1947. On a test of whether one of two random variables is stochastically larger than the other. Annals of Mathematical Statistics 18, 50-60.

Mayerhöfer, T.G., Mutschke, H. and Popp, J., 2016. Employing Theories Far beyond Their Limits – The Case of the (Boguer-) Beer–Lambert Law. ChemPhysChem 17 (13), 1948-1955 doi:10.1002/cphc.201600114.

McGill, R., Tukey, J.W. and Larsen, W.A., 1978. Variations of box plots. The American Statistician 32, 12-16.
Nanni, M. R., Pinheiro Povh, F., Melo Demattê, J. A., de Oliveira, R. B., Chicati, M. L. and Cezar, E. (2011). Optimum size in grid soil sampling for variable rate application in site-specific management. Scientia Agricola, 68 (3), 386-392.
Rundle, C.C., 2011. A Beginners Guide to Ion-Selective Electrode Measurements. Nico2000 Ltd, London. http://www.nico2000.net/Book/Guide1.html (accessed February 2019)
Sakamoto, Y., Ishiguro, M. and Kitagawa G., 1986. Akaike Information Criterion Statistics. Dordrecht, the Netherlands: D. Reidel Publishing Company.
Schirrmann, M., Gebbers, R., Kramer, E., Seidel, J., 2011. Soil pH Mapping with an On-The-Go Sensor. Sensors 11, 573-598.
Wallor, E., Kersebaum, K.C., Lorenz, K. and Gebbers, R., 2018. A comprehensive dataset demonstrating the within-field variability of soil properties and crop growth conditions in northwestern Germany. Open Data Journal for Agricultural Research, 5, 1-10.

Section 9
Site-specific canopy and harvest quality

Satellite-based modelling of protein content in winter wheat and malting barley

T. Börjesson[1*], S. Wolters[2] and M. Söderström[2]
[1]*Agroväst Livsmedel AB, Box 234, 53223 Skara, Sweden; thomas.borjesson@agrovast.se*
[2]*Swedish Univ of Agric Sciences, Dept of Soil and Environment; Box 234, 53223 Skara, Sweden*

Abstract

In this project, the aim was to develop satellite image-based models for prediction of protein concentration in malting barley and winter wheat at harvest. Using data from two out of three sampled regions in 2017, it was possible to predict the protein content in validation points in fields not included in the calibration with a mean absolute error (MAE) of less than 1% protein in all but one winter wheat field using a combination of early and late season satellite data. For 15% of the validation points, the predictions deviated more than 1% from the observations. In malting barley, it was sufficient to use only predictors from the late stage to achieve a similar prediction accuracy. For 2018, models similar to the ones from 2017 were obtained in only one of the regions. In the validations, MAE for fields not included in the calibration set were always below 1% in wheat and in 3 out of 4 fields in barley.

Keywords: satellite imagery, protein content, malting barley, winter wheat, decision support system

Introduction

The relationship between crop reflectance at certain wavelengths and crop nitrogen (N) status has been extensively reported (Lee, *et al.*, 2008; Reusch, 2003) and important wavelengths are found in the red – red edge – near infrared (NIR) region of the electromagnetic spectrum. In practice, different sensors can be used for registration of reflectance within fields. In Europe, the tractor-borne Yara N-Sensor is the most common system (Reusch, 2003), having the advantage compared to optical satellites in that it is an active sensor. Satellites on the other hand, cover large areas and, during recent years, the accessibility of data has become much improved, e.g. through the European Space Agency (ESA) Sentinel-2 program, which delivers free data in 10 bands from 400 to 2,200 nm with 10-20 m spatial resolution. In addition, there are a number of other satellites that have been used, some with higher spatial resolution (e.g. Quickbird, DigitalGlobe; or RapidEye, Planet Labs). Remote and proximal sensors have been successfully applied to predict protein contents in mature cereals, and an optimal crop development stage has been found to be late in the season, around flowering (DC 69 – Zadoks *et al.*, 1974) (Basnet *et al.*, 2003; Börjesson and Söderström, 2003). Söderström *et al.* (2010) found that better protein prediction models could be constructed using Yara N-sensor data (45 wavebands in the passive version of the sensor, 10 nm wide between 450 and 890 nm) compared with satellite data used at that time (SPOT 5, Airbus; IRS-P6; Indian Space Organization). Predictions of protein content at flowering is after the period for supplementary fertilization but still of interest both for farmers and the grain industry. With reliable information, the farmer can separate different batches in order to deliver a produce of intended quality. Grain buyers would have better control of quality available at harvest. The current project (2017-2019) aims to assess the possibility of developing prediction models of protein content in winter wheat and spring barley at flowering to early milk development (DC 69-73) using satellite imagery. Data from 2017 and 2018 covering three of the most intensively cultivated regions in Sweden are reported. If successful, the models could be implemented in a decision support system for precision agriculture used in Sweden.

DOI 10.3920/978-90-8686-888-9_72, © Wageningen Academic Publishers 2019

Materials and methods

Grain samples

Samples from 12 winter wheat and 11 malting barley fields were taken in 2017 and from 10 wheat and 10 barley fields in 2018. In each field, 6-8 samples were collected. The fields were located in three of the major agricultural regions in Sweden (Figure 1).

Altogether 175 samples were collected in 2017 and 125 in 2018. Samples were taken by cutting straw and ears from about 1 m^2 divided into 7 subsamples to represent an area of 20 m^2. After drying, the grains were manually threshed and protein content was determined using NIR instrumentation (Inframatic 9500 Grain Analyzer, Perten Instruments, Stockholm, Sweden). The wavelengths used were 850-1,050 nm and the calibration range for wheat was 6.9-21.1% protein and 7.1-17.7% for barley. For wheat, an ANN (artificial neural network) calibration and for barley a PLS-calibration were used. The abilities of the calibrations to predict Kjeldahl protein have been published (European Committée for Standardization, 2015). Validation results for wheat were a standard error of prediction (SEP) of 0.21% with r^2 of 0.99. For barley the corresponding figures were SEP=0.26% and r^2=0.98 (r^2 is defined below).

Satellite data

Data from both the DMC-satellite and the Sentinel-2 were used in 2017, whereas only data from Sentinel-2 were used in 2018. After selecting scenes without clouds, the ones specified in Table 1 were used.

All satellite data from 2017 were recalculated to reflectance and atmospherically corrected through dark object subtraction (Chavez, 1988). Images from 2018 were downloaded in L2A format and, hence, showed bottom-of-atmosphere reflectance directly. Satellite data in all bands (Table 2) were extracted for the grain sample locations (from the pixel coinciding with the sample co-ordinate). For the evaluation, both original wavelengths but also all possible band quotients of the type $(band_1\text{-}band_2)/(band_1+band_2)$ were calculated using data from Sentinel 2 satellites in 2018. In 2017, all original wavelengths were used, but only a fraction of the possible quotients. From the DMC-

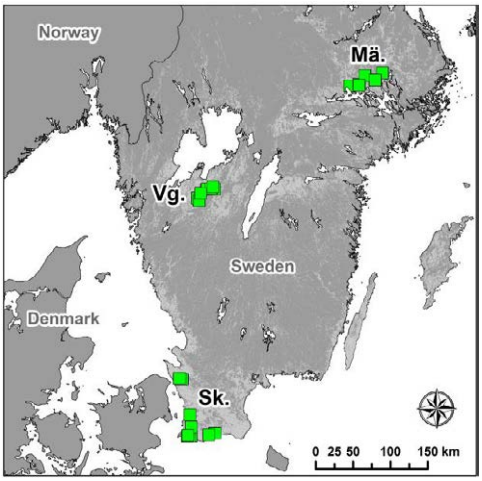


Figure 1. Farm locations (squares) where grain samples were collected in south Sweden. The region names used in the text are: Sk. = Skåne; Vg. = Västergötland; Mä. = Mälardalen. Arable land in Sweden is shown in light grey.

Table 1. Dates (mm-dd) for satellite scenes used during the years 2017 and 2018 respectively in the different regions.

	2017		
Satellite	Skåne	Västergötland	Mälardalen
DMC	05-21, 06-06	05-21, 06-06	05-21
Sentinel	06-19, 07-11*	07-01, 07-04, 07-09	07-01, 07-06
	2018		
Satellite	Skåne	Västergötland	Mälardalen
Sentinel	05-15, 06-01, 06-09, 06-14**	05-15, 06-01, 06-04, 06-19	06-08, 06-28

* One location 07-06. ** Four locations have data from 05-15 and 06-09 or dates close by, 2 locations have data from 06-01 and 06-14.

satellites, normalized difference vegetation index (NDVI) and modified soil-adjusted vegetation index (MSAVI2) were used:

$$NDVI = (NIR - Red)/(NIR + Red) \quad (1)$$

$$MSAVI2 = \frac{(2 \times NIR+1-\sqrt{(2 \times NIR+1)^2-8 \times (NIR-Red)}}{2} \quad (2)$$

and from the Sentinel 2 satellites, in addition, NDVI-type quotients with the following band combinations were used: bands 7 and 6; 7 and 5; 8a and 7; 8a and 6; 8a and 5; 11 and 12; 8a and 12; 7 and 12; 6 and 12; 5 and 12; 8a and 11; 7 and 11; 6 and 11; 11 and 5. These are the 20-m spatial resolution bands and cover the red edge – short-wave infrared (SWIR) region (Table 2).

Table 2. Spectral and spatial resolution of the satellite imagery.

Satellite	Band no	Spectral region	Spatial resolution (m)	Wavelengths (nm)
Sentinel 2	2	Blue	10	443-546
Sentinel 2	3	Green	10	537-583
Sentinel 2	4	Red	10	646-684
Sentinel 2	8	NIR	10	762-908
Sentinel 2	5	Red edge/NIR	20	684-724
Sentinel 2	6	Red edge/NIR	20	730-749
Sentinel 2	7	Red edge/NIR	20	766-796
Sentinel 2	8a	NIR	20	848-881
Sentinel 2	11	SWIR	20	1,540-1,685
Sentinel 2	12	SWIR	20	2,067-2,323
DMC	1	Green	22	520-600
DMC	2	Red	22	630-690
DMC	3	NIR	22	770-900

Multivariate statistics

Models to predict protein content were constructed using the software Unscrambler X 10.5 (Camo, Oslo, Norway) and partial least squares (Martens and Naes, 1989). To start with, all variables listed above were used as input variables. In many cases, however, only a fraction of the variables was valuable for the calibrations. In cases where the models could be substantially improved by reducing the number of input variables, this is documented in the results. Full cross validations and validation sets were used. For calibration using validation sets, all data in one region(s) except samples from one of the fields were used and samples from the omitted field were used as validation set one at a time. For evaluating models using validation sets, only the statistics mean absolute error (MAE) and absolute error (AE) for individual sampling points were used.

For comparing the different cross-validated models, besides MAE and AE, the three statistics square of the Pearson correlation coefficient (r^2), standard error of cross validation (SECV) and ratio of performance to deviation (RPD) were calculated. These statistics are explained below together with bias and root mean squared error of cross validation (RMSECV) that are needed for the calculation of SECV. All these statistics except RPD are received directly from the Unscrambler program. It was not considered relevant to report these statistics when using validation sets, since the number of validation samples were very low in these cases.

- r^2 (Pearson). The Pearson r^2 value is the square of the correlation value and expresses correlation on a positive scale between 0 and 1.
- RMSECV. A measure of the dispersion of the samples around the regression line when cross validation is used. Unlike MAE, the errors are first squared and then the square root is calculated. In this investigation, units are given in % protein content and is the standard deviation of the difference between the predicted and the reference values.
- SECV. The RMSECV corrected for bias.
- Bias. The mean value of all points that either lie above or below the regression line.
- RPD. Calculated by dividing the standard deviation (SD) of the sample set by SECV. If SECV is equal to SD, RPD value is 1 and the model is not predicting reference values. It is a valuable statistic for comparing models (Williams and Sobering, 1993).

Table 3. Statistics for measured protein contents (% of dm) in samples.

Year	Wheat								
	Skåne			Västergötland			Mälardalen		
	Mean	Stdev	n	Mean	Stdev	n	Mean	Stdev	n
2017	11.9	1.4	31	10.9	0.9	32	11.9	1.3	27
2018	11.9	1.2	23	12.2	0.8	21	11.9	0.6	22
	Barley								
	Skåne			Västergötland			Mälardalen		
Year	Mean	Stdev	n	Mean	Stdev	n	Mean	Stdev	n
2017	9.8	1.5	29	10.6	0.7	35	11.1	1.2	20
2018	12.3	1.1	17	13.1	1.8	18	12.6	0.8	21

Results

The protein contents were higher in 2018 than in 2017 and the range was wider for wheat in 2017, while for barley the variation between the years was small (Table 3).

Models based on data from 2017, wheat

The best cross-validated model was obtained using data from Skåne, (N=22, 3 out of 4 locations): r^2=0.76, SECV=0.59 and RPD=2.0. For Mälardalen (N=25, two outliers removed), r^2=0.65, SECV=0.78 and RPD=1.7 were obtained and for Västergötland (N=32), r^2=0.68, SECV=0.50 and RPD=1.8. In the latter case, the best model was obtained using only data from July 1.

A combination of data from Mälardalen and Västergötland (N=58) gave the best combined results with r^2=0.57, SECV=0.78 and RPD of 1.5. In this case the best model was obtained with a reduced set of variables: NDVI from May 21, and NDVI-type indices from July 1: (band 6 – band 11/band 6 + band 11) and (band 11 – band 5/ band 11 + band 5). In all these cases the MAE were below 1% and AE for less than 15% of the individual sampling points were above 1%.

Using this dataset, models that were validated using validation sets with samples from individual fields proved to be able to predict the field mean protein content with high accuracy; MAE of less than 1% protein for all fields except one were obtained and the AE for individual sampling points were always below 1% except for that same field. In four of the eight fields, r^2 values for the validated data points were above 0.1, indicating a possibility to predict within-filed variations.

Models based on data from 2017, barley

The best cross validated models were constructed with data from Skåne, (N=27, two outliers removed), r^2=0.88, SECV=0.52 and a RPD-value of 2.8, indicating a good model. In this case, the model included NDVI and MSAVI2 from the DMC scene June 6, and all the indexes and quotients listed above from the Sentinel-2 scene from July 11.

By combining data from Skåne and Västergötland, N=62, a model with r^2=0.69, SECV of 0.66 and RPD of 1.8 was constructed. A reduction in the number of input variables were tried and a model was obtained using only two variables from the Sentinel 2 scenes from July 11, (band 7 – band 5/ band 7 + band 5) and (band 5 – band 12/ band 5 + band 12), with r^2=0.64, SECV=0.70 and RPD of 1.7. In all these cases, MAE<1%, and AE>1% for less than 15% of the individual sampling points.

This model with reduced numbers of input variables was also used for validations of validation sets. In all these cases, the MAE were below 1%, and in about 10% of the validation points, the predictions deviated more than 1% from the observations. In four of seven fields, r^2 values for the validated data point were above 0.1, indicating a possibility for predicting within-filed variations.

A test version of the protein prediction model in winter wheat was implemented in 2017 in the Swedish version of the free decision support system CropSAT (www.cropsat.se).

Results from 2018

In general, no useful models could be constructed with the exception of models using data from Västergötland. In this case, the best cross-validated models for barley (N=18), r^2=0.66, SECV=1.05 and RPD of 1.7 were obtained and for wheat (N=18) r^2=0.64, SECV=0.41 and RPD of 1.6.

For both crops, the MAE were below 1% and for wheat no sampling points deviated more than 1% between predicted and reference values. For barley, however, the AE for individual sampling points were above 1% in 33% of the cases. This reflects a narrower range in protein contents in wheat.

Using validation sets, similar results were obtained. For wheat MAE were below 1% for all fields and for individual sampling points and the absolute errors were below 1% in 11% of the cases. For barley, MAE were below 1% in 3 out of 4 fields and the absolute errors for individual sampling points were above 1% in 60% of the cases.

Discussion

Models for predicting protein content in both malting barley and winter wheat and for all studied regions, with the exception of barley data from Mälardalen could be constructed using data from 2017. However, using data from 2018, only models based on data from Västergötland could be constructed. Local models worked better when combining data from different regions, which can be explained by the fact that different satellite scenes were often used in different regions. Another reason could be that climatic conditions were more uniform in the separate regions, then when several regions were combined. The SWIR wavelengths were valuable, indicating that the water content of the crops may be important for the predictions. The reasons for the limited success in predicting protein content in 2018 was probably the unusually hot and dry weather during the cropping period in 2018 (Table 4).
In particular, temperatures were much higher than normal in 2018. This probably led to heat-stress and a hampered deposition of starch, which has been described by Pettersson *et al.* (2006). It is thus probable that heat-stress influenced the protein content to a larger extent in 2018 and therefore satellite-based models were more difficult to construct. The best models for 2018 were obtained from Västergötland, the region where the heat stress was not as pronounced compared with the other regions.

Conclusions

Using data from 2017, it was possible to construct models based on satellite imagery to predict protein contents in wheat and malting barley. The phenological stage around early milk stage (DC 73) that occurred late June-early July for wheat and early July-mid July in barley, proved to be an important stage for prediction of protein content at harvest, although earlier stages often gave valuable input. Using models with a combination of data from two regions, it was almost always possible to predict a mean protein content for validated fields and in about half of the fields, within-field variations could also be predicted. During a year with pronounced heat-stress, such as 2018, it was more difficult to construct models to predict protein content based only on satellite data.

Table 4. Average for precipitation (mm) and temperature (°C) at 2-3 weather stations in the vicinity of the trial sites during June and July 2017 band 2018.

	2017		2018		2017		2018	
	Precipitation (mm)				Temperature (°C)			
Region	June	July	June	July	June	July	June	July
Skåne	77	44	29	14	15.4	15.8	17.9	20.3
Västergötland	56	31	62	10	14.8	15.7	16.2	20.2
Mälardalen	39	14	40	22	14.7	16.5	16.3	21.4

Acknowledgements

Lantmännen Research Foundation is acknowledged for funding the project.

References

Basnet, B., Apan, A., Kelly, K., Strong, W. and Butler, B., 2003. Relating satellite imagery with grain protein content. In: Proceedings of the spatial sciences conference, Canberra, Australia: Spatial Sciences Institute. pp. 22-27.

Börjesson, T. and Söderström, M., 2003. Prediction of protein content in cereals using canopy reflectance. In J. V. Stafford & A. Werner (Eds.), Proceedings of the 4th European Conference on Precision Agriculture, Precision agriculture (pp. 89-94). Wageningen, the Netherlands: Wageningen Academic Publishers.

Chavez Jr P. S., 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. Remote Sensing of Environment 24, 459-479.

European Committée for Standardization, 2015. Cereals- Determination of moisture and protein – Method using Near-Infrared-Spectroscopy in whole kernels, European standard EN 15948:2015 E.

Lee, Y.-J., Yang, C.-M., Chang, K.-W. and Shen, Y., 2008. A simple spectral index using reflectance of 735 nm to assess nitrogen status of rice canopy. Agronomy Journal 100, 205-212.

Martens, H. and Naes, T., 1989. Multivariate calibration. Chichester, UK: John Wiley & Sons.

Pettersson, C.G., Söderström, M. and Eckersten, H., 2006. Canopy reflectance, thermal stress, and apparent soil electrical conductivity as predictors for within-field variability in grain yield and grain protein of malting barley. Precision Agriculture, 7, 343-359.

Reusch, S., 2003. Optimisation of oblique-view remote measurement of crop N-uptake under changing irradiance conditions. In J. V. Stafford & A. Werner (Eds.) Proceedings of the 4th European Conference on Precision Agriculture, Precision agriculture Wageningen, the Netherlands: Wageningen Academic Publishers, pp. 573-578.

Söderström, M., Börjesson, T., Pettersson, C.G., Nissen, K. and Hagner, O., 2010. Prediction of protein content in malting barley using proximal and remote sensing. Precision Agriculture 11, 587-599.

Williams, P.C. and Sobering, D.C., 1993. Comparison of Commercial near Infrared Reflectance and Reflectance Instruments for Analysis of Whole Grains and Seeds. Journal of Near Infrared Spectroscopy 1, 25-32.

Zadoks, J.C., Chang, T.T. and Konzak, C.F., 1974. A decimal code for growth stages of cereals. Weed Research 14, 415-421.

Yield sensing technologies for horticultural crops: a short review

L. Longchamps[1*], B. Panneton[1], S. Fountas[2] and R. Khosla[3]
[1]*Agriculture and Agri-Food Canada, 430 Gouin blvd., St-Jean-sur-Richelieu, QC, J3B 3E6, Canada; louis.longchamps@canada.ca*
[2]*Agricultural University of Athens, Iera Odos 75, 11855 Athens, Greece*
[3]*Colorado State University, 307 University Ave, Fort Collins, CO 80521, USA*

Abstract

Few yield monitoring systems exist for horticultural crops. Yield maps can play an important role in the advent of big data in agriculture. The majority of horticultural crops are hand harvested and are diverse, which complicates the development of yield monitoring systems for such crops. Commercial systems for real time yield monitoring of horticultural crops involve use of load cells, which require further improvement to provide information on crop quality and noise filtration. Other technologies covered in this short literature review include machine vision, capacitance, RFID, ultra-wide radar systems and LIDAR for the purpose of yield monitoring of horticultural crops.

Keywords: yield monitoring, yield mapping, load cells, machine vision, capacitance

Introduction

Yield (amount and quality) are the final outcome of crop production systems and yield maps are necessary to conduct studies that empower farmers and lead to improved productivity. Yield data is increasingly becoming important for traceability and to increase safety of the agri-food supply chain (Costa *et al.*, 2013). Yield maps are readily available for combinable crops. However, for horticultural crops (HC), which are in general intensive in farm input use, yield maps are generally not available. This may be due to the complexity of developing yield mapping systems for HC, which are of various shapes, sizes and forms, and require quality aspects of yield as well as repeated harvests (Ruiz-Altisent *et al.*, 2010). Technologies exist to estimate or measure yield in field conditions, from simple load cells to complex machine vision systems. While the utility of yield maps may have been limited to on-farm book keeping up to now, this piece of information may be pivotal for the future of both tactical (i.e. variable rate fertilization) and strategic (i.e. field boundary) crop management. Most HC receive many manual operations; they are mostly hand harvested; quality very often overarching quantity; the field size is generally smaller; plant architectures are more complex; crop protection and fertilisation treatments are also more frequent than in arable crops. In general, HC can be divided into two categories. The first category belongs to the perennial crops, where the planting system remains stable over years, while morphological adaptations of canopy and root develops according to the environment given. The second category are the annual HC, where the variability within the field is important and site-specific operations within the growing season provide potential to enhance productivity and quality (Zude-Sasse *et al.*, 2016). This short review aims at explaining some of the existing technologies that are commercialized or being investigated for yield mapping of HC.

Yield measurement principles

Load cells
Load cells are electromechanical devices converting force to an electrical signal. They can be used to measure weight. Well-designed load cells produce an electrical signal that is linear with respect to weight. Load cells can be installed underneath the belt of a conveyor to measure the weight of

DOI 10.3920/978-90-8686-888-9_73, © Wageningen Academic Publishers 2019

produce supported on a section of the conveyor. Coupling the generated force signal to measurement of the speed of the conveyor, a mass flow rate can be estimated. Dividing this mass flow rate by the ground speed and the collecting width of a harvester, yield can be calculated (i.e. kg/m^2). This approach has been implemented mostly for potatoes and root crops (Panneton and St-Laurent, 1999) and is available commercially (e.g. RiteYield from Greentronics, Elmina, ON, Canada or Advanced Technology and Viticulture (ATV), Joslin, SA, Australia).

The use of load cells is also being investigated for mapping yield of crops not conveyed on a belt. In most cases, this involves weighing harvested product as it is deposited in a cart. Anjom *et al.* (2018) designed a low-cost strawberry yield mapping system based on a load cell installed on a light-weight strawberry-picking cart and connected with a RTK-GNSS. Their system was meant to map the yield of a single strawberry picker as it goes through the field and every time strawberries are added to the cart, the additional weight is recorded along with the geo-coordinates. Even though this system was based on low tech (i.e. PVC pipes and Arduino controller), it enabled a prediction accuracy of 95.2% based on the full tray under field conditions. Other systems use the same concept on larger implements, such as a system for monitoring peanuts falling in a harvesting cart in real time, which provided a prediction accuracy between 97-98% on a trailer load basis under field conditions (Vellidis *et al.*, 2003).

Similar to types of grain yield monitoring systems, certain yield monitoring systems for other crops use load cells attached to impact plates to measure yield in real time. For instance, Maja and Ehsani (2010) developed a yield monitoring system using an impact plate, which converted impact force into citrus fruit mass using a mathematical model. This system enabled an average prediction accuracy of 92.2% between actual and computed fruit weight under field conditions. One common challenge when working with load cells is the background noise due to vibrations (Maja and Ehsani, 2010). For stationary carts, the background noise is minimal. However, for belt conveyors on a harvester, the vibrations induced by the machine (high frequency) and by the uneven terrain typical of crop fields (low frequency) add significant noise (Boschetti *et al.*, 2013). Noise compensation can be accomplished by signal treatment or by vibration attenuation in design (Maja and Ehsani, 2010; Panneton and Brouillard, 2005). Load cells only measure weight while the yield of certain crops greatly depends on quality parameters. In order to mitigate the weaknesses of load cells, other yield monitoring approaches have been investigated in the literature.

Machine vision

Affordable high resolution imagery systems and powerful machine vision and learning algorithms are available. These allow for new types of yield monitoring systems based on machine vision where both, quantity and quality parameters can be acquired. Most yield estimation systems based on machine vision comprise the following steps: (1) image segmentation; (2) counting segmented shapes; and (3) aggregating across multiple images (Boatswain Jacques *et al.*, 2018; Choi *et al.*, 2015; Lee *et al.*, 2018; Roy *et al.*, 2018; Syal *et al.*, 2013). Multiple challenges can be encountered, which may affect results such as fruit occlusion by leaves or other fruits, variability of lighting conditions, plant structure and irregularity of shapes and sizes to be detected (Gongal *et al.,* 2016). Image segmentation of green fruits (e.g. limes, green apples or avocados) on a green tree background may require more powerful algorithms than segmentation of red apples for instance. Choi *et al.* (2015) used gradient vectors in their algorithm to segment green citrus from green tree, which enabled distinction between leaves and the circular pattern of citrus fruits. Once image has been segmented reliably, counting features in the images is simple. The next complex task resides in aggregating counts from multiple images and avoiding recounting objects. One approach consists in avoiding overlap across images. This simplifies image analysis but requires better control of image acquisition to account for the motion of the camera. Another approach consists in adding a step to account for double counts in the image processing analysis. Song *et al.* (2014) used a statistical approach based on location of the object in the image to cluster repeated, incomplete observations. By estimating the probability of finding

an object in a specific location of the image based on the location shift of objects observed in the previous image, they were able to account for the majority of double counts. Another way to avoid double counts is to post process the images. One way could be to generate a mosaic of the fruit wall or of the conveyor belt on which products are transported. However, this may be challenging if wind or vehicle is moving the plants or if objects shift on the conveyor belt. Nuske *et al.* (2014) used a slightly different approach for image post processing which avoided feature recognition required for mosaicing. Instead, they projected images on the fruit wall based on exact location, position and orientation of the camera. By locating fruit bunches in every projected image, they were able to detect fruits that were counted on more than one image.

Yield mapping based on machine vision has been studied for multiple crops and has demonstrated great potential for accurate yield monitoring. Roy *et al.* (2018) used machine vision to develop an end-to-end method based on a monocular camera for mapping apple yield in orchards and were able to achieve a prediction accuracy of 92.0-94.8% across different datasets. Other tree crops can also be monitored with relatively high accuracy. The prediction accuracy for mangos and citrus fruit counting using machine vision was 92.3% for mangos and as high as 98.4% for citrus (Annamalai *et al.*, 2004; Payne *et al.*, 2013). In general, machine vision based fruit counts underestimated actual fruit numbers, mainly due to the location of fruits that were hidden from the camera within the branches of the tree (Annamalai *et al.*, 2004; Payne *et al.*, 2013). A fruiting wall can eliminate the problem of hidden fruits by exposing all fruits in a 2D plane (Dorigoni, 2016; Zhang, 2017).

Beyond counts, other parameters can be drawn from images using machine vision. Estimating the yield of potatoes scattered on the ground on a mass basis, Lee *et al.* (2018) achieved a prediction accuracy of 88%. Chinchuluun *et al.* (2009) determined the number of citrus and their size using machine vision monitoring fruits on a conveyor of a canopy shake and catch harvester. When comparing fruit weights to fruit diameter obtained by machine vision, they reached an R^2 of 0.96. This indicates that crops being commercialized on a mass basis or potentially on a volume basis could be monitored using machine vision approaches. It also demonstrates the potential of machine vision to map yield on both quantity and quality basis such as is required for crops like onion, which are valued according to their calibre (i.e. diameter). One commercial system is expected to be market-ready in 2019, the Harvest Quality Vision™ (Croptracker, Kingston, ON, Canada), which offers the possibility to measure 'color profile, size and quantity' of fruits for single trees in an orchard.

Deep learning or deep neural networks is now a well-established and successful tool in machine vision. One of the main advantage of deep learning compared to other image processing pipelines is that there is no need for feature engineering (Chollet and Allaire, 2018). Feature engineering is the process of defining parameters from image data as inputs for yield estimation models, for example. In traditional image analysis, feature engineering is a tedious task. Deep learning has been demonstrated to be a promising approach to fruit counting and yield estimation (Keresztes *et al.*, 2018; Sa *et al.*, 2016).

Capacitance

Another approach that has been investigated for high throughput yield monitoring is the use of capacitive sensors. Capacitance is the capacity of matter to store electric charge. Capacitance of an object can be measured using two conductive plates (electrodes) separated by a space. There are three factors influencing the capacitance: (1) the size of the electrode plates; (2) the distance between the plates; and (3) the type and quantity of dielectric material, here air and harvested product. For a fixed geometry, the third factor is the one of interest for yield monitoring. While the dielectric constant of air is about 1, the dielectric constant of distilled water is 80. In theory the mass of plant product with a relatively stable water content will be correlated with capacitance. One research group in the Czech Republic has thoroughly investigated the question and obtained success in measuring mass yield for different plant products such as forage crop, sugar beet, hop, potato and chopped maize

biomass (Kumhala *et al.*, 2009, 2010). For instance, Kumhála *et al.* (2013) estimated the weight of wet hop cones using a capacitive sensor with a prediction accuracy of 96% under harvesting conditions.

Ultra-wide band radar systems
Ultra-wide band (UWB) radar systems have been investigated for their potential to map root crop yield while still underground. Konstantinovic *et al.* (2007) used UWB radar to detect the presence of sugar beets in the ground. When coupled with a GNSS, they were able to map sugar beet production. The challenge of such a system resides in differentiating soil aggregates from products. This system was tested under field conditions, attempting to detect location and size of sugar beets in the soil and resulted in a correlation of 80% between mass of sugar beet and amounts of backscattered energy.

High throughput volumetric measurement
Another approach to estimate yield on a conveyor belt involved the use of a Light Detection and Ranging (LIDAR) system. Jadhav *et al.* (2014) measured fruit volume using a LIDAR to estimate total fruit mass of citrus. The LIDAR detected the surface of a conveyor belt perpendicularly to the fruit flow. Under laboratory conditions, the system had a prediction accuracy of 93% while under field conditions, the predictions accuracy was 90%. Lower accuracy was due to belt speed variations.

Harvest georeferencing

Georeferencing harvest provenance
Crops that are organized in a highly structured fashion can be mapped with fairly low technology use. For instance, in a tea plantation, yield can be mapped by recording the weight of leaves harvested from each production block using a paper and pen based registration sheet (Blackmore, 2003). By mapping the weight of tea leaves harvested in each independent production block, it is possible to produce yield maps. Another way to map hand harvested products is to record the location of each harvest bin when they are full and waiting on the ground to be picked up by the collection platform. Full harvest bins are usually left standing very close to where they were filled and thus they are in synlocation with the origin of their content. Fountas *et al.* (2011) measured yield variation in an olive tree orchard. Olives in conventional orchards were picked by hitting fruit branches with sticks. Olives fell on plastic sheets placed underneath each tree. The olives were placed in bags and left in groups where they were filled, for loading on platform. Each bag was weighed and geo-referenced using DGNSS. Each group of bags was considered to present the yield of surrounding trees and was the basis for yield mapping. Spatial variability was observed. Similarly, Liakos *et al.* (2017) and Vatsanidou (2018) created yield maps on hand-harvested apple and pear orchards.

Using Radio Frequency Identification systems
Radio Frequency Identification (RFID) is a system that allows precise identification of an electronic tag using electromagnetic waves at a short distance. It consists of a small silicon computer chip coupled with a small antenna, which is triggered when within the range of a remote reader. The unique identifier of RFID chips can help track the work of each worker independently in a field for instance. Ampatzidis *et al.* (2009) used RFID technology in an attempt to quantify fresh fruit yield on a per tree basis. They compared two approaches. The first approach consisted of installing RFID chips on harvest bins and a DGNSS connected RFID reader on a tractor. The geolocation of each bin filled with fresh fruits would be recorded when loaded on the tractor platform. When co-registered with a map of tree locations, the bins are associated with the nearest tree using a nearest neighbor algorithm. It is thus possible to map the number of bins harvested per tree and a map of harvested weight if bins are weighed at the packinghouse. The second approach consisted of attaching a RFID tag on each tree as well as on each bin. The information was recorded by a RFID reader located on the tractor collecting bins on a platform. When the tractor stops at a location to pick up filled

bins, the RFID reader records tree IDs as well as bin IDs. The tree IDs were associated with location co-ordinates in a preliminary step and thus the bin IDs recorded at the same time as tree IDs are associated and a map can be created on a bin or on a weight basis provided bins are weighed at the packing house. Ampatzidis *et al.* (2009) found that the second approach was preferred when the DGNSS signal was obstructed by dense canopy. They also found that mapping error (0.32%) came from RFID reader missing one bin. Otherwise, the system was reliable. Ampatzidis *et al.* (2016) also used RFID chips. However in this study, the RFID chip was worn (i.e. bracelet) by the picker. An RFID reader associated with a GNSS and a scale weighing the collection bin recorded the transactions (i.e. when picker brings fruits to the collection bin) of all pickers with the following data: picker's ID, fruit weight, time and location. While this Portable Labor Monitoring System (PLMS) was intended to monitor field work and accurately credit pickers for fruits that they harvested, it also enabled yield mapping due to the use of a GNSS system connected to the RFID reader.

Discussion

The yield sensing methods in this review can be categorized based on three measurements; by mass, volume or count. Methods for mass measurement include weighing by load cells or hydraulic pressure gauge, capacitance and impact plates. The volume can be measured using Lidar, UWB radar, 3D vision or laser profilometry. Products can be counted by machine vision, light obscuration particle count or by hand. Therefore, the choice of the right yield monitoring approach can be based on how the produce is marketed (by weight, volume or count) in order to provide a meaningful input to farmers. This choice should also be based on the type of product. For instance, fragile fruits cannot be weighed using impact plates. Another important aspect is the harvesting machinery design and process. Machine vision can perform weakly if installed in a part of the machinery where product is covered in soil, for instance. Research is needed to develop reliable commercial yield mapping systems for the majority of HC in order to increase precision and accuracy of these systems. Advances may also come from the fusion of multiple technologies.

Acknowledgements

Thank you to Éléonore Tremblay for reviewing the scientific literature databases.

References

Ampatzidis, Y., Tan, L., Haley, R. and Whiting, M.D., 2016. Cloud-based harvest management information system for hand-harvested specialty crops. Computers and Electronics in Agriculture 122: 161-167.

Ampatzidis, Y.G., Vougioukas, S.G., Bochtis, D.D. and Tsatsarelis, C.A., 2009. A yield mapping system for hand-harvested fruits based on RFID and GPS location technologies: field testing. Precision Agriculture 10 (1): 63-72.

Anjom, F.K., Vougioukas, S.G. and Slaughter, D.C., 2018. Development and application of a strawberry yield-monitoring picking cart. Computers and Electronics in Agriculture 155: 400-411.

Annamalai, P., Lee, W.S. and Burks, T.F., 2004. Color vision system for estimating citrus yield in real-time. Paper No. 043054, St Joseph, MI, USA: ASAE

Blackmore, S., 2003. The role of yield maps in precision farming. PhD Thesis. Cranfield University, Cranfield, UK.

Boatswain Jacques, A., Adamchuk, V.I., Cloutier, G., Clark, J.J. and Miller C., 2018. Development of a Machine Vision Yield Monitor for Shallot Onion Harvesters. In: Proceedings of the 14th International Conference on Precision Agriculture, (https://www.ispag.org/Proceedings).

Boschetti, G., Caracciolo, R., Richiedei, D., and Trevisani, A., 2013. Model-based dynamic compensation of load cell response in weighing machines affected by environmental vibrations. Mechanical Systems and Signal Processing 34(1): 116-130.

Chinchuluun, R., Lee, W.S., and Ehsani, R., 2009. Machine vision system for determining citrus count and size on a canopy shake and catch harvester. Applied engineering in agriculture 25(4): 451-458.

Choi, D., Lee, W.S., Ehsani, R., Schueller, J.K. and Roka, F., 2015. Machine vision system for early yield estimation of citrus in a site-specific manner. Paper No. 152181863, St Joseph, MI, USA: ASABE

Chollet, F. and Allaire, J.J., 2018. Deep Learning with R.Shelter Island, NY, USA: Manning Publications Company.

Costa, C., Antonucci, F., Pallottino, J., Aguzzi, F., Sarriá, D. and Menesatti, P., 2013. A Review on Agri-food Supply Chain Traceability by Means of RFID Technology. Food Bioprocess Technology 6 (2): 353-366.

Dorigoni, A., 2016. Innovative fruit tree architecture as a nexus to improve sustainability in orchards. Acta Horticulturae 1137: 1-10.

Fountas, S., Aggelopoulou, K., Bouloulis, C., Nanos, G., Wulfsohn, D., Gemtos, T., *et al.*, 2011. Site-specific management in olive tree plantation. Precision Agriculture 12(2): 179-195.

Gongal, A., Silwal, A., Amatya, S., Karkee, M. Zhang, Q. and Lewis, K., 2016. Apple crop-load estimation with over-the-row machine vision system. Computers and Electronics in Agriculture 120: 26-35.

Jadhav, U., Khot, L.R., Ehsani, R., Jagdale, V. and Schueller, J.K., 2014. Volumetric mass flow sensor for citrus mechanical harvesting machines. Computers and Electronics in Agriculture 101: 93-101.

Keresztes, B., Abdelghafour, F., Randriamanga, D., Da Costa, J.-P. and Germain, C., 2018. Real-time fruit detection using deep neural networks. In: Proceedings of the 14th International Conference on Precision Agriculture, (https://www.ispag.org/Proceedings).

Konstantinovic, M., Woeckel, S., Schulze Lammers, P. and Sachs, J., 2007. Evaluation of a UWB radar system for yield mapping of sugar beet. Paper No. 071051, St Joseph, MI, USA: ASABE.

Kumhála, F., Kavka, M. and Prošek, V., 2013. Capacitive throughput unit applied to stationary hop picking machine. Computers and Electronics in Agriculture 95: 92-97.

Kumhala, F., Prosek, V. and Blahovec, J., 2009. Capacitive throughput sensor for sugar beets and potatoes. Biosystems engineering 102(1): 36-43.

Kumhala, F., Prosek, V. and Kroulik, M., 2010. Capacitive sensor for chopped maize throughput measurement. Computers and Electronics in Agriculture 70(1): 234-238.

Lee, Y.-J., Kim, K.-D., Lee, H.-S. and Shin, B.-S., 2018. Vision-based Potato Detection and Counting System for Yield Monitoring. Journal of Biosystems Engineering 43(2): 103-109.

Liakos, V., Tagarakis, A., Aggelopoulou, K., Fountas, S., Nanos, G.D., Gemtos, T., 2017. In-season prediction of yield variability in an apple orchard. European journal of horticultural sciences 82 (5): 251-259.

Maja, J.M. and Ehsani, R., 2010. Development of a yield monitoring system for citrus mechanical harvesting machines. Precision Agriculture 11(5): 475-487.

Nuske, S., Wilshusen, K., Achar, S., Yoder, L., Narasimhan, S. and Singh, S., 2014. Automated Visual Yield Estimation in Vineyards. Journal of Field Robotics 31(5): 837-860.

Panneton, B. and Brouillard, M., 2005. Precision agriculture for baby-greens – An efficient management tool. In Proceedings of the Fruits and Nuts and Vegetable Production Engineering TIC (Frutic05) Conference, Montpellier, pp. 571-582. http://www.symposcience.eu/exl-doc/colloque/ART-00001715.pdf (last accessed February 2019).

Panneton, B. and St-Laurent, G., 1999. Performance of yield monitors for root crops. Paper No. 99-1054. St. Joseph, MI, USA: ASAE.

Payne, A.B., Walsh, K.B., Subedi, P.P. and Jarvis, D., 2013. Estimation of mango crop yield using image analysis – Segmentation method. Computers and Electronics in Agriculture 91: 57-64.

Roy, P., Kislay, A., Plonski, P.A., Luby, J. and Isler, V., 2018. Vision-based preharvest yield mapping for apple orchards. arXiv preprint arXiv:1808.04336.

Ruiz-Altisent, M., Ruiz-Garcia, L., Moreda, G.P., Lu, R., Hernandez-Sanchez, N. and Correa, E.C., 2010. Sensors for product characterization and quality of specialty crops – A review. Computers and Electronics in Agriculture 74(2): 176-194.

Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T. and McCool, C., 2016. DeepFruits: A Fruit Detection System Using Deep Neural Networks. Sensors 16(8): 1222-1244.

Song, Y., Glasbey, C.A., Horgan, G.W., Polder, G., Dieleman, J.A. and van der Heijden, G.W., 2014. Automatic fruit recognition and counting from multiple images. Biosystems Engineering 118: 203-215.

Syal, A., Garg, D. and Sharma, S., 2013. A survey of computer vision methods for counting fruits and yield prediction. International Journal of Computer Science Engineering 2(6): 346-350.

Vatsanidou, A., Nanos, G.D., Fountas, S., Baras, J., Castrignano, A., Gemtos, T.A., 2018. Nitrogen replenishment using variable rate application technique in a small hand-harvested pear orchard. Spanish Journal of Agricultural Research 15(4): 209-222.

Vellidis, G., Durrence, J.S., Hill, R.W., Perry, C.D., Thomas, D.L. and Hamrita, T.K., 2003. Crop yield monitoring system. U.S. Patent No 6,525,276, 25 Feb. 2003.

Zhang, Q., 2017. Automation in Tree Fruit Production: Principles and Practice. Boston, USA: CABI 304 pp.

Zude-Sasse, Fountas, S., Gemtos, T.A., Abu-Khalaf, N., 2016. Applications of precision agriculture in HC. European Journal of Horticultural Science, 81, 78-90.

Preliminary study for weed biomass prediction combining visible images with a plant-growth model

J. Merienne, A. Larmure and C. Gée*
Agroécologie, AgroSup Dijon, INRA, Univ. Bourgogne, Univ. Bourgogne Franche-Comté, 21000 Dijon, France; christelle.gee@agrosupdijon.fr

Abstract

This preliminary study was designed to test an image approach (RGB images) to estimate crop and weed aerial biomass (BM) and predict their growths. From an image processing, based on crop/weed discrimination, a calibration was done between photographed leaf area (PLA) and BM deduced from a standard (destructive) approach. Then, a plant growth model component (AZODYN) was used with PLA as input, considering wheat and weeds separately. For wheat, the results confirm that to initialize the model, the image approach may replace the standard one with no impact on the values of predicted BM whereas weed results are more mixed.

Keywords: agroecology, competitiveness, visible image, prediction, weed control, wheat

Introduction

In conventional agriculture, weed control is often achieved by the use of herbicides as curative treatment raising environmental and health problems and, sometimes, leading to resistance to herbicide (Heap, 2013). To reduce their negative impacts (Mézière *et al.*, 2015), new weed management strategies need to be implemented. Usually, the main factors affecting decision-making of farmers in weed control are based on the presence of certain weed species and their density (plants/m^2). However, weed management at the scale of the crop cycle refers to other more reliable competition indicators such as plant biomass (BM) or leaf area index (LAI). Most of the time, measurement of these attributes are obtained by destructive methods of sampling which are time consuming and which are unworkable in practice for farmers. Hence, an alternative to traditional destructive methods is the development of non-destructive methods that are time and cost effective. They usually refer to multispectral imaging systems (visible and near-infrared range) with the development of vegetation indices (Meyer and Neto, 2008) to characterize vegetation at different spatial scales (from plant to field) or to map weeds in crop fields (Peña *et al.*, 2013). Although many studies report on site-specific-weed management or effect of weed, predicting weed biomass is more complex to attempt because it relies on all the components of the cropping system. This preliminary study is a part of a larger project which aims to develop a decision support tool to determine the ultimate date of weeding intervention in a wheat crop taking into account the effect of weeds on wheat crop through a biomass prediction. One promising way is to estimate daily aerial biomass during the vegetation phase for crop and weed respectively under known climate conditions with a user-friendly plant growth model initially fed by vegetation parameter (photographed leaf area, PLA) deduced from images. An accurate weed dynamics model is available in the literature (FLORSYS, Colbach *et al.*, 2014) but it aims to simulate weed dynamics over the years where the basic rotational pattern is repeated (e.g. oilseed rape/wheat/barley).
So, the first objective of this paper was to test a user-friendly model of biomass prediction for a monoculture (wheat *vs* weed) to evaluate biomass during a short-time (at the 2 leaf stage up to tillering stage) of crop cycle to propose a date of weeding using digital tools. Then, from Monteith equations (derived from Beer-Lambert law), a component of the Azodyn plant model (Jeuffroy and Recous, 1999) was used considering PLA (rather than BM) as initial parameter to simulate aerial

DOI 10.3920/978-90-8686-888-9_74, © Wageningen Academic Publishers 2019

biomass either for wheat crop or weeds. Usually applied to wheat crops, from the end of winter until flowering, the Azodyn-wheat model aims to improve nitrogen management. A second objective was thus to validate the use of RGB images to easily extract the initial (crop and weeds) parameters (PLA) required to feed a dynamic plant growth model.

Materials and methods

Plant material and growth conditions

A field experiment was sown on 12th November 2017 at Dijon (47.301°N; 5.066°E; France) on a chalky-clay deep soil. Wheat plants (*Triticum aestivum* L.) cv. *Apache* were grown on a micro-plot (15×1.20 m): 7 rows, 0.15 m apart.

Measurements

Plants (wheat and weeds) have been sampled at 4 different dates expressed as Julian day numbers, from the middle of March until the middle of April, with three repetitions (0.342 m^2-quadrat), at each date. At the end of the experiment, the wheat phenological stage is the end of tillering (BBCH stage N°29). Weeds were not controlled. Weeds were identified: mostly dicotyledonous and annual species; then wheat and weed plants were characterized separately according to two approaches, a standard (destructive) and an image one.

- the standard approach was based on a destructive determination of the dry matter biomass of aerial plant parts (BM, g/m^2): at each sampling date, wheat and weed plants were collected separately and weighed after oven drying at 80 °C for 48 h.
- the image approach is based on image analysis of the Photographed Leaf Area (PLA, dimensionless). Before the destructive approach, RGB images (image size was equal to quadrat size) were acquired using a digital camera Canon EOS 450D reflex (f=35 mm; image resolution: 2,848×4,272 px) under natural sunlight at solar noon. The image spatial resolution with this digital camera mounted one meter above the plant targets translated into 0.016 mm per pixel. Vegetation (wheat + weeds) images were obtained using the excess green vegetation index (ExG) developed by Woebbecke *et al.* (1995). Then a conversion of the index near-binary to a full-binary image is done by an Otsu threshold value (Otsu, 1979). Total Photographed Leaf Area is evaluated by counting the fraction of white pixels relative to the total number of pixels of the vegetation image. A subdivision of the vegetation image into two classes, a wheat image and a weed image, was obtained from an unsupervised classification based on Hough transform (Hough, 1962) and a region-based segmentation as explained in Jones *et al.* (2009). From wheat crop image (respectively weed image), the PLA is deduced as the ratio of the number of white pixels to the total number of pixels in the image.

Daily photosynthetically active radiation was measured by amorphous silicon sensors.

Description of the plant-growth model component

A plant model component was used in order to predict separately the growth of wheat crop and weeds (Jeuffroy and Recous, 1999). It calculates dry matter biomass production of aerial plant parts, at a daily time-step during the vegetative phase, according to Monteith (1972, 1977), starting from an initial crop or weeds aerial biomass measurement (BM_initial, g/m^2). Main parameters used by the model are presented in Table 1. Biomass production is based on the daily photosynthetically active radiation (Input, MJ/m^2), the radiation intercepted by the plants (parameters: εimax-maximum radiation interception efficiency, k-radiation extinction coefficient), and the radiation use efficiency (parameter εb).

Table 1. Main parameters of the crop model component used for wheat crop and weeds.

Parameter	Wheat crop value	Weeds value
εb	2.8[(a)]	3.49[(b)]
k	0.72[(a)]	0.83[(c)]
εimax	0.96[(a)]	0.96[(a)]

(a) Jeuffroy and Recous, 1999; (b) Zhu *et al.*, 2008, (c) Monteith, 1969

Results

Wheat and weeds growth characterization.
About fifteen weed species have been observed in the wheat crop, mostly dicotyledonous and annual species. So, as a first approximation, weeds will be considered as a single homogenous emergent cohort to predict their aerial biomass by the plant-growth model. It should be noticed that most of weeds are representative of maize crop rather than wheat crop. This particular weed flora can be explained by the history of the plot with maize crop during the last three years and no tillage prior to wheat seeding. Dry matter biomass (BM) and Photographed Leaf Area (PLA) of wheat crop and weeds are presented in Figure 1A and B. Considering wheat, both variables increased over time from 28 to 78 g/m^2 and from 0.15 to 0.33 respectively. Considering weeds, BM also increased over time except for the third date of sampling (12th of April) (Figure 1B). Hypothesis is that in the case of the third date, the experimental values are very low due to the presence of fewer plants. With PLA values, surprisingly a small decrease is observed for the second date with a flat signal for the last dates. However, values of both variables were much lower for weeds than for wheat from 0.8 to 8.8 g/m^2 and from 0.03 to 0.13 respectively for BM and PLA. Moreover, the variability between the 3 repetitions at each date is high for both variables (i.e. high standard deviation in Figure 1B). These results reveal the strong spatial heterogeneity of weeds, even in a micro-plot.

Calibration process based on linear correlations between PLA and BM
The value of an initial measured aerial biomass is need for the simulations by the plant growth model component used in this study (BM_initial, input data). The objective of the calibration was to determine if it is possible to feed this model with image data (PLA) instead of destructive measured data (BM). Considering wheat, PLA and BM are highly linearly correlated (R^2=0.87) (Figure 2A).

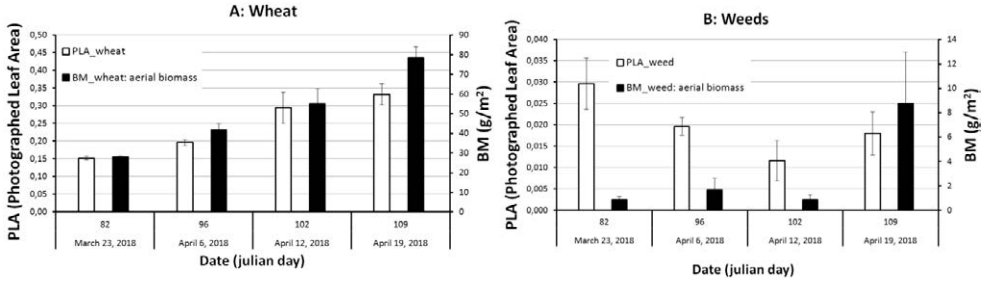


Figure 1. Evolution of dry matter biomass (BM) and photographed leaf area (PLA) for wheat (A) and weeds (B) for both approaches, standard and image, respectively. For each date, the mean value and standard deviation of the three repetitions is represented.

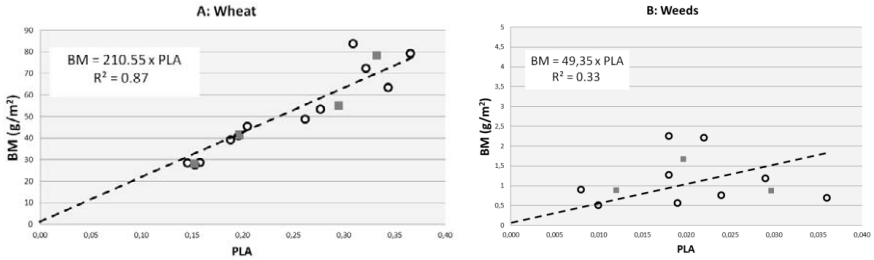


Figure 2. Linear regression (dashed line) between PLA and BM values (unfilled black circles) displayed for wheat crop (A) and weeds (B). Grey solid squares represent mean values for each date.

Considering weeds with their high spatial variability and the presence of few plants per sample, the linear correlation (Figure 2B) between PLA and BM has been done on mean values. The result is not so good (R^2=0.33) compared to results on wheat crop but the approach is worth pursuing. Thus PLA can be a pertinent input variable representing aerial biomass for the initialization of the growth model but for weeds, some improvements are required.

Plant-growth model to predict the biomass evolution of wheat and weed

Simulations were done with the model with the aim of verifying that the predicted results were consistent with observed data. The plant growth model has been tested separately for wheat crop and weeds. The model was fed with an initial biomass obtained at Date 1, by the image approach (PLA) for wheat (Figure 3A) and weeds (Figure 3B) respectively. The linear calibration (see above) has been used to transform PLA data into BM data as initial parameters. Then the BM values predicted by the model were compared with experimental BM values deduced from the standard approach. The predicted biomass was always superior to the experimental data probably because it reflects plant-growth under optimal conditions.

For wheat crop, the predictions of the model are quite good especially in the early dates (Date 2 and Date 3) when the plants are young with a relative error e_r (e_r = 100 × BM_simulated / BM_exp_mean), of 27.3 and 33.56% for Date 2 and 3 respectively. These values are consistent with the literature (Jeuffroy and Recous, 1999).

It is also noted that the errors between model results and those of the experiments are amplified over time because the model is based on the daily accumulation of the biomass translating cumulative errors and so their propagation.

The overall agreement between field and predicted values for wheat crop was verified through a linear regression between simulated data and experimental data, with a R^2 of 0.90. Moreover, the predicted data by the model with an initialization by the standard approach (BM) or by the image approach (PLA) are very similar (R^2=1). In conclusion, these results confirm that, in the case of wheat crop, to initialize the model the standard approach (BM) may be replaced by the image approach (PLA) with no impact on the values of predicted BM.

Concerning the weed population, the predicted values are quite consistent in comparison with experimental values (BM) and a maximum difference (relative error, e_r=86%) between model and experience has been observed for the third date (julian day=102). At this date, it was noticed that the collection has occurred at a location with fewer plants (Figure 1B), suggesting a significant spatial heterogeneity of weeds. It is emphasised that whatever the approach (standard or image) to initialize the model, the predicted BM values are very similar (R^2=0.99).

But in contrast to the case of wheat, the image approach needs improvements to replace the standard approach to initialize the model. Despite the important value of e_r in Date 3, these results confirmed

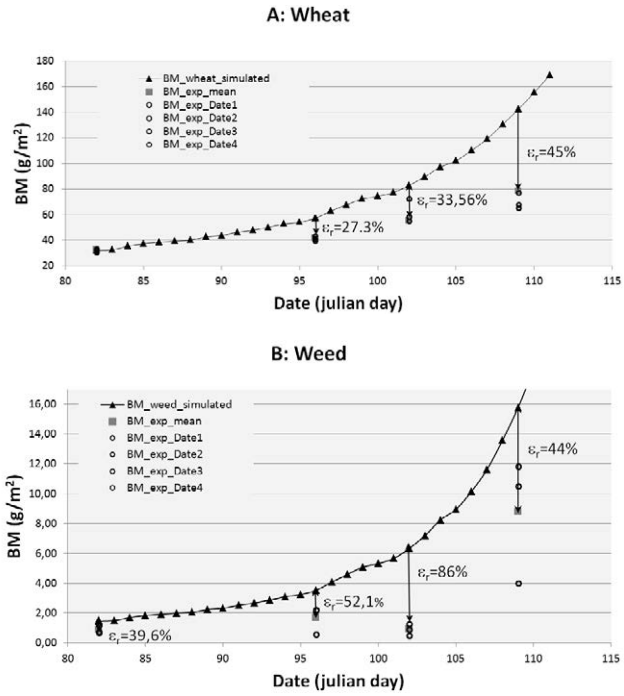


Figure 3. Predicted biomass (black line with black triangles) starting from the initial aerial biomass deduced from PLA with image approach for wheat crop (A) and weeds (B). Grey solid squares represent mean experimental values for each date and unfilled black circles represent the BM_exp_values deduced from destructive approach. The relative errors (e_r) between predicted and experimental mean values are shown on the figure.

the overall trend between field and predicted values with a R^2 of 0.81 deduced from a linear regression between simulated data and experimental data.

Discussion

Validation of the image approach

First of all this work showed the good results of the use of PLA data from the image approach to feed a plant-growth model. This approach will greatly simplify experimental fieldwork because it is time effective compared with the traditional method. Destructive measurements (BM) both tedious and time consuming were transposed into PLA, deduced from RGB images, with operational image processing, especially efficient at early stages of vegetation growth. Indeed, the model which predicts the daily potential biomass is very efficient at the end of winter, where the vegetation is young. Moreover, in a context of spring weed control, the experimental measurements were carried out too late (mid-March) for early weed control. For future work, it would be necessary to realize the first experimental measurements in February. At higher growth stages, model results are substantially moving away from experimental measurements. To explain this increase in errors, different interpretations can be proposed: (1) the model accumulates the daily potential biomass (without stress) and therefore if there are stress and errors, they will spread and amplify over time; (2) as regards the PLA values resulting from image processing, the discrimination algorithm crop/weed becomes less and less efficient over time due to the greater overlap of the plant leaves. The weed images issued of the discrimination algorithm results indicate that PLA values of weed is always

underestimated for the third and fourth date implying difficulties in the calibration process. That is why a manual counting of weed pixels is investigated to quantify algorithm errors and determine the exact PLA values to validate the calibration process on weed plants (work is in progress). Nevertheless, the algorithm needs to be optimized for a better crop/weed discrimination.

A model to estimate weed biomass

The model used in this study is usually developed to estimate the biomass of a crop stand and particularly wheat, barley or rapeseed crop (Jeuffroy and Recous, 1999; Jeuffroy *et al.,* 2006; Beillouin *et al.*, 2018). The aim of this work was to estimate a weed stand and therefore to test this simple and easy to use model and see its potential on such a stand. Beyond an experimental error in locating the repetitions of each Date (three nearby sites chosen instead of distant sites from each other), it is found that for young growth stages of vegetation, the results of the model are acceptable and, in the short-term, they could help the farmer in his spraying strategy. In order to work towards this idea, it would then require to class weeds by species and to calibrate the model for each of them rather than considering all the weed species as a single stand. This would then provide a best estimate of the biomass for spot-spraying management. Nonetheless, the image analysis will have to be greatly improved to identify and class weed plants into different groups associated with herbicides depending on their spectra of action. Finally, other, much more complex models (e.g. STICS intercrop; Corre-Hellou *et al.*, 2009), taking into account the complexity of species interactions must be investigated.

A new indicator to optimize weeding strategy: weed/crop biomass ratio?

The sustainable management of weeds is most often based on a historical knowledge of weeds in the plot and on another indicator, the plant density (number of plants per m^2). However, recent studies have questioned the use of weed density as an indicator of weed competitiveness (Colbach and Cordeau, 2018). Indeed, the number of plants as an indicator does not account as well as the pressure of weeds on the crop as other variables. Hence, a new indicator based on wheat and weed biomass measurement could be studied by looking at the weed/crop biomass ratio. In this article, results about weed are still preliminary and further work is required to study the temporal evolution of this ratio so as to estimate the wheat yield loss as a function of the weed biomass. Then a critical threshold of weed competitiveness could be determined by a simple image analysis (estimation of PLA). Thus, using a plant-growth model fed by a few visible images, the farmer will be able to have an automatic estimate of the weed competitiveness in his plot in order to help him for decision making, avoiding as far as possible the use of spring weed control.

Conclusions

This work explored the use of numerical tools (images and plant model) to estimate the aerial biomass of wheat and weeds using Photographed Leaf Area (PLA). This biomass has been estimated separately for wheat and weeds using a crop/weed discrimination algorithm, which is particularly effective at early stages of growth. Then, these data fed a plant-growth model based on Monteith equations to predict BM over time in order to determine a new competitive indicator, weed/crop biomass ratio. In view of these encouraging results, it is necessary to progress on the image processing algorithm and the judicious choice of a model that integrates the growth of two populations (crop and weed) at the same time.

Acknowledgements

Many thanks to Vincent Durey and Annick Matéjicek who were involved in this project, one on PAR sensor control and the other for plant identification.

References

Beillouin D., Leclère M., Barbu C. M., Bénézit M., Trépos R., Gauffreteau A., M-H., 2018. Azodyn-Barley, a winter-barley crop model for predicting and ranking genotypic yield, grain protein and grain size in contrasting pedoclimatic conditions. Agricultural and Forest Meteorology 262, 237-248

Colbach, N., Collard, A., Guyot, S.H.M., Mézière, D., Munier-Jolain, N., 2014. Assessing innovative sowing patterns for integrated weed management with a 3D crop: weed competition model. European Journal of Agronomy 53, 74-89.

Colbach, N., Cordeau, S., 2018. Reduced herbicide use does not increase crop yield loss if it is compensated by alternative preventive and curative measures. European Journal of Agronomy 94 (2018) 67-78

Corre-Hellou G., Faure M., LaunayM. Brisson N., YvesCrozat Y., 2009. Adaptation of the STICS intercrop model to simulate crop growth and N accumulation in pea–barley intercrops. Field Crops Research 113 (1), 72-81

Heap, I., 2013. The International Survey of Herbicide Resistant Weeds. Online Internet. Available: www.weedscience.org (last accessed February 2019).

Hough, P.V.C., 1962. Method and means for recognizing complex patterns. U.S. Patent Office No. 3069654.

Jeuffroy M-H., Valantin-Morison M., Champolivier L., Reau R., 2006. Azote, rendement et qualité des graines: mise au point et utilisation du modèle Azodyn-colza pour améliorer les performances du colza vis-à-vis de l'azote (Nitrogen, yield and seed quality: Design and uses of the Azodyn-rape model to improve rape performance for nitrogen). Oilseeds and Fats, Crops and Lipids 13 (6) 388-392 https://doi.org/10.1051/ocl.2006.0090

Jeuffroy, M.-H., Recous, S., 1999. Azodyn: a simple model simulating the date of nitrogen deficiency for decision support in wheat fertilization. European Journal of Agronomy 10, 129-144.

Jones, G., Gée, Ch., Truchetet, F., 2009. Modelling agronomic images for weed detection and comparison of crop/weed discrimination algorithm performance. Precision Agriculture 10:1-15. https://doi.org/10.1007/s11119-008-9086-9

Meyer, G.E., Neto, J.C., 2008. Verification of color vegetation indices for automated crop imaging applications. Computers and Electronics in Agriculture 63, 282-293. https://doi.org/10.1016/j.compag.2008.03.009

Mézière, D., Petit, S., Granger, S., Biju-Duval, L., Colbach, N., 2015. Developing a set of simulation-based indicators to assess harmfulness and contribution to biodiversity ofweed communities in cropping systems. Ecological Indicators 48, 157-170.

Monteith, J.L., 1969. Light Interception and Radiative Exchange in Crop Stands. In Physiological Aspects of Crop Yield, pp. 89-91. Madison, WI, USA: American Society of Agronomy & Crop Science Society of America.

Monteith, J.L., 1972. Solar Radiation and Productivity in Tropical Ecosystems. Journal of Applied Ecology 9, 747-766. https://doi.org/10.2307/2401901

Monteith, J.L., 1977. Climate and the efficiency of crop production in Britain. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 281, 277-294.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man and Cybernetics 9, 62-66.

Peña, JM., Torres-Sánchez, J., de Castro, AI., Kelly, M., López-Granados, F., 2013. Weed mapping in Early-Season Maize Fields Using Object-Based Analysis of Unmanned Aerial Vehicle (UAV) Images. PLoS ONE. 8: e77151. pmid:24146963

Woebbecke, D.M., Meyer, G.E., Von Bargen, K., Mortensen, D.A., 1995. Color indices for weed identification under various soil, residue, and lighting conditions. Transactions of the ASAE 38, 259-269.

Zhu, X.-G., Long, S.P., Ort, D.R., 2008. What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? Current Opinion in Biotechnology, 19(2), 153-159. https://doi.org/10.1016/j.copbio.2008.02.004

Using remote sensing to map in-field variability of peanut maturity

A.F. Santos[2], L.N. Lacerda[1], S. Gobbo[1], A. Tofannin[1], R.P. Silva[2] and G. Vellidis[1*]
[1]*University of Georgia, Tifton Campus, Tifton, Georgia, USA; yiorgos@uga.edu*
[2]*São Paulo State University (UNESP), Jaboticabal Campus, Jaboticabal, São Paulo, Brazil*

Abstract

A study was conducted to assess if vegetation indices (VIs) could be used as indicators of peanut maturity. A 24 ha block of a rainfed field with clearly visible soil and crop variability was used. An unmanned aerial vehicle (UAV) equipped with a multispectral camera captured spectral reflectance from the peanut canopy. The reflectance data were used to evaluate several VIs as potential indicators of peanut maturity. Pearson's correlation and linear regression were used to evaluate the response of the VIs as well as the sensitivity of individual reflectance bands to peanut maturity. The red edge band was the most sensitive. The most responsive VIs were the non-linear index (NLI) and the modified non-linear index (MNLI) when red edge reflectance was substituted for red reflectance.

Keywords: vegetation index, multispectral images, reflectance, red edge, UAV

Introduction

The peanut, or groundnut (*Arachis hypogaea*), is a low growing crop which produces its fruit, known as a peanut pod, under-ground. After fertilization of the flower, a stalk commonly referred to as a peg, elongates with the ovary and grows down into the soil. The peg's tip develops into a mature peanut pod (Smith, 1950). Depending on variety, peanut pods are 20 to 40 mm long and typically contain one to three seeds. Because peanut is an indeterminate crop, it produces flowers for extended periods during the growing season. As a result, peanut pods of varying maturity are present on the plant as harvest time approaches. Harvesting a field at peak maturity results in both the best quality and yields. Immature pods can have decreased seed quality and oil content (Fincher *et al.*, 1980), while overly mature pods result in yield losses of eight to 40% as the peanut pods detach from the pegs (Lamb *et al.*, 2004). Determining peak maturity is difficult as there are no obvious aboveground indicators.

Williams and Drexler (1981) found that as peanut pods mature, the pods' mesocarp (second layer of the pod's skin) changes color from white to black. They then developed the hull-scrape method for determining peanut maturity. The exocarp can be removed by pressure washing with water, using abrasion, or hand scraping the hull (Williams and Drexler, 1981; Williams, 2003). Today, peanut growers apply this method by collecting plants from four or five locations within a field and pressure washing approximately 200 peanut pods. The color of the pods is then used to place the peanuts in one of six major color classes on a peanut profile board (Figure 1). Where the leading edge of the sample profile intersects the projection line on the board predicts the approximate days until the peanuts must be inverted in preparation for harvest. Peak maturity occurs when the leading edge of the sample profile reaches the slope line on the peanut profile board.

An alternative method for predicting maturity using the hull-scrape method is the peanut maturity index (PMI; Rowland *et al.,* 2006). The PMI is calculated by dividing the number of mature pods (brown + black) by the sum of the total pods in the sample. Optimal maturity occurs when a sample has a PMI of 0.7. Overall, the hull-scrape method is labor intensive, subjectively based on a person's ability to place pods in the correct color category, and subject to field variability that may affect maturity of individual plants. Alternatives are needed that can estimate the maturity of entire fields rather than individual plants.

DOI 10.3920/978-90-8686-888-9_75, © Wageningen Academic Publishers 2019

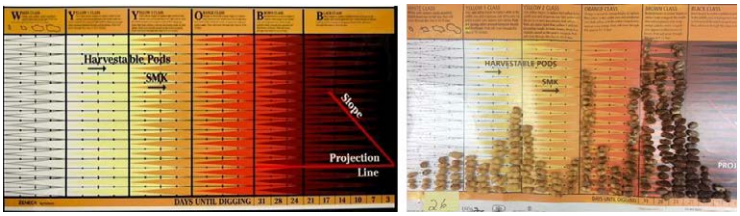


Figure 1. Peanut Profile Board used for rating peanut maturity (left) and with a sample of pressure-washed peanuts with exposed mesocarp (right).

Researchers across the globe have established that variability exists in all aspects of crop response within most fields. This also holds true for the rate at which peanut matures in different areas of a field (Vellidis and Beasley, 2013). The use of remote sensing to predict peanut maturity at a field scale has been explored in the past. Rowland *et al.* (2008) found that the reflectance in the near infrared (NIR) band is best correlated to peanut maturity when compared to other bands. Changes in peanut reflectance in the RE (red edge) band as peanuts approach maturity indicate that these wavelengths are also sensitive to pod maturity (Robson *et al.*, 2006). VIs have been shown to be a good tool for monitoring maturity in various crops (Yu *et al.*, 2016).
Sanders *et al.* (2002), Robson *et al.* (2006), Carley *et al.* (2008) and Rowland *et al.* (2008) evaluated VIs as an indicator of peanut maturity but they were unable to identify an index that consistently predicted peanut maturity. Vellidis and Beasley (2013) used a tractor-mounted optical reflectance sensor to evaluate the response of VIs to peanut maturity on a plot-scale. They found that NLI was a good predictor of maturity for many peanut varieties. The objective of this study was to determine if these methods can be used to predict peanut maturity variability at the field scale.

Materials and methods

The study was conducted in a rainfed commercial peanut field (31°27'38.20'N, 82°38'15.10'W) located in Coffee County, Georgia, USA. The peanut cultivar used was GA-06, a Runner-type variety with dark green foliage that typically requires about 140 days to maturity. Historical satellite images (RGB composition) of the field were analyzed to select a block of the field with the most soil variability. The block was divided into twenty-four 1 ha sampling cells (Figure 2).
Ten peanut plants were collected from an area approximately 10 m in radius at the center of the sampling cells weekly for the 6 weeks prior to harvest (Figure 3). Two hundred peanut pods were removed from each group of plants and pressure-washed to expose the mesocarp. The pods were placed on the peanut profile board and the number of nuts in each color class recorded. The PMI was also calculated for each sample.
During each sampling day, a Sequoia multispectral camera (Parrot, Paris, France) attached to a 3DR Solo quadcopter UAV (3DR, Berkeley, USA) was used to capture multispectral images in 4 different bands; Green (530-570 nm), Red (640-680 nm), NIR (770-810 nm) and RE (730-740 nm) (Figure 3). All UAV flights were within 2 h of solar noon from an altitude of 90 m with 70% overlap. At 90 m, the spatial resolution of the camera was 0.95 m. The Tower Drone Control App (3DR, Berkeley, USA) was used to create flight plans. Before and after each flight, the camera was used to take images of a calibration panel that were later used to calibrate the images acquired over the field. Black and white 1×1 m targets were located at the corners of the 24 ha block to serve as ground control points (GCPs). The geographic co-ordinates of the four GCPs were recorded with a Trimble R8 GNSS (Trimble, Sunnyvale, USA). The Pix4D software (Parrot, Paris, France) was used to create mosaics of reflectance maps in each one of the four bands, to perform radiometric calibration, and perform



Figure 2. The study was conducted in a 24 ha block of a rainfed field located in Coffee County, Georgia, USA. The block was divided into 1 ha sampling cells.



Figure 3. Collecting peanut plant samples (left) and the 3DR Solo quadcopter UAV with the Parrot Sequoia camera used to collect multispectral images (right).

georeferencing corrections. After the mosaics were created, ArcGIS software (Redlands, USA) was used to extract an average reflectance value for the pixels in each of the 24 grid cells.

Seven VIs that have been used by other studies to predict crop maturity were selected for evaluation (Table 1). These VIs are hereafter referred to as the 'original' VIs. The average reflectance values for each grid cell were then used to calculate the response of the seven VIs for each sampling date. Pearson's correlation (P<0.001) and regression analysis were used to compare the response of the VIs to PMI.

A sensitivity analysis was conducted to identify which of the four reflectance bands was most sensitive to peanut maturity. Pearson's correlation (P<0.001) and regression analysis were used for the sensitivity analysis and the band with the highest R^2 value when F was significant was selected. The highest correlation with maturity was found for the RE band. Red reflectance was replaced with RE reflectance in three of the original VIs which included Red reflectance in their formula. These VIs are thereafter referred to as 'modified' VIs (Table 1). In total, ten different VIs were evaluated.

Results and discussion

Figure 4 shows the response of six of the original VIs. All VIs had a negative linear relationship with PMI where the value of the VI decreased as PMI increased. The normalized difference vegetation index (NDVI) (Figure 4A) and green NDVI (GNDVI) were correlated to PMI (P<0.001) but their

Table 1. Original and modified VIs used in the study.

VI	Formula	Reference
Original VIs		
NDVI	(NIR – RED) / (NIR + RED)	Rouse *et al.* (1974)
GNDVI	(NIR – Green) / (NIR + Green)	Gitelson *et al.* (1996)
NLI	$(NIR^2 – RED) / (NIR^2 + RED)$	Goel and Qin (1994)
MNLI	$(NIR^2 – RED) – (1 + L) / (NIR^2 + RED + L)$*	Gong *et al.* (2003)
NDRE	(NIR – RE) / (NIR + RE)	Gitelson & Merzlyak (1996)
SR	NIR / RED	Jordan (1969)
MSR_{RE}	$(NIR / RE – 1) / ((NIR / RE)^{1/2} + 1)$	Wu *et al.* (2008)
Modified VIs		
SR_{RE}	NIR / RE	
$MNLI_{RE}$	$(NIR^2 – RE)(1 + L) / (NIR^2 + RE + L)$*	
NLI_{RE}	$(NIR^2 – RE) / (NIR^2 + RE)$	

*L=0.5.

R^2 values were very low likely because these VIs saturate with full canopy closure (Corti *et al.*, 2018). In fact, there was little change in NDVI and GNDVI values during the 6-week sampling period since sampling began after canopy closure in most areas of the field. Similar results were found by Carley *et al.* (2008). In contrast, Sanders *et al.* (2002) identified NDVI as a promising index to predict peanut maturity in Runner and Valencia peanut types. Rowland *et al.* (2008) found that GNDVI calculated from reflectance data measured with a proximal optical sensor increased the longer plants stayed in the field. The Rowland *et al.* (2008) finding is opposite to the GNDVI response found by the current study.

NLI (Figure 4B) and MNLI (Figure 4C) had significant negative correlation ($P<0.001$) with PMI but poor R^2 values. In Vellidis *et al.* (2013), NLI showed good response in predicting maturity, with the slope of NLI consistently decreasing and approaching zero at maturity. However, a consistent change in slope was not observed in the current study.

Among all seven indices, normalized difference RE (NDRE) (Figure 4D) had the strongest relationship with PMI ($R^2=0.534$). The modified simple ratio RE (MSR_{RE}) which includes the RE band (Figure 4E), had the second strongest relationship with PMI ($R^2= 0.506$). However, none of the original VIs provided an inflection point or indicator of peanut maturity other than a steady decline in value.

A strong negative Pearson correlation and $R^2>0.4$ was observed for all modified VIs, which used the RE reflectance in place of red reflectance (Figure 5). These results indicate that including RE reflectance in the VIs adds significant power to their ability to track maturity in peanut. This is likely because RE reflectance is more sensitive to small physiological changes in the plant leaves as the pods mature.

The improved performance of the modified VIs can be related to the penetration capacity of RE wavelengths into the leaf and their ability to respond to small changes in chlorophyll *a* even in conditions of elevated biomass (Li *et al.*, 2014) in which many other VIs are prone to saturate. However, changes in chlorophyll *a* absorption can be caused by many different factors including water stress and disease and are not necessarily directly related to maturity.

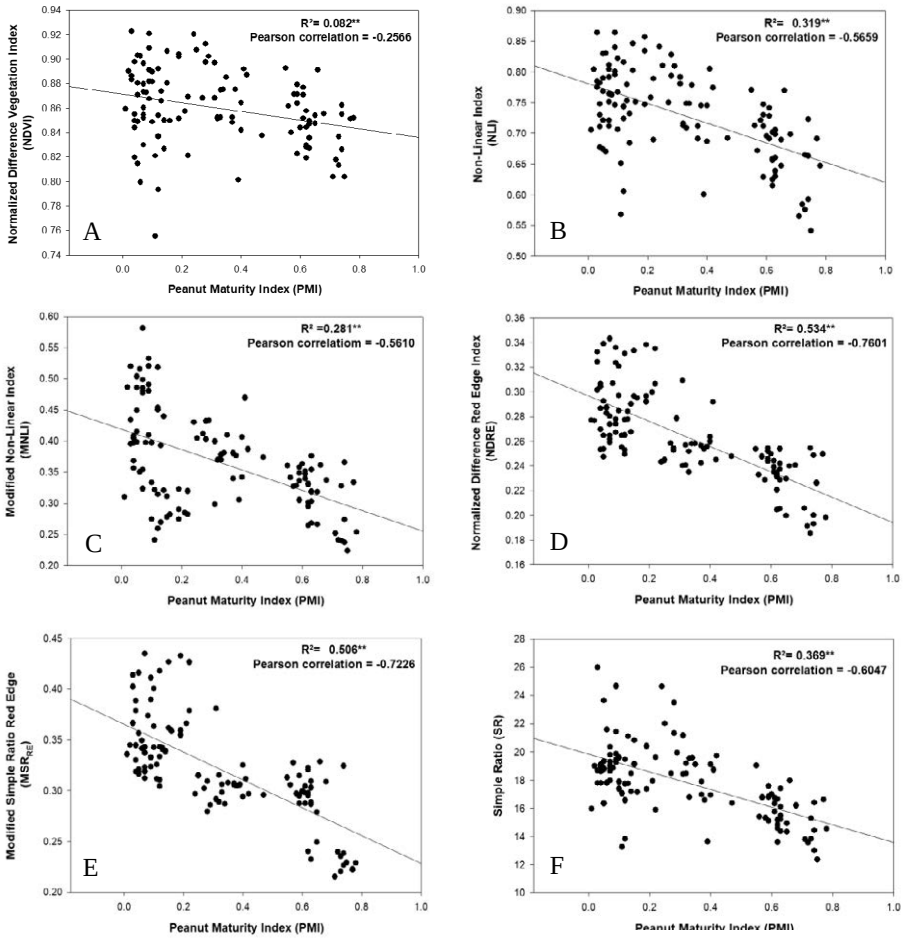


Figure 4. Regression analysis between (A) NDVI, (B) NLI, (C) MNLI, (D) NDRE, (E) MSR_{RE} and (F) SR and Peanut Maturity Index (PMI). (Pearson Correlation: $P<0.001$).

After replacing red reflectance with RE reflectance, NLI_{RE} (Figure 5A) showed an even higher correlation with maturity (Pearson=-0.8181) and an R^2 of 0.671. This modified version of NLI was the best performing VI among the ten VIs evaluated. A similar trend was observed for $MNLI_{RE}$ (Figure 5B) and SR_{RE} (Figure 5C) as both resulted in higher correlation and R^2 values than their original versions.

Conclusions

The use of remote sensing has a great potential to predict peanut maturity, thus improving farmers' efficiency in selecting the right time to harvest peanuts. The majority of VIs studied were correlated to maturity although most had poor R^2 values. NDVI and GNDVI had the worst response likely due to saturation. NLI_{RE} and $MNLI_{RE}$ were strongly related to PMI and may be good indicators of peanut maturity. However, a specific VI value cannot yet be used to predict optimal peanut maturity. Additional research is needed to replicate this work in more fields over two or more years with

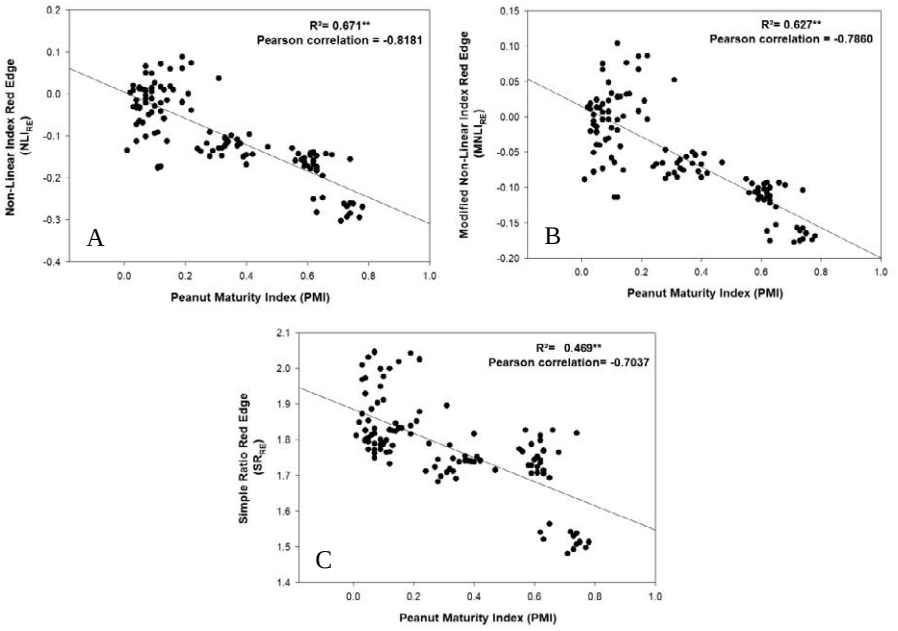


Figure 5. Regression analysis between (A) NLI_{RE}, (B) $MNLI_{RE}$, (C) SR_{RE} and Peanut Maturity Index (PMI). (Pearson Correlation: $P<0.001$).

different varieties to confirm these results. A larger number of data collection days may also assist in evaluating the trend lines of the VIs. Nevertheless, the results are encouraging and indicate that remote sensing may provide an alternative to the Hull-Scrape Method.

Acknowledgements

This project was funded by the Georgia Agricultural Commodity for Peanuts in partnership with the Brazilian Research Funding Agency CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior). We thank Mr Drew Walker for allowing us to conduct the study in his field.

References

Carley, D. S., Jordan, D. L., Dharmasri, L. C., Sutton, T. B., Brandenburg, R. L. and M. G. Burton, 2008. Peanut Response to Planting Date and Potential of Canopy Reflectance as an Indicator of Pod Maturation. Agronomy Journal 100 376-380

Corti, M., Cavalli, D., Cabassi, G., Gallina, P.M., Bechini, L., 2018, Does remote and proximal optical sensing successfully estimate maize variables? A review, European Journal of Agronomy 99 37-50.

Fincher, P.G., Young, C.T., Wynne, J.C. and Perry, A., 1980. Adaptability of the arginine maturity index method to Virginia type peanuts in North Carolina. Peanut Science 7 83-87.

Gitelson, A.A., Kaufman, Y.J., Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment 58 289-298.

Gitelson, A.A., Merzlyak, M.N., 1996. Signature analysis of leaf reflectance spectra: algorithm development for remote sensing of chlorophyll. Journal of Plant Physiology 148 494-500.

Goel, N. S. and Qin, W., 1994. Influences of canopy architecture on relationships between various vegetation indices and LAI and FPAR: A computer simulation. Remote Sensing Review 10 309-347.

Gong, P., Pu, R., Biging, G.S. and Larrieu, M. R., 2003. Estimation of Forest Leaf Area Index Using Vegetation Indices Derived from Hyperion Hyperspectral Data. IEEE Transactions on Geoscience and Remote Sensing 41 (6) 1355-1362.

Jordan, C.F., 1969. Derivation of leaf area index from quality of light on the forest floor. Ecology 50 663-666.

Lamb, M.C., Masters, M.H., Rowland, D., Sorensen, R.B., Zhu, H., Blankenship, P.D., *et al.*, 2004. Impact of sprinkler irrigation amount and rotation on peanut yield. Peanut Science 31 108-113.

Li, F., Miao, Y., Feng, G., Yuan, F., Yue, S., Gao, X., *et al.*, 2014. Improving estimation of summer maize nitrogen status with red edge-based spectral vegetation indices. Field Crops Research. 157 111-123.

Robson, A.J., Wright, G. and Phinn S., 2006. Using field spectroscopy and Quickbird imagery for the assessment of peanut crop maturity and aflatoxin, Journal of Spatial Science, 51 (2) 151-162.

Rouse, J.W., Haas, J.R.H. Schell, J.A. and Deering, D.W., 1974. In Proc. ERTS-1 Symp., Greenbelt, MD. 10-15 Dec. 1974. NASA, Washington, D.C., 1974. Monitoring vegetation systems in the Great Plains with ERTS 309-317.

Rowland, D.L., R.B. Sorensen, C.L. Butts, and W.H. Faircloth, 2006. Determination of maturity and degree day indices and their success in predicting peanut maturity. Peanut Science 33 125-136.

Rowland, D.L., Sorensen, R.B., Butts, C.L. Faircloth, W.H., and Sullivan, D.G., 2008. Canopy characteristics and their ability to predict peanut maturity. Peanut Science 35 43-54.

Sanders, J.C., Reynolds, D.B., Wilson, D.G., and Bruce, L.M., 2002. Utilizing remote sensing to determine crop maturity. Proceedings – Southern Weed Science Society 55 147-147.

Smith, B. W., 1950. *Arachis hypogaea*: Aerial flower and subterranean fruit. American Journal of Botany 37 (10) 802-815, doi:10.2307/2437758.

Vellidis, G. and Beasley, J., 2012. Using vegetation indices to determine peanut maturity. Report to the Georgia Agricultural Commodity Commission for Peanuts. GPC 2013 Research Reports – http://www.gapeanuts.com, (last accessed November 2018).

Williams, E. J., 2003. A simple, quick, inexpensive peanut blaster. University of Georgia. https://t.uga.edu/4B4, (last accessed December 2018).

Williams, E.J. and Drexler, J.S., 1981. A non-destructive method for determining peanut pod maturity. Peanut Science 8 134-141.

Wu, C., Niu, Z., Tang, Q. Huang, W., 2008. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation, Agricultural Forest Meteorology 148 (8-9) 1230-1241.

Yu, N., Li, L. N., Schmitz, L.F. Tian, Greenberg, J.A., Diers, B.W., 2016. Development of methods to improve soybean yield estimation and predict plant maturity with an unmanned aerial vehicle based platform. Remote Sensing of Environment 187 91-101.

Section 10
Precision pasture management

Delineation of management zones in an agrosylvopastoral ecosystem based on the Rasch model

F.J. Moral[1*], F.J. Rebollo[2] and J.M. Serrano[3]
[1]*Departamento de Expresión Gráfica, Escuela de Ingenierías Industriales, Universidad de Extremadura. Avenida de Elvas s/n, 06006 Badajoz, Spain; fjmoral@unex.es*
[2]*Departamento de Expresión Gráfica, Escuela de Ingenierías Agrarias, Universidad de Extremadura. Avenida de Adolfo Suárez s/n, 06007 Badajoz, Spain*
[3]*Departamento de Engenharia Rural, Instituto de Ciências Agrárias e Ambientais Mediterrânicas (ICAAM), Escola de Ciências e Tecnologia, Universidade de Évora, Apartado 94, 7002-554 Évora, Portugal*

Abstract

Estimating the pasture soil fertility and, later, analysing its spatial pattern, the probabilistic Rasch model is a new application in pasture fields. A case study was carried out in a 6 ha Mediterranean evergreen oak woodland. Ten soil properties were measured at 76 locations in the experimental field and, after their integration in the model, all sampling locations were classified according to the pasture soil fertility. Moreover, the influence on the soil fertility of each soil property was highlighted, with soil moisture, clay and sand content, and total nitrogen being the most influential properties. Later, ordinary kriging was applied to estimate pasture soil fertility across the field and homogeneous zones were delineated from the kriged map. Normalized difference vegetation index (NDVI) data at each sampling location were used to verify the suitability of the Rasch measures to differentiate within-field zones. The analysis of zonal differences in pasture systems can lead to an optimal application of inputs and a more effective management, with many associated benefits.

Keywords: probability maps, objective model, soil fertility

Introduction

The 'Dehesa' and 'Montado' ecosystems have great economic and social importance on the Iberian Peninsula because of both the large amount of land involved and its importance in maintaining rural populations. The tree component is oak trees, mainly cork (*Quercus suber*) and holm (*Quercus ilex*), with an understory vegetation, shrubs and pastures, supporting livestock production and other crops and uses.

Soil fertility is the main factor that determines pasture yield and quality. Consequently, the appropriate management of this ecosystem, related to soil fertilisation or grazing intensity, requires the identification of homogeneous areas, with similar permanent characteristics (Schellberg *et al.*, 2008), termed site-specific management zones (SSMZ). The use of geographical information systems (GIS), global navigation satellite systems (GNSS), and remote sensing can be used for pasture management with the aim of improving decision-making. Delineation of SSMZ requires collecting and analysing data for different locations in a field. Better decision-making may provide environmental, economic, energetic and social benefits.

Measurement of the spatial variability of pasture soil and vegetation is the basis for grazing management (Serrano *et al.*, 2010). Soil properties have to be considered to delineate homogeneous zones. Many studies have been devoted to different methods to delineate within-field management zones (e.g. Gavioli *et al.*, 2016; Fortes *et al.*, 2015). However, as Trotter *et al.* (2014) indicated, there is no similar research in pasture systems. To properly define SSMZ, the same methodology used in agricultural fields can be used in pasture soils, combining different layers of information, which

DOI 10.3920/978-90-8686-888-9_76, © Wageningen Academic Publishers 2019

integrate the most significant variables, and using different algorithms, such as a cluster procedure using the fuzzy c-means methods, principal component analysis, alone or in combination with the cluster method, or the use of the coefficient of variation of each data layer (e.g. Xin-Zhong *et al.*, 2009; Ortega and Santibañez, 2007).
Recently, the use of an objective and probabilistic model, the Rasch model (Rasch, 1980), was proposed as a new technique to integrate the data from different soil properties in agricultural fields (Moral *et al.*, 2011; Rebollo *et al.*, 2017). In this study, the same technique was proposed to consolidate some soil properties in pasture soils. Consequently, the objectives of this study were to obtain the soil data that influence pasture soil fertility by using the Rasch model, and generate SSMZ utilising the Rasch measures and geostatistical techniques.

Materials and methods

The case study field, covering 6 ha, at Revilheira farm (38° 27' 51.6' N, 7° 25'46.2' W), was located about 40 km Northeast of Evora, in Southern Portugal. An overview of the boundary of the site is given in Figure 1.
The topography across the field is dominated by gentle hills. The predominant soil is classified as a Leptic Luvisol (FAO, 2006), and the soil profile is characterised by a pedogenetic differentiation of clay content with a lower content in the topsoil (0.15 to 0.50 m) and a higher content in the subsoil.

Data collection

The soil spatial variability of the experimental field was characterised with 76 samples (Figure 1) taken from a 28×28 m grid spacing, in a depth range of 0-0.30 m, and georeferenced with a Trimble 4700 GPS-RTK receiver (Trimble Navigation Limited, Sunnyvale, CA, USA). Texture, pH, organic matter (OM) content and macronutrients (total nitrogen –TN-, phosphorus, and potassium) were measured at each sampling location using conventional methods (Serrano *et al.*, 2010).
The apparent soil electrical conductivity (ECa) survey of the experimental field was conducted in February 2007, using a Veris 2000 XA contact-type sensor (Veris Technologies, Salina, KS, USA). This sensor was equipped with a GNSS antenna, generating one set of georeferenced readings, weighted from 0 to 0.30 m depths. Each 28 m by 28 m square was covered twice in opposite directions, with a spacing of about 14 m between the two consecutive passages. The average ECa value of each



Figure 1. Experimental field and soil sampling location.

square was obtained using the arithmetic average of 28-30 registered values. On the same day, soil samples for soil moisture content (SMC) determination were collected in a depth range of 0-0.30 m. Additionally, NDVI was calculated using two OptRx active crop sensors (Ag Leader, Ames, IA, USA), mounted on a purpose built frame attached to an all-terrain vehicle, at a height of 0.75 m above the ground. Each 28 m by 28 m square was covered twice in opposite directions, with a spacing of about 14 m. The measurements of both crop sensors were georeferenced by means of a Trimble GNSS GeoExplorer 6000 series, model 88951 with sub-meter precision (Trimble GmbH, Raunheim, Germany). This sensor simultaneously measures the surface reflectance in three infrared bands: RED- 670 nm with a range of 20 nm; Red Edge- 728 nm with a range of 16 nm; and NIR- 775 nm with basically everything under 750 nm being filtered out. In consequence, NDVI can be calculated. This survey was performed in March 2007, before cutting the pasture.

Data treatment

One of the simplest and powerful item response theory model for measurement is the probabilistic Rasch model. It constitutes the most viable approach for practical testing, as it can be applied in the context in which individual, soil samples in this case, interacts with items, soil properties. The Rasch model has only one measurement parameter, in a single dimension and scale to measure the classification of both the subjects, sampling locations, and the considered items, soil properties. All data are synthesised by means of a common adimensional referent, defining the latent variable. Thus, in this case study, measures related to some soil properties (10) taken at different sampling locations (76) were consolidated into a global variable which highlighted the interpretation of pasture soil fertility (latent variable), considering that the production potential of pasture biomass is influenced by soil fertility. The unrelated units of the soil properties were categorised with uniform rating scales (Andrich, 1988) and transformed to common logit units with Rasch measurements, eliminating the prior metric constraints. Soil fertility was considered as a straight line, ranging from less soil fertility to more, along which items, soil properties and sample locations were located.

The paradigm underpinning the Rasch model is distinct from the perspective underpinning statistical modelling. When the Rasch model is employed, the objective is to obtain data which fit the model (Andrich, 2004). The Rasch model is a general one and can be applied wherever data are obtained with the intention of measuring a quantitative attribute or trait. The Rasch model gives place to results which can be easily understood and there is no the necessity of defining any kind of weights; furthermore, original variables can be related or unrelated, that is, there are no initial constraints about the variables (Tristán, 2002).

Chi-square fit statistics, known as Infit and Outfit Mean Square (Infit and Outfit MNSQ), ratios of observed residual variance to expected residual variance, were computed to determine the contribution of each item to the measurement of pasture soil fertility. Their expectations are 1. According to Bond and Fox (2007), items that fall between the Infit and Outfit MNSQ limits of 0.6 and 1.5 are accepted.

The formulation of the Rasch model was performed by the Winsteps v. 4.0 computer program (Linacre, 2009), and more information about the mathematical formulation can be obtained, for instance, in Tristán (2002).

The first stage of the formulation of the Rasch model is a transformation of the data to a common scale. In this case, soil properties measures were categorically coded on a scale between 1 and 5 for each property at each location, as in previous studies (Moral *et al.*, 2012, Moral and Rebollo, 2017). Some parameters and the probability curves confirmed the correct use of this assignment scale (not shown).

Values of the pasture soil fertility with the formulation of the Rasch model were obtained. Later, the pasture soil fertility had to be estimated at any point where direct measurements were not carried out. In this study, the ordinary kriging was utilised as interpolation algorithm. The extension Geostatistical Analyst of ArcGIS (version 10.3, ESRI Inc, Redlands, California, USA) was used

to conduct geostatistical analysis and maps of kriged estimates were produced with the ArcMap module of the ArcGIS.

Results and discussion

After processing all data, the raw score, that is, the sum of points of all rating scale categories for each soil property, and the measure value estimated with the Winsteps program for the sampling location, according to the raw score, were obtained (Table 1). When the Rasch measure, that is, the measure of soil fertility, was compared with the NDVI value at each sampling location, as shown in Table 1, it was apparent that in general higher NDVI values were measured at those places in which higher Rasch measures were obtained.

Both mean Infit and Outfit MNSQ for soil samples and properties were 1.02 and 0.99, respectively, very close to the expected value, 1. Furthermore, the reliability statistics, which estimate the degree to which measures are free from error and, consequently, yield consistent results, were 0.71 and 0.92 for samples and items respectively. These values are higher than 0.70 which is the limit considered acceptable (Sekaran and Bougie, 2016). In consequence, data properly fitted the model.

Figure 2 shows in the upper half of the diagram the relative distribution of the soil samples, according to the associated soil fertility achieved by means of all soil properties taken into account. Similarly, the soil properties are provided in the lower half of the diagram, classified according to the pasture soil fertility measure of the soil samples. Silt soil content is the property more to the right in the continuum. In consequence, among all soil properties, its influence on the soil fertility is the least important. In contrast, to the left, soil moisture, sand, and clay content, as well as N content, are situated. Therefore, they were the most influential properties on soil. As sand, clay, and N content are practically at the same position in the straight line, it could be possible to consider dropping one or two of them as redundant. The relationships between these soil properties, particularly between sand and clay, can explain this fact (e.g. Moral *et al.*, 2011). Similar results were found in agricultural soils where clay and sand contents are fundamental properties related to soil fertility (e.g. Moral and Rebollo, 2017; Rebollo *et al.*, 2017).

With respect to the soil samples, they were distributed across the straight line (Figure 2). Those samples located to the left in the continuum had very low pasture soil fertility, but those located to the right were prone to inducing good soil fertility. Thus, soil samples could be sorted according to their pasture soil fertility, estimated as the Rasch measure. Since all samples were georeferenced, those

Table 1. Results obtained after applying the Rasch model: sum of points of the common scale for all individual soil properties (raw score) and pasture soil fertility (measure). Only some sampling points, those with higher and lower Rasch measure, are shown. NDVI values at these sampling locations are also shown. In total there are 76 samples.

Sampling location number	Raw score of soil properties	Measure of soil fertility	NDVI
60	39	0.92	0.75
62	38	0.82	0.76
55	36	0.62	0.75
67	36	0.62	0.67
…	…	…	…
28	19	-1.12	0.60
19	18	-1.27	0.62
2	17	-1.43	0.60
11	17	-1.43	0.58

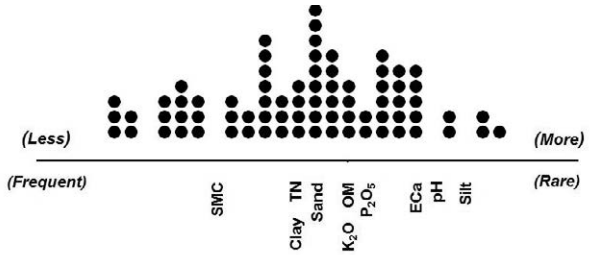


Figure 2. Soil samples and properties on the same scale. The straight line represents the latent variable: pasture soil fertility. Distribution of soil samples (points) is above the line: to the right those more; to the left those less fertile. Soil properties are below the line: to the right less common (rare) properties, with lower influence on pasture soil fertility; to the left more common (frequent) properties, with higher influence on pasture soil fertility.

locations in which higher pasture yield was expected could be visualised. The maximum score was 39 (Table 1) and only 23 samples reached a score higher than 30 points, that is, around 30% of the soil samples, denoting the existence of zones in the field that had higher suitability for pasture fertility. With the aim of estimating pasture soil fertility at any location, ordinary kriging was applied. Initial data were all values at sample locations (Table 1), which were computed with the application of the Rasch model. After assuming isotropic conditions because, according to Goovaerts (1997), the anisotropy ratio was low (1.78), an omnidirectional experimental variogram was computed using the Rasch measure as the random variable. Later, a spherical theoretical variogram was fitted to the experimental variogram (range=298.75 m; sill=0.51; nugget effect=0.04) and this information was considered in the ordinary kriging algorithm. Before producing the kriged map, the validation process provided an idea about the reliability of the final spatial distribution of the Rasch measure. Thus, part of the initial data, a test dataset (consisting of 25% of the whole dataset, 19 sampling locations) was removed and the rest of the data (57 sampling locations) was used as a training dataset to develop the experimental variogram considered in the kriging algorithm. Later, some prediction error statistics were computed and used as diagnostics that indicate whether the kriged map was reasonable. It was found that the root mean square error (RMSE) was 0.40, the mean standard error (ME) was 0.33, and the mean standardised error (MSE) was 0.21. Since all these statistics were very low and the RMSE was close to the ME, the kriged map was considered as appropriate. Moreover, predictions throughout the field were effective since the standard deviation of the measured values of the test samples was 0.62, higher than the RMSE. Finally, three zones were delimited from the kriged map considering the mean value of the Rasch measure, -0.26, plus and minus one standard deviation as the thresholds between intervals (Figure 3).

When the Rasch measures and the NDVI values at each sampling location were plotted on a graph (Figure 4), it was clearer that higher Rasch measures correspond with higher NDVI. Consequently, pasture yield can be more important in the dark zone since soil fertility is higher, and the field can be properly managed taking into account the different conditions of the three zones.

Conclusions

The use of an objective and probabilistic model (the Rasch model) to estimate a measure of pasture soil fertility, integrating different soil variables (texture, SMC, ECa, pH, OM, TN, phosphorus and potassium), has been successful. Data reasonably fitted the model and the considered soil properties had an important influence on the latent variable, pasture soil fertility. Sand and clay content, TN and

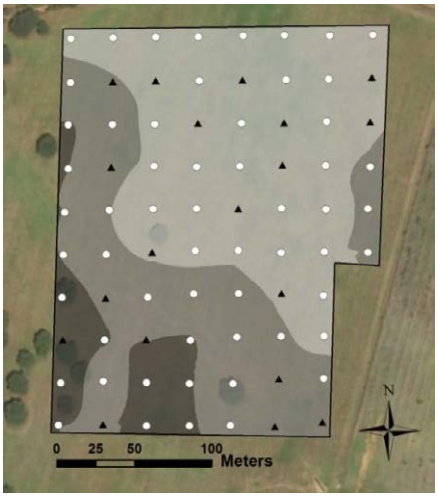


Figure 3. Management zones map based on the kriged map of pasture soil fertility. The three zones were delineated considering the mean Rasch measure plus and minus one standard deviation (-0.26±0.57) as the threshold values. Dark zone is the more fertile and light zone is the less fertile. White circles are the sample points used as training dataset and black triangles are the sample locations used as test dataset.

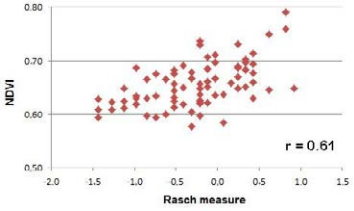


Figure 4. Rasch measure vs NDVI at each sampling location. The correlation coefficient (r) is also shown.

SMC were found the most determining soil variables on the pasture soil fertility in this particular field. A classification of the sampling locations according to their soil fertility level was also obtained. The use of ordinary kriging to interpolate at unsampled locations provided an accurate representation of the spatial distribution of the pasture soil fertility. This was used to delineate homogeneous zones, which is very important in pasture fields like the one considered in the present case study, where pasture soil fertility level is low but zonal differences are apparent.

Acknowledgements

This research was co-financed by the Junta de Extremadura and the European Regional Development Fund (ERDF) through the project GR18086 (Research Group TIC008).
The authors are also very grateful to two anonymous reviewers for providing constructive comments which contributed to improve the final version of this paper.

References

Andrich, D., 1988. Rasch model for measurement. Newbury Park, CA, USA: Sage Publications.

Andrich, D., 2004. Controversy and the Rasch model: a characteristic of incompatible paradigms? Medical Care 42 1-16.

Bond, T. G., Fox, C. M., 2007. Applying the Rasch Model: Fundamental Measurement in the Human Sciences, 2nd ed., Mahwah, NJ, USA: Lawrence Erlbaum Associates Inc.

FAO, 2006. World reference base for soil resources. World Soil Resources Reports N. 103, Rome, Italy: Food and Agriculture Organization of the United Nations.

Fortes, R., Millán, S., Prieto, M. H. and Campillo, C., 2015. A methodology based on apparent electrical conductivity and guided soil samples to improve irrigation zoning. Precision Agriculture 16 441-454.

Gavioli, A., Souza, E. G., Bazzi, C. L., Guedes, L. P. C. and Schenatto, K., 2016. Optimization of management zone delineation by using spatial principal components. Computers and Electronics in Agriculture 127 302-310.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation, New York, USA: Oxford University Press.

Linacre, J. M., 2009. Winsteps (Computer program and manual), Chicago, USA: MESA Press.

Moral, F. J., Rebollo, F. J., 2017. Characterization of soil fertility using the Rasch model. Journal of Soil Science and Plant Nutrition 17 (2), 486-498.

Moral, F. J., Rebollo, F. J. and Terrón, J. M., 2012. Analysis of soil fertility and its anomalies using an objective model. Journal of Plant Nutrition and Soil Science 175 912-919.

Moral, F. J., Terrón, J. M. and Rebollo, F. J., 2011. Site-specific management zones based on the Rasch model and geostatistical techniques. Computers and Electronics in Agriculture 75 223-230.

Ortega, R.A., Santibáñez, O.A., 2007. Determination of management zones in corn (Zea mays L.) based on soil fertility. Computers and Electronics in Agriculture 58 49-59.

Rasch, G., 1980. Probabilistic models for some intelligence and attainment tests. Revised and expanded ed., Chicago, USA: University of Chicago Press.

Rebollo, F. J., Moral, F. J., Campillo, C., Marques da Silva, J. R., Serrano, J. M. and Pérez-Rodríguez, J. M., 2017. Delineation of management zones based on the Rasch model in an olive orchard. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences 8 (2), 610-614.

Schellberg, J., Hill, M. J., Gerhards, R., Rothmund, M. and Braun, M., 2008. Precision agriculture on grassland: Applications, perspectives and constraints. European Journal of Agronomy 29 (2-3) 59-71.

Sekaran, U., Bougie, R., 2016. Research methods for business: A skill building approach, Chichester, UK: John Wiley & Sons.

Serrano, J., Peça, J., Marques da Silva, J. and Shahidian, S., 2010. Mapping soil and pasture variability with an electromagnetic induction sensor. Computers and Electronics in Agriculture 73 7-16.

Tristán, A., 2002. Análisis de Rasch para todos (Rasch analysis for all). Ed. Ceneval, México.

Trotter, M., Guppy, C., Haling, R., Trotter, T., Edwards, C. and Lamb, D., 2014. Spatial variability in pH and key soil nutrients: Is this an opportunity to increase fertiliser and lime-use efficiency in grazing systems? Crop and Pasture Science 65 817-827.

Xin-Zhong, W., Guo-Shun, L., Hong-Chao, H., Zhen-Hai, W., Qing-Hua, L., Xu-Feng L., *et al.*, 2009. Determination of management zones for a tobacco field based on soil fertility. Computers and Electronics in Agriculture 65 168-175.

Accurate segmentation of grass-sward structures in RGB images using machine learning algorithms

C. Rueda-Ayala[1], V. Rueda-Ayala[2*], M. Höglind[2] and D. Andujar[1]
[1]*Centre for Automation and Robotics, CSIC, Arganda (Madrid), Spain*
[2]*Norwegian Institute of Bioeconomy Research, NIBIO Særheim, Postvegen 213, 4353 Klepp Stasjon, Norway; patovicnsf@gmail.com*

Abstract

This work aimed to develop an automatic color image processing tool for segmenting complex ley sward grass structures, i.e. green-leaves, straw and bare soil leaf area index (LAI), using artificial intelligence techniques. Good classification results were achieved with Gaussian mixture models (GMM) + nearest centroid classifier (NCC) trained with data from color scale transformations from all RGB channels to the color space achromatic luminance and blue and red chromatic channels (*Y'CbCr*), to the color model cyan, magenta, yellow and key (*CMYK*). Also vegetation indices were included. A good correspondence between computer-based estimations and the visual- and destructive-estimated field data, with R^2 above 0.85 and even 0.95. This work provides a tool for rapid and accurate determination of parameters which otherwise would require a large amount of working hours.

Keywords: pasture, image analysis, color scales, vegetation indices, supervised classifiers, clustering

Introduction

Current technologies, such as remote and proximal sensing have been practical tools for grassland research, to assess chlorophyll levels through reflectance measurements (Clevers and Gitelson, 2013), and to elaborate grassland management strategies, especially regarding fertilization (Adjorlolo *et al.*, 2015). Imaging and sensing tools provide sound field assessments, sometimes more authentic than ground truth itself, because sensors and platforms can measure huge areas in short time and at low cost. Image analysis has enormously improved in the last decade for plant-shape segmentation and discrimination. Nevertheless, precise determination of simple features, such as coverage of green leaves or bare soil might be extremely challenging (Sadeghi-Tehran *et al.*, 2017), as in the case of a homogeneous grassland layer. RGB and multispectral images also allow calculating vegetation indices (Hamuda *et al.*, 2016), which help to identify environmental and other stresses, such as diseases, weeds or nutritional deficiencies. Moreover, RGB image analysis improve when gray-scale and other color-scale transformations, such as hue, saturation, value (HSV) or color space of achromatic luminance and blue- and red chromatic channels (*Y'CbCr*), are used to prioritize a single color-based vegetation characteristic (Guijarro *et al.*, 2011).

Implementation of PA technologies is challenging due to engineering, ecological and especially economic aspects (Schellberg and Verbruggen, 2014). Furthermore, field sampling is necessary for contrasting sensor data with the reality. Specifically for grasses, proper determination of plant coverage or leaf area index has mostly been done by manual and destructive methods. Although cutting vegetation samples is precise, it is extremely costly, time consuming and can be subject to human error (Doust, 2007). Consequently, non-destructive methods for characterization of grass-mixture plants are highly valuable for grassland research and for a sustainable production and successful management.

This study aimed to develop an objective, fast and low-cost sampling tool for in-field detection of grass plant structures, to facilitate grassland research and in the long-term to provide a tool for

DOI 10.3920/978-90-8686-888-9_77, © Wageningen Academic Publishers 2019

decision-making about production management. RGB-image analysis and various machine-learning processes were applied for the discrimination and segmentation of living and dead vegetative structures from bare soil. This work is a part of the project 'Improved cost-efficiency in Norwegian forage production (ForEff, WP2: Cost-efficient grassland renovation methods to sustain high forage yields over years', NIBIO, in cooperation with the Centre for Automation and Robotics, CAR, CSIC.

Materials and methods

The engineering methodology followed to build up the imaging analysis tool was that of mechatronics in 'V' form (Gausemeier and Moehringer, 2003), with the corresponding components, system design, specific domain design and modeling and model analysis, and the system integration. The system had to accomplish an automatic segmentation function of vegetative structures, and this was the main problem to be solved. In order to do that, some requirements were identified as input for the system: image acquisition, general image pre-processing, and acquisition of useful segmentation parameters. Finally, the image processing enabled the segmentation of 3 structure classes, vegetative alive (green) and dead (yellow-brownish) textures and bare soil (brown), including calculation of the corresponding coverage.

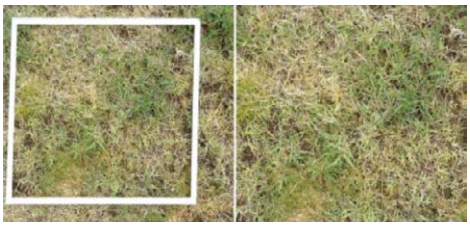
Equipment, experimental site and image acquisition

A commercial camera SONY DSC-HX60V was used to capture RGB images, trying to aim for cloudy conditions to avoid distortions by shadows. The camera resolution was 4,896×3,672 pixels. The camera was located parallel to the ground plane, at the Zenith, at 1.60 m of height. With this camera setting, a net frame of 1 m^2 was cropped for analysis, which corresponded a pixel size of 0.0036 mm). Four sets of images were taken from two fields of timothy and ryegrass grass-mixtures, at different growth stages during the growing period between May and July 2018, and using four repetition blocks. The fields were cultivated in the NIBIO-Særheim research station, located at Klepp Stasjon, Norway, lat. 58.76 N, long. 5.65 E. The soil was of a silty sand type. Each field had an area of 400 m^2, and a variable range of soil coverage between 0 and 100% was artificially created by randomly spraying a non-specific contact herbicide, using various patches of 0.018 m^2. Between 2 and 3 images were taken for each soil coverage range of 10%. A white frame of 1 m^2 was used and all calculations were based on this known surface. Approximately, 600 images were acquired and processed.

Image pre-processing and segmentation

The components to develop algorithms for image analysis consisted of specification of the region of interest (ROI), feature extraction (i.e. segmentation parameters) and implementation of artificial intelligence techniques for image processing. The ROI corresponded to the 1 m^2 area inside the white frame, in order to correlate image data with the ground truth assessments. The ROI was obtained by segmentation of the white frame by transforming the images to HSV color scale, binarization, erosion and dilation (opening), followed by dilation and erosion (closing), using a kernel defined by the matrix of rows and columns [3, 5] and [5, 5] pixels, respectively. For this purpose, the Shi-Tomashi corner detection methods were used to draw the polygon and facilitate the ROI segmentation (Shi and Tomasi, 1993). When the frame was covered by grass leaves, a convex hull algorithm followed by a Canny border detection method (Canny, 1986) were used. The ROI was ortho-rectified (Figure 1), using a vertical plane through calculation of a homography matrix, supported with the RANSAC algorithm (Vincent and Laganiére, 2001). This resulting area was cropped, eliminating the white frame pixels, as well.

For feature extraction, three approaches were tested: (i) combination of the color scales RGB and color spaces *Y'CbCr* (being *Y'* the achromatic luminance, *Cb* the blue- and *Cr* the red chromatic channels), and the color model *cyan*, *magenta*, *yellow* and *key* (*CMYK*), using Equation 1. Each



A B

Figure 1. Cropping ROI from the original image (A) and ortho-rectification (B).

value of *C*, *M*, *Y* and *K*, were calculated on-the-go for each pixel and used by the computer, while analysing each image. Therefore, these values were continuously changing and not saved. *Y'CbCr* and *CMYK* spaces and model contribute to avoid variations due to illumination changes. (ii) RGB, *Y'CbCr*, *CMYK* combined with calculation of the modified excess-green index (*MExG*), using Equation 2, where *MExG* was computed with Equation 3 (Burgos-Artizzu *et al.*, 2011). (iii) RGB to *Y'CbCr* transformation, combined with *MExG*, Equation 4. All vector data generated did not require normalization, because the values remained within the [0, 255] scale. For uniformity of analysis, normalization of *MExG* was also done to fit the 0-255 scale. Dimensionality reduction was applied by principal component analysis (PCA), to select only the influential vector data. PCA was applied to all features, even though the combinations were linear, because *Y'CbCr* enhances the color representation of the RGB scale compared to RGB alone, and *MExG* recognizes better green living leaves. Two principal components were determined to identify the 3 classes, living plants, dead plants and bare soil.

$$\text{Pixel}_{(ij)} = (R,G,B,Y',Cb,Cr,C,M,Y) \quad (1)$$

$$\text{Pixel}_{(ij)} = (R,G,B,Y',Cb,Cr,C,M,Y,MExG) \quad (2)$$

$$MExG = 1.262G - 0.884R - 0.311B \quad (3)$$

$$\text{Pixel}_{(ij)} = (R,G,B,Y',Cb,Cr,MExG) \quad (4)$$

Image processing and segmentation, was done with unsupervised and supervised machine learning algorithms, including a combination of both. Unsupervised algorithms were applied first, because a previously classified image dataset was not available. For this purpose, the Gaussian mixture model (GMM) based on the expectation-maximization algorithm with random samples was used (Reynolds, 2015). The resulting groups from GMM were arranged into a set of 30 artificial images, created pixel by pixel through computational combination of 912×912 pixels for each artificial image. This artificial dataset served as training data and binary masks for labelling the living plant, dead plant and bare soil classes. Each artificial image consisted of labelled segments resulting from GMM. Labelling was done with the software GIMP, via visual supervision of the known classes, and eliminating erroneously segmented pixels. Then, an algorithm was programmed to read the values on each artificial image and assign a value of 0 to bare soil, 1 to living structures and 2 to death material (Figure 2).

For image processing and classification on the acquired image datasets, the algorithms used were GMM, K-nearest neighbour (KNN) (Zhang and Zhou, 2007), and NCC (nearest centroid classifier) (Tibshirani *et al.*, 2002), the latter always as a complementary algorithm to GMM. GMM, KNN and GMM+NCC had a classification efficiency higher than 90%, when applied on the artificial dataset. With this approach, four combination of algorithms were used for image processing: (1)

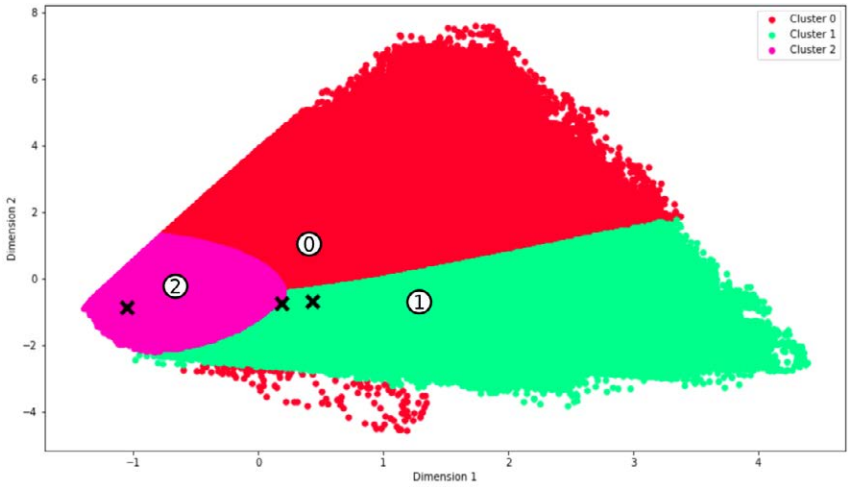


Figure 2. Grouping of three classes with the unsupervised GMM classifier. Clustering was done on PCA-reduced data. Centroids are marked by the numbers, and transformed sample data are marked with the black 'x'.

GMM+NCC trained with the set of features from 'i', (2) GMM+NCC trained with features from 'ii', (3) GMM+NCC trained with features from 'iii', and (4) KNN trained on features from 'ii'. All computational operations were programmed using the Python 2.7 distribution and a computer with an Intel Core Processor i3-8100 (4 Core, 6 MB Cache, 3.6 GHz w/ HD Graphics 630, and 64 GB RAM.

Results

The results from applying the GMM as unsupervised classifier (Figure 2), allowed the artificial creation of images with labels for the three classes, living plant, dead plant and bare soil. This class separation was acceptable and especially accurate for the living plant class (Figure 3), while keeping an undistorted plant shape when segmenting the classes. The use of KNN and GMM+NCC classifiers on the artificial image dataset (supervised method) produced a good classification. However, the shape of the living plant class was affected at the segmentation stage. This distortion occurred because classification was applied at pixel level, which affected the real aspect of living plant structures, especially when shadows were present in an image segment. Therefore, both methods were mixed to improve the classification and shape of the segments. The resulting algorithms were 'A1: RGB + *Y'CbCr* + *CMYK* data processed with GMM + NCC', 'A2: RGB + *Y'CbCr* + *CMYK* + *MexG* data processed with GMM + NCC', 'A3: RGB + *Y'CbCr* + *MexG* data processed with GMM + NCC', and 'A4: RGB + *Y'CbCr* + *MExG* data processed with KNN. Figure 4 shows the results from image processing and segmentation of the best algorithm, namely A2, on a Timothy sward at intermediate growing stage. Color spaces *Y'CbCr* and model *CMYK* were highly useful to diminish the classification error, in which some shadows could be wrongly classified as bare soil. This misclassification occurred minimally in the present study (Figure 4) when pixel values had similar values for both, weed shadows and bare soil, because they were located close in the image space. The output was similar for all three other algorithms.

Analysis of RGB images, including the color spaces *Y'CbCr* and model *CMYK* requires a huge amount of computing power, if the analysis is done by pixel. This requirement would be due to the number of features calculated per pixel in each image, i.e. features resulting from combining R, G,

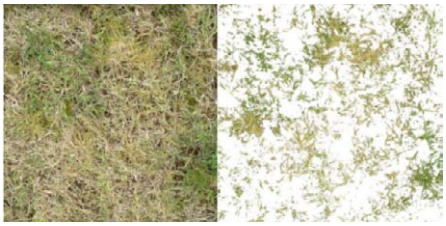


Figure 3. Results of application of GMM after supervised labeling on a sample Timothy image. (A) original image; (B) segmentation of living plants class.

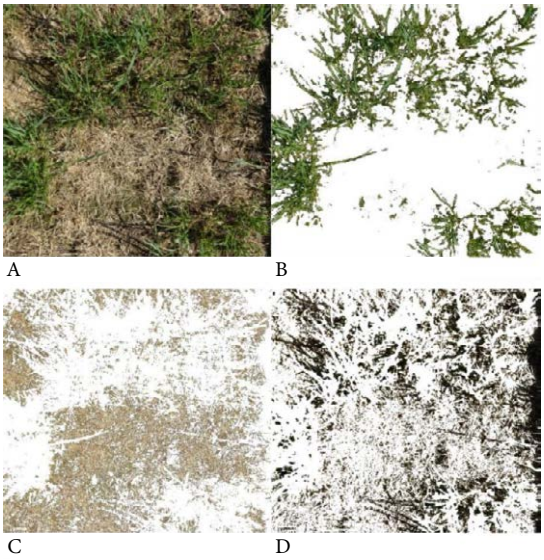


Figure 4. Results of application of A2 on a sample Timothy image. (A) original image; (B) living plant; (C) dead plants; (D) bare soil.

B, *Y'*, *Cb*, *Cr, C, M, Y,* and *K*. For instance, if 9 features were selected per pixel × 912×912 would yield approximately 7,500,000 features to be analyzed in one image. Reducing the dimensionality to only 2 features per pixel × 912×912, resulted in 1,500,000 features per image that went through analysys and segmentation, thus significantly reducing the computing power required. When classification is carried out at pixel level, the supervised algorithm KNN performs well. However, there is a drawback produced when neighbouring pixels to the one that is being analysed do not provide enough information to assign an accurate class label. For example, if neighbouring pixels did not clearly show that the class was living structures, KNN excludes that class, even when the pixel should have been classified as living plant. Contrarily, using a mixture such as GMM+NCC, the classification was done firstly according to the spatial location of pixels, i.e. neighbouring pixels with information of one class were grouped, and then, the classification was done based on the probability that a group of pixels belong to the same class.

Comparison of the assessments made with destructive methods and the developed image processing tool showed an acceptable correlation, and indicated the high potential of A2 and A4 for achieving a good classification, with an R^2 of 0.71 and 0.66, respectively. Computer-based estimations and the

visual- and destructive-estimated field data of LAI and coverage on the three tested growth stages were highly accurate, with R^2 between 0.85 and 0.95. A1 and A3 showed a less efficient performance, with R^2 values of 0.54 and 0.51, respectively.

Conclusions

This study has shown that image analysis for determination of green coverage and bare soil percentage of grassland fields is complex, but there are promising algorithms for achieving this purpose, if they were smartly combined. Inclusion of color spaces *Y'CbCr* and *CMYK* improves the precision in the discrimination of living and plant structures from bare soil, enhancing the limitations of vegetation indices, which only use the RGB scales. Reducing the feature selection dimensionality helps to reduce the computing requirements. In this study, processing each high resolution image took less than 18 seconds when PCA-reduction was applied. A mixture of unsupervised machine learning algorithms with supervised classification improves enormously the discrimination accuracy of the sampling tool, while keeping important plant features, such as shapes. The algorithm combination of GMM+NCC applied on data obtained via feature extraction from color scales RGB+*Y'CbCr*+*CMYK* plus the vegetation index *MExG* achieved a 95% classification accuracy. This algorithm had the advantage of maintaining the vegetation shapes non-distorted, unlike with the supervised algorithm KNN, which also had a good classification accuracy. Other issues, such as discrimination between cultivated grass species and weeds, was not part of this study, but it requires attention for further improvements of the proposed image sampling tool.

Acknowledgements

This research was supported by the Rogaland Fylkeskommune through the project 'Improved cost-efficiency in Norwegian forage production, FOREFF', NIBIO. We thank all technicians and scientists from the research station NIBIO-Særheim.

References

Adjorlolo, C., Mutanga,O. and Cho, M. A., 2015. Predicting C3 and C4 grass nutrient variability using in situ canopy reflectance and partial least squares regression. International Journal of Remote Sensing 36 (6):1743-1761.

Burgos-Artizzu, X. P., A. Ribeiro, M. Guijarro and G. Pajares., 2011. Real-time image processing for crop/weed discrimination in maize fields. Computers and Electronics in Agriculture 75 (2):337-346.

Canny, J., 1986. A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence, 6:679-698.

Clevers, J. G. and A. A. Gitelson, 2013. Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and-3. International Journal of Applied Earth Observation and Geoinformation 23:344-351.

Doust, A. N., 2007. Grass architecture: genetic and environmental control of branching. Current Opinion in Plant Biology 10 (1):21-25.

Gausemeier, J. and S. Moehringer, 2003. New Guideline VDI 2206 A Flexible Procedure Model for the Design of Mechatronic Systems. In Folkeson, A; Gralen, K; Norell, M; Sellgren, U (Eds.) DS 31: Proceedings of ICED 03, the 14th International Conference on Engineering Design. Scotland: The Design Society. pp 599-699. https://tinyurl.com/y4xz8g9q (last accessed February 2019).

Guijarro, M., G. Pajares, I. Riomoros, P. Herrera, X. Burgos-Artizzu and A. Ribeiro, 2011. Automatic segmentation of relevant textures in agricultural images. Computers and Electronics in Agriculture 75(1):75-83.

Hamuda, E., M. Glavin and E. Jones, 2016. A survey of image processing techniques for plant extraction and segmentation in the field. Computers and Electronics in Agriculture 125:184-199.

Reynolds D., 2015. Gaussian Mixture Models. In: Li S.Z., Jain A.K. (Eds) Encyclopedia of Biometrics. Boston, MA, USA: Springer pp 827-832.

Sadeghi-Tehran, P., N. Virlet, K. Sabermanesh and M. J. Hawkesford, 2017. Multi-feature machine learning model for automatic segmentation of green fractional vegetation cover for high-throughput field phenotyping. Plant methods 13 (1):103.

Schellberg, J. and E. Verbruggen, 2014. Frontiers and perspectives on research strategies in grassland technology. Crop and Pasture Science 65 (6):508-523.

Shi, J. and C. Tomasi, 1993. Good features to track. Technical Report, Cornell University. Ithaca, NY, USA.

Tibshirani, R., T. Hastie, B. Narasimhan and G. Chu, 2002. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proceedings of the National Academy of Sciences 99 (10):6567-6572.

Vincent, E. and R. Laganiére, 2001. Detecting planar homographies in an image pair. In Proceedings of the 2nd International Symposium on Image and Signal Processing and Analysis ISPA 2001., pp. 182-187. IEEE.

Zhang, M.-L. and Z.-H. Zhou, 2007. ML-KNN: A lazy learning approach to multi-label learning. Pattern recognition 40 (7):2038-2048.

Long-term evaluation of the Grassmaster II probe to estimate productivity of pastures

J. Serrano[1], S. Shahidian[1], F. Moral[2] and J. Marques da Silva[1,3*]
[1]*ICAAM, Departamento de Engenharia Rural, Universidade de Évora, P.O. Box 94, Évora 7002-554, Portugal; jmrs@uevora.pt*
[2]*Departamento de Expresión Gráfica, Escuela de Ingenierías Industriales, Universidad de Extremadura. Avenida de Elvas s/n, 06006 Badajoz, Spain*
[3]*Agroinsider Lda. (spin-off da Universidade de Évora), PITE, R. Circular Norte, NERE, Sala 18, 7005-841 Évora, Portugal*

Abstract

The standard method of assessing pasture biomass is based on cutting the forage and requires great effort and expense to collect enough samples to accurately represent a pasture. This work presents the results of tests carried out to calibrate a capacitance probe (Grassmaster II). A monthly regression analysis was performed between the capacitance (CMR) measured by the probe and the values of pasture green matter and dry matter (respectively, GM and DM, in kg/ha). The results showed that the more favorable period for the use of this probe in dry land pastures in Alentejo Portuguese region coincides with the end of winter and beginning of spring (February-March). It can be concluded that the capacitance probe is an expedient tool that can enable the farm manager to estimate pasture productivity and support the decision-making process in the management of dry land pasture and its respective grazing.

Keywords: capacitance probe, permanent pastures, prediction models, dry matter, differential management

Introduction

Estimates of pasture productivity at various stages of the growth cycle are an important element for planning and decision-making by the farmer. The management of animal grazing, in terms of animal stocking at each parcel, animal rotation through the various fields, calculation of forage needs and animal feed supplement is based on productivity. Measurement of pasture productivity can be carried out through direct or indirect methods. The conventional (direct) method is based on harvesting green matter at specific sampling areas, which becomes a lengthy and expensive process given the large number of samples necessary (Cauduro *et al.*, 2006). Since this methodology is not practical at the farm level, other indirect techniques have emerged, that provide an estimate of productivity and its spatial variability in large areas in a timely manner, fulfilling one of the important pre-requisites of implementing innovative precision agriculture strategies (Serrano *et al.*, 2011; Pullanagari *et al.*, 2013). One of the methods proposed for quantifying and mapping the pasture production variability is based on the measurement of vegetation indices, mainly the NDVI (normalized difference vegetation index; Equation 1) calculated by measuring the reflectance of the radiation emitted by the plants, using satellite images (Edirisinghe *et al.*, 2011; Cicore *et al.*, 2016; Louhaichi *et al.*, 2018).

$$NDVI = NIR - Red / NIR + Red \quad (1)$$

where: NIR- Near infrared radiation; Red- visible red radiation

DOI 10.3920/978-90-8686-888-9_78, © Wageningen Academic Publishers 2019

Although use of satellite imagery is a very promising, low cost and non-destructive technique, it has its own limitations in terms of spatial resolution and the presence of clouds, as well as the spatial and temporal specificity of the associated algorithms (Handcock *et al.*, 2016).
Another alternative for calculating vegetation indices is to use unmanned aerial vehicles (UAV) equipped with the appropriate sensors. In either case, whether using satellites or UAV, an important limitation persists in the case of *montado* agro-silvo-pastoral system (named *dehesa* in Spain): it is not possible to access information about the pasture under a tree canopy, which has led to the use of proximal sensors (Handcock *et al.*, 2016). In the last few years, many such sensors have been developed and marketed, such as 'Crop Circle', 'Yara N-Sensor', 'GreenSeeker' or 'OptRx' (Serrano *et al.*, 2018).
Despite the growing use of optical sensors for monitoring the vegetation cover of the globe (either through remote or proximal detection), NDVI tends to saturate for high values of plant leaf area indices, which is when productivity has reached high values (Cicore *et al.*, 2016; Schaefer and Lamb, 2016).
Another specific proposal for estimating pasture productivity has been introduced by Novel Ways Electronic, a New Zealand company that has developed a capacitance probe (Grassmaster II), which was described by Serrano *et al.* (2011). This probe is based on the pioneering work of Vickery and Nicol (1982) who described the theory and the operating principles of such equipment. Capacitance instruments are generally composed of an electronic circuit which generates a signal of a certain frequency, and then carries out a measurement of the capacitance of the air-herbage mixture. The probe produces an electric field that is modified by the pasture close to it. The modified field is detected as a change in capacitance by the electronic circuit within the probe. This capacitance change is proportional to the water content of the grass. Because the dry matter content is highly correlated with the water content, the probe capacitance change is calibrated to the pasture yield, with linear calibration equations used to correlate readings taken on pasture over a short seasonal time, with measured dry matter production (Serrano *et al.*, 2011). The manufacturer has presented an equation for estimating pastures productivity based on the readings from the probe (Equation 2). Nonetheless, this equation was calculated for New Zealand pastures, which are composed of legumes and grasses with an average dry matter of 14-16%.

$$DM = 0.72 \times CMR - 2{,}200 \quad (2)$$

where: DM is the estimated productivity of the pasture, in kg of dry matter per hectare; CMR is the capacitance measured by the Grassmaster II probe.
The results published by Serrano *et al.* (2011, 2016a, 2016b) indicate that the relation between pasture productivity and capacitance measured by the probe is influenced by factors such as the mixture of plant species, their phenological stage or pasture moisture content. This variability requires the calibration tests to span the diversity of parameters related to permanent dry land pastures of Alentejo area of Portugal. The growth and development of these pastures is conditioned by a series of factors, such as soil fertility, the grazing strategy and, mainly, by the rainfall pattern which introduces an inter-annual variation in terms of both growth and development. Thus, it is difficult to conceive that a generic equation can adequately portray the seasonal variability associated with the Mediterranean climate, which is accentuated by the complexity of the *montado* ecosystem (Pullanagari *et al.*, 2013). This article, based on the results obtained in experiments carried out between 2007 and 2018 at three *montado* fields in Alentejo region of Portugal, covering a total of 1,279 sampling points, has the objective of developing monthly calibration equations for the Grassmaster II probe, in order to help identify the best period of year to use this equipment.

Materials and methods

Characteristics of the experimental sites

The experiments were carried out between 2007 and 2018 at three fields located in the Évora District. The characteristics of these sites are presented in Table 1.

These bio-diverse pastures (composed of legumes, grasses, composites and other species) share the common characteristic of being inserted in the context of a holm oak montado, grazed by sheep in a rotational or permanent system. These are bio-diverse pastures (composed of different botanical species), representative of the regional dry land pastures. The fact that they are located on soils that are poor, relatively acid and deficient in phosphorous leads some farmers to try to improve the productivity through application of gypsum-based amendments and phosphorous fertilizers.

Pasture sampling

Pasture sampling was carried out at each location, with repetitions through the growth cycle, totaling 1,279 samples, spread across three fields. The sampling process consisted of measuring the capacitance with the Grassmaster II probe, followed by placing a metal rim with an area of 0.1 m^2 (0.25×0.40 m in size) over the pasture. The pasture contained in this area was cut with electric shears and preserved in numbered plastic bags. Once in the laboratory, the pasture sample was weighed to establish total biomass produced by unit area (GM in kg/ha), dried in an oven (72 h at 65 °C) and weighed again to establish productivity in terms of dry matter per unit of area (DM in kg/ha), according to standard methodology (Serrano *et al.*, 2018). Given that the probe, due to its operating principle (Serrano *et al.*, 2011), does not provide any readings when the pasture is dry (during summer) or when the pasture is wet (rain, frost or dew), most of the measurements were taken during the Spring (April, May and June).

Statistical analysis of the results

The statistical analysis of the results included a descriptive analysis with calculation of average and standard deviation (SD) of each data set. Analysis of the linear correlations between capacitance and pasture productivity was carried out with the MSTAT-C software, with a significance level of 95% (P<0.05). The Pearson (r) correlation coefficient was used to measure the degree of correlation or the linear dependence between variables and the coefficient of determination to measure the proportion of the total variation of the dependent variable explained by the variation of the independent variable.

Table 1. Characteristics of the three experimental fields used in this study.

Site	Revilheira	Silveira	Mitra
Co-ordinates	38°27.9'N; 7°25.7'W	38°62.2'N; 7°94.8'W	38°32.2'N; 8°01.1'W
Area (ha)	6.0	7.0	27.3
Pasture type	Biodivers (predominance of composites)	Biodivers (predominance of grasses)	Biodivers (predominance of grasses and legumes)
Trees (number/ha)	Holm oak (1)	Holm oak and olive trees (16)	Holm oak (8)
Animal species and grazing system (number/ha)	Sheep in rotational grazing (24)	Sheep in permanent grazing (2)	Sheep in permanent grazing (8)
Sampling years	2007, 2013	2013-2016	2007, 2013, 2015-2018
Total of sample points	199	486	594

The rigor of the resulting regression models was evaluated by the absolute value of the root mean square error (RMSE) and its relative value (CV_{RMSE}).

Results

Variability pattern of the measured parameters

The average values and the standard deviation of the parameters measured monthly in the pastures monitored in this study are presented in Table 2. In order to provide farmers with options for estimating pasture growth at two critical moments in the management of the pasture, data collected during two month periods are clustered in Table 2: (1) February and March, which marks the transition from Winter to Spring, when productivity starts to increase significantly due to increase in the air temperature; and (2) April and May, in mid Spring, with average temperatures above 20 °C and, normally, with a reasonable amount of soil moisture, factors that combine to trigger the maximum peak of pasture production.

Correlations between capacitance and pasture productivity

Table 3 and 4 show, respectively, the monthly regression equations between GM and capacitance and between DM and capacitance. For each regression equation, the correlation coefficient (r) and the root mean error (RMSE), in absolute value and in percentage are presented.

Discussion

A reference should be made to the difficulty in contextualizing and discussing the results obtained in this work, since no other research has been published on the use of Grassmaster II probes in pastures in Portugal, beyond the studies carried out by this team (Serrano *et al.*, 2011, 2016a, 2016b). The spatial variability, detected by the high coefficients of variation of the pasture productivity (around 50-80%; Table 2), shows the need for application of differentiated management techniques. On the other hand, the coefficients of variation of capacitance (20-40%) and of the percentage of pasture dry matter (30-45%) are of the same order of magnitude, which reflects the operating principle of the probe, which is very sensitive to pasture moisture content (Serrano *et al.*, 2016a). Table 2 depicts the typical seasonal variability pattern of dry land pastures in Alentejo region of Southern Portugal (Efe Serrano, 2006): low productivity in the months of Autumn-Winter (in

Table 2. Mean±standard deviation of the parameters monthly measured in the pastures of the three fields of this study.[1]

Month	CMR	GM, kg/ha	DM, kg/ha	DM, %
Nov (n=23)	3,978±697	3,930±2,995	842±671	22.0±8.5
Dec (n=37)	5,132±1,391	3,858±2,038	659±307	17.7±7.8
Feb (n=70)	5,436±1,609	8,492±6,141	1,532±851	21.6±8.7
Mar (n=198)	5,629±1,680	6,717±4,320	1,299±641	23.2±9.6
Apr (n=308)	6,989±2,554	12,718±7,468	2,297±1,201	19.8±8.5
May (n=390)	7,547±2,992	13,712±1,0576	2,821±1,454	25.6±10.6
Jun (n=253)	6,248±2,219	10,509±6,347	3,460±1,884	37.0±11.6
Feb-Mar (n=268)	5,579±1,661	7,295±5,299	1,383±749	22.8±9.6
Apr-May (n=608)	7,301±2,819	13,556±9,626	2,621±1,387	23.1±10.2

[1] CMR = capacitance; DM = dry matter; GM = green matter; n = number of samples.

Table 3. Linear monthly equations proposed to estimate the pasture green matter yield (GM): correlation coefficients (r), root mean square error (RMSE) and coefficient of variation of the root mean square error (CV_{RMSE}).[1]

Month	Linear equation	r	RMSE kg/ha	CV_{RMSE} %
Nov	2.9285 × CMR – 7,549.2	0.686**	2,126	54.1
Dec	0.8522 × CMR – 551.3	0.597*	1,611	41.8
Feb	3.8787 × CMR – 12,033	0.919**	2,404	28.3
Mar	2.3358 × CMR – 6,205.6	0.886**	1,971	29.3
Apr	2.6155 × CMR – 4,904	0.831**	4,223	33.2
May	2.6438 × CMR – 6,078	0.745**	7,045	51.4
June	2.3002 × CMR – 3,469.8	0.744**	4,295	40.9
Feb-Mar	2.6804 × CMR – 7,330.9	0.804**	3,377	46.3
Apr-May	2.6129 × CMR – 5,182.3	0.751**	5,759	42.5

[1] CV_{RMSE} = coefficient of variation of root mean square error; GM = green matter; r = correlation coefficient; RMSE = root mean square error; * = correlation significant at the 0.05 level; ** = correlation significant at the 0.01 level.

Table 4. Linear monthly equations proposed to estimate the pasture dry matter yield (DM): correlation coefficients (r), root mean square error (RMSE) and coefficient of variation of the root mean square error (CV_{RMSE}) of these equations and of the equation proposed by the Grasmaster II manufacturer (MAN).[1]

Month	Linear equation	r	RMSE kg/ha	CV_{RMSE} %	$RMSE_{MAN}$ kg/ha	$CV_{RMSE\text{-}MAN}$ %
Nov	0.5281 × CMR – 1,169.7	0.393*	601	71.4	848	100.7
Dec	0.0919 × CMR + 182.87	0.420*	275	41.7	1,248	189.4
Feb	0.5242 × CMR – 1,262.4	0.936**	297	19.4	977	63.8
Mar	0.3052 × CMR – 418.03	0.807**	377	29.0	1,114	85.8
Apr	0.3309 × CMR + 9.586	0.678*	881	38.4	1,842	80.2
May	0.3009 × CMR + 560.53	0.617*	1,143	40.5	1,907	67.6
June	0.4525 × CMR + 656.88	0.506*	1,622	46.9	2,253	65.1
Feb-Mar	0.3433 × CMR – 521.58	0.758**	495	35.8	958	69.3
Apr-May	0.3083 × CMR + 383.28	0.619*	1,079	41.2	1,426	54.4

[1] CMR = capacitance; CV_{RMSE} = coefficient of variation of root mean square error; DM = dry matter; MAN = based in the equation proposed by the Grasmaster II manufacturer; (DM = 0.72 × CMR – 2,200); r = correlation coefficient; RMSE = root mean square error; * = correlation significant at the 0.05 level; ** = correlation significant at the 0.01 level.

this example, November and December), due to reduced air temperature; followed by increasing productivity at the end of Winter and beginning of Spring; and a peak of production at the end of Spring (May in the case of GM and June in the case of DM), a behavior also observed at the level of capacitance. The DM content, in %, remained at around 18-23% between November and April, with an important increase from May onwards (around 26%) and, especially June (37%), coinciding with the final stages in the growth cycle of the pasture.

February corresponds to the maximum correlation coefficients and minimum CV_{RMSE} (r=0.936; P<0.01; RMSE=297 kg/ha and CV_{RMSE}=19.4% between capacitance and DM; r=0.919; P<0.01; RMSE= 2,404 kg/ha and CV_{RMSE}=28.3% between capacitance and GM). Strong correlations were also obtained in March (r=0.807; P<0.01; RMSE= 377 kg/ha and CV_{RMSE}=29.0% for DM; r=0.886; P<0.01; RMSE=1,971 kg/ha and CV_{RMSE}=29.3% for GM). According to Jamieson *et al.* (1991), the estimate can be considered excellent, good, acceptable or poor, if CV_{RMSE} for <10, 10-20, 20-30 or >30%, respectively. Based on this criterion, the estimate of DM was good in February and acceptable in March, while the estimate of GM was acceptable in February and March.

As the growth cycle of the pasture progressed, a tendency was observed for diminishing correlation coefficients between capacitance and pasture productivity, decreasing the reliability of the estimate at the time of peak production (April-May), although with significant correlation coefficients, but large root medium errors (>30%). In the last column of Table 4 are included the square error obtained when using the equation provided by the manufacture of the probe ($RMSE_{FAB}$ and $CV_{RMSE-FAB}$). The errors obtained by the use of this equation ($CV_{RMSE-FAB}$>60%) are due, on the one hand, to the diversity of Mediterranean pastures when compared to the *standard pasture* of New Zealand, and on the other hand, the effect of profound annual and inter-annual seasonality of the Mediterranean climate, which requires adaptation of the slope of the regression line, which cannot be expressed in a linear equation prepared for general use.

Thus, it is clear, that the more favorable period for the use of the Grassmaster II probe in dry land pastures in Alentejo coincides with the end of winter and beginning of spring (February-March). The probe can thus help the farmer organize the rotation of the animals through the various parcels, dynamically adapting the animal stocking to the evolution of the pasture over the year as a response to the rainfall pattern, especially in the critical spring period. If rainfall is distributed in a balanced pattern during the spring period, accompanied by the normal gradual rise in the air temperature, then the maximum production peak will be boosted and the quality-quantity binomium will be sustained for a longer period, providing the animals with ample feed supply, as well as the possibility of cutting and storing part of the pasture (Jamieson *et al.*, 1991).

Ideally, the probe will maintain reasonable sensitivity until the production peak (which normally takes place in the months of April or May, which was the case of the fields considered in this study), allowing the farmer to estimate the quantity of pasture that can be cut from each field and stored as hay or silage. The corresponding regression equations obtained at two bimonthly periods (February-March and April-May) are presented in Table 3 and 4, seeking to simplify, without generalizing, the adoption of the equations to two critical periods of dry land pasture growth in Alentejo. These equations have significant correlation coefficients, although, naturally, with less rigor than the monthly ones and show that the capacitance probe is an important tool for the management of dry land pastures and their grazing. As demonstrated in Table 4, the equation for estimating MS from capacitance provided by the manufacturer of Grassmaster II overestimates the pasture productivity in both periods (February-March and April-May).

Conclusions

The results obtained in this work show strong correlations between pasture productivity and the capacitance measured by the Grassmaster II probe, especially in the months of February. As the pasture growth cycle progressed during spring, a tendency was observed for a decrease in the correlation coefficients between the parameters, and the reliability of the estimates decreased at the peak of production (April-May), although the correlation coefficients remained significant. The management of the animal grazing, in terms of animal stocking in each field, of animal rotation through the different fields, the calculation of the supplementary feed and forage needs for the animals is based on productivity. It can be concluded that the capacitance probe is an expedient tool

that can enable the farm manager to estimate pasture productivity and support the decision-making process in the management of dry land pasture and its respective grazing.

Acknowledgements

This work was funded by National Funds through FCT (Foundation for Science and Technology) under the Project UID/AGR/00115/2019) and by the projects PDR2020-101-030693 and PDR2020-101-031244 ('Programa 1.0.1-Grupos Operacionais').

References

Cauduro, G. F., Carvalho, P., Barbosa, C., Lunardi, R., Pilau, A., Freitas, F. *et al.,* 2006. A comparison of methods used to determine forage mass in italian ryegras pasture (*Lolium Multiflorum* Lam.). Ciência Rural 36 1617-1623.

Cicore, P., Serrano, J., Shahidian, S., Sousa, A., Costa, J.L. and Marques da Silva, J., 2016. Assessment of the spatial variability in tall wheatgrass forage using LANDSAT 8 satellite imagery to delineate potential management zones. Environmental Monitoring and Assessment 188 1-11.

Edirisinghe, A., Hill, M.J., Donald, G.E. and Hyder, M., 2011. Quantitative mapping of pasture biomass using satellite imagery. International Journal of Remote Sensing 32 2699-2724.

Efe Serrano, J., 2006. Pastures in Alentejo: technical basis for characterization, grazing and improvement. Ed. Universidade de Évora, ICAM, Gráfica Eborense: Évora, Portugal.

Handcock, R.N., Gobbett, D.L., González, L.A., Bishop-Hurley, G.J. and McGavin, S.L., 2016. A pilot project combining multispectral proximal sensors and digital cameras for monitoring tropical pastures. Biogeosciences 13 4673-4695.

Jamieson, P.D., Porter, J.R. and Wilson, D.R., 1991. A test of the computer simulation model ARCWHEAT1 on wheat crops grown in New Zealand. Field Crops Research 27 337-350.

Louhaichi, M., Hassan, S., Clifton, K. and Johnson, D.E., 2018. A reliable and non-destructive method for estimating forage shrub cover and biomass in arid environments using digital vegetation charting technique. Agroforestry Systems 92 1341-1352.

Pullanagari, R. R., Yule, I. J., Tuohy, M. P., Hedley, M. J., Dynes, R. A. and King, W.M., 2013. Proximal sensing of the seasonal variability of pasture nutritive value using multispectral radiometry. Grassland Forage Science 68 110-119.

Schaefer, M. T. and Lamb, D. W., 2016. A combination of plant NDVI and LiDAR measurements improve the estimation of pasture biomass in Tall Fescue (*Festuca arundinacea* var. Fletcher). Remote Sensing 8 1-10.

Serrano, J., Peça, J., Silva, J.M. and Shahidian, S., 2011. Calibration of a capacitance probe for measurement and mapping of dry matter yield in Mediterranean pastures. Precision Agriculture 12 860-875.

Serrano, J., Shahidian, S. and Silva, J. M., 2016a. Calibration of Grassmaster II to estimate green and dry matter yield in Mediterranean pastures: effect of pasture moisture content. Crop & Pasture Sciences 67 780-791.

Serrano, J., Shahidian, S. and Silva, J.M., 2016b. Monitoring pasture variability: optical OptRx crop sensor versus Grassmaster II capacitance probe. Environmental Monitoring and Assessment 188 1-17.

Serrano, J., Shahidian, S., Marques da Silva, J. and Carvalho, M., 2018. A holistic approach to the evaluation of the montado ecosystem using proximal sensors. Sensors 18 570.

Vickery, P. J. and Nicol, G.R., 1982. An improved electronic capacitance meter for estimating pasture yield: construction details and performance tests (pp. 1-22). Australian Animal Research Laboratories technical paper no. 9. Armidale, Australia: CSIRO.

Species distribution mapping of grass clover leys using images for targeted nitrogen fertilization

S. Skovsen[1*], M.S. Laursen[1], R. Gislum[2], J. Eriksen[3], M. Dyrmann[1], A.K. Mortensen[1], S. Farkhani[1], H. Karstoft[1], N.P. Jensen[4] and R.N. Jørgensen[1]
[1]*Aarhus University, Department of Engineering, Finlandsgade 22, 8200 Aarhus N, Denmark; ssk@eng.au.dk*
[2]*Aarhus University, Department of Agroecology, Forsøgsvej 1, 4200 Slagelse, Denmark*
[3]*Aarhus University, Department of Agroecology, Blichers Allé 20, 8830 Tjele, Denmark*
[4]*I-GIS, Voldbjergvej 14A, 8240 Risskov, Denmark*

Abstract

Targeted nitrogen fertilization of grass clover mixtures, based on the local distribution of grass and clover, can potentially increase yield, improve yield quality, and reduce nitrate leaching. In this work, the grass and clover distributions of 150 ha were mapped based on image recognition in colour images. With a high-speed down-facing image acquisition platform mounted on an ATV, the fields were systematically traversed in October 2018 to capture 17,759 geotagged and spatially distributed images. Using a trained deep convolutional neural network, the images were automatically analysed to classify every pixel and estimate the species distributions in each image. Interpolation of the clover distribution into maps showed a high variation of clover content between the fields, and frequently within them. Site-specific nitrogen fertilization strategies were then produced based on the maps.

Keywords: deep learning, remote sensing, plant classification, semantic segmentation, species mapping

Introduction

Grass and clover are often grown in mixtures to increase the yield stability and herbage quality, compared with pure grass leys (Phelan *et al.*, 2015). As clover is a nitrogen fixing crop, able to fix nitrogen from the air, the competition between the species is influenced by the applied fertilizer in the soil. Thus, an increase in nitrogen fertilization decreases the clover fraction in the biomass due to competition from faster growing grass. Since part of the herbage quality is based directly on the clover percentage, it is important preserve the clover in areas of high clover contents. A spatially sensitive fertilization scheme can achieve this, when nitrogen fertilization is focused on areas where clover presence is low. This way, biomass quantity and quality could simultaneously be improved.
While the true grass and clover distribution can be determined in small plots by hand, this approach is not feasible. To realize the potential gains in targeted fertilization and to support further research in the field, a viable, robust and accurate method to map the local composition of species at a large scale is required.
Grass clover leys are traditionally fertilized in the early spring and after each cutting of the season. As 40-50% of the annual nitrogen fertilization is applied in the early spring, the first mapping must be available early to impact entire fertilization strategy. As a result, the mapping is best carried out at an early growth stage or in the preceding autumn. Major changes in plant characteristics occur during overwintering and as a result, overwintering success is variable (Wachendorf *et al.*, 2001). However, for fertilizer decision purposes in the early spring, autumn determination is in practice the best estimate available for clover content.

DOI 10.3920/978-90-8686-888-9_79, © Wageningen Academic Publishers 2019

Mapping the grass and clover distributions using images of field canopies is actively researched. A manual segmentation of grass and clover in canopy images shows a high correlation with distributions in the biomass (Rayburn, 2014).
The combination of edge detections and morphological operations in grass clover images is the first successful automatic approach for determining the grass and clover distributions (Bonesmo *et al.*, 2004; Himstedt *et al.*, 2012). But, the reliance on low level image processing significantly hinders the robustness with regards to scale invariance, field conditions and clover species.
Using deep learning, it is possible to overcome the lack of robustness of previous approaches, and segment grass clover canopy images in plot trials with high accuracy (Skovsen *et al.*, 2017). But the labor-intensive acquisition of high quality images at a large scale limits the application.
The aim of this work is to present a fast and robust approach to map the local distribution in grass clover mixtures at a large scale using high resolution colour images and deep learning. The data acquisition was carried out in the autumn, preceding the first season of the majority of the fields. This initiated a full season of targeted nitrogen fertilization and improved feed management of the dairy farm. To ease the use of deep learning, and to encourage the research in mixed crop distribution mapping, a free to use web service was launched by I-GIS (Aarhus, Denmark). Based on the principles of this work, geotagged colour images uploaded to www.cloversense.net are automatically segmented into grass, clover and weeds, and translated into species distribution maps, ready to import into a GIS program. A random subset of 500 analysed images from one of the surveyed fields is available for closer inspection through a username made for this publication[1].

Materials and methods

In preparation for large scale targeted nitrogen trials in 2019, two farmers growing mixtures of perennial ryegrass and white clover in the Jutland region of Denmark were selected. Two farms in central Jutland and northern Jutland were surveyed on the 12th and 14th of October 2018 respectively, preceding the final cut of the season. Utilizing a high-speed camera acquisition platform mounted on an ATV (Laursen *et al.*, 2017), eight fields totalling 150 ha were traversed and sampled with an average density of 119 images per ha. A summary of the image acquisition per field is shown in Table 1.

Image acquisition platform

An image acquisition platform, originally developed for weed detection in cereal (Laursen *et al.*, 2017), was reconfigured to fit the task of surveying mixed crops. This was achieved by moving

Table 1. Summary of the image acquisition of the eight surveyed fields.

Field	Area (ha)	Acquisition time (mm:ss)	Images	Density (images/ha)	Speed (ha/hour)
A	16.9	49:25	2,423	143	20.5
B	14.6	46:49	2,170	148	18.7
C	12.5	44:11	2,163	173	17.0
D	11.3	34:38	1,380	122	19.6
E	45.8	28:16	1,422	31	97.2
F	9.4	47:06	1,878	200	12.0
G	20.5	78:12	3,324	162	15.7
H	18.8	58:46	2,999	160	19.2

[1] Username: guest@cloversense.net; Password: cloversense.

the camera mount to the front of the ATV to ensure a non-perturbed sampling of the crop. This configuration is shown in Figure 1A.
Using the built-in GPS, SPS851, Trimble (Sunnyvale, USA), the system was setup to automatically trigger the 5 Mpixel camera, GrassHopper 3, FLIR (Richmond, Canada) and the surrounding ring flash, Alienbees ABR800, Paul C. Buff Inc. (Nashville, USA). This system was set to trigger every 5th metre, based on the Euclidean distance to the last trigger location, but no faster than 1 image per second. The image quality is demonstrated in the sample image shown in Figure 1B. By using the existing acquisition platform, images were collected ensuring satisfactory quality for image analysis with existing methods.

Image segmentation into plant species
Automatic image segmentation was based on previous work in the field at Aarhus University (Skovsen *et al.*, 2017; Larsen *et al.*, 2018). A fully convolutional deep neural network for pixelwise classification (Shelhamer *et al.*, 2015) was previously trained on artificial images to distinguish grasses, clovers, weeds and soil (Larsen *et al.*, 2018) and employed to perform the image analysis. To match the resolution of the collected images to the model, the images were initially resized from 5 to 2.3 Mpixel. The output of the fully convolutional neural network is a four channel image, with the same spatial resolution as the input image. Each channel represents one of the four trained classes, and each pixel value of the channel corresponds to the probability of the given class in the pixel. Applying a threshold on the probability of classifying grass, clover and weeds in the collected images, every image was automatically partitioned into grass, clover, weeds and background pixels. This technique is illustrated in Figure 2 where three of the collected images of varied clover content are analysed automatically.
Based on previous studies in grass clover mixtures (Himstedt *et al.*, 2012; Skovsen *et al.*, 2017), the clover fraction of the dry matter biomass is estimated using a linear fit between the clover content in the visible canopy leaf area and that in the biomass.
Changing the visualization scale from pixels in an image to images in a field reveals clear trends in the grass clover fields. Figure 3A illustrates the geolocated images of field D in a universal transverse Mercator (UTM) co-ordinate system with a colouring corresponding to the predicted dry matter clover content in each image. The boundaries of the field are clearly dominated by black dots, corresponding to pure grass. Figure 2A originates from this region. A large part of the centre of the field is consistently rich in clovers, corresponding to Figure 2C.



Figure 1. (A) The image acquisition platform mounted in front of the ATV. (B) Sample image from the survey collected at a speed of 5 m/s with a ground sampling distance of 5.6 pixel/mm. The image is cropped to illustrate the image quality. Clover leaves typically varies from 12 to 130 pixels in length within the images.

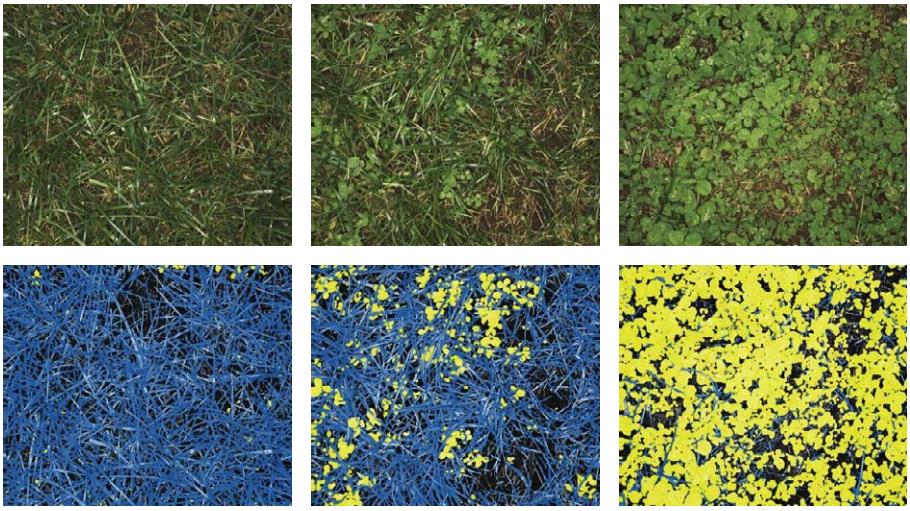


Figure 2. Three collected image samples (top) from field D with increasing levels of clover from left to right and their corresponding automatic segmentations (bottom). Yellow is clover, blue is grass, black is background. Each image covers an area of 0.44×0.37 m.

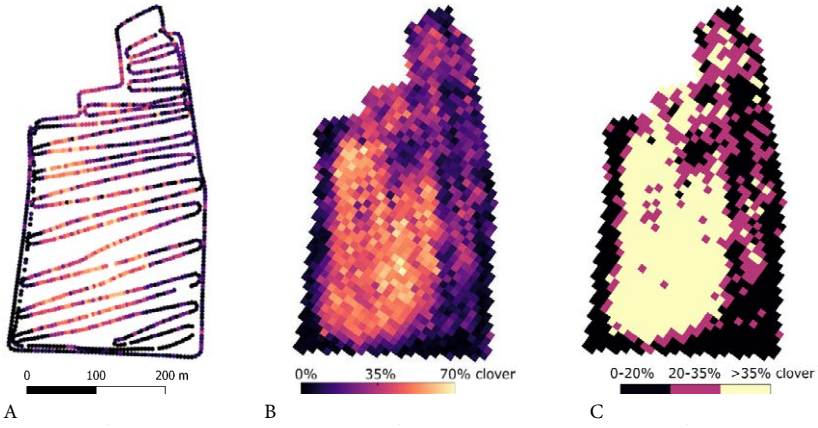


Figure 3. Visualizations of the clover content in field D. (A) Each dot represents an image location in field D. The colour represents the predicted dry matter clover content in the image. (B) Interpolated map of the clover content in the dry matter. (C) Discretized map of the clover content in the dry matter into three levels, corresponding to the current Danish guidelines for targeted nitrogen fertilization of grass clover.

Interpolation into species composition maps

Field D was surveyed with 1,380 images corresponding to 122 images per hectare. Taking the image coverage into consideration, these images cover only 0.2% of the entire field. To map the in-between clover contents, the large-scale trends are exploited with a linear 2d interpolation between the image locations. Due to the sparsity of the image samples, and to supress small scale fluctuations,

the interpolated resolution was reduced to 10×10 m grids. The resulting interpolated map of the clover content in the dry matter of field D is shown in Figure 3B.

Targeted fertilization strategies

Current guidelines for nitrogen fertilization of grass clover mixtures in Denmark differentiate between three levels of clover. Therefore, discretizing the mapped clover content into regions of 0-20, 20-35 and 35-100% directly states the targeted nitrogen fertilization strategy. This is demonstrated for field D in Figure 3C. Here, the black area should receive the maximum permitted amounts of nitrogen, whereas the light area should receive none.

Results

Five images were selected for manual annotation based on their visual clover grass distribution to cover a large range of variation between them. A subimage of 1000×1000 pixels were cropped from each image, manually annotated and compared to the output of the neural network (Figure 4). The image segmentation appears robust, and the large variation in size, depending on the growth stages, is handled well. Overall, the leaves of grasses and clovers are correctly detected and classified. But visible clover stems in are frequently misclassified as grass, and grasses are often segmented wider in the automatically segmented images than the manually annotated images. This leads to a general

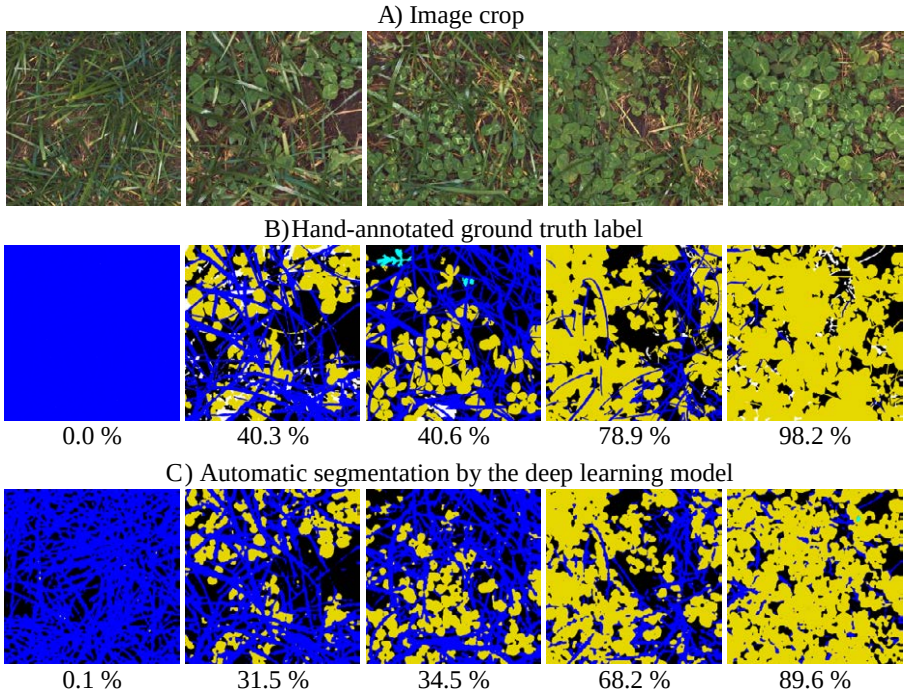


Figure 4. Evaluation of 5 annotated image crops of 1000×1000 pixels of increasing clover content. (A) Annotated image crop. (B) Ground truth label. (C) Automatically segmented image based on the deep learning model. The percentages below each image specify the percentage of clover in the vegetation pixels. Yellow is clover, blue is grass, teal is weeds, black is background, white is unknown vegetation.

decrease in the estimated visual clover content by 6-9 percent points when clovers are present in large numbers.
The mapped clover distributions in the eight surveyed fields are shown in Figure 5. Field A, F, G and H have very low amounts of clover. Field A and B, although located adjacently, show a clear difference in the average clover content. Field D and E demonstrate the highest levels of within-field variance with defined regions ranging from pure grass to pure clover.

Discussion

The uncertainty of the composition mapping is difficult to quantify due to multiple scales of inhomogeneity in the mixed crop. Previous research in the segmentation approach has stated a standard deviation of 7.8% in estimating the clover content of the dry matter within 0.25 square meter plots (Skovsen *et al.*, 2017). Utilizing the spatial context of the images, it is expected to suppress the sub-meter variation and gain an overall improved accuracy in the composition map.
The within-field variance in the clover content is most prominent in field D and E. The high ratios of clover are possibly a result of a fertilization trial on those fields in the spring of 2018. This led to a stretch of increased clover content 30 m wide extending throughout the two fields in the first cut of 2018.

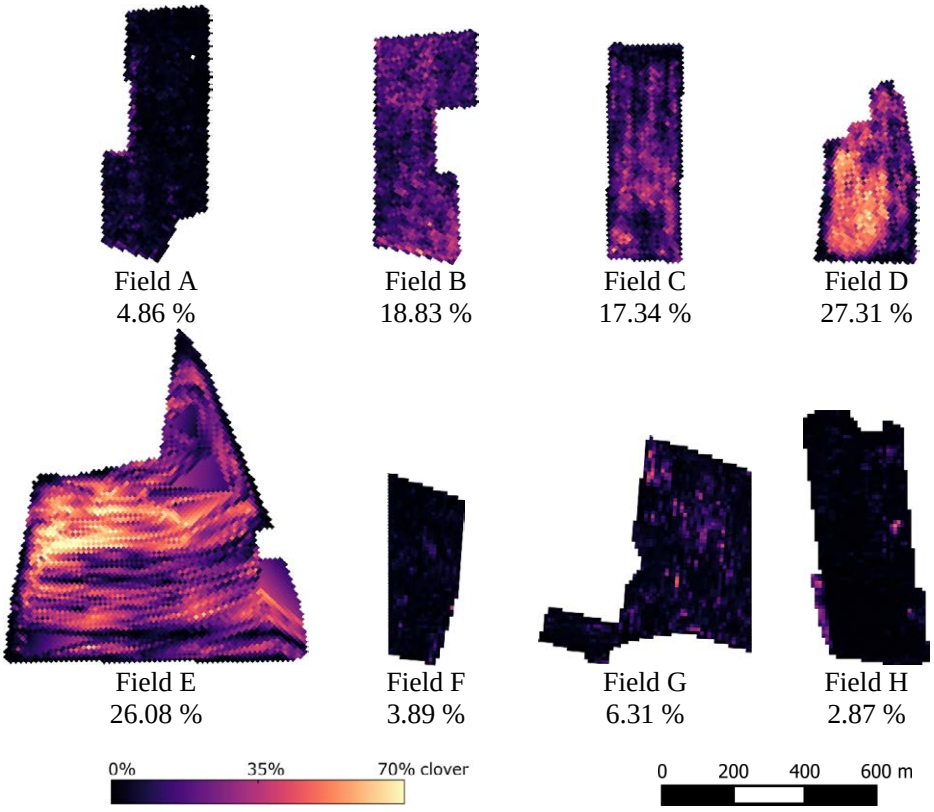


Figure 5. The mapped clover distributions in surveyed fields. The average clover distribution of each field, assuming a uniform yield, is shown below each subfigure.

The fields were surveyed with varied image densities and driving patterns. Only the lowest sampling density of 31 images per hectare in field E led to clear interpolation artefacts. This was amplified by the high variance present in the field. The highest variation has generally been observed perpendicular to the paths of the tractor. To effectively make use of the high quantity of images in the driving direction, it is advisable to navigate the ATV perpendicular or angled related to the paths of the tractor, despite a lower sampling speed.

Conclusions

An ATV traversing 150 ha of grass clover mixtures in 6.5 h successfully led to detailed mapping of the spatial grass and clover distributions. A sampling density of 120 images per hectare provided detailed distribution maps of the grass and clover content. In the spring of 2019, these will be directly applicable for applying a targeted nitrogen fertilization. The ATV survey approach presented here has the potential to deliver detailed maps at 20 ha/h. This would allow the farmer to target the nitrogen fertilization and optimize the feed management, and supports future research in the field. To minimize the workload of the farmer, the camera acquisition platform could be mounted directly on the disc mower to map the species distributions when mowing the field.

Acknowledgements

The project was funded by Green Development and Demonstration Programme (GUDP) under the Danish Ministry for Food, Agriculture and Fisheries, and Innovation Fund Denmark.

References

Bonesmo, H., Kaspersen, K., Bakken, A.K., 2004. Evaluating an image analysis system for mapping white clover pastures. Acta Agriculturae Scandinavica. Section B, Soil and Plant Science, 54:2, 76-82.

Himstedt, M., Fricke, T., Wachendorf, M., 2012. The benefit of colour information in digital image analysis for the estimation of legume contribution in legume-grass mixtures. The Crop Science Society of America, 52, 943-950.

Larsen, D., Skovsen, S., Steen, K. A., Grooters, K., Eriksen, J., Green, O., *et al.,* 2018. Autonomous mapping of grass-clover ratio based on unmanned aerial vehicles and convolutional neural networks. In 14th International Conference on Precision Agriculture. www.ispag.org/proceedings/?action=abstract&id=5107

Laursen, M.S., Jørgensen, R.N., Dyrmann, M., Poulsen, R., 2017. RoboWeedSupport – sub millimeter weed image acquisition in cereal crops with speeds up till 50 km/h. International Journal of Agricultural and Biosystems Engineering, 11 (4) 311-315.

Phelan, P., Moloney, A.P., McGeough, E.J., Humphreys, J., Bertilsson, J., O'Riordan, E.G., *et al.,* 2015. Forage legumes for grazing and conserving in ruminant production systems. Critical Reviews in Plant Sciences, 34, 281-326.

Rayburn, E.B., 2014. Measuring legume content in pastures using digital photographs. Forage and Grazinglands, 12 (1), DOI: 10.2134/FG-2011-0143-MG

Shelhamer, E., Long, J., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431-3440.

Skovsen, S., Dyrmann, M., Mortensen, A.K., Steen, K.A., Green, O., Eriksen, J., *et al.,* 2017. Estimation of the botanical composition of clover-grass leys from RGB images using data simulation and fully convolutional neural networks. Sensors, 17:12, 2930.

Wachendorf, M., Collins, R.P., Connolly, J., Elgersma, A., Fothergill, M., Frankow-Lindberg, B. E, *et al.,* 2001. Overwintering of Trifolium repens L. and Succeeding Growth: Results from a Common Protocol carried out at Twelve European Sites. Annals of Botany, 88:1, 669-682.

Section 11
Drainage optimization and variable rate irrigation

Evaluating and improving soil sensor-based variable irrigation scheduling on farmers' fields in Alabama

L. Bondesan[1,2], B.V. Ortiz[2*], G.T. Morata[2], D. Damianidis[2], A.F. Jimenez[2,3], G. Vellidis[4] and F. Morari[1]
[1]*Department of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova, Italy*
[2]*Crop, Soil, and Environmental Sciences Department, Auburn University, Auburn, AL, USA; bortiz@auburn.edu*
[3]*Department of Mechanics and Mechatronics Engineering, Univ. Nacional de Colombia, Colombia*
[4]*Crop and Soil Sciences Department, University of Georgia, Tifton, GA, USA*

Abstract

During recent decades, crop yields in the Southeast USA have shown an increasing trend which in some cases coincides with irrigation adoption. This increase in irrigated areas has not always included the use of water-smart irrigation technologies; however, ongoing on-farm projects are evaluating their benefits. During 2017 and 2018, a study conducted in south Alabama, USA focused on the comparison between the farmer's standard irrigation practices and technology-based irrigation strategies combining variable rate irrigation (VRI) and soil sensor-based scheduling (SS). Data showed that VRI and SS could result in water savings, preventing under- or over- irrigation without adversely affecting yield.

Keywords: corn, irrigation, soil sensors, irrigation scheduling, variable rate irrigation

Introduction

Irrigation water management is one of those practices that farmers either pay a lot of attention to or overlook, which could cause impacts to the crop and the environment. The use of irrigation for row crop production in Alabama is increasing; however, the adoption of water-smart irrigation technologies is low. In Alabama, of the 1,022 farms using irrigation scheduling methods in 2012, only 7% used soil moisture sensors and 4% used reports on daily crop-water ET (USDA-NASS, 2014). Attendees at the 2016 Alabama Corn and Wheat short course indicated lack of knowledge on interpretation of soil moisture sensor data (58% of respondents) and application of the correct rate (53%), as well as the lack of use of irrigation scheduling methods (41%). Several studies have shown positive benefits from the use of sensor-based irrigation scheduling and variable rate irrigation strategies. Variable rate irrigation (VRI) is an effective method to reduce yield variability and increase overall productivity, as well as reduce leaching and soil erosion by increasing water use efficiency (Sadler *et al.*, 2005; King *et al.*, 2009). Water conservation benefits that can result from the use of VRI include significant water savings, reduction or elimination of irrigation in wet areas, and spatial management of within-field water deficits (Evans and King, 2012). A complement to VRI is the use of soil sensors to improve irrigation scheduling (SS). Distributed, in-field sensor networks with wireless communications can be used to measure soil water and plant response to irrigation or rainfall and can assist in the development and implementation of site-specific irrigation water management (Vellidis *et al.*, 2013; Vellidis *et al.*, 2008). Irmak *et al.* (2012) conducted on-farm demonstrations of SS irrigation scheduling using Watermark® (Irrometer Co., Inc., Riverside, CA, USA) sensors that resulted in 33% fewer irrigation applications while maintaining crop yields and energy savings of $44-67/ha due to less water withdrawals. Several on-farm demonstrations are currently being conducted in Alabama to demonstrate the benefits of various irrigation management

DOI 10.3920/978-90-8686-888-9_80, © Wageningen Academic Publishers 2019

approaches. The main objective of this study was to evaluate soil sensor-based irrigation scheduling and to compare VRI, aided by soil sensors, with farmers' irrigation management strategies.

Materials and methods

Study field description and data collection

The data used for this study was collected from a precision irrigation project, conducted on a 25 ha field located in Samson, Alabama, USA (31°07'02.9'N 86°05'32.0'W) during the 2017 and 2018 growing seasons. In 2018, the field was planted with the Dyna-Gro 57VP51 (Dyna-Gro Seed, Loveland Products Inc., Loveland, CO, USA) corn hybrid of 117 days relative maturity. During both years, the corn crop was planted in mid-March and harvested in early August. Flowering occurred on May 19th and June 5th in 2017 and 2018, respectively. The seeding rate used was 79,000 seeds/ha and the total nitrogen rate applied was 308 kg/ha and 393 kg/ha in 2017 and 2018, respectively. A John Deere grain combine 9570 STS with a yield monitor was used for harvest and to record yield data. Prior planting in 2017, the variability of the field was assessed by collecting apparent soil electrical conductivity (Soil EC_a), terrain elevation, and soil samples for soil texture and nutrients assessment. A small tractor pulling a VERIS 3100® (Veris Technologies Inc., Salina, KS, USA) sensor for the collection of Soil EC_a data was also fitted with a real time kinematic (RTK) Trimble GPS (Trimble Inc., Sunnyvale, CA, USA) to collect terrain elevation data. These data sets were used to study within-field variability and identify locations for soil sampling used for the determination of soil texture and soil water retention characteristics. Five locations were selected for collection of undisturbed soil samples at soil depths of 150, 300 and 600 mm. Every location-sampling depth was used for determination of soil water retention curves (SWRC), field capacity (FC), and permanent wilting point (PWP) values. The SWRCs were determined using a Hyprop® (METER Group Inc., Pullman, WA, USA) sensor, while soil water content values at high water tension values close to the PWP value, approximately -1.5 MPa, were determined using a dew point potentiometer WP4C sensor (METER Group Inc., Pullman, WA, USA).

Management zones and soil water monitoring for irrigation

The Management Zone Analyst (MZA) software (Fridgen, *et al.*, 2004) was used to delineate irrigation zones initially based on soil EC_a and terrain elevation data. The management zones (MZ) used in the soil sensor-based variable rate irrigation treatment (SS-VRI) were later refined using 2017 yield data and true-colour soil imagery. The differences between zones were ground truthed with soil texture data collected from different locations within the field and with farmer's field variability knowledge. The co-operating farmer started using variable rate irrigation (VRI) only 2 years prior to this study. The farmer delineated irrigation zones based on soil map units of the SSURGO soil survey (NRCS, 2011). Data of saturated hydraulic conductivity, which is correlated with infiltration, and available water storage in the soil profile were also available on the SSURGO survey and used by the farmer to determine maximum irrigation rate and irrigation frequency by soil type. Soil sensors were installed in each zone-treatment combination to monitor differences in soil matric potential (SMP), and to determine irrigation scheduling for each SS-VRI zone managed by Auburn University. The soil sensor probes installed comprised three Watermark® (Irrometer Co, Inc., Riverside, CA, USA) sensors for monitoring SMP changes at the depths of 0.15, 0.3, and 0.6 m of the same location. Those sensors were arranged in a wireless mesh network to allow communication and data transfer between sensors and to a base station installed at the field (Vellidis *et al.*, 2013). Although soil water tension data is correlated with plant water stress (Kramer, 1963) (it provides measurements analogous to the force (suction) needed by plants to extract water from the soil), these values were converted to soil water content, which farmers are more familiar with.

Comparison between irrigation practices

Irrigation rate, timing and placement decisions were compared by imposing irrigation treatments under a 355 m Zimmatic (Lindsay Mfg. Co., Omaha, NE, USA) Centre Pivot which had a Lindsay® Growsmart precision VRI system. Although the farmer uses several strategies to decide when and how much to irrigate, for the purpose of this study, those will be described as the evapotranspiration-based water balance method. The rain and irrigation are water inputs and weekly crop water use (evapotranspiration), deep percolation, and runoff are outputs. The farmer recorded changes of weekly crop growth and data of crop water use was gathered from extension bulletins. Because this research was trying to show the advantages of soil sensor-based irrigation scheduling, data from soil sensors was used to estimate soil water depletion and the irrigation rate was estimated as the amount necessary to reach the soil field capacity water content. The comparison of irrigation strategies was conducted by establishing pair-comparison irrigation strips, 40 m in 2017 and 80 m in 2018, spanning the length of the field. The strip trial compared farmer irrigation strategies (Farmer treatments) with soil sensor-based smart irrigation strategies. Within his strips, the farmer applied different rates of water in 2017 (F-VRI) but uniform irrigation during 2018 (F-URI) growing season. The SS-VRI treatment consisted in the application of MZ-specific irrigation rates using Watermark® soil sensors' data for real time evaluation of the soil moisture conditions. However, the VRI of both years only started at flowering time and prior to that, the SS-VRI treatments received the same rate as the farmer's treatments (FTs). In most cases, the strips ran across various irrigation zones, therefore soil sensors were installed not only in each treatment comparison but also in each treatment-MZ combination. Although the goal was also to evaluate irrigation timing, these decisions were limited because irrigation on SS-VRI was applied in correspondence with the FTs due to the time it takes for the pivot to cover the whole field. However, irrigation rate by location changed when needed. The SS-VRI irrigation rate was determined based on real-time SMP values measured by each sensor probe installed in each zone. The SMP data values were converted to soil water content using measured SWRC, FC and PWP values of a respective MZ. Therefore, hourly values of soil moisture content up to 600 mm depth were calculated and used to estimate soil water depletion and irrigation rate. Differences between irrigation strategies were evaluated by comparing irrigation decisions by zone and treatments within a zone with respect to: in-season irrigation rates and timing, total irrigation applied at the end of the season, yield, irrigation water use efficiency (IWUE = yield/total irrigation applied), water use efficiency (WUE = yield/irrigation plus rainfall), and water savings (what could have been done if the whole field was managed with SS-VRI).

Results and discussion

Field variability and farmer irrigation strategies

Within-field variability in soil texture, soil colour and soil water characteristics were identified through collection of soil samples, soil imagery, soil EC_a (Sudduth *et al.*, 2005) (Figure 1A) and SWRC (Table 1). Although the SSURGO soil survey identified three different soil types over the study area, the soil EC_a data (Figure 1A) showed larger spatial variability. In 2017 for example, the farmer treated zones 1, 2, and 5 as a single unit for irrigation purposes; however, data of elevation, soil EC_a, soil colour and texture as well as 2017 SMP indicated that this area should be managed as three different zones (Figure 1A). For the purpose of this paper, the discussion of results will be focused on the data collected from zones 1 and 4: the two most contrasting zones (Table 1).

Comparison of irrigation treatments within a zone

2018 – The period of June 5th to 27th was selected for comparison of irrigation prescriptions between the FTs and SS-VRI treatments. This period corresponded with silking and early grain filling period: the highest crop water demand period. Irrigation prescriptions were compared within zone 4, characterized by low soil water holding capacity (Table 1). Figure 2 shows the differences in soil water

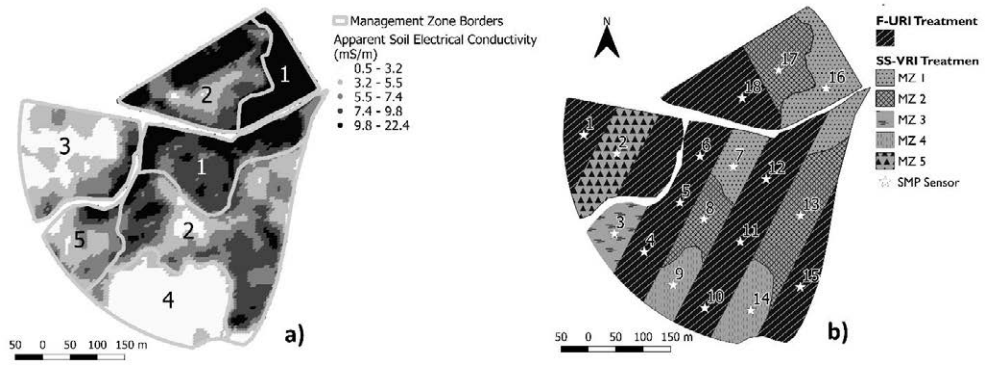


Figure 1. Within-field variability represented by changes in apparent soil electrical conductivity (ECa) with relative irrigation management zones delineated (1A). Experimental design of the treatment comparison with relative soil sensor locations for the year 2018 (1B).

Table 1. Soil properties differences between irrigation management zones in terms of soil texture and soil water holding capacity at field capacity (FC) and permanent wilting point (PWP) at 300 mm depth.

Zone	Sensor #	FC (m^3/m^3)	PWP (m^3/m^3)	% Sand	% Clay	Soil textural class
1	7	0.359	0.255	46.8	34.4	sandy clay loam
2	8	0.204	0.076	73.2	8.4	sandy loam
3	3	0.191	0.098	77.6	4.4	loamy fine sand
4	9	0.153	0.060	87.2	4.4	fine sand
5	2	0.070	0.020	88.8	4.4	fine sand

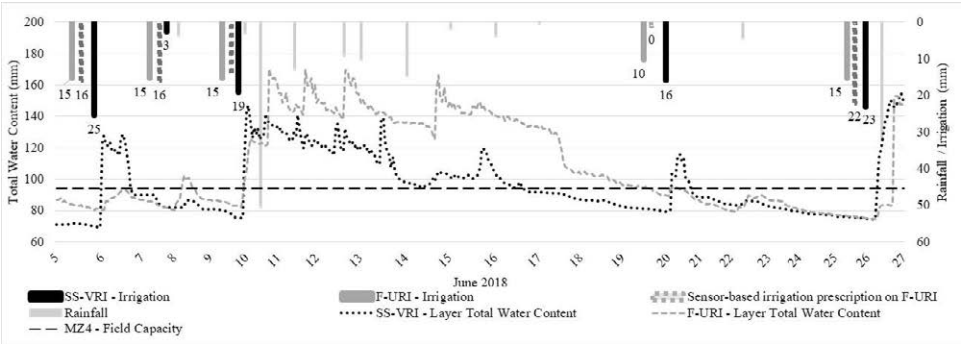


Figure 2. Differences in soil water content and irrigation prescriptions between both F-URI (Grey line) and SS-VRI (black line) MZ 4 during 2018 growing season.

content and irrigation prescriptions between both treatments. Five irrigation events occurred on the F-URI treatment during that period (grey solid bars); the farmer applied irrigation when the soil was at 16% soil water depletion from FC (dashed horizontal line). However, the farmer over-irrigated

and under-irrigated at times. When the SMP data collected on the F-URI was input into the soil-sensor based irrigation prescription method developed by Auburn University, the rate selected by the farmer on three out of five events agreed with the system prescriptions (Figure 2). On June 19th in contrast, the water content in the F-URI was close to field capacity; however, the farmer applied 10 mm of water, but the soil-sensor based irrigation scheduling system did not prescribe irrigation that day (Figure 2). On June 26th, even though the water content on both treatments was similar, the farmer applied 7 mm less than what the soil sensors installed in his treatment were prescribing, indicating that an under-irrigation occurred that day.

Evaluation of irrigation strategies between zones

2017 – Table 1 shows higher soil water holding capacity in MZ 1 (Figure 3 – grey horizontal dash line) compared with zone 4 (Figure 3 – black horizontal dash line). The FC at 600 mm soil depth in MZ 1 was around 216 mm contrasting with 94 mm in MZ 4. These differences suggested the need for VRI and sensor-based irrigation scheduling. The period of data shown in Figure 3 corresponds to corn flowering (May 19th, 2017) when two irrigation events were scheduled on the SS-VRI treatment under the conditions of MZ 1 (grey vertical bars) and three events in MZ 4 (black vertical bars). The frequent rainfall events during the period reduced the need for irrigation. Differences of irrigation frequency and rate between the F-VRI and the SS-VRI treatments were observed in both MZ 1 and 4. The farmer scheduled four irrigation events in both MZ 1 and 4 while only two events were prescribed in MZ 1 and three in MZ 4 using the soil sensor data (grey and black vertical bars, Figure 3).

2018 – Differences of irrigation strategies among zones were also observed in 2018. The SS-VRI treatment in MZ 4 (fine sandy soil) was irrigated five times compared to three times on zone 1 (sandy clay loam soil) (Figure 4). Besides the differences in irrigation frequency, there were differences in irrigation rates. On June 9th, the SS-VRI in MZ 4 received 19 mm of irrigation and only 7 mm were prescribed and applied in MZ 1. Similar differences in irrigation amounts were observed on June 20th. On June 5th and 25th, irrigation was not applied on zone 1; however, zone 4 received irrigation within a range of 23-25 mm. In 2018, the SS-VRI demonstrated also the ability to better define the irrigation amounts based on real-time soil moisture conditions, which could help farmers with irrigation decision-making process.

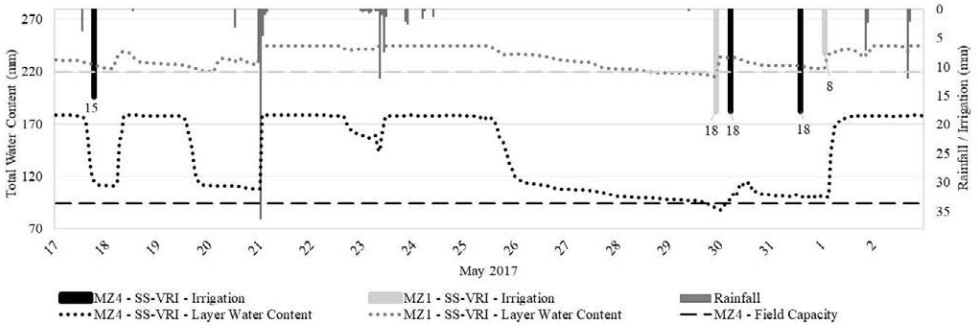


Figure 3. Differences in soil water content and irrigation rates between MZ 1 and 4 within the SS-VRI treatment using soil sensor-based irrigation scheduling during 2017 growing season.

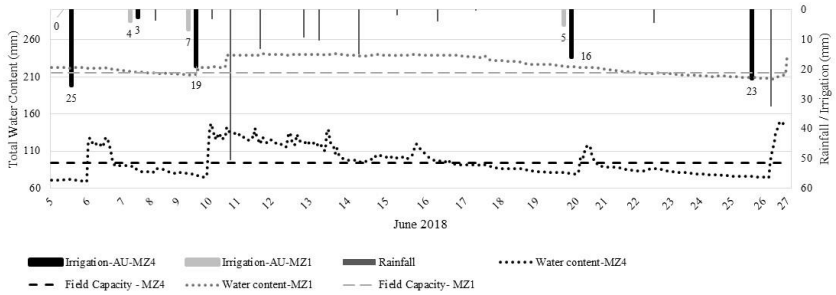


Figure 4. Differences in soil water content and irrigation rates between MZ 1 and 4 within SS-VRI treatment using the soil sensor-based irrigation scheduling during 2018 growing season.

Increasing irrigation water use efficiency with soil sensors and VRI

In both years, the use of soil-sensor based irrigation scheduling resulted in water savings and higher IWUE and WUE in MZ 1. The water savings in that zone were 26 and 35% in 2017 and 2018 respectively (Table 2 and 3). These savings resulted in higher IWUE during both years because of the use of soil sensor data for irrigation scheduling compared with the scheduling strategies used by the farmer. In 2018, even though corn yield for the SS-VRI treatment was slightly lower compared to the F-URI, higher efficiencies were recorded. MZ 1 is the zone with the highest soil water holding capacity within the field, and the results of this study show that the sensor-based irrigation scheduling can prevent over-irrigation under these conditions, especially after frequent rainfall events (June 10th to 20th, 2018 – Figure 2). Overall, less irrigation was applied to the field in 2018 than in 2017. The rainfall occurred in May and July 2018, corresponding to plant vegetative stages and grain filling period, was almost double the amount received in 2017. In 2018, higher water savings and IWUE were overall achieved, with higher efficiencies in MZ 2 and 3, which cover a large portion of the study area (Table 2). The most challenging area for irrigation management was zone 4 with a fine sandy soil type. During both years, an increased irrigation rate did not result in important changes in IWUE or yield. As explanation of such event was hypothesized that the coarse soil texture along with the rainfall and high irrigation rates applied could have impacted on soil fertility and nutrient availability. The results of the 2018 season showed the benefits of VRI for irrigation management, especially in terms of IWUE. In 2018, the farmer was asked to use only uniform rate irrigation on

Table 2. Irrigation efficiency comparison between treatments and MZs in 2017.

Zone	Treatment	Irrigation (mm/season)	IWUE (kg/mm)	WUE (kg/mm)	Yield (kg/ha)	Water saved (%)
1	SS-VRI	116.8	102.4	16.2	11,965	26
	F-VRI	157.5	73.3	14.8	11,547	
2	SS-VRI	147.3	82.5	15.8	12,152	6
	F-VRI	157.5	79.2	16.0	12,480	
3	SS-VRI	142.2	89.2	16.6	12,681	9
	F-VRI	157.5	80.0	16.2	12,603	
4	SS-VRI	228.6	46.7	12.6	10,684	0
	F-VRI	215.9	51.3	13.2	11,075	
5	SS-VRI	228.6	50.1	13.5	11,456	0
	F-VRI	215.9	52.7	13.6	11,383	

Table 3. Irrigation efficiency comparison between treatments and MZs in 2018.

Zone	Treatment	Irrigation (mm)	IWUE (kg/mm)	WUE (kg/mm)	Yield (kg/ha)	Water saved %
1	SS-VRI	81.5	138.6	14.0	11,299	35
	F-URI	125.5	91.4	13.5	11,469	
2	SS-VRI	71.1	172.1	15.3	12,236	43
	F-URI	125.5	99.5	14.7	12,485	
3	SS-VRI	54.4	170.3	11.9	9,259	57
	F-URI	125.4	96.0	14.1	12,042	
4	SS-VRI	140.5	71.0	11.5	9,977	0
	F-URI	125.5	69.1	10.2	8,674	
5	SS-VRI	84.3	119.9	12.5	10,109	33
	F-URI	125.5	71.9	10.6	9,018	

his treatments while the irrigation rate changed between zones on the SS-VRI treatments with the help of soil sensors data. Even though VRI was only established after flowering, data are showing that each zone had different irrigation requirements. Although the farmer applied different irrigation rates in 2017, he is currently managing zones 1 to 3 as a single unit or zone; however, two years of study indicate that the irrigation requirements over that area are different and therefore it should be treated as three different zones. When the area covered by zones 1, 2 and 3 were used to calculate water savings in 2018 as a result of VRI and soil sensor-based scheduling, savings of 35, 43 and 57% were found in those zones, respectively.

Conclusions

Two years of on-farm research show that soil-sensor based irrigation scheduling and variable rate irrigation could result in water saving and improved irrigation water use efficiencies. The use of soil sensor data for irrigation scheduling could prevent under or over irrigation and support the application of the right rate at the right location with the use of a variable rate irrigation system.

Acknowledgements

We want to acknowledge our co-operating farmer, Jim Lewey, for his collaboration. Funding for this project was provided by the Alabama Soil and Water Conservation Committee, Alabama Agricultural Experiment Station, Alabama Association of Conservation Districts, and Natural Resources Conservation Service (NRCS).

References

Evans, R. G. and King, B. A., 2012: Site-specific sprinkler irrigation in a water-limited future. Transactions of the ASABE 55, 493-504.

Fridgen, J.J., Kitchen, N.R., Sudduth, K.A., Drummond, S.T., Wiebold, W.J., Fraisse, C.W., 2004. Management zone analyst (MZA): software for subfield management zone delineation. Agronomy Journal 96 (1), 100-108.

Irmak, S., Burgert, M. J., Yang, H. S., Cassman, K. G., Walters, D. T., Rathje, W. R. *et al.,* 2012. Large-scale on-farm implementation of soil moisture-based irrigation management strategies for increasing maize water productivity. Transactions of the ASABE 55, 881-894.

King, B. A., Wall R. W. and Karsky, T. F., 2009: Center-pivot irrigation system for independent site-specific management of water and chemical applications. Applied Engineering in Agriculture 25, 187-198.

Kramer, P. J., (1963). Water Stress and Plant Growth 1. Agronomy Journal 55(1), 31-35.

Natural Resources Conservation Service, United States Department of Agriculture, 2011. Soil Survey Geographic (SSURGO) Database. Available online at https://sdmdataaccess.sc.egov.usda.gov. Accessed 12/08/2018.

Sadler, E. J., Evans, R. G., Stone K. C. and Camp, C.R., 2005: Opportunities for conservation with precision irrigation. Journal of Soil and Water Conservation 60, 371-378.

Sudduth, K. A., Kitchen, N. R., Wiebold, W.J., Batchelor, W.D., Bollero, G.A., Bullock, *et al.,* 2005. Relating apparent electrical conductivity to soil properties across the north-central USA. Computers and Electronics in Agriculture 46, 263-283.

USDA – NASS, 2014: Farm and ranch irrigation survey (2013). Series; 2012 Census of Agriculture. United States Department of Agriculture – National Agricultural Statistics Service.

Vellidis, G., Tucker, M., Perry, C., Kvien, C. and Bednarz, C., 2008: A real-time wireless smart sensor array for scheduling irrigation. Computers and Electronics in Agriculture 61, 44-50.

Vellidis, G., Tucker, M., Perry, C., Reckford, D., Butts, C., Henry, H. *et al.*, 2013: A soil moisture sensor-based variable rate irrigation scheduling system. In: J. V. Stafford (Ed.) Precision '13. Proceedings of the 9th European Conference on Precision *A*griculture. pp. 713-720. Wageningen, the Netherlands: Wageningen Academic Publishers.

Artificial neural networks for irrigation management: a case study from southern Alabama, USA

A. Jimenez[1,2,3], B.V. Ortiz[1*], L. Bondesan[1,4], G. Morata[1] and D. Damianidis[1]
[1]*Department of Crop, Soil, and Environmental Sciences, Auburn University, Auburn, AL, USA; bortiz@auburn.edu*
[2]*Department of Mechanics and Mechatronics Engineering, Universidad Nacional de Colombia, Bogota, Colombia*
[3]*Department of Mathematics and Physics, Macrypt R.G., Unillanos, Colombia*
[4]*Department of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova, Padova, Italy*

Abstract

Irrigation scheduling and application at the right time and rate is a key aspect in precision irrigation. This study proposed the 'Long Short-Term Memory' (LSTM) Neural Network model to predict irrigation prescriptions. Training data for LSTM were collected from a precision irrigation study. The estimations of hourly irrigation prescriptions used soil matric potential data collected from two contrasting soil types. Performance of the LSTM models were evaluated by comparing neural network parameters and prediction capability by soil type. The LSTM Neural Network showed good prediction capabilities for both soil types (R^2 ranging between 0.82 and 0.98).

Keywords: artificial neural network, corn, irrigation prescription, soil water tension

Introduction

Irrigation can reduce crop production risks and even increase crop yield especially on areas with degraded soils or irregular rainfall periods (Gutzler *et al.*, 2015). Adopting improved land and water management practices among other sustainable practices can assist farmers in the definition of irrigation scheduling, and therefore an increase of water use efficiency (Levidow *et al.*, 2014). Precision irrigation is the sustainable management of water resources which involves the application of water to the crop at the right time, right amount, right place and right manner, to meet specific plant water demands, while avoiding excessive or insufficient irrigation (Adeyemi *et al.*, 2018). Ensuring adequate soil water for crop growth could be done through real-time monitoring of soil moisture at several soil depths (Seidel *et al.*, 2016). However, this monitoring and the estimations derived represent a complex non-linear problem due to the relationship between climatic parameters, soil hydraulic properties and soil moisture dynamics (Mashayekhi *et al.*, 2016). Spatial variability of soil moisture is related to within-field variability of soil properties and elevation and its assessment is necessary to support variable rate irrigation (Levidow *et al.,* 2014). Temporal variability of water content is useful in forecasting next irrigation (Pandey and Pandey, 2010). Methods for predicting irrigation amounts using simple approaches, less demanding in data collection and processing, are needed to implement effective irrigation scheduling. Artificial intelligence (AI) can be used to accomplish this goal because it uses computational algorithms to perform tasks that normally require human intelligence (Kumar, 2018). In addition, AI has been applied in many areas where sensor information was collected, processed and interpreted (Gonzalez and Yu, 2018). Artificial Neural Networks (ANN) are AI models used in predictions using data from sensors (Qi *et al.*, 2018). An ANN is a computational model that is inspired by the human brain and consists of a set of processing units called neurons. The ANNs can identify and learn correlations between input data and corresponding target output values, being a useful tool in predictive modeling. Some

DOI 10.3920/978-90-8686-888-9_81, © Wageningen Academic Publishers 2019

ANN models, such as the multilayer perceptron neural network (MP), do not have the ability to learn in sequential problems because they cannot store previous information, which limits their prediction capabilities in time series forecasting (Zhang *et al.*, 2018). In the MP, the 'input layer' accepts the data, then feeds the 'hidden layers' where all the calculations are made and, finally, the information is sent to the 'output layer' to produce the outcome (Adewale *et al.*, 2018). A Recurrent Neural Network (RNN) model, an ANN technique, is useful in dynamic modelling tasks (Kumar, 2018). The term 'recurrent' is because computations derived from earlier input are fed back into the network, which gives them a kind of 'memory'. The output of such networks at a given instant reflects the current input as well as previous inputs and outputs (Goodfellow, 2016). However, in practice, training of such systems remains difficult because of a problem called 'vanishing gradient', resulting in the capture of short-term history, but dependencies spanning over longer intervals are much harder to learn (Hochreiter, 1998). Long Short-Term Memory (LSTM) network is an extension of RNN with successful application in non-linear dynamic systems (Gonzalez and Yu, 2018). The LSTM includes 'cells' that can maintain information in memory and remember their inputs for long periods. The 'internal memory' means outputs of the network are conditional on the recent context in the input sequence, not what has been presented as input to the network as the MP (Kumar, 2018). Some applications of LSTM are in hydrology (Zhang *et al.*, 2018), but few studies have shown the use of these models to define irrigation prescriptions specially using soil matric tension sensors. In summary, LSTM have been used in modelling sequential data of volumetric water content (VWC) using remote sensing (Lu *et al.*, 2015), prediction of water table depth (Zhang *et al.*, 2018) and VWC based on the past soil moisture, precipitation and climate measurements (Adeyemi *et al.*, 2018). The goal of this study was to assess the ability of LSTM neural networks as a potential irrigation scheduling tool evaluating performance and processing time. If this technique shows promising results, it could be used to support irrigation scheduling on different types of irrigation systems but especially on variable rate irrigation.

Materials and methods

Study field description and data collection

This study used data collected from a variable rate irrigation project conducted on a 25 ha field located in Samson, Alabama, USA (31.115940° N, 86.093257° W; elevation 67 m a.s.l.; 23.6 ha) during 2017 and 2018 (Figure 1). The mean annual precipitation is 1473.2 mm and the period from March to July which corresponds to the corn growing season, has a historical average of 137.16 mm. During both years, the field was planted with the Dyna-Gro 57VP51 corn hybrid (Dyna-Gro Seed, Loveland Products Inc., Loveland, CO, USA), in mid-March and harvested at the beginning of August. Flowering occurred on May 19th and June 5th in 2017 and 2018, respectively. Irrigation zones were delineated to apply water at different rates and times over different locations within the study field. In 2017, four irrigation zones were defined using apparent soil electrical conductivity (soil EC_a) and terrain elevation data analyzed through the Management Zones Analyst software (USDA-ARS, CO, USA). Data of soil water tension and yield collected in 2017 suggested that five zones better described crop growth and water needs (not shown). Then, a new irrigation zones map was delineated including 2017 yield data, soil texture and a true-color soil imagery in addition to the initial variables. In 2018, five locations within the five zones were selected for collection of both undisturbed and disturbed soil samples at 150, 300 and 600 mm depths for determination of soil water retention curves (SWRC), field capacity (FC) and permanent wilting point (PWP), by location and soil depth. The SWRCs were determined using the Hyprop® sensor (METER Group Inc., Pullman, WA, USA) and soil water content at high water tension values close to the PWP value, were determined using a dew point potentiometer WP4C sensor (METER Group Inc., Pullman, WA, USA). For the purposes of this study, the analysis focused on data obtained of the two most contrasting zones: Zone 1 (sandy clay loam soil – 47.2% Sand and 30.4% Clay at 150 mm depth) and

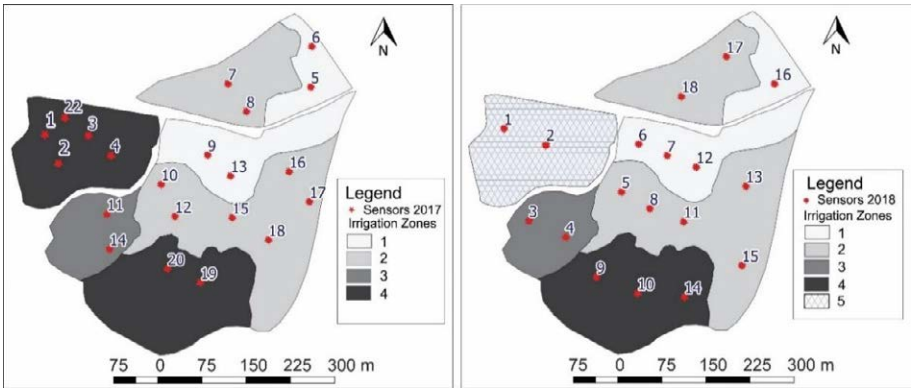


Figure 1. Study area showing irrigation zones and location of soil sensor probes for Left: 2017 and Right: 2018.

Zone 4 (loamy fine sand soil – 85.6% Sand and 6.4% Clay at 150 mm depth) (NRCS, 2019). Because of the differences in soil type in these zones, it was expected that, if the irrigation prescriptions methodology was able to predict irrigation under contrasting conditions, it would be able to do it for intermediate conditions that would be different between zones.

Assessment of soil water content and irrigation prescriptions

Changes in soil matric potential (SMP) within each zone were monitored using a wireless soil sensor system, a smart sensor array (SSA) (Vellidis *et al.*, 2016), that included multiple soil sensor nodes each one with three Watermark® sensors (Irrometer® Company, Riverside, CA, USA). Each node monitored soil matric potential (SMP) at soil depths of 150, 300, and 600 mm. Depending on the area covered by each zone, between one to three nodes were installed. The sensors installed in each zone provided data on irrigation needs. As shown in Figure 1, during the 2017 season, sensors 9 and 13 were installed in Zone 1, and sensors 19 and 20 in Zone 4. In 2018, sensors 6, 7 and 12 were installed in Zone 1 and sensors 9, 10 and 14 in Zone 4. Rainfall data were obtained using a WatchDog 2000 weather station (Spectrum Technologies, Aurora, IL, USA). Data of irrigation dates and rates by zone were also collected. Irrigation prescriptions targeting a soil depth of 600 mm were calculated using soil moisture content after transforming soil matric potential data and information of soil water characteristics such as field capacity and permanent wilting point.

Long Short-Term Memory Network (LSTM) architecture and parameters

A many-to-one LSTM model Goodfellow *et al.*, 2016) was studied in this research, which means that multiple data inputs as a time series, can be used to predict a single value (Kumar, 2018). Figure 2 shows the many-to-one LSTM architecture, where there are a single 'input layer', a single 'hidden layer' of blocks and a single 'output layer'. A block is a memory with a mechanism that controls the input, output and storage of data.

A LSTM was trained with data containing examples of the input and output pairs. Normally, the dataset is split into a training set (60%), a validation set (20%) and a test set (20%). During the training phase, the training data is used to update the network parameters. The validation data is used during training to monitor the training process and to detect overfitting. Inference is when the trained model is tested with new unseen data (Kumar, 2018). In the LSTM, one epoch corresponds to one pass of all the training examples and, in each epoch, a loss function is used to evaluate how well a specific algorithm models the dataset. If the predictions deviate too much from the actual

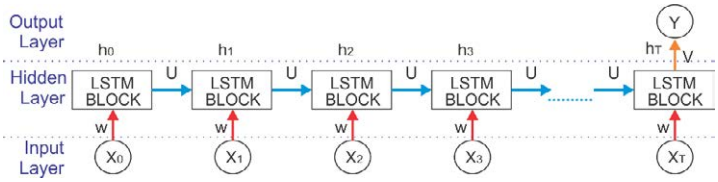


Figure 2. Many to one long short-term memory network (LSTM) architecture.

results, the loss function will produce a higher number. In addition, dropout is a technique used under training to avoid overfitting. This means that some neurons are temporarily removed from the network (Gonzalez and Yu, 2018).

Data pre-processing

The LSTM training data consisted of a time series of hourly SMP data collected at three soil depths, rainfall and irrigation amounts (Figure 3). The complete SSA nodes dataset consisted of 16,024 SMP records at three soil depths, 9,328 records from Zone 1 and 6,996 records from Zone 4. The dataset was transformed into a supervised learning problem. Then, the variables were normalized to ensure they remain on the same scale. This pre-processing guaranteed a stable convergence of parameters in the model. The standardization formula used in the dataset studied was a Min-Max Scaler as follows:

$$x^j_{i_{new}} = \frac{x^j_i - min(x^j)}{max(x^j) - min(x^j)} \quad (1)$$

where x_i represents data in i^{th} of variable j, $min(x^j)$ is the minimum value of x in the variable j and $max(x^j)$ is the maximum value of x in the variable j, where j = 1,2,..6.

Model evaluation

Several experiments were developed to define the optimal LSTM model parameters (the number of blocks, layers, dropout and epochs). Combinations of 20, 40 and 60 blocks in the first hidden layer, dropout values of 0 and 0.5, and number of epochs of 30, 60 and 100 were evaluated for

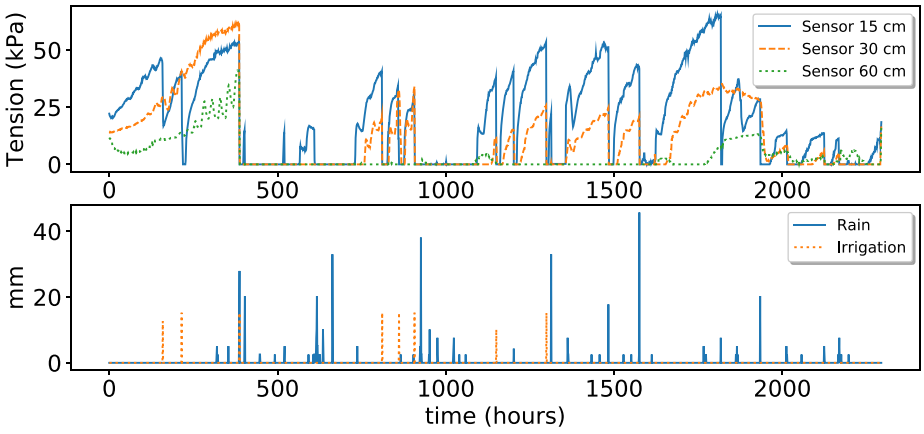


Figure 3. Top: soil water tension – three depths. Bottom: rain and irrigation amounts time series data collected from May 03 – Aug 06 in Zone 1 (Sandy Clay Loam) in 2018.

each experiment using the whole dataset. Experiments were defined as: experiment 1 (inputs: Soil Matric Potential – 3 depths (SMP), management zone (Z), rain (r), irrigation (i) output: irrigation prescription (ip), experiment 2 (inputs: SMP and Z, outut: ip) and experiment 3 (inputs: SMP, output: ip). Each experimental scenario was run 10 times using Keras Deep Learning library on the python programming platform (Kumar, 2018), using data from Zone 1 and Zone 4 to one hour ahead irrigation prescription. mean squared error (MSE) was used as loss function. Moreover, the prediction performance of the LSTM models was studied by calculating the coefficient of determination (R^2) and RMSE (root mean square error).

Results and discussion

The evaluation of LSTM performance for the one-hour-ahead prediction of irrigation prescription across different parameters is summarized in Table 1. Results showed that the experiments were able to predict irrigation prescription with RMSE<1.4 mm and R^2>0.9. The highest R^2 and minimum RMSE were obtained with 60 neurons, dropout of 0.5 and 100 epochs for experiment 1 (RMSE=0.86 mm and R^2=0.975); 40 neurons, dropout of 0 and 100 epochs for experiment 2 (RMSE=0.92 mm and R^2=0.968) and 20 neurons, dropout of 0 and 100 epochs for experiment 3 (RMSE=0.91 mm and R^2=0.972). These three models had nearly the same performance statistics and thus the selection of one of these models over the other should be dependent upon available data. It was also found that experiments conducted with 20 hidden neurons, dropout of 0 and 30 epochs were able to predict accurately irrigation prescriptions. Nevertheless, additional blocks, dropout

Table 1. Prediction performance of LSTM models in training stage.[1]

B	D	E	Experiment 1			Experiment 2			Experiment 3		
			RMSE (mm)	R^2	PT (s)	RMSE (mm)	R^2	PT (s)	RMSE (mm)	R^2	PT (s)
20	0	30	0.91	0.9725	16.9	1.02	0.9721	12.7	0.94	0.9710	8.4
20	0	60	0.87	0.9749	27.6	0.92	0.9559	20.4	0.92	0.9719	16.3
20	*0*	*100*	0.87	0.9751	39.9	0.92	0.9573	31.4	*0.91*	*0.9725*	*27.9*
20	0.5	30	1.15	0.9565	17.6	1.16	0.9650	12.8	1.11	0.9592	10.7
20	0.5	60	1.31	0.9432	29.1	1.14	0.8234	25.0	1.04	0.9647	19.1
20	0.5	100	0.99	0.9676	43.1	1.03	0.8400	34.4	0.93	0.9712	28.5
40	0	30	0.90	0.9735	18.9	0.94	0.9721	16.9	0.93	0.9713	10.2
40	0	60	0.88	0.9745	30.0	0.94	0.9591	23.1	0.93	0.9713	21.1
40	*0*	*100*	0.88	0.9748	40.2	*0.92*	*0.9687*	*38.3*	0.92	0.9721	29.0
40	0.5	30	1.40	0.9353	22.3	1.11	0.9718	15.2	0.98	0.9687	10.8
40	0.5	60	0.93	0.9718	31.5	0.98	0.9267	28.0	0.92	0.9721	19.5
40	0.5	100	0.96	0.9699	47.2	0.93	0.9313	45.5	0.94	0.9709	40.6
60	0	30	0.89	0.9742	24.5	0.94	0.9720	19.5	0.95	0.9703	13.2
60	0	60	0.88	0.9743	44.8	0.95	0.9677	34.0	0.93	0.9713	21.7
60	0	100	0.87	0.9753	67.0	0.92	0.9712	56.4	0.92	0.9719	36.6
60	0.5	30	1.08	0.9612	26.2	0.99	0.9722	20.5	0.95	0.9703	14.0
60	0.5	60	1.06	0.9633	45.4	0.94	0.9514	34.7	0.92	0.9721	24.8
60	*0.5*	*100*	*0.86*	*0.9759*	*64.2*	0.92	0.9611	51.6	0.94	0.9712	45.8

[1] B = blocks; D = dropout; E = epochs; P = processing time; R^2 = coefficient of determination; RMSE = root mean square error.

values or epochs, did not further improve the learning capabilities of the LSTM. The increment in blocks number, maintaining however the number of epochs in the training stage with dropout zero, did not showed important variations in R^2 but changes were found among RMSE values with a maximum difference of 0.021 mm. The same analysis with dropout of 0.5, showed that processing time increased, but with differences in RMSE values between 0.1 to 0.38 mm and between 0.01 to 0.12 in R^2. In conclusion, the increase in the number of blocks not only does not provide any significant improvements in the results, but also increase the number of parameters that need to be optimized into the LSTM, increasing the processing time. Because of that, the number of blocks should be kept as low as possible in the LSTM model. Results for the LSTM models in terms of dropout changes maintaining the same block numbers and epochs, also are visualized in Table 1. For changes in dropout value from 0 to 0.5, 81% of the cases increased in RMSE, 89% decreased in R^2 and 85% increased processing time. This behaviour was different for experiment 1 (60 blocks, 100 epochs) in which an increase from 0.9753 to 0.9759 was obtained in R^2 with also reduction in processing time from 67 s to 64.2 s. Therefore, it is important to keep the dropout value in zero. The increase in the number of epochs, maintaining the same number of blocks and dropout, showed that there was a decrease of RMSE in 93% of the cases and an R^2 increase in 53.3% of the cases. In experiment 2, there is a decrease of R^2 from 30 to 60 epochs and then an increment from 60 to 100 epochs, corresponding to 33.33% of the cases analyzed. The rest 13.33% has a behavior of increase R^2 from 60 to 100 epochs and a decrease from 60 to 100 epochs. According to these results, the number of epochs does not guarantee a better model performance. Finally, as shown in Table 1, the processing time increases with the number of epochs.

According to these results, parameters selected for training the LSTM models were: 20 blocks, 0 dropout and 100 epochs, obtaining RMSE=0.87 mm and R^2=0.975 in Experiment 1, RMSE=0.92 mm and R^2=0.96 in experiment 2 and RMSE=0.91 mm and R^2=0.9725 in experiment 3. The previous parameters were selected because with fewer blocks it is possible to obtain successful results using these models, reducing also the processing times. The model obtained in Experiment 3 was implemented to evaluate the performance in Zone 1 and 4 independently. Based on the results summarized in Table 1, it seems that SMP played a more important role than other inputs in temporal variations for the irrigation prescription predictions. Results of experiment 3 (Table 1), showed that the same level of predictability and error with data from experiment 1 can also be achieved with only SPM data. This can be explained because the information acquired from the SMP sensors at different depths, allowed the estimation of soil moisture changes in accordance with the soil water balance. The LSTM model's performance with respect to the prediction of irrigation prescription was studied using the loss function. The loss function for each selected LSTM model decreased and stabilized around a small value for the training and validation datasets without under-fitting or over-fitting. This further highlights the robustness of the LSTM models, which were able to sufficiently learn the underlying function approximating the dynamic data, making them useful for generating irrigation predictions. The predictive capability of the LSTM models was evidenced by their performance with the test dataset. Figure 4 shows de behavior between original and predicted irrigation prescriptions for Zone 1 and Zone 4. The LSTM implemented were able to accurately model the irrigation prescription dynamics in comparison with the original data for each zone. According with Figure 4, RSME values for Zone 1 were lower than Zone 4. An explanation for this outcome is that a bigger dataset could better train the model to learn differences between inputs and outputs. Despite this, the LSTM model works properly for the purpose of predicting the irrigation prescription in Zone 1 as well for Zone 4.

Conclusions

The usefulness of Long Short-Term Memory neural networks in assisting irrigation prescriptions was evaluated for sandy clay loam and loamy fine sand soil. The LSTM models developed for the

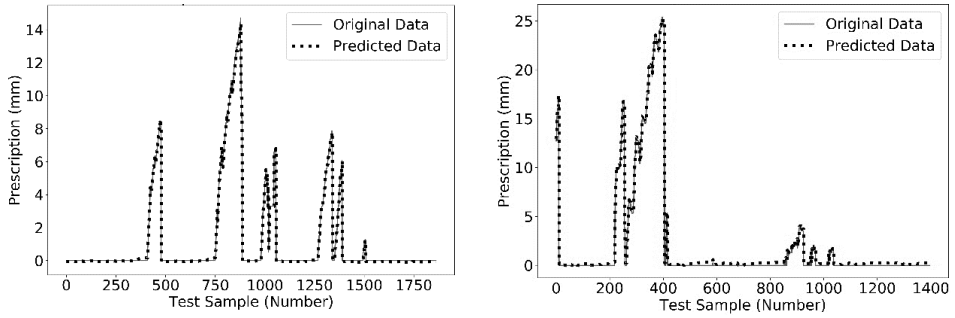


Figure 4. Prescription comparison between test and predicted data in Experiment 3. Left: Zone 1 (sandy clay loam soil). Right: Zone 4 (loamy fine sand soil).

study site achieved coefficients of determination ranging from 0.82 to 0.98 for irrigation prescriptions predicted one-hour-ahead. The LSTM is a powerful strategy to model irrigation prescriptions using data from soil matric potential sensors since the soil water content behavior can be studied effectively as a sequential problem. This ability is demonstrated by the performance obtained in the minimization of the number of categorical attributes from six inputs to three. The number of blocks, dropout and epochs defined, allowed the optimization of the training procedure: obtaining successful results with less processing time. The outcomes obtained demonstrate that LSTM models with SMP data acquired in real time could be positively used, reducing the number of sensors and improving the irrigation management.

Acknowledgements

We want to acknowledge farmer Jim Lewey for his collaboration. Project funding was provided by NRCS, Alabama, Soil and Water Conservation Committee, Alabama Association of Conservation Districts and the Agricultural Experimental Station.

References

Adewale, A., Bamidele, A. and Lateef, U., 2018. Predictive modelling and analysis of academic performance of secondary school students: ANN approach. International Journal of Science and Technology Education Research 9 (1) 1-8.

Adeyemi, O., Grove, I., Peets, S., Domun, Y. and Norton, T., 2018. Dynamic neural network modelling of soil moisture content for predictive irrigation scheduling. Sensors, 18(10) 3408.

Gonzalez, J. and Yu, W., 2018. Non-linear system modeling using LSTM neural networks. IFAC-PapersOnLine 51 (13) 485-489.

Goodfellow, I., Bengio, Y., Courville, A. and Bengio, Y., 2016. Deep learning (Vol. 1). Cambridge, MA, USA: MIT press

Gutzler, C., Helming, K., Balla, D., Dannowski, R., Deumlich, D., Glemnitz, M. *et al.,* 2015. Agricultural land use changes scenario-based sustainability impact assessment for Brandenburg, Germany. Ecological indicators 48 505-517.

Hochreiter, S., 1998. The vanishing gradient problem during learning RNN and problem solutions. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 6 (02) 107-116.

Kumar, N., 2018. Deep learning with applications using python. Chatbots and face, object, and speech recognition with TensorFlow and Keras. Berkeley, CA, USA: Apress.

Levidow, L., Zaccaria, D., Maia, R., Vivas, E., Todorovic, M., Scardigno, A. *et al.,* 2014. Improving water-efficient irrigation: Prospects and difficulties of innovative practices. Agricultural Water Management 146 84-94.

Lu, Z., Chai, L., Ye, Q., and Zhang, T., 2015. Reconstruction of time-series soil moisture from AMSR2 and SMOS data by using recurrent nonlinear autoregressive neural networks. In IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 980-983.

Mashayekhi, P., Ghorbani, S., Mosaddeghi, M., Shirani, H. and Nodoushan, A., 2016. Different scenarios for inverse estimation of soil hydraulic parameters from double-ring infiltrometer data using HYDRUS-2D/3D. International Agrophysics 30(2) 203-210.

NRCS, 2019. Soil Texture Calculator. https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/?cid=nrcs142p2_054167 (last accessed 05/12/2018).

Pandey, V. and Pandey, P., 2010. Spatial and temporal variability of soil moisture. International Journal of Geosciences 1(02) 87

Qi, Y., Huo, Z., Feng, S., Adeloye, A. and Dai, X., 2018. Prediction of consumptive use under different soil moisture content and soil salinity conditions using artificial neural network models. Irrigation and Drainage 67(4) 615-624.

Seidel, S., Werisch, S., Barfus, K., Wagner, M., Schütze, N. *et al.,* 2016. Field evaluation of irrigation scheduling strategies using a mechanistic crop growth model. Irrigation and drainage 65(2) 214-223.

Vellidis, G., Liakos, V., Porter, W., Tucker, M. and Liang, X., 2016. A dynamic variable rate irrigation control system. In 13th International Conference on Precision Agriculture. https://www.ispag.org/proceedings/?action=abstract&id=2066&search=years (last accessed February 2019).

Zhang, J., Zhu, Y., Zhang, X., Ye, M., and Yang, J., 2018. Developing a LSTM based model for predicting water table depth in agricultural areas. Journal of Hydrology 561 918-929.

Using computational optics for agricultural monitoring with an emphasis on irrigation management zones

I. Klapp[1*], O. Brand[1,2], P. Yafin[1,3], S. Papini[1,2], N. Oz[1,3], I. Bahat[1,5], Y. Cohen[1], V. Alchanatis[1] and N. Sochen[2]
[1]*Institute of Agricultural Engineering, Agricultural Research Organization (ARO),Volcani Center, Rishon LeZion, Israel; iftach@volcani.agri.gov.il*
[2]*Department of Mathematics, Tel Aviv University, Tel Aviv 69978, Israel*
[3]*School of Electrical Engineering, Tel Aviv University, Tel Aviv 69978, Israel*
[4]*Department of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel*
[5]*The Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture, The Hebrew University of Jerusalem, The Robert H. Smith Faculty of Agriculture, Food and Environment, Rehovot 76100, Israel*

Abstract

The increasing global water deficit and the constant demand for yield improvement require higher imaging resolution of irrigation management zones (IMZs). Infrared radiometric remote sensing is efficient for IMZ delineation. To provide high spatial resolution, radiometric remote sensing is conducted from a relatively low altitude, raising two constraints: the radiometric camera's price, and the lower altitude that forces lower coverage in a given time. This work presents progress in using methods associated with computational optics to stabilize low-cost thermal cameras and improve the resolution of radiometric cameras. Performance results are promising.

Keywords: radiometry, computational imaging, remote sensing and sensors, super resolution

Introduction

The increasing global water deficit and the constant demand for yield improvement require higher imaging resolution in irrigation management. A field's water status is generally variable in the spatial dimension, and therefore requires variable irrigation rates. To apply variable-rate irrigation, the field has to be divided into irrigation management zones (IMZs). Infrared (IR) radiometric remote sensing is efficient for IMZ delineation. To provide high spatial resolution, the radiometric remote sensing is conducted from a relatively low altitude, resulting in low area coverage and high cost.
The new technology of uncooled microbolometer focal plane array (FPA) is characterized by small size, lower power consumption and a reduced price; these cameras are widely used in military and commercial applications (Bhan *et al.*, 2009) and are the most commonly used cameras in thermography (Tempelhahn *et al.*, 2016). A big advantage of this technology for agriculture is the ability to mount the camera on a microdrone. However, previous work has reported these cameras' sensitivity to environmental conditions which affect their readings (Kusnierek and Korsaeth, 2014). The drift in the FPA due to environmental loads limits the acquisition of informative radiometric images from the thermal camera. Thus, calibrating the camera response to stabilize its readout is an important challenge. The image of FPA cameras is a function of their optics and detector response; the image formation in lexicographic form is:

$$y(T)=\left[G_i(T)\right]\cdot[A]x+d(T)+n \tag{1}$$

where y is the FPA image, [G] is the gain matrix, [A] is the point spread function (PSF) matrix representing the optical blur, x is the object radiation, n is the noise, and (x,d,y,n) are vectors in lexicographic form. The gain and the offset depend upon the FPA temperature and changes in

DOI 10.3920/978-90-8686-888-9_82, © Wageningen Academic Publishers 2019

environmental temperature (Nugent and Shaw, 2014). These rapid changes require constant calibration against a black body (Tempelhahn *et al.*, 2016). While the gain calibration is relatively simple, as it is a linear function of the FPA temperature, the offset calibration is tedious and difficult.

Field coverage and image resolution come one at the expense of the other. SR is a collection of methods in which information redundancy or prior knowledge are exploited to improve the resolution of a given image. Such an improvement, when possible, can eliminate the above tradeoff, enabling higher and faster flight, and thus potentially lowering the cost of the aerial survey. There are a few leading approaches to SR. Elad and Feuer (1997) and Farsiu *et al.* (2004) showed that when the image formation is known, a sequence of non-identical images can be used to obtain an image with improved resolution. Another SR methodology is single-image SR. In this approach, prior knowledge, such as gradient distribution in low resolution and high resolution, is exploited to estimate a high-resolution image from a low-resolution one (Sun *et al.*, 2008). Single-image SR can be tackled as a learning process, i.e. a parametric system is tuned to learn the transformation from low resolution to high resolution. One emerging learning technique is the convolutional neural network (CNN) (Dong *et al.*, 2014; Rivenson *et al.*, 2017). This process is performed while minimizing the parametric for the best fitting low-resolution input image and its matched high-resolution image over typically thousands of couples.

This paper presents computational optics methods to decrease the cost of IR radiometric image acquisition. This effort is divided into two complementary approaches. The first attempts to lower the cost of the camera itself. In the second, super-resolution (SR) algorithms are developed to enable higher acquisition capacity by imaging from higher altitude.

Materials and methods

Data acquisition

Data were acquired in a commercial vineyard located in Mevo-Beitar, Israel. Plot size was approximately 2.49 ha, with co-ordinates at its center of latitude/longitude 31.729531/35.103165, respectively. Measurements were performed through midday on 15 Aug 2017. For IR image acquisition, a FLIR SC655 camera, (FLIR, Wilsonville, Oregon, USA) was mounted on a drone (Self-made Quad-Copter) hovering at heights of 15, 30, 45 and 60 m above the field. A series of thermal images were acquired at a frame rate of (6 fps), and stored as one file (*.SEQ format). The algorithms presented below were fed with a sequence of images acquired by post-processing the SEQ video saved in .jpg format using (FLIR Tools,2018) software.

The single-image SR was performed as a deep learning process, realized in the visible range. Images taken from a drone hovering 15 m above a cotton field was used as a case study. The images were taken with a Sony alpha 5100 camera (Sony, Tokyo, Japan) and saved in .jpg format. All images were taken in nadir orientation.

First approach: stabilizing the readout of an uncooled FPA

Method 1: FPA offset unknown

The goal here was to restore radiometric information from an uncalibrated camera with unknown offset. In previous work (Klapp *et al.*, 2017), it was shown that using pairs of images of the same object at different blur levels provided the information required to estimate the unknown object and drift when the gain is known. Thus two sequential images with two different PSFs (optical blur) were taken. Since the camera integration time is much faster than changes in thermal load, assuming that the gain and offset do not change and therefore, image formation of the images is:

$$\begin{aligned} y_1(T) &= \lfloor G_1(T) \rfloor \cdot [A_1]x + d_1(T) + n_1 \\ y_2(T) &= \lfloor G_2(T) \rfloor \cdot [A_2]x + d_2(T) + n_2 \\ S.T.: \quad & \lfloor G_1(T) \rfloor = \lfloor G_2(T) \rfloor, d_1(T) = d_2(T), [A_1] \neq [A_2] \end{aligned} \tag{1}$$

The use of two blur levels creates the required difference between the equations. The unknown offset was estimated, by substituting one equation into another. This scheme is beneficial because gain calibration is much easier to perform than offset calibration. The estimation for the offset is (Klapp *et al.*, 2017):

$$\tilde{d}(T) = inv\left\{[G][A_1]inv\left\{[G][A_2]\right\} - [I]\right\}\left([G][A_1]inv\left\{[G][A_2]\right\} y_2 - y_1\right) \tag{2}$$

Solving the inverse problem of the image formation by substituting Equation (2) into the image formation equation Equation (1) such that get:

$$\tilde{x} = inv\left\{[G_2][A_2]\right\} \cdot \left(y_2 - \tilde{d}(T)\right) \tag{3}$$

For a detailed account of the proposed algorithm, see Klapp *et al.* (2017).

Method 2: FPA offset and gain unknown

The goal here was to restore radiometric information from an uncalibrated camera where both offset and gain were unknown. Thus, Method 1 was extended to the general case where the gain [G] and drift (*d*) are unknown. For an image with *M* pixels, Equation (1) is a system of *M* linear equations, each of which has unknown object value, offset value and gain value; thus, the two equations in Equation 1 are not sufficient to obtain a solution.

To have a sufficient number of equations, at least one additional sharp and blurred pair of a different object x* with different image y* (Papini *et al.*, 2018) should be added, such that:

$$\begin{aligned} &[G][A_1]inv\left\{[G][A_2]\right\} d - d = [G][A_1]inv\left\{[G][A_2]\right\} y_2 - y_1 \\ &[G][A_1]inv\left\{[G][A_2]\right\} d - d = [G][A_1]inv\left\{[G][A_2]\right\} y_2 * - y_1 * \end{aligned} \tag{4}$$

Solving this system poses two additional obstacles. The first is the additive noise (*n*), which randomly changes from one image to the next. To overcome this, (*N*) number of pairs were used.

In addition, since the unknown gain [G] and unknown object multiply one another, their solution is determined up to a scale (Papini *et al.*, 2018). A unique solution is determined by using prior knowledge of solution behavior (Papini *et al.*, 2018). The solution of the system is obtained as the minimization of a cost function:

$$S\left([\mathrm{G}],\mathrm{d},\{x_i\}\right) = \sum_{i=1}^{N}\sum_{j=1}^{2}\left\| y_{i,j} - [G]\cdot\left(\left[A_j\right]x_i\right) - d\right\|_2^2 \tag{5}$$

Since [G],d,{x_i} are not known, to solve Equation (5), the following algorithm was suggested: (1) perform two-point correction and bad pixel removal (once); (2) assign the sharp value of (y_i) for each pair as the initial guess for (x_i); (3) solve the resulting regression connection in Equation (5) between [G] and (d); (4) assign the resulting [G], (d) and {y_i} back into Equation (5) and solve it using a least square QR algorithm (LSQR, 2018). For a detailed account of the proposed algorithm, see Papini *et al.* (2018).

Second approach: improving imaging capacity by means of SR

Method 1: N-frame SR

A typical remote-sensing survey was performed with the camera hanging at nadir toward the ground. The natural results of this dynamic are a sequence of (*N*) frames of the same resolution. Since typical frame rate, flight velocity and flight height result in a significant overlap in field of view between sequential frames, under the assumption that the overlapping area is non-perfectly

identical, the difference may be used to estimate SR by minimizing the following cost function (Elad and Feuer, 1997):

$$E(x_{hr}) = \arg\min\left[\sum_{k=1}^{N}\left\|D_k H_k F_k x_{hr} - y_k\right\|_2^2\right] \rightarrow x_{hr} \tag{6}$$

where k is the frame number, x_{hr} is the estimated high-resolution object, y_k is the ($_k$) low-resolution frame, D_k is the 'decimation' matrix associated with the scaling factor between the low-resolution frames and the estimated high-resolution object, and H_k is the PSF matrix. Typically, it is assumed that all frames are sharp with equal minimal blur. F_k is the 'shift' matrix describing the transformation between the k^{th} frame and a reference frame. Correct estimation of x_{hr} relies on an accurate estimation of the models. The most problematic factor to estimate is the relative 'shift' between frames. By a suggested new method for estimating this shift accurately enough by filtering the results of the ORB descriptor (Rublee, 2011) allow the required accuracy.

Method 2: toward transfer learning single-image SR by deep learning

Following the above, a method architecture initially developed for SR in microscopy (Rivenson *et al.*, 2017) was used for agricultural scenery. First, the system was tuned to agricultural features in the visible range, resulting in neural network SR (CNNSR). For the training set, COCO dataset was used (Lin *et al.*, 2018). The training images were divided into 180×180 (pixels) and downscaled by a bicubic kernel to 60×60 (pixels) images. Data set of 90,000 images was used for training, divided into training batches of 64 pairs of low- and high-resolution images. 20 epochs were performed, which is roughly 100,000 iterations. Next, the system was tuned to the IR regime.

Results

First approach: stabilizing the readout of an uncooled FPA

First method: In this section, simulation results are presented under the assumption that thermal images were taken consecutively. The first image is sharp while the second image is blurred with a defocus level equal to one wavelength of 10 μm. The temperature range was 24.0 to 54.5 °C, spread over 249 gray levels (gl) such that sensitivity was 0.122 °C/gl. Following Equation (1), the blurred image was subjected to gain and offset variation. A white Gaussian noise with signal-to-noise ratio of 1000 was added to the images. Following model 1, known gain, the unknown offset and object were assumed. Both offset and gain are quadratic such that:

$$\begin{aligned} G(r) &= \frac{1}{255^2}\cdot\left[\left(x-0.3*255\right)^2+\left(y-0.7*255\right)^2\right]+0.3, \\ d(r) &= \frac{1}{0.4*255^4}\left[\left(x-0.5*255\right)^4+\left(y-0.4*255\right)^4\right]-0.3, \qquad r=\sqrt{x^2+y^2} \end{aligned} \tag{7}$$

Simulation results are presented in Figure 1. The pair of sharp and blurred images (Figure 1A and B, respectively) were subjected to the same gain and offset function.

The original ground truth is presented in Figure 1C and the recovered image is presented in Figure 1D. Comparing the recovered image to the original one, the difference is seamless with root mean square error (RMSE) of 0.27 gl equal to 0.033 °C.

Second method: In this simulation, N=13 images of the field were used simultaneously to restore object radiometric values when gain, offset and the object values are unknown. All images were subjected to blur levels as in the previous data; without loss of generality, the unknown gain and offset functions were slightly different, such that:

$$\begin{aligned} G(r) &= \left[\left(x-0.3\cdot W\right)^2+\left(x-0.7\cdot H\right)^2\right]/\left[1.942\cdot\left(W\cdot H+0.3\right)\right] \\ d(r) &= 6.25\cdot\left(\sigma_x\cdot\left[\left(\left(x-0.5\cdot W\right)^4+\left(y-0.4\cdot H\right)^4\right)\right]/\left[\left(0.4\cdot H^2\cdot W^2\right)-0.3\right]+4.6411\right) \end{aligned} \tag{8}$$

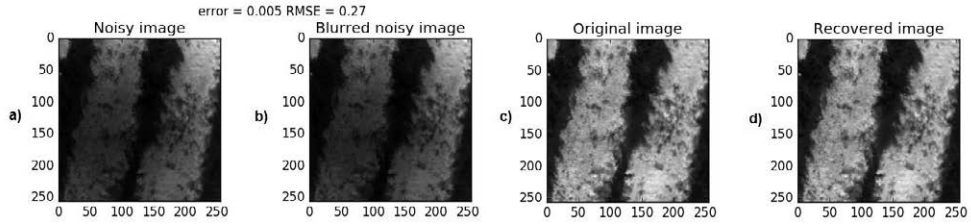


Figure 1. Recovery of radiometric data from drifted information with known gain and unknown offset.

Restoration results are presented in Figure 2. Restoration of gain offset and the 13 objects was performed jointly. The 13 distracted pairs of images are given one below the other in rows 1 and 2, sharp and blurred, respectively. The restored object is presented in row 3 and the ground-truth object is presented in row 4. Comparing the restoration to the ground truth, RMSE was in the range of 1.3 to 1.44 °C, which is acceptable accuracy for practical uses.

Second approach: improving imaging capacity by means of SR

First method: Following the *N* frames SR, four successive images were used to create a two-time-folded (in two directions) SR image. In Figure 3, one of the original images (Figure 3A) and the resultant SR image (Figure 3B) are presented side by side. Zoomed in portion of the images are presented in Figure 3C. Comparing the recovered SR image to the original, the image is clearly resolved in finer detail.

Second method: Following the above, a three-time-folded (in two directions) SR process in the visible range, performed by CNNSR is presented in Figure 4. The input for this demonstration was a cotton bush in the visible range. To test the model, the high-resolution image was downscaled by a factor of 3. The original image is presented in Figure 4A. The resultant low-resolution image is presented in Figure 4B. The output of the CNNSR is the given low-resolution input upscaled by a factor of 3. For the sake of detailed observation, a single leaf of each image was zoomed in. The detailed images are shown in Figure 4A, B and C for high resolution, low resolution and SR, respectively. The RMSE of the SR image relative to the high-resolution image is 11.1 gl. In comparison, the RMSE of a bicubic interpolation of the low-resolution relative to the high-resolution image is only 14.18 gl.

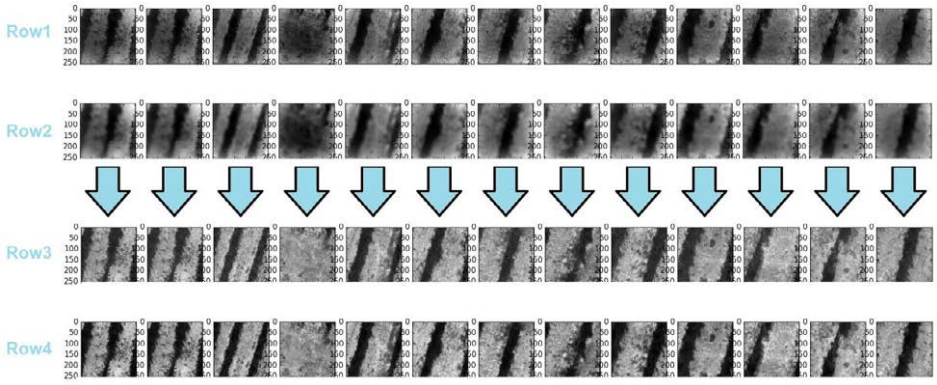


Figure 2. Recovery of radiometric data from drifted information with unknown gain and offset

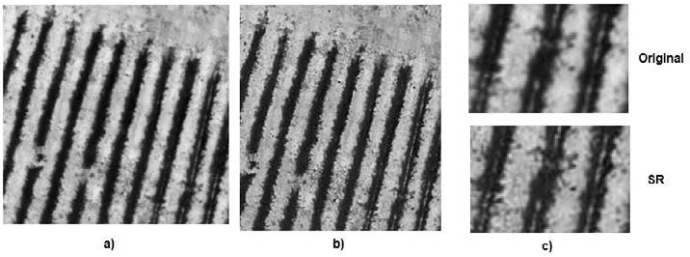


Figure 3. Recovering SR image from a series of low-resolution radiometric images.

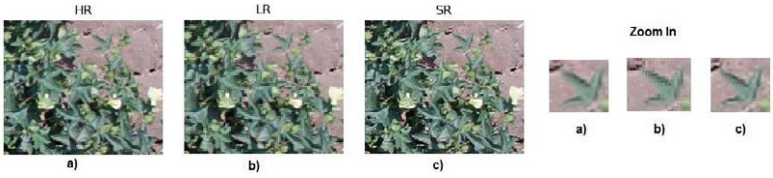


Figure 4. Recovery of SR image from a series of low-resolution radiometric images.

Discussion

The high price and low spatial resolution of radiometric imagery are key technological barriers to extensive implementation of IR remote sensing in IMZ delineation. Progress in these two directions by means of computational optics methods was presented, It was shown that redundant information, which is relatively easy to obtain, can help restore radiometric information from uncalibrated thermal data associated with low-cost thermal cameras and increase the resolution of the IR images which are limited to relatively low resolution. With respect to resolving drift in camera calibration, in principle, when both gain and offset are unknown, an RMSE of ~1.4 °C can be obtained, and an RMSE of only a fraction of a degree Celsius is obtained when the gain is also known. With respect to spatial resolution, SR was shown to increase spatial resolution by a magnitude of 2 to 3 for agricultural data. The accuracy of the data shown in one case resulted in RMSE of ~11.4 gl, which is 4.7% of the full-scale value, in comparison to 5.6% for the full-scale interpolation.

Conclusions

Cheaper radiometric imagery is essential for extensive IMZ. Results showed that systems composed of SR and low resolution uncooled radiometric cameras with two level focusing mechanism allowed performance which may replace an expensive radiometric camera.

Acknowledgements

The authors would like to thank the Chief scientist of the Israeli Ministry of Agriculture (Kandel Program) for funding this research under grant no. 20-12-0030.

References

Bhan, R. K., Saxena, R. S., Jalwania, C. R. and Lomash, S. K., 2009. Uncooled Infrared Microbolometer Arrays and their Characterisation Techniques, Defence Science Journal 59 (6), 580-589.

Dong.C, Loy.C.C, He.K, and Tang.X., 2014. Learning a Deep Convolutional Network for Image Super-Resolution. In European Conference on Computer Vision. Cham, Switzerland: Springer. pp 184-199.

Elad.M and Feuer.A., 1997. Restoration of a Single Super resolution Image from Several Blurred, Noisy, and Undersampled Measured Images, IEEE Transactions on Image Processing 6, 1646-1658.

Farsiu.S, Robinson.M.D, Elad.M. and Milanfar.P., 2004. Fast and Robust Multiframe Super Resolution 2004 IEEE Transactions on Image Processing 13, 1327-1343.

FLIR tools, https://www.flir.com/products/flir-tools (last accessed 14 Jan 2019).

Klapp, I., Papini, S. and Sochen, N., 2017. Radiometric Imaging by Double Exposure and Gain Calibration, Applied Optics 56, 5639-5647.

Kusnierek, K. and Korsaeth, A., 2014. Challenges in Using an Analog Uncooled Microbolometer Thermal Camera to Measure Crop Temperature, International Journal of Agricultural and Biological Engineering 7, 60-74.

Lin, T-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R. J., Hays, J. *et al.,* 2018. Microsoft COCO: Common Objects in Context, https://arxiv.org/abs/1405.0312 (last accessed 04 Dec 2018).

LSQR, Sparse Equations and Least Squares http://web.stanford.edu/group/SOL/software/lsqr (last accessed 04 Dec 2018).

Nugent, P.W. and Shaw, J.A., 2014. Calibration of Uncooled LWIR Microbolometer Imagers to Enable Long-Term Field Deployment. In Proceedings of the SPIE, Vol. 9071 90710V, pp. 1-10.

Papini, S., Yafin, P., Klapp, I. and Sochen, N., 2018. Joint Estimation of Unknown Radiometric Data, Gain and Offset from Thermal Images. Applied Optics 57 (36) 10390-10401

Rivenson,Y., Göröcs, Z., Günaydin, H., Zhang,Y., Wang, H. and Ozcan, A., 2017. Deep Learning Microscopy. Optica 4 (11), 1437-1443.

Rublee, E., Rabaud, V., Konolige, K., Bradski, G., 2011. ORB: An efficient alternative to SIFT or SURF. IEEE, International Conference on Computer Vision. DOI: 10.1109/ICCV.2011.6126544

Sun, J., Xu, Z. and Yeung, H., 2008. Shum Image Super-Resolution Using Gradient Profile Prior. IEEE, Conference on Computer Vision and Pattern Recognition.

Tempelhahn, A., Budzier, H., Krause, V. and Gerlach, G., 2016. Shutter-Less Calibration of Uncooled Infrared Cameras. Journal of Sensors and Sensor Systems 5, 916.

A variable rate irrigation decision support system for corn in the US eastern coastal plain

K. Stone[1*], P. Bauer[1], S. O'Shaughnessy[2], M. Andrade[2] and S. Evett[2]
[1]*USDA-ARS, Coastal Plains Soil, Water, and Plant Research Center, Florence, SC, USA; ken.stone@ars.usda.gov*
[2]*USDA-ARS, Conservation & Production Research Laboratory, Bushland, TX, USA*

Abstract

Variable rate irrigation (VRI) systems can apply different water depths both in the direction of travel and along the irrigation system length. In this research, a VRI system was evaluated using the USDA-ARS Irrigation Scheduling and Supervisory Control and Data Acquisition System (ISSCADA) to spatially manage corn irrigation in the US Eastern Coastal Plain. The ISSCADA system utilizes remote sensing of plant canopy temperature, soil moisture and microclimate to schedule VRI irrigations. Utilizing these measurements, it calculates an integrated crop water stress index and provides a spatial irrigation recommendation. Results from 3-years of field evaluation are discussed.

Keywords: variable rate irrigation, decision support system, crop water stress

Introduction

Variable rate irrigation (VRI) systems are irrigation systems capable of applying different water depths both in the direction of travel and along the length of the irrigation system (Evans *et al.*, 2010). Thus, VRI systems can be tools for conserving water and spatially allocating limited water resources while potentially increasing profits (Evans and King, 2012). Spatially variable water applications with VRI systems attempt to overcome site-specific problems that include spatial variability in topography, soil type, soil water availability, and other landscape features, but may also be used in response to site-specific crop water requirements. These VRI systems could also be used as a tool for improving crop water management and efficiency by delivering water to plants where needed, when the crop demands it, and in the appropriate amounts (O'Shaughnessy *et al.*, 2015). Although spatially variable water application technology is available and has high interest among growers, there has been limited adoption of VRI systems (Evans *et al.*, 2013). One potential reason for this limited adoption of VRI systems is the lack of science-based information on how to precision-apply water with these systems (Sadler *et al.*, 2005). Using VRI, spatial water application can be either static or dynamic during the growing season based on soil and crop monitoring. Static water application maps are usually based on soil survey maps, aerial photographs, producer knowledge of the fields, and on soil electrical conductivity (EC) field maps together with historic yield maps (Lund *et al.*, 2001, Perry *et al.*, 2007). Static water application maps are problematic in that spatial variability of crop water stress throughout the growing season is often overlooked. This spatio-temporal crop variability can be a consequence of spatially variable precipitation, topography, soil infiltration rates, soil water storage capacity (Sadler *et al.*, 2005) as well as other factors such as variable plant stand, or disease and pest infestation (Falkenberg *et al.*, 2007). For dynamic spatial water applications, Sadler *et al.* (2005) and Evans *et al.* (2013) identified critical research needs that included the development of decision support systems and integrated management systems to sense within-field variability in real time and dynamically define irrigation management zones. Dynamic management zones for VRI system management can be estimated using remote sensing methods including crop canopy temperatures.

DOI 10.3920/978-90-8686-888-9_83, © Wageningen Academic Publishers 2019

Canopy temperature measurements are widely used as indicators of crop water stress. An early study by Idso *et al.* (1981) used remotely sensed canopy temperature to calculate a foliage-air temperature differential that was related to crop water stress. Jackson *et al.* (1981) also reported on the use of remotely sensed crop canopy temperatures and related them to a crop water stress index (CWSI). Canopy temperatures have also been used to measure spatial crop responses to irrigation. Sadler *et al.* (2002) mounted infrared thermometers (IRT) onto a center pivot irrigation system to measure spatial variation in water stress for a corn crop with four irrigation treatments and were able to determine significant differences among the irrigation treatments as well as within and among the soil mapping units. In the Texas High Plains, Peters and Evett (2004a) developed a similar system to measure spatial soybean canopy temperatures; and they automated their center pivot irrigation system using the temperature-time-threshold (TTT) irrigation scheduling method based on a method of modeling diurnal canopy temperatures based on one-time-of day measurements (Peters and Evett, 2004a, 2004b, 2008). Later, O'Shaughnessy and Evett (2010) developed a wireless sensor network for monitoring crop canopy temperatures using a moving irrigation system. O'Shaughnessy *et al.* (2013) combined the CWSI with the method of Peters and Evett (2004b) into an integrated CWSI (iCWSI) that better represented crop water stress than did a single time of day CWSI. This iCWSI approach was then used in automating irrigations and developing dynamic prescription maps for VRI management using the ARS Irrigation Scheduling and Supervisory Control and Data Acquisition System (ISSCADA) (Andrade *et al.*, 2015, O'Shaughnessy *et al.*, 2013).
These studies have demonstrated the effectiveness of canopy temperature measurement in observing spatial crop response to water stress and in automating VRI systems. These VRI automation methods need to be evaluated in different environments, particularly more humid environments where interactions with rainfall make irrigation scheduling a more complex process. In this research, the ARS Irrigation Scheduling and Supervisory Control and Data Acquisition System (ISSCADA) was evaluated for potential use in managing variable rate irrigation of corn in the humid Southeastern US and compared it to a method used in the region based on soil water potential measurements.

Methods and materials

The experiment was conducted from 2016-2018 under a variable rate center pivot irrigation system located at the USDA-ARS Coastal Plains Soil, Water, and Plant Research Center in Florence, South Carolina (34° 14' 43.48', 79° 48' 26.44'). The experiment consisted of three treatments: (1) Irrigation management using the ARS Irrigation Scheduling and Supervisory Control and Data Acquisition System (ISSCADA) system; (2) Irrigation management based on maintaining soil water potentials (SWP) in the rooting zone <-30 kPa (~50% depletion of plant available water); and (3) Non-irrigated. Treatment plots were established under the VRI system in a randomized complete block plot design with eight replicates. The pivot area was divided into eight 45° segments with a complete block containing each of the three treatments in each segment. The plot width along the VRI lateral was approximately 15 m wide. To reduce impacts from overlap between management zones, two zones were combined into one management zone approximately 30 m wide. If multiple soil types occur within a plot, irrigation management using the SWP management was based on the plots' dominant soil type while irrigation management with the ISSCADA system used an average of the entire plots' canopy temperature.

Irrigation management

The ISSCADA system integrated soil water, plant and weather data together to produce prescription map recommendations for managing variable-rate irrigation systems (Andrade *et al.*, 2015, O'Shaughnessy *et al.*, 2013, O'Shaughnessy *et al.*, 2016). The ISSCADA system calculated an integrated crop water stress index (iCWSI) to estimate when to initiate irrigation and the soil water depletion to determine the amount of irrigation. Prior to canopy closure when the ISSCADA IRTs

were reading both ground and vegetation temperatures, the ISSCADA treatments irrigations were based on measured soil water depletion. After canopy closure and the IRTs were only reading vegetation, irrigations for the ISSCADA system were initiated based on IRT measurements and calculated iCWSI. A schematic of the ISSCADA process for determining a spatial irrigation application is shown in Figure 1. For the SWP treatment, irrigation was initiated when the SWP at the 0.3 m depth decreased below -30 kPa. When this condition was met, the SWP plots received an irrigation application of 12.5 mm. The non-irrigated plots did not receive irrigation applications.

Measurements

In the SWP treatment plots, soil water potential sensors (Watermark, Irrometer, Riverside, CA, USA) were installed at 0.3- and 0.6-m depths and recorded at least three times per week. In the ISSCADA plots, soil water content readings using Acclima TDR-315 sensors (Acclima, Inc., Meridian, ID, USA) were measured at four depths (0.15, 0.3, 0.45, 0.6 m). The soil water contents were recorded on an hourly basis coupling the TDR-315 sensors with a Campbell Scientific CR206X data logger (Campbell Scientific, Inc., Logan, UT, USA). Crop temperature was recorded using Dynamax wireless infrared thermometers (IRT) (Dynamax Inc, Houston, TX, USA) mounted on the VRI pivot at each management zone border with opposing, oblique views of the management zone (two IRTs for each management zone). IRT measurements were recorded at 1-minute intervals and collected at the ISSCADA controller. At defined intervals, the VRI system was rotated without irrigating over the treatment plots to record the crop temperatures. These crop temperatures were then used for calculating iCSWI and managing irrigation applications in ISSCADA plots. In 2018, the ISSCADA treatment plots were divided into either an ISSCADA plot (based on crop temperature only) and a hybrid treatment (based on crop temperature after soil water was depleted below a set threshold). At harvest, grain yield in each plot/subplot was measured from equal areas near the center of each plot. An on-site weather station recorded rainfall and provided environmental parameters

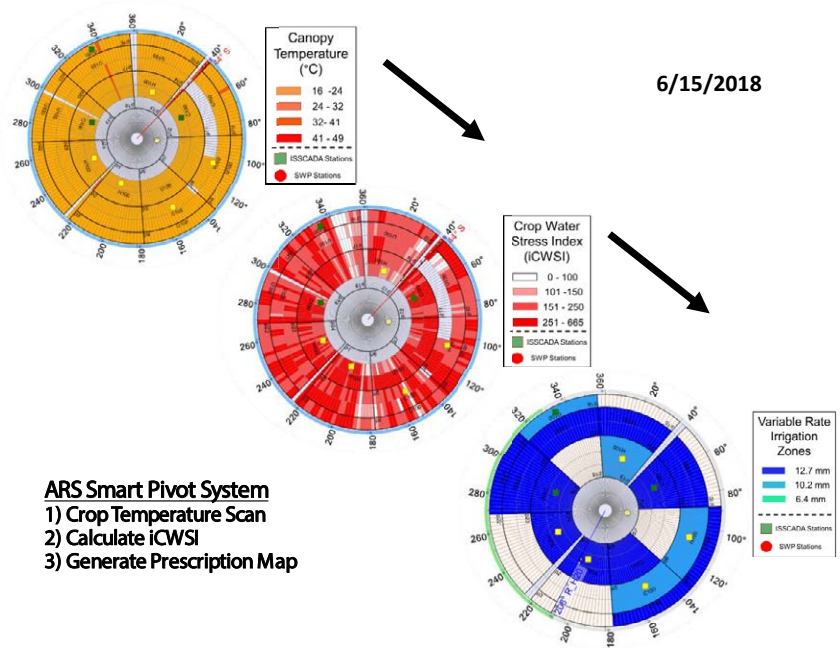


Figure 1. Schematic of the ISSCADA flow process for determining a spatial irrigation prescription.

used to calculate daily reference evapotranspiration (ETo) using the ASCE standardized equation (Allen *et al.*, 2005). The corn nutrient and weed management was based on Clemson University Extension recommendations. The field was planted in circles to facilitate the delineation of irrigation management zones and plots.
The corn yield and total water applied (rainfall + irrigation) were analyzed as a randomized complete block design using analysis of variance. Blocks and irrigation treatment were fixed effects. Means were separated with standard mean separation techniques using the Statistical Analysis System (SAS; SAS Institute Inc., Cary, NC, USA).

Results

In 2016 and 2017, rainfall was higher than normal during the growing seasons. The seasonal rainfall was 541 and 615 mm in 2016 and 2017 (Table 1) and was well distributed through the tasseling and throughout grain fill (Figure 2). In 2018, the seasonal rainfall was 532 mm and was not as well distributed during the growing season (Figure 2) resulting in the need for greater irrigation

Table 1. Rainfall, irrigation, and corn yields for the 2016-2018 growing seasons.

Year	Treatment	Rainfall (mm)	Irrigation (mm)	Mean yield (mg/ha)
2016	SWP	541	38	9.7 A
	ISSCADA	541	25	9.4 A
	Rainfed	541	-	8.7 B
2017	SWP	615	47	10.8 A
	ISSCADA	615	45	10.2 A
	Rainfed	615	-	10.6 A
2018	SWP	532	118	9.9 A
	ISSCADA	532	126	10.4 A
	Hybrid	532	113	10.7 A
	Rainfed	532	-	7.7 B

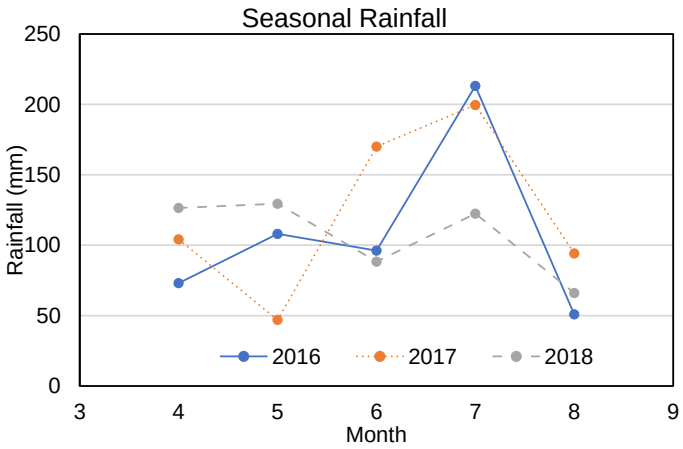


Figure 2. Seasonal rainfall for the 2016, 2017, and 2018 corn growing seasons.

applications. In both 2016 and 2017, the soil water potentials were generally below the threshold of -30 kPa. The SWP treatments required 38 mm of irrigation in 2016 and 47 mm of irrigation in 2017. The ISSCADA treatment in 2016 and 2017 required 25 mm and 45 mm of irrigation, respectively. In 2018, rainfall was considerably lower than normal from tasseling through grain fill (Figure 2). The 2018 mean SWP-based irrigation application was 118 mm. The ISSCADA treatment required 126 mm of irrigation and the hybrid-ISSCADA treatment required 113 mm of irrigation (Table 1). The irrigation depths for the hybrid-ISSCADA treatment was lower than the ISSCADA only treatment because the soil water depletion had not reached a critical threshold needed for initiating an irrigation event.

The corn grain yields were not significantly different for the irrigation treatments or from the rainfed treatment in either 2016 or 2017 (Table 1). In 2018, the irrigation treatment corn yields were significantly greater than the rainfed treatment. The irrigation treatments ranged from 9.9 to 10.7 mg/ha and the rainfed treatment was 7.7 mg/ha. The rainfed treatment in 2018 was considerably lower than rainfed yields in either 2016 or 2017. The ISSCADA-hybrid treatment had greater yield than the ISSCADA only treatment while requiring less irrigation. The use of both canopy temperatures with soil moisture feedback improved the water management. Figure 3 and 4 show a progression of canopy temperature/iCWSI/irrigation prescription maps for the 2018 growing season from June 15 – July 16. During this time period, there was little rainfall and irrigations were required frequently. These figures show the progression of increasing canopy temperatures and corresponding increases in the calculated iCSWI readings.

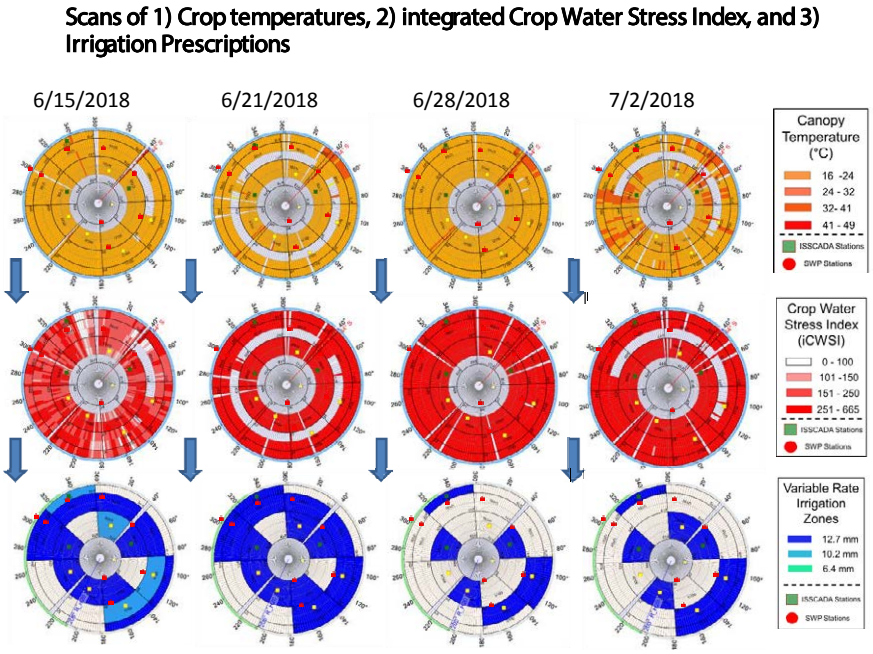


Figure 3. Canopy temperatures, integrated crop water stress indices, and irrigation prescription maps for the variable rate irrigation system from June 15 to July 2, 2018.

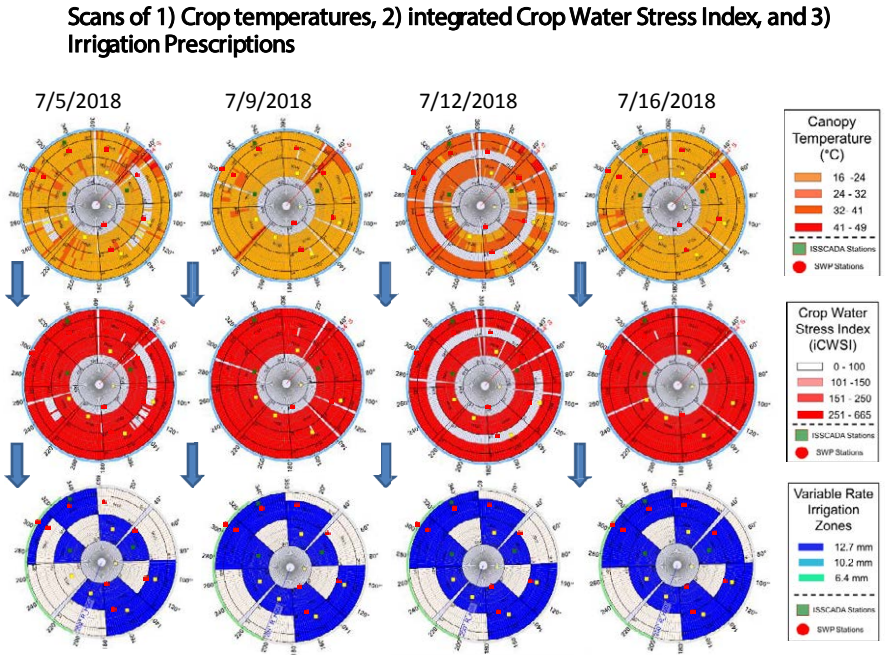


Figure 4. Canopy temperatures, integrated crop water stress indices, and irrigation prescription maps for the variable rate irrigation system from July 5-16, 2018.

Conclusions

In both the 2016 and 2017 irrigation seasons, there were no statistical differences in corn yields across the treatments due to the well distributed rainfall distribution. In 2018, rainfall was below average during the growing season and required considerably more irrigation than in the previous years. The 2018 corn yields for the irrigation treatments were significantly greater than the rainfed yield but were not different from each other. Results of this and ongoing research into the use of a decision support system that incorporates crop stress and available soil moisture will improve and optimize the utilization of water for crop production.

References

Allen, R.G., Walter, I.A., Elliott, R.L., Howell, T.A., Itenfisu, D., Jensen, M.E. *et al.,* 2005. The ASCE Standardized Reference Evapotranspiration Equation. American Society of Civil Engineers – Technology & Engineering, 216 pages.

Andrade, M.A., O'Shaughnessy, S.A. and Evett, S.R., 2015. ARSmartPivot V.1-Sensor Based Management Software for Center Pivot Irrigation Systems. Paper No: 152188736, St. Joseph, MI, USA: ASABE.

Evans, R.G. and King, B.A., 2012. Site-Specific Sprinkler Irrigation in a Water-Limited Future. Transactions of the ASABE 55:493-504.

Evans, R.G., Iversen, W.M., Stevens, W.B. and Jabro, J.D., 2010. Development of Combined Site-Specific Mesa and Lepa Methods on a Linear Move Sprinkler Irrigation System. Applied Engineering in Agriculture 26:883-895.

Evans, R.G., LaRue, J., Stone, K.C. and King, B.A., 2013. Adoption of Site-Specific Variable Rate Sprinkler Irrigation Systems. Irrigation Science 31:871-887.

Falkenberg, N.R., Piccinni, G., Cothren, J.T., Leskovar, D.I. and Rush, C.M., 2007. Remote Sensing of Biotic and Abiotic Stress for Irrigation Management of Cotton. Agricultural Water Management 87:23-31.

Idso, S.B., Jackson, R.D., Pinter Jr, P.J., Reginato, R.J. and Hatfield, J.L., 1981. Normalizing the Stress-Degree-Day Parameter for Environmental Variability. Agricultural Meteorology 24:45-55.

Jackson, R.D., Idso, S.B., Reginato, R.J. and Pinter Jr, P.J., 1981. Canopy Temperature as a Crop Water Stress Indicator. Water Resources Research 17:1133-1138.

Lund, E.D., Wolcott, M.C. and Hanson, G.P., 2001. Applying Nitrogen Site-Specifically Using Soil Electrical Conductivity Maps and Precision Agriculture Technology. The Scientific World Journal 1:767-776.

O'Shaughnessy, S.A., Evett, S.R., Andrade, M.A., Workneh, F., Price, J.A. and Rush, C.M., 2016. Site-Specific Variable-Rate Irrigation as a Means to Enhance water Use Efficiency. Transactions of the ASABE 59(1):239-249.

O'Shaughnessy, S.A. and Evett, S.R., 2010. Developing Wireless Sensor Networks for Monitoring Crop Canopy Temperature Using a Moving Sprinkler System as a Platform. Applied Engineering in Agriculture 26:331-341.

O'Shaughnessy, S.A., Evett, S.R. and Colaizzi, P.D., 2015. Dynamic Prescription Maps for Site-Specific Variable Rate Irrigation of Cotton. Agricultural Water Management 159:123-138.

O'Shaughnessy, S.A., Evett, S.R., Colaizzi, P.D. and Howell, T.A., 2013. Wireless Sensor Network Effectively Controls Center Pivot Irrigation of Sorghum. Applied Engineering in Agriculture 29:853-864.

Perry, C., Vellidis, G., Page, W., Milton, A. and Sullivan, D., 2007. Use of Veris Soil Ec Sensor for Mapping Soil Texture in Georgia Cotton Fields. Beltwide Cotton Conference, p 895, Cordova, TN, USA: National Cotton Council,.

Peters, R.T. and Evett, S.R., 2004a. Complete center pivot automation using the temperature-time threshold method of automatic irrigation scheduling. Paper No: 042196, St. Joseph, MI, USA: ASABE.

Peters, R.T. and Evett, S.R., 2004b. Modeling diurnal canopy temperature dynamics using one-time-of-day measurements and a reference temperature curve. Agrononomy Journal 96(6):1553-1561.

Peters, R.T. and Evett, S.R., 2008. Automation of a Center Pivot Using the Temperature-Time-Threshold Method of Irrigation Scheduling. Journal of Irrigation and Drainage Engineering 134:286-291.

Sadler, E.J., Camp, C.R., Evans, D.E. and Millen, J.A., 2002. Corn Canopy Temperatures Measured with a Moving Infrared Thermometer Array. Transactions of the ASAE 45:581-591.

Sadler, E.J., Evans, R.G., Stone, K.C. and Camp, C.R., 2005. Opportunities for Conservation with Precision Irrigation. Journal of Soil and Water Conservation 60:371-379.

Use of soil electrical conductivity mapping in variable rate irrigation

R. Sui
USDA-ARS Crop Production Systems Research Unit, Stoneville, Mississippi, USA; ruixiu@ars.usda.gov

Abstract

Soil electrical conductivity (EC) was used to generate variable rate irrigation (VRI) prescriptions. Two 7-ha fields under a centre pivot VRI system were selected for the research. Soil apparent EC of the fields was mapped. VRI prescriptions were created based on the soil EC, and irrigation water was delivered to corn and soybean crops according to the VRI prescriptions. Crop yield and irrigation water productivity in the VRI treatments were determined and compared to those in the uniform rate irrigation (URI). Results indicated that there was no significant difference in crop yields between VRI and URI treatments. The VRI treatments reduced irrigation water by up to 30% for two years. Results demonstrated that soil EC maps could be used to establish VRI management zones and VRI could improve irrigation water use efficiency.

Keywords: variable rate irrigation, soil electrical conductivity, water management

Introduction

Irrigation plays a critical role in crop production. Irrigated crops produced higher and more stable yields than dryland crops. Irrigated agriculture in the US is a major consumer of freshwater, accounting for 80% of the nation's consumption of water (Schaible and Aillery, 2015). Limited water resources are becoming an increasing constrain in agriculture. To meet global demands for food and fiber while maintaining agricultural production sustainability, crop water use efficiency must be increased.

VRI technologies allow the producers to site-specifically apply irrigation water at variable rates within the field to adjust for the temporal and spatial variability in soil and plant characteristics. Adoption of VRI has the potential to increase water use efficiency.

Soil physical properties in the Mississippi Delta (USA) region can vary significantly within a single field from excessively drained loamy sands to the poorly drained clays (Cox *et al.*, 2006, Thomasson *et al.*, 2001). The variability in soil properties results in differing water storage capabilities and amounts of water available to the crop, contributing to spatial variability of crop growth, and creating challenges in crop water management. Due to within-field soil variability, plants in one location may need more water than plants in another location in the field. Treating the plants differently based on their needs is necessary for optimizing water use efficiency.

Many research results have demonstrated that soil electrical conductivity (EC) was related to soil properties and crop yield potential, and it could be used for site-specific management in precision agriculture (Johnson *et al.*, 2003, Corwin and Lesch, 2003). Veris 3100 soil EC system (Veris Technologies, Salina, Kansas, USA) is one of the devices commercially available and widely adopted for mapping soil EC. This system measures soil EC using the direct contact method. With a GPS receiver, Veris 3100 system is able to map soil EC at two depths, 0-250 mm and 0-750 mm, simultaneously, described as shallow EC (EC_{sh}) and deep EC (EC_{dp}), respectively.

The objectives of this study were to evaluate the feasibility of using soil EC to create VRI prescriptions and test the response of crop yield to VRI management in humid region.

DOI 10.3920/978-90-8686-888-9_84, © Wageningen Academic Publishers 2019

Materials and methods

Experimental site

The study was conducted for 2 years in 2016 and 2017 in two adjacent fields (Field A and Field B) at the USDA-ARS Crop Production Systems Research Unit research farm in Stoneville, Mississippi, USA (latitude: 33°26'30.86', longitude: -90°53'26.60'). The experimental layout in the fields is shown in Figure 1. Each field is 6.7 ha with a 1% slope from west to east. Soil samples were taken from Fields A and B in a 0.3-ha grid and 150-mm depth and analysed for soil physical properties. Though silt loam was the predominant soil type, variability in clay and sand content existed across the fields. Fields A and B were under the coverage of a VRI centre pivot irrigation system, and occupied half of the pivot's full circle between 0 to 180 degrees (clockwise from north). Field A was in the circular angle 0° to 90° while Field B was in 90° to 180°.

Two treatments, variable rate irrigation (VRI) and uniform rate irrigation (URI) with four replicates were used in the study. Each of Field A and Field B were divided into 8 plots. All plots had the same area. One plot was assigned to one treatment. The remaining area not covered by the pivot in each field was employed as the rainfed plot.

Centre pivot VRI system

The irrigation system used in this study consisted of a Valley 8000 Standard Pivot coupled with the Valley VRI zone control package (Valmont Irrigation, Valley, NE, USA). Field tests showed that this centre pivot VRI system had a coefficient of uniformity of 86.5% with constant rate application and 84.3% with variable rate application (Sui and Fisher, 2015). The system was configured in 4 spans with a total length of 233 m. Sprinklers along the length of the centre pivot were divided into 10 control zones, with each zone covering the same surface area of 1.7 ha. The Valley VRI controller included the zone control units, solenoid valves, a GNSS (Global Navigation Satellite System) receiver, and software. The zone control unit controlled the duty cycle of the sprinklers by turning electric solenoid valves on and off to achieve desired application depths in individual control zones. The GNSS receiver determined the pivot's position in the field for identification of control zones in real time. VRI prescriptions were created using the software provided with the VRI system.

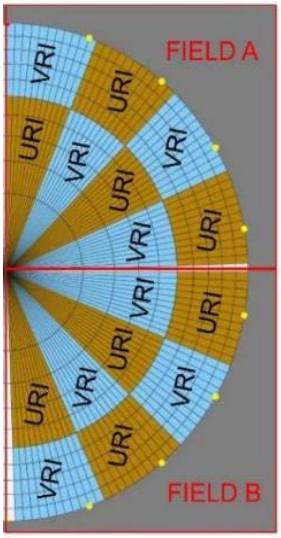


Figure 1. Experimental layout with two treatments (VRI and URI) and four replicates in each field.

Management zone delineation using soil electrical conductivity

Soil EC of Field A and Field B was measured in April 2012 and March 2016 using the Veris 3100 soil EC mapping system. The EC map generated using the data collected in 2016 had the very similar pattern to the EC map in 2012. The EC map of 2016 was used to delineate irrigation management zones for VRI in 2016 and 2017 seasons.

In the EC measurement in 2016, EC_{dp} and EC_{sh} were measured simultaneously. The EC_{dp} varied from 11.5 to 119.2 mS/m. The EC_{dp} was linearly related with the EC_{sh} (r^2=0.68). With consideration that more soil water is stored and more plant roots are grown in the soil horizon of 0 to 750 mm than in the soil horizon of 0 to 25 cm, EC_{dp} was selected for use to delineate VRI management zones. An EC_{dp} map of Field A and B was created using software ArcMap (version 10.2.1, ESRI, Redlands, CA, USA) (Figure 2). The EC_{dp} was classified into three categories in each field as shown in Figure 2.

Three management zones were created based on the soil EC_{dp}. In Field A and B, areas in EC_{dp} category 1, 2 and 3 were assigned as management zones A, B and C, (MZ-A, MZ-B and MZ-C) respectively. On account of their soil properties under the EC_{dp} categories and previously observed yield potential, irrigation rates of 100% (R100), 80% (R80), and 60% (R60) were respectively applied to MZ-A, MZ-B and MZ-C in the VRI treatment. Irrigation rate R100 was applied to the entire URI treatment. No irrigation was applied to the rainfed treatment. Irrigation rate R100 represented the irrigation rate that was determined using soil water content measured by soil moisture sensors, and the application rates of the other management zones were scaled based on their percentages. With soil EC_{dp} map as the background image, a VRI prescription was generated using software provided by the VRI system manufacturer (Valmont Irrigation, Valley, NE, USA). In the VRI prescription map, various depths of irrigation water were applied to different management zones according to the irrigation rate assignments (Figure 3).

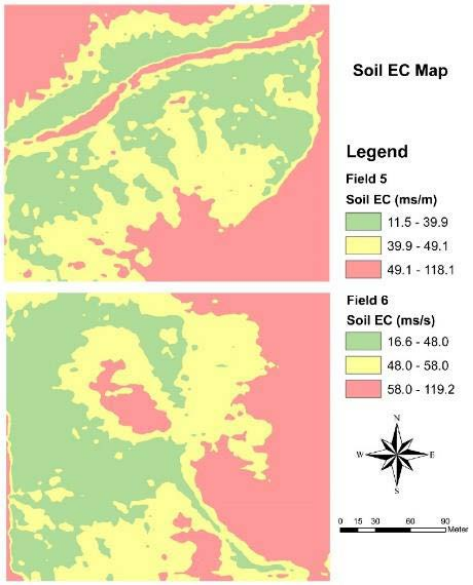


Figure 2. Soil electrical conductivity map of Field A (top, Field 5) and Field B (bottom, Field 6). The filled contours correspond to soil EC_{dp} categories 1-3 (in green to red on legend).

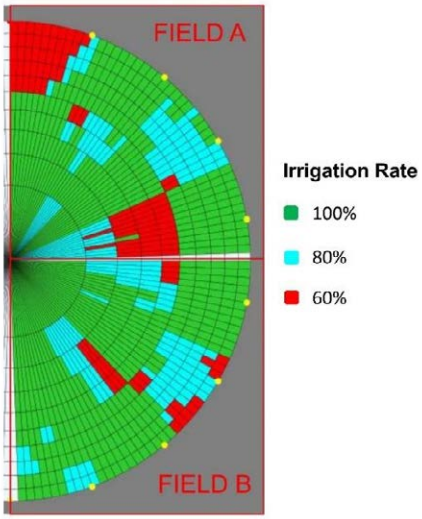


Figure 3. Prescription map for variable rate irrigation in 2016 and 2017. Irrigation water application rates were indicated by different colours on the map.

Field management

Corn (24BHR93) was planted in Field A on March 29, 2016 and March 20, 2017, respectively. Soybean (HBK4850LL) was planted in Field B on April 26, 2016 and April 7, 2017, respectively. Nitrogen fertilizer at 224 kg N/ha was applied as a urea-ammonium nitrate solution (N-sol, 32% N) to the corn field 43 days after planting (DAP) in 2016 and 28 DAP in 2017. Insects and weeds in both soybean and corn fields were controlled with generally recommended procedures in the region throughout the growing seasons. The corn and soybean were harvested using a combine equipped with a yield monitor so that the yield in each location within the field was recorded along with the spatial information. The corn was harvested on August 29 in 2016 and on August 24 in 2017. The soybean was harvested on September 30 in 2016 and on September 8 in 2017.

Irrigation scheduling and application

Soil water content sensors were installed at depths of 150, 300 and 610 mm in the predominant soil of the field to measure soil water content (SWC). The sensors were calibrated with the soil from the field. The weighted average of the soil water contents at the 3 depths was used for irrigation scheduling. Considering plant root distribution in the soil profile, the weight assigned to the measurement at depth of 150, 300 and 610 mm was 0.45, 0.35, and 0.2, respectively (Sui and Yan, 2017). Percent plant available water (PPAW) is calculated using Equation 1 to trigger irrigation events.

$$PPAW = \frac{sensor\ measured\ SWC - SWC\ at\ wilt\ point}{field\ capacity - SWC\ at\ wilt\ point} \quad (1)$$

Irrigation was triggered when PPAW dropped approximately to 50%.

In each irrigation event, a 19 mm depth of water was applied to the R100 zone; water depth applied to the other zones was scaled down according to the rate assigned. Irrigation water was delivered using the centre pivot VRI system described above. In 2016, three irrigation events were scheduled for corn and soybean with a total water depth of 57 mm applied to the R100 zone in each crop. In the 2017 season, 38 mm of irrigation water was applied in two irrigation events to the R100 zone in corn while 76 mm of water applied in four events to the R100 zone in soybean.

Data collection and analysis

The amount of irrigation water used in the VRI and URI treatments was measured using a water flow meter installed at the inlet of lateral pipeline of the centre pivot. Crop yield data from each plot were collected and analysed using PROC GLIMMIX procedure (SAS Institute Inc., Cary, NC, USA) to compare the effect of the irrigation treatment on yield and irrigation water productivity (WP). WP was defined as follows.

$$WP\ (\frac{kg}{m^3}) = \frac{Amount\ of\ grain\ produced\ with\ Irrigation\ Water\ (kg)}{Amount\ of\ Irrigation\ Water\ used\ (m^3)} \quad (2)$$

Results and discussion

Crop yield

In general, there was no significant yield difference between VRI and URI treatment for corn and soybean in 2016 and 2017. However, the VRI treatment used 30 and 13% less irrigation water than the URI in corn and soybean, respectively.

In 2016 corn, the yield in VRI, URI, and rainfed treatment was 11,329, 11,545 and 10,904 kg/ha, respectively. The corn yield in 2017 was similar to 2016, the yield in VRI, URI, and rainfed was 10,997, 11,428 and 10,324 kg/ha, respectively. In both years, no significant yield difference between VRI and URI was observed though URI yield was slightly high than VRI yield.

Soybean yield of VRI, URI, and rainfed treatment in 2016 was 4,497, 4,500, and 4,399, kg/ha, respectively. Soybean yield of VRI, URI, and rainfed treatment in 2017 was 3,966, 4,014, and 3,619 kg/ha, respectively. Statistical analysis indicated that there was no difference in yield between the VRI and URI treatments.

Compared to the rainfed treatment, irrigation increased corn yield by 4.9% in 2016 and 8.4% in 2017, respectively, and increased soybean yield by 2.2% in 2016 and 10.2% in 2017, respectively. During the growing season, rainfed corn and soybean plants experienced more water stress in 2017 than that in 2016, which resulted in more yield increase by irrigation in 2017 than in 2016.

Irrigation water productivity

Irrigation water use and the water productivity (WP) in VRI and URI treatments was calculated and is given in Table 1. The WP equals the amount of grain produced by irrigation water divided by the amount of irrigation water applied (Equation 2).

In corn, the WP of the VRI treatment was 16 and 26% lower than the WP in the URI in 2016 and 2017, respectively. In the 2-year averages, WP in corn was 1.67 kg/m^3 in the VRI management and 2.06 kg/m^3 in the URI. In soybean, the WP of VRI treatment was 11% higher than the URI in 2016. The WP of VRI and URI was almost the same in 2017. The average WP in soybean in the two years was 0.36 kg/m^3 for both VRI and URI treatment. It was observed that WPs in 2017 were higher

Table 1. Irrigation water use and water productivity of VRI and URI treatment for corn and soybean in 2016 and 2017.

Crop	Year	VRI water use (m$^{3)}$	URI water use (m^3)	VRI water saving (%)	VRI WP (kg/m^3)	URI WP (kg/m^3)
Corn	2016	918	1,191	30	0.99	1.15
	2017	612	794	30	2.34	2.96
Soybean	2016	1,058	1,191	13	0.20	0.18
	2017	1,411	1,588	13	0.52	0.53

than that in 2016. This could be due to the drier weather in 2017 summer that caused the irrigation to have greater impact on the yield.

Conclusions

Results of this 2-year study with corn and soybean crops showed that there was no difference between the yields in the VRI and URI treatment while the VRI saved 30% irrigation water in corn and 13% irrigation water in soybean. Compared with rainfed crops, irrigation increased yield in corn and soybean, especially in a dry year. The URI in corn had a slightly higher WP than the VRI. In soybean, the WP was not different between the VRI and URI. Results demonstrated that VRI management was able to save irrigation water in the Mississippi Delta region, which has a humid sub-tropical climate.

With a large spatial variability of soil EC in a field and understanding the relationships among the soil EC, soil properties and yield potential of the field, the method reported in this article has the potential to be used in other climates and fields to improve irrigation management. However, even though the use of soil apparent electrical conductivity to generate irrigation management zones could be an easy-to-use method in VRI management, research on the algorithms with multiple input variables for delineating VRI management zones and determining VRI application rates are needed because there are many factors affecting crop water requirements for irrigation.

Acknowledgements

Author thanks Mr Jonnie Baggard for his assistance in this study.

Disclaimer

Mention of a trade name, proprietary product, or specific equipment does not constitute a guarantee or warranty by the U.S. Department of Agriculture and does not imply approval of the product to the exclusion of others that may be available.

References

Corwin, D. L., Lesch, S. M., 2003. Application of soil electrical conductivity to precision agriculture: Theory, principle, and guidelines. *Agronomy Journal,* 95, 455-471.

Cox, M. S, Gerard, P. D., Abshire, M. J., 2006. Selected soil properties' variability and their relationships with yield in three Mississippi fields. *Soil Science,* 171, 541-551.

Johnson, C. K., Mortensen, D. A., Wienhold, B. J., Shanahan, J. F., Doran, J. W., 2003. Site-specific management zones based upon soil electrical conductivity in a semiarid cropping system. *Agronomy Journal*, 95, 303-315.

Schaible, G., Aillery, M., 2015. Irrigation & water use. http://www.ers.usda.gov/topics/farm-practices-management/irrigation-water-use.aspx. Accessed on January 30, 2019.

Sui, R. and Yan, H., 2017. Field Study of Variable Rate Irrigation Management in Humid Climates. *Irrigation and Drainage*, 66:327-339. doi:10.1002/ird.2111.

Sui, R., Fisher, D. K., 2015. Field test of a center pivot irrigation system. *Applied Engineering in Agriculture.* 31(1): 83-88. DOI 10.13031/aea.31.10539.

Thomasson J. A., Sui, R., Cox, M. S., Al-Rajehy, A., 2001. Soil reflectance sensing for determining soil properties in precision agriculture. *Transaction of ASAE* 44(6), 1445-1453.

Modeling spatio-temporal variations in crop water stress for variable-rate irrigation

J.D. Svedin[1*], N.C. Hansen[2], R. Kerry[3] and B.G. Hopkins[2]
[1]*Soil, Environmental, and Atmospheric Sciences University of Missouri, Columbia, MO, USA; jeffreysvedin@mail.missouri.edu*
[2]*Plant and Wildlife Sciences Brigham Young University, Provo, UT, USA*
[3]*Geography Brigham Young University, Provo, UT, USA*

Abstract

Critical to variable-rate irrigation (VRI) management, is measuring the temporal variation in crop water demand in response to spatially variable soil properties. The objective of this study was to model soil water and crop water stress at 85 locations in a field of winter wheat (*Triticum aestivum* L.) to provide data for future prescription maps. Soil moisture samples were collected at spring green-up and post-harvest to validate soil water modeling for two growing seasons near Grace, Idaho, USA. The range between the first and last points to experience modeled water stress was 21 and 56 days under uniform and variable-rate irrigation treatments, respectively.

Keywords: variable-rate irrigation, crop water stress, soil water depletion modeling

Introduction

Demand for freshwater resources worldwide are increasing due to a growing global population, declining groundwater levels, deteriorating water quality, environmental regulation and other factors (Evans *et al.*, 2013). In addition to a rising freshwater demand, the Food and Agriculture Organization of the United Nations (FAO) estimated that by 2030 withdrawals for irrigation alone would need to increase by 14% to meet world food demand (Abdullah, 2006). These future projections are clear evidence that farmers need to improve irrigation efficiency and food production per unit of applied water (Evans *et al.*, 2013).

Variable-rate irrigation (VRI) is a promising tool to improve food production per unit of water by spatially matching irrigation application rates to crop water demand. Lo *et al.* (2016) found that effective VRI techniques using stored soil moisture resulted in at least a potential 25 mm/yr irrigation reduction for 13% of agricultural fields with center-pivot systems in their area of study. Hedley and Yule (2009) demonstrated the ability of VRI to conserve up to 26% irrigation compared to uniform irrigation using a water balance approach in each irrigation zone, delineated from apparent electrical conductivity.

Coupling daily estimates of crop ET with spatial soil water holding capacity (SWHC) can be a basis to develop VRI before soil moisture falls below a crop specific threshold (Allen *et al.*, 1998). As soil moisture falls below crop specific soil moisture thresholds during important growth stages, crop transpiration is reduced and yield is affected. The objective of this study was to investigate the field-scale spatio-temporal variation of soil water depletion in the root zone and crop water stress to provide data for future VRI prescriptions.

Materials and methods

Site description and management practices

The study was conducted for two winter wheat growing seasons (2016-2017) in a commercial production field (22 ha) near Grace, ID, USA, (42.611113 N, 111.784244 W; 1,690 m elevation). The

DOI 10.3920/978-90-8686-888-9_85, © Wageningen Academic Publishers 2019

soil is a silty clay loam coarse-silty, mixed, superactive, frigid Calcic Haploxerolls with 1-4% slopes and 0.3 ha of unfarmed rock outcroppings (Figure 1A). There is some variability in topography but limited in terms of soil properties with texture analysis at 42 sites all being classified as silty clay loam. The mean annual precipitation is 392 mm with much of the precipitation falling as winter snowfall and spring rain. Irrigation is from a 380 m long center pivot sprinkler equipped with a VRI system (GrowSmart Precision VRI, Lindsay Zimmatic, Omaha, NE, USA).

In 2016, the VRI irrigation zones were determined by the farmer's experience with irrigation events occurring nearly every four days. The VRI in 2016 decreased irrigation in <10% of the model sites and did not significantly alter the soil moisture model because the difference between irrigation rates were minimal. As such, irrigation in 2016 is considered uniform. In 2017, the irrigation was spatially variable with areas applying the grower standard practice rate (GSP), and low and high irrigation treatments that applied ±30% relative to the GSP. Irrigation treatments (Figure 1) were based on historic yield and crop water productivity (Svedin, 2018) and irrigation timing and GSP rates were determined by the farmers experience. Planting dates were 5 and 10 October 2015 and 2016, respectively, and harvest dates were 16 and 30 August 2016 and 2017, respectively.

Soil water depletion modeling

Soil water depletion modeling followed Allen *et al.* (1998) and the FAO Penman-Monteith estimation of ET to adjust daily soil water levels. Soil core samples were collected at 85 points for both seasons at spring green-up and post-harvest to measure soil water content and estimate plant available soil water. Daily root zone soil water depletion was modeled at each sample point beginning at the spring green-up sampling and continuously through post-harvest sampling. Spatially uniform non-stressed reference evapotranspiration (ET_o) and a crop coefficient (K_c) were used to estimate the wheat crop specific ET (ET_c) and the remaining soil water in the root zone. When soil moisture fell below a site-specific readily available water content (RAW) threshold, ET_c was adjusted with a crop stress coefficient (K_s) to estimate the adjusted crop water stressed ET ($ET_{c\text{-}adj}$). The model was validated by comparing the summed modeled daily $ET_{c\text{-}adj}$ to the ET measured from the following seasonal water balance approach:

$$ET = P + I + \Delta S - D \quad (1)$$

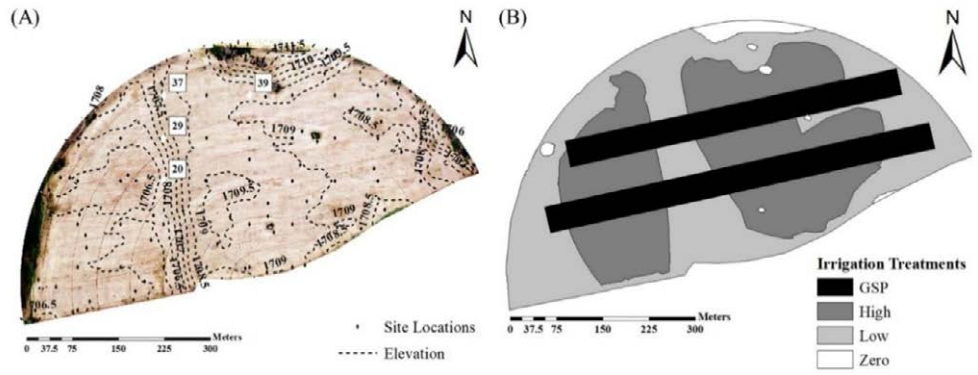


Figure 1. (A) Bare soil image with elevation contours and modeled point locations- note: dark colored areas are exposed bedrock; and (B) 2017 irrigation treatments (note: GSP = grower's standard practice, High = rates 30% > GSP, Low = rates 30% < GSP, and Zero = no irrigation).

where P represents the seasonal precipitation, ΔS represents the difference in measured soil moisture between spring green-up and harvest, I is irrigation, and D is drainage during the season. Spring green-up and harvest soil samples were collected on April 18 and August 20, 2016 and on May 4 and September 1, 2017. Soil cores were collected at four depths 0-0.3 m, 0.3-0.6 m, 0.6-0.9 m, and 0.9-1.2 m with a 0.051 m diameter auger to describe the moisture removed from the soil at each site (Figure 1A). Drainage was assumed to occur when modeled soil moisture content exceeded SWHC. Statistical measures of model fit included: (1) relative error (RE); (2) normalized objective function (NOF); (3) root mean square error (RMSE); and (4) R-squared. Where RE represents model bias, NOF indicates the fit of the model (where NOF<1 is less than one standard deviation from the mean), and RMSE is the difference between predicted and measured values in the same units as the modeled variables.

Weather data were collected at a weather station ~2,000 m from the field site (42.51496 N, –111.73606 W; Cooperative Agricultural Weather Network AgriMet). The weather data were used to calculate daily evapotranspiration (ET_o) for a reference crop using the American Society of Civil Engineers (ASCE) Standard Penman-Monteith Evapotranspiration model (Allen *et al.*, 1998). The crop coefficient approach (Allen *et al.*, 1998) was used to calculate daily $ET_{c\text{-}adj}$ using the following equation:

$$ET_{c\ adj} = K_s * K_c * ET_o \quad (2)$$

where $ET_{c\ adj}$ is the adjusted ET from water stress and winter wheat crop specific conditions, K_s is the transpiration reduction factor reflecting the available soil water, K_c is the crop coefficient that reflects the winter wheat crop development stage, and ET_o is the reference crop ET. The K_c curve was estimated from table values in a similar environment (Allen *et al.*, 1998) and validated with field observations. The dimensionless transpiration reduction factor is represented as K_s and calculated as:

$$Ks = \frac{\mathrm{SWHC} - \mathrm{Dr}}{(1 - \mathrm{p}) * \mathrm{SWHC}} \quad (3)$$

where SWHC is the soil water holding capacity in the root zone (mm), p is the table value 0.55 from Allen *et al.* (1998) that represents the average fraction of SWHC that can be depleted before crop stress, and Dr is the current root zone water depletion in mm.

Daily soil moisture was modeled based on SWHC and readily available soil water using:

$$\mathrm{SWHC}=1000* (\theta_{FC}\text{-}\theta_{WP})*Z_r \quad (4)$$

$$\mathrm{RAW} = p*\mathrm{SWHC} \quad (5)$$

where SWHC is defined as above and calculated for the crop rooting depth in mm, θ_{FC} is the volumetric water content (m^3/m^3) at field capacity (FC), θ_{WP} is the volumetric water content at wilting point (m^3/m^3), and Z_r is the estimated rooting depth of the crop (1.2 m). RAW is the readily available soil water content in the root zone (mm), and *p* is defined above. Field capacity measurements followed the guidelines of Martin *et al.* (1990) which state that sampling 1-3 d after a thorough soil-wetting event when crop water use is small is a good indication of FC. The spring green-up sampling used the snow melt as the wetting event and FC was assumed to be the greatest measured soil water content between the two spring green-up measurements for each point. It was assumed autumn soil samples were at WP because the soils had been dried down for harvest with 15 and 7.1 mm of rain since the last irrigation event 18 and 16 days earlier for 2016 and 2017, respectively. Daily root zone depletion was estimated using the following equation:

$$D_{r,i} = D_{r,i-1} - P_i - I_i + ET_{c,i} + DP_i \quad (6)$$

where $D_{r,i}$ is the root zone soil moisture depletion at the end of d_i, $D_{r,i-1}$ is the root zone depletion at the end of the previous day, P_i is the precipitation on d_i, I_i is the irrigation depth on d_i, $ET_{c,i}$ is the crop evapotranspiration on d_i, and DP_i is deep percolation out of the root zone on d_i.

Results and discussion

Model evaluation

Modelled seasonal ET values matched seasonal water balance ET values (RMSE 73 mm), with a moderate bias towards over-prediction (Table 1). Modeled predictions of soil water content at harvest also matched well with measured values (RMSE 81 mm; Table 1), with bias towards under prediction (RE -5%; Table 1). For both parameters, model predictions were within one standard deviation of the mean for each season (NOF <1). Model fit was better in 2016 than in 2017 likely because daily ET was calculated using a spatially uniform crop coefficient, which disregarded the 2017 irrigation treatments. The irrigation treatment had a significant influence on yield production and adjusting the crop coefficient to match crop growth and staging would have improved accuracy.

2016 Uniform irrigation

Sampled spring soil moisture values demonstrated substantial variation of soil water depletion ranging from zero (at FC) to ~50 mm (Figure 2). Large variations in early season soil water depletion were not expected considering how shortly after snowmelt sampling was conducted. Because of the uniform irrigation, the points with the most depleted soil moisture at spring sampling were the first points to experience crop water stress later in the season. Measuring this early season spatial variability in soil moisture can help growers address the variability in early season soil moisture with VRI and avoid crop water stress during sensitive growth stages in many field locations.

Irrigation events began on day 151 (May 1) with 198 mm applied over 11 events (Table 2) ending on day 193. Peak crop water demand covered days 170-191 and corresponded with the grain development growth stage. The first modeled soil moisture fell below RAW on day 175 (point 20; Figure 1; Table 2) with the last on day 196 (point 37; Figure 1; Table 2). The onset of water stress for point 20 began early in these growth stages while point 37 did not experience water stress until after the growth stage had ended. The $ET_{c\text{-}adj}$ for point 37 on day 196 was 8.4 mm compared to 5.4 mm for point 20; a 36% decrease in $ET_{c\text{-}adj}$ representing severe ET restrictions and yield loss from water stress. The observed three week range in the onset of crop water stress is reflective of the variability of plant available water spring soil moisture levels (Figure 2). This variability was not anticipated but demonstrates the inherent variability found in some fields (Longchamps *et al.*, 2015).

Table 1. Model fit statistics used to evaluate model predicted values of seasonal evapotranspiration and soil moisture at harvest. Model statistics include root mean square error (RMSE), R-squared, relative error (RE) and normalized objective function (NOF).

Year		Seasonal evapotranspiration			Soil moisture at harvest		
		2016	2017	Mean	2016	2017	Mean
RE	%	0.10	2.4	1.3	8.0	-13	-5.0
NOF		0.10	0.18	0.14	0.20	0.80	0.50
R^2		0.61	0.83	0.79	0.49	0.10	0.35
RMSE	mm	53	91	73	27	132	81

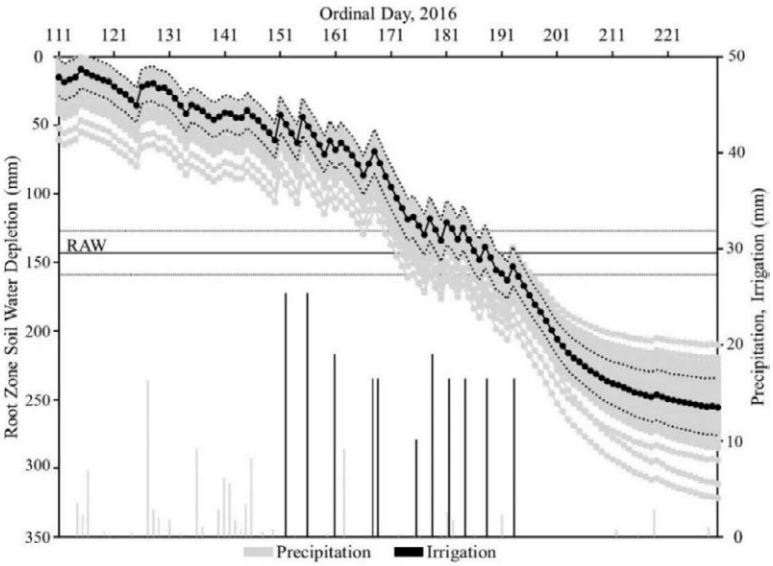


Figure 2. Variation of soil moisture depletion in the root zone at 85 modeled sites in a commercial field (grey lines). Modeling began at spring green-up until harvest in 2016, under uniform irrigation. The field average daily soil water depletion is identified in black with one standard deviation (SD) from the mean (dashed lines). Rain and irrigation rates identified with bar graphs on the secondary y-axis. Field average readily available water (RAW) and one SD from the mean are identified with horizontal black lines and represent the temporal variability in the onset of crop water stress found in this field.

2017 Variable-rate irrigation

Variable irrigation treatments were applied during each event with a total irrigation of 158, 225 and 291 mm for low, GSP, and high treatments in the 2017 season, respectively (Figure 1B). Each

Table 2. Precipitation, irrigation, seasonal ET, and the ordinal day of the first and last field sites modeled to experience crop water stress by season and irrigation rate.

Season	Precipitation	Irrigation	ET	Onset of crop water stress	
				First	Last
	---------- mm ----------			Ordinal day	
2016	95	198	520	175	196
2017	90			155	211
2017 Irrigation treatments					
Low		158	437	155	187
GSP		225	509	158	195
High		291	573	186	211

irrigation event, 14 total, and GSP rates were producer decisions based on GSP test strips. At each irrigation, the high and low treatment received ±30% relative to that application rate. While this irrigation strategy is sub-optimal because of the test strips, in total, ~25% less irrigation was applied to the field relative to a uniform GSP application.

The first onset of stress was modeled on day 155 (point 29; Figure 1) and the last occurred day 211 (point 39; Figure 1) – spanning 56 days. $ET_{c\text{-}adj}$ for point 39 on day 211 was 7.0 mm while point 29 $ET_{c\text{-}adj}$ was 2.9 mm; a 59% reduction in daily $ET_{c\text{-}adj}$. This temporal range in the onset of crop water stress was significantly larger than the previous season, largely in part, from the irrigation treatments. The onset of crop stress for the low, GSP and high irrigation rates were 32, 37 and 25 days respectively (Figure 3; Table 2). The range in the onset of crop water stress was greater in each treatment compared to the field range in 2016 (Figure 3). This is likely because the variability in soil water depletion at spring sampling was much greater in 2017. The greater variability in soil moisture depletion is attributed to the differences in snow accumulation patterns and spring weather conditions.

For both seasons, the variations in observed soil moisture at spring sampling influenced root zone soil moisture depletion the remainder of the season (Figure 2 and 3). The treatment averages for the onset of crop water stress were days 176, 184 and 205 for the low, GSP and high treatments respectively. There was an average 29 day difference between the onset of crop water stress between the low and high treatments (Figure 3). The peak ET demand for the crop spanned days 166-206 and, similar to 2016, the earliest modeled points experienced crop stress early in the season. The average high irrigation treatment points experienced water stress near the end of grain-fill, which did not significantly harm yield (Svedin, 2018). During the final two irrigation applications, many of the high irrigation modeled points were not water-stressed, suggesting that yield was not improved

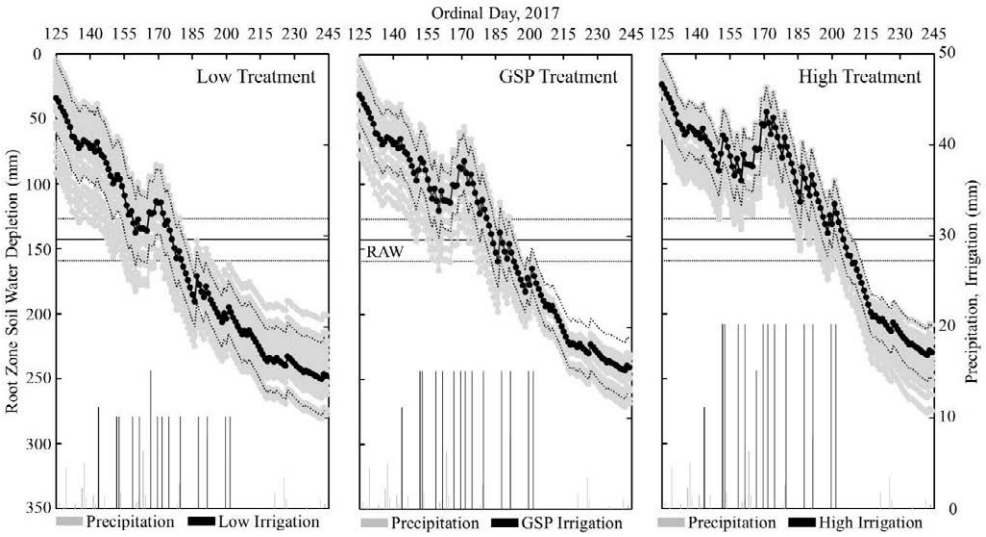


Figure 3. Modeled root zone soil moisture depletion for 85 sampling points, in grey, from spring green up until harvest for 2017. Each irrigation treatment soil moisture depletion average is identified in black with one standard deviation (SD) from the mean (dashed lines). Rain and irrigation events identified on the secondary y-axis. Field average and one SD of readily available water (RAW) are identified as black horizontal lines to represent the temporal variation in the onset of crop water stress within each treatment.

by the final two irrigation applications. Not applying irrigation to the high treatment sites could conserve water and has been suggested by Lo *et al.* (2016).
The difference between the treatment average onset of crop water stress was much larger between the GSP and high treatment than the GSP and low treatment (Table 2). This is anticipated because the supplementary early season irrigation acted as a buffer against crop water demand during peak ET rates. This suggests that early season irrigation could be used to build a soil water reservoir to postpone water stress when it is difficult logistically for irrigation managers to apply large irrigation rates to match soil water depletion. The adoption of VRI prevents the oversaturation of the soil profile in some field locations when filling these soil profile reservoirs. Finally, this data demonstrates how critically important it is to create accurate VRI zones. Inaccurate VRI zones can lead to significant plant water stress and subsequent yield reductions as observed in the 2017 season (Table 2).

Conclusions

The results of this work present significant field-scale spatial and temporal variability in modeled soil moisture and crop water stress at 85 sample sites. The range in the onset of crop water stress was 21 days under uniform irrigation compared to 56 days under VRI treatments. Spring soil samples measured soil moisture depletion ranging 0-90 mm soon after snow melt, with late season crop water stress often reflecting variability found in spring soil moisture. Soil water dynamics during these two seasons demonstrate the potential for VRI in this field and suggests early season soil moisture sensing could provide data for late season irrigation management. Further work is needed to confirm the observed variability at other field locations with different soil and environmental conditions and identify the properties governing spatial and temporal crop water demand.

Acknowledgements

We thank the Charles Redd Grant Center for Western Studies for funding, Christensen Farms for collaboration, and BYU student employees for sample collection and processing.

References

Abdullah, K.B., 2006. Use of water and land for food security and environmental sustainability. Irrigation and Drainage, 55 219-222.

Allen, R. Pereira, L.S. Raes, D., and Smith M., 1998. Crop evapotranspiration. Guidelines for computing water requirements. FAO irrigation and drainage paper 56. Rome, Italy: FAO.

Evans, R.G., LaRue, J., Stone, K.C., and King, B.A., 2013. Adoption of site-specific variable-rate sprinkler irrigation systems. Irrigation Science, 31 871-887.

Hedley, C.B., and Yule, I.J., 2009. Soil water status mapping and two variable-rate irrigation scenarios. Precision Agriculture, 10 342-355.

Lo, T., Heeren, D.M., Martin, D.L., Mateos, L., Luck, J.D., and Eisenhauer, D.E., 2016. Pumpage reduction by using variable-rate irrigation to mine undepleted soil water. Transactions of ASABE, 56 1285-1298.

Longchamps, L., Khosla, R., Reich, R., and Gui, D.W., 2015. Spatial and temporal variability of soil water content in leveled fields. Soil Science Society of America Journal, 79 1446-1454.

Martin, D.L., Stegman, E.C., and Feres, E., 1990. Irrigation scheduling principles. In Hoffman, G.J., Howell, T.A., and Solomon, K. H. (Eds.), *Management of Farm Irrigation Systems*, St. Joseph, MI, USA: ASAE, pp. 155-206.

Svedin, J.D., 2018. Characterizing the spatial variation of crop water productivity for variable-rate irrigation management. MS thesis. Brigham Young University, UT, USA: Department of Plant and Wildlife Sciences.

Comparison of precision and conventional irrigation management of cotton

E. Vories[1*], S. O'Shaughnessy[2] and M. Andrade[3]
[1]*USDA-ARS CSWQRU, 147 State HWY T, Portageville, MO 63873, USA; earl.vories@ars.usda.gov*
[2]*USDA-ARS CPRL, 300 Simmons RD, Unit 10, Bushland, TX 79012, USA*
[3]*ORISE, 100 ORAU Way, Oak Ridge, TN 37830, USA*

Abstract

Soil textural variability diminishes the effectiveness of conventional irrigation management. Variable rate irrigation (VRI) can address soil variability; however, users need guidance to prepare prescriptions for optimal water application. A study was conducted at Portageville, MO, USA, in 2016 and 2017 with the objective to compare rainfed cotton and cotton irrigated based on the USDA-ARS Irrigation Scheduling Supervisory Control And Data Acquisition (ISSCADA) system and a water balance method. Although the trend was for the rainfed treatment to have the lowest yield in both years, the yield differences among all treatments in 2017 were not significant. The study is continuing at Portageville and other locations and the ISSCADA system is constantly being improved to better meet the needs of agricultural producers.

Keywords: irrigation, irrigation scheduling, cotton

Introduction

Irrigated agriculture in the US is a major consumer of freshwater, accounting for approximately 80% of the nation's consumptive water use (Schaible and Aillery, 2015). While irrigation is essential for crop production in arid and semi-arid regions, in recent years, irrigation has increased rapidly in more humid areas. Uncertainty in the amount and timing of precipitation is one of the most serious risks to crop production in those areas and timely irrigation has been shown to increase yields of cotton (Vories *et al.*, 2007). Producers have become increasingly reliant on irrigation to ensure adequate yields and reduce production risks.

Irrigation scheduling systems include weather-based, soil moisture-based and plant-based methods. Weather-based methods schedule irrigation based on the estimated amount of water lost by plant evapotranspiration (ET) and the amount of effective rainfall and irrigation water entering into the plant root zone. The Arkansas Irrigation Scheduler (AIS) has been in use for almost forty years in Arkansas and surrounding states to aid farmers in managing irrigation. The AIS uses a water-balance approach to schedule irrigation based on the soil water deficit (SWD). Rooting depth is not used explicitly in the program, but is implicit in the choice of a maximum allowable SWD. Cahoon *et al.* (1990) provided information about the program and Vories *et al.* (2009) reported on changes after the original publication.

Soil textural variability within many irrigated fields diminishes the effectiveness of conventional irrigation management with respect to irrigation uniformity, scheduling and water use efficiency. Irrigation scheduling methods that assume uniform soil conditions may produce less than satisfactory results on highly variable soils, contributing to lack of interest or commitment to irrigation scheduling even though, in general, scheduling improves overall crop production. Furthermore, benefits of variable-rate application of agrochemicals, seeds and nutrients can be partially masked by applying inappropriate amounts of water. The development of variable rate irrigation (VRI) technology provides the opportunity to address some of these problems, but several management questions remain unanswered (Stone *et al.*, 2010).

DOI 10.3920/978-90-8686-888-9_86, © Wageningen Academic Publishers 2019

Commercial VRI systems have been tested and shown to perform dependably (e.g. Sui and Fisher, 2015). The USDA-ARS team in Bushland, TX, recently patented an Irrigation Scheduling Supervisory Control And Data Acquisition (ISSCADA) system (Evett *et al.*, 2014) that can develop prescriptions and control VRI systems. With integrated sensor network subsystems, the ISSCADA system detects variable crop water needs and provides spatially variable recommendations for watering rates. An integrated CWSI (iCWSI) was presented by O'Shaughnessy *et al.* (2012) as a way to reduce some of the limitations of the traditional CWSI measurements. With the ISSCADA system, an irrigation event is triggered when an iCWSI threshold is exceeded, and a fixed amount of water is applied. Conversely, no irrigation is applied if the iCWSI threshold is not exceeded. This approach to triggering irrigation worked well in field tests in the U. S. Southern High Plains. Grain yields and water use efficiency (WUE) for crops irrigated at the highest treatment amounts resulted in similar or better yields and WUE than irrigations accomplished using a neutron probe (NP) (O'Shaughnessy and Evett, 2010; O'Shaughnessy *et al.*, 2012). However, it is uncertain how ISSCADA systems that rely solely on plant feedback will perform under extended periods of cloud cover and in climates with high humidity, because the CWSI is less responsive to irrigation needs under these conditions (Bockhold *et al.*, 2011). Furthermore, although the infrared thermometer (IRT) sensors view the crop at oblique angles and thus can report accurate canopy temperature data even under less than full cover, there is a lower limit of canopy cover below which temperature data may be influenced by soil background (e.g. soil is usually warmer than canopy); this may result in abnormal iCWSI values and false irrigation triggers during partial canopy cover. The current version of the ISSCADA system includes soil water sensors to provide more direct measurements of crop irrigation requirements in space and time, and for different climates. The iCWSI and soil water content are both accounted for in the resulting irrigation recommendations. The goal of this ongoing research is to test the ISSCADA system in a range of climates and determine ways to improve it. The specific objective of this study was to compare the yield of rainfed cotton and irrigated cotton based on the ISSCADA system and the AIS, and to compare the irrigation water use efficiency (IWUE) of the cotton between the two irrigation scheduling methods.

Materials and methods

A field study was conducted at the University of Missouri Fisher Delta Research Center Marsh Farm at Portageville (36.41° N, 89.70° W) during the 2016 and 2017 growing seasons to investigate irrigated cotton production. The rectangular field is approximately 5 ha, 320 m by 156 m, with the primary slope in the S-N (320 m) direction. It is located roughly 14 km west of the Mississippi River and lies within the New Madrid Seismic Zone. The combination of alluvial, aeolian and seismic activity over the years has resulted in highly variable soils in the region. Farming activities, including precision land grading, have made the effects less obvious but they still exist.

The test location had been in corn production for three years prior to establishing this experiment. Weather data were collected on a 5-minute interval approximately 600 m from the most distant part of the study site and placed on the University of Missouri Agricultural Electronic Bulletin Board (AgEBB; http://agebb.missouri.edu).

Soil mapping units within the study field included Tiptonville silt loam (Oxyaquic Argiudolls), which made up the majority of the field, and Reelfoot loam and sandy loam (Aquic Argiudolls) (USDA-SCS, 1971) (Figure 1). However, the field contains areas of high sand content that are too small to show up in the soil survey map. To provide higher resolution information, mobile soil apparent electrical conductivity (EC_a) data were collected on 13 April 2016. A Veris MSP3 system (Veris Technologies, Inc., Salina, KS, USA) was used to collect data on a 1-s interval and a 10-m transect spacing. Soil texture data (sand, silt and clay fractions) from laboratory analysis were obtained by soil horizon from 8 calibration locations within the study field and an adjoining 5-ha field. These locations were chosen to cover the range in mobile EC_a data values obtained. Veris EC_a data were inverted using the

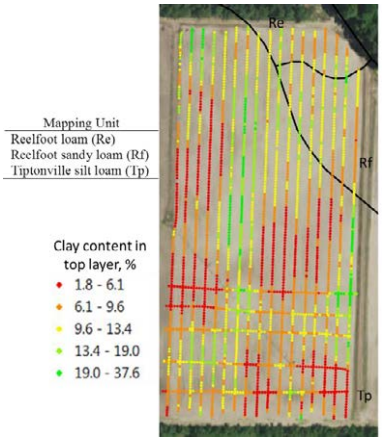


Figure 1. Aerial image of study field with overlaid clay content from EC_a transects.

invVERIS software package (EMTOMO, Odivelas, Portugal) and layer EC values at the calibration locations were related to measured soil properties using linear regression. A detailed description of the procedure, including validation of the findings, was reported by Sudduth *et al.* (2017).

The field was bedded in the north-south direction with a row spacing of 0.97 m and the cotton variety PHY 333WRF was seeded at 13 seed/m on 23 May 2016 and 15 May 2017. Nitrogen (134 kg N/ha) was broadcast applied near the first square growth stage, normally approximately 30 d after planting. Reduced tillage was utilized and standard cultural practices for weed and insect control for producing irrigated cotton in Missouri were employed. Growth regulator, insecticide, and defoliant were applied as blanket treatments to the whole field.

The study field and adjoining field were irrigated using a 160 m Valley 6000 center pivot irrigation system (Valmont Irrigation, Valley, NE, USA) with a Valley Zone Control VRI system. The system included 7 independently controlled zones of approximately equal area, except for the outermost zone, which consisted of the 3 sprinklers on the 5-m overhang beyond the outer drive tower. The 3 experimental plots each consisted of two adjacent management zones. The seventh zone was managed the same as the adjacent treatment but the yield data were not included in the analyses.

Twelve Dynamax SapIP-IRT sensors (Dynamax, Inc., Houston, TX, USA) were mounted to the pivot lateral and two were placed at stationary locations within the field. The IRTs were included in a wireless network with a remote computer where the data were collected and stored. The observed canopy temperatures, together with the local weather data, were used to calculate iCWSI values for the irrigation management zones. Soil moisture content was not considered in creating the irrigation prescriptions.

The AIS (version 2.3.4) was updated daily from before planting until 17 August, the end of the irrigation season. A 19 mm application was applied to the AIS plots whenever the estimated SWD exceeded 44 mm. After a few days without rain, the center pivot system was operated without water application to allow the ISSCADA system IRTs to scan the dry canopy. The daily 5-minute weather was downloaded at approximately midnight and used with the canopy temperature to prepare a VRI prescription for the following day. The ISSCADA system was set up to recommend an application of 0, 6, 13 or 22 mm within a plot and each replication was managed independently of the others. In 2016, there were four replications, each comprising an arc of 45° (Figure 2). In 2017, there were five replications, each comprising an arc of 36°.

The crop was harvested on 10 Nov. 2016 and 6, 7 Oct. 2017 with a Case IH 2155 cotton spindle picker (Case IH, Racine, WI, USA) equipped with an Ag Leader Insight (Ag Leader Technology,

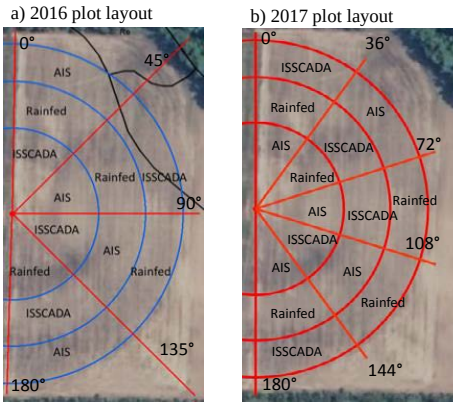


Figure 2. Plot layouts for (A) 2016 and (B) 2017.

Ames, IA, USA) yield monitor system with sensors for every row. The yield monitor was calibrated by placing the cotton from each harvest load in a boll buggy equipped with scales (2016: Short Line Manufacturing, Shaw, MS, USA; 2017: Master Scales, Greenwood, MS, USA). Yield data were obtained with the program SMS Basic 19.00 (Ag Leader Technology). Erroneous data points were removed from the yield data with Yield Editor 2.07 (Sudduth and Drummond, 2007; Sudduth *et al.*, 2012)). The large, spatially referenced data sets were managed with ArcGIS for Desktop 10.4.1 (ESRI, Redlands, CA, USA). Spatial data analyses were conducted with GeoDa 1.12 (Center for Spatial Data Science, Univ. of Chicago, Chicago, IL, USA) using the spatial error model as recommended for yield monitor data (Griffin *et al.,* 2007).

To account for overspray from adjacent nozzles, a 7-m buffer was placed on each side of each plot and the yield obtained in those areas was not used in the subsequent analyses. For the outermost zone, the overhang sprinklers (zone 7) served as the buffer. To estimate the IWUE, the average yield for the rainfed treatment was subtracted from the yield for each irrigated plot. That marginal yield value was divided by the total irrigation application for the plot.

Results and discussion

The 2016 and 2017 growing seasons were similar and typical for the region (Table 1). However, while the total rainfall through the irrigation period (planting through 17 August) was similar, the timing differed, resulting in more irrigations in 2017. For the AIS, each replication was irrigated the same amount on the same day; however, the ISSCADA-managed plots were irrigated independently from each other. Even in cases where the total seasonal application was the same for two replications, the timing of the applications was sometimes different.

The soil variability illustrated in Figure 1 is reflected in the yield maps from the study field for both growing seasons (Figure 3A and 3B). Some differences can be seen, likely due to the different randomization of irrigation treatments between the two years (Figure 2). The highest quintile for both years was approximately the same, with >3,600 and >3,610 kg seed cotton/ha in 2016 and 2017, respectively. However, the lowest quintile contained much lower yield in 2017, with <2,090 and <1,160 kg seed cotton/ha in 2016 and 2017, respectively. Figure 3C and 3D shows the yield after the buffer areas between plots and the areas outside the center pivot system were removed. The yield quintiles were more similar after the area outside the irrigation system were removed.

Table 1. Short grass reference evapotranspiration (ET_o) and water applications to field from planting to 17 August, when irrigation scheduling was terminated.

Water application	ISSCADA	AIS	Rainfed
2016			
Planting	23 May		
Growing degree days (15.6 °C base)	1,009		
Reference evapotranspiration (ET_o; mm)	390		
Total rainfall (mm)	321		
Number of irrigations	1.25*	2	0
Total irrigation application (mm)	24*	38	0
Total water applied (mm)	303	318	279
2017			
Planting	15 May		
Growing degree days (15.6 °C base)	937		
Reference evapotranspiration (ET_o; mm)	434		
Total rainfall (mm)	309		
Number of irrigations	3.8*	6	0
Total irrigation application (mm)	65*	112	0
Total water applied (mm)	371	418	306

* Value is average of all replications.

The statistical comparisons for the seed cotton yield for the two years are included in Table 2. In 2016, yield of the AIS treatment was not significantly different from either the ISSCADA or rainfed treatment. Yield of the ISSCADA treatment was significantly greater than the rainfed treatment. When the water application (Table 1) was considered, the IWUE for the ISSCADA treatment was significantly greater than for the AIS treatment. In 2017, the range of treatment yields was greater than in 2016, with 419 and 622 kg seed cotton/ha in 2016 and 2017, respectively. However, due to greater variability among the plot yields, there were no significant differences among the treatments (Table 2). With the larger total water application associated with the AIS treatment, the IWUE for the ISSCADA treatment was significantly greater than that for the AIS in 2017. While the value for the AIS treatment was similar both years, the ISSCADA treatment had a lower IWUE in 2017. In 2016, the rainfed treatment yield was not significantly different from the AIS treatment yield (Table 2). In 2017, there was a non-significant trend for rainfed treatment yield to be lower than the yield of either irrigated treatment, but the differences were not significant. This is observed in some years when a large rainfall occurs shortly after an irrigation. In 2017, the AIS plots and two of the ISSCADA plots were irrigated on 4 August and 25 mm of rainfall were recorded on 6 August. The value of an increase in IWUE without an associated increase in yield (i.e. the increased IWUE resulted from water savings) often depends on the location of the farm. For example, in the Portageville, MO, area where this study was conducted, irrigation water is pumped from a shallow aquifer. With no direct cost for the water and no limits on the amount of pumping, most producers would opt for the higher yield if given a choice. In areas where the producers must pay for the water they use and/or are limited in their pumping, the economically best practice could be different. The study is continuing at Portageville and other locations and the ISSCADA system is constantly being improved to better meet the needs of agricultural producers.

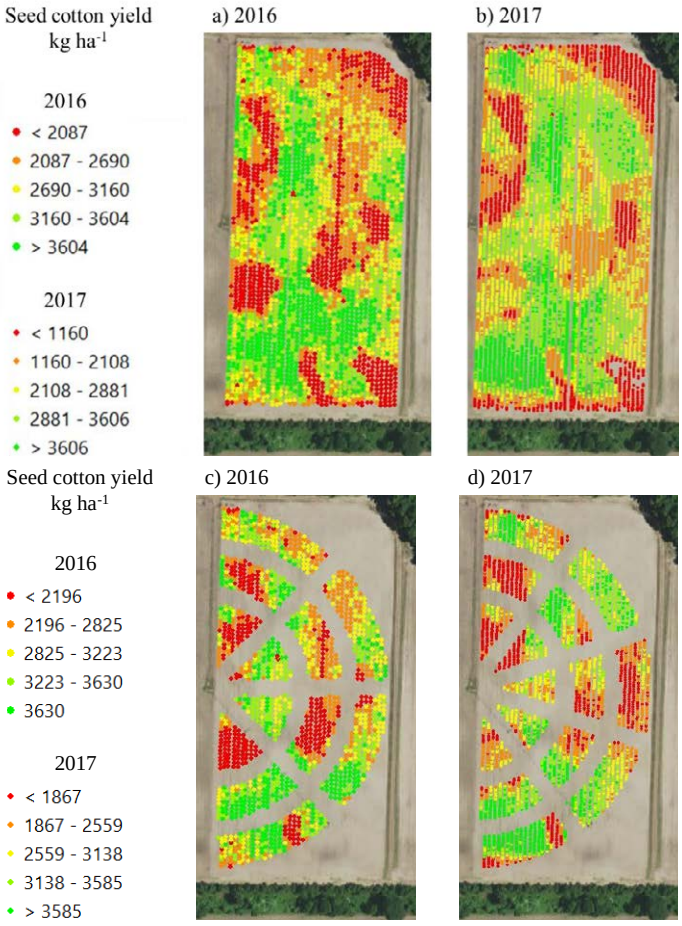


Figure 3. Seed cotton yield quintiles for whole field (kg/ha) for (A) 2016 and (B) 2017; seed cotton yield quintiles (kg/ha) in study plots only for (C) 2016 and (D) 2017.

Table 2. Seed cotton yield and irrigation water use efficiency (IWUE).

Irrigation treatment	Seed cotton yield (kg/ha)	IWUE (kg/m^3)
2016		
ISSCADA	3,132 a	1.6 a
Arkansas Irrigation Scheduler	2,853 ab	0.4 b
Rainfed	2,713 b	----
2017		
ISSCADA	2,686 a	1.1 a
Arkansas Irrigation Scheduler	2,865 a	0.5 b
Rainfed	2,243 a	----

Conclusions

This 2-year study compared rainfed to irrigated cotton based on 2 irrigation scheduling methods in southeastern Missouri, USA. The AIS recommended a uniform irrigation application based on an estimated SWD, while the ISSCADA system made site specific recommendations based on an iCWSI. The principal findings of the study were:

While the total rainfall during the irrigation period of planting until mid-August was similar both years, the timing of the rainfall resulted in more irrigation applications and thus a greater total irrigation application in 2017.

Although the trend was for the rainfed treatment to have the lowest yield in both years, the yield difference between rainfed and AIS in 2016, and the yield differences among all treatments in 2017 were not significant.

Even though the yields between the two irrigation scheduling treatments were not significantly different in either year, the lower total irrigation applications resulted in significantly greater IWUE for the ISSCADA treatment in both years.

Acknowledgment and disclaimer

We acknowledge the support of Valmont Industries and Jake LaRue in this research, as well as the contributions of Andrea Jones.

Mention of trade names or commercial products in this paper is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the United States Department of Agriculture.

References

Bockhold, D. L., Thompson, A. L., Sudduth, K. A. and Henggeler, J. C., 2011. Irrigation scheduling based on crop canopy temperature for humid environments. Transactions of the ASABE 54(6):2021-2028.

Cahoon, J., Ferguson, J., Edwards, D. and Tacker, P., 1990. A microcomputer-based irrigation scheduler for the humid mid-South region. Applied Engineering in Agriculture 6(3):289-295.

Evett, S. R., O'Shaughnessy, S. A. and Peters, R. T., 2014. US Patent No. 8,924,031, 'Irrigation Scheduling and Supervisory Control and Data Acquisition System for Moving and Static Irrigation Systems', issued 30 December 2014.

Griffin, T. W., Brown, J. P. and Lowenberg-DeBoer, J., 2007. Yield monitor data analysis protocol: a primer in the management and analysis of precision agriculture data. http://www.agriculture.purdue.edu/SSMC. (last accessed 6/12/18).

O'Shaughnessy, S. A. and Evett, S. R., 2010. Canopy temperature based system effectively schedules and controls center pivot irrigation of cotton. Agricultural Water Management 97(9):1310-1316.

O'Shaughnessy, S. A., Evett, S. R., Colaizzi, P. D. and Howell, T. A., 2012. A crop water stress index and time threshold for automatic irrigation scheduling. Agricultural Water Management 107:122-132.

Schaible, G. and Aillery, M., 2015. Irrigation & water use. http://www.ers.usda.gov/topics/farm-practices-management/irrigation-water-use.aspx. (last accessed February 2019).

Stone, K. C., Bauer, P., Busscher, W., Millen, J., Evans, D. and Strickland, E., 2010. Variable-rate irrigation management for peanut in the eastern coastal plain. Paper No. IRR10-8977. St Joseph, MI, USA: ASABE

Sudduth, K. A. and Drummond, S. T., 2007. Yield editor: Software for removing errors from crop yield maps. Agronomy Journal 99(6):1471-1482.

Sudduth, K. A., Drummond, S. T. and Myers, D. B., 2012. Yield Editor 2.0: Software for automated removal of yield map errors. Paper No. 121338243. St Joseph, MI, USA: ASABE

Sudduth, K. A., Kitchen, N. R. and Drummond, S. T., 2017. Inversion of soil electrical conductivity data to estimate layered soil properties. In: J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Bioscience 8(2): 433-438

Sui, R. and Fisher, D. K., 2015. Field test of a center pivot irrigation system. Applied Engineering in Agriculture 31(1):83-88.

USDA-SCS, 1971. Soil survey of Pemiscot County. USDA, Missouri, Washington, D.C., USA

Vories, E. D., Hogan, R., Tacker, P., Glover, R. and Lancaster, S., 2007. Estimating the impact of delaying irrigation for Midsouth cotton on clay soil. Transactions of the ASABE 50(3):929-937.

Vories, E., Tacker, P. and Hall, S., 2009. The Arkansas Irrigation Scheduler. In: S. Starrett (Ed.) World Environmental and Water Resources Congress 2009: Great Rivers, Reston, VA, USA: ASCE. pp. 1-10.

Section 12
Geostatistics, mapping and spatial data analysis

A geophysical and spectrometric sensor data fusion approach for homogeneous within-field zone delineation

G. Buttafuoco[1*], R. Quarto[2], F. Quarto[3], M. Conforti[1], A. Venezia[4], C. Vitti[5] and A. Castrignanò[1]
[1]*National Research Council of Italy, ISAFOM, Rende (CS), 87036, Italy; gabriele.buttafuoco@cnr.it*
[2]*Earth and Geoenvironmental Sciences department, University of Bari A. Moro, 70125, Italy*
[3]*Geoprospector srl, Bari, 70121, Italy*
[4]*CREA Research Centre for Vegetable and Ornamental Crops, Pontecagnano, 84098, Italy*
[5]*CREA Research Centre for Agriculture and Environment, Bari, 70125, Italy*

Abstract

The paper describes a geostatistical approach for combining multi-source data with different support for field delineation into homogeneous soil zones. Geophysical and spectrometric sensor data were used with soil chemical properties measured in laboratory on 50 samples collected in a field cropped with Tomato (*Solanum lycopersicum*, L. cv *San Marzano*). The approach consisted in performing Gaussian anamorphosis with support correction, multi-collocated block cokriging and factorial block cokriging to jointly analyse all data. Two regionalised factors at different spatial scales were retained to split the field into homogeneous zones for site specific management.

Keywords: hyperspectral data, support change, block cokriging, factorial cokriging

Introduction

Sustainable agricultural production requires adopting site specific crop management based upon assessment of within-field variation (Corwin, 2003). Therefore, any precision agriculture (PA) system must first measure and then understand variability. Traditional soil sampling is not often viable because it is costly and labour intensive and generally requires a large number of soil samples to achieve an accurate representation of soil properties. Proximal sensing technology and data spatial analysis techniques can make the application of PA feasible. Moreover, it would be beneficial if different instruments were used in conjunction with one another (data fusion), so to take advantage of their complementary characteristics (Castrignanò *et al.*, 2017). In these cases, a process of combining information from heterogeneous sources into a single composite picture would be needed, such that the composite picture were more accurate and complete than that derived from any single source alone. Such a combination process is called data fusion (Hall and McMullen, 2004).
Electromagnetic induction (EMI) sensors are commonly used to measure apparent soil electrical conductivity (ECa) in the field because of their speed, ease of use, relatively low cost and the volume of fine-scale data collected (Doolittle and Brevik, 2014). Electromagnetic induction sensors such as GEM300® (Geophysical survey Systems, Inc., Nashua, NH, USA) can measure ECa and the operating frequency can be set at different values as a function of the surveyed soil depth of interest.
Laboratory soil spectroscopy in the visible and near infrared (400-2,500 nm) region of the electromagnetic spectrum, combined with chemometric data analysis, has been suggested as a cost- and time-saving procedure to characterize soil chemical properties (Conforti *et al.*, 2018; Summers *et al.*, 2011; among many others). Spectral data, besides being directly used for statistically predicting soil properties, can be integrated in a geostatistical analysis (Zovko *et al.*, 2018) and used as an auxiliary (covariate) measure of soil quality.
Generally, no mention of data support sizes is made but geophysical and spectral data as well as soil properties determined in the laboratory are associated with different support sizes and their combination results in the problem of change of support (Emery, 2007; Matheron, 1976). This

DOI 10.3920/978-90-8686-888-9_87, © Wageningen Academic Publishers 2019

problem is really crucial because correlations among variables are scale-dependent, which implies that inferences obtained at a given scale cannot be transferred to another scale. In addition, to manage effectively such multi source data, a sort of homogenization is necessary to make them compatible (Castrignanò *et al.*, 2017; Gotway and Young, 2002). There is no single and well defined methodology for combining heterogeneous data in a proper way. Geostatistics (Matheron, 1971) enables rigorous treatment of spatial prediction by using a mathematical model of spatial dependence and is flexible enough to jointly analyse data of different types. Block cokriging (Webster and Oliver, 2007) is a multivariate geostatistical approach which allows handling heterogeneous data and solving the critical issues of combining multi source data with different support sizes. Block cokriging requires fitting a linear model of coregionalization averaged over a block and produces the prediction of each variable over the same block.

The objective of this paper is to propose a geostatistical approach for field delineation into homogeneous soil zones by integrating data from multiple sensors with sparse soil sample data and taking into account spatial dependence and the change of support from input data to predictions. The approach was applied using data from an EMI sensor, a hyperspectral sensor in the visible and near infrared (400-2,500 nm) region, and the conventional laboratory measurements of some soil chemical properties.

Materials and methods

Study area and measurements

The study area (Figure 1) is located in southern Italy (Lat. 40°48'22"N; Long. 14°34'12"E) on the alluvial plain of the Sarno river and it is characterized by white stony-sandy soils in silt matrix. The climate is Mediterranean with rather mild and rainy winters and very hot and dry summers (Longobardi *et al.*, 2016). In 2015, the soil of a field cropped with tomato (*Solanum lycopersicum* L. cv *San Marzano*) in Campania region (south-western Italy) was sampled with a regular auger at 50 nodes of a quasi-regular grid with mesh of about 9 m and down to, 0.30 m-depth, so to evenly cover the 5,500 m^2- field. The location of soil samples was measured using a GNSS system with differential correction (DGNSS, HiPer 27 Pro, Topcon, Tokyo, Japan) and sub-metric accuracy. The projected co-ordinate system is universal transverse Mercator (UTM zone 33N).

The soil samples were analysed in laboratory for total organic carbon content (TOC) and concentrations of nitrogen (N), potassium (K), and phosphorous (P) (Violante, 2000).

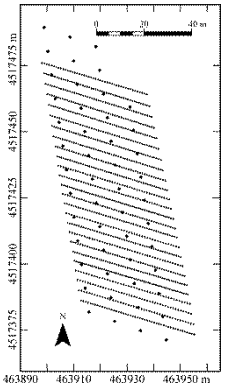


Figure 1. Study area and survey scheme: the black dotted lines are the 24 transects where the geophysical surveys were carried out; the diamonds are the soil sample locations.

In addition, reflectance of each soil sample, in the radiometric range between 350-2,500 nm, was measured in laboratory using an ASD FieldSpec IV spectro-radiometer and pre-processed according to the procedures described by Conforti *et al.* (2015). Principal component analysis (PCA) (Jackson, 2003) of spectral data was performed to reduce the dimensionality of the data.
A soil survey was carried out to measure ECa (mS/m) along 24 parallel transects with a multiple frequency electromagnetic profiler GEM300®. Only ECa data measured with the highest operating frequency (19,975 Hz) were used in this study to characterise the topsoil. The maximum depth of investigation of this operating frequency was 1.59 m (Huang, 2005) but the effective depth (the depth above which 50% of the signal contribution effect is derived) was about 0.40 m (Sudduth *et al.*, 2005). Therefore, the vertical supports of the three different types of measurements can be considered comparable, differently the surface supports that were quite different: punctual (a circle of about 0.008 m^2 area) for soil sample data including reflectance data, and about 1.2 m (distance between the two transmitter-receiver coils) × 1 m (the mean distance travelled in 1 sec) for EMI data.

Geostatistical approach for data fusion

A data fusion approach based on geostatistics is proposed to take advantage of the complementary features of the different types of data by combining the data sets in a manner that is statistically robust. Multivariate geostatistics offers a set of optimal linear (cokriging) estimators (unbiased and with minimum error variance) and requires a spatial statistical model of dependence. Moreover, block cokriging allows to refer the estimates of all the variables to the same support (block) even if the input data were measured on different supports (Castrignanò *et al.*, 2017). However, the derivation of cokriging estimates requires the inversion of the covariance matrix describing the relationships between the observations. Because proximal sensing datasets are often very large, inversion of the covariance matrix is difficult if not impossible. Multicollocated cokriging (Rivoirard, 2001) is a simplified version of full cokriging, which makes use of the auxiliary variable(s) only at the target point and at all locations where the primary variable of interest is available (Castrignanò *et al.*, 2009). This approach relies on the availability of auxiliary variable(s) known (collocated) at all the blocks of the interpolation grid, which requires to interpolate the even numerous observations (vectorial format) of auxiliary variable(s) on a grid (raster format). Moreover, it combines effectively the differences among the sensor data sets and reinforces the complementary value of proximal sensing and ground-based observations of the variable of interest (Castrignanò *et al.*, 2012).
Another multivariate geostatistical tool, called factorial cokriging, (Goovaerts, 1997) applies a multivariate geostatistical data-fusion technique to synthesize multi-type data at multiple spatial scales. It creates simple summary scale-dependent indices that can be used to delineate homogeneous management zones for site-specific management (Castrignanò *et al.*, 2017).
To apply the above data fusion approach, using soil sample data, ECa, and the retained principal components of reflectance, several steps of data processing were required: (1) the selected auxiliary variable ECa was previously interpolated using point kriging, so to be known at each node of a 1×1 m interpolation grid; (2) ECa data were migrated to the locations of sample data so to obtain a coregionalization data set, needed to calculate the matrix of experimental point-based direct and cross-variograms; (3) all the variables were transformed into Gaussian-shaped variables by fitting a mathematical function of Hermite polynomials (Gaussian anamorphosis) (Wackernagel, 2003), when variables showed skewed data distributions; (4) the anamorphosis coefficients, were corrected, due to change of support, to get an anamorphosis function defined on a block support (Chiles and Delfiner, 2012); (5) the calculated point experimental (direct and cross-) variograms were submitted to regularization (averaging) (Castrignanò *et al.*, 2017) over the 1 m by 1 m-block; (6) a linear block model of coregionalization (LMC) was fitted to all experimental variograms; (7) using the previously estimated LMC multicollocated block cokriging for prediction of all variables, and (8) factorial block cokriging, to extract few synthetic scale-dependent indicators to be used for field partition, were performed; (9) finally the block estimates of the Gaussian transformed variables

were back transformed using the corrected block anamorphosis function to provides estimates in the original units. More details on the procedures can be found in Chiles and Delfiner (2012) and Castrignanò *et al.* (2017).
The spectra pre-processing and PCA were performed with Unscramble X 10.5 (https://www.camo.com/unscrambler); all geostatistical techniques and mapping were implemented in ISATIS® 2018.1 (www.geovariances.com).

Results and discussion

Principal component analysis of soil spectra showed that the first principal component (PC1) accounted for about 92% of the total variance; therefore it was retained and used as a synthetic spectral index to be processed together with the other variables. Loading coefficients of PC1 (data not reported) showed that PC1 was mainly influenced by the reflectance of the bands between 750 and 2,250 nm, because the main absorption wavelengths of most chemical soil components fall within this range. The data distributions of EC_a, K and TOC showed significant deviations from normality (Table 1).
To perform multivariate analysis, it was then decided to transform all data into standardized Gaussian-shaped variables to mean 0 and variance 1, using the point Gaussian anamorphosis. This transformation was required also for the quasi-Gaussian variables to reduce the impact of the outliers on variogram fitting. After that, the block anamorphosis was calculated with support correction for each variable and in Table 2 the correction coefficients together with the punctual and block variance are reported for each variable. The Table 2 highlights the reduction of variance due to averaging over the block, especially for PC1, N and P.
This result is extremely critical for the site-specific management. The preliminary condition is to fix the spatial scale the phenomena under study refer to. If, for example, the absorption of N and P by roots occurs on a scale smaller than 1×1 m, producing a soil composite sample over 1-m^2 or more of the field might cause the filtration of more than 40% of the variance, which might negatively impact the optimization process of local fertilization. What above said underlines the importance of properly selecting the sensor with a support (footprint) comparable to that of the study process, because all predictions can be done at this scale when using proximal sensing.
An anisotropic nested model was previously fitted to the punctual variogram of the Gaussian transformed ECa data consisting of the following three spatial structures: a nugget effect; an isotropic spherical model with range of 20 m and a zonal anisotropic K-Bessel model (Chilès and Delfiner, 2012) along the direction of 107° from north with range = 40 m and parameter = 2.

Table 1. Summary statistics for apparent electrical conductivity (ECa, mS/m), first principal components of spectral data (PC1), potassium (K, mg/kg), nitrogen (N, mg/kg), phosphorous (P, mg/kg), and total organic carbon (TOC, %). Number of samples=50.

	ECa	PC1	K	N	P	TOC
Mean	90.69	0.00	692.81	1.22	135.33	16.94
Median	88.24	-0.36	678.32	1.23	134.32	16.55
Minimum	80.25	-32.26	556.74	0.90	113.00	14.18
Maximum	106.69	38.58	863.18	1.61	171.09	21.74
Stand. dev.	6.37	14.12	70.06	0.18	13.88	2.22
Skewness	0.75	0.23	0.73	-0.06	0.27	0.76
Kurtosis	2.81	3.29	3.05	2.19	2.19	2.73

Table 2. Block support correction calculations for apparent electrical conductivity (ECa, mS/m), first principal component of spectral data (PC1), potassium (K, mg/kg), nitrogen (N, mg/kg), phosphorous (P, mg/kg), and total organic carbon (TOC, %).

	ECa	PC1	K	N	P	TOC
Punctual variance	40.51	198.43	4892.58	0.03	192.03	4.88
Block variance	39.90	57.31	4741.44	0.01	74.14	4.39
Block support correction	0.99	0.54	0.99	0.65	0.63	0.96

The use of point kriging is justified by the fact that the surface support size of GEM sensor outcomes is about 1.2×1 m, therefore it is comparable with the one of the grid-cell.
The experimental variograms of the Gaussian data were regularized over 1×1 m block and the fitted block LMC consisted of an isotropic linear model of coregionalization including a cubic model (Chilès and Delfiner, 2012) with a range of about 35 m and a K-Bessel model with a range of about 14 m and a parameter of 10.46 (Figure 2).

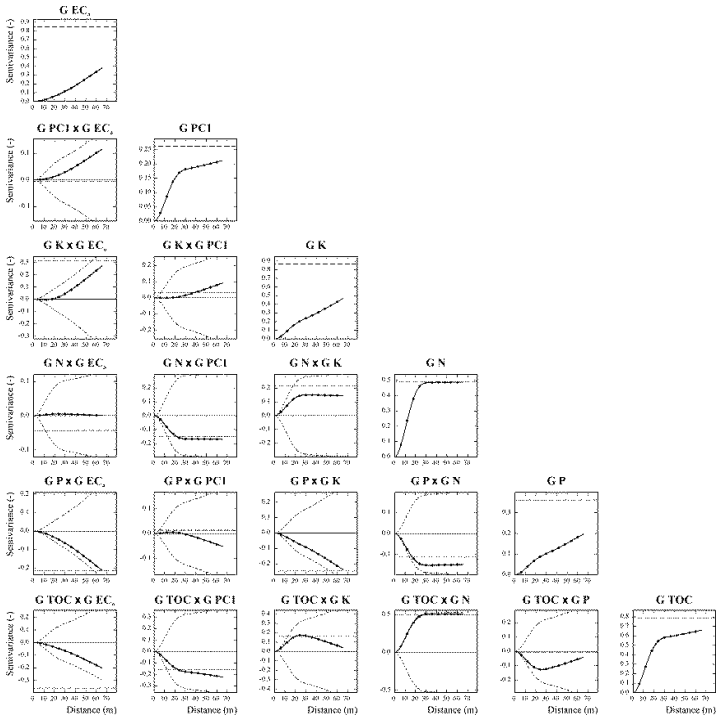


Figure 2. Auto- and cross-variograms of the block linear model of coregionalization for the Gaussian data of apparent electrical conductivity (ECa), first principal components of spectral data (PC1), potassium (K), nitrogen (N), phosphorous (P) and total organic carbon (TOC). The plotted points are the experimental values whereas the solid lines are the model of coregionalization. The dashed-dotted lines are the hull of perfect correlation while the horizontal dashed lines are the experimental variances.

The visual inspection of the cokriged maps of TOC, N, P, K, and PC1 (Figure 3), the last one assumed as a synthetic indicator of soil quality, reveals a quite wide central area of the field characterised by lower chemical fertility.

The greater visible continuity (uniformity) along the direction 170° from the north in the variation of P, K and TOC, perpendicularly to the rows of cultivation, is to be attributed to intrinsic variation of the soil rather than due to an effect of agricultural operations.

A completely different spatial pattern is shown by the ECa map. Actually the field could be coarsely split into two, northern and southern, areas with different dielectric properties of soil. This diversification might be due to physical characteristics of soil, such as texture, porosity and/or moisture.

Among the six regionalised factors corresponding to each spatial scale (14 m and 35 m) only the first factor for each scale was retained because it was associated to an eigenvalue greater than 1. Therefore, the two factors explained a proportion of the variance at each scale greater than the one explained by each variable standardised to unity variance. In particular, the first factor at shorter range explained a proportion of variance equal to 95.7% and was mainly affected by the Gaussian transformed variables of ECa and K; whereas the first factor at longer range explained a proportion of variance equal to 82% and on it the Gaussian transformed variables of TOC and N weighed more and negatively.

The two first regionalized factors have a different pattern (Figure 3H, G). The map of the first regionalized factor at short range (Figure 3G) shows a clear partition of the field into northern and southern zones, quite similar to the one produced by ECa.

The first factor at long range shows a pattern similar to nitrogen map but in inverse relation as expected (Figure 3D and 3H).

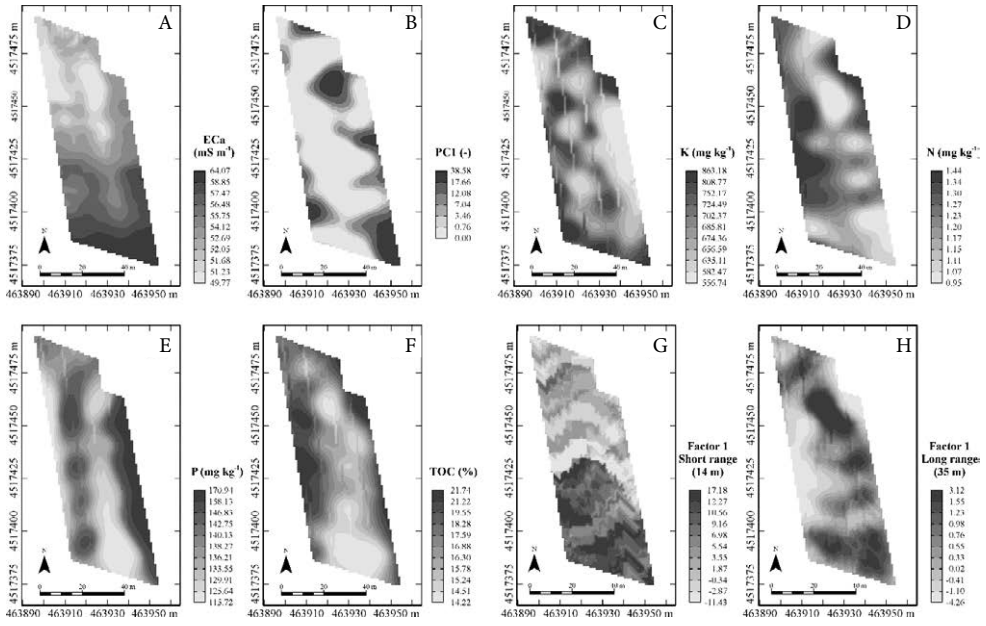


Figure 3. Maps of (A) apparent electrical conductivity (ECa), (B) first principal components of spectral data (PC1), (C) potassium (K), (D) nitrogen (N), (E) phosphorous (P), (F) total organic carbon (TOC), first regionalized factor at short (G) and long (H) range.

The two previous maps show how the field delineation may vary as a function of the spatial scale, because the soil processes involved are scale dependent and as a consequence also the relevant soil properties. All this warns against using traditional statistical clustering techniques to delineate homogeneous areas in the field. These do not take into account the proximity of the samples and their spatial correlation, and jointly process data regardless of the spatial scale and support.
Therefore, the factor at short-range could be used to direct agricultural operations, such as variable rate application (VRA) for seeding or irrigation, more related to physical properties of soil, whereas the factor at long range could support site-specific N fertilization.

Conclusions

A geostatistical approach, for combining heterogeneous data to produce a field delineation into homogeneous zones, has been described. Block cokriging has been used to jointly interpolate the geophysical, spectral, and traditional laboratory data on the same block despite their different original supports.
Currently combining heterogeneous spatial data is a novel and active area of statistical research. However, it is an issue of great interest also for PA practitioners, therefore it is hopeful that the data fusion methods are quickly incorporated into GIS software packages.

Acknowledgements

Financial support for this work comes from the project 'M2Q' PON03PE_00180_1 co-funded by the National Operational Programme for Research and Competitiveness (PON R&C) 2007-2013, through the European Regional Development Fund (ERDF) and national resource (Revolving Fund – Cohesion Action Plan MIUR). D.M. MIUR n. 738/05.03.2014.

References

Castrignanò, A., Buttafuoco, G., Quarto, R., Vitti, C., Langella, G., Terribile, F. *et al.,* 2017. A combined approach of sensor data fusion and multivariate geostatistics for delineation of homogeneous zones in an agricultural field. Sensors 17 2794 1-20.

Castrignanò, A., Costantini, E.A.C., Barbetti, R., Sollitto, D., 2009. Accounting for extensive topographic and pedologic secondary information to improve soil mapping. Catena 77 28-38.

Castrignanò, A., Wong, M. T. F., Stelluti, M., De Benedetto, D., Sollitto, D., 2012. Use of EMI, gamma-ray emission and GPS height as multi-sensor data for soil characterization. Geoderma 175-176 78-89.

Chilès, J.P., Delfiner, P., 2012. Geostatistics: Modelling spatial uncertainty. 2nd edition. New York, USA: Wiley.

Conforti, M., Castrignanò, A., Robustelli, G., Scarciglia, F., Stelluti, M., Buttafuoco, G., 2015. Laboratory-based Vis-NIR spectroscopy and partial least square regression with spatially correlated errors for predicting spatial variation of soil organic matter content. Catena 124 60-67.

Conforti, M., Matteucci, G., Buttafuoco, G., 2018. Using laboratory Vis-NIR spectroscopy for monitoring some forest soil properties. Journal of soils and sediments 18 1009-1019.

Corwin, D. L., Lesch, S. M., 2003. Application of Soil Electrical Conductivity to Precision Agriculture. Agronomy Journal 95 455-471.

Doolittle, J. A., Brevik, E. C., 2014. The use of electromagnetic induction techniques in soils studies. Geoderma 223-225 33-45.

Emery, X., 2007. On some consistency conditions for geostatistical change-of-support models. Mathematical Geology 39 205-223.

Goovaerts, P., 1997. Geostatistics for Natural Resource Evaluation. New York, USA: Oxford University Press.

Gotway, C. A., Young, L. J., 2002. Combining incompatible spatial data. Journal of the American Statistical Association 97 632-648.

Hall, D. L, McMullen, S.A.H., 2004. Mathematical Techniques in Multisensor Data Fusion. 2nd edition. Norwood, MA, USA: Artech House, Inc.

Huang, H., 2005. Depth of investigation for small broadband electromagnetic sensors. Geophysics 70 G135-G142.

Jackson, J. E., 2003. User's guide to principal components. New York, USA: John Wiley & Sons.

Longobardi, A., Buttafuoco, G., Caloiero, T., Coscarelli, R., 2016. Spatial and temporal distribution of precipitation in a Mediterranean area (southern Italy). Environmental Earth Sciences 75 1-20 189

Matheron, G., 1971. The theory of regionalized variables and its applications. Les Cahiers du Centre de Morphlogie Mathematique (Vol. 5). Fontainebleau, France: Ecole Nationale Superieure des Mines de Paris.

Matheron, G., 1976. Forecasting block grade distributions: the transfer functions. In M. Guarascio, M. David, C. J. Huijbregts (Eds.), Advanced Geostatistics in the Mining Industry (pp. 237-251). Dordrecht, the Netherlands: Springer.

Rivoirard, J., 2001. Which Models for Collocated Cokriging? Mathematical Geology 33 117-131.

Sudduth, K. A., Kitchen, N. R., Wiebold, W. J., Batchelor, W. D., Bollero, G. A., Bullock, D. G., *et al.,* 2005. Relating apparent electrical conductivity top soil properties across the North-Central USA. Computer and Electronics in Agriculture 46 263-283.

Summers, D., Lewis, M., Ostendorf, B., Chittleborough, D., 2011. Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties. Ecological Indicators 11 123-131.

Violante, P. (Ed.), 2000. Metodi di Analisi Chimica del Suolo (Chemical methods of soil analysis). Milan: Italian Ministry of Agriculture, Franco Angeli (in Italian).

Wackernagel, H., 2003. Multivariate Geostatistics: an introduction with applications. 3rd ed. Berlin, Germany: Springer-Verlag.

Webster, R., Oliver, M.A., 2007. Geostatistics for Environmental Scientists. 2nd ed. Chichester, UK: Wiley.

Zovko, M., Romić, D., Colombo, C., Di Iorio, E., Romić, M., Buttafuoco, G. *et al.,* 2018. A geostatistical Vis-NIR spectroscopy index to assess the incipient soil salinization in the Neretva River valley, Croatia. Geoderma 332 60-72.

Wheat yield forecast using contextual spatial information

M. Fajardo*, B. Whelan, P. Filippi and T. Bishop
Precision Agriculture Laboratory, Sydney Institute of Agriculture, University of Sydney, 1 Central Av., Eveleigh, 2016 Australia; mario.fajardo@sydney.edu.au

Abstract

In recent years, the sources of data on farming practices have increased exponentially. On-the-go fuel consumption, grain moisture and yield per hectare are now easily stored by farmers. Fine resolution (5 m) wheat yield forecast is presented here using two machine learning approaches: (1) bootstrapped regression trees (BRT) where predictions are pixel-wise; and (2) convolutional neural networks (CNN) where predictions use neighbouring pixels. This study made use of publicly available data (e.g. Sentinel 2 imagery) and farm-owned data (e.g. yield data). Results showed better performance of CNN over BRT, and when higher resolution data was included.

Keywords: convolutional neural networks, bootstrapped regression trees, yield forecasting, remote sensing

Introduction

As mechanization of agricultural practices increases, different types of information become easily available such us vehicle fuel consumption, speed, harvested grain protein and moisture. Despite this increasing amount of data, its practical use in agricultural decision-making tools remains minimal. There is also an ever-increasing amount of publicly available sources of information created daily e.g. worldwide satellite imagery.

One of the more common uses of remotely sensed data, is the calculation of normalized difference vegetation index (NDVI) using multispectral imagery (Irons *et al.*, 2012) with the subsequent application of a clustering process (K-means for example) for the purpose of establishing management zones to optimize the use of fertilizers. However, the development of more complex analyses is still in the early stages. (e.g. yield forecasting).

Within-paddock and seasonal variability of crop yield is the reason why forecasting yield is still a challenge (Hasan *et al.*, 2019). The main studies that have explored the possibility of yield forecasting are those from different governments or international organizations related to food-security or policy-making and their spatial coverage is state-wide or even country-wide (Baruth, *et al.*, 2008; Bauer, 1975; Doraiswamy *et al.*, 1979; Erickson, 1984).

Within-paddock yield varies from location to location and is largely seasonally influenced (due to sudden rainfall events for example). Traditionally, farmers will estimate their yield goals based on previous experience and personal knowledge about their farms. Recent advances in machine learning algorithms and machine vision have catalysed the improved use of the ever increasing amount of digitized data (Dempewolf *et al.*, 2014; Krizhevsky, *et al.*, 2012; You, *et al.*, 2017).

In this study, forecasting within-paddock yield variability was attempted at a fine resolution (5 m) using two of the increasingly popular machine learning approaches. On the one hand a BRT approach is undertaken, where the predictions are made pixel-wise, and on the other hand a CNN approach is used where 2D convolutions are included as sliding windows, which adds spatial 'patterns' to the model. The goal of the study was to test and report the results of a basic modelling example, using only a few layers of information in order to test the potential for forecasting wheat yield just using publicly available sources of information.

DOI 10.3920/978-90-8686-888-9_88, © Wageningen Academic Publishers 2019

Material and methods

Study area

The study was conducted on a large aggregation of several farms (with 47 fields or paddocks in total) owned by a single corporation and located in the southern agricultural region of Western Australia (120.726° / -33.698° Lat/Lon, WGS 84). The soils of the area are typically sandy with notable amounts of gravel, and one of the most widespread soil types are sodosols (Isbell, 2002). Only paddocks with wheat yield information were used for this investigation.

The study area is characterised by a Mediterranean climate, with cool, wet winters, and hot, dry summers and average annual rainfall of the aggregations ranges from 420 to 542 mm. This study used data from 2013 to 2017 growing seasons, with a variation in the amount of total rainfall received between each year ranging from 389.0 mm to 689.3 mm for the different fields (Bureau of Meteorology, 2019).

Data layers

Multiple data layers were available (up to 44 layers including satellite imagery, rainfall and terrain attributes) from different public data-sources. All the layers were compiled in a space-time array of ~80 M pixels, covering a total area of ~6,800 ha across a time-span of 4 years, i.e. 2013-2017. A total of 35,948 pixels with wheat yield data were selected from this array as the initial training dataset. For this initial study, only 8 of those layers were used in order to test the efficacy of a very basic example of a deep network analysis, and to keep within a reasonable computational time frame. The specific layers were chosen to represent well known properties used in yield predictions, namely biomass, soil physical and chemical properties and terrain attributes.

The models used two calculated NDVI layers, either from Landsat 8 (30 m resolution) (Irons *et al.*, 2012) or Sentinel 2 (10 m resolution) (Drusch *et al.*, 2012) satellites, from the clearest (cloud-free) images of both autumn and winter seasons. In addition, surface clay content (%) and total nitrogen (%) (both 90 m resolution) were included as soil attributes modelled by the terrestrial ecosystem research network (TERN) project (Rossel *et al.*, 2015) and a digital elevation model (DEM) (30 m resolution) layer used as a terrain attribute (Farr *et al.*, 2007). Finally, surface electromagnetic induction data (DualEM 21S, DualEM, Ontario, Canada) (5 m resolution) was included in the modelling as a soil attribute. All the layers were first set to a uniform resolution by a nearest neighbour algorithm and later processed using the Google earth engine platform (Gorelick *et al.*, 2017).

Model specifications

Two modelling types were employed, the first followed a BRT approach and the second a CNN approach, and both modelling approaches were validated against an independent dataset, following a 'leave one field out' validation. Samples from all the fields but one, were used in creating a model that then was used for predicting the one field excluded. Using this procedure, a distribution of predictions was obtained rather than a single performance parameter.

Bootstrapped regression trees (BRT)

A 'cubist' algorithm was selected as the regression trees model (Kuhn *et al.*, 2014a; Quinlan, 1986). This model builds successive discrimination points that segregate the data into different 'branches', and at the end of each branch a different multivariate linear regression is created to model the relationship between the variables in each branch. These regression models are combined to form an algorithm that predicts the final outcome.

To make the process more robust, a bootstrapping technique was applied to the BRT process. Using this technique, 50 different regression trees were created using a 90% subset of the original dataset, but for each tree a different subset was randomly chosen. Subsets are used to create different samples

of the data population so that a distribution of predicted outcomes can be achieved. At the end, the mean of the predictions from the 50 trees was used as the final result.

Convolutional neural networks (CNN)

CNNs take most of the concepts of an artificial neural network (ANN) (LeCun *et al.*, 1989) which are an attempt to use concepts from neurobiology, where many *neurons* (or nodes) are linked by *synapses* (or weights) between them. A typical neural network design will involve an input layer with n *nodes*, and one or more *hidden* layers that will carry and modify information from the input layer to a final output layer, which results in a function of the multiple weights of each of the hidden layers. If, in the input process, a *tensor* (n-dimensional array) with at least 2 dimensions is used instead of a one dimensional vector (as is the case of regression trees), then information from more than one point source (e.g. groups of pixels rather than single pixels) can be used. The approach is graphically depicted in Figure 1. As shown in Figure 1, the input layer corresponds to the extracted values from each of the raster layers (Sentinel2, clay, etc.) from a 21×21 pixel window, (Figure 2). The next layer receives this array and applies 16 different filters (i.e. a 3×3 pixel kernel moving throughout the array), and this process is the one that gives the 'convolution' name to this kind of network. As this convolution uses a moving window, then the resulting output is a smaller 19×19 stack.

The next step involves a batch normalization followed by a dropout layer (30% of the information is left aside), which has been shown to improve the generalization of the model (Srivastava *et al.*, 2014). The next steps involve a second convolution + normalization + dropout layers which increases the 'depth' of the network (i.e. more layers). Deep networks have been shown to produce better results when using convolutional processes (Krizhevsky *et al.*, 2012).

Finally, after a third dropout, a 'flatten' layer process is used to deliver a one-dimensional output of a single value of wheat yield. The architecture of a deep network represents a whole area of research by itself, and the one used in this study follows previously used approaches that achieved some success when predicting crop yield (You *et al.*, 2017). More information about neural network layers and filters can be found in Chollet and Allaire (2018).

Sampling design and modelling approaches

In order to optimize the samples and the observed yield variability, a conditioned Latin hypercube design (Minasny and McBratney, 2006; Roudier, 2011) was employed to select the sampling locations. Within each field, a fixed percentage (2%) of the total number of pixels (5 m resolution) was sampled. Hence, using this design, each field will have a balanced contribution to the model in terms of within-paddock variability (Table 1).

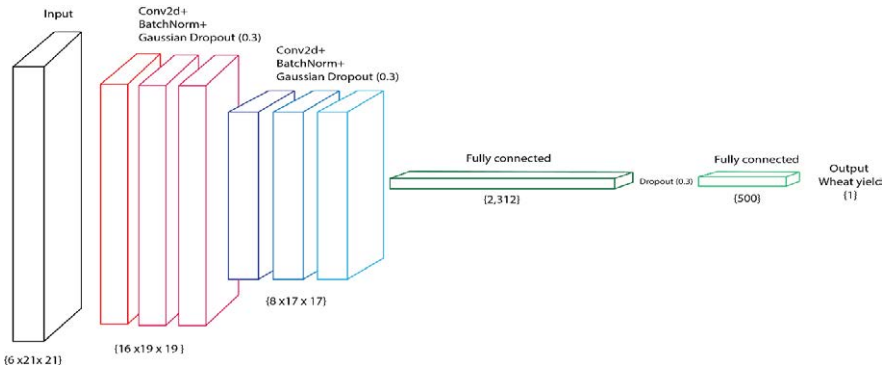


Figure 1. CNN architecture used for wheat yield prediction, numbers in brackets represent dimensions of each array.



Figure 2. Example of eight sampling locations (top). Detail of the same locations with central pixel highlighted in white (middle) and detail of two sample locations with a 21×21 pixels contextual sample (bottom).

Three different sets of information for both BRT and CNNs were used, resulting in six different modelling approaches. The main difference between approaches was the use of either Landsat 8 or Sentinel as the 'NDVI' input and the use of DualEM. The first approach used all the observed fields with a total of 35,948 samples with Landsat 8 imagery as the 'NDVI' input, clay, total nitrogen and DEM (5 covariates) and was simply named 'RegTrees' or 'CNN' depending on the type of modelling framework (Table 1). The second approach (with a total of 16,132 samples) employed Sentinel 2 imagery instead of Landsat 8 (5 covariates as well) and was named either 'RegTrees-Sent' or 'CNN-Sent' and finally, the third (with a total of 7,961 samples) used Sentinel 2 imagery plus DualEM information, this final approach was named as 'RegTrees-DualEM' or 'CNN-DualEM' (Table 1). As mentioned, CNN modelling requires contextual information, hence for each of the selected locations, a window of 21×21 pixels was selected as input for the CNN modelling and only the central pixel of this matrix was used as input for the regression trees. In this way, both training datasets have the same number of samples (Figure 2).

Table 1. Different model designs and respective number of samples.

Model type	Covariates	Years	Number of fields	Number of samples
CNN	5	2013-2017	47	35,948
CNN-Sent	5	2016-2017	10	16,132
CNN-DualEM	6	2016-2017	7	7,961
RegTrees	5	2013-2017	47	35,948
RegTrees-Sent	5	2016-2017	10	16,132
RegTrees-DualEM	6	2016-2017	7	7,961

All the analyses were undertaken using Python (for the raster extraction using Google earth engine) and R for the sampling design and modelling. BRT were implemented using R package 'Cubist' (Kuhn *et al.*, 2014b) and CNN modelling was implemented using the R package 'Keras' with a 'TensorFlow' back-end (Chollet and Allaire, 2018).

Model performance

In order to compare the different modelling procedures, four metrics were employed. These were the well-known coefficient of determination (R^2), root mean squared error (RMSE), Lin's concordance correlation coefficient and a bias statistic (mean predicted values – mean observed values). All yield information was scaled between 0 to 1 for privacy reasons.

Results

It was observed that as the resolution (Landsat 8 < Sentinel 2) and complexity of the modelling (RegTrees < CNN) increased, the performance increased accordingly (Figure 3).

These results seem consistent with the amount of information included in each of the modelling approaches i.e. the BRT approach uses just the central pixel of a CNN observation; this is 1 pixel vs 441 pixels (21×21 window). In terms of RMSE, the results agreed with previous attempts at modelling (at bigger spatial scales). Even though, as far as the authors know, there is no comparable study modelling within-paddock forecasting, You *et al.* (2017) reported a substantial improvement of accuracy when using CNNs compared with regression trees at county scale in the U.S.A.

The results also showed the benefit of farm-owned data such as electromagnetic induction studies (DualEM), to final predictions. There was an obvious improvement for both BRT and CNNs compared with models without DualEM data in terms of higher R^2, less bias and less dispersion of the predictions (Figure 3). In terms of RMSE for the different modelling approaches, the differences were not as obvious, with all the models presenting RMSEs values around 10 to 15% of the total yield production (i.e. for a field with a yield of 4 t/ha an error of about 400~600 kg/ha). In terms of total yield, this result seems useful, but still a rough estimate.

In order to estimate the quality of the predictions for decision-making purposes, a spatial assessment is needed. Figure 4 shows an example of the spatial predictions derived from the CNN-Sent model for a randomly selected field (from all the fields of Figure 3). There are spatial patterns evident that are captured by the model which details the potential yield and its distribution. It is important to understand that the current modelling approach only uses information from the previous harvest (autumn-winter imagery) excluding harvest information from that particular field.

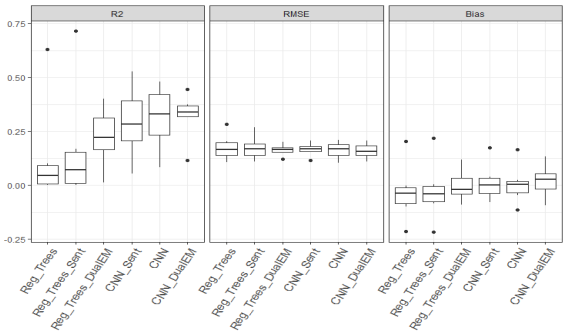


Figure 3. Performance of the different modelling approaches.

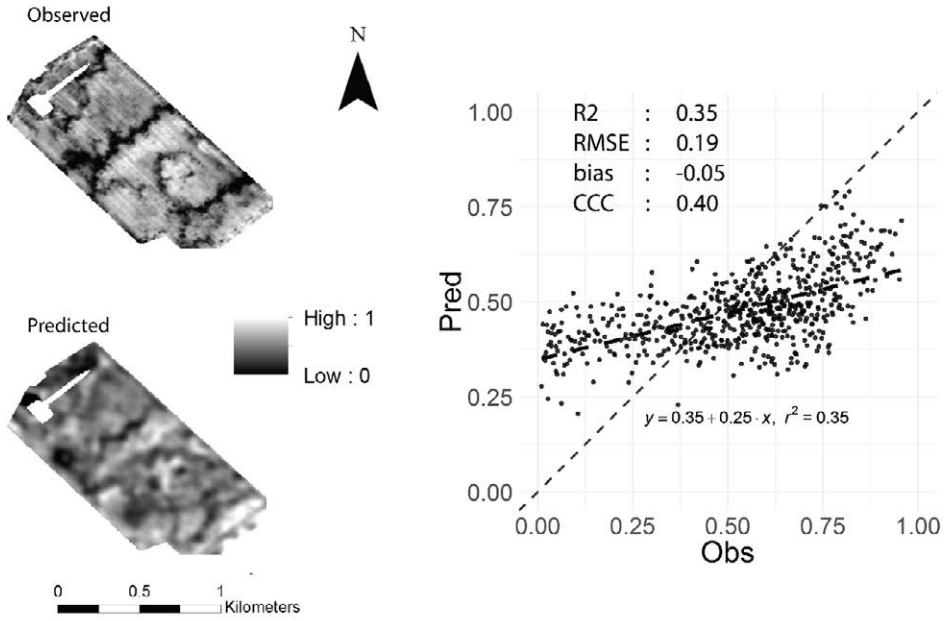


Figure 4. Spatial prediction using the CNN-Sent model for a randomly selected field in the data set (yield data standardised to 0-1 for privacy reasons).

Finally, considering the limited information included in the current models, within-paddock forecasting using publicly available data shows potential, especially if a more complex modelling approach (e.g. deep networks) such as CNNs is used. Future work will involve the use of data with greater yield variability (inclusion of more farms from different agro-climatic regions) and the addition of more layers in the modelling process, such as gamma-radiometric surveys and climate (temperature and rainfall) data.

Conclusions

The study showed encouraging results for within-paddock wheat yield forecasting using only machine learning techniques. The model performance increased as more detailed information was used (higher spatial resolution) and when contextual information was considered in the modelling exercise. However, there is still room for improvement in the use of CNNs coupled with high-resolution yield information.

References

Baruth, B., Royer, A., Klisch, A., & Genovese, G., 2008. The use of remote sensing within the MARS crop yield monitoring system of the European Commission. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Beijing, 37, 935-940.

Bauer, M. E., 1975. The role of remote sensing in determining the distribution and yield of crops. Advances in Agronomy, 27, 271-304.

Bureau of Meteorology, 2019. Australian Government. http://www.bom.gov.au/climate/data (last accessed February 2019).

Chollet, F., & Allaire, J. J., 2018. Deep Learning with R. New York, USA: Manning Publications Company, 360p.
Dempewolf, J., Adusei, B., Becker-Reshef, I., Hansen, M., Potapov, P., Khan, A., *et al.,* 2014. Wheat yield forecasting for Punjab Province from vegetation index time series and historic crop statistics. Remote Sensing, 6, 9653-9675.
Doraiswamy, P.C., Hodges, T. and Phinney, D.E., 1979. Crop yield literature for AgRISTARS crops corn, soybeans, wheat, barley, sorghum, rice, cotton and sunflowers. AgRISTARS Technical Report SR-L9-00405, Houston, Texas: Lockheed Electronics Company, Inc., 105 p.
Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., *et al.,* 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sensing of Environment, 120, 25-36.
Erickson, J., 1984. The LACIE experiment in satellite aided monitoring of global crop production. In: G. M. Woodwell (Ed.) The role of terrestrial vegetation in the global carbon cycle: measurement by remote sensing, Chichester, UK and New York, USA: John Wiley & Sons Ltd. pp 191-217.
Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., *et al.,* 2007. The Shuttle radar topography mission. Reviews of Geophysics, 45, RG2004, pp33, doi:10.1029/2005rg000183
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18-27.
Hasan, M.M., Croke, B.F.W. & Karim, F., 2019. Spatial and seasonal variations and inter-relationship in fitted model parameters for rainfall totals across Australia at various timescales. Climate, 7, 4-15. doi.org/10.3390/cli7010004
Irons, J. R., Dwyer, J. L., & Barsi, J. A., 2012. The next Landsat satellite: The Landsat data continuity mission. Remote Sensing of Environment, 122, 11-21.
Isbell, R. F., 2002. The Australian soil classification. CSIRO Publishing, 141p.
Krizhevsky, A., Sutskever, I., & Hinton, G. E., 2012. Imagenet classification with deep convolutional neural networks. In: NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, Nevada, USA: Curran Associates Inc., pp 1097-1105.
Kuhn, M., Weston, S., & Coulter, N., 2014a. C50: C5.0 Decision trees and rule-based models. R package version 0.1.0-19. https://CRAN.R-project.org/package=C50. (last accessed February 2019).
Kuhn, M., Weston, S., Keefer, C., & Coulter, N., 2014b. Cubist: Rule- and instance-based regression modeling. https://CRAN.R-project.org/package=Cubist. (last accessed February 2019).
LeCun, Y., Boser, B. E., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. E. *et al.*, 1989. Back-propagation applied to handwritten zip code recognition. Neural Computation, 1, 541-551.
Minasny, B., & McBratney, A. B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Computers & Geosciences, 32, 1378-1388.
Quinlan, J. R., 1986. Induction of decision trees. Machine Learning, 1, 81-106.
Rossel, R. V., Chen, C., Grundy, M., Searle, R., Clifford, D., & Campbell, P., 2015. The Australian three-dimensional soil grid: Australia's contribution to the GlobalSoilMap project. Soil Research, 53, 845-864.
Roudier, P., 2011. clhs: a R package for conditioned Latin hypercube sampling. Retrieved from http://code.scenzgrid.org/index.php/p/clhs (last accessed Febuary 2019).
Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, 15, 1929-1958.
You, J., Li, X., Low, M., Lobell, D., & Ermon, S., 2017. Deep gaussian process for crop yield prediction based on remote sensing data. In: AAAI-17 Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, USA: Association for the Advancement of Artificial Intelligence, pp 4559-4565.

Spatial analysis of mycotoxins in stored grain to develop more precise management strategies

R. Kerry[1*], B. Ingram[2], E. Garcia-Cela[3] and N. Magan[3]
[1]*Department of Geography, Brigham Young University, Provo, UT, USA; ruth_kerry@byu.edu*
[2]*Cranfield Soil and Agrifood Institute, Cranfield University, Cranfield. MK43 0AL, United Kingdom*
[3]*Applied Mycology Group, Cranfield University, Cranfield. MK43 0AL, United Kingdom*

Abstract

Mycotoxins are harmful to human and animal health so legislative limits in stored cereals must not be exceeded. The spatial distribution of many different mycotoxins in grain silos is largely unknown. Spatial analysis is needed for precise management of stored grain to minimize waste, prevent rejection of lots and maximize profit. This study re-analysed existing datasets where spatial analysis was hindered by highly-skewed data. Here the normal score transform was used. Dataset 1: Cross-variograms and bi-variate local Moran's I (LMI) analysis showed spatial structure and co-variation in two toxins. Dataset 2: 3D spatial analysis showed spatial structure was stronger and concentrations were higher in the outer layers of the grain and where moisture collects under the influence of gravity.

Keywords: mycotoxins, grain storage, spatial analysis, variogram, local Moran's I

Introduction

Precision farming is a production line with many stages. One of the later stages in this line is grain storage before it is sold or fed to animals. Mycotoxins are natural toxic secondary metabolites produced by fungi in a range of staple crops both pre- and post-harvest. While in-field variability in fungal development could be addressed with precise fungicide application, cost-effective methods to identify areas with fungal growth in the field still need to be developed. This, however, would not solve the problem of mycotoxin production in stored grain. As mycotoxins are harmful to human and animal health, strict legislative limits for the maximum concentrations in stored cereals exist, which must not be exceeded to avoid rejection of the crop. Mycotoxins are usually measured after harvest by collecting one large bulked sample composed of several smaller samples taken systematically from the stored grain to give an average contamination level. Previous studies (Biselli *et al.*, 2008 and Rivas Casado *et al.,* 2009) have used variograms to characterize the spatial variability of mycotoxins to inform how this sub-sampling for bulking should be conducted. However, applying the principles of precision agriculture to this problem suggests that a better use of systematic sampling data from stored crops would be to characterize and improve the understanding of the spatial variation of toxins and to identify the most heavily contaminated areas. Heavily contaminated grain could then be removed or remedial drying regimes implemented to reduce crop contamination, and minimise the risk of the crop being rejected. Effective real time monitoring of grain respiration using CO_2 sensors placed in an informed way in a grain silo could allow early detection of the initiation of spoilage and mycotoxin contamination and the identification in 3D of pockets requiring remedial action. This would meet some key precision agriculture goals by minimizing waste and maximizing profits. This study re-analysed two existing datasets. The first measured deoxynivalenol (DON) and ochratoxin A (OTA), at 100 locations in 2D in a 26 t truck of previously stored wheat (Biselli *et al.*, 2008). Rivas Casado *et al.* (2009) found spatial structure in DON but not in OTA using variograms. Data for the latter are highly skewed. While there is no explicit assumption of normality for computing typical method of moment (MoM) variograms, they are based on variances and can be

DOI 10.3920/978-90-8686-888-9_89, © Wageningen Academic Publishers 2019

sensitive to a few large outliers or asymmetry in the distribution (Webster and Oliver, 2007). Typical methods of dealing with asymmetry in data for variogram computation involve transforming the data to logarithms, removing outliers or computing normal scores (Webster and Oliver, 2007).
The global Moran's I (GMI), another method of investigating spatial autocorrelation, standardizes data to a mean of zero and unit variance and this can be problematic when data are highly skewed. However, in using the mean and standard deviation of a local neighbourhood to standardize data, the local Moran's I (LMI) is less likely to encounter problems with skewed data. The aim in re-analysing this dataset was to determine if dealing with the non-normality in the OTA data could reveal spatial structure in the OTA data and spatial association with DON.
The second dataset measured fumonisin (FB_1) and (FB_2) contamination of maize in 3D (Rivas Casado *et al.*, 2010). The initial study found no spatial structure in the data, which were skewed and variograms for individual layers were based on only 50 data points. Webster and Oliver (2007) recommend at least 100 data points to compute reliable MoM variograms. The aim here was to re-analyse this data set to determine if dealing with the skewed data could reveal any spatial structure and if Moran's I analysis could detect this, in individual layers, as it may not need as many data points as MoM variogram computation.

Materials and methods

Data collection

Dataset 1: A 2.5×10 m truck containing wheat was sampled in the horizontal plane at a spacing of 0.5 m producing a grid of 100 points (Biselli *et al.*, 2008). A five-aperture probe sampler was vertically inserted into the grain at each grid point to take a single incremental sample containing grain from five depths. Each sample was mixed and subsampled before the DON and OTA concentrations (μg/kg) were measured. Full details of the lab procedures used are given in Biselli *et al.* (2008). As concentrations were averaged with depth, there was no insight into how DON or OTA changed with depth.
Dataset 2: Data were collected from a bed of maize in a large grain store with an exposed surface area of ~3,800 m^2 and a height of ~10 m. The samples were taken in three planar layers inclined at 45°, parallel to the open face of the bed, providing a three-dimensional grid of points (Rivas Casado *et al.*, 2010). There were 50 points in each layer on a 10 by 5 point grid with a 0.5 m spacing. There was also a spacing of 0.5 m between each of the layers (Figure 1A). Grain samples were taken using

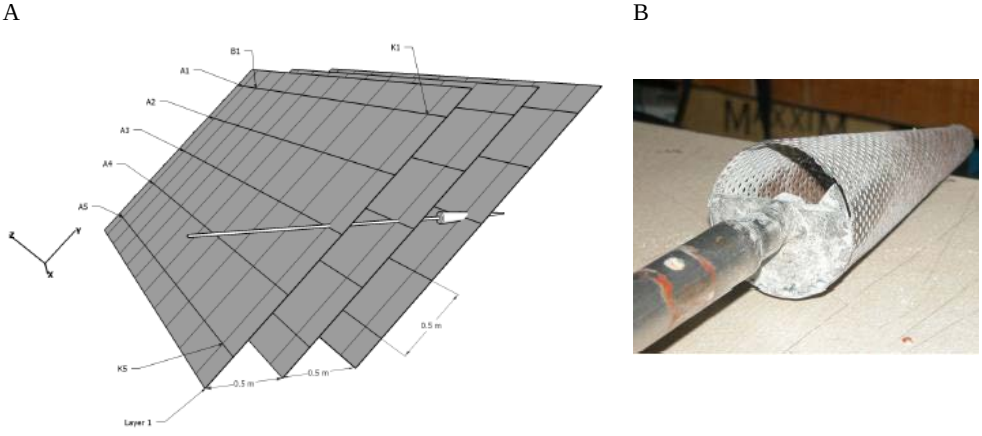


Figure 1. (A) Sampling scheme and (B) sampling probe.

a probe with a conical chamber at the tip that could be opened and closed using a lever at the end of the shaft (Figure 1B), collecting about 200 g of grain. A 0.5 m square metal frame was placed on the grain face to locate the sampling points where the probe would be inserted. Samples from the three different layers were taken at each point before moving the frame to minimize disturbance of the grain (Rivas Casado *et al.*, 2010), nevertheless there may be some inaccuracies resulting from mixing of the grain between layers during probe insertion

The FB_1 and FB_2 contents in maize were analyzed by the following method, full details of which are given in Rivas Casado *et al.* (2010). Samples were ground, passed through a 1 mm sieve and homogenized. A solution of acetonitrile:methanol:water (25:25:50, v/v/v) was used to extract fumonisins with a grain to solution ratio of 1:5. Following further sample preparation concentrations of FB_1 and FB_2 (µg/kg) were determined by liquid chromatography-mass spectrometry, (Thermo-Fisher Scientific, San Jose, CA, USA).

Data analysis

The OTA (Dataset 1) and FB_1 and FB_2 data (Dataset 2) were transformed to normal scores (NST). Normal score ties were broken using the local mean of the eight nearest neighbors. Omni-directional MOM variograms in two dimensions with a lag separation of 0.5 m were computed and modelled for Dataset 1 and each layer of Dataset 2 using SpaceStat (Jacquez *et al.*, 2014). Residual maximum likelihood (REML) variograms were also computed for each layer of Dataset 2 using GeoR (Ribeiro *et al.*, 2001). A 2D cross-variogram with a 0.5 m lag was computed between DON and OTA NST using SGems (Remy *et al.*, 2009). Omni-directional variograms in 3D with a lag separation of 0.5 m were computed for FB_1 and FB_2 in SGems (Remy *et al.*, 2009).

Univariate GMI and LMI analysis (Anselin, 1995) was done using SpaceStat for DON and OTA from Dataset 1 and for FB_1 and FB_2 in individual layers for Dataset 2. The global statistic determines if there is significant clustering of similar or dissimilar values indicating positive or negative spatial autocorrelation, respectively, compared to a random distribution. The LMI statistic computes the statistic for a moving window to determine if a given observation is similar or different to the mean of neighbouring values (here within 1 m). Maps were produced to determine whether points were part of significant (P=0.05) clusters of low (LL- low values surrounded by low values) or high (HH – high values surrounded by high values) values, were significant spatial outliers (LH – low values surrounded by high values or HL – high values surrounded by low values) or were not significantly (NS) different from a random distribution. Monte Carlo simulation was used to determine the P-value of the LMI statistic for each observation. The number of tests was large and the risk of false positives was increased, so the Simes correction (Jacquez *et al.,* 2014) for multiple testing was used. Bivariate GMI and LMI analysis between DON and OTA NST and between different layers for FB_1 and FB_2 in Dataset 2 indicates if clusters in the two variables or layers are collocated by using the standardized values of the second variable as the neighbours for each point in the first variable. This means that HH and LL clusters are significant clusters with high or low values, respectively of both variables or in both layers. Any HL or LH clusters indicate significant clusters of high values in the first variable and low values in the second variable, or the reverse, respectively. Finally, 3D LMI analysis of FB_1 and FB_2 was performed following the approach of Kerry *et al.* (2017).

Results and discussion

Dataset 1: Figure 2 shows the histograms for DON and OTA. DON had a relatively normal distribution with an elevated skew caused by a large outlier and slight asymmetry (Figure 2A). DON had a large range of values of up to ~3,000 µg/kg and there were no points with 0 µg/kg. In contrast, OTA (Figure 2B) showed a highly skewed distribution, approximating a Poisson distribution, with a large number of points with 0 µg/kg and the highest concentration being ~8 µg/kg.

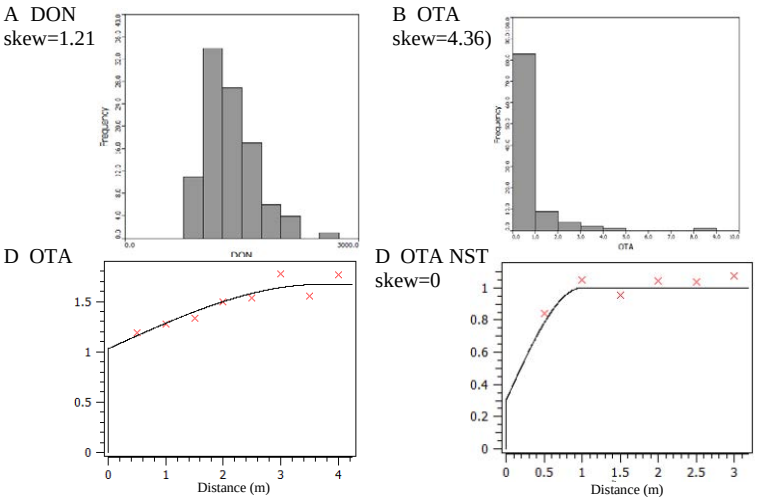


Figure 2. Histograms of (A) DON, (B) OTA and variograms of (C) OTA and (D) OTA NST.

Table 1 shows the variogram parameters for DON and OTA. The raw DON values showed good spatial structure with a variogram range of 4 m and nugget:sill ratio of about 13%. The variogram for OTA (Figure 2C) however, had a similar range but a large nugget effect showing a large proportion of spatially random variation. However, the variogram of OTA NST (Figure 2D) had a nugget:sill ratio of 30% and the range decreased to 1 m. Rivas Casado *et al.* (2009) suggested that there was no spatial structure in OTA but these findings show more spatially random variation in OTA than DON and smaller areas with similar values. The GMI values (Table 1) backup these findings. Raw DON values showed strong (0.468) positive and significant (P=0.001) spatial autocorrelation whereas raw OTA values showed negative spatial autocorrelation that was not significant (P=0.494). The OTA NST data showed moderate positive and significant spatial clustering (P=0.001).

Figure 3A and C show the spatial patterns of the raw values of DON and OTA. These show most spatial variation occurring along the long axis of the truck which is consistent with the method of loading the truck. Indeed, examination of directional variograms (not shown) for the long and short axis of the truck showed spatial structure for the long axis with parameters almost identical to the omni-directional variogram and predominantly random variation across the short axis of the truck. Figure 3B and d show how DON and OTA values are clustered in space using the univariate LMI. These maps backup the findings of the variogram and GMI analysis in that they show larger significant clusters of HH and LL points for DON than for OTA. The last row in Table 1 reports the parameters of a cross-variogram between DON and OTA NST. This showed spatial structure and

Table 1. Univariate and bi-variate GMI values and variogram and cross-variogram parameters for DON and OTA.

Variable(s)	GMI	*P* value	Model	Nugget:sill ratio (%)	Range (m)
DON	0.468	0.001	Spherical	13	4.00
OTA	-0.012	0.494	Spherical	62	3.57
OTA NST	0.124	0.001	Spherical	30	1.00
DON and OTA NST	0.157	0.001	Gaussian	0	3.00

co-variation in the two toxins with patches of similar values with a diameter of ~3 m. The bivariate GMI value (Table 1) also showed significant (P=0.001) spatial association between DON and OTA NST as did the bivariate LMI (Figure 3E). The areas with significant values showed both positive (HH and LL) and negative spatial (HL and LH) association between DON and OTA. Thus, correlation between the two toxins is missed by standard correlation coefficients (Spearman's r<0.1, P>0.05). As DON is usually produced in the field and OTA during storage, the spatial analyses suggest that OTA may well develop from foci of high DON concentrations and spread out from there, hence, the lower values of OTA and smaller clusters and co-location of some OTA clusters with DON clusters. This suggests that developing methods to identify DON contaminated grain in the field or at harvest, before storage using methods such as hyperspectral imaging is an important goal for future research. *Dataset 2:* The histograms in Figure 4A, B show the distribution of FB_1 and FB_2. Both showed marked asymmetry and FB_2 had a larger skew. Variograms (3D) of raw and log transformed data were erratic or pure nugget but 3D variograms for NST FB_1 and FB_2 showed spatial structure with areas of similar values of 4 and 3 m diameter, respectively and ~ 50% spatially random variation (Table 2 and Figure 4C). The effect of skew on data can be more pronounced when datasets are small, so it is not surprising that the MoM variograms for the individual layers were predominantly erratic or pure nugget. Nevertheless, the REML variograms for layer 1 (FB_1 and FB_2) and layer 2 (FB_1) showed some spatial structure as did the MoM variograms for layer 3 (FB_1 and FB_2) (Table 2). These variograms showed a decreasing range with depth of layer from the exposed grain face (Figure 1A) for FB_1 and for FB_2 a similar pattern was apparent but there was no spatial structure in layer

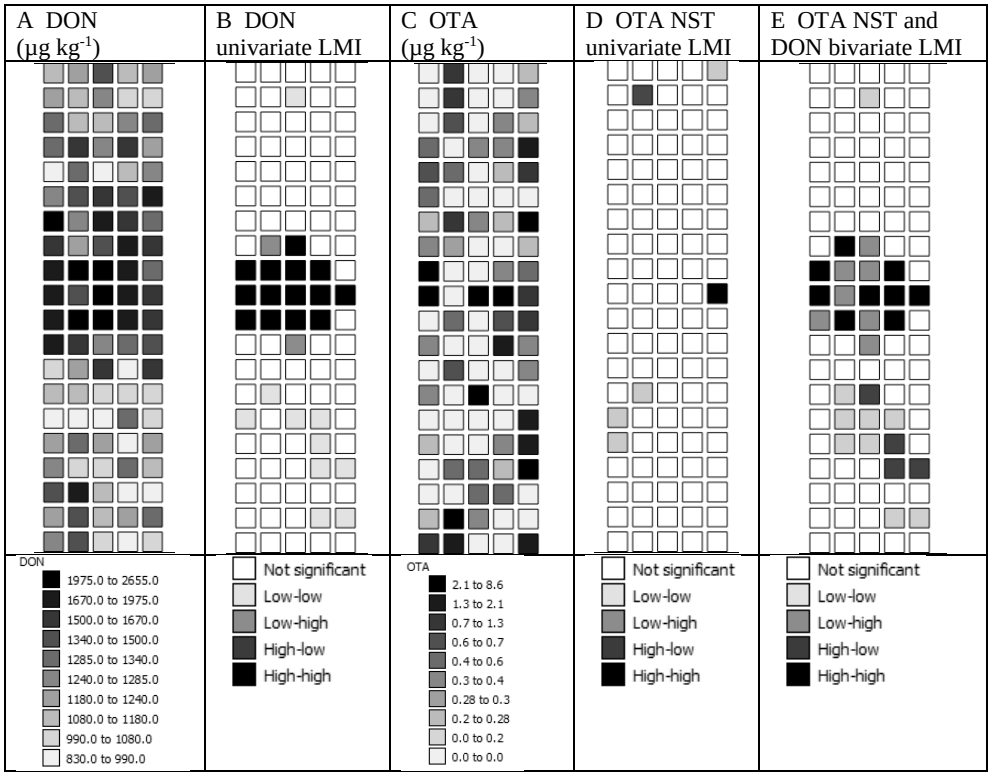


Figure 3. Maps of DON (A) and OTA (C) distribution and clusters from univariate (B, D) and bivariate (E) LMI analysis.

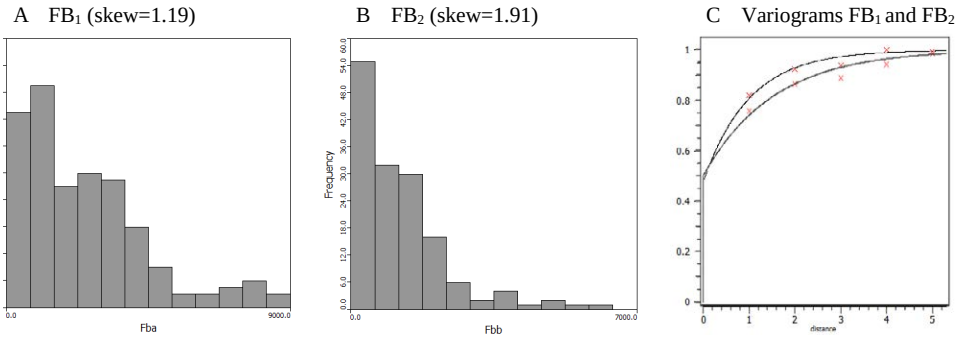


Figure 4. Histograms (A-B) and 3D variograms (C) of Mycotoxins FB_1 and FB_2.

2 and the patches of FB_2 were smaller than those of FB_1 in both layers 1 and 3. This suggests larger patches of both toxins in the outer layers of the stored grain that are more exposed to the air and smaller patches or random variation in the microaerophilic conditions closer to the wall of the silo. The GMI and variogram analysis (Table 2) show large proportions of spatially random variation and the spatial structure present was closest to significant (P=0.05) in layers 1 and 3. The bivariate GMI values between layers showed more spatial association between layers 2 and 3: P=0.003 and P=0.016, for FB_1 and FB_2, respectively. Correlations between FB_1 and FB_2 for each layer and all layers together were strong r>0.9 and 3D LMI maps showed (not shown) significant (P=0.05) HH clusters were in rows 4 and 5 of layers 1 and 2 (Figure 1A) and LL clusters in rows 1 and 2 of layers 1-3. The spatial analysis showed that FB_1 and FB_2 behave and are located similarly. The shorter variogram ranges for FB_2 suggest that it might develop from points of FB_1 contamination. The distribution of HH and LL clusters in the stored grain suggested that toxins develop most in the outer layers at the

Table 2. Univariate and bivariate GMI for NST FB_1 and FB_2 within and between layers and parameters of 2D and 3D variograms for single layers and whole data.[1]

Variable(s)	GMI	*P*-value	Model	Nugget:sill ratio (%)	Range (m)
FB_1 (layer 1)	0.029	0.158	*Gaussian*	*50*	*3.86*
FB_1 (layer 2)	-0.021	0.500	*Spherical*	*67*	*2.78*
FB_1 (layer 3)	0.047	0.054	Spherical	52	1.02
FB_2 (layer 1)	-0.060	0.170	*Spherical*	*67*	*1.98*
FB_2 (layer 2)	-0.009	0.400	---	---	---
FB_2 (layer 3)	0.019	0.164	Spherical	48	0.98
FB_1 layers 1 and 2	0.069	0.115	---	---	---
FB_1 layers 2 and 3	**0.119**	**0.003**	---	---	---
FB_1 layers 1 and 3	0.022	0.326	---	---	---
FB_2 layers 1 and 2	0.035	0.204	---	---	---
FB_2 layers 2 and 3	**0.111**	**0.016**	---	---	---
FB_2 layers 1 and 3	-0.003	0.499	---	---	---
FB_1 all layers (3D)	0.048	---	Exponential	50	4.00
FB_2 all layers (3D)	0.068	---	Exponential	48	3.00

[1] Bold values significant at P=0.05, *Italics* show REML variogram parameters, --- missing/not valid.

base of the grain pile where oxygen supply, and moisture collecting under the influence of gravity, are greatest. These areas of high contamination risk and could be monitored and contamination grain from them could then be discarded to ensure that values in the bulked samples that are taken before sale are below critical thresholds.

Conclusions

Variogram, GMI and LMI analysis of existing data using the NST to deal with highly skewed data allowed detailed spatial analysis that revealed some key findings.
Dataset 1: Variogram, GMI and LMI analysis detected spatial structure for OTA with smaller clusters than DON. Cross-variograms and bi-variate LMI analyses showed structure indicating co-location of DON and OTA clusters. As DON usually develops in the field and OTA during storage, OTA may be developing from foci of high DON concentrations.
Dataset 2: Variograms, GMI and LMI analysis for individual layers and in 3D showed larger clusters and stronger spatial autocorrelation in the outer grain layers with HH clusters at the base and LL clusters at the top of the stored grain. This is consistent with FB_1 and FB_2 tending to develop in moister, more oxygen rich locations at the base and in the outer layers of the stored grain.

References

Anselin, L., 1995. Local Indicators of Spatial Association – LISA, Geographical Analysis, 27, 93-114.

Biselli, S., Persin, M. and Syben, M., 2008. Investigation of the distribution of deoxynivalenol and ochratoxin A contamination within a 26 t truckload of wheat kernels. Mycotoxin Research. 24:2, 98-104.

Jacquez, G. M., Goovaerts, P., Kaufmann, A., and Rommel, R., 2014. SpaceStat 4.0 User Manual: Software for the space-time analysis of dynamic complex systems (4th ed.). Ann Arbor, MI, USA: BioMedware.

Kerry, R. Ingram, B.R., Goovaerts, P., Meza, F. and Gimenez, D., 2017. 3D LISA: A Flexible Program for Calculating the Local Moran's I in 1, 2 or 3D Illustrated with Examples from Soil Science. GeoComputation 2017, http://www.geocomputation.org/2017/papers/63.pdf. (last accessed February 2019)

Remy, N, Boucher, A. and Wu, J., 2009. Applied Geostatistics with SGeMS: A User's Guide. 1st Edition. New York, USA: Cambridge University Press,. 284 pp.

Ribeiro Jr., P.J. and Diggle, P.J., 2001. geoR: A package for geostatistical analysis. R- NEWS, 1: (2), 15-18.

Rivas Casado, M., Parsons, D. J., Weightman, R. N., Magan, N. and Origgi, S., 2009. Geostatistical analysis of the spatial distribution of mycotoxin concentration in bulk cereals, Food Additives and Contaminants, 26 (6), 867-873.

Rivas Casado, M., Parsons, D. J., Magan, N., Weightman, R. M., Battilanic, P. and Pietric, A., 2010. A short geostatistical study of the three-dimensional spatial structure of Fumonisins in stored maize. Mycotoxin Journal, 3 (1): 95-103.

Webster, R., Oliver, M.A., 2007. Geostatistics for Environmental Scientists, 2nd Edition. Chichester, UK: Wiley,. 330pp.

An iterative region growing algorithm to generate fuzzy management zones within fields

C. Leroux[1,2*], H. Jones[2] and B. Tisseyre[2]
[1]*SMAG, Montpellier, France; cleroux@smag.tech*
[2]*ITAP, Montpellier SupAgro, Irstea, Univ Montpellier, Montpellier, France*

Abstract

Management zones are within-field spatial units that are more homogeneous than the rest of the field and that have different factors or values than other areas of the field. While most algorithms tend to delineate fixed management zone borders, it is acknowledged that within-field variations between management zones are more gradual than sharp, mainly because of smooth variations in soil and landscape characteristics. In this work, an approach is proposed to account for the transient observations between management zones by delineating within-field fuzzy management zones. Observations within transition zones could belong to more than one zone, and this perhaps should be accounted for during the delineation process. Theoretical fertilization-based case studies, i.e. high and weak environmental constraints, were considered to show the interest of membership frequencies to refine yield-based management zones when user's expertise and management decisions come into play.

Keywords: fuzzy delineation, management zones, membership frequency, transition areas

Introduction

The delineation of within-field zones is widely reported as a means to improve the management of cropping systems within the scope of precision agriculture. Many studies generated homogenous within-field units, mostly using classification- (Li *et al.*, 2007; Peralta *et al.*, 2015) or segmentation-based (Leroux *et al.*, 2017; Roudier *et al.*, 2008) algorithms. However, in agriculture, management zones are not well-defined spatial units that can be clearly distinguished from other surrounding spatial entities. Indeed, the variations in agronomic information between two management zones are not sharp but rather gradual, mainly because of smooth variations in soil and landscape characteristics across fields. While the borders between management zones are generally considered fixed in space, the presence of transient information between zones would be better taken into account if these borders were defined from a fuzzy perspective (Zadeh, 1964). Indeed, the membership degree of observations inside these transition areas with regard to a specific zone is not clear. These observations could effectively be associated with more than one zone.

As suggested by Paoli *et al.* (2007), a fuzzy delineation of management zones would be of interest to agricultural professionals because those zones could be adapted to their will and expertise, and to specific within-field applications. For instance, a wine producer might be interested in harvesting only the grapes within the top-quality zones, leaving behind those that could belong to transition zones to make sure that the vine-making process is optimized. Jones *et al.* (2016) have already tackled this problem by using simulated datasets by allowing the borders between management zones to either expand or shrink depending on a quantile separation of the data. Multiple studies have also proposed using the fuzzy variant of the classical k-means algorithm to generate management zones (Li *et al.*, 2007; Peralta *et al.*, 2015; Tagarakis *et al.*, 2013). However, the main problem is that, in most cases, the observations' membership degree to each zone is seldom considered, meaning that the fuzzy approach has not much advantage with respect to a more classical k-means algorithm. Note however that Urretavizcaya *et al.* (2013) proposed an approach to account for membership degrees

DOI 10.3920/978-90-8686-888-9_90, © Wageningen Academic Publishers 2019

obtained with a fuzzy c-means algorithm to generate compact areas with significant oenological differences between them.
For the analysis reported here, it is considered that within-field data have already been zoned, i.e. an initial zoning is available. The objective of this study was to test and report the result of an iterative region growing algorithm to evaluate to what extent an initial zoning can be adapted to the user's expertise and management decisions.

Material and methods

Dataset used
The approach was tested on a real within-field yield dataset arising from a farm located near Verdun, in the north-eastern part of France (WGS84: E:5.568, N:49.378). This 10-ha field was cropped in wheat and harvested with a Claas (Harsewinkel, Germany) combine. Yield data were first filtered and the zoning algorithm of Leroux *et al.* (2017) was then applied on these within-field yield data, which led to the generation of three within-field zones (Figure 1). Note that the high-yielding area in the bottom left-hand corner of zone 1 was considered too small to be delineated as a specific zone by the algorithm of Leroux *et al.* (2017). For visualisation purposes, yield data were separated into five quantiles.

An iterative region growing algorithm
Within-field initial zones as shown in Figure 1 were generated by the region growing and merging algorithm proposed in Leroux *et al.* (2017). In simple terms, a set of observations, referred to as the seeds, was selected in homogeneous regions within the field as the core of within-field zones. Observations were aggregated to these seeds following a similarity measure to form the zones (region growing). As the number of resulting regions is generally too high for management purposes, these zones were merged by maximizing an opportunity criterion (region merging). Each observation is necessarily associated to one and only one zone. Interested readers are referred to Leroux *et al.* (2017) for more information.

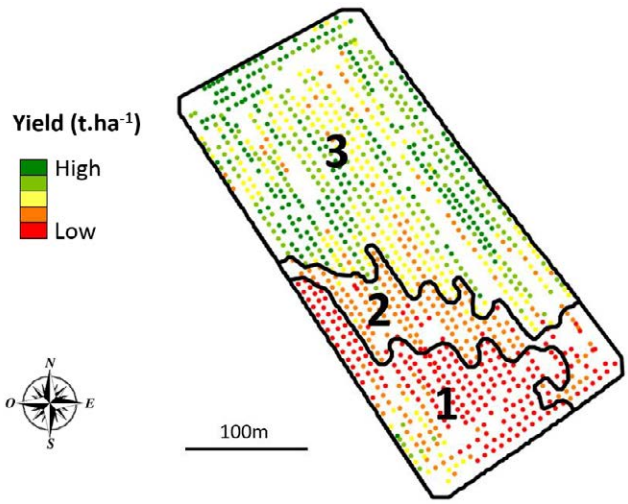


Figure 1. Within-field yield dataset under study with the associated initial zoning. Yield data are separated into five quantiles.

To generate fuzzy maps, the region growing algorithm of Leroux *et al.* (2017) was run multiple times with a fixed number of seeds, i.e. the number of zones in the initial zoning (three seeds in this case study). At each iteration, the region growing algorithm was run with a different set of seeds (each one being selected in one of the initial zones), so that in the end, all the zonings were somehow different but the number of zones were necessarily the same given that the number of seeds was fixed. At the end of all the iterations, each observation was given a membership frequency to each one of the initial zones. When observation x_i always belonged to zone Z_j across iterations, then x_i effectively belonged to the core (or kernel) of Z_j In contrast, if x_i belonged to multiple zones, then x_i was classed transition area (or support of a zone). A step by step description of this approach is presented below:

Algorithm 1: Membership map generation

1. Select randomly one observation within each one of the N initial zones. This set of observations was referred to as the set of seeds S (S_1, S_2,…, S_N).
2. Run the region growing algorithm of Leroux *et al.* (2017) with the previously selected set of seeds.
3. Repeat n times steps 1 to 2.
4. For each observation x_i compute the membership frequency $\mu_{i,j}$ to each zone Z_j as the ratio of the number of times x_i belongs to Z_j to the number of iterations n. Note, $\mu_{i,j}$ lies between 0 (observation x_i never belongs to Z_j) and 1 (observation x_i always belongs to Z_j).
5. Generate the membership frequency map.

In this study, the number of iterations n was set to 30. By construction, each observation was necessarily associated to one and only one zone at each iteration, $\sum_{j=1}^{N} \mu_{i,j} = 1.$.

Adapting the initial zoning to users' expertise and management decisions

The membership map that can be generated with the previously defined algorithm is interesting in itself as it can be used to visualise to what extent an observation belongs to a zone rather than another. The membership frequency of each observation to the initial zones could be used to refine the initial zoning for specific management decisions. Two theoretical fertilization-based use-cases (UC) were considered:

- UC1: The field was subjected to strong environmental constraints and the farmer cannot accept any risk of overdosing. Zones should therefore be limited to their core or kernel.
- UC2: The farmer was using a low-cost fertilizer for which the risks on environment or the impact on yield were limited. Zones could therefore be extended to their support to a certain level.

To provide a zoning that copes with these two different management strategies and users' expertise, the following approach was proposed.

Algorithm 2: Zoning adaptation

1. Choose the zone Z_j for which the border should be refined.
2. Choose a membership threshold α above which observations definitely belong to Z_j.
3. Update the membership frequency of each observation as follows:

$$if\ \mu_{i,j'} > \alpha\ then \begin{cases} \mu_{i,j'} = 1 \\ \mu_{i,j(j \neq j')} = 0 \\ \mu_{i',j'(i \neq i')} = 0 \\ Rescale\ \mu_{i',j}\ so\ that \sum_{j=1}^{N} \mu_{i',j}\ remains\ equal\ to\ 1 \end{cases}$$

4. For each observation x_i compute a weighted yield value $\overline{Y_{Z_j}}$ where $Y_i^w = \sum_{j=1}^{N} \mu_{i,j} \times \overline{Y_{Z_j}}$ is the average yield of the initial zone Z_j.

5. Select one seed per zone as the one with the maximum membership value.
6. Run the region growing algorithm of Leroux *et al.* (2017) using the weighted yield attribute and the previously selected seeds.

Results and discussion

Membership maps and fuzzy zoning

Figure 2 shows the membership frequency of each observation for the three zones of interest. It appears clearly that the spatial extent of membership frequencies for zone 2 was much larger than for the two other zones, meaning that zone 2 could be seen as a transition zone between zones 1 and 3. The core of zone 3 seems to be located on the western, northern, and eastern parts of the field as the southern region of zone 3 (centre of the field) has a colour ranging from purple to blue. It can be seen that this is the region where observations have a lower yield value compared to the rest of zone 3 (Figure 1). Figure 2 highlights that observations near the border of each zone have a significantly lower membership frequency than the remaining observations. Interestingly, observations within the region on the bottom right hand corner of the field, belonging to the initial zone 2, have a similar membership frequency for zones 1 and 2 (Figure 2). The membership frequency map is also interesting as the area of the transition zones can be evaluated.

Figure 3 shows how the zoning could be refined for zones 1 and 3 depending on the management strategies defined by users. As the threshold increases, zones 1 (top) and 3 (bottom) shrink as transient observations are removed from the zones. These new zonings would be more adapted to the first use-case (UC1) previously defined where environmental constraints are strong and the farmer must not take any risk with respect to the fertilization of regions out of zones 1 or 3 (Figure 3, right). In contrast, as the threshold α decreases, zones 1 (top) and 3 (bottom) expand as transient observations are aggregated to the zones, i.e. the zone's support is considered to be part of the zone (Figure 3, left). In this case, the zonings would be more suitable for UC2 as the fertilization-based risk is much lower. Figure 3 demonstrates that the expansion and shrinking of the zones was not simply a buffer applied on the borders as the new zones follow the yield values across the field, e.g. when α is set to 0.7 for zone 3.

Further considerations

Initial zoning is important in the proposed methodology. Initial zoning conditions the number of zones and the possible locations where the seeds can be selected in each zone. Different initial zonings

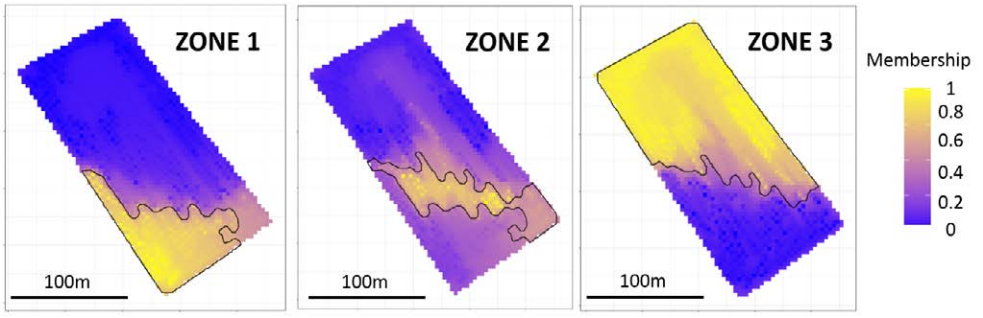


Figure 2. Membership frequency map to each zone of interest. For visualisation purposes, membership values were interpolated via inverse distance weighting with a power distance of 2 on a 2×2 m grid. Interpolated values were solely used for mapping. All the algorithms make use of raw (non-interpolated) data.

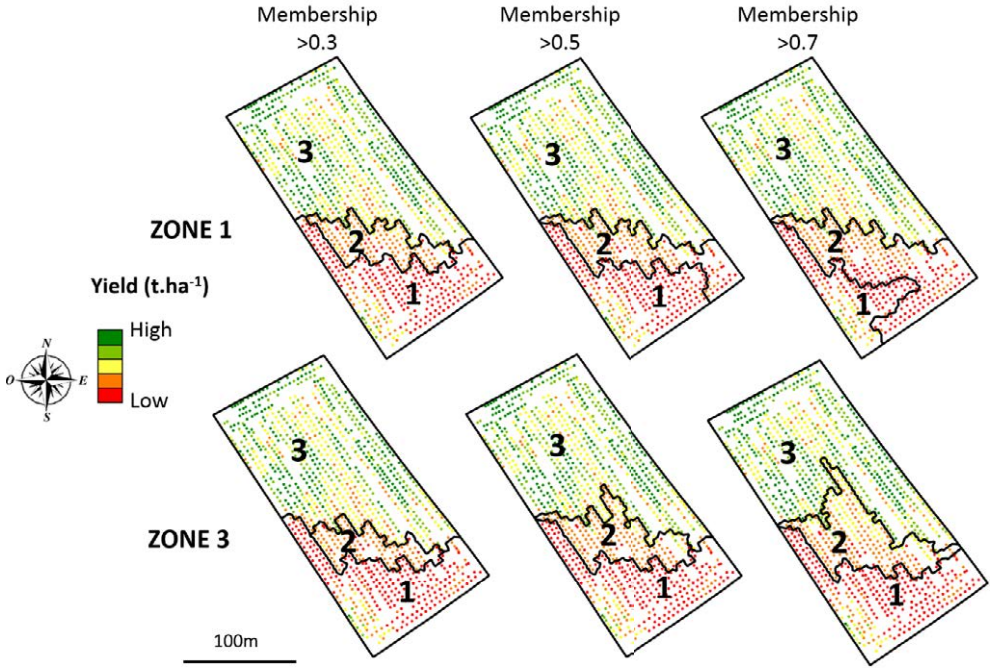


Figure 3. Adaptation of the zoning to specific management decisions. Zones were either expanded or shrunk depending on the membership threshold that was chosen. Plots at the top represent zoning with respect to zone 1. Plots at the bottom represent zoning with respect to zone 3.

would necessarily lead to different fuzzy zonings. The initial zoning, derived from the approach of Leroux *et al.* (2017), was motivated by operational constraints and was considered reliable in this case. The high-yielding region in the south-west portion of the field was effectively considered too small to be delineated.

The concept of fuzzy zones has been tackled in this study with a frequentist or probabilistic approach. Indeed, the membership of each observation x_i to a zone Z_j was defined as the number of times x_i belonged to Z_j. This frequency membership was different from the more common degree membership that can be obtained with a specific fuzzy algorithm such as the fuzzy k-means. Further work should be carried out to evaluate the similarities or differences between these two approaches.

The membership frequency maps presented in this study were obtained after the application of an iterative region growing algorithm, which used seeds randomly selected within the initial zones Z_j. It is clear that the choice of the algorithm to generate the zones and the way seeds are selected within the zones will have an impact on the membership frequency of each observation to the zones. Other seed sampling processes might be considered. For instance, the seed selection could be constrained to regions near the zone border. This would certainly allow zones to expand further (zones would have a larger support). Another possibility could be to select seeds with a yield value near the centre of the yield distribution of each zone to ensure that noisy observations are not selected.

For now, no rules have been set up for the selection of the membership threshold α. A set of maps with different thresholds was shown to evaluate how the zoning could be adapted to specific management decisions. The threshold should be chosen in accordance with the user's expertise in the field and with respect to the management decisions to be made. There might be cases where some zones

have multiple cores, i.e. transition areas are not solely at the border between zones but also within the zones. It might be a problem to adapt the zoning as there would be a need to account for these multiple cores. These multiple cores would however be seen relatively easily on the membership frequency map. If such phenomenon happens, there might be a need to work on the initial zoning to make sure that the zone with multiple cores is separated into multiple zones.
For a given membership threshold α, the proposed approach generates one map per zone of interest. As the number of zones increase, the number of maps will consequently increase, which will make the reading and interpretation of the zoning more difficult by end users. Further work will have to investigate how to combine these maps to facilitate the decision-making process.
Finally, fuzzy zonings would also benefit from quality criteria that would help assess how well the zones were treated. Even if zones considered homogeneous, it is clear that the natural intrinsic variability of plants that go beyond the environmental and management information used to create the zones, and will thus lead to some treatment errors. It must be noted that this criterion will help to evaluate the application error in the small region in the bottom left hand corner of zone 1, which was considered too small to be delineated in the initial zoning.

Conclusions

This investigation used fuzzy set theory to improve the delineation of management zones. Maps of membership frequency were built with an iterative region growing algorithm to evaluate to what extent observations belonged to specific within-field zones. These membership frequencies were then used to refine the borders of management zones so as to cope with theoretical fertilizer-based case studies where environmental constraints were either strong or weak. So far, the membership threshold above which observations were considered to be part of the core of the zones was not defined automatically nor by expertise. Maps with different threshold settings were simply displayed to assess how the zoning could be possibly adapted. Future work will combine the multiple fuzzy zonings that were generated to facilitate the decision-making process.

References

Jones, H., Guillaume, S., Loisel, P., Charnomordic, B., & Tisseyre, B., 2016. Generation of Plateau-Approximated Fuzzy Zones. In Bailly J-S, Griffith, D, Josselin D.(Eds.), Proceedings of the Conference on Spatial Accuracy, Montpellier, France. pp 31-37.

Leroux, C., Jones, H., Clenet, A., & Tisseyre, B., 2017. A new approach for zoning irregularly-spaced, within-field data. Computers and Electronics in Agriculture, 141, 196-206.

Li, Y., Shi, Z., Li, F., & Li, H.Y., 2007. Delineation of site-specific management zones using fuzzy clustering analysis in a coastal saline land. Computer and Electronics in Agriculture, 56, 174-186.

Paoli, J.N., Strauss, O., Tisseyre, B., Roger, J.M., & Guillaume, S., 2007. Spatial data fusion for qualitative estimation of fuzzy request zones: Application on precision viticulture. Fuzzy Sets and Systems, 158, 535-554.

Peralta, N.R., Costa, J.L., Balzarini, M., Franco, M.C., Córdoba, M., & Bullock, D., 2015. Delineation of management zones to improve nitrogen management of wheat, Computer and Electronics in Agriculture, 110, 103-113.

Roudier, P., Tisseyre, B., Poilvé, H., & Roger, J., 2008. Management zone delineation using a modified watershed algorithm. Precision Agriculture, 9, 233-250.

Tagarakis, A., Liakos, V., Fountas, S., Koundouras, S., & Gemtos, T.A., 2013. Management zones delineation using fuzzy clustering techniques in vines. Precision Agriculture, 14, 18-39

Urretavizcaya, I., Santesteban, L. G., Tisseyre, B., Guillaume, S., Miranda, C., and Royo, & J.B., 2013. Oenological significance of vineyard management zones delineated using early grape sampling, Precision Agriculture, 9, 101-103.

Zadeh, L., 1964. Fuzzy sets. Information and Control, 8, 338-353.

Investigating the harmonization of highly noisy heterogeneous datasets hand-collected over the same study domain

L. Pichon[1*], C. Leroux[1,2], V. Geraudie[3], J. Taylor[1] and B. Tisseyre[1]
[1]*ITAP, Univ Montpellier, Montpellier SupAgro, Irstea, Montpellier, France; leo.pichon@supagro.fr*
[2]*SMAG, Montpellier, France*
[3]*Pellenc, Pertuis, France*

Abstract

The objective of this paper is to propose an approach to harmonise noisy spatial data acquired by different operators using (low-cost) hand-held sensors over the same spatial domain. In such cases, datasets need to be harmonised first to be comparable before decision making. This work proposes a methodology to address this issue in the case of nested and noisy spatial data. First, it proposes the implementation of a non-parametric test of Kolmogorov-Smirnov to determine if harmonisation is needed. Then, it proposes an aspatial harmonization method based on standardization. The method was applied on grape sugar content datasets collected by 2 hand-held spectrometers. Results showed that harmonizing a less confident dataset is interesting solely if the size of the trusted one is too small.

Keywords: sugar content, spectrometer, precision viticulture, crowdsourcing

Introduction

The use of hand-held measurement systems, particularly smartphones, is flourishing in agriculture. This is particularly true for perennial crops, such as vines, where many operations are still manual and data collection by operators with hand-held sensors is feasible (Aquino *et al.*, 2018, Aquino *et al.*, 2015, Geraudie *et al.,* 2010). Spatial observations collected with such systems may be abundant and are following an increasing trend as the acquisition is simple and cost and time effective. In a production context, this permits the acquisition of multiple datasets by several operators simultaneously over the same spatial domain *D*, ranging from large (e.g. watershed) to small (e.g. sub-field/plot) domains, but sometimes under varying acquisition conditions, e.g. different operators, sensor model or calibration. These datasets have unique features:

1. Spatial observations are usually very noisy because measurements are performed on a small spatial support and the short range variability is often high on perennial crops (from one leaf/fruit to another or from one plant to another),
2. The source of variability may be caused by the operator, who may be more or less aware of how the acquisition conditions affect the quality of the measurements. The resulting dataset quality may differ from one operator to another depending on the level of training, skill or attention paid to the calibration procedure, if such a procedure is necessary,
3. The data collected by several operators over the same domain are generally intimately intertwined (nested). Operators effectively do not necessarily know the location of measurements performed by other operators.
4. No reference dataset is generally available. In a production context, it is therefore not possible to assess the quality of data collected with hand-held measurement systems by comparing them to a reference.

The resulting datasets are not comparable and need to be harmonized before used in decision-making. While some research has addressed the general issue of data harmonization (Brenning *et al.*, 2008, Sams *et al.*, 2017), very few studies have directly considered the case of noisy spatial datasets collected over the same spatial domain *D*, such as agricultural crowdsourcing data (Minet

DOI 10.3920/978-90-8686-888-9_91, © Wageningen Academic Publishers 2019

et al., 2017). Some studies have proposed advanced data fusion approaches (Ge *et al.,* 2014) but generally considering that reference datasets are available. This study proposed a methodology to address this specific issue. The proposed methodology was applied on real grape sugar content data collected by two different operators with a hand-held spectrometer each.

Materials and methods

General approach

Consider several spatialized datasets (Y_1, Y_2,…Y_i) relating to the same agronomic phenomenon and collected by different operators using hand-held measurement systems on the same study domain *D* and with similar spatial distribution. These datasets are considered as estimations of the same regionalized variable (agronomic phenomenon) whose first and second order moments are expected to be consistent over the domain *D*. The set of conditions c_1, c_2… c_i under which these datasets were collected correspond to the characteristics of the operator (level of experience, level of attention, interest, etc.) and the sensor (calibration, model, etc.). The proposed approach assumes that there is at least one set of conditions corresponding to a properly calibrated sensor and to the best possible conditions of acquisition in a given production context. The dataset acquired under these conditions is considered as the best possible dataset (BPD) even though its quality cannot be assessed properly. In this paper, this set of conditions is noted c_1 and the resulting dataset is noted Y_1. In a concern for simplicity, the rest of the approach is described for two datasets Y_1 and Y_2. Note however that, the approach remains transferable to *i* datasets (Y_1, Y_2, …Y_i).

The objective of this paper is to propose a simple harmonisation approach for highly noisy heterogeneous datasets hand-collected over the same study domain *D*. This paper aims at testing this simple approach on a real case study and to evaluate its relevance for a conjoint use of Y_1 and Y_2 in a decision support context. This relevance was evaluated through different scenarios.

Harmonization approach

First, the approach proposes to evaluate whether the conjoint use of Y_1 and Y_2 requires a pre-processing harmonization step. As a first approach a simple method based on attribute distribution was proposed. More sophisticated approach based on the covariance between Y_1 and Y_2 was not considered in this first approach mainly because of the noise and the number of available measurements which make it difficult to perform a proper spatial covariance analysis. Considering the common characteristics of Y_1 and Y_2 (same agronomic phenomenon observed, same domain *D* studied, similar spatial distribution), it is expected that their attribute distributions have to be similar. If not, it is considered necessary to harmonize Y_2 with Y_1. The Kolmogorov Smirnov statistical test (K–S test) was proposed to compare Y_2 and Y_1 distributions. Harmonization was considered necessary when the null hypothesis is rejected ($P<0.05$).

When harmonisation is necessary, it is considered that the acquisition conditions c_2 have generated a transformation of the attribute values of Y_2. This transformation function can be approximated by a function *f* such that:

$$Y_2 = f(Y_1) \quad (1)$$

Depending on the conditions that generated this transformation, function *f* can be simulated in several different ways. As a first approach, it was considered that function *f* was a linear function (Equation 2).

$$Y_2 = a \times Y_1 + b \quad (2)$$

In Equation 2, the linear function makes it possible to account for the main sources of transformation of Y_2. The parameter a can be seen as a change in the sensitivity of the measuring system. It allows modelling changes in the sensor calibration or in the level of expertise and in operator's care when performing the measurement. Parameter b can be seen as a bias of the measuring system. It allows accounting for potential systematic bias due to the operator or sensor calibration.
The objective of the proposed harmonization approach was to estimate parameters a and b. As observations of Y_1 and Y_2 were not located on the same sites, a linear regression between these two datasets was not possible. In this respect, the following simple standardization method was proposed. First, a was estimated as the value that could minimize the difference in variance between datasets Y_1 and Y_2 using a minimization function (Nelder and Mead, 1965). This parameter a was used to transform the dataset Y_2 into a new dataset Y_2' (Equation 3).

$$Y_2' = a \times Y_1 \tag{3}$$

In a second step, parameter b was estimated as the value that could minimize the difference in mean value between Y_2' and Y_1 (centering) using the same minimization function. Parameter b was finally used to calculate $Y_{2harmonized}$ (Equation 4)

$$Y_{2harmonized} = Y_2' + b \tag{4}$$

Description of the case study
The use case corresponds to two datasets Y_1 and Y_2 of grape sugar content measurements collected using 2 portable spectrometers (Geraudie *et al.,* 2010). Observations were collected in 2015 on a 0.5 ha vineyard block. The block was located about 20 km north of Aix en Provence in the Provence region of south of France. Observations were collected by 2 different operators. Operators have gone through the same 4 vine rows. Each operator randomly chose the bunches on which measurements were performed. Therefore, observations collected by the two different operators were not collocated in space. Grape sugar contents measured by the first operator, referred to as Y_1, were considered as the BPD. This choice may seem subjective, however in a production context, indicators (often well known) relating to the knowledge of operators, their level of training, their rigour in field measurements, etc. but also on the sensors they used (sensor model, renewal, calibration procedure, etc.) may be used to decide. Y_1 and Y_2 both contain 200 observations.

Studied scenarios
In the study case, Y_1 and Y_2 contain the same number of observations. Nevertheless, in many real cases, Y_1's observations are more complex and expensive to acquire than Y_2's. It is therefore common for Y_1 to contain fewer observations than Y_2. This is the case, for example, when Y_1 is collected by a trained operator using a reliable but expensive sensor and Y_2 is collected by many operators using a low-cost but less reliable sensor.
To tackle this issue, the sensitivity of the proposed approach to the diminution of the amount of data in Y_1 was tested. In this respect, the harmonization approach was tested with Y_{1n} as the reference dataset. Y_{1n} corresponds to Y_1 but with only a certain percentage of data from the original dataset, the rest being removed. 'n' refers to this percentage, with 'n' taking the values 10, 30, 50, 70 and 90. Five iterations were performed for each value of 'n'.
Three scenarios were considered. In scenario 1, the considered dataset was only Y_{1n}. In scenario 2, it was considered Y_{1n} to which was added the whole dataset Y_2 (without harmonization). In scenario 3, it was considered Y_{1n} to which was added the whole dataset $Y_{2harmonized}$.

Evaluation of the scenarios

Datasets of the different scenarios were used to produce sugar content maps following a block-kriging procedure on a regular 10 m square grid. Scenarios were evaluated by comparing these maps to the reference map. The RMSE was chosen as an indicator to summarise the observed pixel-to-pixel differences between both maps. Spatial observations were block-kriged in order to: (1) compare observations that were not collected at identical sites; (2) smooth the information of this relatively noisy signal; and (3) work over a spatial support that is consistent with what is usually done in an operational context. This paper aims at comparing the quality of the estimation of the within-field sugar content for the different scenarios as a function of the quantity of data contained in Y_1. It is not intended to define an absolute prediction error for these scenarios. The objective was therefore to define a reference map that was relatively easy to build and that would allow scenarios to be compared to each other. The whole dataset was chosen as a reference and the reference map was then built following the same block kriging procedure described above.

Results

Figure 1A shows that the spatial distribution of the two datasets (Y_1 and Y_2) are nested in space. The attribute distribution of Y_2 presents a very similar shape but biased when compared to the one of Y_1 (Figure 1B). Regarding their spatial structure, both Y_1 and Y_2 present a high erratic variance (nugget effect around 0.7) compared to their total variance (sill around 0.9).

The application of the non-parametric K–S test to Y_1 and Y_2 led to a P-value of 2.2×10^{-3}. Distributions are considered significantly different and the data harmonisation is required. The K–S test was considered to be the most relevant in the context of this study. However, particular attention should be paid in the case of small datasets for which the P-value may be high even if the two datasets are different.

Figure 2 represents the RMSE of the sugar content prediction as a function of the size of the subset of Y_1 (from 10 to 90% of data points). The red, blue and green boxplots represent respectively the RMSE of the five iterations of the predictions obtained from scenarios 1 (only Y_{1n}), scenarios 2 (Y_{1n} + Y_2) and scenarios 3 (Y_{1n} + $Y_{2harmonized}$).

Note that no estimation was made for n=10 in scenario 1. The number of observations available (20 in this case) was too small for the variogram calculation to produce a map by block-kriging. For n=30, the number of observations available (60) was considered large enough. Regarding scenario 1, with an increasing n, the RMSE of the prediction decreased (Figure 2). This result is not surprising

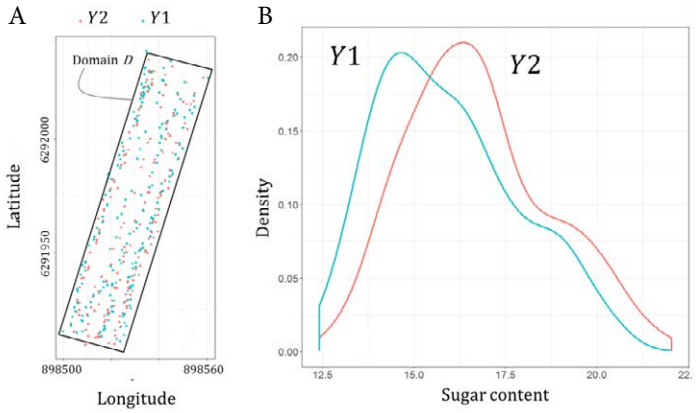


Figure 1. Spatial (A) and attribute (B) distribution of the two datasets studied.

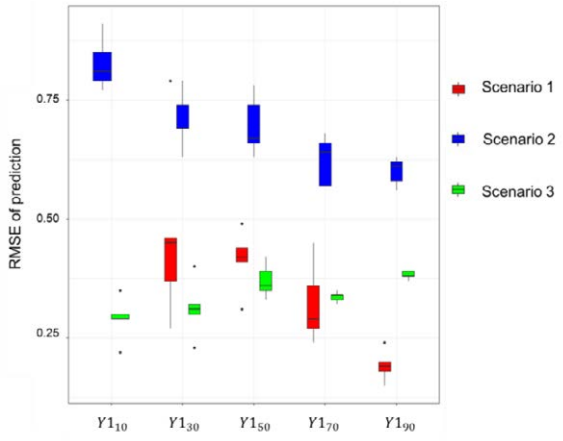


Figure 2. RMSE of prediction for the 3 scenarios depending on the quantity of data contained in Y_{1n}.

and simply reflects the fact that as more relevant data is considered, the estimation errors decrease. The RMSE of the predictions obtained in scenario 2 follows the same trend but with significantly higher RMSE values whatever the value of n (Figure 2). The addition of the uncorrected (not harmonized) dataset Y_2 significantly increases the estimation errors in any case. This proves that there is no value, and in fact a cost, in merging both datasets with different confidence levels if those datasets are left as such, i.e. raw data. The RMSE obtained in scenario 3 remains relatively constant when n varies. The observed RMSE slightly oscillates between 0.3 and 0.4 (Figure 2). This result can be explained by the fact that the $Y_{2harmonized}$ dataset ensures a lower estimation error when the number of relevant data in Y_1 is small. In contrast, when the subset of Y_1 increases, the addition of the dataset $Y_{2harmonized}$ may limit the quality of the prediction. The low RMSE value (around 0.2) observed with only $Y_{1n=90}$, is never reached (Figure 2).

More generally, for high values of n (>70%), the RMSE observed with Y_{1n} (scenario 1) is always lower than that observed with Y_{1n} + Y_2 (scenario 3) (Figure 2). This trend changes when n decreases and shows that there is a quantity of data in Y_1 below which the addition of the data set $Y_{2harmonized}$ improves the quality of the resulting prediction. In this particular case, this amount of data corresponds to a value of 50% of Y_1. Therefore, when n is lower than this value (n=50%), results show that it makes sense to use $Y_{2harmonized}$ to produce a map of the grape sugar content.

Figure 3 represents some of the block-kriged maps that were computed for different sizes of subsets of Y_1. For n equal to 90%, the spatial patterns obtained in scenario 1 (Figure 3B) are the closest to those of the reference map (Figure 3A). The southwestern/northeastern gradient of the sugar content is similar. There are also the three main zones, respectively low, medium and high sugar contents along this gradient. On the other hand, for n equal to 30%, the spatial patterns of scenario 1 are different from those of the reference. In this case, scenario 3 (Figure 3G) presents the closest spatial patterns to the reference. It also has the lowest RMSE (0.32 compared to 0.39 for scenario 1). In both cases (n=30% and n=90%), the spatial patterns of scenario 2 are the most different of the reference map. The maps (Figure 3) of the two example values of n (30 and 90%) confirm the results presented in Figure 2: the use of the $Y_{2harmonized}$ dataset improves the quality of sugar content maps only if the subset of Y_1 is small. In this study, and as a first approach, block-kriging was carried out with the same weight given to Y_1 and Y_2. In future work, it seems relevant, for scenarios 2 and 3, to be able to take into account the error variance of Y_2 in the weight assigned to these data as described by Chiles and Delfiner (1999).

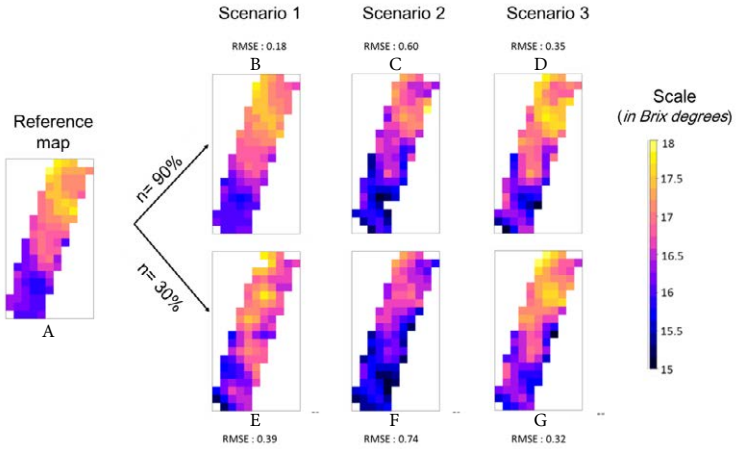


Figure 3. Reference map (A) and estimations maps produced for n=90% (B, C, D) and n=30% (E, F, G) according to the 3 scenarios.

In this paper, the reference map corresponds to the entire dataset of Y_1. This is acknowledged to be a limitation. Ideally, the reference map should have been constructed from independent data, such as from laboratory measurements of berry sugar contents. The choice of this reference map is a first approach. It allows tackling this issue by taking into account the available data in an operational context. The resulting constraints do not allow the use of more traditional methods such as independent reference maps to be used. The choice of methods like cross validation or data splitting could have offered an alternative. However, they were not considered here mainly because the amount of data available would have been too small. Results must be therefore tempered, particularly for high n values. Indeed, when n increases, on the one hand Y_{1n} tends towards Y_1. The RMSE of scenario 1 (only Y_{1n}) then decreases towards 0. On the other hand, regarding scenario 3, $Y_{1n} + Y_{2harmonized}$ tends towards $Y_1 + Y_{2harmonized}$. $Y_1 + Y_{2harmonized}$ is different from the reference (Y_1) so the RMSE is higher than 0, whatever the quality of the data contained in $Y_{2harmonized}$. This difference is due to the choice of the reference map. The differences in RMSE that are observed between scenarios 1 and 3 for n equals to 90% are therefore partly due to this choice and not only to the quality of the prediction. The choice of a better reference map would likely reduce the observed differences in RMSE. However, the trend would certainly be maintained.

This discussion raises the more global question of the choice of reference data. Indeed, this choice is often a major issue when evaluating the quality of datasets collected manually by different operators. When reference data are available, there are many methods available to assess the quality of the data collected (Flanagin and Metzger, 2008; Senaratne *et al.*, 2016). On the other hand, in the literature, only a few papers have addressed the case where reference data were not available. Muller *et al.* (2015) provided an interesting approach by using the spatial and temporal coherence of observations (temperatures) to assess data quality. This approach seems particularly relevant to crowd-sourced datasets in the context of precision agriculture, where the studied phenomena are often spatially structured. Future work will also focus on the comparison of spatial and non-spatial harmonisation methods along with their advantages and disadvantages. In particular, attention should be paid to the sensitivity of these methods to the noise of the initial datasets.

Conclusions

The development of hand-held measurement systems, and in particular smartphones, can lead to the collection of multiple datasets over the same spatial domain, but sometimes unfortunately under different acquisition conditions (operator, sensor, timing, etc.). It is therefore likely to generate nested and noisy datasets that require data harmonization before interpretation. This paper studied a simple harmonisation approach: standardisation. In this study, two datasets were compared and harmonized, one of which was considered much more trustworthy than the other. The proposed aspatial approach was tested on real grape sugar content datasets by varying the size of the confident dataset. Results showed that harmonizing a less confident dataset with respect to a more trustworthy one is interesting solely if the size of the reliable dataset is too small.

Acknowledgements

We thank the Pellenc Company for making datasets available and for their expertise in their analysis.

References

Aquino, A., Barrio, I., Diago, M. P., Millan, B., & Tardaguila, J., 2018. vitisBerry: An Android-smartphone application to early evaluate the number of grapevine berries by means of image analysis. Computers and Electronics in Agriculture, 148, 19-28.

Aquino, A., Millan, B., Gaston, D., Diago, M. P., & Tardaguila, J., 2015. vitisFlower®: development and testing of a novel Android-smartphone application for assessing the number of grapevine flowers per inflorescence using artificial vision techniques. Sensors, 15(9), 21204-21218.

Brenning, A., Koszinski, S. & Sommer, M., 2008. Geostatistical homogenization of soil conductivity across field boundaries. Geoderma, 143(3-4), pp.254-260.

Chiles, J. & Delfiner, P., 1999. Geostatistics: Modelling Spatial Uncertainty. New York, USA: John Wiley,

Flanagin, A.J. & Metzger, M.J., 2008. The credibility of volunteered geographic information. GeoJournal, 72(3-4), 137-148.

Ge, Y., Avitabile, V., Heuvelink, G.B.M., Wang, J & Herold, M., 2014. Fusion of pan-tropical biomass maps using weighted averaging and regional calibration data. International Journal of Applied Earth Observation and Geoinformation 31, 13 24

Geraudie, V., Roger, J.M. & Ojeda, H., 2010. Développement d'un appareil permettant de prédire la maturité du raisin par spectroscopie proche infra-rouge(PIR). (Development of a sensor to predict grape maturity by near infrared spectroscopy (NIR)). Revue Française d'Oenologie,. pp 2-8, <hal-00647105>

Minet, J., Curnel, Y., Gobin, A., Goffart, J.P., Mélard, F. & Tychon, B., 2017. Crowdsourcing for agricultural applications: A review of uses and opportunities for a farmsourcing approach. Computers and Electronics in Agriculture, 142, 126-138.

Muller, C.L., Chapman, L., Johnston, S., Kidd, C., Illingworth, S. & Foody, G., 2015. Crowdsourcing for climate and atmospheric sciences: Current status and future potential. International Journal of Climatology, 35(11), 3185-3203.

Nelder, J. A. & Mead, R., 1965. A simplex algorithm for function minimization. Computer Journal, 7, 308-313.

Sams, B., Litchfield, C., Sanchez, L. & Dokoozlian, N., 2017. Two methods for processing yield maps from multiple sensors in large vineyards in California. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences 8, 530-533

Senaratne, H., Mobasheri, A., Loai Ali, A., Capineri, C & Haklay, M., 2016. A review of volunteered geographic information quality assessment methods. International Journal of Geographical Information Science, 31(1), 139-167.

Can mapping of within-vineyard variability be facilitated using data from multiple vineyards?

B. Sams[1,2*], R. Bramley[3], V. Pagay[1], L. Sanchez[2], C. Ford[1] and N. Dokoozlian[2]
[1]*School of Agriculture, Food, and Wine, University of Adelaide, Urrbrae, SA, Australia*
[2]*Department of Viticulture, Chemistry, and Enology, E&J Gallo Winery, Modesto, CA, USA; brent.sams@ejgallo.com*
[3]*CSIRO, Waite Campus, Urrbrae, SA, Australia*

Abstract

Variogram estimation requires extensive sampling that may limit its practical uses in agriculture. During the 2017 harvest, 125 grape samples were collected for the analysis of chemical composition from each of four vineyards in the Lodi area of the Central Valley of California. Maps of grape colour (anthocyanins) of each block were produced using variograms derived from data from each individual vineyard (n=125), as well as from the aggregated dataset containing all samples from all four vineyards (n=500). The resulting grape colour maps show differences in spatial patterns, although the impact of these on vineyard decision making would be minimal. Aggregation of data from multiple vineyards in a region may therefore offer a means of cost-effectively examining within-vineyard variation in individual vineyards.

Keywords: vineyard variability, kriging, precision viticulture

Introduction

The cost-effectiveness of collecting and analysing sufficient samples for characterization and mapping of within-vineyard variation in crop attributes is a major constraint to understanding variability in commercial vineyards. However, several studies have mapped specific attributes of grape composition, chiefly anthocyanins (Bramley, 2005; Stamatiadis *et al.,* 2006; Bramley *et al.*, 2011; Baluja, *et al.,* 2012) or of other attributes of fruit quality (Cortell, *et al.*, 2005; Reynolds *et al.*, 2007; Brillante *et al.*, 2017). However, according to Cressie (1993) and Webster and Oliver (2007), too few samples have been collected in a majority of these studies for sufficient rigour in variogram estimation and thus, map interpolation using kriging. Cressie (1993) suggested 30-50 pairs of samples for each distance class, whilst Webster and Oliver (2007) recommended that 100 samples should be regarded as the minimum for adequate variogram estimation. Most of the aforementioned studies fail to meet these criteria. Exceptions were those of Bramley *et al.* (2011) using a prototype on-the-go sensor and Bramley (2005), whose vine sampling design allowed for adequate variogram estimation and subsequent kriging of maps of grape composition; this can be used as a guide for future work in this area. However, anecdotal evidence suggests that such a sampling design presents an impediment to commercial operations, especially when, in addition to collection of samples, some sort of laboratory analysis is also required. There would therefore be much value if an understanding of variability in an attribute of interest in one vineyard could inform similar understanding in another, especially if this understanding could assist in reducing the sampling and analytical demand. McBratney and Pringle (1999) attempted to address this problem in the case of soil mapping by using the magnitude of different soil properties to estimate proportional and average variograms for other properties. Similarly, Walter *et al.* (2007) explored using combined data from different years and regions to better estimate variograms to predict oyster abundance. However, to the authors' knowledge, no study of this type has been applied to agriculture. The objective of this study was to compare the predictions between two variogram estimation methods and their impact on characterisation and

DOI 10.3920/978-90-8686-888-9_92, © Wageningen Academic Publishers 2019

mapping of spatial variability in four vineyards in the Lodi region of California. In particular, the main interest was to see whether samples collected in one vineyard could assist in the mapping of variability in another from the same region.

Materials and methods

In the 2017 harvest season, four Cabernet Sauvignon vineyard blocks were selected for high density sampling of grape anthocyanins as shown in Figure 1. All four vineyards are in the American Viticultural Appellation (AVA) of Lodi, California (38° 7' 44' N, 121° 16' 51' W) and occur within <25 km of each other.

A detailed scheme for fruit sampling was developed for each vineyard as shown in Figure 2. This sampling scheme assumed no prior knowledge of vineyard variability and contained enough points, or 'data vines', for adequate variogram estimation (n=125) based on the criteria of Webster and Oliver (2007). To develop this scheme, the points were first spread out in a common square grid across each vineyard. Next, each point was given a random offset in both *x* and *y* directions – both to randomise aspects of data vine selection, and to create a sampling grid with different distances between points for more robust variogram estimation, especially at shorter distances. Finally, approximately 20 sample points from each sampling grid were then randomly removed and randomly reassigned to a position adjacent to another data vine (i.e. either one row space or one vine space from the adjacent data vine). This method allowed for the estimation of an experimental variogram for grape anthocyanin content in each vineyard over lags ranging from a single vine spacing to as far as the farthest two points from each other. Importantly, it provided adequate numbers of pairs of points at short distances for characterising short-range spatial variation. Total anthocyanins were measured by collecting 20 clusters from each data vine and subsequent analysis according to Iland *et al.* (1996). Experimental variograms were then estimated for the anthocyanin data of each



Figure 1. Each of the four selected vineyards in the Lodi area of interest.

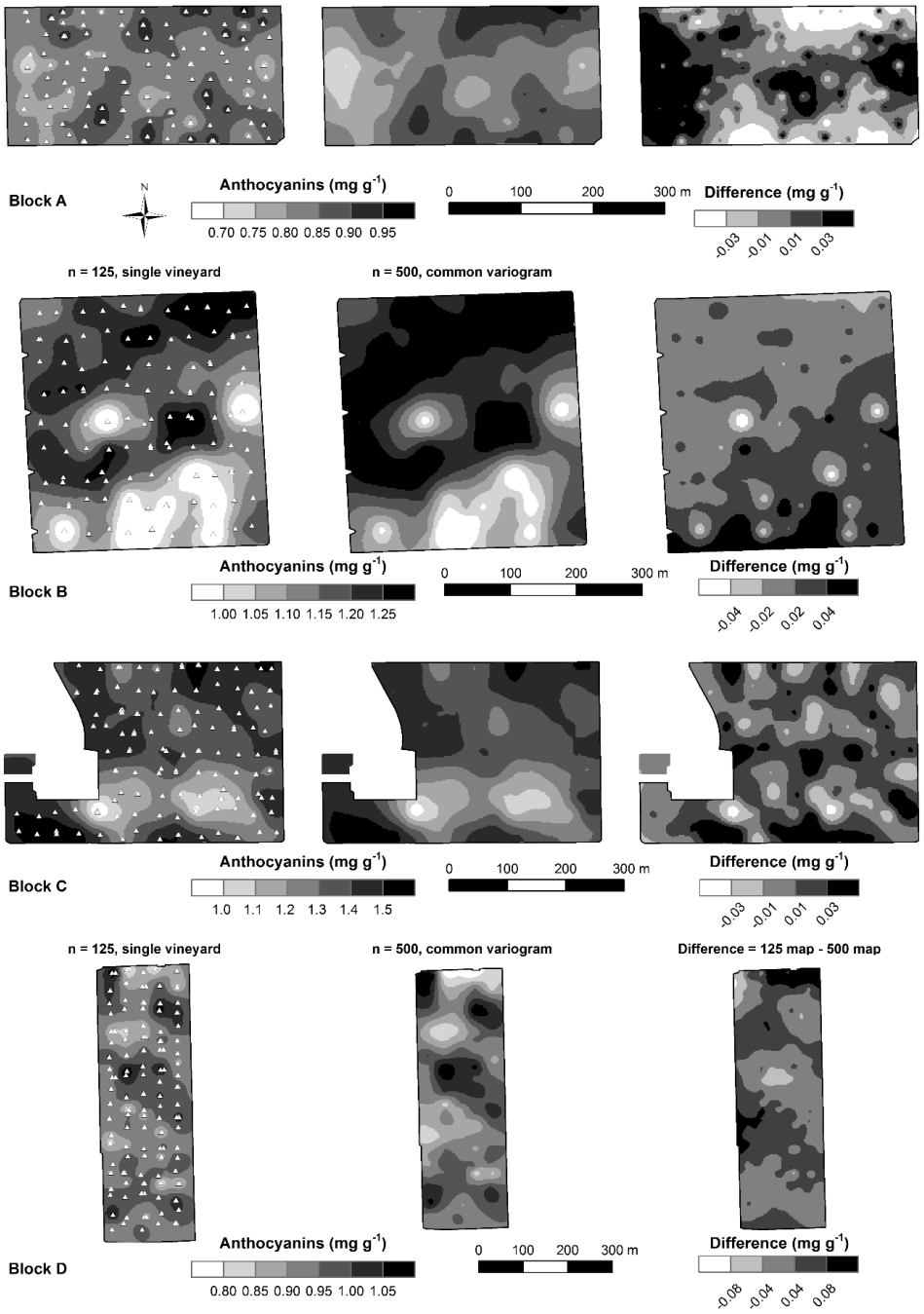


Figure 2. Anthocyanin maps using variograms derived either from only sample points within the vineyard (n=125) or from all four vineyards (n=500) and the difference between these maps.

vineyard using VESPER (Minasny, *et al.,* 2006). Preliminary exploratory variogram fitting led to the assumption of the exponential model being appropriate for characterising spatial structure in anthocyanin variation. Maps were then interpolated for each vineyard in VESPER using global point kriging, based on the 125 samples within.

Aside from the assumption of an exponential variogram, several constraints were applied to the variogram estimation. The number of lag classes (20) and the lag tolerance (10%) were set for each model, as was a maximum distance for variogram estimation (250 m); that is, the largest lag class was set at 250 m. The choice in variogram settings was based on initial exploratory analysis and to provide a consistent set of parameters for method comparison. A further variogram was then produced by combining all 500 points from all four vineyards into a 'common variogram' which was nevertheless estimated using the same settings in VESPER as for the vineyard-specific variograms, viz. number of lag classes, lag tolerance and maximum distance. From the variogram parameters fitted in VESPER, (nugget, sill and effective range), the nugget/sill ratio, sometimes referred to as the 'Cambardella Index', was calculated to give a measure of spatial structure, where a value less than 25 was deemed to represent 'strong' spatial structure, and a value greater than 75 representing 'low' spatial structure (Cambardella, *et al.,* 1994).

Results

Differences between the maps generated using either the vineyard-specific or common variograms were small (Figure 2 and 3). Absolute differences in mapped anthocyanin concentration were generally of the order of analytical error or smaller, and the difference ranges in all four vineyards were similarly spaced from 0 (Figure 2). However, in terms of providing a platform for winery decision making, when the maps were classified on a 'low', 'medium' and 'high' anthocyanin basis (derived from 33rd percentiles), the misclassification of 'anthocyanin zones' when using the common variogram was minimal (Figure 3). Block C represents the smallest level of misclassification, over the entire area of 11.9 ha, where only 0.7 ha (i.e. 5.9%) were assigned to a different zone when using the map derived from the common variogram (n=500). In Blocks A, B, and D, the misclassifications represented 1.3 ha 10.0%), 0.9 ha (12.4%), and 1.6 ha (18.4%) respectively. Figure 3 also shows that the misclassifications generally occurred on zone boundaries, as might be expected.

Table 1 shows the difference between parameters of both variograms and the effect on spatial structure. Effective range was between 12.5 m and 49.5 m when using the individual within-vineyard variograms, while the effective range in the common variogram was higher (66.6 m) than all four within-vineyard variograms. The change in Cambardella index in Block D was the highest (10.3), with the lowest in Block B (1.18). All four values of the Cambardella index were in the medium to low range indicating moderate to strong spatial structure. These small changes would have no impact on the interpretation of each map.

Discussion

The difference maps in Figure 2 and 3 show the largest difference in anthocyanins to be less than 0.1 mg/g, which would be considered insignificant to winemakers and vineyard managers. The key finding of this work is that, in practice, vineyard managers would do nothing differently based on the changes between maps produced by the two models, yet the maps derived from the common variogram reduce the need for expensive and time-consuming sampling in an individual vineyard. While the overall variance amongst the 500 data points used to construct the common variogram was greater than in the individual blocks (n=125), the effect on spatial structure in the resulting map and on the characterisation of spatial structure (i.e. Cambardella index) was minimal. The difference in predicted anthocyanins was also very low, further making the case for use of the common variogram. The increase in the number of points in the shortest distance classes in the

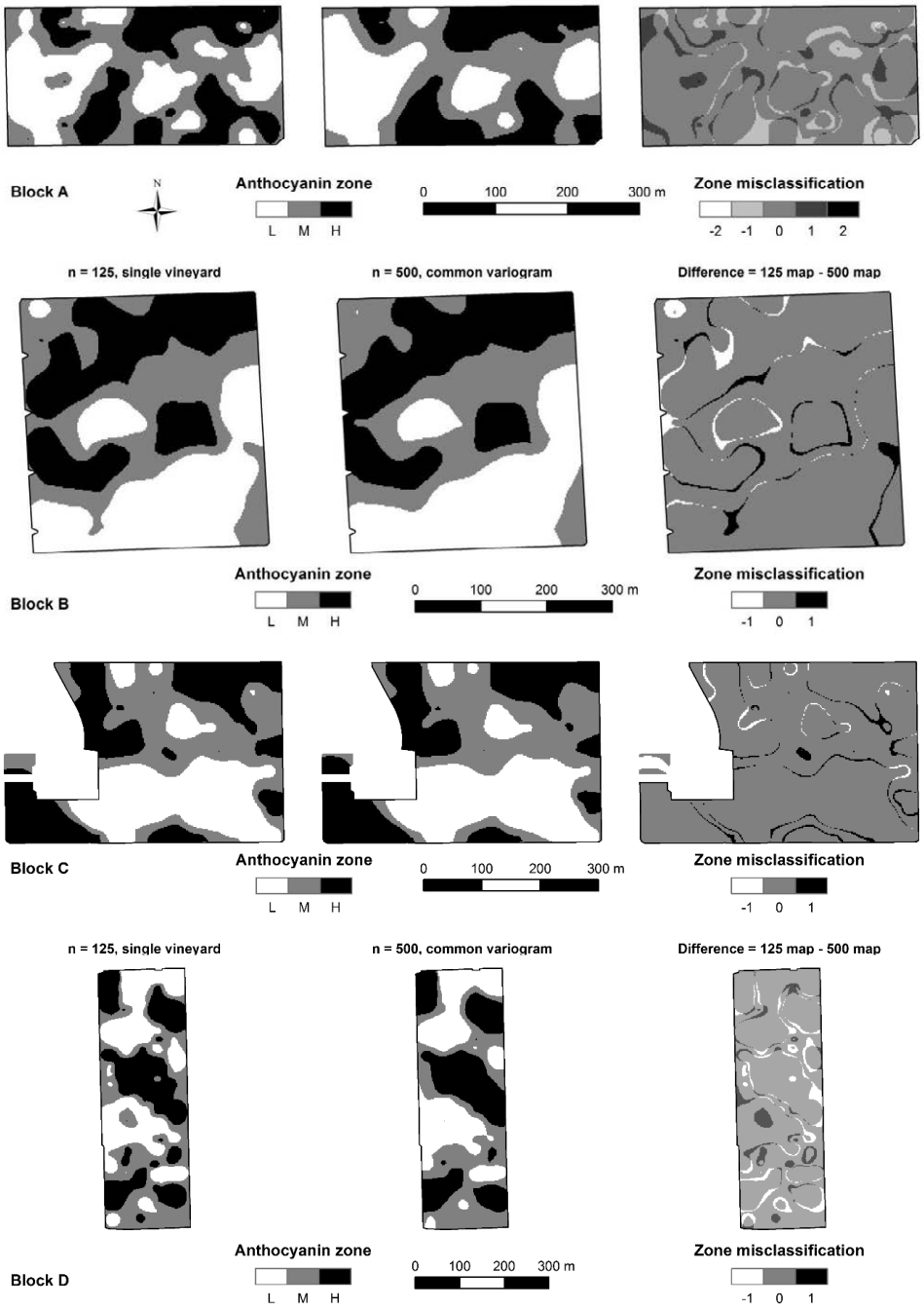


Figure 3. Maps of 'high', 'medium', and 'low' zones of anthocyanins derived either from only sample points within the vineyard (n=125) or from all four vineyards (n=500) and the misclassification of zones due to use of the common variogram.

Table 1. Comparison of descriptive and spatial statistics from each of the four vineyard blocks.

		Vineyard block			
		A	B	C	D
Variogram 125	Mean (μ)	0.7907	1.1460	1.3433	0.8913
	Variance (σ^2)	0.0106	0.0350	0.0477	0.0244
	c_0 (nugget)	0.00423	0.01303	0.01676	0.01162
	$c_0 + c_1$ (sill)	0.00978	0.03414	0.04790	0.02334
	Effective range (m)	12.50	35.05	49.46	16.23
	Cambardella index	43.25	38.17	34.99	49.79
Variogram 500	Mean (μ)	1.0428			
	Variance (σ^2)	0.0762			
	c_0 (nugget)	0.01204			
	$c_0 + c_1$ (sill)	0.03060			
	Effective range (m)	66.57			
	Cambardella index	39.35			
		A	B	C	D
Difference	Mean (μ)	0.2521	0.1030	0.3010	0.1515
	Variance (σ^2)	0.0656	0.0412	0.0285	0.0518
	c_0 (nugget)	0.00781	0.00100	0.00470	0.00042
	$c_0 + c_1$ (sill)	0.02082	0.00350	0.01730	0.00730
	Effective range (m)	54.07	31.52	17.11	50.34
	Cambardella index	3.90	1.18	4.36	10.44

common variograms improved the robustness of the prediction due to the importance of properly characterising variation at lags less than the range of spatial dependence. Thus, this approach could be used as a default for anthocyanin prediction in Lodi vineyards. Samples could be continuously added to the common variogram for an increasingly robust predictive capability. The Cambardella index in all four blocks indicated moderate to high spatial structure, and the difference between the two variograms showed a very small change in each block. Differences in the effective range could have the greatest effect on vineyard decision making or sample design, since this distance could potentially be used in optimizing any measurements or tissues to sample.

The common variogram approach offers several promising improvements to sample based design and future sampling schemes. As more and more data are collected with geospatial analysis in mind, they could be combined in multiple ways. This type of data combination, along with normalisation, could also be used in temporal analysis by combining data from multiple seasons for a different type of common variogram. The inclusion of additional data over time could greatly increase the effectiveness of smaller sampling campaigns, as the relationships over multiple distances could be integrated into the same variogram. In cases where samples are taken prior to harvest, these values could be normalised and included into the regional variogram for an estimate of where to expect the relative colour ranges in that region and then streamed into appropriate wine programs. These additions could also be used to build spatial databases of key vineyard attributes to include, among many others, fruit quality, soil characteristics or plant nutrition.

Conclusions

The finding that a common variogram may be applied to mapping grape anthocyanins in single vineyards may enhance the robustness of further anthocyanin mapping, especially in previously unsampled Cabernet Sauvignon vineyards in Lodi. Using 500 data values from four vineyards enabled more robust variogram construction than when the data from individual blocks were used separately, with very small differences seen between maps kriged using either individual or common variograms; these differences were inconsequential for decision making.

Acknowledgements

The authors would like to thank the Viticulture, Chemistry and Enology staff at E&J Gallo Winery for their help and support in the collection and processing of the samples.

References

Baluja, J., Diago, M.P., Goovaerts, P., and Tardaguila, J., 2012. Spatio-temporal dynamics of grape anthocyanin accumulation in a Tempranillo vineyard monitored by proximal sensing. Australian Journal of Grape and Wine Research 18, 173-182.

Bramley, R.G.V., 2005. Understanding variability in winegrape production systems 2. Within vineyard variation in quality over several vintages. Australian Journal of Grape and Wine Research 11, 33-42.

Bramley, R.G.V., Moigne, M.L., Evain, S., Ouzman, J., Florin, L., Fadaili, E.M., *et al.,* 2011. On-the-go sensing of grape berry anthocyanins during commercial harvest: development and prospects. Australian Journal of Grape and Wine Research 17, 316-326.

Brillante, L., Martínez-Luscher, J., Yu, R., Plank, C.M., Sanchez, L., Bates, T.L., *et al.,* 2017. Assessing Spatial Variability of Grape Skin Flavonoids at the Vineyard Scale Based on Plant Water Status Mapping. Journal of Agricultural and Food Chemistry 65, 5255-5265.

Cambardella, C., Moorman, T., Parkin, T., Karlen, D., Novak, J., Turco, R., *et al.,* 1994. Field-scale variability of soil properties in central Iowa soils. Soil Science Society of America Journal 58, 1501-1511.

Cortell, J.M., Halbleib, M., Gallagher, A.V., Righetti, T.L., and Kennedy, J.A., 2005. Influence of Vine Vigor on Grape (*Vitis vinifera* L. Cv. Pinot Noir) and Wine Proanthocyanidins. Journal of Agricultural and Food Chemistry 53, 5798-5808.

Cressie, N., 1993. Spatial Prediction and Kriging. Statistics for Spatial Data, Revised Edition, Chichester, UK: John Wiley & Sons, Ltd,. pp 105-209.

Iland, P.G., Cynkar, W., Francis, I.L., Williams, P.J., and Coombe, B.G., 1996. Optimisation of methods for the determination of total and red-free glycosyl glucose in black grape berries of Vitis vinifera. Australian Journal of Grape and Wine Research 2, 171-178.

McBratney, A. and Pringle, M., 1999. Estimating average and proportional variograms of soil properties and their potential use in precision agriculture. Precision Agriculture 1, 125-152.

Minasny, B., McBratney, A., and Whelan, B., 2006. Vesper version 1.62. Australian centre for precision agriculture, Mcmillan building A05, The University of Sydney. NSW.

Reynolds, A.G., Senchuk, I.V., van der Reest, C., and de Savigny, C., 2007. Use of GPS and GIS for elucidation of the basis for terroir: Spatial variation in an Ontario Riesling vineyard. American Journal of Enology and Viticulture 58, 145-162.

Stamatiadis, S., Taskos, D., Tsadilas, C., Christofides, C., Tsadila, E., and Schepers, J.S., 2006. Relation of Ground-Sensor Canopy Reflectance to Biomass Production and Grape Color in Two Merlot Vineyards. American Journal of Enology and Viticulture 57, 415-422.

Walter, J.F., Christman, M.C., Hoenig, J.M., and Mann, R., 2007. Combining data from multiple years or areas to improve variogram estimation. Environmetrics 18, 583-598.

Webster, R. and Oliver, M.A., 2007. Characterizing spatial processes: the covariance and variogram. Geostatistics for Environmental Scientists, Second Edition, Chichester, UK: John Wiley & Sons, Ltd,. pp 47-76.

Assessing the spatial variability of winter wheat yield in large-scale paddy fields of Japan using structural equation modelling

T.S.T. Tanaka*, Y. Kono and T. Matsui
Faculty of Applied Biological Sciences, Gifu University, 1-1 Yanagido, Gifu, Japan; takashit@gifu-u.ac.jp

Abstract

Wheat is often grown in upland fields converted from paddy fields in Japan. These recently consolidated large-scale paddy fields may show high spatial heterogeneity in soil properties. Thus, the spatial variability in soil properties and their impacts on wheat yield were evaluated using structural equation modelling and semivariogram. The results showed that the coarse sand content had a significantly positive effect on the number of established seedlings, which in turn contributed to the number of panicles and grain yield. Furthermore, the coarse sand content showed high spatial dependence. These results indicate that variable-rate seeding might be an effective method to improve wheat yield in accordance with the spatial variability in coarse sand content in large-scale paddy fields.

Keywords: large-scale paddy field, semivariogram, structural equation modelling, winter wheat

Introduction

Within-field variability in soil properties causes spatial variability in crop yield. Site-specific management aims to improve crop yield without increasing environmental loads by optimizing inputs in accordance with the spatial variability in soil properties (Cassman, 1999). Site-specific management is performed by prescriptions based on crop growth status and soil properties. However, the lack of knowledge of the spatial variation in crop yield limits the development and adoption of site-specific management technologies in Asian countries.

In Japan, it is recommended to cultivate wheat or soybean in upland fields converted from paddy fields in a multiple cropping system because of the low price of rice and the high demand for domestic production of wheat and soybean. However, the relatively high moisture content of paddy soil after rice cropping hinders the establishment, tillering and yield of winter wheat in the rice-wheat cropping system (Timsina and Connor, 2001; Inamura *et al.*, 2010). Furthermore, conventional small-scale paddy fields have been increasingly consolidated into large-scale paddy fields to increase labour efficiency. Recently consolidated large-scale paddy fields may have high spatial heterogeneity in soil properties. Conventional small-scale paddy fields are managed by small-holder farmers in Japan. However, crop yield cannot be improved by conventional management in large-scale paddy fields. Site-specific management would not be profitable in conventional small-scale paddy fields because of the high investment, but an economic return would be expected if site-specific management was introduced to large-scale paddy fields. Thus, it is essential to understand the spatial variability in soil properties and crop yields in large-scale paddy fields for the future development of site-specific management practices particularly suited to Japanese agricultural conditions.

The objectives of this study were (1) to characterize the within-field spatial variability in soil properties and wheat yields; and (2) to understand the hierarchical processes determining wheat yields in large-scale consolidated paddy fields. For this purpose, structural equation modelling (SEM – Awang, 2012) was applied. SEM is a statistical approach for the analyses of intercorrelated complex data, and allows both the direct and indirect causal relationships between variables to be inferred. The final goal was to identify yield-limiting factors that can be manipulated or controlled by site-specific management.

DOI 10.3920/978-90-8686-888-9_93, © Wageningen Academic Publishers 2019

Materials and methods

Two upland fields converted from consolidated paddy fields (fields A and B) were surveyed at Sotohama, Kaizu, Gifu, Japan (35°11' N, 136°40' E), from November 2017 to June 2018. Small paddy fields with an area of approximately 0.3 ha were consolidated into large-scale paddy fields with an area of approximately 1 ha in the 1990s. Four paddy fields were consolidated for field A, and 3 paddy fields were consolidated for field B. The land use system of these fields involved the cropping of three crops (paddy rice-winter wheat-soybean) in two years. The seeding rate of wheat was 100 kg/ha. Both fields received the same amounts of basal fertiliser (8.4 g N/m^2, 5.9 g P_2O_5/m^2, and 5.9 g K_2O/m^2) and top dressing (5.00 g N/m^2, 1.0 g P_2O_5/m^2, and 4.0 g K_2O/m^2) at the tillering stage. Immediately before wheat seeding (25 November 2017), surface soil (0-150 mm) samples were collected from 4 transects at each large-scale paddy field. The sampling points occurred at 5-m intervals along the long-side transect (East-West axis) to investigate the spatial variability among the pre-consolidated paddy fields and at 2.5-m intervals along the three narrow-side transects (North-South axes) to investigate spatial variability within each pre-consolidated paddy field (Figure 1). At each location, 3 random partial soil samples weighing approximately 0.5 kg were collected within one square metre and mixed to obtain one composite sample. Soil pH, electrical conductivity (EC), total carbon (C) and nitrogen (N) contents, particle size distribution (coarse sand, fine sand, silt, and clay), and available N were measured. Soil pH and EC were measured using a glass electrode after mixing the soil with water (1:5, w/v). Total C and N were determined using a CN analyser (Sumigraph NC-95A, Sumitomo Chemical Co., Tokyo, Japan). The particle size distribution was measured using the pipette method and sieving (Gee and Bauder, 1986). Available N was measured using the hot KCl extraction method (Gianello and Bremner, 1986).

At each soil sampling point, the number of established seedlings per square metre was quantified at the seedling stage (14 February 2018), and plant materials were collected from a 0.6-m^2 area at the maturity stage (4 June 2018). The plant materials were divided into grain and straw. The dry matter weight was determined after the subsamples of grain were oven-dried at 70 °C to a constant weight. Wheat yield was calculated at 13.0% moisture content. A total of 184 samples, composed of 95 samples from field A and 89 samples from field B, were collected for both soil and plant materials.

Data analysis

All statistical analyses were performed using R version 3.5.0 (R Development Core Team, 2018). A *P*-value<0.05 was considered to be statistically significant for all statistical analyses. Welch's t test was used for comparison of means between fields. Segmented linear regression analysis was used

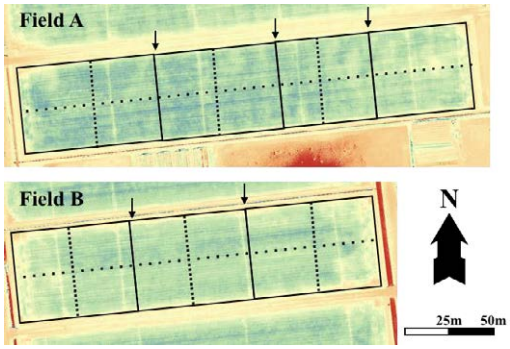


Figure 1. Map of the sampled fields. Black points represent the soil and plant sampling points. Black arrows represent boundaries between fields before consolidation (pre-consolidated fields).

to evaluate the relationships between the number of established seedlings and that of panicles using the 'segmented' package implemented in R 3.5.0. The breakpoint was tested by the Davies's test.
To elucidate the causal relationships affecting the within-field variability in wheat yield, SEM was performed with the 'lavaan' package version 0.6-3 (Yves, 2012) implemented in R 3.5.0. SEM is a powerful statistical approach for the analysis of complex intercorrelated data with wide range of potential applications in plant sciences (Lamb *et al.,* 2011). SEM is a multivariate analysis of the causal relationships between variables, which developed from the methods of path analysis. The initial model of SEM represents a hypothesis on the causal relationships in the form of path diagram, which can be tested against the actual variances and covariances in the observed data. Thus, the construction of the initial model is an essential step in SEM. The model is constructed based on theoretical knowledge. The seedling establishment ratio is known to be affected by the particle size distribution, soil moisture and total C which are relevant to the determination of clod size during the sowing season, as suggested by Inamura *et al.* (2010). The seedling establishment ratio was calculated by dividing the number of established seedlings by the product of seeding density and seed germination ratio. In this study, there were no variations in seeding density and seed germination ratio, and clod size were not measured. Thus, the particle size distribution and total C is hypothesized to affect the number of established seedlings in the initial model. The particle size distribution is hypothesized to affect soil moisture which in turn affects active growth and each yield component (the number of panicles, the number of grains per panicle and the 1000-grain weight) because the high soil moisture content of paddy soil often negatively affects wheat growth. Furthermore, soil available N is hypothesized to affect each yield component. The model fit was measured by the χ^2 test with a null hypothesis of adequate fit. When the initial model did not adequately fit the data (P>0.05), new paths were added or non-significant paths were removed from the model according to the modification indices to improve model fit. The model fit was further evaluated by the goodness of fit index (GFI), comparative fit index (CFI) and root mean square error of approximation (RMSEA). Values of the GFI and CFI greater than 0.9 and a value of the RMSEA less than 0.08 indicate an acceptable model (Awang, 2012). All variables were standardized prior to the SEM application.
A semivariogram was used to evaluate the spatial variability in parameters used in SEM using the 'gstat' package (Pebsma, 2004) implemented in R 3.5.0. Either spherical or exponential model was used for the estimation of semivariogram parameters (nugget, sill, and range) using sum of squares errors (SSE). The semivariogram was just applied for the long-side transect (east-west axis) of each field because of the anisotropy and insufficient maximum sampling distance in the narrow-side transects (north-south axes) for estimating ranges. The Q value was used to evaluate the spatial dependence of each variable (Goerres *et al.*, 1997). This value is defined as follows:

$$Q = (S - N) / S \quad (1)$$

where S and N are the sill and nugget variance, respectively. The Q value ranges between 0 and 1. Q values closer to 1 indicate a more developed spatial structure. The nugget-to-sill ratio indicates spatial dependence as described by Cambardella *et al.* (1994). Therefore, strong spatial dependence is indicated if the Q value is greater than 0.75. Moderate spatial dependence is indicated if the Q value is between 0.25 and 0.75. Weak spatial dependence is indicated if the Q value is less than 0.25.

Results

Structural equation modelling was undertaken for each field to explain the processes affecting within-field variability in wheat yield. The coarse sand content was selected as a parameter to perform SEM because of the more stable and higher goodness of fit among the other particle size distribution for both fields. The results of the χ_2 test indicated that the initial model was not an adequate fit for either field A (χ_2=57.270, P<0.001) or B (χ_2=37.047, P=0.001). There were no significant effects of available

N or total C on any of the yield components (the number of panicles, the number of grains per panicle and 1000-grain weight) in the initial models for either field. Therefore, the non-significant paths were removed from the initial model. Furthermore, the paths of 1000-grain weight were also removed from the initial model because the standardized coefficient of the 1000-grain weight on yield was extremely small, which contributed to the inadequate model fit. The modified model had adequate fit according to the χ_2 test for fields A (χ_2=6.905, P=0.141) and B (χ_2=7.797, P=0.099) (Figure 2). Most of the other indices indicated a satisfactory fit for field A (GFI=0.967, CFI=0.993) and field B (GFI=0.960, CFI=0.988), although the values of the RMSEA did not meet the criteria for fields A (RMSEA=0.087) and B (RMSEA=0.103). The effect of the number of panicles on yield was stronger than that of the number of grains per panicle in both fields. There were significant direct effects of coarse sand content on the number of panicles for both fields, whereas there was a significant indirect effect of coarse sand content on the number of panicles through a change in the number of established seedlings in field A but not in field B. The direct effect of coarse sand content on the number of panicles (0.22) was almost the same as the indirect effect (0.48 × 0.37 = 0.17) in field A. To evaluate the processes affecting the number of panicles between fields, the relationships between the coarse sand content, number of established seedlings and number of panicles are shown in Figure 3. The number of established seedlings and panicles increased with coarse sand content for both fields. The number of established seedlings significantly contributed to an increase in the number of panicles (P<0.001) until it reached a breakpoint 71 seedlings/m^2 according to the result of the Davies' test (P<0.001). The relatively smaller number of panicles was observed at the sampling points of 5 m away from the edge although the number of established seedlings exceeded the breakpoint.

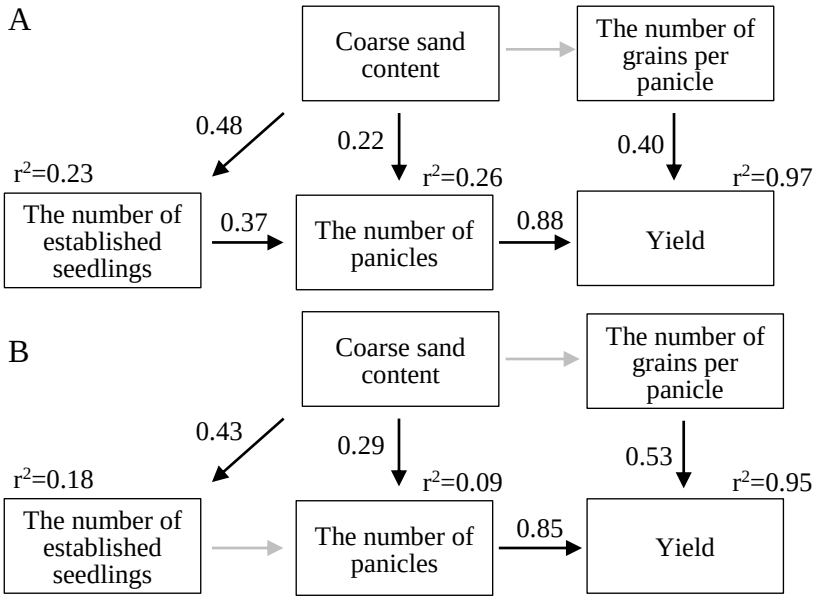


Figure 2. Fitted observed model modified from the initial model for fields A (a) and B (b). The black arrows represent significant paths (P>0.05). The grey arrows represent non-significant paths. The values associated with the arrows are significant standardized path coefficients. The values of r^2 are presented for three endogenous variables (the number of established seedlings, the number of panicles and wheat yield).

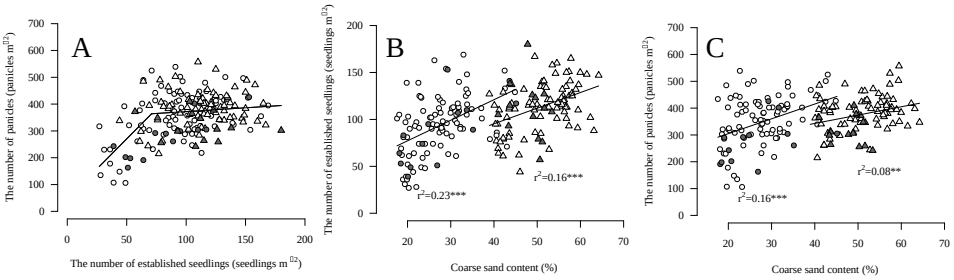


Figure 3. The relationships between the coarse sand content and, (A) the number of established seedlings, (B) the number of panicles, and (C) the number of established seedlings and the number of panicles. The circles represent field A, and triangles represent field B. The grey symbols represent sampling points at a distance of 5 m from the edge. **P<0.01, ***P<0.001.

The descriptive statistics and geostatistical parameters of the SEM parameters are summarized in Table 1. The mean value of coarse sand content was approximately twice as large in field B than in field A. The coarse sand content showed strong dependence at both field (Q>0.79). The coefficient of variance of wheat yield was greater than 20% for both fields, indicating that the yield greatly varied within the fields. The mean values of the number of established seedlings and panicles were significantly higher in field A than in field B, although the coefficients of variance were greater than 17% in both fields. There was no significant difference in the number of grains per panicle and yield between the two fields. The Q value of the number of established seedling indicated relatively strong spatial dependence for both fields (Q>0.64). However, there were inconsistencies in the spatial dependence of the number of panicles and yield between the fields because there were no convergence for field B. Furthermore, the range exceeded the sampling distance for yield and the spatial variability increased linearly with distance in field A (data not shown).

Table 1. Descriptive statistics and geostatistical parameters for the SEM parameters.[1]

Parameters	Field	Mean		Max	Min	CV (%)	N	S	R (m)	Q	SSE
Coarse sand (%)	A	27.6	***	45.1	17.6	24.4	5.1	43.5	128	0.88	3.20
	B	50.7		64.22	40.08	12.5	13.7	64.6	79	0.79	12.62
The number of established seedlings (seedlings/m^2)	A	94	***	169	27	33.8	262	1,209	93	0.78	4,998
	B	113		180	44	24.1	525	1,458	106	0.64	21,182
The number of panicles (panicles/m^2)	A	347	*	539	106	27.4	1,331	10,661	125	0.88	814,405
	B	376		557	215	17.6	-	-	-	-	-
The number of grains per panicle (grains/panicle)	A	30.61	NS	45.8	19.26	12.3	Pa	-	-	-	-
	B	30.39		39.41	19.78	10.9	-	-	-	-	-
Yield (g/m^2)	A	371	NS	624	92	33.3	4,294	29,182	307	0.85	2,358,147
	B	393		632	198	22.5	-	-	-	-	-

[1] The semivariogram was applied only for the long-side transect (east-west axis) of each field. A spherical model was the best fit for all parameters. Only converged model are shown. * P<0.05; *** P<0.001; NS not significant (Welch's t test). N nugget; S sill; R range; SSE sum of squares errors.

Discussion

The processes affecting within-field variability in wheat yield were evaluated by SEM. As predicted in the initial model, coarse sand content had significant direct and indirect effects on the number of panicles, which was the most important yield component in field A (Figure 2). However, there was no significant effect of the number of established seedlings on the number of panicles, whereas the effect of coarse sand content on the number of established seedlings was significant in field B (Figure 2). The wheat yield and the number of panicles are known to increase with plant density to a certain extent (Spink *et al.*, 2000). The mean value of coarse sand content was significantly lower in field A (27.6%) than in field B (50.6%), and the coefficient of variance was higher in field A (24.4%) than in field B (12.5%) (Table 1). The number of established seedlings did not contribute to an increase in the number of panicles when it exceeded a threshold value of 71 seedlings/m^2 (Figure 3). The results indicate that the coarse sand content may not be an important factor determining the number of panicles in field B (Figure 3). This implies that the within-field yield-limiting factors varied among the fields in accordance with the amplitude rage of soil properties. Thus, variable-rate seeding would be an effective method for improving wheat yield in fields in which the coarse sand content greatly varies. The coarse sand content would be a great indicator for creating variable rate seeding prescriptions because the strong spatial dependence with relatively long ranges (79-129 m) could contribute to save time and costs for soil test. Furthermore, the coefficients of determination of the number of panicles were not high in fields A (r^2=0.26) and B (r^2=0.09) (Figure 2). The available N was not an important factor affecting the yield components. The smaller number of panicles occurred for the sampling points of 5 m from the edge even at more than the breakpoint (71 seedlings/m^2) (Figure 3). The farmers used fertiliser broadcaster without GPS positioning and spreader control computer, and there were open-channel drainages at the edge of the fields. This might indicate an edge effect where potentially received less fertiliser or more wet damage near the edge of the fields. Furthermore, there were strong spatial dependences for the number of panicles and yield in field A, whereas there were no developed spatial dependence for the number of panicles and yield in field B. Specifically, the spatial variability of yield increased linearly with distance in field A. Thus, there might be different yield-limiting factors between the fields. Further studies with more information on spatial variability in management practices and soil properties affecting wheat yield are required to obtain a conclusion on the best management practices.

Conclusions

The coarse sand content had a significantly positive indirect effect on the number of panicles and grain yield through a change in the number of established seedlings when the coarse sand content varied greatly within the fields. Furthermore, the coarse sand content had a significantly positive direct effect on the number of panicles. These results indicate that variable-rate seeding might be an effective method to improve wheat yield in large-scale paddy fields where the coarse sand content varies substantially.

Acknowledgements

This work was supported by a JSPS KAKENHI Early-Career Scientists Grant (number 18K14452).

References

Awang, Z., 2012. Structural Equation Modeling Using Amos Graphic. Malaysia: Universiti Teknologi MARA Press.

Cambardella, C. A., Moorman, T. B., Novak, J. M., Parkin, T. B., Karlen, D. K., Turco, R. F. *et al.*, 1994. Field-scale variability of soil properties in Central Iowa soils. Soil Science Society of America Journal 58(5) 1501-1511.

Cassman, K. G., 1999. Ecological intensification of cereal crop production systems: Yield potential, soil quality, and precision agriculture. Proceedings of the National Academy of Sciences of the United States of America 96(11) 5929-5959.

Gee, G.W. and Bauder, J.W., 1986. Particle size analysis. In: Methods of Soil Analysis: Part I (2nd edition), Agronomy Monograph, vol. 9, edited by A. Klute, Madison, Wisconsin, USA: ASA and SSSA, pp. 383-411.

Gianello, C. and Bremner J. M., 1986. A simple chemical method of assessing potentially available organic nitrogen in soil. Communications in Soil Science and Plant Analysis 17(2) 195-214.

Goerres, J. H., Dichiaro, M. J., Lyons, J. B. and Amador, J. A., 1997. Spatial and temporal patterns of soil biological activity in a forest and an old field. Soil Biology and Biochemistry 30(2) 219-230.

Inamura T., Yoshikawa A., Ikenaga S. and Iida M., 2010. Path Analysis of Tiller Density of Winter Wheat Demonstrates the Importance of Practices that Manipulate Clod Size Based on Soil Moisture at Seeding in the Rice–Wheat Cropping System. Plant Production Science 13(1) 85-96.

Lamb, EG, Shirtliffe, S.J. and May, W.E., 2011. Structural equation modeling in the plant sciences: An example using yield components in oat. Canadian Journal of Plant Science 91(4) 603-619.

Pebesma, E.J. 2004. Multivariable geostatistics in S: The gstat package. Computers & Geosciences 30(7) 683-691.

R Development Core Team, 2018. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org. (last accessed 23/4/18)

Spink, J.H., Semere, T., Sparkes, D.L., Whaley, J.M., Foulkes, M.J., Clare, R.W. *et al.*, 2000. Effect of sowing date on the optimum plant density of winter wheat. Annals of Applied Biology 137(2) 179-188.

Timsina, J. and Connor, D. J., 2001. Productivity and management of rice–wheat cropping systems: issue and challenges. Field Crop Research 69(2) 93-132.

Yves R, 2012. lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software 48(2) 1-36.

Section 13
Robotics, guidance and automation

Overfitting a convolutional neural network to support annotations of weeds

M. Dyrmann*, R.N. Jørgensen, M.S. Laursen and S. Skovsen
Department of Engineering, Aarhus University, Finlandsgade 22, 8200 Aarhus N, Denmark; madsdyrmann@eng.au.dk

Abstract

Convolutional neural networks have gained interest in the field of precision agriculture including automated weed detection. However, plenty of annotated training data is required to regulate the network, which is a labour-intensive task to achieve. This paper proposes a field-specific model training, that updates during annotation to adapt to a given field and thereby support the annotator after only a few annotated plants. It was found that after only annotating the species of 30 plants from a field, the computer was on average able to classify more than 80% of the remaining plants correctly.

Keywords: deep learning, weed classification

Introduction

The optimal dosages of herbicides when spraying fields, depends on which weeds are present in the field. Studies have shown that savings can be achieved when the choice of herbicides is based on a dense inspection of weed populations in a field. Further savings can be achieved when the concentrations are optimized locally (Rydahl *et al.*, 2018); the herbicide dosages can be increased in highly weed infested areas and reduced when the weed pressure is low.
Inspecting fields should, however, be automated in order to be cost-effective.
Precision agriculture is increasingly using convolutional neural networks (Kamilaris *et al.,* 2018) to monitor fields due to their ability to recognize objects in environments containing great variety. Thus, convolutional neural networks have been able to detect weeds despite plants that are occluded (Dyrmann *et al.*, 2018, Lottes *et al.,* 2018); distinguishing mixed-crops (Mortensen *et al.*, 2016, Skovsen *et al.*, 2017); and classification of crops from satellite imagery (Ruβwurm and Körner, 2018).
Sufficient training data is, however, hard to achieve: Most implementations of convolutional neural networks require a large amount of training data to cover the variation within the field of operation. Therefore, when classifying weeds, the training data must cover the variations in growth-stage, leaf orientation and environmental variations. To comprise the common weed species, tens of thousands of labelled images are needed. However, some unforeseen plant variations can cause the neural network to fail at predicting the species correctly. In a fully automated application, where the herbicide dosages are based on the detected weed, a confusion between harmless weeds and harmful weeds can result in inexpedient herbicide use. Therefore, more than 60,000 images of weeds from Danish fields have so far been annotated manually to make a robust weed classification system. However, the classifier can sometimes fail at recognizing a known species if the new sample differs significantly from previously seen samples of that species.
The normal mantra in the field of machine learning, when a classifier is unable to classify a sample is: 'You just need more training data!'. However, annotating is a labour-intensive task that should be made as efficient as possible.
Therefore, a semi-supervised annotation system was set-up, which would off-load the annotator: An automated classifier predicts the species of a plant in an image, and a human expert would only have to correct the answer given by the computer if it is incorrect. In some cases, this system works

DOI 10.3920/978-90-8686-888-9_94, © Wageningen Academic Publishers 2019

well for a whole field but, at other times, a field is significantly different from previous fields and the classifiers can only distinguish monocots from dicots, but not the species of the weeds.
Even though there is some spatial variability in fields, the within-field variation in visual appearance of weeds is usually small compared to the intra-field variation as illustrated in Figure 1. Weeds of a given species are thus often present at the same development stage and have the same colour because they have undergone the same treatment (tillage, spraying) and grow in the same type of soil.
Therefore, an automated classifier is likely to either classify most of the weeds in a field correctly or fail at classifying most of the weeds in a field.
When new weed samples are aggregated for training the classifier, focus should be on weeds that the classifier fails at predicting rather than weeds that it already handles well. However, since the classifier fails at predicting them, there is not support from the semi-automated system when annotating weeds from the field from which samples are needed.
This paper proposes a field-specific model training, that gradually adapts to a given field during annotation. Since weeds in a single field have grown under comparable conditions, the variation between weeds is often limited, which is taken advantage of by 'overfitting' the trained model to the characteristics of only a few weed samples. The trained model will therefore support the annotator when annotating the weeds that otherwise were to be annotated completely manually.
The goal when training a classifier is usually to make it generalize, but here the goal is to train a model specifically for a single field, which can then be thrown away when the annotation is finished.

Methods and materials

Data

In this study, ten fields were selected for training/evaluation. These fields were all cereal fields, except for one maize-field. The images were taken top-down either using cell phones or using a flash-supported machine-vision camera (Laursen *et al.*, 2017) mounted onto an ATV. From these images, single weeds were cropped out manually or automatically using a weed detector algorithm



Figure 1. Weeds of a given species from a single field (Top row) tend to carry the same characteristics and, therefore, look more alike than the same species from other fields (Bottom row). A classifier that is trained on samples from one field is therefore more likely to classify new weed samples from the same field correctly than weeds from other fields.

(Dyrmann *et al.*, 2018), which resulted in between 701 and 4,127 annotated weeds per field (growth stage BBCH 10-19), that were used for this experiment. Random samples are shown in Figure 2 to illustrate the diversity in the data-set. The total number of annotated weeds in each of the ten fields is shown in Table 1.

Methods

To investigate how well a network can be trained to recognize plants in a field when training only a few samples from the field in question, the following training scheme was drawn up: First, a classifier was trained on 5 plants and the other plants from the field were used for testing the ability to recognize weed species. The training was then started over, but with an additional 5 plants in the training set. This was repeated until there were 100 plants in the training set, i.e. 20 training passes per field. As there were 10 fields, a total of 200 individually trained models were made.
Prior to the training, all weed images in a field were randomized and the samples were drawn randomly. Optimally, the species of the samples drawn for the training set should have the same distribution as those in the rest of the field, but this criteria can hardly be met, when drawing only a few plants. The procedure, therefore, resembles a normal annotation procedure, where there is no prior knowledge about the weed in the field. However, in order to compensate an imbalance in the number of training samples per weed species, the under-represented species in the training set contribute relatively more to the training loss.



Figure 2. Random weed-samples from the 10 test fields.

Table 1. Number of samples in each of the ten fields (A-J).

Field	A	B	C	D	E	F	G	H	I	J
Weed count	933	701	1,173	1,070	749	979	1,172	4,127	2,775	923

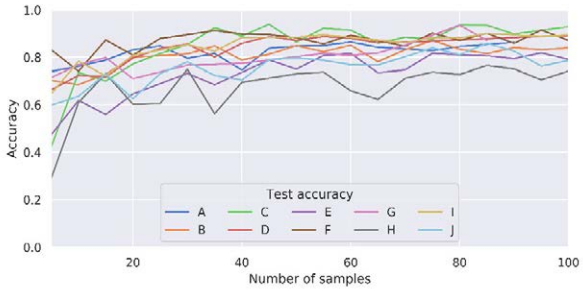
Each model was trained for the same number of iterations independently of the number of samples. Since the number of training samples varied, the number of epochs varied too.
Normally, in the given application, the network would be pre-trained on all previously annotated plants from other fields. However, to ensure reproducibility for other researchers, the network was firstly pre-trained on the publicly available ImageNet dataset (Russakovsky *et al.*, 2014), and then trained for five epochs on the publicly available plant seedlings dataset (Giselsson *et al.,* 2017).

Network architecture
A standard off-the-shelf residual neural network (Resnet50) (He *et al.*, 2015) convolutional neural network for classification was used as the classifier. A Resnet is characterized by a series of residual layers, which allow it to forward features from first hierarchical layers to the deeper layers while preventing vanishing gradient problems that normally occur with deep architectures (He *et al.*, 2015). During each training-pass, the final layers were modified to reflect the number of species in the training set. The weights of those new layers were initialized randomly.

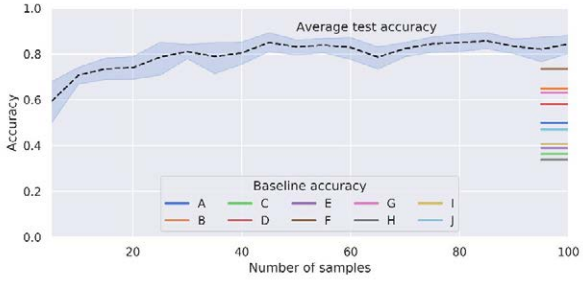
Data augmentation
Since few samples can hardly represent weeds of a whole field, random augmentation was added to the training images. This increased the variance in an otherwise limited training set, and the network could thus be rendered more immune to environmental changes or camera parameters, for example, intrinsic parameters, colour or saturation. All training images were therefore augmented as follows:
- Images were rotated and flipped randomly in 90 degree intervals to increase the rotation invariance.
- Image intensity, saturation and hue were adjusted to increase invariance towards changing light conditions.
- Images were sheared up to ±20 degree to increase the contextual dependency.
- Images were randomly cropped and scaled between 0.8 and 1.0 of the original size to increase the scale invariance.

Results and discussion

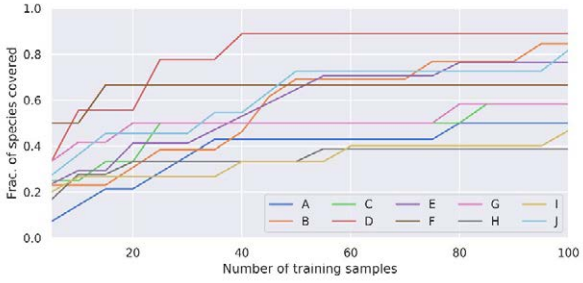
The results of each of the 200 trained models are summarised in Figure 3A. This figure shows that when only a few samples were annotated in a field, the species of the remaining weeds could be predicted correctly most of the time.
On average, more than 30 plants should be annotated manually before the classification accuracy exceed 80%, as depicted in Figure 3B, and the human expert would early on get time-saving support from the computer, as he would only have to correct the errors. However, a higher number of training samples benefits the classifier: the likelihood of covering all species increases with the number of training samples and it helps the classifier to differentiate between species, as it will not predict a species it has not yet seen.
The distribution of species in the test-set sets the baseline of the classifier, which is the accuracy that is achieved when always guessing on the most frequent species (even though that would be unknown in a real-world setup). The baselines for each field are marked in Figure 3B along with the average accuracy of the classifier. This Figure shows that with only a few annotated samples, the classifier outperforms the baselines significantly. The baselines change slightly (less than 2%) when increasing the number of training data, as the number of test-images decreases. Therefore, only the highest baselines are plotted.
The fraction of weed-species in the test-sets that are covered by the species in the training-sets is plotted in Figure 3C. The classifier can only predict species contained in the training set, and it will, therefore, always misclassify species that it has not seen before. However, from Figure 3A and Figure 3C, it can be seen that the overall classification accuracy was high when only a few samples



(a)



(b)



(c)

Figure 3. Test accuracy when increasing the number of training samples for field A - J. (A) accuracy on test sets when varying the number of samples for each field. (B) The average accuracy across fields. (C) The fraction of species covered in the test set, when varying the number of samples in the training set.

were used for training, while the fraction of species covered was low. The reason for this is dominant weed species in the selected fields, which account for most of the weeds.

Conclusion

Annotating data for a machine-learning algorithm can be tiresome. Especially in an agricultural field, where the variations in plant appearance make great demands on large datasets. Here a semi-automated annotation procedure is presented, which benefits from a classical convolutional neural network that can be trained online while annotating plants' species.

Instead of training a general model that covers all species in all fields, it has been shown, that by 'overfitting' a model on a few samples from single fields, the models can predict the classes correctly of most of the remaining samples from those fields.

Therefore, after only a few manually annotated weeds, the model can support annotating images that otherwise would have to be annotated manually. Tests show that for the ten selected fields, with only 30 manually annotated images, the computer can distinguish 80% of the remaining weed images in the fields.

Acknowledgements

This work was funded by The Innovation Fund, Denmark

References

Dyrmann, M., Skovsen, S., Laursen, M.S., and Jørgensen, R.N., 2018. Using a Fully Convolutional Neural Network for Detecting Locations of Weeds in Images from Cereal Fields. In The 14th International Conference on Precision Agriculture, 1-7. https://www.ispag.org/proceedings/?action=download&item=5081

Giselsson, T.M., Jørgensen, R.N., Jensen, P.K., Dyrmann, M., and Midtiby, H.S., 2017. A Public Image Database for Benchmark of Plant Seedling Classification Algorithms, November. https://vision.eng.au.dk/plant-seedlings-dataset Last accessed 2018-12-01

He, K., Zhang, X., Ren, S., and Sun, J., 2016. Deep Residual Learning for Image Recognition, In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA: IEEE,), pp 770-78, https://doi.org/10.1109/CVPR.2016.90.

Kamilaris, A., and Prenafeta-Boldú, F.X., 2018 Deep Learning in Agriculture: A Survey, Computers and Electronics in Agriculture 147 70-90, https://doi.org/10.1016/j.compag.2018.02.016.

Laursen, M.S., Jørgensen, R.N., Dyrmann, M., Poulsen, R.N., 2017. RoboWeedSupport – Sub Millimeter Weed Image Acquisition in Cereal Crops with Speeds up till 50 Km/H. International Journal of Biological, Biomolecular, Agricultural, Food and Biotechnological Engineering 11 (4):317-21. https://doi.org/10.5281/zenodo.1130297.

Lottes, P., Behley, J., Milioto, A., and Stachniss, C., 2018. Fully Convolutional Networks with Sequential Information for Robust Crop and Weed Detection in Precision Farming, IEEE Robotics and Automation Letters (RA-L), 3, 3097-3104. doi:10.1109/LRA.2018.2846289

Mortensen, A.K., Dyrmann, M., Karstoft, H., Jørgensen, R.N., and Gislum, R., 2016. Semantic Segmentation of Mixed Crops Using Deep Convolutional Neural Network, CIGR-AgEng Conference (Aarhus), 1-6.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., *et al.*, 2014. ImageNet Large Scale Visual Recognition Challenge, September. http://arxiv.org/abs/1409.0575.

Ruβwurm, M., and Körner, M., Multi-Temporal Land Cover Classification with Sequential Recurrent Encoders, ISPRS International Journal of Geo-Information 7, 4) 129, https://doi.org/10.3390/ijgi7040129.

Rydahl, P., Bojer, O. M., Jorgensen, R. N., Dyrmann, M., Andersen, P., Jensen, N.P., *et al.*, 2018. Spatial variability of optimized herbicide mixtures and dosages. In International Conference on Precision Agriculture. International Society of Precision Agriculture, https://www.ispag.org/proceedings/?action=abstract&id=5040

Skovsen, S.; Dyrmann, M.; Mortensen, A.K.; Eriksen, J.; Gislum, R.; Jørgensen, R.N. *et al.*, 2017. Estimation of the Botanical Composition of Clover-Grass Leys from RGB Images Using Data Simulation and Fully Convolutional Neural Networks. *Sensors 17*, 2930.

Hybrid topological location and mapping for autonomous agricultural robots

L. Emmi*, J. Dufour, V. Cadenat and M. Devy
LAAS, CNRS, 7 Avenue du Colonel Roche, 31400 Toulouse, France; luis.emmi@laas.fr

Abstract

This paper considers the problem of autonomous navigation in an agricultural field. It proposes a location and mapping framework based on semantic place classification and key location estimation, which together build a hybrid topological map. The approach has been assessed through off-line data, recorded in real conditions, on diverse fields, at different seasons. The results have shown the interest of the approach, which allows: (1) obtaining asimple and easy-to-update map; (2) the use of artificial landmarks to be avoided, thus (3) improvement of autonomy of agricultural robots.

Keywords: hybrid topological map, semantic identification, autonomous navigation

Introduction

In the last twenty years, an important number of robots have been developed to help in daily tasks (Royakkers and Van Est, 2015), including the management and care of crops. The development of research prototypes and commercial robots for agriculture has been recently growing (Ampatzidis *et al.*, 2017). Their autonomy has also been improved, reducing the need for human intervention. Many of these developments focus on solving the problem of navigation within the cultivated areas of a farm, where two main approaches are used (Bechar and Vigneault, 2016): GNSS-based navigation and sensor-based navigation. The first one gives the robot information about the path to be followed, although a precise prior mapping operation is required, commonly obtained through the use of real-time kinematic GNSS (RTK-GNSS). Despite its advantages, some drawbacks can be identified: (1) high cost; (2) need for up-to-date maps; and (3) vulnerability in certain areas (GNSS-denied zones). Sensor-based navigation relies on proximity sensors, such as cameras or LIDARs, for avoiding obstacles and following crop rows (Durand-Petiteville *et al.*, 2018). Also, although this topic has been deeply addressed in recent years, it still requires a prior knowledge of the beginning and end of the rows to ensure robust navigation. This crucial information is generally provided by the farmer, making most of the solutions specific to the considered field. Although some proposals have tackled this situation, there is still a need for general navigation strategies, less dependent on: (1) human intervention; (2) absolute location; and (3) controlled environment, allowing long-range navigation to be achieved.

This problem is at the core of this paper, motivated to improve the autonomy of Oz, a commercial weeding robot developed by Naïo Technologies (Anon., 2018), that requires the use of artificial landmarks for row boundary identification. Thus, this paper presents a location and mapping framework that takes advantage of the particular structure of the field to extract natural landmarks for precise and autonomous navigation. The framework presented consists of a perception system (LIDAR and RGB camera) that acquires and labels the information, which then is analysed by a semantic classifier to extract natural features. The perception system and the semantic classifier are accompanied by a map update process.

A field is generally structured by a sequence of rows connected through two headlands. This allows the working environment to be divided into a limited number of places, which are related to each other by specific rules. This type of problem can be described by a topological map, a commonly used strategy in in-door navigation to estimate the state (position and orientation) of a robot

DOI 10.3920/978-90-8686-888-9_95, © Wageningen Academic Publishers 2019

in a partitioned and semi-structured environment (Thrun and Bücken, 1996). In this sense, the framework presented includes a general modelling of the field expressed throughout a topological map. The semantic classifier benefits from this model to compare the acquired information and thus estimate the area where the robotic system is currently located, and the exact points where the transitions occur. Both elements together build a hybrid topological map. Some advantages can be highlighted for this type of map: (1) it captures the environment connectivity, while limiting the use of metric information; (2) the map is 'lighter' and easier to update; (3) absolute location sensors and previous mapping are no longer necessary; and (4) artificial landmarks are no longer required.

Related work

Currently, there are numerous companies that offer diverse mapping services, from delimiting the farm zones, to obtaining in detail several vegetation and soil indices. The mapping task is primarily done by the use of UAVs, satellite images or manually. Most solutions use dedicated software to analyze the data, and mainly rely on RTK-GNSS. Moreover, most of the maps that are created in these solutions use the contour of the field given by an external source or acquired with their own system, and generate the path that the tractor driver should follow for optimal operation. These are called full coverage maps, and are created independently of the seeded crop. Most of the maps that could be available via this kind of service might provide some information to support mobile robotic systems to autonomously navigate on a farm, such as seeder-extracted maps. However, the acquired data needs to be manually processed to be adapted for: (1) a specific robotic system; (2) a specific task; and (3) a specific environment.

One alternative to absolute mapping is the fusion of proximity sensors (i.e. LIDAR) with odometry and/or camera systems, which has been a broad subject of research. Most of the developed works seek to solve the problem of row following, and only some of them address the detection of the beginning and end of the crop row. Normally, artificial landmarks are installed for the detection of these limits (Zhang *et al.*, 2014). Although it is a suitable solution for perennial crops, manual work is required for their installation and removal in non-perennial crops. This restrains the autonomous capabilities of robotic systems. An alternative is using the information already provided by the natural environment, through the semantic extraction of key elements and an adequate modeling of the working environment. Related to this subject, a specific work has aroused interest, where semantic classification of the meaning of a field is proposed (Weiss and Biber, 2010). They used a particle filter to determine the probability distribution of the robot state in the different locations where it is applied. It should be noted that in this approach: (1) there are no limits on the meaning of open field; (2) the row-start and row-end areas are not specific points; and (3) it is not clear the difference between row-start and row-end areas, which hinders its use to accurately define the transition between the row and the open field.

The field model

In general, a field is a cultivated area consisting of one type of crop, planted in parallel and at a relatively constant distance between them. A field can be divided into four places, named: headland, lane, alley, and gate (Figure 1A). The headland is where the robot switches between lanes or alleys, or exits the field. The alley and lane represents the cultivated area, although the alley corresponds to the lateral boundaries. Both could include more than one inter-row space. Moreover, their width corresponds to the width of the tool. Finally, the gate allows the robot to enter/quit the field.

These places represent the possible robot *states* (Figure 1B), i.e. a semantic representation of the location of the robot in the field. A fifth *state* called 'unknown' has been included to handle the uncertainties when sensory data is not conclusive. At this step, the proposed topological map only allows the system to instantaneously distinguish in which type of place (*state*) the robot is present, but it is not able to predict future *states* or where transitions are located. This kind of map may not bring a sufficient accuracy to efficiently control the tools. Therefore, a metric map is included

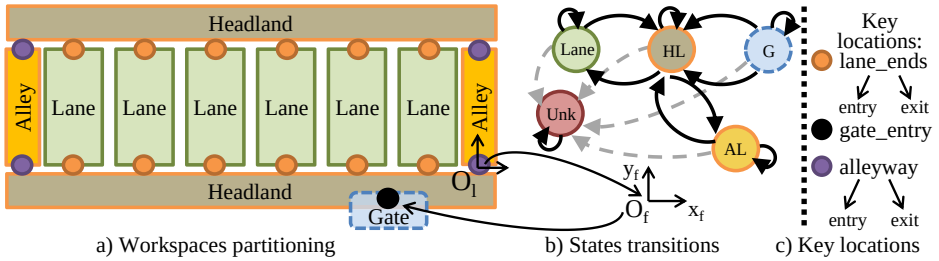


Figure 1. The proposed field model. (A) The semantic field place partitioning. (B) Topological representation of the workspaces. (C) Key location types.

by estimating the exact coordinates of the transitions, called key locations (Figure 1C). The key locations are represented by points $p(x_p, y_p, z_p, \vec{v}) \in \mathbb{R}^3$, where $\vec{v}$ is a unit vector that defines the orientation of p, related to $F_f = (O_f, \vec{x}_f, \vec{y}_f, \vec{z}_f)$. F_f is the field coordinate frame located at the gate. A temporary reference frame $F_l = (O_l, \vec{x}_l, \vec{y}_l, \vec{z}_l)$, related to F_f, serves to keep the perception system information referenced close to the last key location identified. Then, F_l is relocated whenever a *state* transition is detected. This procedure allows the drift caused by the odometry to be minimized, with the drawback that the previous key locations lose validity within F_f. On the other hand, in the map update process, the key locations can be referenced to a global reference frame. The fusion of the semantic identification and the estimation of the key locations builds the hybrid topological map.

Algorithm for semantic location and mapping

Perception system: data acquisition and labelling

Figure 2 presents the algorithm developed. The algorithm looks for valid key locations by analysing the distribution of the crops, encompassed in clusters of laser point. The type of crop is the sole required configuration input, and the laser points are labelled depending on if they belong to that

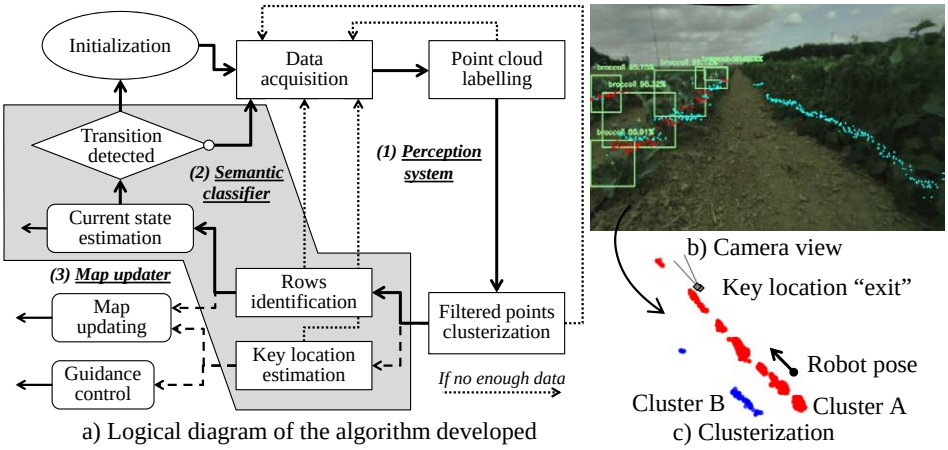


Figure 2. Algorithm for semantic location and mapping. (A) Logical scheme for state estimation and key location estimation. (B) Laser points labelling based on crop detection. (C) Laser points clusterization and key location estimation.

crop (Figure 2B). This allows the alleys to be identified. The laser points are placed in the same reference frame as the camera. Then, the points that are within the positive detection are labelled (Figure 2B). Since the labelling process requires certain computational load, several epochs of laser data are labelled from a single camera image. The labelled points are referenced into F_l which is relocated each time the algorithm is initialized. Then, the points are clustered based on DBSCAN (density-based spatial clustering of application with noise).
Each cluster represents a crop row, or part of one, given that gaps can be found (Figure 2C). The clusters are kept stored until the system detects a change in the *state*. Therefore, every time new data is acquired, the stored clusters are updated. When a change in *state* is detected, the process is restarted, i.e. the clusters are discarded and a new F_l is defined. This procedure provides enough information to enrich the detection process of the key locations, even if the robot location system is noisy.
For laser points labelling, an object identifier based on a classical convolutional neural network (CNN) called YOLOv3 (Redmon and Farhadi, 2016) has been developed. It has been trained to identify an important variety of market garden crops. The training was performed using approximately 14,000 examples (4/5 of the data for training, and 1/5 for testing). The samples were extracted from the on-board camera of the Oz robot, performing real operations, in different seasons, farms, light conditions, etc. (crops and number of samples: green lettuce: 8,258; red lettuce: 1,808; cabbages: 1,301; leeks/onions: 1,444; rapeseed: 1,135). Each sample was annotated by a hand-drawn bounding box and each annotation consists of an image where only the desired crop appears. To evaluate the selected network and the ability to correctly classify the crops, a comparative study has been performed with RetinaNet – Resnet152 (Lin *et al.*, 2017), another common network for object classification, which was also been trained with the same data. Similar performances have been obtained on both networks, in terms of average precision (AP) and mean average precision (MAP). A threshold of about 0.5 of the intersection over union (IoU) was used for performance analysis. Nevertheless, YOLOv3 offers better results in terms of performance P (on a CPU and a GPU), which allows it to be implemented on the Oz robot without installing any extra or expensive hardware. Table 1 presents results in terms of detection capacity and processing time.

Semantic identification and key location estimation
For this work, only the identification of the headland (n=H), lane (n=L) and alley (n=A) are considered. For the two situations (Figure 3), left and right points are estimated, placing them at a distance d of the nearest cluster. If both points are located, a key location is set at the midpoint between both (*lane_end*). Non-valid key locations are discarded, and only those that are more likely to represent the *state* transitions are maintained. They are named *valid key locations*, and are classified as entry (t=En) or exit (t=Ex). Two sorts of indices are calculated: key_idx_t, corresponding to a correlation between the same types of *valid key location*, and angle_idx_t, corresponding to the correlation between the robot orientation and the last *valid key location*. For every *state*, an index c_idx_n is calculated, given by Equations (1) and (2):

Table 1. Comparison of two CNNs trained for crop detection. Detection is considered true positive (TP) if the IoU between ground truth and detection is greater than 50%.

CNN	AP-.50					MAP-.50	P [ms]	
	cabbage & broccoli	rape-seed	green lettuce	red lettuce	onion / leek		i7 6600U	GTX 1080
YOLOv3	0.871	0.676	0.966	0.957	0.499	0.794	1,460.7	56.9
RetinaNet	0.86	0.73	0.966	0.958	0.488	0.8	6,124.8	72.4

$$\begin{cases} c_idx_H = \frac{key_idx_En}{angle_idx_En} \times cluster_idx \\ c_idx_A = \frac{angle_idx_Ex}{cluster_idx} \times alley_idx \\ c_idx_L = \frac{angle_idx_Ex}{cluster_idx} \times lane_idx \end{cases} \begin{matrix} if\ left \cup right \\ alley_idx = 0 \\ lane_idx = 1 \\ \\ else \\ alley_idx = 1 \\ lane_idx = 0 \end{matrix} \quad (1)$$

$$cluster_idx = abs(B_{dist} - F_{dist})/euc_dis(FC, BC) \quad (2)$$

Equation 2 expresses how misaligned the robot is with respect to the nearest cluster, calculated by projecting a forward *FR* and backward *BR* points, and the nearby cluster points are located, designated as *FC* and *BC* respectively (Figure 3A). Then, the orthogonal distances (B_{dist},F_{dist}) between the projected points and the cluster are calculated, where *euc_dis* represents the Euclidean distance between two points. These indices give an idea of how probable the robot is located in the topological map. Also, the *state* estimation is improved by taking into account the possible transitions allowed and the previous estimation. For this, a simplification of the state transition probability given a motion presented in Weiss and Biber (2010) is used, since the representation of the environment and the relationship between the different *states* is similar.

Guidance control and map update

Given the current *state* estimation and the *valid key location*, the navigation strategy can be adjusted to meet the objective provided by the topological map and the general mission. Identification of the *valid key location* allows the exact co-ordinate where a *state* transition occurs to be predicted. Work within the field is also performed with greater efficiency and precision. The navigation strategy then relies on the search of possible lane candidates that require to be treated. Therefore, each time a *state* transition is identified, the last *valid key location* is stored into the map, referenced to F_f (Figure 1A). This is only possible if a global location system is available (i.e. GNSS), even if the uncertainty of this measurement is bigger than the distance between two nearest key locations. This map can be used to give the robotic system an idea of the necessary orientation to find a specific node within the hybrid topological map. Although, once close to that location, the approach presented should be used for precise navigation.

Results and discussion

The approach proposed has been tested and validated off-line with seven diverse dataset extracted from the Oz robot, recorded in real working conditions. Each dataset contains a set of asynchronous raw images, as well as the laser point-cloud referenced with the odometry of the robot. Five were collected in the same farm, on two different days, for two different types of crop (cabbage and

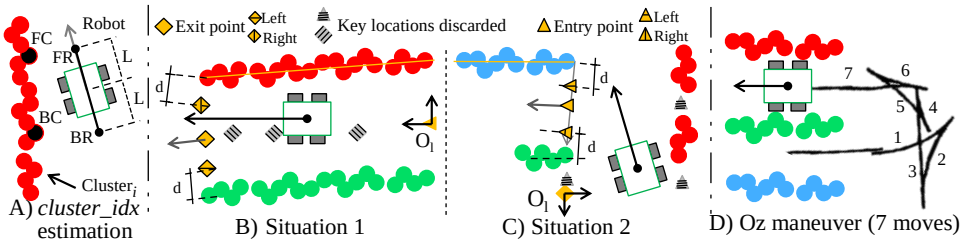


Figure 3. Example of key location extraction. (A) *a_index* points definition (*L*=1 m). (B) Looking for exit. (C) Looking for entry (*d*=0.1 m). (D) Example of Oz maneuvers.

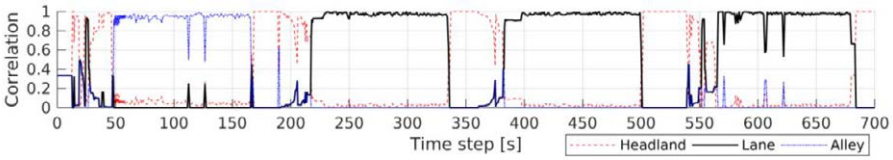


Figure 4. Correlation index c_idx_n for each state studied (dataset B1).

broccoli). The remaining set was obtained for different seasons in a field of leeks located in a greenhouse. Thus, these data have been acquired during various real situations: high changes in lighting conditions (different time of day), different crops, different locations, etc. Figure 4 shows the correlation indices c_idx_n for one of the passes (a broccoli field consisting of four crop rows with a length of about 50 m). The maxima of the indices are used to estimate the *state*, including previous estimation and transitions allowed. Figure 5 presents the *state* estimation based on c_idx_n and compared it with the ground truth.

To define the ground truth, the positions of the artificial landmarks (red sticks) required by the Oz robot to navigate are used. It must be noted that, in some cases, these marks might not exactly be placed on the last crop, inducing a small error. In addition, every time the robot enters a lane, it performed a maneuver composed of several forward and backward movements (between five and seven, Figure 3C). Thus, in some cases, the perception system successively gets in and out of the lane, which generates false positive detections (see red circles on Figure 5). Indeed, this approach assumes forward motion only (there is no information about the motion direction). Despite these perturbations, the system is able to works properly. Along with the semantic robot location, the key locations are also estimated. Figure 6 shows the odometry of the robot and the *valid key locations* for entry and exit. For the purpose of comparing the results, the *valid key locations* are also referenced with the GNSS sensor on-board the Oz robot, which may allow creating a map of the field to be used in the future. This map will not guarantee accurate and safe navigation given the GNSS accuracy limitations (±0.6 m of positioning error), but in conjunction with the semantic location in the topological map, the navigation strategy through the entire farm should be possible. Table 2 presents the analysis of the detection success rate (DSR), in addition to the mean distance error (MDE) between the detected key locations and the ground truth. Moreover, the orthogonal distance error (Orth_E) is analysed to identify if a key location was estimated in a different lane. Even with the error introduced in the definition of the ground truth, the results obtained are encouraging, demonstrating that it is possible to locate the rows boundaries without the use of artificial landmarks.

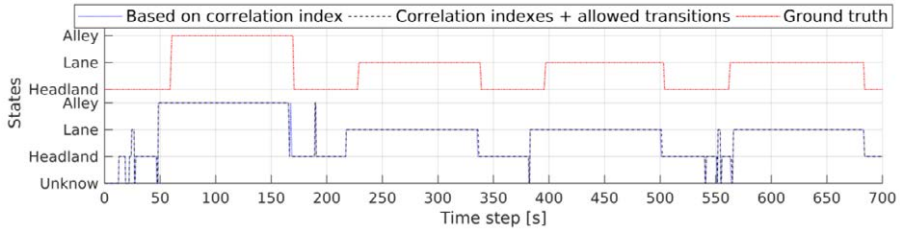


Figure 5. State estimation based on only the correlation index (continuous line) and on the previous estimation and transitions allowed (dotted line). Dataset B1.

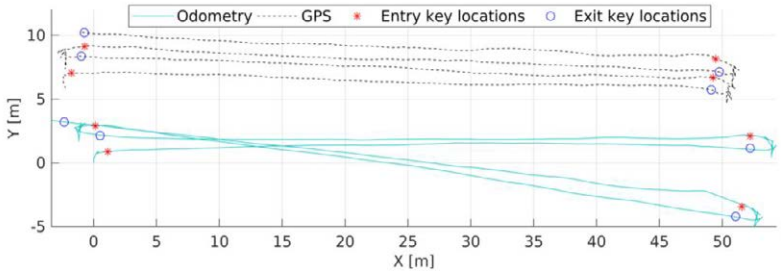


Figure 6. Odometry of the robot and GNSS with key locations estimated for dataset B1.

Conclusions

This paper considers the problem of autonomous navigation through an agricultural field. First, a model of the field is introduced to integrate it into a hybrid topological map. This map is part of the location and mapping framework presented. This approach introduces an alternative for absolute location or the use of artificial landmarks, which are not always reliable systems for navigation. The approach presented benefits from the structure of the field to identify indistinguishable characteristics that allow the instantaneous robotic pose to be estimated. Moreover, the hybrid attribute of the map also makes possible the state transitions to be predicted. This approach has been validated using off-line data, recorded in real working conditions, from diverse fields with different crops. The results have demonstrated the efficiency and the interest of the approach proposed, paving the way for low-cost and long-range navigation in agricultural environments.

Acknowledgements

The authors would like to thank colleagues from Naïo Technologies, Toulouse, for their participation to this work, through the collaborative project DESHERB'EUR funded by the program 'Investment for the Future' of the French government.

Table 2. State estimation and key locations results, for dataset: B for broccoli; C for cabbage; and L for leeks. For every key location type, a MSE of its position as well as the orthogonal projection error (Orth_E) are calculated compared with the ground truth.

Set	MDE [m]			Orth_E [m]			Transitions		DSR
	Entry	Exit	Mean	Entry	Exit	Mean	Entry	Exit	
B1	0.36	0.151	0.256	0.083	0.038	0.06	4	4	0.829
B2	0.829	0.533	0.681	0.3	0.187	0.244	4	4	0.875
C1	0.478	0.473	0.475	0.006	0.115	0.089	3	4	0.86
C2	0.655	1.153	0.904	0.052	0.137	0.095	3	4	0.923
C3	0.647	0.275	0.461	0.334	0.089	0.212	3	4	0.804
L1	0.921	0.868	0.895	0.148	0.458	0.301	8	7	0.805
L2	0.062	0.343	0.203	0.037	0.038	0.034	2	2	0.637

References

Ampatzidis, Y., De Bellis, L., and Luvisi, A., 2017. iPathology: robotic applications and management of plants and plant diseases. *Sustainability*, 9 (6), 1010.

Anon., 2018. Robots agricoles Naïo Technologies. https://www.naio-technologies.com (last accessed October 2018).

Bechar, A., and Vigneault, C., 2016. Agricultural robots for field operations: Concepts and components. *Biosystems Engineering*, 149, 94-111.

Durand-Petiteville, A., Le Flecher, E., Cadenat, V., Sentenac, T., and Vougioukas, S., 2018. Tree Detection With Low-Cost Three-Dimensional Sensors for Autonomous Navigation in Orchards. *IEEE RAL*, 3(4), 3876-3883.

Lin, T., Goyal, P., Girshick, R., He K., and Dollár, P., 2017. Focal Loss for Dense Object Detection. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2999-3007.

Redmon, J., and Farhadi, A., 2016. Yolo9000: Better, faster, stronger, *arXiv preprint arXiv: 1612.08242.*

Royakkers, L., and van Est, R., 2015. A literature review on new robotics: automation from love to war. *International Journal of Social Robotics*, 7 (5), 549-570.

Thrun, S., and Bücken, A., 1996. Integrating grid-based and topological maps for mobile robot navigation. In *Proceedings of the National Conference on Artificial Intelligence.* Menlo Park, CA, uSA: AAAI Press/MIT Press, pp 944-951.

Weiss, U., and Biber, P., 2010. Semantic place classification and mapping for autonomous agricultural robots. In *Proceedings of IROS Workshop on Semantic Mapping and Autonomous, Knowledge Acquisition*, Taipei, Taiwan.

Zhang, J., Maeta, S., Bergerman, M., and Singh, S., 2014. Mapping orchards for autonomous navigation. Paper No. 14-1838567, St Joseph, MI, USA: ASABE

Determination of a field boundary using ISO 11783 and D-GNSS data acquired during a ploughing operation

A. Heiβ*, D.S. Paraforos and H.W. Griepentrog
University of Hohenheim, Institute of Agricultural Engineering, Technology in Crop Production, Garbenstr. 9, 70599 Stuttgart, Germany; andreas.heiss@uni-hohenheim.de

Abstract

Easily available and detailed area-related information is very valuable for the optimization of crop production processes. In the present paper, historical data from ploughing were analysed to gain precise information about the field boundary. The data had been acquired from a tractor's ISO 11783 communication network and georeferenced using a low-cost differential global navigation satellite system (D-GNSS) receiver. After the exclusion of non-work sequences, single cultivated tracks were reconstructed which represented, as a whole, the cultivated area of a field. Based on that, the boundary was automatically detected in terms of location and size, with a deviation of 1.4% from the documented field size.

Keywords: ISO 11783, low-cost D-GNSS, process optimization, fleet management

Introduction

Management processes in agriculture become more complex by e.g. implementing precision farming technologies, machinery sharing or contracting (Steinberger *et al.*, 2009). Furthermore, machinery operation has a remarkable effect on the total cost of crop production (Adamchuk *et al.*, 2004) and it is affecting other cost factors like time and farm inputs. Easily available and detailed area-related information could help farmers, contractors or machinery rings to facilitate documentation and invoicing. Beyond that, it is helpful to analyse crop production processes and thus uncover inefficiencies. Tsiropoulos *et al.* (2013) emphasized the value of performance analyses for tractor-implement combinations.

The increasing implementation of ISO 11783 (commonly designated as ISOBUS) compliant machine communication networks, as well as affordable positioning systems using global navigation satellite systems (GNSS), are promising data sources for an information-driven crop production. However, there is still a need for algorithms and methods to handle the increasing amount of data that can and will be collected to benefit from this progress (Kortenbruck *et al.*, 2017).

The boundary encompassing the arable area within a field is a valuable information for e.g. documentation purposes and it can be set in relation to the cultivated area to assess the overlapping. Usually, this field boundary complies with the official field size and location as it is documented in the farm's field catalogue. Recent scientific works have already dealt with automated field recognition (e.g. Paraforos *et al.*, 2017b). However, there is still a lot of manual effort necessary for farmers to define and respectively deposit the field boundaries in digital field records. These tasks are usually inconvenient, because they are time-consuming and error-prone and partly demand ever new familiarisation with the operation. Commercial solutions for an automated derivation of field boundaries from continuous GNSS data logs can already facilitate these processes. As they usually don't use additional information about the employed machine, they have limited accuracy. To meet the requirements for e.g. subsequent area-related analyses, deviations of a few percent from the actual size would be desirable.

With the aim of automatically obtaining precise information about the field boundary during fieldwork, historical ISOBUS and D-GNSS position data from ploughing with a mounted reversible

DOI 10.3920/978-90-8686-888-9_96, © Wageningen Academic Publishers 2019

mouldboard plough were analysed in this work. As the determination of the boundary's location and size was supposed to be based on the actual covered area, a high accuracy, which is mainly limited by the used georeferencing, should be achieved. The main contribution of this work lies in providing methods for an automated derivation of highly accurate information about the field boundary from continuous data records.

Materials and methods

Instrumentation

Data were acquired during ploughing with a four-furrow reversible VariOpal 7 mouldboard plough (Lemken GmbH & Co. KG, Alpen, Germany) mounted on a 6210 R tractor (Deere & Company, Moline, IL, USA). The working width of the plough was hydraulically adjustable to a maximum value of 2.16 m. To record specific J1939 and ISOBUS communication data, a GL2000 CAN-Bus data logger (Vector Informatik GmbH, Stuttgart, Germany) was configured and connected to the diagnostic interface of the tractor. The acquired data were georeferenced as the logger was recording also absolute positioning from a low-cost D-GNSS receiver with a specified Circular Error Probable (CEP) of 2 m. The receiver's antenna was placed on the tractor cabin, on the longitudinal axis of the tractor. The correction data originated from the European Geostationary Navigation Overlay Service (EGNOS). The logged data were stored on a 2 GB storage card and when triggered by the operator, they were wirelessly transmitted to a cloud-based server for further processing. The raw data analysed for the present research work originate from the work performed by Paraforos *et al.* (2017a,b), where also the data acquisition system is described in detail.

Experimental data

The logged messages were available in batches in files in the American Standard Code for Information Interchange (ASCII) format, with one message per row. For the whole post-processing of the data, the numerical computing environment MATLAB R2016b was used (The MathWorks Inc., Natick, Massachusetts, USA). The files were parsed according to timestamp, channel, parameter group number (PGN) according to ISO (2015) or specific ID, mode, number of bytes and data bytes. Afterwards, the necessary messages were split to get information about the parameters of interest. Those were the wheel-based machine speed (WBMS) deduced from the message wheel-based speed and distance (WBSD) with the PGN 'FE48', the rear hitch position (RHP) deduced from primary or rear hitch status (RHS) with the PGN 'FE45', as well as the D-GNSS latitude and longitude deduced from Msg2 (specific ID from the logger's manufacturer: '1FFFFFD1'). Furthermore, the time stamps of those parameters were extracted.

According to ISO (2015), the WBMS has the suspect parameter number (SPN) '1862' and it is indicated in m/s. The RHP (SPN '1873') is defined as the measured position of the rear three-point-linkage, expressed as a percentage of full travel, whereby 0% indicates the full down position, and 100% the full up position. The D-GNSS latitude and longitude were defined in degrees according to the World Geodetic System 84 (WGS 84). A latency of eight seconds was noticed in the D-GNSS recorded data and thus, the algorithm was adapted to subtract these from every D-GNSS time stamp. The raw D-GNSS position data of the examined data sets are presented in Figure 1, indicated as green dots. They can be assigned to the field 'Schlag 5' of the farm examined, with a size of 2.90 ha indicated in the field catalogue. In Figure 2, the flowchart for the main steps of the data analysis is presented. The single steps are outlined in the subsequent subchapters.

Data editing

Both, RHP and WBMS data had been recorded with a frequency of 10 Hz, whereas D-GNSS data were available at a frequency of 1 Hz. To enable a concurrent analysis at a specific point and to increase the resolution of the D-GNSS information, a linear interpolation of the WBMS as well as



Figure 1. Satellite view with the raw D-GNSS position data from the examined data sets.

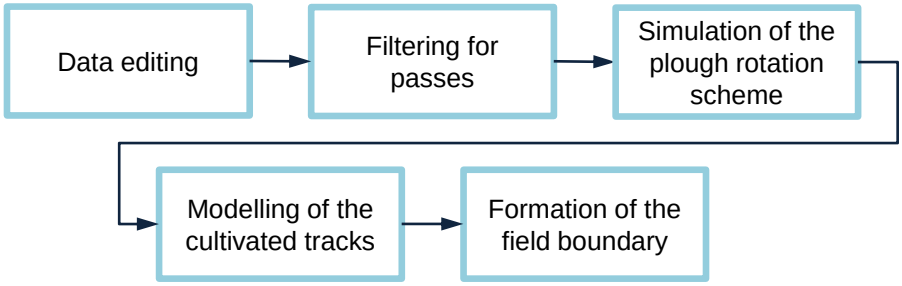


Figure 2. Flowchart for the main analysis steps.

the DGNSS latitude and longitude was performed. The RHP time stamps were the query points, whereas the sample points were the time stamps of the WBMS and D-GNSS. As the positioning error of the D-GNSS receiver, and excitation of the tractor cabin, caused a jittering of the recorded track, a Savitzky-Golay filtering was applied separately on the lat/long values.

Filtering for passes

A pass was supposed to contain all the in-work points from a lowering up to a lifting of the plough. The in-work points were defined as the timestamps with their corresponding RHP, WBMS and D-GNSS information, where the plough was assumed to be set in operation. The analysis was starting with the second value of each data set, as every RHP value was compared to the preceding one. As soon as it fell off a threshold of 60% from one value to the next, this event was recognized as a lowering point. As an in-work condition for the subsequent points, the WBMS at a query point had to exceed a threshold of 0.134 m/s. That helped to exclude points with an erroneous D-GNSS positioning. As soon as the RHP exceeded again the threshold of 60%, a lifting event was detected. To facilitate geometric calculations, the D-GNSS information assigned to in-work points was transformed from latitude and longitude to northing and easting using the Universal Transverse Mercator (UTM) projection (Schimel, 2011).

Simulation of the plough rotation scheme

An optimal adjustment of the plough during the data acquisition was assumed. As the logged data couldn't deliver exact information about the work state, assumptions were made also concerning the main parameters. The working width was assumed to be set always to its maximum during the data acquisition. Concerning the rotation status, a simplified scheme was developed and applied to simulate this special characteristic.

In preparation to this scheme, an identification of the different field parts was performed. First visual assessments of the data indicated that the main part passes were running parallel to the longer field edge, whereas the passes of two headlands were running transversely. This is a common pattern for soil tillage operations. The differentiation of these parts was performed as follows: First, the rough direction of every pass was determined by creating a vector from the lowering to the lifting point. Second, the angle between each vector of a pass and the vector of the very first pass, which was at the same time the first main part pass, was determined using basic vector algebra. In the case that the angle was between 45 and 135°, the arbitrary pass was declared as a headland pass. In any other case, it was declared as a main part pass. Figure 3 is schematically illustrating the field part identification. To distinguish the headlands, a k-means clustering was performed on their D-GNSS data.

In the next step to apply the simplified rotation scheme, for every field part, a differentiation of passes having the same direction as the first pass and passes running in the opposite direction was performed. The angle between the vector of the very first and all other passes of a field part was determined. If the angle was between 0 and 45°, the pass was declared as running in the same direction as the first pass. Then, the plough was assumed to be rotated right, which means clockwise in relation to the forward direction. If the angle between the current pass and the first pass was between 135 and 180°, it was assumed to be rotated left, which means counter-clockwise in relation to the forward direction. A differentiation of the two rotation positions is important, because they represent two mirror-inverted arrangements of the working tools.

Modelling of the cultivated tracks

As a preparatory step, relevant sizes were taken from the real tractor-implement combination. For every rotation position, two points were defined that represented the lateral edges of the working width. They were at the same height above the ground as the D-GNSS antenna. To determine their

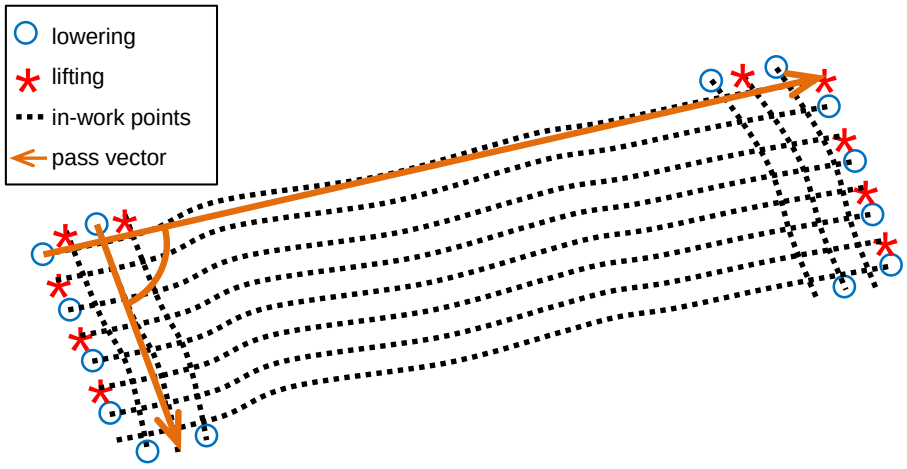


Figure 3. Schematic illustration of the field part identification.

relative position, vectors were created that originated in the position of the antenna and pointed to the representative points. The length of the vectors was calculated, as well as the angle to the forward direction of the tractor. For this, basic trigonometric equations, as well as the Theorem of Pythagoras, were applied. Figure 4 is schematically illustrating the procedure of the relative positioning of the representative plough points.

In Figure 5, the modelling of the cultivated tracks is schematically illustrated for three example passes. To determine the direction of movement, for every point of a pass, a vector was created, which originated in the D-GNSS position and pointed to the position of the chronologically subsequent point. Using the previously outlined measures for the points representing the edges of the plough and setting them in relation to a static unit vector, which was parallel to the UTM easting and pointing eastwards, the absolute position of the representative points could be determined for every antenna position within a pass. To model the movement of the plough during work and thus smooth the track, again a Savitzky-Golay filter was applied on the D-GNSS positions of the representative plough points. All the points could finally be connected to a polygon that was supposed to represent the cultivated track of a specific pass.

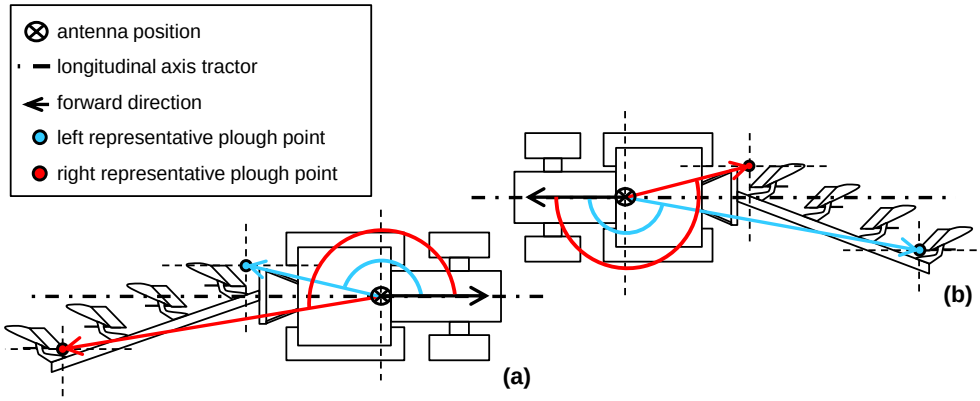


Figure 4. Schematic illustration of the relative positioning of the representative plough points (A) for the plough rotated left and (B) for the plough rotated right.

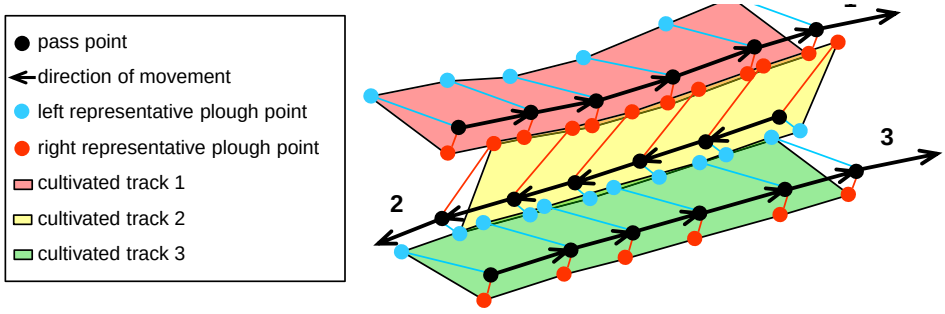


Figure 5. Schematic illustration of the modelling of the cultivated tracks.

Formation of the field boundary

Automatic determination of the boundary was configured to give D-GNSS co-ordinates of the outermost representative plough points encompassing all cultivated tracks within the field. A polygon was formed using the boundary points and its area was calculated.

To enable an examination of the boundary within a geographic information system (GIS), a Keyhole Markup Language (KML) file was created, which contained all the WGS 84 co-ordinates of the boundary points. Thus, the UTM northing and easting of the boundary points had to be transformed back (Palacious and Nindl, 2013).

Results and discussion

In Figure 6, the modelled cultivated tracks as well as the boundary are presented (A) for the whole field and (B) for a zoomed area in the north. The cultivated tracks of the main part run roughly along a southwest-northeast axis. In Figure 6B, it is obvious that the tracks of the headland in the northeast

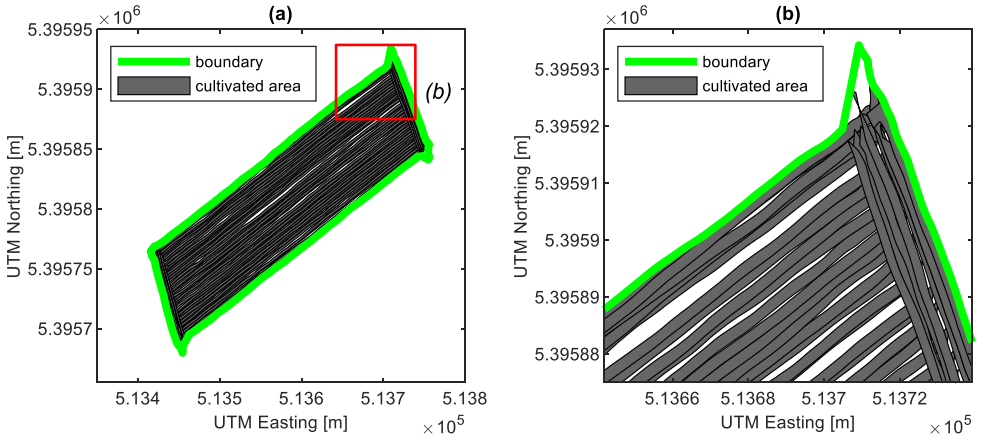


Figure 6. Modelled cultivated tracks and detected field boundary (A) for the entire examined field and (B) for a zoomed area in the north part of the field.



Figure 7. Satellite view with the detected field boundary.

are running roughly perpendicular. The boundary is encompassing the area defined by the cultivated tracks. The size of the boundary, as this was calculated by the developed algorithm, had a value of 2.94 ha. This resulted in an overestimation of 1.4% compared to the field size as documented in the farm's field catalogue. Figure 7 shows a satellite view with the detected field boundary indicated as a green line. The determined boundary fits to the actual field location.
Main influences negatively affecting the determination of the boundary are apparent in Figure 6B. For the formation of the boundary, the algorithm couldn't deliver the closest fit possible to the outermost representative plough points. This has caused an overestimation of the boundary's size. Beyond that, there are partly noticeable distortions at the tops of several headland tracks, which was a further source of error. They can be traced back to an increasing positioning error with decreasing rover speed. This phenomenon was examined in detail by e.g. Ehsani *et al.* (2003). Furthermore, the relative error of the D-GNSS positioning caused significant overlaps between the tracks alternating with gaps that seem to be left uncultivated. The behaviour of the pass-to-pass error of low-cost receivers was treated by e.g. Han *et al.* (2004). At the outermost tracks defining the boundary, this had a further negative effect. Context-related matter is e.g. a rolling of the tractor while driving on a slope along the field edge. Finally, the incomplete information about the work state of the plough is assumed to have an influence. Especially at the beginning and end of a field part, a deviation from the plough rotation scheme or a variation in working width is not unusual. The error sources could not be monitored and corrected. However, they can be minimized in future work by applying e.g. an improved positioning, as well as deriving more detailed information about the actual work state of the implement using additional sensors, if necessary.

Conclusions

A methodology was implemented that enabled a differentiation of work sequences from non-work sequences using a few parameters retrieved from a tractor's ISOBUS. A field part identification was applied that enabled the implementation of a simplifying assumption concerning the plough rotation scheme. An algorithm was implemented that automatically detected the field boundary in terms of location and size, based on the modelled cultivated area. Incomplete information about the actual work state of the implement, an erroneous D-GNSS positioning as well as a limited fitting by the algorithm were identified as the main sources of error for the determination of the field boundary. Despite these errors, the boundary's size was automatically detected with an error of 1.4%, which indicates the efficiency of the algorithm.
As it reduces manual input, an automated, detailed detection of a field boundary can be very valuable for contractors and farmers. An added value can be expected especially for e.g. documentation of agricultural tasks within networked software infrastructures. Furthermore, the boundary size can serve as a reference value to assess e.g. overlapping. A transfer of the methodology to other plant production processes seems feasible.

Acknowledgements

This work was financially supported by the German Federal Ministry of Food and Agriculture (BMEL) through the Federal Office for Agriculture and Food (BLE), grant numbers 2815ERA01H and 281B300116.

References

Adamchuk, V. I., Grisso, R. D. and Kocher, M. F., 2004. Machinery performance assessment based on records of geographic position. Paper No 041149. St. Joseph, Michigan, USA: ASAE.

Ehsani, M. R., Sullivan, M. D., Zimmerman, T. L. and Stombaugh, T., 2003. Evaluating the dynamic accuracy of low-cost GPS receivers. Paper No 031014. St. Joseph, Michigan, USA: ASAE.

Han, S., Zhang, Q., Noh, H. and Shin, B., 2004. A dynamic performance evaluation method for DGPS receivers under linear parallel-tracking applications. Transactions of the ASAE 47 (1) 321-329.

ISO, 2015. ISO 11783-7:2015 Tractors and machinery for agriculture and forestry -- Serial control and communications data network -- Part 7: Implement messages application layer.

Kortenbruck, D., Griepentrog, H. W. and Paraforos, D.S., 2017. Machine operation profiles generated from ISO 11783 communication data. Computers and Electronics in Agriculture 140 227-236.

Palacious, R. and Nindl, E., 2013. Matlab File Exchange: utm2deg (version 1.0). www.mathworks.com/matlabcentral/fileexchange/10914-utm2deg (last accessed 06/02/19).

Paraforos, D. S., Vassiliadis, V., Kortenbruck, D., Stamkopoulos, K., Ziogas, V., Sapounas, A. A. *et al.,* 2017a. Automating the process of importing data into an FMIS using information from tractor's CAN-Bus communication. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences 8 (02) 650-655.

Paraforos, D. S., Vassiliadis, V., Kortenbruck, D., Stamkopoulos, K., Ziogas, V., Sapounas, A. A. *et al.,* 2017b. Multi-level automation of farm management information systems. Computers and Electronics in Agriculture (142) 504-514.

Schimel, A., 2011. Matlab File Exchange: wgs2utm (version 2). www.mathworks.com/matlabcentral/fileexchange/14804-wgs2utm--version-2- (last accessed 06/02/19).

Steinberger, G., Rothmund, M. and Auernhammer, H., 2009. Mobile farm equipment as a data source in an agricultural service architecture. Computers and Electronics in Agriculture 65 (2) 238-246.

Tsiropoulos, Z., Fountas, S., Gemtos, T., Gravalos, I. and Paraforos, D., 2013. Management information system for spatial analysis of tractor- implement draft forces. In: J.V. Stafford (Ed.) Precision agriculture '13: Proceedings of the 9th European Conference on Precision Agriculture, Wageningen, the Netherlands: Wageningen Academic Publishers, pp. 349-356.

Proof-of-concept modular robot platform for cauliflower harvesting

F.B. Klein[1*], A. Wilmot[1], V.F. de Tejada[1], B.L. Rodriguez[1], I. Requena[3], S. Busch[1], A. Rondepierre[1], T. Auzeeri[1], T. Sauerwald[3], W.F.P. Andrews[1], H. Rihan[2], M.P. Fuller[2] and M.F. Stoelen[1,3]
[1]*School of Computing Electronics and Mathematics, University of Plymouth, United Kingdom; frederico.klein@plymouth.ac.uk*
[2]*School of Biological and Marine Sciences, University of Plymouth, PL4 8AA, United Kingdom*
[3]*Fieldwork Robotics Ltd., PL4 8AA, United Kingdom*

Abstract

This paper presents a proof-of-concept platform for demonstrating robotic harvesting of summer-varieties of cauliflower, and early tests performed under laboratory conditions. The platform is designed to be modular and has two dexterous robotic arms with variable-stiffness technology. The bi-manual configuration enables the separation of grasping and cutting behaviours into separate robot manipulators. By exploiting the passive compliance of the variable-stiffness arms, the system can operate with both grasping and cutting tool close to the ground. Multiple 3D vision cameras are used to track the cauliflowers in real-time, and to attempt to assess the maturity. Early experiments with the platform in the laboratory highlight the potential and challenges of the platform.

Keywords: robotics, selective harvesting, variable-stiffness actuators, cauliflowers

Introduction

The cost of rural workforce labour drives automation. A broad range of harvesting robot designs have been investigated over the last 30 years to enhance selective picking operations in agriculture. To name a few recent examples, in the EU-funded CROPS (Clever Robots for Crops) project, an autonomous robot platform for harvesting sweet-peppers in greenhouses was developed. The design concept implied a custom-design 9 degree-of-freedom (DOF) manipulator, 2 types of end-effectors, several RGB and depth cameras and a lighting rig. The overall performance of the robot system on simplified experiments was evaluated under typical crop environment conditions (Bac *et al.,* 2017). A proof-of-concept robot system aimed at picking apples in modern orchards was investigated by Davidson *et al.* (2016). The robotic design comprised a custom 6 DOF manipulator mounted on a travelling vehicle and a tendon-driven end-effector. A global vision system consisting of an RGB camera and a depth camera was used for detection and 3D location of the fruit. Laboratory testing was conducted, and the performance evaluation was discussed. Despite the wide range of robotic designs explored, versatile robotic solutions and extensive field experiments along with performance evaluations are required to make harvesting robots commercially available in agricultural environments.
Most harvesting operations for Brassica crops are still performed with manual labour, which drives up the costs for the farmer. Large combined harvesters for broccoli and cabbage are available, but the current technology has several drawbacks. First, it will harvest all crops in one part of the field at any time, mature or not, leading to extensive waste. No selective harvesters are currently commercially available. For broccoli, there is ongoing research on 3D vision (Kusumam *et al.*, 2017) and harvesting (Wageningen, 2018), and some industrial testing is underway (KMS projects, 2018). Cauliflower are more challenging due to the extensive foliage, and the difficulty in reaching the curd. The potential economic benefit to the producer is a considerable reduction in costs compared to manual labour, less hygiene issues and potentially less waste.

DOI 10.3920/978-90-8686-888-9_97, © Wageningen Academic Publishers 2019

In this paper, a proof-of-concept platform is presented that, when complete, will aim to reproduce all the steps in cauliflower harvesting, namely identifying maturity, grasping the curd, cutting it from its stem and placing on to a moving platform for packaging.

Materials and methods

Mobile test platform

The robotic platform is shown in Figure 1. It consists of an inverted U extruded aluminium frame with caster-wheels, fitted with two dexterous variable-stiffness robotic arms, and sensors. The arms are heavily modified versions of the open-source GummiArm (Stoelen *et al.*, 2016) (Figure 2). Each arm is made up of three major sections, the shoulder, the upper arm and the lower arm, where an end-effector for a specialised task can be attached. The movement of the different sections of the arm is realised by employing an agonist-antagonist tendon pulley system. That is, by emulating the pairs of opposing muscles found around joints in the human body. This enables control of movement and stiffness (through co-contraction of opposing actuators). As earlier versions of the GummiArm were used for demonstrations or selective harvesting of delicate and light fruits, the payload needed to be increased for selective harvesting of cauliflowers. This was done by implementing a bi-directional set-up of the pulley system for the most demanding joints in all three sections. End-effectors and the major sections can be connected effortlessly, due to a dovetail connection system, which makes it easy to remove them. One of its robotic arms is fitted with a cutting end-effector and another fitted with a picking end-effector. Additionally, it has 3 RGB-D cameras to perform the visual tasks necessary to control the state machine, a top mounted white LED array and a dual power system that can run on either batteries or mains.

Gripping end-effectors

For the task of gripping the cauliflower, there are two major gripper designs: the belt gripper and the three fingered gripper. The belt gripper consists of a bendy, but stable 3-ply PVC belt and its driving mechanism. The gripper will approach the cauliflower from above, push down any distant foliage, and enclose the cauliflower including its leaves. Gripping the leaves by tightening the belt provides an even pressure on the crop, reducing the risk of damage. The soft-fingered gripper consists of a modular base shape and three ends, driving the fingers with a rack and rail system. This makes it

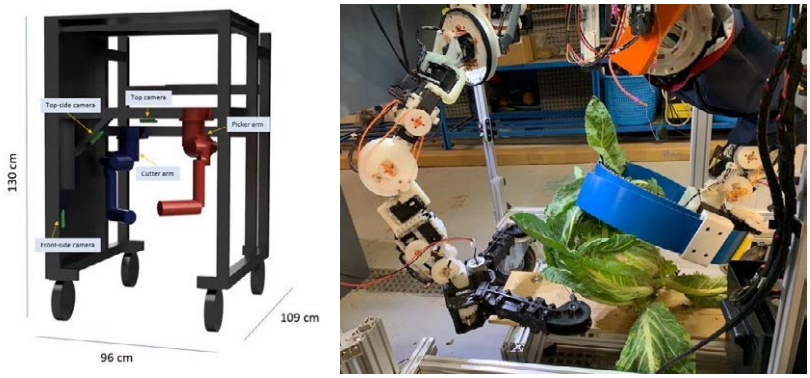


Figure 1. Left: Schematic render of the robotic cauliflower harvester color-coded for clarity. Red robotic arm is the cutting arm. Blue is picker arm. Green are the three D400 Intel Realsense cameras. Right: Image of cutting and gripping arm during lab testing.

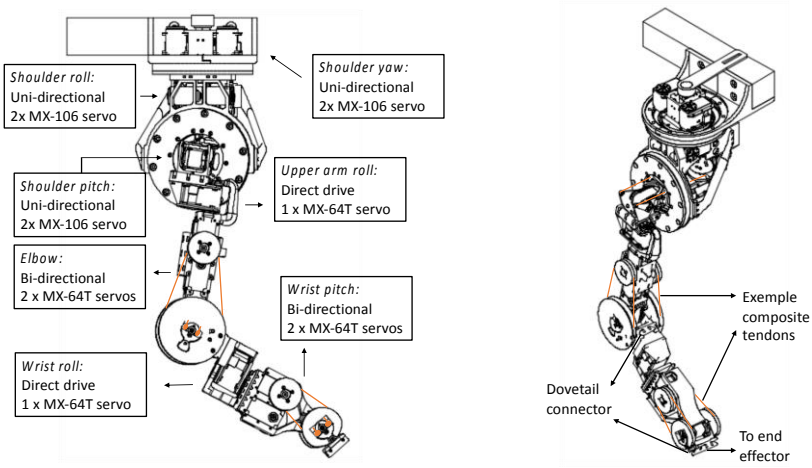


Figure 2. Two views of the modified version of the 7 jointed variable-stiffness GummiArm (Stoelen *et al.*, 2016) in development for cutting and picking cauliflowers.

possible to independently adjust the torque needed or the spread needed for the crop to harvest. The gripper can approach the curd from any angle and will also grip a crop through foliage.
The arms are controlled using a software framework based on ROS MoveIt! (Sucan, 2018) also used for other applications such as tomato and raspberry picking, previously and currently being explored by the authors' group (Stoelen *et al.*, 2018) (University of Plymouth, 2018). The soft end-effectors in an open position would be moved to the best estimate or the curd's centre in open loop, and by having it grasp around anything inside its grasping volume. Here a useful design feature of the passively compliant arm is exploited by setting the arm to be fully compliant/soft in every joint. This means it will no longer try to oppose any passive deflection of the end-effector caused by the belt-gripper tightening its grip around the rooted cauliflower head.

Computer vision

In the approach presented, the vision subsystem oversees two different tasks, namely: (1) detecting where a cauliflower is present in the 3D space; and (2) whether it is a cauliflower that should be picked or not. The cameras chosen were the Intel Realsense D400 series (Intel, 2018) due to the its low cost, high resolution and speed, allowing for a standard cross platform, cross application solution. To control lighting a set of ten 0.9 m long white LED strips, was evenly placed underneath the top of a fully covered platform.
A critical challenge in detecting cauliflowers is the partial or total occlusion of the curd by its leaves. The leaf growth pattern in Brassica's were exploited, namely that its leaves grow from the stem in a radial fashion and that the curd is on the top of this stem. To detect maturity curd flower separation was employed. A developing cauliflower has a very tight curd with no gaps and holes, and an overmatured one presents significant gaps in the curd. As such, the identification and measurement of the dark section in the curd of the vegetable can indicate ripeness, and the harvesting can be done. A diagram of curd and maturity detection can be seen in Figure 3.

Results

In order to test the vision algorithms as an independent part of the whole system, a small dataset of cauliflowers of different varieties was manually collected. This was done using both the D415 and the

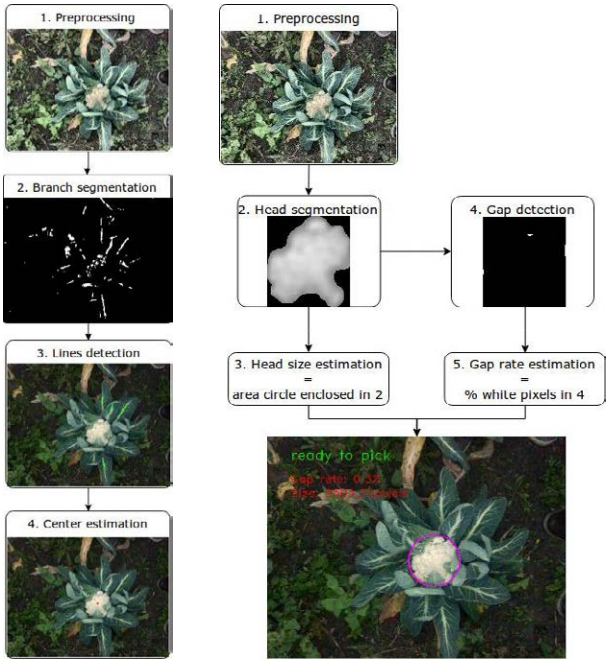


Figure 3. Overview of vision algorithms for detection of (A) curd, and (B) maturity.

D435 cameras and the images were collected with colour, infrared and depth channels. Each plant was imaged multiple times from different angles, at different times of day under different weather conditions, between March and May of 2018 on various occasions, mainly on a small university garden in the University of Plymouth campus and once in a commercial farm in Cornwall, UK and consisted mostly of around 5-10 plants of 3-4 different varieties, each plant being imaged from 6 to 50 times (median=20, average=24.16, σ=13.54), resulting in a total of 1,015 RGB images, most of which presented an exposed curd when mature. The results were qualitative, considering the great variability of the photos taken (different varieties, inconsistent lighting conditions, different cameras). The imaging was very different from the imaging obtained from the rig cameras. See Figure 4 and 5 for example results on localisation of centre and basic maturity estimation.

For the cutting tests it was important that the robotic arm could perform the whole movement, approximately 0.3 m at the base of the cauliflower. The arm also needed to produce sufficient correction forces to maintain the direction of the cut. An image of the cutting setup can be seen in Figure 6. A force-sensitive surface was used, with a one-axis load-cell. The resulting cut cauliflower curd can be seen in Figure 6C.

As can be seen in Figure 7, the cutting produces side forces around 5-10 N on the cauliflower. The cutting took about 15-20 s in this test due to the constrained mobility of the arm. Typically <10 s is possible with the current setup. A test was also done to make sure the arm could grasp an actual cauliflower. This test was meant to evaluate whether the arm's geometry had a suitable range to reach a cauliflower and position itself to grasp it (Figure 8). One of the main goals of the lifting test sequence was to ascertain if the picking end-effector can lift the payload of a cauliflower curd and a part of its foliage. See an example lifting sequence in Figure 9.

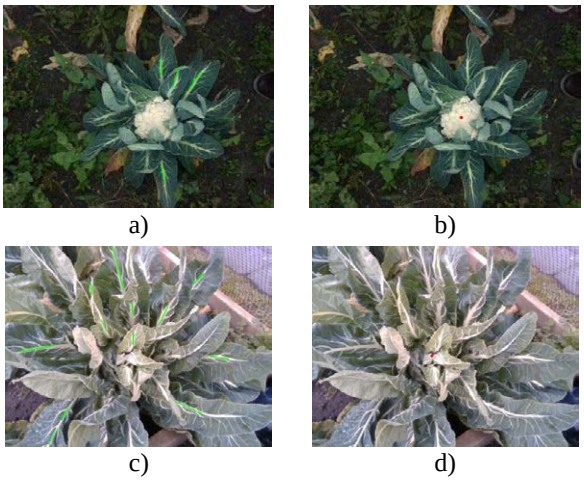


Figure 4. Example results from localisation of the centre of the cauliflower curd, (A) leaf vein detection (light green lines) for an open curd cauliflower, (B) centre (red dot) from extrapolated intersection of those lines, (C) a completely occluded curd with leaf veins detected, and (D) centre extrapolated from vein lines for occluded curd (in red).

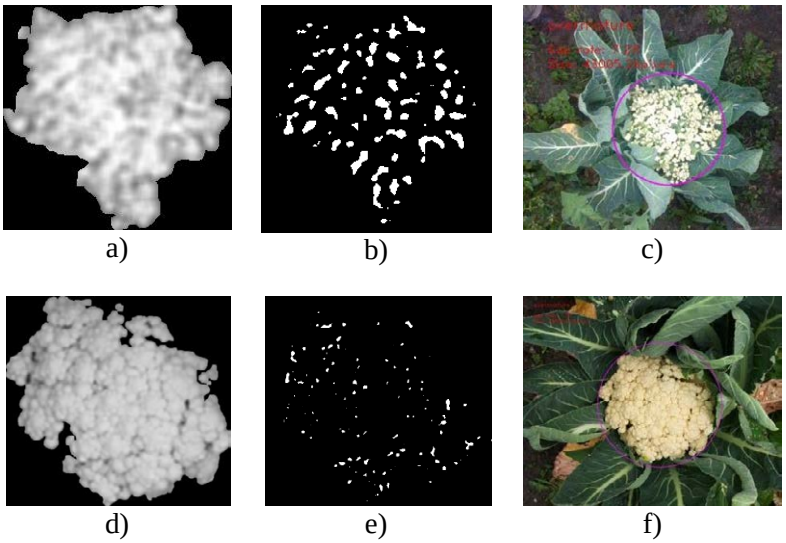


Figure 5. Example results from basic maturity estimation based on gap detection, (A) mask focused on the exposed curd, (B) gap detection, (C) visible gaps in overly matured curd, (D) masking on the exposed curd of younger cauliflower, (E) less prominence of white spots in this example, and (F) less visible gaps on this less overmatured curd.



a) b) c)

Figure 6. Experiment setup for teleoperated cutting tests with force-sensitive surface, (A) platform with stabilized cauliflower, (B) wire cutter and force-sensitive surface, and (C) the resulting cut in the cauliflower.

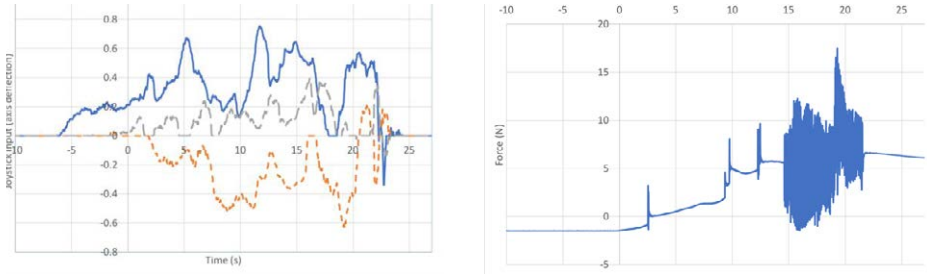


Figure 7. The profiles of joystick signals and forces during cutting test, (A) joystick axes signals, x: blue solid, y: orange dotted, z: grey dashed, and (B) force signal from the force-sensitive surface. Cutting from time 0.



a) b) c)

Figure 8. (A) initial position, (B) approach and (C) final grasping position reached.



a) b) c) d) e)

Figure 9. A lifting test sequence (A through E). The test was started with the cauliflower end-effector already closed around the curd, lift and angle it upwards.

Conclusions

This paper described a proof-of-concept platform for bi-manual robotic cauliflower harvesting. The platform can currently perform the tasks of: (1) identifying the approximate cauliflower maturity for visible curds; (2) localising and grasping the cauliflower curd; and (3) cutting it at a desired location along the stem. For these tasks, it relies on computer vision algorithms for identifying a cauliflower suitable for picking, and passively compliant robotic arms for dealing gracefully with possible collisions with the ground and plant material. These are the first steps towards a fully autonomous and selective cauliflower harvester, which should be run as an integrated system. The system architecture is designed to also be adaptable to other similar crops, by replacing the modular end-effectors and the vision/control algorithms.

Acknowledgements

This work presented here was part of the Automated Brassica harvesting in Cornwall (ABC project), a sub-project of Agritech Cornwall, part-funded by the European Regional Development Fund (*ERDF*), and with match-funding from Cornwall Council. The work was also partially supported by CNPq Brazil (scholarship 232590/2014-1). Thanks to the Cornish industrial partners, including Riviera Produce (Hayle, UK), Teagle Machinery Ltd (Truro, UK), Robotriks (Par, UK) and CNC Design Ltd (St Columb Major, UK).

References

Bac, C.W., Hemming, J., van Tuijl, B.A.J., Barth, R., Wais, E. and van Henten, E.J., 2017. Performance evaluation of a harvesting robot for sweet pepper. Journal of Field Robotics, 34(6), pp. 1123-1139.

Davidson, J.R., Silwal, A., Hohimer, C.J., Karkee, M., Mo, C. and Zhang, Q., 2016, October. Proof-of-concept of a robotic apple harvester. In International Conference on Intelligent Robots and Systems (IROS), IEEE/RSJ (pp. 634-639). IEEE.

Intel, 2018. Overview of the Intel® RealSense™ Depth Camera. https://software.intel.com/en-us/realsense/d400. (accessed December 2018).

Kamilaris, A. and Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, pp. 70-90.

KMS projects, 2018. 'Breakthrough' automated broccoli harvester turns to FANUC for a helping hand. https://www.fanuc.eu/uk/en/who-we-are/news/uk-kms-automated-harvester-08-2018 (accessed December 2018.

Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems. Neural Information Processing Systems Foundation, pp. 1097-1105.

Kusumam, K., Krajník, T., Pearson, S., Duckett, T. and Cielniak, G., 2017. 3D-vision based detection, localization, and sizing of broccoli heads in the field. Journal of Field Robotics, 34(8), pp. 1505-1518.

ROS. n.d. Flexbe – ROS Wiki. http://wiki.ros.org/flexbe (accessed December 2018).

Stoelen, M.F. *et al.,* 2018. China Robot Harvest++. http://www.agritechinchina.org/LG017 (accessed December 2018).

Stoelen, M.F., Bonsignorio, F., Cangelosi, A., 2016. Co-exploring Actuator Antagonism and Bio-inspired Control in a Printable Robot Arm, In: Tuci, E., Giagkos, A., Wilson, M., Hallam, J. (Eds.), From Animals to Animats 14. Basel, Switzerland: Springer International Publishing, pp. 244-255.

Sucan, I.A. and Chitta, S., 2018. MoveIt!. http://moveit.ros.org (accessed December 2018.

University of Plymouth, 2018. University spinout signs agreement to field test harvesting robots. https://www.plymouth.ac.uk/news/university-spinout-signs-agreement-to-field-test-harvesting-robots (accessed December 2018).

Wageningen University, 2018. Robotic Broccoli Harvester. https://www.wur.nl/en/project/Robotic-Broccoli-Harvester.htm (accessed December 2018).

Deep learning-based image segmentation for grape bunch detection

R. Marani[1*], A. Milella[1], A. Petitti[1] and G. Reina[2]
[1]*Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing, National Research Council, via G. Amendola 122D/O, 70126 Bari, Italy; roberto.marani@stiima.cnr.it*
[2]*University of Salento, Department of Engineering for Innovation, Via Arnesano, 73100 Lecce, Italy*

Abstract

This paper presents a method for automatic fruit detection in vineyards through the inspection of colour images obtained by a low cost RGB-D sensor placed on-board an agricultural vehicle. Image segmentation is obtained by using a pre-trained convolutional neural network, which receives input data, as sub-patches of known size, and performs the classification in a few classes of interest. Output scores are then used to create the probability maps for each class and, thus, pixel-by-pixel segmentation of the grape clusters. Field experiments prove the ability of the proposed processing to successfully segment grape clusters, with accuracy of 87.5%, despite the poor quality of the input images.

Keywords: deep learning, image processing, grape bunch segmentation, RGB-D sensing

Introduction

Accurate knowledge of the crop characteristics at sub-field level is crucial for precision agriculture applications. Conventional approaches based on a few random samples for visual or destructive inspection by human operators are time consuming, subjective and prone to human errors. On the other hand, high-throughput phenotyping platforms (HTPPs) mainly using vision and LIDAR sensors, have attracted much interest in the last decade (Reina *et al.*, 2016). Recently, portable consumer-grade RGB-D sensors, such as the Microsoft Kinect, have been proposed as an alternative, cost-effective solution to recover 3D models of plants. However, it can effectively work only under laboratory conditions or in controlled environments, like greenhouses (Narvaez *et al.*, 2017; Paulus *et al.*, 2014; Chéné *et al.*, 2012). In the last years, different sensors, based on active stereovision assisted by light projectors, have been proposed. Among them, the Intel RealSense R200 has attracted increasing interest since it can work in outdoor under different conditions of ambient light, producing RGB and infrared images, together with 3D depth data. These features make it suitable for its use on-board an agricultural vehicle.

As known, visual data acquired with any setup has to be processed automatically for extracting informative contents. As an example, accurate and reliable fruit detection systems are the basis for autonomous robotic platforms devoted to precision agriculture. Recently, standard machine learning approaches have been replaced by deep learning-based ones (LeCun *et al.*, 2015). In this case, artificial networks are made of consecutive layers of filters which generate a hierarchical representations of the phenomena (or of the corresponding data), by producing maps of lower size but of increased meaning.

With reference to the agricultural context, deep learning has become the most promising technique in applications targeted to fruit detection and counting (Sa *et al.*, 2016; Bargoti and Underwood, 2017; Kamilaris and Prenafeta-Boldú, 2018). In this scenario, this paper demonstrates the application of deep learning for the detection of grape bunches in images captured in an actual vineyard by the low-cost consumer-grade Intel RealSense R200. Specifically, a pre-trained convolutional neural network (CNN), namely the AlexNet (Krizhevsky *et al.*, 2012), is adapted to obtain pixel-by-pixel segmentation, with final results of good segmentation accuracy, despite the low quality of the available data.

DOI 10.3920/978-90-8686-888-9_98, © Wageningen Academic Publishers 2019

Materials and methods

The proposed methodology aims to carry out detection and segmentation of images acquired by a consumer grade hardware placed on a mobile platform (e.g. an agricultural vehicle) which operates in a commercial vineyard, named Räuschling, located in Switzerland. The mobile robot moves within rows with an average speed of 1.5 m/s; row spacing is in the range between 1.5 and 2 m. The Intel RealSense R200 (Santa Clara, CA, USA) used in the experiments frames the grape row laterally and returns RGB images (8-bit-depth for each channel) and depth maps (16-bit-depth) with resolution of 640×480 pixels at a frame rate of 5 Hz. An example of a standard RGB image is in Figure 1.
As a first step, only the RGB images have been used in this study. These input images suffer from motion blurring and compression artefacts, and their low resolution, together with the distance between the vehicle and the target, makes their analysis of high complexity. In addition, grapes can have similar colours to the leaves of the same plants and to the background, often made of plants of other rows.
Following the standard approach of image classification, input images are randomly split into two datasets: training and test dataset. The first one, made of 60 images, creates a set of known examples, used to setup the methodology. The second set is made of 25 images used to evaluate the methodology. All images are first labelled manually in 5 different classes of interest: grapevine bunches (*Bunch class*); vineyard grape poles (*Pole class*); grapevine trunks, cordons and canes (*Wood class*); leaves (*Leaves class*); all other objects than those of the previous classes (*Background class*). This labelling produces the sample results of Figure 2.
The task of grape cluster detection and segmentation is carried out by dividing the input image into sub-patches of known size. Each patch encloses pixels belonging to a few classes. Depending on the number of labelled pixels in the patch, the whole area is labelled with a single class. Actually, this selection is performed by a weighted summation aimed at emphasizing minor classes, such as wood, pole and bunch. For instance, if the summation of bunch pixels in the patch is higher than the 20% of the whole number of pixels, the patch is labelled as bunch, regardless of the remaining 80% of the patch. The results of this approach are highlighted in Figure 3, where sample patches of the five classes of interest are shown.

Training

This input data is processed by a pre-trained convolutional neural network (CNN), namely the well-known AlexNet. This architecture has been preferred among a set of possibilities, such as



Figure 1. A RGB image acquired by the mobile vehicle with the proposed hardware.

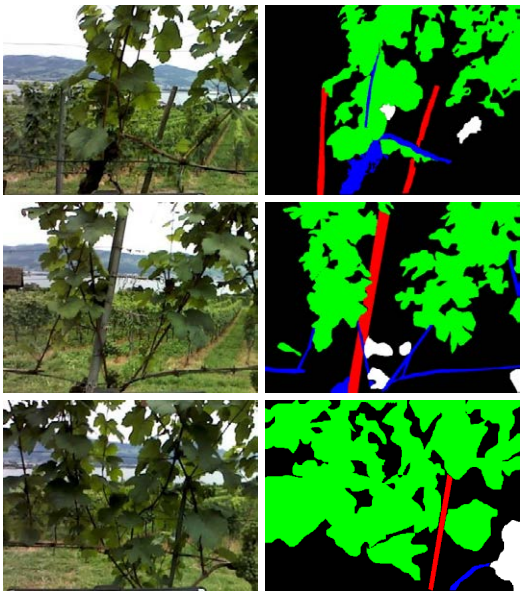


Figure 2. Manual labelling of input images. The left column reports three examples of input images. The right column shows the corresponding segments. Black, green, blue, red and white refer to background, leaves, wood, pole and bunch classes, respectively.

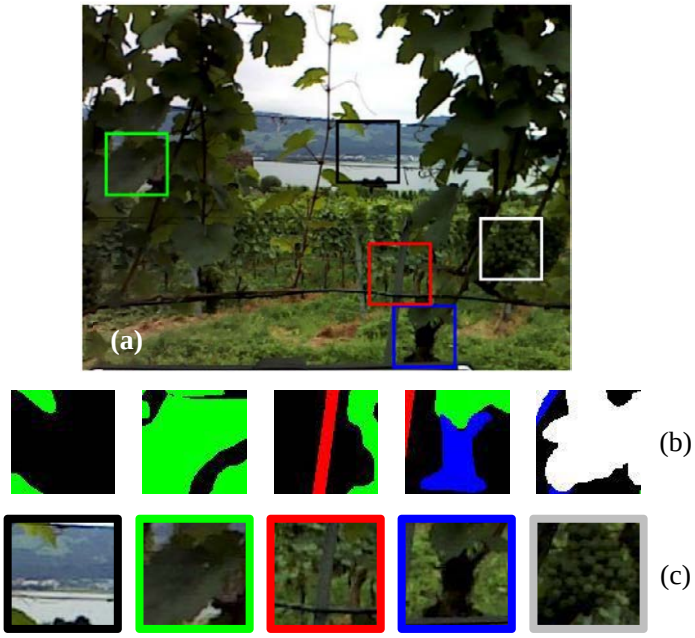


Figure 3. Examples of patch extraction. (A) Input image with patches of size 80×80; (B) corresponding ground truth; (C) patch magnification (the colours of the patch boundaries refer to a specific class, in agreement with the legend of Figure 2).

VGG16, VGG19 and GoogleNet, since it has been demonstrated that it returns good results with lower requirements in terms of computational time (Milella *et al.*, 2019).
This CNN has been proposed to label input RGB images among 1000 different classes of objects. It is made of a sequence of 5 convolutional layers and 3 fully-connected layers, whose weights have been determined during a training phase performed on a dataset of 1.2 million images.
Convolutional layers are arranged to perform a cascade of simple operations, such as:

- for all layers: to filter the input representations of the data, applying kernels of different size. Each output activates its corresponding neuron through a rectified linear unit (ReLU);
- for the first and second layers: to normalize the local response of each output of ReLUs on the adjacent 5 channels, thus improving generalization capabilities;
- for the first, second and fifth layers: to group outputs of neighbouring neurons in order to create higher levels of abstraction of the input representation (max-pooling).

The first two fully-connected layers, which follow the convolutional ones, are made of 4,096 neurons each, whereas the third is made of 1000 neurons. In addition, the first two fully-connected layers are followed by dropout layers, which randomly set input elements to zero with a probability of 0.5, to reduce overfitting. The deepest layer of the net architecture applies a softmax function to the input. It returns the classification score of the specific input image over the 1000 labels. This architecture results in more than 60 million parameters to be defined during training, which requires significant datasets of labelled images to be completed. This constraint can be hardly overcome with actual datasets devoted to the recognition of specific classes, e.g. grape bunches.
Deep learning networks, including the Alexnet, show common characteristics: although they are devoted to solve specific tasks, the first convolutional layers of the net perform standard processing. As a consequence, there is no need to train the whole architecture without initialization, but it is possible to adapt its parameters to the new task. Following the paradigm of transfer learning (Yosinski *et al.*, 2014), the architecture of the AlexNet is modified to classify inputs on the five reference classes, i.e. reducing the last fully-connected layer to 5 neurons. Then, its parameters are tuned in a further training phase, performed by using a smaller, but much more focused dataset. It is obtained by extracting sets of 80×80 patches from the 60 images of the training set. Patches are then upscaled to 227×227 pixels to match the size of the input layer.

Segmentation
After training, the final goal of image segmentation can be reached by gathering classification scores into probability images. As a first step of the processing, a moving window, of square size of 80×80, extracts the patches. Since the window frames the image with a unit step size in both directions, every single pixel will fall into many patches (in this case in 79×79=6,241 patches). The AlexNet can thus classify the patches (640×480=307,200 patches), producing specific labels and the corresponding classification scores. These are vectors of five entries, each one describing the probability (in the range [0,1]) that the patch belongs to a specific class. As remarked previously, every pixel of the input image has 6,241 vectors of classification scores of 5 entries. By averaging the entries of the classification scores, which lie on the same pixel and rearranging the results in space, five probability maps can be created from the single input colour image. Then, the comparison of the probability maps produces the final segmented image, whose pixels are labelled with the class having the highest score among the five maps. Consequently, the segmented image includes information about the exact area where the grape clusters are located.

Results

As stated above, the input dataset of 84 images acquired by the Intel RealSense R200 sensor is divided into training and test sets. The first set is used to train the weights of the network. This phase has been performed on a single GPU (nVidia Quadro K5200). The net training is defined in terms of a

stochastic gradient descent with momentum (SGDM) optimization with momentum equal to 0.9, and different learning rates for the first layers (10^{-4}) and the last fully-connected (2×10^{-3}). Thus the training is mostly focused on the last layers, while the previous ones are less affected. These parameters allow the convergence of the learning phase in 17 epochs, which are evaluated in about 4 hours (241 minutes) (Milella *et al.*, 2019).

Test images, which have never been proposed to the classifier, are finally used to test the performance of the proposed processing. As shown in the previous section, each test image is divided into 307,200 sub-patches, which are individually processed to find the probability score of each class. Then, output redundancy is filtered by averaging results belonging to the same pixels. The whole processing on a single image takes about 11 minutes, producing the outcomes of Figure 4. It reveals that probability values for background, leaves, wood and bunch classes reaches unity, thus suggesting that the corresponding regions can be labelled with high reliability. In contrast, although the input image shows a thin pole, the corresponding region of the probability map of the pole class has values of about 0.64. In the same region, the probability map of the background reaches about 0.34, thus suggesting that the area includes also further objects, different from the pole. Anyway, the final labelling will result in the pole class, which has the highest classification score.

The final segmentation of the input images can be obtained by comparing the probability maps of the five classes and by selecting the one with the highest score. Labelled pixels are then grouped into

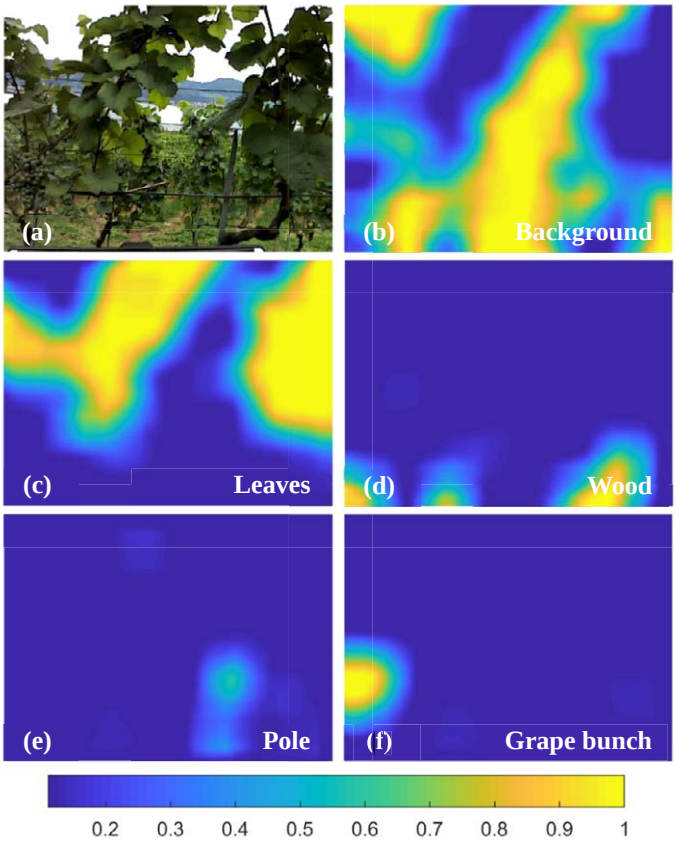


Figure 4. (A) Input for the processing and corresponding probability maps for (B) background, (C) leaves, (D) wood, (E) pole, and (F) grape bunch class.

connected components having significant extension, i.e. with area higher than a threshold (set to 9). The increase of the value of the threshold area can reduce the number of isolated groups of pixels. This processing returns the results of Figure 5, which shows the input data, the corresponding ground truth (manually labelled), and the segmented images. The inspection of Figure 5 denotes that the segments of the bunch, wood and pole classes are overestimated, i.e. their extension is higher than the actual one in the image. This behaviour is due to the preliminary selection of the labels of the patches of the training dataset. As stated previously, each training patch is labelled with a single class, as a result of a weighted summation of the labels of its pixels. It is aimed at giving more emphasis to pixels of the bunch, wood and pole classes, whose size within the patches may be much smaller than the one of leaves and background objects. In turn, the classification scores of these classes are high even if the corresponding patches have only few pixels of that specific segments. However, the grape bunch detection, which is the final aim of the processing, is not altered.

Quantitative metrics are finally proposed to test the net performance in bunch detection. Specifically, detections are correct (true positives) if the intersection of detected bunch regions with the corresponding of the ground truth is not void. The analysis of the 24 test images, which enclose 56 bunches, has produced 49 true positives, corresponding to an overall accuracy of 87.5%. The remaining 7 bunches are not detected (false negatives), whereas 1 bunch is wrongly segmented (false positive). Future works will consider the projection of the detected segments on the depth maps produced by the RGB-D sensor, thus enabling the exact volume estimation of grape clusters.

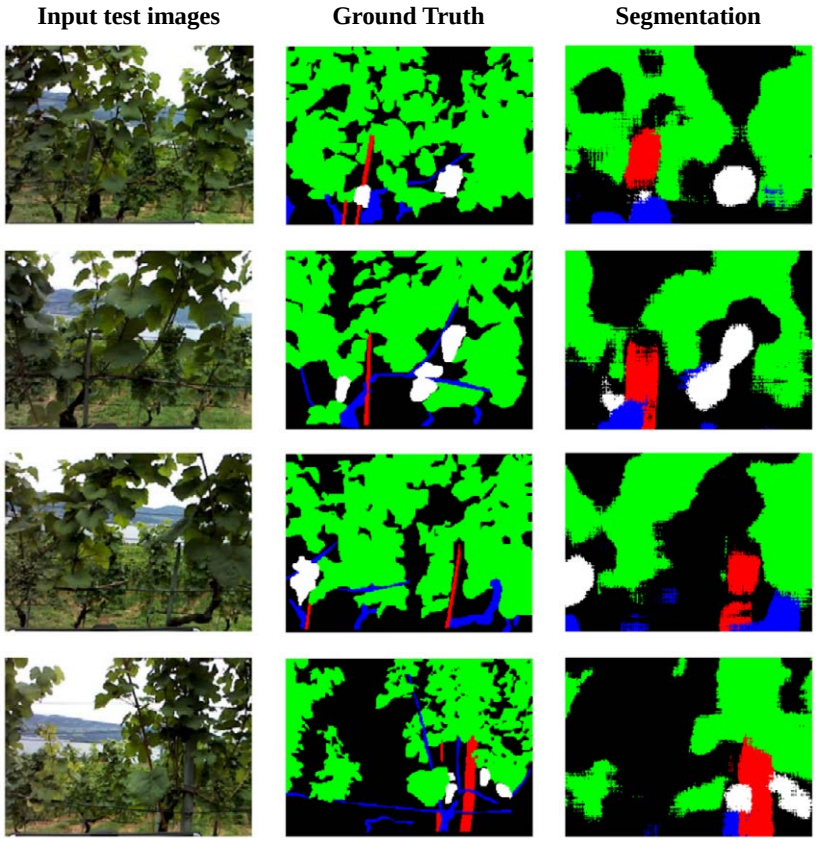


Figure 5. Results of the segmentation process. First column: input images from the test set; second column: corresponding ground truth (manually labelled); third column: results of the segmentation.

Conclusions

This paper has presented a methodology for segmenting grape bunches in images captured from a consumer grade RGB-D sensor, placed on-board an agricultural vehicle. This goal is achieved by using a pre-trained CNN, which receives input data, computes classification scores and rearranges results to create the probability maps for each class, including the one of grape bunches. In-field experiments have proven an accuracy of 87.5% in grape cluster segmentation, despite the poor quality of the input images. Grape clusters can be exactly located in an automatic and non-invasive way, thus overcoming time and cost problems related to traditional manual measurements.

Acknowledgements

The financial support is acknowledged: Simultaneous Safety and Surveying for Collaborative Agricultural Vehicles (S3-732 CAV), ERA-NET ICT-AGRI2 (Grant No. 29839), Autonomous DEcision Making in very long traverses (ADE), H2020 (Grant No. 821988), and Electronic Shopping & Home delivery of Edible goods with Low environmental Footprint (E-SHELF), POR Puglia FESR-FSE 2014-2020(Id. OSW3NO1). The authors are grateful to M. Attolico, G. Bono, S. Rilling, P. Frölich and M. Nielsen for the support in performing the experiments and gathering the data.

References

Bargoti, S., Underwood, J., 2017. Deep fruit detection in orchards. In: Arai, F., Arrichiello, F., Burschka, D., Castellanos, J., Hauser, K., Isler, V., *et al.* (Eds.) IEEE Int. Conference on. Robotics and Automation (ICRA), pp. 3626-3633.

Chéné, Y., Rousseau, D., Lucidarme, P., Bertheloot, J., Caer, V., Morel, P., *et al.*, 2012. On the use of depth camera for 3D phenotyping of entire plants. Computers and Electronics in Agriculture 82 122-127.

Kamilaris, A., Prenafeta-Boldú, F., 2018. Deep learning in agriculture: A survey. Computers and Electronics in Agriculture 147 70-90.

Krizhevsky, A., Sutskever, I., Hinton, G., 2012. Imagenet classification with deep convolutional neural networks. In: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q., (Eds.) 25th Int. Conference on Neural Information Processing Systems 1, Lake Tahoe, NV, USA: Curran Associates Inc., pp. 1097-1105.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 436- 444.

Milella, A., Marani, R., Petitti, A., Reina, G., 2019. In-field high throughput grapevine phenotyping with a consumer-grade depth camera. Computers and Electronics in Agriculture 156 293-306.

Narvaez, F., Reina, G., Torres-Torriti, M., Kantor, G., Cheein, F., 2017. A survey of ranging and imaging techniques for precision agriculture phenotyping. IEEE/ASME Transactions on Mechatronics 22 2428-2439.

Paulus, S., Behmann, J., Mahlein, A., Plümer, L., Kuhlmann, H., 2014. Low-cost 3D systems: Suitable tools for plant phenotyping. Sensors 14 3001-3018.

Reina, G., Milella, A., Rouvere, R., Nielsen, M., Worst, R., Blas, M., 2016. Ambient awareness for agricultural robotic vehicles. Biosystems Engineering 146 114-132.

Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., McCool, C., 2016. Deep-fruits: A fruit detection system using deep neural networks. Sensors 16 1222.

Yosinski, J., Clune, J., Bengio, Y., Lipson, H., 2014. How transferable are features in deep neural networks? In: Eds. Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N.D., Weinberger, K.Q., 27th Int. Conference on Neural Information Processing Systems 2, Cambridge, MA, USA: MIT Press. pp. 3320-3328.

Soft manipulator robot for selective tomato harvesting

A. Mohamed[1*], J. Shaw-Sutton[1], B.M. Green[1], W. Andrews[1], E.J. Rolley-Parnell[1], Y. Zhou[2], P. Zhou[2], X. Mao[3], M.P. Fuller[4] and M.F. Stoelen[1,5]
[1]*University of Plymouth, School of Computing, Electronics and Mathematics, Drake Circus, PL4 8AA Plymouth, United Kingdom; abdulla.mohammad@plymouth.ac.uk*
[2]*Shanghai Jiaotong University (SJTU), School of Agriculture and Biology, Dongchuan Road 800, 200240 Shanghai, China P.R.*
[3]*Shanghai Sunqiao Modern Agricultural United Development Company, 185 Mianbei Road, The China (Shanghai) Pilot Free Trade Zone, 201210 Shanghai, China P.R.*
[4]*University of Plymouth, School of Biological and Marine Sciences, Drake Circus, PL4 8AA Plymouth, United Kingdom*
[5]*Fieldwork Robotics Ltd, Drake Circus, PL4 8AA Plymouth, United Kingdom*

Abstract

The harvesting of soft fruits and vegetables is a labour-intensive process, often representing more than 50% of the total costs for the producer. In this paper, a harvesting robot is proposed for tomato picking. The robotic solution is developed to address the needs of tomato producers in the Shanghai region in China, one that faces population growth and therefore a higher demand on food supply. The robotic system presented consists of a variable-stiffness manipulator arm, a soft robot gripper, and different types of sensors that are used to identify and locate in 3D and pick the tomatoes. The implemented variable compliance enables the robot manipulator to work in a semi-structured environment without damage to itself, the crop or the surrounding. The hardware and software of the robot is described in detail. Early results from the first testing of a proof-of-concept on fresh tomatoes placed on artificial stems in Shanghai are presented, as well as picking UK tomato varieties in greenhouse conditions.

Keywords: selective harvesting, soft robotics, tomatoes, variable-stiffness actuators

Introduction

The increasing demand on farmers motivates research on agriculture robots for different applications such as seeding, weed control and harvesting. (Reddy *et al.*, 2016). Many robots have been developed to work in harvesting for different fruits and vegetable. For example, Hemming *et al.*, (2014) designed a robot for sweet-pepper, van Henten *et al.*, (2002) proposed a robot for harvesting cucumbers and Leu *et al.*, (2017) designed a robot for the harvesting of green asparagus.
Feng *et al.*, (2015), proposed a robot for picking tomato within a greenhouse that uses an installed rail to travel between the tomato vines. A 4 degree-of-freedom (DOF) arm was used with a perception system that consisted of a CCD camera attached to the linear laser projector. The detection algorithm used to detect the tomato applied a threshold to HSI colour space. The same group (Feng *et al.*, 2015) proposed a design for tomato grasping (Wang *et al.*, 2016). A high success rate was reported (up to 83.9%), but experimental conditions and procedures were not detailed. Zhao *et al.*, (2016), proposed another robot for tomato harvesting using dual-arms consisting of 3 degrees of freedom for each arm. Each arm had a different end-effector, one to hold the tomato using a pneumatic suction cup and the other equipped with a cutter to cut the stem of the tomato. A stereo vision system was integrated with the robot to locate the tomato positions. The robot proposed in Zhao *et al.*, (2016) required a human worker to identify the ripe tomato.
Tomatoes are the world's largest vegetable category (Eurofresh, 2016), with a total production of around 130 Mt. China is among the 5 largest tomato producers. The project described in this paper

DOI 10.3920/978-90-8686-888-9_99, © Wageningen Academic Publishers 2019

is a collaboration between UK- and China-based research groups to design and develop a soft robot manipulator for tomato harvesting. The project aimed to explore a soft end-effector for gentle picking and separation of tomatoes without a cutting mechanism, and a variable-stiffness robot arm for operating in this complex horticultural environment. The proposed system was intended to have a fully autonomous detection, planning and picking loop, and used point clouds and a watershed algorithm for detection. The first proof-of-concept robot was tested briefly in greenhouse environments in both China and the UK, awaiting longer-term quantitative experiments.

Methods and materials

Robot arm

The GummiArm is a robotic arm with seven DOFs inspired by the mechanisms of human and animal sensorimotor systems (Stoelen *et al.*, 2016). The GummiArm can be considered soft robotics owing to the joint mechanisms that use agonist-antagonist actuators connected to the joints via flexible tendons. These actuators are able to control the stiffness of the joints during operation (Figure 1). Agonist-antagonist actuators can be described as a pair of linear actuators where the agonist actuator causes movement in a given direction, and the antagonist actuator opposes the torque induced by the agonist actuator to increase stiffness.

The GummiArm was originally developed at the University of Plymoth, UK, for research groups working on robots operating in complex and unpredictable environments. It is largely 3D printable to enable quick iterations on hardware, and the base version is available open source (GummiArmCE, 2018). It utilises Dynamixel (Robotis, Seoul, Korea) digital servos. The environment during harvesting can be unpredictable, and detecting all hazards in the environment using sensors can be difficult (e.g. in most fruit plants, a thin string is used to support the plants). The GummiArm is therefore an interesting option to employ in this type of environment.

Gripper

The proposed gripper is mounted as an end-effector on the robot arm. It is a soft gripper with three fingers that deform to the shape of the object upon grasping. The body of the gripper is made of rigid Polylactic Acid (PLA) whilst the fingers are made from an elastic deformable polymer coated with high friction silicone. This decreases the likelihood of damage to the tomato. A Red, Green, Blue (RGB) camera integrated in the gripper is used for visual servoing and to provide data about a cluster of tomatoes. An Infra-Red (IR) sensor is used to measure the distance of the tomato to the gripper.

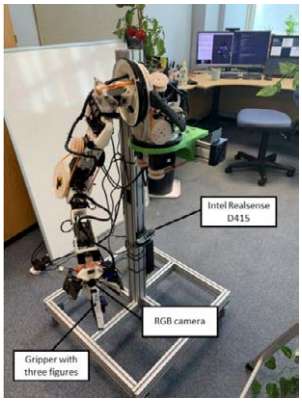


Figure 1. The robot arm, gripper, and sensors.

Camera systems

There are two cameras integrated with the robotic manipulator. The first sensor is a Realsense D415 (Intel, CA, USA) color/depth camera, see Figure 1. This camera is used to detect and 3D locate the tomato in front of the robot. The second vision sensor is a USB camera integrated with the gripper (Figure 1).

Software architecture

The software of the robot is based on Robotic Operation System (ROS) (Quigley *et al.*, 2009). ROS is an open source system comprising a collection of libraries and software that are used to accelerate prototyping robots and share the experience between developers.

The state machine for the system was based on ROS and specifically tuned to tomato harvesting. In general, the state machine is a cycle of different processes as shown in Figure 2. The cycle starts by: (1) locating the 3D position of the fruit; (2) command the arm to move closer to the target; (3) a visual servoing process for precise positioning relative to target fruit; (4) close the gripper and pull the fruits to separate it from the plant; and (5) place the fruit in a tray. The following subsections describe each stage in detail.

Detection and 3D location of tomatoes

The first process in the state machine is the 3D detection of the tomato relative to the arm base. This process computes the position of the tomato using RGB image and depth data provided by Intel Realsense D415. Figure 3 shows the pipeline of the depth detection algorithm. The process starts by detecting the tomato in RBG image by applying a linear regression model to detect the tomato. The linear regression model applies to each pixel to compute the probability of said pixel corresponding to a tomato or not $P_{RGB}\varepsilon(0,1)$.

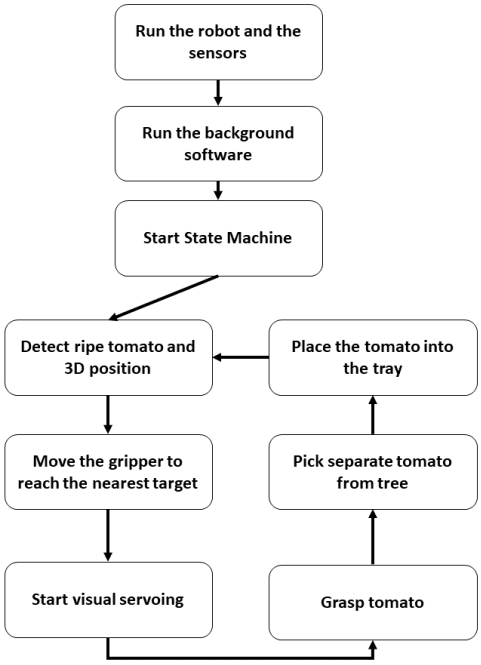


Figure 2. State machine of tomato harvesting.

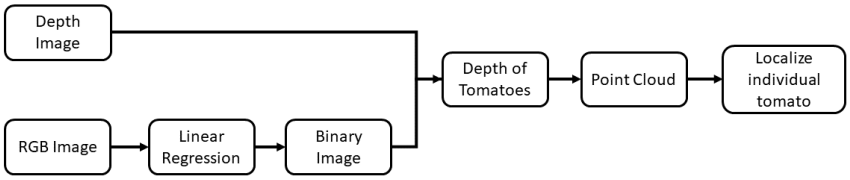


Figure 3. Detection and 3D localisation pipeline used.

$$P_{RGB} = X \times w + B \quad (1)$$

The output of the linear regression process is a binary image used to mask the depth image using the bitwise process. The bitwise process is used to irrelevant content from the depth image. The depth map is converted to point cloud format which is then used to compute the centroid of the detected tomato.

Visual servoing

The visual servoing is used in order to approach the tomato with greater precision than can be achieved from a more remote camera. The visual servoing node was designed to grip the tomatoes which are clustered so tightly that the 3D detection could not manage to detect the individual targets. The visual servoing uses the camera in the gripper where the process starts by detecting the tomatoes and then determines the tomato nearest to the centre of the image. The visual servoing detection algorithm is designed to identify each tomato separately using a watershed algorithm (Bleau and Leon, 2000). Figure 4 shows the input image from the gripper camera and the output image of the algorithm where the tomatoes detected are outlined. The output of the detection algorithm provides a 2D pixel co-ordinate (x,y), whereas the end effector manipulates in 3D, $P_e(X,Y,Z)$. An exponential function was used to compute the movement in the axis of the image frame by computing the Euclidean distance (q) which is biased for greater movement with smaller distances.

$$z = \exp^{(q \times \lambda)} \times \beta \quad (2)$$

where β is the scaling constant which limits the range of the output to meet the image size and λ is the constant which describes the shape of the output.

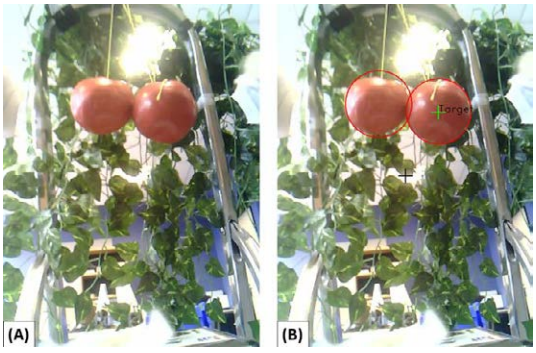


Figure 4. Visual servoing detection of individual tomato in the cluster. (A) camera image and (B) the output of the detection algorithm where the black cross is the centre of the image and the green cross indicates the target that is closest to the centre.

The output of the detection process is the difference between the position of the target and the centre of the camera image. A proportional controller is used to compute the error of the end effector position relative to the tomato position. This error added to the current position of the end-effector then feeds into the inverse kinematic model to compute the new joint angles that feed to the joint controller to move the arm to the new position. The K_p was selected based on trial and error.

Motion planning

As the robot manipulator is integrated with ROS the position of the end-effector is computed and controlled using the MoveIt! motion planning framework. MoveIt! provides an inverse kinematics planner and motion controller for the GummiArm by reading the Universal Robotic Description Format (URDF) file of the manipulator. The motion planning and control are used in different stages within the state machine, including positioning the end-effector in front of the detected tomato and placing the tomato in the tray after picking.

Experiment setup

The system has been tested in greenhouses in the UK and China. Figure 6 shows the setup. The robot is placed in front of a tomato plant and initialised to pick the tomato with the shortest absolute

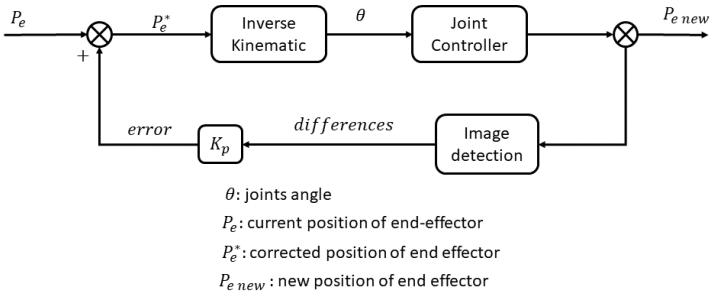


Figure 5. The visual servo controller used for the last part of the tomato-picking process.

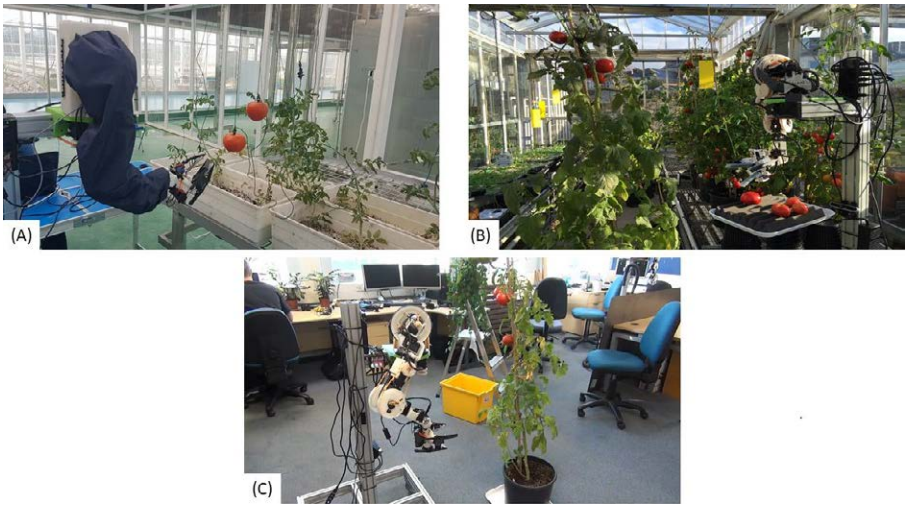


Figure 6. Experiment setup (A) China, (B) UK, (C) Lab.

distance to the robot base. This testing has been done to evaluate the robot performance in terms of picking success rate and the speed of picking by measuring the entire circle of the state machine. The time of each process in the state machine has been measured separately during the experiment in order to analyse the durations and identify any bottlenecks of the cycle. However, the first test in China has a more simplified setup due to the lack of ripe tomato plants at the time the test was performed (the setup was designed to simulate an authentic scenario).

The principal target for the robotic system was the Chinese market, and the robot was therefore primarily designed for larger Chinese tomatoes (approx. 75 mm diameter). A range of UK varieties were also tested, and it was found that the solution can be adapted to regular UK varieties with small changes to the size of the gripper fingers. Further experimental work is needed to truly assess the quantitative performance and robustness of the system under commercial growing conditions, in terms of detection, maturity classification, and picking performance. This work is underway winter 2020 in Shanghai, China, with an improved version of the system.

Result and discussion

The main advantage of using the robot in harvesting is accelerating the picking process and reduction of labour cost where the robot may work continuously. Therefore, the robot was tested and evaluated based on the speed of picking and the success rate of picking the tomatoes. Figure 7 shows the time of each step in the harvesting process starting with the 3D detection and ending with placing the tomato in to the tray. From Figure 7, the timing leads to understanding bottlenecks in the state machine process which is currently in the visual servoing that take the longest time to finish. The main reason is the complexity of placing the gripper on the tomato in the best position. However, in the next iteration of the system, a depth camera will be integrated into the end effector in order to generate improved approach paths and more accurately position the gripper with regard to the target tomato.

The complete picking process for picking one tomato was 30±5 s. The 5 s margin error is due to the varying position of the tomato relative to the arm base. However, the time is comparable to the work of Feng *et al.*, (2015), at 25 seconds. More work is needed to define benchmarking standards to make such comparisons more relevant.

In future work, the visual servoing process will integrate with a depth camera to improve both the speed and confidence of motion and determine the best pose of the gripper around the tomato. The state machine of the robot will be rewritten in order to improve the speed and the reliability of the picking process. More extensive benchmarking will also be performed on production crop.

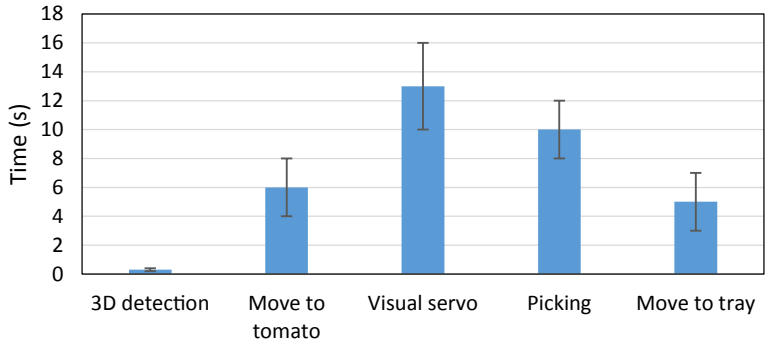


Figure 7. Time for each process in the state machine.

Conclusions

This paper presents a soft manipulator robot for selective tomato harvesting, using human inspired motor mechanisms. Passive compliance was introduced in the picking process to help increase robustness to unforeseen impacts during autonomous operation. The robot integrates various types of sensor in order to detect, locate and pick ripe tomatoes. The software architecture was based on ROS, while using the MoveIt! framework for motion planning. A custom soft robot hand was designed to enable one-handed picking without a cutting mechanism. The robot has been through first tests in greenhouses in both China and the UK. The full cycle time for picking a tomato was 30±5 s, but more extensive experiments are needed to assess performance and robustness on commercial crops.

Acknowledgement

The work presented here was funded by two grants from the Agri-Tech in China: Newton Network+ (ATCNN). The work also received support in-kind from Detao Shanghai Centre, Shanghai, China and Fieldwork Robotics Ltd, Plymouth, UK.

References

Bleau, A., Leon, L.J., 2000. Watershed-Based Segmentation and Region Merging. Computer Vision and Image Understanding 77, 317-370. https://doi.org/https://doi.org/10.1006/cviu.1999.0822

Eurofresh, 2016. Around the World: Tomatoes. https://www.eurofresh-distribution.com/news/around-world-tomatoes Accessed February 2019.

Feng, Q., Wang, X., Wang, G., Li, Z., 2015. Design and test of tomatoes harvesting robot, in: IEEE International Conference on Information and Automation. pp. 949-952. https://doi.org/10.1109/ICInfA.2015.7279423

GummiArmCE, 2019. GummiArm Community Edition on GitHub. https://github.com/GummiArmCE Accessed February 2019.

Hemming, J., Bac, C.W., Tuijl, B., Barth, R., Bontsema, J., Pekkeriet, E., *et al.*, 2014. A robot for harvesting sweet-pepper in greenhouses. In AgEng Conference, Zurich, Switzerland.

Leu, A., Razavi, M., Langstädtler, L., Ristić-Durrant, D., Raffel, H., Schenck, C., *et al.*, 2017. Robotic Green Asparagus Selective Harvesting. IEEE/ASME Transactions Mechatronics 22, 2401-2410. https://doi.org/10.1109/TMECH.2017.2735861

Quigley, M., Gerkey, B., Conley, K., Faust, J., Foote, T., Leibs, J., *et al.*, 2009. ROS: an open-source Robot Operating System, In: IEEE International Conference on Robotics and Automation (ICRA), Workshop on Open Source Software. 3 (3.2)

Reddy, N.V., Vishnu, A. V, Reddy, V., Pranavadithya, S., Jagadesh Kumar, J., 2016. A critical review on agricultural robots. International Journal of Mechanical Engineering and Technology, 7, 183-188.

Stoelen, M.F., Bonsignorio, F., Cangelosi, A., 2016. Co-exploring Actuator Antagonism and Bio-inspired Control in a Printable Robot Arm, In: Tuci, E., Giagkos, A., Wilson, M., Hallam, J. (Eds.), From Animals to Animats 14. Cham, Switzerland: Springer International Publishing, pp. 244-255.

van Henten, E.J., Hemming, J., van Tuijl, B.A.J., Kornet, J.G., Meuleman, J., Bontsema, J., *et al.*, 2002. An Autonomous Robot for Harvesting Cucumbers in Greenhouses. Autonomous Robots 13, 241-258. https://doi.org/10.1023/A:1020568125418

Wang, G., Yu, Y., Feng, Q., 2016. Design of End-effector for Tomato Robotic Harvesting. IFAC-PapersOnLine 49, 190-193. https://doi.org/10.1016/j.ifacol.2016.10.035

Zhao, Y., Gong, L., Liu, C., Huang, Y., 2016. Dual-arm Robot Design and Testing for Harvesting Tomato in Greenhouse. IFAC-PapersOnLine 49, 161-165. https://doi.org/10.1016/j.ifacol.2016.10.030

Analysis of historical auto-steering position data for optimising headland turning time

D.S. Paraforos*, R. Hübner and H.W. Griepentrog
University of Hohenheim, Institute of Agricultural Engineering, Technology in Crop Production, Garbenstr. 9, 70599 Stuttgart, Germany; d.paraforos@uni-hohenheim.de

Abstract

This study is focused on optimising the headland-turning time, which is needed to traverse all parallel tracks to cover the entire field area. A 4.8 m cultivator and a 2-track tractor equipped with an auto-steering system were utilised for a total area of 8,000 ha during four consecutive years, i.e. 2014 to 2017. Four drivers performed various turning patterns from driving to the adjacent track (T_1), up to driving to the sixth parallel track (T_2-T_6). The cumulative distribution function of the turns was used to perform Monte Carlo simulations considering a field with 10 passes. The optimal path sequence required 68.9% less turning time than performing always a T_1-turn.

Keywords: auto-steering, headland-turning, parallel tracking system, time optimisation

Introduction

Farmers and machine developers strive to minimise the necessary time of agricultural operations and especially the non-working time that is spent on a field. Such an example is the time needed for the turning manoeuvres that take place at the headlands (track-to-track turn) in order to perform the in-field parallel tracks until the entire field area is covered. With the increase in automatic steering systems development, more complex steering paths can be followed including adjacent and non-adjacent pass headland turns.
Most of the performed studies that can be found in the literature propose methods for minimising the non-working travelled distance concluding that this minimisation could also minimise the non-working time spent at the headlands for the necessary manoeuvres. Bochtis and Vougioukas (2008) used an algorithmic approach to find the most suitable sequence of tracks which minimises the non-working distance travelled by the machine. Bochtis *et al.* (2013) compared B-patterns with conventional fieldwork patterns by using the operation simulation. Their study showed a range of non-working distance savings to be between 3.73 and 58.65%, and a range in the increased area capacity to be between 0.50 and 19.23%. An optimal path planning algorithm that was offered in a study by Jin and Tang (2010) was aimed at finding an optimal division of the field into sub-regions with a consequent determination of the coverage directions. Hameed *et al.* (2011) developed an approach for solving the problem related to field area coverage using genetic algorithms.
Time minimisation has been also examined. In a study performed by Seyyedhasani and Dvorak (2018) for reducing field work time using fleet routing optimisation, a conventional human-operator routing was compared to a routing that was improved by the Tabu Search optimisation, in terms of the field completion time. It was proved that a reduction in completion time (by 17.3%) and a total operating time of the vehicle (by 11.5%) can be attained by the computer-optimised routings. Although, in the work of Bochtis *et al.* (2015), the optimisation criterion was the non-working distance, they also examined non-working time as a side effect and indicated a time reduction from 17.5 to 32.4%. Spekken and De Bruin (2013) tried to minimise the non-operational time of agricultural vehicles. They applied the heuristic Clarke-Wright savings algorithm (Clarke and Wright, 1964) resulting in a reduction of up to 50% in turning time.

DOI 10.3920/978-90-8686-888-9_100, © Wageningen Academic Publishers 2019

The main aim of the paper is to find the sequence of a field's parallel tracks that require the minimum headland turning time. This should be achieved without using complex modelling or a high level of automation, but using a thorough analysis of historical headland turning data. The methodology assumes that an auto-steering system with parallel tracking is available but not necessarily automatic headland turning support, i.e. the driver takes control of the steering during turning at the headlands. The contribution of the present work lies in the fact that it uses collected historical auto-steering data from different drivers that perform track-to-track turn with a different number of skipped rows.

Materials and methods

Instrumentation

A 2-track 8345RT 254 kW tractor with a GreenStar™ 3 2630 (GS3) touchscreen display (John Deere, Moline, Illinois, USA) was used in this study. A Tiger LT 5 cultivator (HORSCH Maschinen GmbH, Schwandorf, Germany) with 15 tines and 0.32 m tine spacing, with a total working width of 4.8 m, was pulled by the tractor to perform tillage operations. The tractor was equipped with an AutoTrac™ automatic guidance system with parallel tracking functionality. The GS3 display was connected to a StarFire™ dual-frequency GPS receiver with integrated Terrain Compensation (iTC). The receiver was receiving the John Deere-exclusive free StarFire™ 1 (SF1) differential correction signal, as well as the European Geostationary Navigation Overlay Service (EGNOS) correction signal, providing a 0.15 m pass-to-pass position accuracy.

Parallel tracking system

The tracking mode of the AutoTrac system was configured to follow a straight track using an AB line (A+B method in the GS3 display). In this mode, the driver set an initial point A, usually at the headland and more precisely when entering the field, and then drove the first part of the track manually. After some driven distance, the driver set point B. The necessary distance to point B in the examined case was 3 m but this figure varies from manufacturer to manufacturer. After setting point B, an AB line was created but also the parallel lines to this initial AB line.

On the GS3 display, the currently driven track is usually indicated with a white line. Parallel to this white line, all alternative parallel tracks are shown in blue colour. In the operations performed, the drivers manually drove to one of them and then they aligned the tractor (along with the implement) in the direction of the desired line and activated the autopilot using the toggle switch or the touchscreen. The tractor was then turned onto the selected line and followed it accordingly. Using this technology, as soon as the drivers reached the field boundaries to perform a headland turn, they could choose to continue to the adjacent track (1-track turn) or skip a number of tracks.

Acquired data

Data were collected while using the tractor and the implement cultivating an area of around 800 ha in the northern part of Germany. The recorded data were latitude and longitude in degrees and the tractor heading also in degrees (0-360°). This information was time-stamped in order to facilitate post-processing. The data were collected only when the cultivator was under working operating mode. In order to achieve this, the GS3 display was configured to stop recording data when the operator was giving the command to lift the cultivator. As soon as the driver was engaging the cultivator to enter the soil, the data recording was starting again.

Historical data from four successive years (2014 until 2017) were acquired and analysed resulting in a total cultivated area of around 3,000 ha. In total 1,262,773 records (position data) were acquired from these four years. Considering the 1 Hz acquisition rate, this resulted in 350.8 h of cultivator usage. The tractor was operated by four different persons in order to reduce driver-dependent data. The drivers were only instructed to cultivate the entire area of the field and were left free to decide the starting point and the turning pattern. In terms of experience, the tractor drivers varied from

beginners to experienced operators. The former were internship students with no to some driving experience, while the latter were farm employees with several years of experience in that specific field. The examined headland turns involved all combinations from turning to the adjacent track (T_1), forming an Ω-turn, up to turning to the sixth parallel tack (T_6) forming a U-turn. In Figure 1, an example of all examined headland turning patterns is presented. A fishtail turn was not performed as this is mostly indicated when there is a limited area at the headlands and more specifically for mounted implements. For the specific tractor-implement combination, the Ω-turn (T_1) proved to be an inconvenient way to turn as the tracked tractor created a sand wall in the soil that was difficult to flatten afterwards. This is why most of the operators tried it only a number of times but then stopped doing it.

Turning type and turning time determination

Since the collected data indicated only infield working positions, it was necessary to extract the type of each turn and the corresponding duration. In order to achieve that, the heading information was utilised. One turn was determined by a change of around 180 degrees in direction. After all complete turns had been identified, the remaining incomplete turns were removed. All positions that had a direction difference higher than one degree from the previous position were regarded as neither turns nor straight passes and were removed. This restriction arose from the focus on the headland turns and it was decided in order to avoid passes that would not be classified correctly (e.g. small curves around an obstacle).

Only complete turns on the headland were considered and not, for example, the linear passes performed at the headlands. Each turn was produced based on the orientation as described above and was comprised of a starting and an ending point. Based on the produced turns, the tracks could be indexed with a consecutive number analogous to the entire route. This was done chronologically

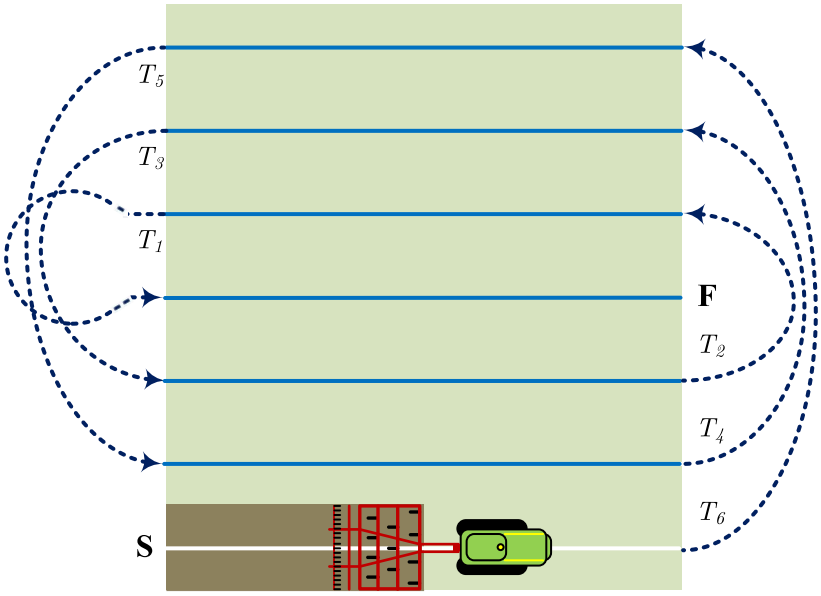


Figure 1. Examined headland turning patterns.

by the provided timestamp. Using this numbering as a basis, an algorithm, which is described in the following paragraph, was developed to calculate the distances between consecutive tracks.
For an accurate calculation of the distances between two consecutive tracks, it was decided not to use only two co-ordinates for the straight paths. Instead, two new co-ordinates were calculated for each one of the two tracks that were involved in one turn. One co-ordinate was from the first half and the second co-ordinate from the second half of the track. Figure 2 presents an example of a turn where the driver had skipped one track (T_2 U-turn) and turned from the first to the third track. The latitude and the longitude of the points L_{A1}, L_{A2}, L_{B1} and L_{B2} were calculated from the mean values of the latitude and longitude of the points that belonged to the corresponding segment. Each segment corresponded to 50% of the points of the straight path. This ensured that a single shifted value had a minimum effect on the distance calculation. All calculated distances were divided using the working width, to give the type of the performed turn.

Monte Carlo simulation
It was necessary to perform a sensitivity analysis that related uncertainties in each headland-turning pattern to uncertainties in the results of the total required time for all headland turns. The uncertainty originated from the various field conditions but it was also influenced by the different driving style of the four drivers. This was done by using a Monte-Carlo simulation as this approach is a well-established technique in metrology and has been used extensively in various domains (Cox *et al.,* 2003).
A field with 10 parallel passes was considered in order to examine the sequence that resulted in the minimum time spent at the headlands. The main assumption was that the driver had to start from pass number 1 close to the boundary of the field and then to drive through all the other nine passes until the entire field area was cultivated. Of course, the restriction was that the maximum allowed turn should have been a T_6 U-turn (i.e. skipping five passes) since this was the maximum turn examined.

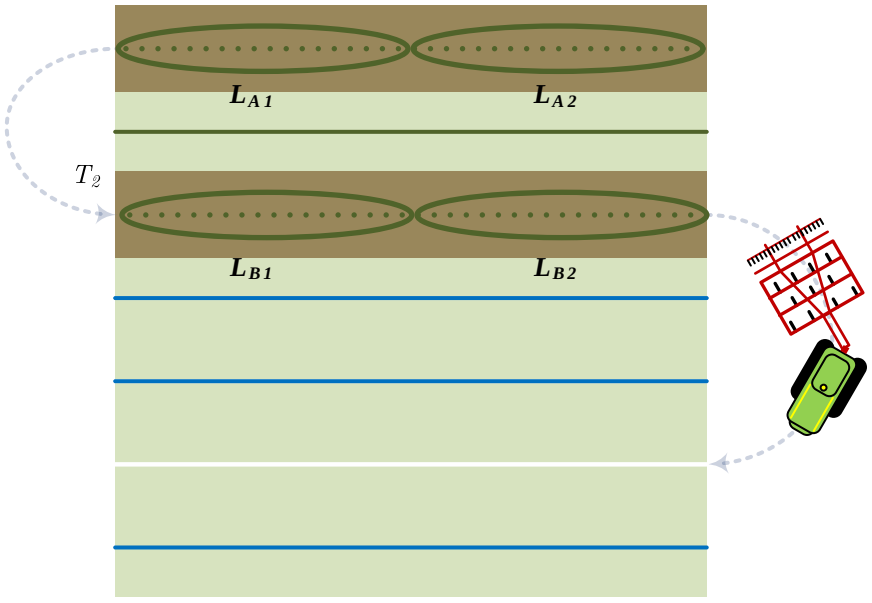


Figure 2. Track-to-track distance determination. In this example, the driver had performed a T_2 headland U-turn. Each of the two tracks was divided into two equal parts.

In total 103,320 permutations were considered that fulfilled the above description. For each one of these permutations, 500 Monte Carlo simulations were performed, while for each turning pattern, the sample was selected from the distribution that best fitted the corresponding dataset. Since the focus of the study was on the headland turning time, this time was only assessed, as the time that would be spent at the parallel passes would have been the same for every different driving sequence (assuming a constant speed).

Results and Discussion

Statistical analysis

The summary statistics for all resulting datasets are presented in Table 1. The maximum time was needed to perform the T_1-turn (38.1 s) while the minimum time was required to perform the T_3-turn with a mean value of 22.3 s. Nevertheless, the standard deviation and the 95th percentile of the T_3-turn, which were equal to 5.4 s and 31.0 s, respectively, were not the lowest since, for the T_5-turn, these figures were equal to 3.31 s and 26.6 s, respectively. The results confirm the need to perform Monte Carlo simulations towards finding the sequence of turning types that gives the minimum time, as by the summary statistics presented in Table 1, it was not straightforward to identify the turning pattern with the overall shortest time.

It was found that the generalised extreme value (GEV) best fitted the data for all turning patterns. The empirical cumulative distribution juxtaposed with the cumulative distribution function (CDF) of the GEV distribution for all six headland-turning patterns (T_1-T_6) are presented in Figure 3A-F, respectively. The results from Table 1 and Figure 3 highlight the increased time that was necessary to perform the T_1-turn and provides an explanation of why the tractor drivers reported that the Ω-turn was not the best pattern to turn at the headlands with the specific tractor-implement configuration. After some turns, all drivers stopped performing it and started using the parallel tracking system to skip some rows, as this was more convenient for them. Table 1 reveals a similar aspect for the T_2-turn as this pattern was the one with the lowest number (60 turns).

Monte Carlo simulation

In Table 2, the five sequences that resulted in the shortest headland turning time for the examined field with 10 passes are presented. It can be noticed that the majority of the turns (five to six turns) had been T_3 turns while the rest in most cases were T_5 turns and some T_4 turns. The sequence with the longest time had all headland turns as T_1 Ω-turn and is also indicated in Table 2. In Figure 4, the Monte Carlo simulation results for the sequences with the minimum (σ_{min}) and maximum (σ_{max}) headland turning time are presented in terms of probability histogram. By comparing the mean values, it is apparent that the σ_{min} required 68.9% less time than the σ_{max}. For the 95th percentile, the difference was almost identical with a value equal to 68.6%.

Table 1. Summary statistics for all datasets (headland-turning patterns).

Turning pattern	No of performed turns	Mean (s)	Standard deviation (s)	95th percentile (s)
T_1	63	38.1	7.3	50.7
T_2	60	27.1	7.3	42.0
T_3	531	22.3	5.4	31.0
T_4	302	23.5	5.5	32.0
T_5	158	23.1	3.3	26.6
T_6	90	25.7	4.1	34.0

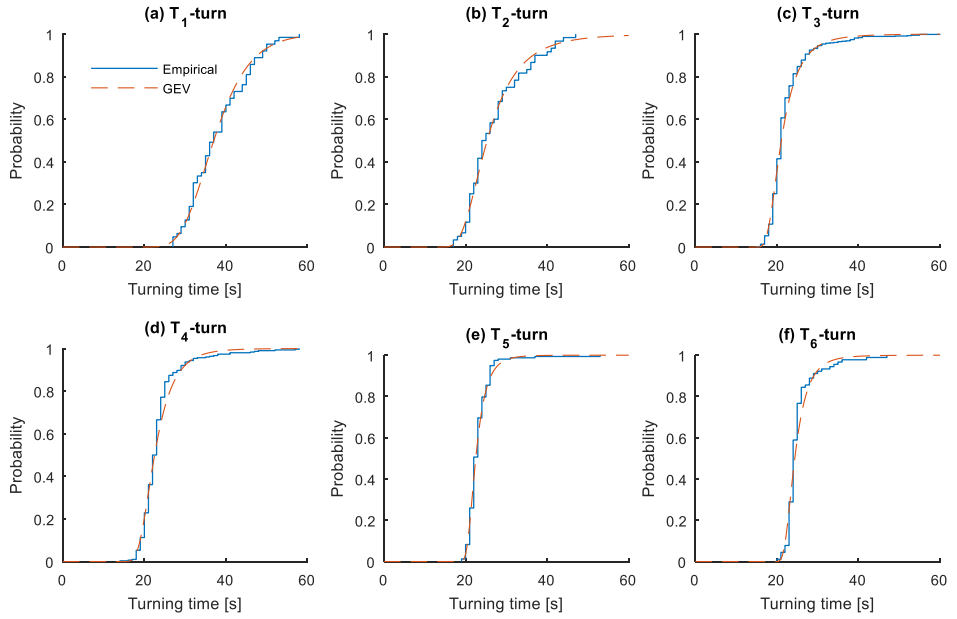


Figure 3. Empirical (blue solid line) and the fitting (red dashed line) GEV cumulative distribution function for all examined turning patterns.

Table 2. The five sequences that resulted in the shortest headland turning time and the one with the longest time.

Sequence	No of turns			Mean (s)	Standard deviation (s)	95th percentile (s)
	T_3	T_4	T_5			
1,4,9,6,3,8,5,10,7,2	5	0	4	202.44	11.6	223.97
1,4,9,6,10,7,2,5,8,3	5	1	3	202.95	12.6	224.35
1,4,9,6,3,8,5,2,7,10	6	0	3	203.06	11.8	224.18
1,4,7,10,5,2,8,3,6,9	6	0	3	204.17	12.1	223.63
1,4,9,6,10,7,3,8,5,2	5	2	2	204.17	11.8	225.85
				⋮		
1,2,3,4,5,6,7,8,9,10	0	0	0	341.93	21.7	377.67

Conclusions

A methodology was developed based on the analysis of historical auto-steering position data, in order to provide the sequence of parallel passes, in terms of skipped tracks during headland turning, which required the minimum time. The T_3-turn pattern, i.e. skipping two tracks and entering the third parallel track, gave the best results. Nevertheless, a thorough statistical analysis using Monte Carlo simulations proved indispensable to reach a concrete result for a field with a specific number

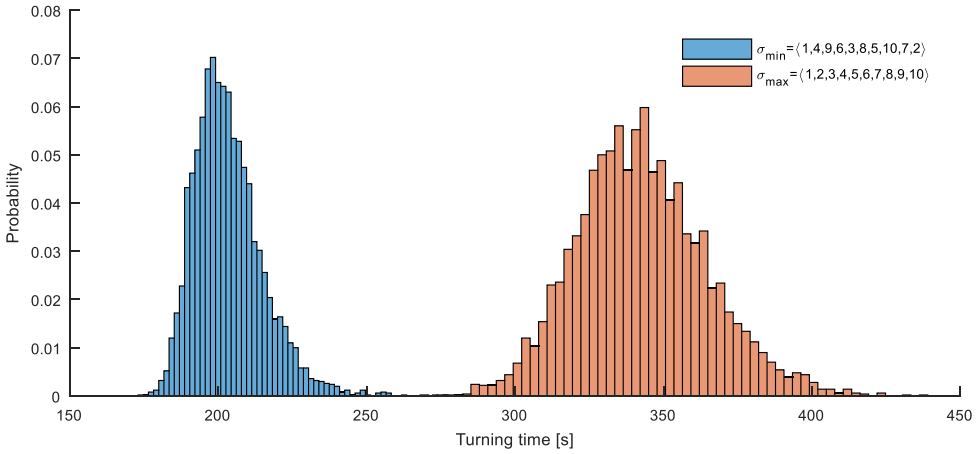


Figure 4. Probability histogram plots of the Monte Carlo simulation results for the sequence with the minimum (σ_{min}) and maximum (σ_{max}) headland turning time.

of passes as using only the summary statistics would not have acknowledged the uncertainty in the turning time.

Distance optimisation or other turning patterns (e.g. fish-tail or slope turn) were not considered in this work but they should be examined if the field shape, the tractor-implement combination, or any other field-specific conditions requires them. In terms of position accuracy, if the farmer desires a higher level of accuracy, an RTK (real time kinematic) solution could be utilised, considering, of course, the related cost. Furthermore, future research should be dealing with the relative contributions of the different sources of variability.

The field with 10 passes was regarded as a simplified case in order to examine the uncertainty that was introduced by the measured data. For a field with a larger number of passes, the number of all possible permutations can be significantly increased, as the number of permutations increases exponentially with each added pass, making an analytical approach not possible. In such a case, heuristic optimisation algorithms should be examined for finding a local optimal solution. Another sub-optimal solution would have been to divide the field into blocks of maximum 10 passes and follow the methodology presented.

References

Bochtis, D. D., Griepentrog, H. W., Vougioukas, S., Busato, P., Berruto, R., & Zhou, K., 2015. Route planning for orchard operations. *Computers and Electronics in Agriculture*, *113*, 51-60. doi:10.1016/j.compag.2014.12.024

Bochtis, D. D., Sørensen, C. G., Busato, P., & Berruto, R., 2013. Benefits from optimal route planning based on B-patterns. *Biosystems Engineering*, *115*(4), 389-395. doi:10.1016/j.biosystemseng.2013.04.006

Bochtis, D. D., & Vougioukas, S. G., 2008. Minimising the non-working distance travelled by machines operating in a headland field pattern. *Biosystems Engineering*, *101*(1), 1-12. doi:10.1016/j.biosystemseng.2008.06.008

Clarke, G., & Wright, J. W., 1964. Scheduling of Vehicles from a Central Depot to a Number of Delivery Points. *Operations Research*, *12*(4), 568-581. doi:10.1287/opre.12.4.568

Cox, M., Harris, P., & Siebert, B. R.-L., 2003. Evaluation of Measurement Uncertainty Based on the Propagation of Distributions Using Monte Carlo Simulation. *Measurement Techniques*, *46*(9), 824-833. doi:10.1023/B:METE.0000008439.82231.ad

Hameed, I. A., Bochtis, D. D., & Sorensen, C. G., 2011. Driving angle and track sequence optimization for operational path planning using genetic algorithms. *Applied Engineering in Agriculture*, *27*(6), 1077-1086. doi:10.13031/2013.40615

Jin, J., & Tang, L., 2010. Optimal coverage path planning for arable farming on 2D surfaces. *Transactions of the ASABE*, *53*(1), 283-295. doi:10.13031/2013.29488

Seyyedhasani, H., & Dvorak, J. S., 2018. Reducing field work time using fleet routing optimization. *Biosystems Engineering*, *169*, 1-10. doi:10.1016/j.biosystemseng.2018.01.006

Spekken, M., & De Bruin, S., 2013. Optimized routing on agricultural fields by minimizing maneuvering and servicing time. *Precision Agriculture*, *14*, 224-244. doi:10.1007/s11119-012-9290-5

Section 14
Economics of precision agriculture

The economic value of on-farm precision experimentation

D.S. Bullock[1*] and T. Mieno[2]
[1]*Univesity of Illinois Department of Agricultural and Consumer Economics, 326 Mumford Hall, Urbana, IL 61801, USA; dsbulloc@illinois.edu*
[2]*University of Nebraska-Lincoln Department of Agricultural Economics, 626 Arbor Dr., Lincoln, NE 68583, USA*

Abstract

Researchers have begun showing that the exact same technology that allows farmers to implement variable rate N application can also be used to run agronomic experiments on their own farms. Thus, farmers may run agronomic experiments and then refine their current N application rates based on the information obtained through the experiments. Monte Carlo simulations were conducted of the dynamic optimization problem of when to switch from on-farm field trials to site-specific management to maximize the net present value of the stream of profits for a given period of time. Results were that on-farm field trials can be highly valuable to farmers, depending on their current N application practice.

Keywords: on-farm experimentation; economic value

Introduction

Variable-rate input application technology has been commercially available for more than two decades. The idea of using truly space-age methods to customize input management to small parts of farm fields was and still is exciting. However, actual adoption of site-specific agricultural technology remains far below the expectations expressed in the late 1990s. Addressing this subject amidst the early excitement, Bullock, *et al.* (1998, 2002) predicted that variable rate technology would be neither profitable nor widely adopted in the then-foreseeable future. They argued that the complementarity between information about yield response and variable rate technology made it infeasible to use the new technology without much more information than was available. Bullock, *et al.* (2009) provided empirical support for this hypothesis.

The economic complementarity of precision technology equipment and yield response information has important research implications and, ultimately, important practical implications, which provide both bad and potentially good news about precision technology. The bad news: unless more information can be produced about yield response, variable rate technology may never be very profitable on a wide scale for major crops. The potentially good news: precision technology itself can be used to inexpensively run large-scale agronomic field trials that increase the supply of the information needed to make precision technology profitable. The data thus generated could be used in concert with other agricultural 'Big Data,' a topic generating much excitement in the media.

The objective of this article is to examine the practicality of the idea that precision technology can be used to sufficiently lower the costs of acquiring the information necessary to make precision agriculture (PA) profitable. In pursuit of this objective, this article offers a detailed discussion about the types of information demanded by precision technology users. Then it discusses how precision technology itself can be used to inexpensively conduct agronomic field trials that increase the supply of that information. Finally, it presents Monte Carlo simulations to gauge the practicality of this idea of using precision technology to increase the information needed to increase the demand for precision technology.

DOI 10.3920/978-90-8686-888-9_101, © Wageningen Academic Publishers 2019

Demand for information about site-specific yield response: what farmers need to know to use precision technology profitably

Figure 1 illustrates to farmers that they have insufficient knowledge about yield response. Following Bullock and Bullock (2000), yield is denoted y, and it is supposed that yield depends on three classes of variables: $y = f(x,c,z)$. The elements of the vector of variables x are managed inputs. These might include the quantities of various types of fertilizer applied during the year, the seeding rate and the tillage system used. The elements of the vector of variables care field characteristics, which might describe the plot's topography, soil type, organic matter content, or soil electoral conductivity. The elements of the vector of variables z are temporally stochastic, non-managed factors of production, for example weather and pest variables. Figure 1 shows 'production functions' like those typically featured in undergraduate agricultural economics textbooks. Because the diagrams are two-dimensional, the values of the c and z vectors are held constant along the curves showing relationships between x_1 and y. Also for the purposes of illustration, it is assumed that the field under consideration has four distinct rectangular 'sites,' labeled A, B, C, and D. Vector c^i denotes the field characteristics of site $i = A,B,C,D$, and each site's characteristics are assumed to differ from each other site's characteristics: $c^A \neq c^B$, etc. These inequalities mean that even if every site receives the exact same weather, the sites' yield responses to the input x_1 differ. Thus, the four curves labeled $f(x_1,c^A,z^{2018})$, $f(x_1,c^B,z^{2018})$, $f(x_1,c^C,z^{2018})$, and $f(x_1,c^D,z^{2018})$ in Figure 1 can be called 'site-by-year-specific response functions'. Figure 1 also depicts the profit-maximization problem's first order conditions; the optimal application rates for x_1 at the sites, given an input-price/output-price ratio of w_1/p are x_1^{A18}, x_1^{B18}, x_1^{C18}, and x_1^{D18}. If yield response to managed inputs did not vary spatially, there would be no need for precision agriculture technology, since the economically optimal rates of managed input application would be equal throughout a field in any given year. In contrast, because response does differ spatially, precision agriculture has the potential to increase farm profits. However, the profitability of precision agriculture technology depends crucially on the farmer's knowledge of the many site-specific production functions in his/her field. Farmers who do not know how yields respond to managed inputs, and how those responses differ spatially, are quite limited in their abilities to make good decisions about those inputs' application rates and how to vary them spatially. That is, farmers need to know the response function 'map,' as illustrated in Figure 1.

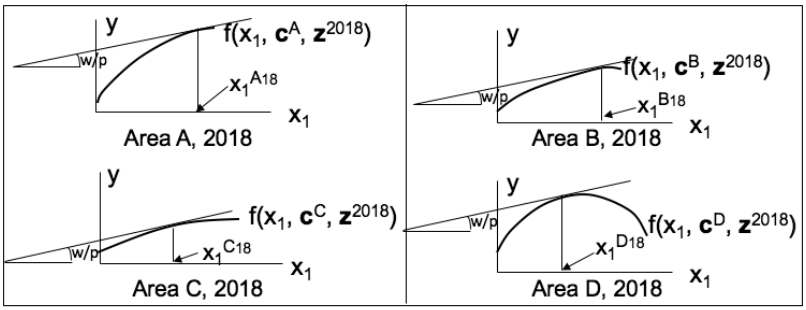


Figure 1. A "map" of site-and-year-specific yield response curves (y-axis yield), and profit-maximizing input application rates (x-axis).

Supply of information about site-specific yield response: how precision technology can be used to make precision technology profitable

Crop scientists and agricultural economists have been attempting for generations to estimate yield response functions with small amounts of data (e.g. Heady and Pesek, 1954; Cerrato and Blackmer, 1990). However, that research has been insufficient to provide farmers with the detailed knowledge needed about their yield response maps, especially now that precision technology is available. Yield-response studies were based on very limited data because until recently, conducting field trials was very labor-intensive and very expensive. Field trials have been implemented using extremely labor-intensive techniques; for example, researchers have marked off plots of land using measuring tapes and orange flags, applied inputs by hand or with specialized equipment at varying rates on different plots, and harvested without the benefit of large-scale farm machinery. This labor intensity meant that it was generally only financially feasible to run trials on very small areas of land, at few locations, usually only for a few years. Consequently, despite generations of research, little is known about optimal fertilizer management; university- and industry-provided fertilizer management recommendations have been based on 'rules-of-thumb' (e.g. Hoeft and Peck, 2007) that are, at best, based only loosely on data analysis and science (Rodriguez, *et al.*, in press).

The idea of on-farm research is not new. Brouder and Nielsen (2000) offer practical steps to farmers wanting to obtain data from randomized input application trials on their own fields. They recommended (p. 103) that farmers develop experimental designs that attempt to answer simple 'yes/no' management questions, such as whether one hybrid provides higher yields than does another, or whether one fertilization date offers higher profits than does another. They explain fundamental concepts of randomization and replication, review strip-trial design strategies, and discuss the importance of statistical analysis. Agronomic research projects have recently demonstrated that precision technology can be used to conduct very large, more complex field trials at very low cost, and so may make it possible to gather the information needed to make variable rate profitable in commercial farming (c.f., Bullock, *et al.,* 2009; Peralta, *et al.,* 2015). Figure 2 depicts one such field trial, conducted in 2018 on an eastern Washington commercial farm in participation with the farmer, who actually put the experiments 'in the ground' by operating the farm machinery in his usual manner while rates of NH_3 nitrogen application were controlled by computer and GNSS equipment according to the researchers' pre-programmed field trial design. As is shown in Figure 2, that design

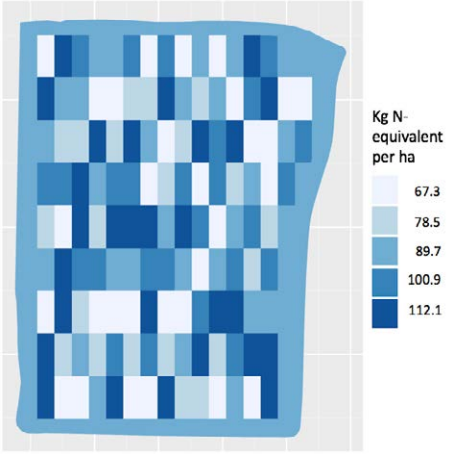


Figure 2. An on-farm nitrogen fertilizer trial, conducted on a 37.9 ha wheat field in eastern Washington in 2018.

included five nitrogen application rates that were used in the randomized design. The depicted field was 37.9 ha in size, with each plot 26.8 m wide by 83.8 m long. In 2019, the Data-Intensive Farm Management (DIFM) program will conduct approximately sixty such large field trials at very low costs in eight U.S. states, through at the University of Illinois, Louisiana State University, Montana State University, University of Nebraska and Washington State University. DIFM is funded with support from the USDA's National Institute for Food and Agriculture.

A Monte Carlo simulation of the relationships between application rate technologies and information from agronomic experimentation

To examine the potential value of conducting field trials such as the one depicted in Figure 2, Monte Carlo simulations were conducted to initiate an empirical examination of the relationships between precision input application rate technologies and information about yield response. The experimental design involved a 'true' meta-response function, a field with spatially autocorrelated characteristics, random assignments of nitrogen fertilizer rates to field blocks, and resultant experimental yields.

The simulations' 'true' meta-response function f

Throughout, the analysis used as the 'true' yield response function was the estimated response function (in kg/ha) reported in table I of Bullock *et al.* (2009), for Farmer *D* in Champaign County, Illinois, USA:

$$f(N,M,S,I,\varepsilon) \equiv \beta_0 + \beta_N N + \beta_D D + \beta_I I + \beta_M M + \beta_R R + \beta_S S + \beta_{NN} N^2 + \beta_{NI} NI + \beta_{NM} NM + \beta_{NS} NS + \beta_{NNI} N^2 I + \varepsilon \quad (1)$$

where β_0=5,524, β_N=154.97, β_{NN}=-0.4461, β_M=-108.5, β_S=-194.7, β_I=52.33, β_D=-2,440, β_R=0, β_{NM}=0.1624, β_{NS}=0.7537, β_{NI}=-0.4866, β_{NNI}=0.001318, and ε is a normally-distributed error term with mean zero and standard deviation of 200 kg/ha.

The variables in Equation 1 are explained in detail in Ruffo, *et al.* (2006), and are *N* (nitrogen fertilization rate, in kg/ha), *D* (a dummy variable, here always equal to one), *M* (May precipitation, in mm), *R* (a dummy variable, in this investigation always equal to zero), *S* (the value of the stream power index (*SPI*)), and *I* [the value of the Illinois Nitrogen Soil Test (*ISNT*) (Khan, *et al.,* 2001) result]. In reality, the *ISNT* is an index of soil-borne nitrogen and changes over time. For the purposes of the simulations, the *ISNT* values are treated as exogenous and constant over time.

The field

In each Monte Carlo run, a farm field with spatially autocorrelated production characteristics was simulated. A rectangular 366 732 m field is assumed. To simulate agronomic experiments, the field was divided into thirty-two 91×91 m square blocks, each containing five 18.3×91 m plots. It was assumed that assumed that the model's geospatial characteristics (the *SPI* and *ISNT* values) varied by block, but did not vary within any block. For each Monte Carlo round separately, a spatially autocorrelated *ISNT* map and a spatially autocorrelated *SPI* map were generated as follows.

First, in reference to Ruffo, *et al.* (2006), thirty-two independent draws were taken from a normal distribution with mean zero and a standard deviation of 40.0. A vector called δ was constructed with these thirty-two values as its elements. The ISNT map was generated through the autoregressive process in Equation 2. The associated data generating process statement is in Equation 3. Equations 2 and 3 are versions of Equations (1.10) and (1.11) of LeSage and Pace (2009, p. 13):

$$V_{ISNT} = \alpha i_{32} + \rho W \times V_{ISNT} + \delta \quad (2)$$

$$V_{ISNT} = [I_{32} - \rho W]^{-1} i_{32}\alpha + [I_{32} - \rho W]^{-1} \delta \quad (3)$$

In Equations 2 and 3, i_{32} is a vector of thirty-two values of 1.0, I_{32} is a 32×32 identity matrix, ρ=0.5 is the dependence parameter, and α = 235(1 − ρ) = 117.5 is a constant. W is a 32×32 normalized matrix of spatial weights, which has zeroes on the diagonal, and the distance between block centroids, raised to the power of -4, elsewhere. The SPI map was similarly generated, but assuming a mean of 1.5 and a standard deviation of 0.3.

Simulated data sets

Using the characteristics map generated in a Monte Carlo run, a thirty-year series of agronomic experiments on the field was created, which resulted in a simulated yield in every block in every year. For every year of modeled experimentation, a value for the random variable May precipitation, in millimeters, was drawn from a normal distribution with mean $\overline{M}$=100 and standard deviation σ_M=10. It was assumed that in a given year *t*, the same amount of rain per ha, M_t, fell on every plot in the field. For each of the 160 plots in the field, a random yield disturbance term was from a normal distribution with mean 0 and standard deviation 200 kg/ha. Next, in every block, to each of the five plots one of five nitrogen fertilization rates was randomly assigned: 125, 150, 175, 200, and 225 kg/ha. In each Monte Carlo run, for any given block *b*, all the plots in the block were assigned the same ISNT value I_b, and the same SPI value S_b, and these values were maintained constant across the years of the simulated experiment. Given the plot's block's assigned levels of ISNT and SPI values, for every plot in every block in every year of the experiment, a yield value was generated according to Equation 1. Thus, every observation in the resultant simulated data set had a value for the applied nitrogen fertilization rate, a value for May precipitation, a value for its random disturbance term, and a yield. Since the experiment was simulated for five plots on each of thirty-two blocks for thirty years, the resultant number of observations was 5×32×30=4,800. To simulate varying lengths of agronomic experimentation, for each Monte Carlo run, years from the all-block, 30-year data set were simply omitted.

Results

One hundred Monte Carlo simulations were conducted using the steps outlined above. Throughout the analysis, a corn output price of p^0=0.246$/kg and a nitrogen fertilizer price of w^0=0.80$/kg for anhydrous ammonia were assumed. The objective was to examine the values and marginal values of information from years of experimentation, given uniform and site-specific technologies, and to measure the value of site-specific technology as compared to uniform technology, given varying levels of information. Ordinary least squares (OLS) was used in all regressions. (An extension of this research would develop a *Mathematica* program using the spatial econometrics methods (Anselin, *et al.,* 2004).)

The causes of the results above have to do with how the presence of 'partial information,' and the advantages of using a particular data set to estimate a whole-field response function instead of many block-specific response functions. In the simulations using the 30-year data set, estimation of the whole-field-response-to-uniform *N* function $f_{wf}(N,M,\varepsilon)$ used 4,800 observations. However, estimations of block-specific response functions $f_b(N,M,\varepsilon)$, were based on only one thirty-second as many observations, that is, 150. As a result, the estimates of the whole field's optimal uniform N rate were far more accurate than the estimates of the optimal N rates for the individual blocks. So, here we see a trade-off. By using site-specific technology, the farmer has the advantage of being less constrained in his/her choices. However, a given number of field trials (that is, as defined, a given amount of information) can be more useful to the uniform-rate farmer than to the site-specific farmer, because the increased number of observations make for better estimations of the optimal uniform N fertilizer rate than for the optimal site-specific rates. Now, these results will depend on the heterogeneity of the spatial characteristics in the field. The field used in the model presented here emulated a central Illinois 'flat and black' field, and the spatial heterogeneities of the SPI and

ISNT values were not great. This leads to similar site-specific inputs rates among sites, and limits the site-specific technology's advantage. Obviously, extending the analysis to examine the values of information and technology types of fields less spatially homogeneous is called for.

Results were that the mean of the value of site-specific technology under full information was $2.19/ha from all one hundred Monte Carlo runs. That is, by using site-specific variable rate technology instead of uniform rate technology, a farmer who knew every block's true response function, $f_b(N,M,\varepsilon)$ could expect to be able to increase annual net revenues (revenues minus N fertilizer costs) by $2.19/ha. In the context of the model used here, in which there is no change in the response functions over time, this advantage provided by the site-specific technology would be gained every year. The present value of a stream of annual net revenue increases of $2.19/ha/yr would be over $40/ha/yr at an assumed 5% discount rate, and over $100/ha at an assumed 2% discount rate. The simulation results suggest that replacing uniform technology with site-specific technology would not pay for itself when information is not full and free. Farmers using site-specific technology with full information make higher net revenues than do those using site-specific technology but possessing only information from agronomic experiments of various lengths. The farmer who had full information and used site-specific technology could expect to enjoy net revenues more than $10/ha greater than those of the farmer using site-specific technology while having only four years of experimental data. Even for farmers with over twenty years of experimental data, net revenues from precision farming would be about $3/ha less than the net revenues of farmers using site-specific technology with full information. Clearly, in the simulations, information about crop response to managed inputs is key to the profitability of site-specific technology. The simulations showed that the information is worth much less to the uniform rate farmer than to a site-specific farmer. The net revenues of the uniform-rate farmer who only has information from an experiment of four years in length are only about $0.60/ha less than those of a uniform rate farmer with full information. In the simulations, the marginal value of the information from the fifth year's experiment (that is the value of the information gained from running five years of experiments instead of four) averaged approximately $1.50/ha over the Monte Carlo runs. After the fifth year, the value of the marginal product of experimentation dropped steeply, and after 15 years was consistently under $0.25/ha. This result occurs because the accuracy of the estimates of the site-specific N fertilization rates is not much improved by an additional year of experimentation once data from about seven or ten years of experimentation is at hand. Finally, Figure 3 shows the Monte Carlo runs' ex-ante values of site-specific technology under various amounts of partial information. For every experiment length, from four to thirty years, this value was negative. For the farmer with only partial information, net revenues are always higher under uniform technology than under site-specific technology, even when the differences in the costs of production under these technologies is not accounted for.

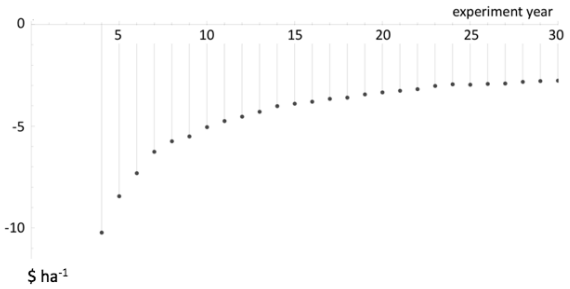


Figure 3. Value of site-specific technology under various amounts of information

Figure 3. Value of site-specific technology under various amounts of information.

Conclusions

PA technology has been commercially available for more than twenty years. Initially, it generated great enthusiasm in the agricultural press. Very gradually, a market for the technology has developed, principally with private consulting firms attempting to increase farm profits by offering site-specific management advice. For the farmer who purchases it, the results of the simulations reported here suggest that the profitability of this advice is open to question. Managing inputs site-specifically may require large amounts of information about how crop yields respond to managed inputs, and how those responses vary over time and space.

Acknowledgements

This research was funded in part by a USDA National Institute of Food and Agriculture Food Security Program grant, award number 2016-68004-24769, and by USDA NIFA's Hatch Project 470-362.

References

Anselin, L., Bongiovanni, R. and Lowenberg-DeBoer, J., 2004. A spatial econometric approach to the economics of site-specific nitrogen management in corn production. American Journal of Agricultural Economics 86 675-687

Brouder, S. and Nielsen, R., 2000. On-farm research. In Lowenberg-DeBoer, J. and Erickson (Eds.). In: Precision Farming Profitability. Site-specific Management Center, School of Agriculture, Purdue University, IN, USA pp. 103-112. https://ag.purdue.edu/agry/courses/Documents/AGRY%20598/Precision%20Farming%20Profitability.pdf. (last accessed January 2019)

Bullock, D, S., Ruffo, M. L., Bullock, D. G., and Bolero, G.A., 2009. The value of precision technology: an information-theoretic approach. American Journal of Agricultural Economics, 91 209-223.

Bullock, D. G., Bullock, D. S., Nafziger, E. D., Peterson, T. A., Carter, P., Doerge, T. *et al.,* 1998. Does variable rate seeding of corn pay? Agronomy Journal 90 830-836.

Bullock, D. S., Lowenberg-DeBoer, J. and Swinton, S., 2002. Adding value to spatially managed inputs by understanding site-specific yield response. Agricultural Economics 27 233-245.

Bullock, D.S., and Bullock, D. G., 2000. From Agronomic Research to Farm Management Guidelines: A Primer on the Economics of Information and Precision Agriculture. Precision Agriculture, 2 (1) 71-101

Cerrato, M. E., and Blackmer, A. M., 1990. Comparison of models for describing corn field response to nitrogen fertilizer. Agronomy Journal 82 138-143.

Heady, E. O., and Pesek, J., 1954. A fertilizer production surface with specification of economic optima for corn grown on calcareous Ida Silt Loam. Journal of Farm Economics 36 466-482.

Hoeft, R. G., and T. R. Peck., 2007. Soil testing and fertility. In: R. G. Hoeft and E. Nafziger (Eds.) Illinois Agronomy Handbook. Urbana-Champaign, IL, USA: Dept. of Crop Sciences, University of Illinois, pp. 91-131.

Khan, S. A., Mulvaney, R. L., and Hoeft, R. G., 2001. A simple test for detecting sites that are non-responsive to nitrogen fertilization. Soil Science Society of America Journal 65 1751-1760.

LeSage, J., and Pace, R. K. 2009. Introduction to Spatial Econometrics. Boca Raton, FL, USA: Chapman & Hall/CRC Press.

Peralta, N. R., Costa, J. L., Balzarini, M., Franco, M. C., Córdoba, M., and Bullock, D., 2015. Delineation of management zones to improve nitrogen management of wheat. Computers and Electronics in Agriculture 110 103-113.

Rodriguez, D. G., Bullock, D. S., and Boerngen, M. A., in press. The origins, implications, and consequences of yield-based nitrogen fertilizer management.' Agronomy Journal. Accepted Oct. 2018. doi:10.2134/agronj2018.07.0479

Ruffo, M. L., Bollero, G. A., Bullock, D. S., and Bullock, D.G., 2006. Site-specific production functions for variable rate corn nitrogen fertilization. Precision Agriculture 7 327-342.

The emerging structure of the U.S. precision agriculture industry[1]

J.R. McFadden* and D.E. Schimmelpfennig
Economic Research Service, U.S. Department of Agriculture, Washington, DC, USA;
jonathan.mcfadden@ers.usda.gov

Abstract

This study provides average trends for the U.S. precision agriculture industry as of mid-2017 using a new dataset of nearly 200 private companies from publicly-available website information. Firms' average age, size and headquarters varied significantly by technology type. On average, field production equipment and guidance companies tended to be among the oldest and largest firms in the industry, with headquarters mainly in the central United States. Imagery service providers, imagery analytics and several automation companies were among the youngest and smallest, with headquarters on the East and West U.S. coasts. This descriptive overview, however, is meant to be illustrative rather than exhaustive.

Keywords: market structure, technology adoption, entry and exit, mergers and acquisitions

Introduction

The evolution of farmers' technology choices are significantly influenced by the development, commercialization and subsequent availability of product innovations. After a new product is commercially introduced, there tends to be a period of rapid increase in the number of technology suppliers, followed by periods of zero – and then negative – net entry, until eventual obsolescence of the technology when suppliers have dropped out of the market (Gort and Klepper, 1982). This characterization assumes that industry boundaries are well-defined and generally static. However, new technologies that initially serve a set of niche uses or users may eventually enter more mainstream markets, bringing about competitive pressure (Adner and Zemsky, 2005). Such 'disruptive' technologies can influence the rate of technological progress, technology market size, pricing and ultimately industry boundaries.

The global precision agriculture industry is experiencing a period of transformative growth (Lowenberg-DeBoer, 2015; Cornell University, 2018). Various factors have contributed to this transformation, including the advent of cheap computer processing and storage and reliance on algorithms using big data (Coble *et al.*, 2016; Coble *et al.*, 2018). This has resulted in increases in the number of technology developers and custom service providers offering precision planting and spraying (Erickson *et al.*, 2017), as well as increases in the scope and quality of precision technologies. Some of these new technologies, like complete or near-complete automation of field operations in row crop production or robotic livestock management, could prove to be 'disruptive' (i.e. significantly reduce farmers' direct and indirect costs of production). However, the breadth of this structural change is unknown, in part because little rigorous research has analyzed the current industry structure: average number, size and age of firms, their product lines, and prevailing technology market prices.

The objective of this research was to provide an illustrative overview of the current organization of the U.S. precision agriculture industry. Unlike much past work that explores farm-level adoption and

[1] This work did not receive extramural funding, and the authors have no financial interests in application of this research, which was supported by the intramural research program of the U.S. Department of Agriculture, Economic Research Service. The findings and conclusions in this preliminary article have not been formally disseminated by the U.S. Department of Agriculture and should not be construed to represent any agency determination or policy.

DOI 10.3920/978-90-8686-888-9_102, © Wageningen Academic Publishers 2019

profitability (e.g. Griffin *et al.*, 2004; Schimmelpfennig, 2016), this study follows a parallel literature that identifies trends in agricultural input market research (e.g. Fuglie *et al.*, 2011). Specifically, this study analyzes new data obtained from publicly-available websites of nearly 200 U.S. precision agriculture firms as of 2017. Within the broader context of structural change and continuing erosion of the traditional distinction between hardware and software companies, trends in firms' age, size and business headquarters are illustrated and discussed.[2] Given the similarities between the U.S. and Europe regarding regional commercialization and adoption of these technologies (Barnes *et al.*, 2019), many of this study's implications apply to Europe.

Materials and methods

Website data were collected on a set of private firms listed in the *Guide to Precision Agriculture and Ag-Tech Companies* (*Guide*) for early summer 2017 published by Meister Media Worldwide. Meister is a U.S. media company that primarily publishes magazines, e-newsletters, and other digital content tailored to U.S. growers of fruit, vegetables, and certain row crops. Major brands include *American Fruit Grower*®, *CropLife*®, *PrecisionAg*®, and other publications on specialty crop production. The *PrecisionAg*® branded media, under which the 2017 *Guide* is published, has a self-reported outreach in excess of 20,000 average monthly website users and 14,000 e-newsletter subscribers (Meister Media Worldwide, 2018). In 2014, roughly 48% of the *PrecisionAg*® audience of 42,000 were corn and soybean farmers managing at least 809 ha, 35% were agricultural retailers and dealers, and 8% were fertilizer or pesticide producers (Meister Media Worldwide, 2014).

Table 1. Summary statistics for sample of U.S. precision agriculture companies in 2017.

Technology category	Technology subcategories: range of products and/or services	Number of firms	Percent that are subsidiary
Automation	robotics, artificial intelligence (AI)	15	13%
Climate-weather	weather reporting & analysis	13	38%
Equipment	large machinery, hardware, GPS	21	10%
Field monitoring	internet of things (IoT), soil health	13	8%
Guidance	GPS, IoT, mapping, data analysis	12	17%
Imagery analytics	data analysis, custom mapping, AI	13	23%
Imagery service provider	aerial/geospatial services, mapping	10	10%
Imagery systems	aerial services, mapping, GPS	6	17%
Input efficiency	data analysis/management, AI, IoT	5	20%
Irrigation hardware	water monitoring, IoT, mapping	12	17%
Sensors	input monitoring, IoT, hardware,	15	13%
Smart irrigation	water monitoring, mobile apps, IoT	14	21%
UAV control platforms	aerial services, small machinery	5	20%
UAV platforms-systems	software, aerial services, AI	16	7%
VRT analytics	software, mapping, soil health	4	25%
VRT applications	management, ROI, hardware	14	29%

[2] Many case studies on precision agriculture adoption have been published in the past 20 years. Given the unique challenges associated with early adoption, these studies have provided valuable insights into the determinants of (and barriers to) technology uptake. Our goal is to extend the literature by illustrating and discussing general trends based on analysis of a greater number of observations than datasets of most case studies.

The *Guide* was designed to be an inventory of major companies in the global precision agriculture industry as of early summer 2017. Of the *Guide*'s 39 technology categories, data were collected on companies falling under 16 of these categories. To be included in the study, companies must have: (1) been headquartered in the U.S. and (2) developed and/or supplied products or services considered to be precision agricultural technologies used in major U.S. row crop production (following Griffin *et al.*, 2004, Schimmelpfennig, 2016).

As such, producers of technologies in categories excluded from this analysis were either headquartered outside of the U.S. or deemed to be outside the scope of conventional precision agriculture used in crop or livestock production (e.g. cold chain monitoring, food recovery).

These criteria led to an initial dataset of 206 precision agriculture companies, most of which are technology developers. Of this initial set, three companies were dropped due to lack of a website and any online information. Five companies had inactive websites but were included in the analysis as information was available from other online sources. Six companies were bought or merged with other companies since early summer 2017. This resulted in a set of 200 companies, though data on relevant market information (size, age, headquarter location, range of products and services, and information about affiliate or parent companies) were not available for 12 of these companies. This resulted in a final dataset of 188 companies.

Significant variation exists in the foundational technologies used to develop precision agriculture equipment and services (Table 1). Automation technologies rely heavily on artificial intelligence (AI), while imagery service providers rely mainly on mapping. Yet there is also substantial overlap in technologies underlying precision products. Besides automation, AI is also a key input for imagery analytics, input efficiency, and unmanned aerial vehicle (UAV) platforms and systems. This overlap is further evidenced by the number of firms with parent companies. Between 7 and 38% of firms in the sample are subsidiaries of other companies, depending on technology category.

Why do some sectors have a greater number of firms, while others have very few? There are a number of potential explanations, though the underlying economics of product innovations plays a crucial role. In sectors with low barriers to entry, like imagery-based companies, the industry can sustain multiple entrants, at least in the short run, if market prices for their technologies or services exceed the firms' average variable costs. In contrast, the UAV control platforms and VRT analytics sectors might have relatively fewer firms because of higher barriers to entry (e.g. specialized engineering expertise required for innovation). Despite differences between U.S. and European precision agriculture policy (e.g. Kritikos, 2017), these economics shape the structure of the European precision agriculture industry.

Data on the number of years of active business operations were inferred from information about the year in which the company was founded. Average age was computed with respect to the number of companies within each technology category. Headquarter locations and employment data were taken from company websites. If exact employment data were not available on these websites, data were taken from the company's LinkedIn profile, which provided interval-coded company sizes (i.e. 2-10 employees, 11-50 employees, 51-200 employees and so on). A small number of very large firms were doing business in several of the 16 technology categories (e.g. Monsanto, John Deere), which inflated estimates of average company sizes.

Results and discussion

Average company ages prevailing in the U.S. precision agriculture industry are consistent with expectations (Figure 1). Among the youngest firms in the sector are imagery service providers, unmanned aerial vehicle (UAV) platforms-systems and imagery analytics – averaging well under 10 years of commercial operations. Companies in these sectors make heavy use of drones, AI and sophisticated data analyses – all of which are more nascent technologies that have only recently begun to be combined. In contrast, equipment, irrigation hardware and variable rate technology

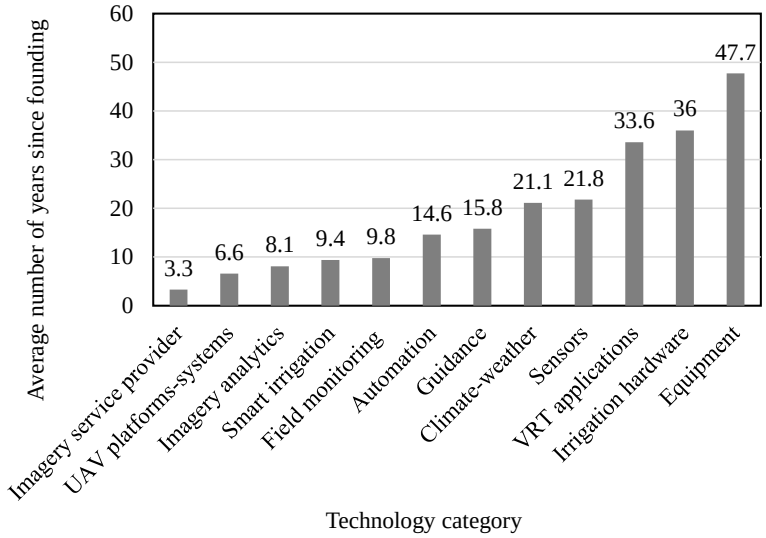


Figure 1. Average number of years of commercial operations, by technology category.

(VRT) applicators are industry incumbents with average ages in excess of 30 years. Between these extremes are automation and guidance, conventional precision agricultural technologies that began to be commercialized in the early 2000s (Schimmelpfennig, 2016).

Fewer inferences can be made about size trends by technology sector because of less granularity in the underlying employment data (Figure 2).[3] Most of the companies in the sample are small: the modal size category (for 63 firms) is 11-50 employees. Among the smallest are smart irrigation and field

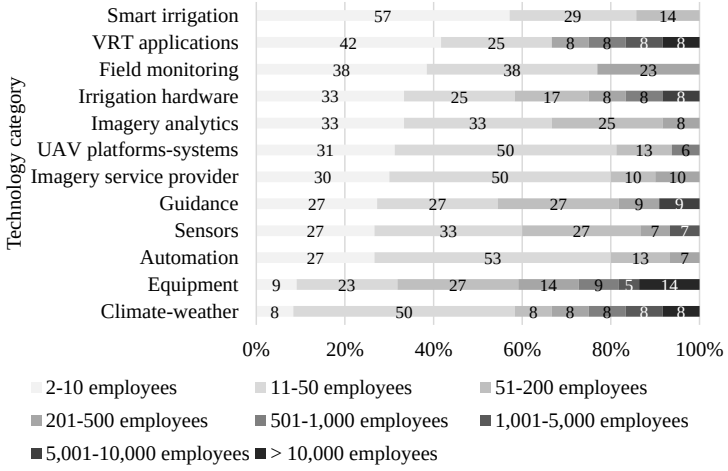


Figure 2. Employment size categories, by technology category.

[3] Four technology categories were represented by fewer than 10 firms (Table 1). These categories were dropped from the analysis in Figure 1 and 2 due to the small sample size.

monitoring companies that specialize in Internet-of-Things (IoT) solutions. These companies rely on data transmission to and from sensors connected to various types of small and large machinery that perform field operations.

The largest companies produce guidance systems, climate-weather predictions and input applications equipment. Many of these companies also manufacture large tractor and combine machinery, requiring sizeable staffing in engineering, operations research, supply-chain management and human resource departments.

Intuitively, some of the oldest companies are also among the largest (e.g. John Deere, Monsanto). Profitability of product sales, technology acquisitions, patent protection of new technologies and financial investments have driven their expansion. Economies of scope in research and development, whereby firms' average costs decrease as more types of products are developed, have also played a significant role in company expansion.

Precision agriculture companies are headquartered in at least 38 U.S. states (not including Washington, DC), but they are highly geographically concentrated (Figure 3). Roughly 29% of sample companies are headquartered in California, which has significant specialty crop area – typically requiring irrigation and specialized machinery – and an ample labor supply of engineers, computer programmers, software designers and data scientists. The technology sectors with the greatest proportion of California-based companies are imagery service providers (80%), imagery systems (66%), UAV control platforms (60%), input efficiency (60%) and automation (40%). Massachusetts is home to roughly 7% of sample companies. Like California, Massachusetts also has a labor force with the kind of engineering technical expertise needed to develop various precision technologies but is mainly represented among automation companies (20%). The Midwestern states of Illinois, Minnesota, Colorado and Iowa collectively account for 24% of sample companies.

Rather than the data-intensive automation and algorithm development occurring on the West and East coasts of the U.S., precision technologies developed in the Midwest are oriented more towards hardware and large machinery. These states are mainly represented among VRT analytics (50%), VRT applications (38%), equipment (33%), guidance (33%) and smart irrigation (29%) companies. One exception to this is sensors, with 33% of sample companies coming from the Midwest. Some of this results from medium-sized and large industrial firms that produce sensors for agricultural use (e.g. Banner Engineering).

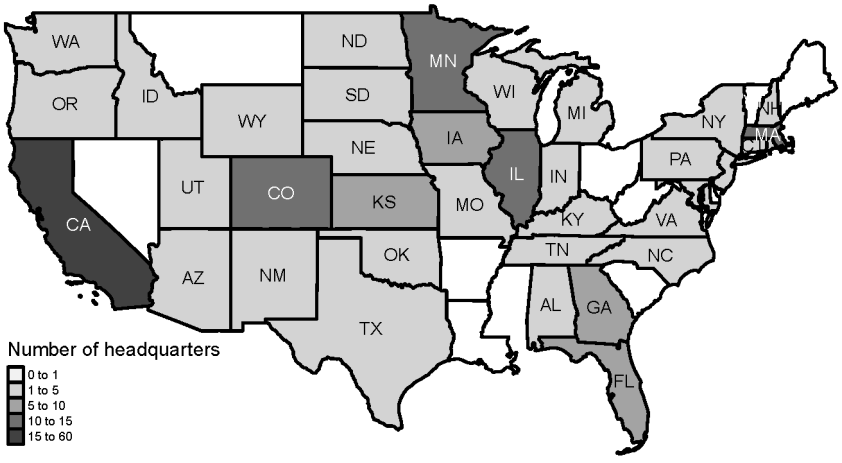


Figure 3. Location of sample company's headquarters. One company is headquartered in Hawaii, not depicted.

One important dimension of the ongoing precision agriculture transformation not apparent from the results is the significant depth of intersecting technologies. Although firms were sorted into one of 16 categories, many would have been appropriately classified as suppliers of 2-4 technologies. For example, John Deere is an agricultural equipment company that also manufactures VRT and guidance machinery. Given its $305 million acquisition of Blue River Technology in September 2017, Deere could also be conceptualized as an automation firm. Boundaries between software and hardware development are expected to dissolve further if usage of IoT and big data algorithms in U.S. agriculture increases as expected. One potential implication could be eventual consolidation among one or more large firms, not unlike recent mergers in agricultural chemicals markets (MacDonald, 2017).

Although this study relied on analysis of U.S. companies, many of the underlying trends are relevant for Europe. As in the U.S., major innovations in precision agriculture technologies are undertaken by European firms with the potential for economies of scale and/or scope under legal regimes with strong intellectual property rights. Unlike U.S. policy provision, rural development provisions of the Common Agricultural Policy can incentivize adoption of precision agriculture in European countries. Differences in financial assistance policies, as well as further differences between the U.S. and European antitrust and regulatory environments, are expected to influence the rate and direction of technological progress in precision agriculture. Consequently, the composition of Europe's precision agriculture industry will differ from the U.S. industry but will have commonalities: many of the oldest firms will be among the largest, located in regions with relatively low technology costs.

Conclusions

This study provided average industry characteristics for the U.S. precision agriculture market as of mid-2017 using a new dataset of nearly 200 private companies from publicly-available website data. Firms' average age, size and headquarters varied significantly by technology type. On average, equipment and guidance companies tended to be among the oldest and largest firms in the industry, with headquarters mainly in the central U.S. Imagery service providers, imagery analytics and several automation companies were among the youngest and smallest, with headquarters on the East and West U.S. coasts.

Acknowledgements

The authors thank Rachel Ganson, financially supported by a 2018 USDA fellowship, for her diligent assistance in compiling data. The authors are solely responsible for any errors.

References

Adner, R. and Zemsky, P., 2005. Disruptive Technologies and the Emergence of CompetitionRAND Journal of Economics 36(2) 229-254.

Barnes, A.P., Soto, I., Eory, V., Beck, B., Balafoutis, A., Sánchez, B., *et al.,* 2019. Exploring The Adoption of Precision Agricultural Technologies: A Cross Regional Study of EU Farmers. Land Use Policy 80 163-174.

Coble, K.H., Griffin, T.W., Ahearn, M., Ferrell, S., McFadden, J., Sonka, S., *et al.,* 2016. Advancing U.S. Agricultural Competitiveness with Big Data and Agricultural Economic Market Information, Analysis, and Research (No. 249847). Washington, DC, USA: Council on Food, Agricultural, and Resource Economics.

Coble, K.H., Mishra, A.K., Ferrell, S., and Griffin, T.W., 2018. Big Data in Agriculture: A Challenge for the Future. Applied Economic Perspectives and Policy 40(1) 79-96.

Cornell University, 2018. Institutional Analysis and Development of Innovation Systems for Digital Agriculture. https://www.digitalagriculture.cornell.edu. (last accessed December 2018).

Erickson, B., Lowenberg-DeBoer, J., and Bradford, J., 2017. 2017 Precision Agriculture Dealership Survey. http://agribusiness.purdue.edu/files/file/croplife-purdue-2017-precision-dealer-survey-report.pdf. (last accessed February 2019).

Fuglie, K.O., Heisey, P.W., King, J.L., Pray, C.E., Day-Rubenstein, K., Schimmelpfennig, D., *et al.*, 2011. Research Investments and Market Structure in the Food Processing, Agricultural Input, and Biofuel Industries Worldwide. Economic Research Report 130. Washington, DC: U.S. Department of Agriculture, Economic Research Service.

Gort, M. and Klepper, S., 1982. Time Paths in the Diffusion of Product Innovations. The Economic Journal 92(367) 630-653.

Griffin, T.W., Lowenberg-DeBoer, J., Lambert, D.M., Peone, J., Payne, T., and Daberkow, S.G., 2004. Adoption, Profitability, and Making Better Use of Precision Farming Data Staff paper #04-06. West Lafayette, IN, USA: Department of Agricultural Economics, Purdue University.

Kritikos, M., 2017. Precision Agriculture in Europe: Legal, Social, and Ethical Considerations. http://www.europarl.europa.eu/RegData/etudes/STUD/2017/603207/EPRS_STU(2017)603207_EN.pdf (last accessed February 2019).

Lowenberg-DeBoer, J., 2015. The Precision Agriculture Revolution. Foreign Affairs 94. Council on Foreign Relations, New York City, NY, USA.

MacDonald, J., 2017. Mergers and Competition in Seed and Agricultural Chemical Markets. Amber Waves (April). Washington, DC, USA: U.S. Department of Agriculture, Economic Research Service.

Meister Media Worldwide, 2014. PrecisionAg 2014 Media Planner. https://www.meistermedia.com/2014-media-planner-precisionag. (last accessed February 2019).

Meister Media Worldwide, 2018. PrecisionAg. https://www.meistermedia.com/precision-ag. (last accessed December 2018).

Schimmelpfennig, D., 2016. Farm Profits and Adoption of Precision Agriculture. Economic Information Bulletin 217. Washington, DC, USA: U.S. Department of Agriculture, Economic Research Service.

Financial and environmental performance of integrated precision farming systems

S.M. Pedersen[1*], M. Medici[2], T. Anken[3], G. Tohidloo[3], M.F. Pedersen[1], G. Carli[4], M. Canavari[2], Z. Tsiropoulos[5] and S. Fountas[5]
[1]*University of Copenhagen, Rolighedsvej 25, 1958 Frederiksberg, Denmark; marcus@ifro.ku.dk*
[2]*Alma Mater Studiorum, Università di Bologna, Viale Giuseppe Fanin 50, 40127 Bologna, Italy*
[3]Agroscope, Tänikon, 8356 Ettenhausen, Switzerland
[4]*Open University, Walton Hall, Milton Keynes, MK7 6AA, United Kingdom*
[5]*Agricultural University of Athens, Iera Odos 75, Athina 118 55, Greece*

Abstract

Smart farming technologies have a wide potential to improve agricultural performance, ranging from improved use of crop nutrients and plant protection products, increased crop quality and reduced overlaps to better production economy. In order to reduce the use of inputs and to diminish the adverse effects on the environment, more and more focus is put on site-specific application of fertilizers, pesticides and irrigation water. However, the costs of implementing these technologies are often quite high, compared to the benefits. Therefore, a comprehensive methodology approach that facilitates the understanding of investments, costs and benefits can provide an overview of the most feasible pathways for farmers to implement precision agriculture (PA) and may offer the chance to significantly enhance the level of adoption of the most suitable technologies. The objective of this study is to describe an overall integrated methodology approach to support cost-benefit analysis related to PA. A number of scenarios and their financial and environmental performance are presented either as single technologies or as a combination of different technologies. Key outputs of the analysis are net present value with selected environmental indicators such as change in fuel, pesticides and water use compared with conventional practices from other similar technologies. Findings from this study indicate that PA is mainly beneficial to large scale farms.

Keywords: cost, benefit, precision agriculture, smart farming, economy

Introduction

Farming employs a wide range of technologies and practices that require continual integration of new knowledge (Oreszczyn *et al.*, 2010). The implementation of new agricultural equipment is essential for farmers to remain competitive in their business. Variable Rate Application (VRA) and auto-steering have a wide potential to improve agricultural performance, ranging from improved use of crop nutrients, increased crop quality, reduced overlaps and better production economy.
Several studies have indicated that it is possible to reduce overlap by using RTK-GPS guidance technologies such as auto-steering. When applying inputs with centrifugal spreaders, fertilizer sprayers or even slurry distributors with boom sections it should be possible to reduce overlap by 3-7% from using auto-guidance systems (Batte and Ehsani, 2006); according to the same authors, saving of spraying materials increases proportionally to un-overlapped area. Estimates of the inputs saved by reducing overlap vary widely; in addition to those showing reduced overlap, also other studies which show no effect (Holpp *et al.*, 2013) can be found. To stay on the safe side, in this study a 3% reduction of agricultural input savings due to reduced overlap was assumed.
Auto section control on a nitrogen (N) liquid sprayer can reduce overlap when turning on headlands; this feature is particularly relevant when fields are irregularly shaped which otherwise would give an overlap with conventional systems. Some studies indicate that the reduced overlap is about 5%

DOI 10.3920/978-90-8686-888-9_103, © Wageningen Academic Publishers 2019

for using pesticides (Pedersen and Pedersen, 2018) but it might be smaller with N application when spreading dry granular products. Other studies indicate that the reduced overlap with auto-steering combined with reduction on the headland is about 5-10% (Lyngvig *et al.*, 2013 and Petersen *et al.*, 2006). The real potential is dependent upon the particular context, which may count various field shape and headland sizes relatively to the entire field size. In this study a 3% reduction of agricultural input due to the use of automatic section control is assumed.

Several approaches have been developed to make prescription maps for variable rate fertilizer application. Some technologies focus on prescription maps with a biomass index and NDVI (Normalized Difference Vegetation Index). These maps can be generated from satellite or drone images and tractor-mounted sensors like the Yara N sensor. Others are based on farmer's own observations, previous years yield maps and/or targeted soil samples. A Danish study of potential increase in yield from variable rate N application with N-sensors in winter wheat, showed modest results with yield increase close to zero (Berntsen *et al.*, 2006). A similar study from Australia in wheat found an average yield increase of 0.8% with variable rate treatment (compared with uniform application) over two years (Mayfield and Trengove, 2009). Bourgain and Llorens (2006) experimented with variable rate applications of N obtaining a 11.1% saving of the agricultural input, and Casa *et al.* (2011) reported a 22% reduction in N; both of these studies were based on cereal crops and uniform and variable rate applications were based on practically the same yields.

The objective of this study is to describe an overall integrated methodology approach to support cost-benefit analysis related to PA. The methodology will embrace 2 sets of evaluations referring to (1) financial performance and (2) environmental impact.

The overall approach covers a broad number of PA technologies that are integrated in a web-app designed in the ICT-Agri project Precision Agriculture Methodologies for Cost Benefit Analysis (PAMCOBA) (http://ict-agri.eu/node/36322). The aim here is to guide farmers over their decisions to invest in selected precision agricultural technologies on their farm depending on crop rotation and farm size. In this study, an example is given for variable rate N application in a traditional cereal crop rotation.

Methodology

The methodology embraces two sets of evaluations referred to as: (1) financial performance; and (2) environmental impact. The financial analysis and environmental performance study is based on the estimation of differential cash flows from selected PA technologies with description of life time, input costs and expected benefits in relation to location and in-field variability. A number of scenarios and their financial net-benefits and environmental performance are presented either as single technologies or as a combination of different technologies. The selected examples presented here focus on variable rate fertilizer application combined with auto-steering and section control on the fertilizer spreader. The environmental performance is based on findings from existing literature on fertilizer application as well as estimates from Denmark on potential nitrate leaching from variable rate N application.

To compare the modelled site-specific fertilization with uniform application of nutrients, a rate of (180 kg N, 25 kg P, 95 kg K per ha) with an average price of 2€/kg fertilizer (incl. labour distribution costs) are assumed. In assessing economic life, a cut off period of 6 years with null residual value and a discount rate equal to 1% (risk free) were assumed. In formulating these assumptions, the risk burden has switched from the discount rate to the economic life and to the residual value of the investment for two reasons: first, the cost of purchasing equipment is relatively low, being affordable by the majority of farms without need of external investors, and secondly, interest rates are very low at the present time. That should not prevent one from paying attention to technological risk, but this is already taken into the economic life, which normally extends to 10-12 years, here accounted to 6 years, with no residual value In addition, the sensitivity relating to discount rate was tested, comparing the change in discount rate with output represented by the net present value (NPV). In evaluating fuel consumption, in this

study a fuel price of 1.50€/l (diesel), and an average fuel consumption for the real-time kinematic (RTK)-GPS equipped tractor of 79 l/ha which is consistent with the median fuel consumption for farm activities measured by Lorencowicz and Uziak (2009). These assumptions are consistent with a conventional cropping system in arable farming.

In this study, three systems characterized by increasing effectiveness in fertilization activity in a common cereal crop rotation are modelled (Table 1). In particular, System 1 consists of a variable rate fertilization spreader, i.e. a solid or liquid spreader, including prescription software, proper flow sensors, and a base GNSS system characterized by a relatively low accuracy (about 60-100 cm). In addition, System 2 includes an integrated auto-steering system that is comparable with a precision RTK-GPS guidance technology with a high accuracy (2-3 cm). System 3 finally includes the features of the first two systems plus automatic section control.

Each of the PA technologies reported in Table 1 shows a cost structure in which the following cost items are represented: a purchase price (€), and possible annual fees due to external service support (€/year); common market values for the systems considered are shown in Table 2. The investment cost is the sum of the purchase prices of each technology considered within each system and does not include annual fees, which affect the differential yearly revenues in the net present value calculation where NPV is the sum of present values of costs and benefit at any time in the future at a given discount rate. The economic values shown in Table 2 are thought to be appropriate for a 50-ha farm; accordingly, when farm size is increased, investment cost should be properly re-adapted. To model this aspect of scale economies it was decided to adopt the '0.6 rule' in the economics literature (Tribe and Alpine, 1986): this rule has its origins in the relationship between the increase in equipment cost (i.e. investment) and the increase in capacity (i.e. farm size) given by $I_i/I_{i+1} = (F_i/F_{i+1})^{0.6}$. Accordingly, the investment cost for a 100 ha farm (i.e. I_{i+1}) is calculated considering the former value of the farm size, i.e. 50 ha, and the relating investment cost (Table 2).

To model the economic benefits in this study, two scenarios were considered: Scenario A, characterized by limited benefits arising from the adoption of the selected PA technologies, and Scenario B, slightly more optimistic. Table 3 shows the base cost reductions, in terms of input used, considered for both scenarios and for each system. It was decided to stay as much as possible on the safe side by adopting reductions that range from 1.5% (Scenario A) to 3.0% (Scenario B) even though some contributions in scientific literature suggest even higher saving rates.

Table 1. Site-specific equipment.

System 1	VR fertilizer spreader
System 2	VR fertilizer spreader, RTK-GPS guidance technology
System 3	VR fertilizer spreader, RTK-GPS guidance technology and auto section control

Table 2. Cost structure of the site-specific equipment.

	Purchase prices (€)				Investment (€)	Annual fees (€/year)
	VR fertilizer spreader	GPS technology	Guidance technology	Auto section control		
System 1	7,300	2,086	-	-	9,386	0
System 2	7,300	6,417	4,078	-	17,795	800
System 3	7,300	6,417	4,078	2,500	20,295	800

Table 3. Yearly benefit structure of the systems.

Agricultural input savings (fertilizer)

	Base cost reduction (%)		w/ RTK-GPS, auto-steering	w/ Auto section control	Total saving (€/ha)	
	Scenario A	Scenario B			Scenario A	Scenario B
System 1	-1.5%	-3.0%	-	-	9.0	18.0
System 2	-1.5%	-3.0%	-3.0%	-	27.0	36.0
System 3	-1.5%	-3.0%	-3.0%	-3.0%	45.0	54.0

Fuel savings

	Fuel use (%)	Total saving (€/ha)
System 1	-0.0%	-
System 2	-4.0%	2.37
System 3	-5.0%	2.96

Yield benefits

		Differential revenues (€/ha)		
	Yield increase (%)	Year 1	Year 2	Year 3-6
System 1	+1.0%	25.0	25.3	25.5
System 2	+2.0%	50.0	51.0	52.0
System 3	+2.0%	50.0	51.0	52.0

With regard to fuel savings, saving rates equal to 4.0 and 5.0% respectively for System 2 and System 3 are assumed, with no fuel saving considered for System 1 because GNSS system is regarded as less accurate than RTK-systems. Increased yield benefits are considered to increase for three years, from year 1 to year 3, before becoming steady in years 4, 5 and 6. In addition, yield increase of 1% for the base VR fertilization system (System 1) and 2% increase for both Systems 2 and 3 are assumed. These values are regarded as fairly conservative, considering that similar systems that are comparable to system 1 gave yield benefits of 1.4-1.5% (Bourgain and Llorens, 2009).

Results

Figure 1 shows the results of the financial analysis (NPV) of the three systems for farm sizes ranging from 50 ha to 500 ha with respect to scenarios A and B.

When only considering Scenario A, it appears that the base system (System 1) is clearly unprofitable for every farm size considered, reaching the worst performance between 350 and 400 ha and tending to a loss of about 12.5 k€. The lowest costs are found at 50 ha (-6.7 k€). System 2 is profitable only for relatively large farm holdings (about 450 ha), with a maximum profit of 5.9 k€ at 500 ha. System 3 shows a positive NPV for a farm size greater than 200 ha, ranging from 2.1 k€ to 36.6 k€ at 500 ha which is the largest farm size considered.

In Scenario B with more favorable boundary conditions, it appears that all systems are profitable at 250 ha. System 1 shows a breakeven point at 250 (0.3 k€) and still unprofitable up to 200 ha while

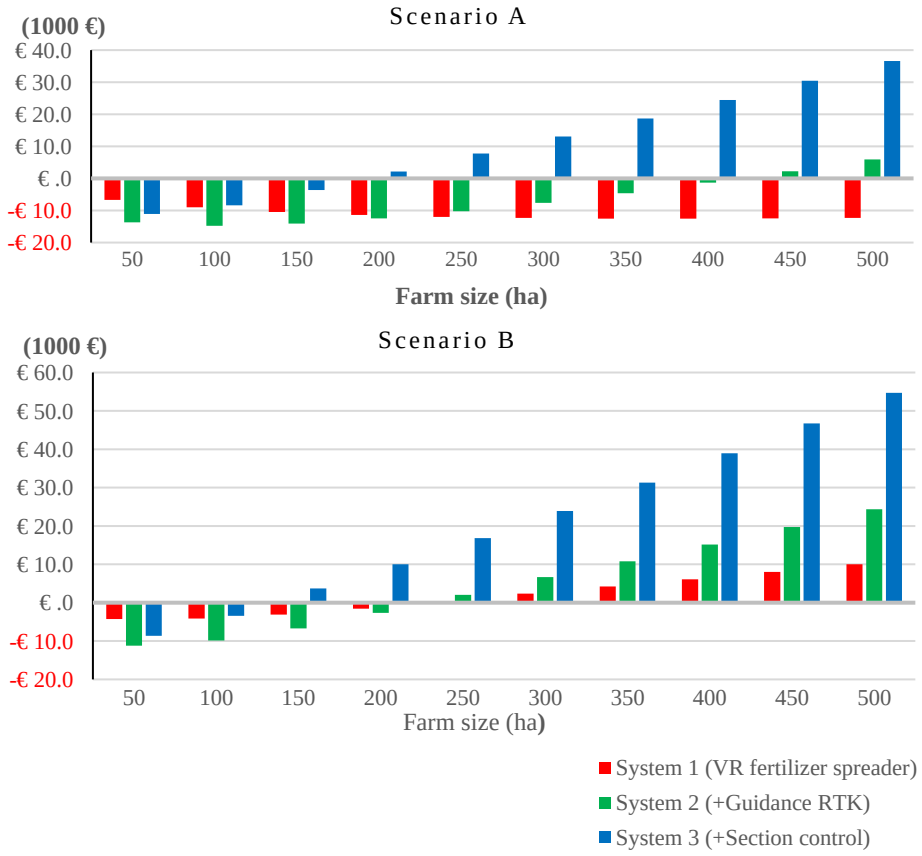


Figure 1. NPV of investments on systems 1, 2 and 3 for 2 scenarios (A and B).

System 2 is profitable for farm size greater than 250 ha. System 3 also provide a positive NPV at 150 ha (3.7 k€) and increases the expected economic benefit almost linearly to 500 ha (55.0 k€).
In general, the most advanced system composed of state-of-the-art technology (i.e. System 3) shows higher scale advantages that vastly outweigh the relatively high initial investment. In contrast, less innovative systems that can execute site-specific treatments with limited accuracy are at the risk of becoming unprofitable in the case of unfavorable environmental contingencies, for almost all farm sizes. The largest item in the cost-benefit analysis is the saving of agricultural inputs, that is, in this study, nitrogen. The average kilograms of nitrogen annually saved as a result of performing a site-specific N-fertilization vary in the range 2.7-5.4 kg/ha for scenario A and B, assuming an average application of 180 kg/ha. Direct financial benefits both in terms of operating cost savings and in reduced application of physical inputs become relevant for large farms, for instance on a 100 ha farm the nitrogen savings are in the range 270-540 kg.
Finally, each of the systems considered saved on fuel with resulting environmental benefits in terms of reduced air emissions. Taking as a reference the emission factor of diesel engines that is 2.64 kg CO_2/l, the amount of CO_2 yearly saved was estimated to be equal to 4.2 and 5.2 kg/ha respectively for Systems 2 and 3. In this study, the impact on nitrate leaching from reduced N application was not considered. It is however likely that by limiting overlap of N application it is possible to reduce overall N leaching at the root zone with selected precision farming technologies compared with

uniform application. Estimates by Knudsen (2017) indicate a reduced nitrate leaching of 3-4% from VRA on Danish farms with only crop production and slightly higher on farms with mixed crops and animals. A simulation with the Danish farm model FASSET in cereals indicate a reduction of nitrate leaching between 0-3% from VRA on fields with moderate soil type variation (Pedersen, 2003). Depending on the location and crop varieties, it may also be possible to gain additional benefits from better grain quality such a higher average protein content, which again may provide an additional price premium. From the sensitivity analysis it was found that the net present value appears to be sensitive to the variation of interest rate. For instance for System 2, Scenario B, NPV decreases on average between 14.8 and 19.5%, when the interest rate passes from 0.5 to 1.0%, then to 1.5%, and finally to 2.0%.

Conclusions

Findings from this study indicate that precision agriculture with targeted fertilization is mainly beneficial to large scale farms. The example provided shows that by combining different tools has a greater benefit than by applying a single tool. The calculation of two scenarios showed, that a slight increase of the efficiency of only 1.5% has an important impact on the profitability of PA technologies. So only a small increase is sufficient to amortize supplementary investments. Apart from direct profits, the study did not account for the fact that these techniques allow a much better management, as much more data over the years are available, allowing the farmer to better understand and to manage his farm. It is evident, that larger farms have the advantage of amortizing the costs over a larger area. Such techniques are mainly interesting for small-scale farms if they can improve their production systems in conditions of irregularly shaped fields (section control) or with heterogeneous soil conditions.

Acknowledgement

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 618123 ERA-NET – ICT-Agri in the project PAMCOBA.

References

Batte M.T., Ehsani M.R., 2006, The economics of precision guidance with auto-boom control for farmer-owned agricultural sprayers, Computers and Electronics in Agriculture, 53 (1) 28-44.

Berntsen J., Thomsen A., Schelde K., Hansen O. M., Knudsen L., Broge N., *et al.*, 2006, Algorithms for sensor-based redistribution of nitrogen fertilizer in winter wheat, Precision Agriculture, 7 (2) 65-83.

Bourgain O., Llorens J.M., 2009, Methodology to estimate levels of profitability of precision agriculture: simulation for crop systems in Haute-Normandie, EFITA Conference 2009, Esitpa, Agricultural Engineering School, Mont Saint Aignan, France (available at: http://www.efita.net/apps/accesbase/bindocload.asp?d=6490&t=0&identobj=YJkgtriS&uid=57305290&sid=57&idk=1, (last access January 2019).

Casa R., Cavalieri A., lo Cascio B., 2011, Nitrogen fertilisation management in precision agriculture: A preliminary application example on maize, Italian Journal of Agronomy, 6 (1) 23-27.

Holpp, M., Kroulik, M., Kviz, Z., Anken, T., Sauter, M. & Hensel, O., 2013, Large-scale field evaluation of driving performance and ergonomic effects of satellite-based guidance systems, Biosystems Engineering,116 (2) 190-197.

Knudsen L., 2017. Skitse til pilotprojekt om Teknologi og viden som virkemiddel til reduktion af kvælstofudvaskning, (Draft pilot project – Technology and knowledge as an instrument for reduced nitrate leaching) SEGES, Arhus, Denmark.

Lorencowicz, E., Uziak, J., 2009. Fuel Consumption in Family Farms, TEKA (Komisji Motoryzacji I Energetyki Rolnictwa) - OL PAN, 9 164-171.

Lyngvig H. S., Hørfarter R and Knudsen L., 2013. FarmTest Centrifugalspredere med sektionskontrol, Videncentret for Landbrug (FarmTest Centrifugal spreaders with section control) (SEGES, Arhus, Denmark). https://www.landbrugsinfo.dk/Tvaerfaglige-emner/FarmTest/Maskiner-og-planteavl/Sider/FarmTest-centrifugalspredere-med-sektionskontrol_pl_13_1528.aspx (Last accessed February 2019)

Mayfield A.H., Trengove S.P., 2009 Grain yield and protein responses in wheat using the N-Sensor for variable rate N application, Crop and Pasture Science, 60 (9) 818-823.

Oreszczyn, S., Lane, A., Carr, S., 2010. The role of networks of practice and webs of influencers on farmers' engagement with and learning about agricultural innovations. Journal of Rural Studies, 26 (4) 404-417.

Pedersen S.M., 2003: Precision farming – technology assessment of site-specific input application in cereals. Department of Manufacturing Engineering and Management, Technical University of Denmark.

Pedersen, M. F., Pedersen, S. M., 2018 Erhvervsøkonomiske gevinster ved anvendelse af præcisionslandbrug (Production economic net-benefits from precision farming) 49 s., IFRO Udredning, Department of Food and Resource Economics, University of Copenhagen, Nr. 2018/02

Petersen, H H, Hansen J. P. og Øllgaard T., 2006. Økonomiske og miljømæssige fordele ved autostyring, NordVest Agro. (Economic and environmental benefits from Auto steering) https://www.landbrugsinfo.dk/planteavl/praecisionsjordbrug-og-gis/sider/1537_bilag3.pdf

Tribe, M.A., Alpine, R.L.W., 1986. Scale economies and the 0.6 rule, Engineering Costs and Production Economics, 10 (1) 271-278.

Section 15
Adoption of precision agriculture

How quickly do farmers adopt technology? A duration analysis

T.W. Griffin* and E.A. Yeager
Department of Agricultural Economics, Kansas State University, Manhattan, Kansas, USA; twgriffin@ksu.edu

Abstract

Precision technologies have been available at the farm level for decades. Some technologies have been readily adopted while others lagged. Analysis of 526 Kansas farms provided insights regarding duration of adoption. The lag, in years, between technologies becoming commercially available and adopted were evaluated using non-parametric duration analysis. Duration for embodied-knowledge technologies were statistically sooner than for information-intensive technologies, indicating farmers adopt automated guidance 'quicker' than yield monitors. Duration was indirectly (directly) proportional to commercialization date of embodied-knowledge (information-intensive) technology. Results are useful to farmers considering adoption, retailers targeting customers, and manufacturers managing supply chains.

Keywords: survival analysis, embodied knowledge, information intensive, commercialization, time, hazard rate

Introduction

Farm-level adoption of agricultural technologies remain a persistent issue among researchers and practitioners in terms of understanding the rate at which technologies are adopted, order in which adoption occurs, and the benefits the technology provides. Many precision technologies have been available since the early 1990s, yet adoption has not plateaued even after nearly 30 years. Adoption of technology has been associated with perceived profitability from farm-level utilization. The benefits of precision agriculture have been said to be 'site-specific'. Given that the economics of technology are a function of the specific grower's fields and their management ability, profitability assessments of specific technologies have been elusive.

The diffusion of technology has been segregated into simpler and more complex versions (Coombs *et al.*, 1987; Davies, 1979; Miller *et al.*, 2019). Miller *et al.* (2019) described technology as two distinct groups, i.e. embodied knowledge and information-intensive, and how adoption rates differed. Embodied-knowledge technologies have been more readily adopted than information-intensive technologies (Griffin and Yeager, 2018). This study builds upon previous work by applying duration analyses (Dinterman and Katchova, 2018) to Kansas farms.

An indirect way that precision agriculture has been found to affect profitability is its ability to substitute information and knowledge for fertilizer, seeds and chemicals given soil and other conditions. Several researchers thus far have examined this savings from an environmental stewardship perspective and the reduction of purchased inputs leading to better sustainability of resources (Bongiovanni and Lowenberg-DeBoer, 2004). Schimmelpfennig and Ebel (2016) examined the distortion between adoption of precision agricultural technologies given expected lower input costs. Their findings indicated that differences in the size of operation, education of operator and type of farm played significant roles in the adoption of technology. There was also an inconsistency in the savings as variable rate technology in some instances could result in increased input usage. The first step in examining profitability of technology is to ascertain the adoption trends. The longest running survey of precision agricultural technology adoption focuses on agricultural service providers rather than within the farm gate (Erickson *et al.*, 2017). Their results of agricultural retail services are consistent with farm level estimates.

DOI 10.3920/978-90-8686-888-9_104, © Wageningen Academic Publishers 2019

The decision to adopt, i.e. technology choice, is an inherently dynamic process. This decision is based on past decisions as well as future expectations. Duration analysis is one means to examine this dynamic nature of the decision and specifically focus on the time-to-adopt decision (Burton *et al.*, 2003; An and Butler, 2012).

An early study by Fuglie and Kascak (2001) examined the adoption of conservation tillage, soil nutrient testing and integrated pest management. They found that farm size, farmer education and land quality could result in adoption lags of up to 20 years. Due to self-reported data limitations, they were somewhat constrained in their analysis, but they did make use of duration analysis methodology. Dadi *et al.* (2004) used duration analysis to examine the adoption of fertilizer and herbicides by smallholder farmers in Ethiopia. They found that the most important determinant of the time farmers waited to adopt was economic incentives. It was likely that farmers were unwilling to adopt until evidence indicated that it was profitable or rewarding to do so (Griffin *et al.*, 2018). Alcon *et al.* (2011) examined the adoption of drip irrigation technology by farmers in Spain using duration analysis. The authors found that educational factors, technological trials, availability of credit, price and information networks were among the most important factors influencing adoption timing.

Burton *et al.* (2003) were one of the first to apply duration analyses to agricultural technology adoption. They provided detailed reviews of several theoretical studies evaluating time-to-adopt of agricultural technologies. One key result illustrated that duration analysis was a proper means to analyze time-to-adopt due to the dynamic nature of the decision and potential impactors. D'Emden *et al.* (2006) examined no-tillage adoption decisions of Australian farmers. They found cost of inputs, specifically herbicides, to be important factors in adoption decisions. Their study was complicated by tradeoffs regarding possible herbicide resistance following no-till practices.

As a first step into understanding farmers' perceived benefits of precision agriculture, the time-to-adopt was evaluated. The overall objective of this study was to determine if time-to-adopt after commercialization was similar between individual technologies and across groups of technology.

Data

The Kansas Farm Management Association (KFMA) has archives of farm-level agronomic and financial data since 1973. In 2015, KFMA economists began collecting and annually updating technology records (Griffin *et al.*, 2017). By September 2018, 608 farms reported having either 'used' or 'never used' at least one of six technologies. Of the 608 farms, 523 (86%) adopted one or more technologies. Technologies included global navigation satellite system (GNSS) enabled yield monitors (GNSSYM), variable rate fertilizer (VRF), precision soil sampling (PSS), lightbar (LB), automated guidance (AGS) and automated section control (ASC).

Duration is the spell, or length of time, after being able to adopt that adoption occurs. Lags are measured in years between the farm adopting technology and when the farm could adopt the technology. Individual farms could adopt technology after the farm began operating and the technology became commercially available (Figure 1).

The year farm operations began was obtained from KFMA Operator Database. The KFMA Operator Database contains 1,776 unique farms including birth year, year began farming and number of dependents. Joining the 1,776 KFMA Operator Database farms and 608 farms with technology yielded 526 farms common to both datasets. Duration was measured for each technology, *j*, on 526 farms, *k* (Equation 1). Farms may have different start dates because some operations began after technology became available.

$$duration = t_{jk} - \max(t_{CA\,j}, t_{BF\,k}) \quad (1)$$

where *duration* is time-to-adopt in years, t_{jk} is date farm *k* adopts technology *j*, t_{CAj} date technology *j* became commercially available, and t_{BFk} the date farm *k* began farming.

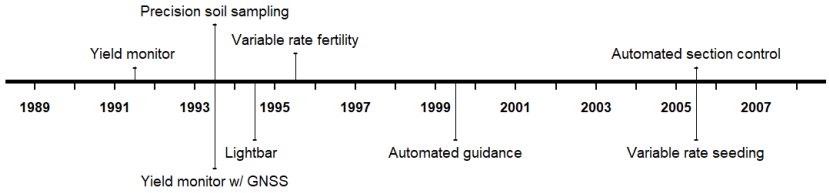


Figure 1. Date technology became commercially available.

Methods

Duration analysis is concerned with time-to-event data (Greene, 2017). The timing when subjects transition from one state of the world to another is of interest (An and Butler, 2012). In this study, the event failure is when farms adopt technology. The specific question addressed by this study is 'given that a farm has not adopted some technology by time *t*, what is the chance they will adopt after time *t*?' (An and Butler, 2012, paraphrased p. 496).

Each KFMA farm was tracked over time until 2018 or the event failure, i.e. adoption of technology. Farms not adopting were censored. Some farms may cease operations, discontinue KFMA membership, or never adopt technology; any of these actions would cause the farm to be censored. Each farm was evaluated for the duration that it remained in the sample up until the adoption event, i.e. survival. As adoption events occur, the survival rate decreases, and farms fall out of the sample. The hazard rate is the probability of event failure at time *t,* i.e. the probability of a farm adopting given the time period.

The dependent variable, *duration*, is the number of years between being able to adopt and adopting technology; and is assumed to have a continuous probability distribution, f(t) (Equation 2). Theoretically, the dependent variable is non-negative but due to recall bias and uncertainty of exact commercialization dates, some farms reported adopting before the first date possible thereby producing negative duration in this sample.

A non-parametric model was estimated to determine the general shape of the hazard (and survival) functions. The probability density function is:

$$\mathrm{f(t)} = \frac{\mathrm{dF(t)}}{\mathrm{dt}} \tag{2}$$

The probability that the duration, in years, will be less than t is a cumulative distribution function expressed as:

$$\mathrm{F(t)} = \mathrm{Pr(T \leq t)} = \int_0^{\mathrm{t}} \mathrm{f(s)ds} \tag{3}$$

where the random variable T has some duration less than t. Equation 2 can be rewritten as a survivor function that is the probability that duration equals or exceeds t (Equation 4) (An and Butler, 2012).

$$\mathrm{S(t)} = \mathrm{Pr(T > t)} = 1 - \mathrm{F(t)} \tag{4}$$

The hazard rate is the probability that the duration will end after time t, given that it has lasted until time t (Equation 5).

$$\lambda(t) = \frac{f(t)}{S(t)} \tag{5}$$

The Kaplan-Meier (KM) non-parametric estimator corrects for censored data in a distribution (Borgan, 2005) and therefore has a decreasing step function with a jump at each discrete event time (Colosimo *et al.*, 2002). Statistical tests were conducted on estimated curves to determine if duration differed between technologies. Duration results can be interpreted to reveal if technologies have been adopted at faster rates than others.

Results

The duration curves for individual technologies and the two main categories of precision agriculture were compared. When the duration curve was to the left of another curve, then it reached failure, i.e. adoption, sooner. Note that duration curves originate from survival probability equal to 1.0 on the y-axis. If the curves do not diverge, then the adoption path was not statistically significantly different. All three embodied knowledge technologies (AGS, LB, ASC) were compared as a collective group to all three information-intensive technologies (GNSSYM, PSS, VRF). Duration curves were statistically significantly (P<0.0001) different from each other with embodied-knowledge technologies adopted sooner after commercialization than their information-intensive counterparts (Figure 2). Roughly one-third of KFMA farms remain likely to adopt embodied-knowledge technology (per the end of the curve where it plateaus). Almost half of KFMA farm subjects remain as potentially adopting information intensive technology.

In addition to comparing groups of technology, duration curves were estimated for individual technologies within each category. The duration curves for the three embodied-knowledge technologies were compared against each other. Results indicated that curves were statistically different (P<0.0001) with automated section control (ASC) being adopted relatively sooner than automated guidance (AGS) or lightbar (LB) (Figure 3). Automated guidance was adopted in a relatively shorter amount of time than lightbar guidance. Automated section control is a relatively newer technology, i.e. with a more recent commercialization date; and has approximately half of farmers remaining as nonadopters. Lightbar has been commercially available longer than AGS or ASC; and has 40% of farms not adopting the technology. Automated guidance has the least number of subjects at risk for adoption at nearly one-third. Duration of adoption was indirectly proportional to the respective commercialization dates.

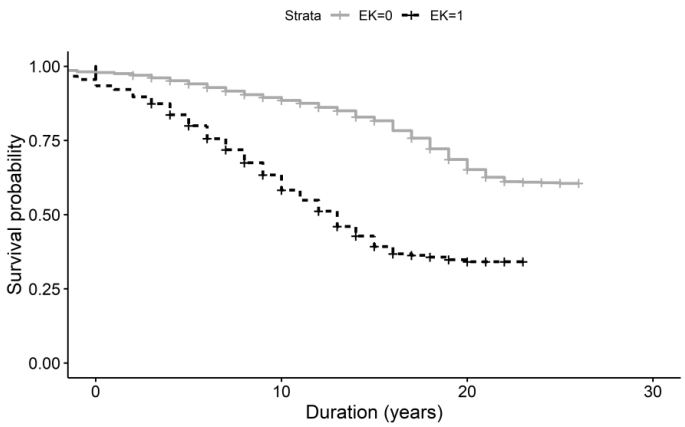


Figure 2. Duration of embodied-knowledge (EK=1) and information-intensive (EK=0) technology adoption.

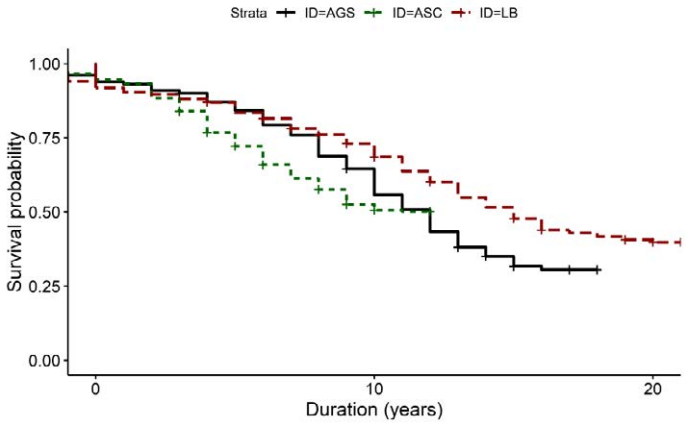


Figure 3. Duration of embodied-knowledge technologies including automated guidance (AGS), automated section control (ASC) and lightbar (LB).

The duration curves for three information-intensive technologies were compared against each other. Results indicated that curves were statistically different (P<0.0001) with GNSS-enabled yield monitor (GNSSYM) being adopted before precision soil sampling (PSS) or variable rate fertility (VRF) (Figure 4). Unlike embodied-knowledge technologies, duration of information-intensive technology adoption was directly proportional to commercialization dates. More than half of farms are considered potential adopters of any information-intensive technologies.

Discussion

Duration curves of embodied-knowledge and information-intensive technologies were expected to substantially differ with the former being adopted sooner than the latter after commercialization. Results confirmed duration statistically differed across these two broad categories. It was expected

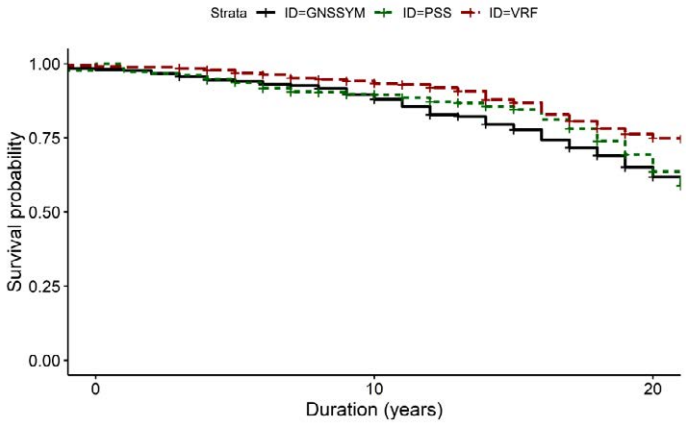


Figure 4. Duration of information-intensive technology including GNSS-enabled yield monitor (GNSSYM), precision soil sampling (PSS) and variable rate fertility (VRF).

that automated guidance and section control were adopted at much higher rates than yield monitors due to differences in human capital costs to use data. These are consistent with nearly all precision agriculture adoption studies.

Duration curves of similar technologies grouped within either embodied-knowledge or information intensive were not expected to substantially differ. It was somewhat unexpected that duration statistically differed among technologies within these two broad categories. Specifically, automated section control was adopted sooner than automated guidance, and automated guidance was adopted sooner than lightbar. Results may partially be explained by commodity prices and farm policies being favorable to expansion or shifts towards utilization of automated section control. Retailer marketing may also have played a role. For information intensive technologies, GNSS-enabled yield monitors were adopted sooner than precision soil sampling or variable rate fertility. This may be due to ease of use or the perceived benefit GNSS-enabled yield monitors provide. Yield monitor adoption may also have been due to acquiring combine harvesters already equipped with the technology.

These results may be used to evaluate whether lightbar will continue to be considered an embodied-knowledge technology since the technology provides operators with visual aid to manually steer the equipment without automating the process. Arguments could be made that even though substantial technology was embodied into the lightbar, it was analogous to information-intensive given that the user must make use of the information albeit the action was reactive rather than proactive.

Duration was indirectly proportional to commercialization dates of embodied-knowledge and directly proportional to dates of information-intensive technology. This was somewhat surprising and warrants further investigation. Future work also includes estimating semi-parametric models including explanatory variables. Model specification will include farm-specific financial performance and agronomic metrics plus exogenous variables such as commodity prices. These analyses should provide insights into farm characteristics useful in predicting when specific farms might adopt technology. Uncertainty exists with respect to the exact dates that technologies became commercially available. Continued effort is being applied to finding historic documents that may push commercialization to earlier dates.

Conclusions

Automated technologies including guidance and section control were adopted sooner after becoming commercially available than more data intensive technologies such as yield monitors and grid soil sampling.

Acknowledgements

The authors appreciate KFMA economists and member farms for supporting our research program. Specifically, we thank Kevin Herbel for making this research possible and Koren Roland for data processing. We appreciate informal reviews from our graduate students including Luke Minnix, Emily Carls, and Jared Cullop.

References

Alcon, F., Dolores de Miguel, M., and Burton, M., 2011. Duration Analysis of Adoption of Drip Irrigation Technology in Southeastern Spain. Technological Forecasting and Social Change 78(6) 991-1001.

An, H. and Butler, L.J., 2012. A Discrete-Time Duration Analysis of Technology Disadoption: The Case of rbST in California. Canadian Journal of Agricultural Economics 60 495-515

Bongiovanni, R. and Lowenberg-DeBoer, J., 2004. Precision Agriculture and Sustainability. Precision Agriculture 5(4) 359-387.

Borgan, Ørnulf, 2005. Kaplan–Meier Estimator. Encyclopedia of Biostatistics. Hoboken, NJ, USA: Wiley. 6100 pp.

Burton, M., Rigby, D., and Young, T., 2003. Modelling the Adoption of Organic Horticultural Technology in the UK using Duration Analysis. Australian Journal of Agricultural and Resource Economics 47(1) 29-54.
Colosimo, E., Ferreira, F., Oliveira, M. and Sousa, C., 2002. Empirical comparisons between Kaplan-Meier and Nelson-Aalen survival function estimators. Journal of Statistical Computation and Simulation 72(4) 299-308.
Coombs, R., Saviotti, P., Walsh, V., 1987. Economics and Technological Change. London, UK: Palgrave, 296 pp. https://doi.org/10.1007/978-1-349-18683-9
Dadi, L., Burton, M., and Ozanne, A., 2004. Duration Analysis of Technological Adoption in Ethiopian Agriculture. Journal of Agricultural Economics 55(3) 613-631.
Davies, S.W., 1979. Inter-firm diffusion of process innovations, European Economic Review 12(4) 299-317 https://doi.org/10.1016/0014-2921(79)90023-0.
D'Emden, F.H., Llewellyn, R.S., and Burton, M.P., 2006. Adoption of Conservation Tillage in Australian Cropping Regions: An Application of Duration Analysis. Technological Forecasting and Social Change, 73(6) 630-647.
Dinterman, R., and Katchova, A., 2018. Survival Analysis of Farm Bankruptcy Filings. 2018 Annual Meeting, August 5-7, Washington, D.C. 274117, Agricultural and Applied Economics Association. https://ideas.repec.org/p/ags/aaea18/274117.html (last accessed 24/1/19)
Erickson, B., Lowenberg-Deboer, J. and Bradford, J., 2017. 2017 Precision Agriculture Dealership Survey. Departments of Agricultural Economics and Agronomy, Purdue University, IN, USA. http://agribusiness.purdue.edu/files/file/croplife-purdue-2017-precision-dealer-survey-report.pdf (last accessed 24/1/19)
Fuglie, K., and Kascak, C., 2001. Adoption and Diffusion of Natural Resource Conserving Agricultural Technology. Review of Agricultural Economics, 23(2) 386-403.
Greene, W.H., 2017. Econometric Analysis (8th Ed). New York, NY, USA: Pearson 1176 pp.
Griffin, T.W. and Yeager, E.A., 2018. Adoption of Precision Agriculture Technology: A Duration Analysis. In Proceedings of 14th International Conference on Precision Agriculture. Monticello, IL: International Society of Precision Agriculture. https://www.ispag.org/proceedings/?action=abstract&id=5271
Griffin, T.W., Miller, N.J., Bergtold, J., Shanoyan, A., Sharda, A., and Ciampitti, I.A., 2017. Farm's Sequence of Adoption of Information-Intensive Precision Agricultural Technology. Applied Engineering in Agriculture, 33(4) 521-527.
Griffin, T.W., Shockley, J.M., and Mark, T.B., 2018. Economics of Precision Farming. In: D.K. Shannon, D.E. Clay, N.R. Kitchen, (Eds.), Precision Agriculture Basics, Madison, WI, USA: ASA, CSSA and SSSA. pp. 221-230.
Miller, N.J., Griffin, T.W., Ciampitti, I., and Sharda, A., 2019. Farm Adoption of Embodied Knowledge and Information Intensive Precision Agriculture Technology Bundles. Precision Agriculture, 20 (2) https://doi.org/10.1007/s11119-018-9611-4
Schimmelpfennig, D., and Ebel, R., 2016. Sequential Adoption and Cost Savings from Precision Agriculture. Journal of Agricultural and Resource Economics 41(1) 97-115

A collective framework to assess the adoption of precision agriculture in France: description and preliminary results after two years

N. Lachia*, L. Pichon and B. Tisseyre
ITAP, Univ Montpellier, Montpellier SupAgro, Irstea, 2 Place P. Viala, 34060 Montpellier, France; nina.lachia@supagro.fr

Abstract

The article presents the results of a project that led to a collective organisation designed to assess and follow the level of adoption of precision agriculture in France. This organisation, called the French Observation Centre of Digital Agriculture adoption (FrOCDA) allowed the design and the set-up of an original and exhaustive study method mainly based on interviews of the main players in precision farming in France. The article presents the results of three studies. The first study dealt with remote sensing. It showed that the adoption of this technology is relatively high in field crops (about 1,000,000 ha) and more marginal in viticulture. The second dealt with soil mapping. The mapping remains marginal since it concerns 130,000 ha mainly in viticulture. Finally, the use of agricultural robotics in France will be presented.

Keywords: adoption, consortium, national survey, technology

Introduction

Reliable information about the adoption of precision agriculture (PA) is important for different stakeholders: (1) for service provider companies, this information makes it possible to have an overview of the current uptake in order to design services that correspond to current needs; (2) for academics and farmer's organisations, it is a prerequisite to define the most appropriate initial and continuing training courses to support farmers and advisors in adoption and use of PA; and (3) for regional or national institutions, it helps to define strategies and policies in order to support the profession (Tey and Brindal, 2012).

Many studies have examined the current uptake of PA and generally find that except for GNSS guidance and related technologies like sprayer boom control and seeder row shutoffs, PA adoption is generally low. Available studies which provide reliable estimates of PA adoption by farmers based on random sample methods were mostly focused on US and Australian farmers (Lewellyn and Ousman, 2014; USDA ARMS, 2018). Within Europe, uptake rates are less well explored and available studies are focused on specific countries (e.g. DEFRA. 2013, Paustian and Theuvsen, 2017). In Europe, the most recent study on PA adoption based on random sample of more than 2000 farmers was performed across four countries: Germany, Greece, Finland and Denmark (Lawson *et al.*, 2011). This short state of the art highlights several issues for the French industry. Firstly, to the authors' knowledge, no study has been carried out in France, one of the main agricultural countries in Europe. Secondly, most of these studies remain punctual over time. Indeed, with the exception of the Services Dealership Survey conducted for more than 15 years in the United States (Erickson *et al.*, 2017), there is no organization that regularly monitors PA uptake with a methodology that allows results to be compared year after year. Finally, most studies are based on a targeted sample of farmers and rely on extrapolation from a small number of surveys to the regional or national scale, causing potential bias.

The objective of this article is to present an original organisation that has been set up in France to tackle these issues. The aim of the organisation is to provide exhaustive and regular figures on the adoption of digital tools and PA services in France for the main crops (arable, viticulture, fruit

DOI 10.3920/978-90-8686-888-9_105, © Wageningen Academic Publishers 2019

production). This organisation was initiated in 2016 and has produced 7 studies since its launch. It focuses on digital agriculture with a framework more general than precision agriculture, including for example farm management systems or mobile applications. This paper will however, focus only on PA uptake in France. It will present, in a first step, the organization and the partnership that has been set up to provide a sustainable framework over time. In the second step, the method used to carry out an exhaustive monitoring of PA uptake will be detailed. Finally, the paper will present some of the results obtained on the adoption of PA. These results focus on the uptake of remote sensing, soil mapping and robotics in France.

Description of the French Observation Centre of Precision Agriculture adoption

Origin and genesis of the project

Official French organisations providing yearly reports about agriculture figures do not include PA in their survey. Consequently, until 2016, it was impossible to find any reliable figures on PA uptake in France. Digital agriculture and PA have, however, been identified as drivers for placing agriculture production in a more sustainable, economically profitable and environmentally friendly context. This lack of objective statistics was raised by various stakeholders (academics, manufacturers and service providers) for whom this information was necessary in order to establish new developments, new services, research programs and/or training courses to support PA adoption. This context has highlighted the need for an observation centre of digital agriculture adoption.

Organisation of the observation centre

The French Observation Centre of Digital Agriculture Adoption (FrOCDA) was created in 2016. Its organization is the result of a convergence between two independent dynamics in digital agriculture, the AgroTIC Business Chair at SupAgro Montpellier (www.agrotic.org) and the #DigitAg convergence lab (www.hdigitag.org). The FrOCDA is co-financed by these 2 consortia (Figure 1) for a total amount of 60 k€ per year, over a renewable period of 3 years.

The AgroTIC Business Chair is a consortium that brings together 1 research institute, 2 universities and 24 companies. One of the challenges of the Chair is to bring together collective dynamics to meet the needs in education and the appropriation of new technologies in agriculture. The #DigitAg convergence lab brings together 4 research institutes, 3 universities, 2 transfer structures and 8 companies. One of the main actions of #DigitAg is to finance research projects in the field of digital agriculture. Both these consortia bring together a wide variety of companies (28 in total). These companies cover a large part of the digital agriculture chain. Through these two consortia, the FrOCDA has access to manufacturers, dealers, users (co-operatives, advisors), technical media and digital agricultural companies, such as software, sensor or robot manufacturers.

Although complex, this organization has several advantages: (1) the involvement of universities and research institutes ensures commercial neutrality and academic legitimacy; (2) the involvement of

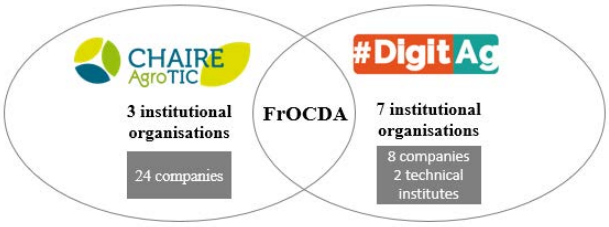


Figure 1. Organisation of the French Observation Centre of Digital Agriculture Adoption (FrOCDA).

agricultural organisations guarantees focusing on high-stakes technologies; and (3) the involvement of digital service companies provides access to important information on sales and adoption; (4) it provides joint public and private funding over renewable periods of 3 years, ensuring the sustainability of the project; (5) the project benefits from public and sponsorship funding. This occurs within the framework of a university foundation affiliated with Montpellier SupAgro, which requires the results not to be confidential.

Study management

The management and execution of the studies is carried out by an operational team led by an engineer, recruited full-time on this project, with strong expertise in the field of Agtech. A project manager engineer (about 10% of time) and a professor (about 5% of time) consolidate the team for the methodological support. This team constitutes the operational team of the FrOCDA (Figure 2). It defines methods, carries out studies and produces deliverables (documents, infographics). It should be noted that this team is part of a larger teaching and research team. This configuration enables the FrOCDA to benefit from a network of scientific partners as well as university work done by students as part of their training.

Methods for the studies

Selection of the studies

When the FrOCDA was launched, it seemed unrealistic to consider producing a general inventory of the adoption of all the digital tools and services available. It was therefore decided to proceed in a step by step fashion by carrying out short and targeted studies on identified technologies at three month intervals. This approach allows quick and consistent results. The hypothesis is that collectively the successive studies should consolidate statistics on the state of adoption of digital agriculture in France.

The choice of the studies to prioritize is based on the expertise and collective needs of all FrOCDA partners. Figure 2 presents the method used to identify high-priority studies. It is based on workshops that bring together all the FrOCDA partners twice a year. During these workshops, the partners work in different small groups, mixing delegates from companies of each part of the PA value chain (co-operative, manufacturer, service provider) in order to bring up (among other things) the themes of interest to be addressed by the FrOCDA.

These themes remain general and the operational team of the FrOCDA translate them into concrete studies. In a second step, the operational team presents the expected studies and deliverables to all the FrOCDA partners for validation.

This organization has the advantage of prioritizing studies of strategic or current interest to the various partners. In this way, it is expected that the concrete studies meet the real needs of the profession and that they remain consistent with each other. Another advantage is that it involves all companies in the choice of the studies. As the latter are informed, it makes it easier for the

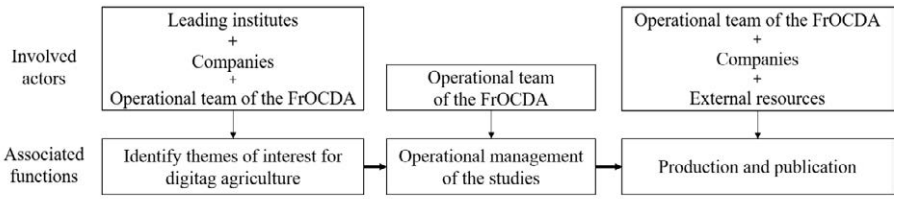


Figure 2. Management and operational structure of the FrOCDA.

operational team to benefit from data or networks of contacts to facilitate studies. This organisation promotes collaboration between the operational team and service companies and makes it easier to access sales figures and adoption by customers (farmers). It should be noted that the rapid and regular production of the FrOCDA deliverables helps to feed and guide the partners' reflections during the workshops.

Methods for the production of studies

FrOCDA produces two types of deliverables: infographics, whose objective is to produce synthetic figures on the adoption of digital tools and services, and thematic brief whose objectives are to produce more precise and qualitative information on a specific agricultural sector and to identify the barriers and drivers of adoption and training needs for farmers and advisors. These two types of deliverables are complementary and make it possible to develop a more systemic vision of the adoption of precision agriculture and digital technologies in general in France. These two deliverables require different methodologies. For the sake of simplicity, this article will only present infographics. This section therefore only describes the methodology that has been implemented for this type of deliverable.

The objective of infographics is to answer the following questions for each service or technology considered: what is the level of use of this solution in France? What are the associated agronomic applications? Are there specificities, particular barriers or drivers of adoption according to the production sectors?

The method used to answer these questions is original since it is not based on a survey made on a sample of farmers. It aims at obtaining the most exhaustive possible view by interviewing service or product suppliers about their market share by region and by type of production. Figure 3 presents the general methodology to achieve these studies. The first step is a comprehensive inventory of all service providers related to the target technology. A follow-up of the main farmers' discussion forums on the web completes this step. This first step aims at identifying the main providers and the main challenges for a given technology. The second step is based on individual interviews with each identified industry actor. These interviews are aimed at obtaining data about their market share, but also qualitative information about adoption and market share evolution. Finally, individual interviews with experts (from technical institutes and advisor companies) and with research and development managers from the main French co-operatives are done to verify the suppliers' data after consolidation at the national scale. A follow-up of the main farmers' discussion forums is used to check the consistency of information.

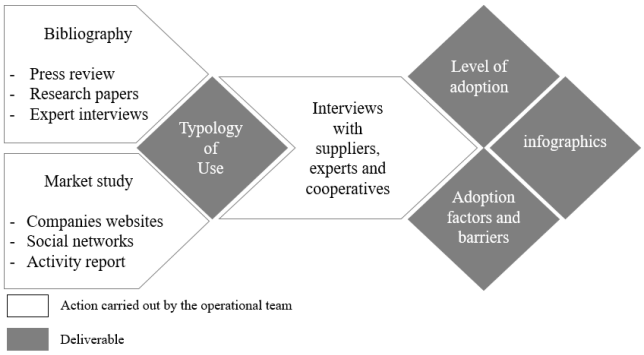


Figure 3. General methodology of the FrOCDA productions.

Results

Choice of studies by the partners of the FrOCDA

Since the launch of the FrOCDA, the partners (research institutes and companies) have prioritized 7 studies corresponding to the following digital tools or services: (1) remote sensing in 2017; (2) smartphone applications; (3) farm management information system; (4) geophysical measurements and soil mapping; (5) remote sensing in 2018; (6) Robotics; and (7) variable rate application technologies (fertilization, seeding, etc.).

Of these 7 studies, the first 6 have been finalised, while the variable rate application technologies study is currently on-going. The reader is invited to visit the FrOCDA website to have access to all the productions (www.agrotic.org/observatoire). It should be noted that the FrOCDA covers the field of digital agriculture, which explains why some studies address areas outside of PA.

This paper will focus only on results relating to PA, i.e. remote sensing in 2017 and 2018, soil mapping and agricultural robotics. It is possible that some PA are studied over two successive years, as is the case here with remote sensing. The possibility to study PA uptake over several years is due to the stability of the organisational structure and reflects the interest of the FrOCDA's partners in understanding and monitoring adoption in key PATs over time, not just as a snapshot.

Remote sensing in France

The area managed with remote sensing in France was estimated from interviews with the main suppliers of remote sensing services. This area was approximately 1 Mha (wheat, barley, canola, vines) over the 2016 and 2017 seasons (Table 1). Over these two years, satellite sensors represented the main technology used, accounting for 85% of the area covered, while aerial and unmanned aerial vehicle (UAV) imagery together account for only 15%.

Remote sensing represented 13% of the area sown to cereals compared to 1% of the area under vines in 2016. Apart from vines and arable crops, remote sensing services are rarely used. For arable crops, the main use of remote sensing is to manage nitrogen fertilization. The services are used for: (1) calculation of the average nitrogen amount to be provided at the field level and (2) manual or automatic variable rate application at the within-field level. For wheat, remote sensing services are used for the 2nd or 3rd nitrogen application, in order to adapt fertilization to the conditions of the current season (weather conditions, diseases). The study shows other very specific applications: (1) detection and assessment of the impact of wild game damage; (2) variable rate seeding; and (3) organization and logistics of harvest sites. In viticulture, it is difficult to make a typology of remote sensing uses. Indeed, from one vineyard to another, the study shows a great diversity, ranging from target sampling (maturity, yield, etc.) to differential harvest.

Comparing 2017 to 2016, a slight decrease in the number of remotely sensed areas was observed for arable crops in 2017 (Table 1). This decrease may be explained by a loss of farm income caused by low yields in 2016 due to climatic conditions. This result shows the fragility of these types of services, which can be strongly dependent on the financial health of the farms. This type of service

Table 1. Level of use of remote sensing in France.

Crop type	Total area measured by remote sensing in 2016 (ha)	Total area measured by remote sensing in 2017 (ha)
Field crops	1,284,238	925,712
Viticulture	7,035	11,200
Other	~0%	~0%

is considered by the farmer as an adjustment variable. The farmer can easily decide not to subscribe to the service in order to maintain financial profitability as it is non-essential for production.

Soil mapping

Within-field soil mapping with apparent electrical conductivity or apparent electrical resistivity, is not widespread in France. Indeed, this service was applied to less than 1% of the French agricultural area for the last 10 years (Table 2).

Soil mapping is used as a support to define soil-sampling strategies in order to consider site-specific potassium and phosphorus applications. For arable crops, it is also used as ancillary information to delineate zones with different textures and water holding capacity. In some rare cases, these zones are for variable rate seeding. Soil mapping is also common in high value crops; such as seed production. To a lesser extent, soil mapping is also used in viticulture where, like remote sensing, it is used for a large diversity of applications from one vineyard to another. The main application remains target sampling to delineate zones of soil calcium content in order to choose the right rootstocks before planting. Other applications, such as variable rate fertilization, differential harvest or targeted implementation of sensors (tensiometers, sap flow) to monitor the vine's water status, were noted. In viticulture, soil mapping is also adopted for 'marketing' issues in relation to 'terroir' characterisation. Having digital soil maps may constitute a promotional tool to explain/justify the uniqueness of the vineyard to consumers. The study highlighted a very low level of adoption in fruit production, mainly associated with the location of sensors (tensiometer, dendrometer) used for irrigation management.

Robotics

Robots are equipped with sensors that can collect data and/or geo-referenced observations (such as temperature, maturity, chlorophyll content). Robots are therefore an interesting platform for precision agriculture. The results of the study showed that agricultural robotics is mainly adopted by livestock farming with about 11,000 robots used in 2018. The adoption for crops, although real, remains very small with approximately 150 robots in use in 2018.

Robots used in livestock farming are mainly milking robots. This represents about 9,000 robots, or nearly 10% of French dairy farms. This number will probably increase rapidly as 70% of new facilities used milking robots in 2018. In breeding systems (bovine, caprine or ovine dairy), about 2,000 robots are used to feed or clean buildings (scrapers and slurry vacuum cleaners). In planted crops, the FrOCDA estimated that almost 150 operational robots were used in 2018. These robots were mainly used for mechanical weeding. It was noted that most of the farms using these kind of robots were vegetable organic farms (~100 robots). The study identified viticulture as second for robot adoption (<10) with these robots also dedicated to mechanical weeding. In addition to these figures, the study identified the emergence of many different robots in the pre-test phase for various applications (such as weeding in arable crops, poultry).

Observed benefits/outcomes

It is difficult to have an exhaustive picture of the impact of FrOCDA on all stakeholders in agriculture in France. However, it is possible to note direct impacts on the partners involved in the FrOCDA. For

Table 2. Cumulative soil area mapped by soil sensors in France as of 2017.

Type of technology	Estimated area cumulated in 2017 (ha)
Soil apparent resistivity	50,000
Soil apparent conductivity	85,000
Total	135,000

production structures (co-operatives), the project has provided a better insight into the adoption of PA by farmers and has made it possible to implement targeted support initiatives (such as farmers' groups, targeted ongoing trainings). For the universities (Montpellier SupAgro and Bordeaux Sciences Agro), FrOCDA results have made it possible to design specific and adapted continuing training courses according to immediate needs of the agriculture industry. More specifically, advisors were identified as the key stakeholder to support farmers in the adoption of PA and, as a result, targeted trainings were specifically designed with contents focusing on remote sensing services for variable rate application and data exchange (ISOBUS in particular). Finally, although the strategies of companies remain difficult to know precisely for reasons of confidentiality, emerging collaborations between manufacturers and software companies in order to promote and facilitate the exchange of information has been noted.

Conclusions

This work presented the organization and functioning of the French Observation Centre of Digital Agriculture adoption (FrOCDA). The FrOCDA has conducted 7 studies on PA uptake in France since its launch in 2016. This organisation is sustainable (renewable every 3 years) and allows, thanks to targeted studies made every 3 months, to consolidate a vision of PA uptake in France. After 2 years of existence, the FrOCDA is becoming a reference organisation for advising official survey structures in order to introduce new questions on the adoption of PA in France and for producing figures for regional and national policies. The studies carried out until now show that PA is used by French farmers. The issues associated with the adoption of PA are economic (better use of inputs) and also related to improving comfort at work for some operations, which explains the growing interest and adoption of robotics.

References

DEFRA, 2013. Farm practices survey October 2012 – current farming issues. Department for Environment, Food and Rural Affairs of UK. https://www.gov.uk/government/statistics/farm-practices-survey-october-2012-current-farming-issues (Last accessed January 2019).

Erickson, B., J. Lowenberg-DeBoer, and J. Bradford, 2017. 2017 Precision agriculture dealership survey. Departments of Agricultural Economics and Agronomy, Purdue University, West Lafayette, IN, USA. http://agribusiness.purdue.edu/precision-ag-survey (Last accessed January 2019).

Lawson, L. G., Pedersen, S. M., Sørensen, C. G., Pesonen, L., Fountas, S., Werner, A., *et al.,* 2011. A four nation survey of farm information management and advanced farming systems: A descriptive analysis of survey responses, *Computers and Electronics in Agriculture* 77: 7-20.

Lewellyn, R. and J. Ousman, 2014. Adoption of precision agriculture-related practices: Status, opportunities and the role of farm advisors. Commonwealth Scientific and Industrial Research Organisation (CSIRO). https://grdc.com.au/resources-and-publications/all-publications/publications/2014/12/adoption-of-precision-agriculture-related-practices (Last accessed November 2018).

Miller, N. J., Griffin, T. W., Bergtold, J., Ciampitti, I. A., & Sharda, A., 2017. Farmers' Adoption Path of Precision Agriculture Technology. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences, 8 (2), 708-712.

Paustian, M., & Theuvsen, L., 2017. Adoption of precision agriculture technologies by German crop farmers. Precision Agriculture, 18 (5), 701-716.

Tey, Y. S., & Brindal, M., 2012. Factors influencing the adoption of precision agricultural technologies: a review for policy implications. Precision agriculture, 13 (6), 713-730.

USDA ARMS, 2018. Agricultural Resource Management Survey. https://www.ers.usda.gov/data-products/arms-farm-financial-and-crop-production-practices/documentation (Last accessed November 2018).

How does European adoption of precision agriculture compare to worldwide trends?

J. Lowenberg-DeBoer[1*] and B. Erickson[2]
[1]*Harper Adams University, Newport, Shropshire, UK TF10 8NB, United Kingdom;* *jlowenberg-deboer@harper-adams.ac.uk*
[2]*Purdue University, West Lafayette, Indiana 47907, USA*

Abstract

European politicians sometimes complain that farmers are lagging on precision agriculture (PA) adoption. The objective of this study was to compare European PA adoption with other industrialised countries in existing studies that used a variety of methodologies. The data showed that global navigation satellite systems (GNSS) guidance is being adopted by mechanised farms worldwide. In Europe, guidance adoption is slightly lower than in the Americas and Australia probably because of smaller farms and the prevalence of horticultural and grazing-based livestock farms. In contrast, variable rate technology (VRT) fertiliser adoption is lagging almost everywhere, but in parts of Europe this is at the upper end of the worldwide range.

Keywords: adoption, guidance, variable rate technology, demonstrated value

Introduction

There is widespread perception that precision agriculture (PA) adoption in Europe is occurring too slowly. Many academic studies of PA use have been motivated by the perceived slow pace of adoption. Agricultural technology conferences have sessions focused on adoption barriers. PA is not usually high on the political agenda, but when agricultural technology is mentioned the concern is often about why the country's farmers are lagging behind in technology adoption. For example, in 2015 the United Kingdom (UK) Parliament Office of Science and Technology stated, 'Precision farming uses technology to improve efficiency. It offers benefits for yields, profits and the environment. However, uptake by farmers has been slow.' The Italian Ministry of Agriculture guidelines for precision agriculture make a similar comment (2015). The objective of this study was to compare European PA adoption data with information from other industrialised countries. Because PA adoption data collection methods vary widely from country to country there are limitations in making direct numerical comparisons. Consequently the methodology is impressionistic comparison that looks at the big picture, rather than making quantitative comparisons. This study will be of interest to PA researchers and educators across all the disciplines involved, to agribusinesses involved in manufacturing and selling PA tools, and policymakers concerned about agricultural productivity and the environment.

The lack of a clear definition of PA makes tracking adoption more difficult. The definition of PA used in this study is: 'Precision agriculture is a management strategy that uses electronic information and other technologies to gather, process and analyse spatial and temporal data for the purpose of guiding targeted actions that improve efficiency, productivity and sustainability of agricultural operations.' In this definition, 'agriculture' should be broadly defined to include all types of biological production systems (e.g. arable cropping, perennial forages, forestry, orchards, vineyards, horticulture). This study focuses on crop PA, especially grain and oilseed production. In general this study refers to satellite-based geographical locational systems as 'global navigation satellite systems (GNSS)'. The term Global Positioning System (GPS) is only used where a survey asked specifically about this US technology.

DOI 10.3920/978-90-8686-888-9_106, © Wageningen Academic Publishers 2019

The general objective was to compare adoption of PA in Europe with the use of these technologies in other industrialized countries using existing studies. Specific objectives were: (1) document adoption of PA in industrialised countries outside of Europe based on existing large random sample studies; (2) identify PA adoption trends in Europe based on a critical evaluation of the sampling and analysis methodologies employed; and (3) determine how PA adoption patterns in Europe have differed from those elsewhere. The hypotheses corresponding to these objectives are: (1) adoption rates differ widely among PA tools, with GNSS guidance having become standard practice in most mechanised farming systems around the world and in Europe; and (2) VRT fertiliser adoption is relatively low in most of Europe consistent with adoption patterns elsewhere.

Materials and methods

Because no country regularly collects statistics on PA. There is no international consensus on a PA definition; PA is often referenced as 'digital', 'site-specific', 'smart', etc. agriculture. PA involves numerous technologies, and much information exists outside academic journals and in a variety of languages. Systematic review techniques would not ensure a thorough assemblage of PA adoption information. Consequently, the authors relied on their network of academic and business contacts, interactions at trade shows and scientific meetings, as well as internet searches. The studies collected data using different methods and at various times. Consequently, the methodology was to look at the big picture and overall patterns in existing studies, rather than numerical comparison. Data quality and representativeness varies widely. Denmark, USA, UK and Australia have PA adoption data collected by government agencies using standard random sampling techniques. Within Europe, many PA adoption surveys used non-random samples (e.g. volunteers who respond to a survey posted on the internet, or attendees at farm shows).

Most PA adoption surveys measure adoption by percent of farms using the technology. This is often the easiest and clearest way to pose the question, but it does not measure intensity of use either on the respondent farm or regionally. It avoids the problem of misreported or inaccurate area measurements. When using percent of farms to measure adoption, a farmer who used VRT on one field is the same as a farmer who uses it on all fields. When the distribution of farm sizes is skewed and/or the sample is not representative, percent of farms may be misleading. This study identifies which surveys used percentage of farmers as compared to those that used percentage of land area. The standard way to ensure representative results is to randomly sample from an exhaustive list of farmers. Non-random selection of respondents often leads to a sample that is biased but other factors can also create representativeness problems. For instance, if the initial list of farmers is drawn from a specific size, farm type, geography or other list, even a random sample will not be representative of all farmers. A systematic non-response can also create unrepresentativeness. For instance, if some farmers in an area consider PA to be a waste of time and money, they may be less likely to answer a PA survey and the results would be dominated by the farmers that consider PA useful.

The methodology was to organize the available PA adoption data in tables with key information about the data collection methods including year of the data collection, collection method and source of respondents. The data collected using standard random sampling methods is discussed first because it is probably the most reliable. Other surveys in Europe are then discussed in the light of those random sample surveys. Data was collected going back more than twenty years, but this report is limited to the last ten years. The older surveys from Europe and those from America, Asia and Latin America are discussed by Lowenberg-DeBoer (2018).

Results and discussion

The USA, UK, Australia, and Denmark governments collect data on PA adoption by commodity crop farmers using random sampling methods (Table 1). This data is probably the most reliable and

representative PA adoption information available. Unfortunately, the sample selection procedures and questions asked differ substantially among countries:

- The US Department of Agriculture (USDA) Agricultural Resource Management Survey (ARMS) (USDA, 2018, and USDA/ERS, personal communication, 2019) is organized by crop for the continental USA, so it excludes Alaska and Hawaii. The USDA definition of a farm is 'any place from which $1000 or more of agriculture products were produced or sold, or normally would have been sold during the year.' Consequently, it includes a wide range of farm sizes. Data is collected and organized by crop and state, and only in states where the focus crop is a major farm product. The maize survey is done in 20 states, but the peanut survey includes just five states. The schedule is irregular; not every crop is surveyed every year. Respondents are randomly selected from a list maintained by the USDA of US farmers producing the focus crop. Response is voluntary, but the response rate is high because the mailed paper surveys are followed by a visit from a government staffer for a personal interview 'at the kitchen table.' The number of farmers surveyed each year depends on the focus crops in a given year and varies from 5,000 to 30,000. Farmers report only for the crop specified, not their entire operation.
- The UK Department of Food and Rural Affairs (DEFRA) farming practices survey in 2012 was done only in England, not in Scotland, Wales or Northern Ireland. To be included in the survey, a farm must have at least one of the following: 50 cattle, 100 sheep, 100 pigs, 1000 poultry, 20 ha of arable crops, 20 ha of orchard, 5 ha of berries or other soft fruit, or 10 ha of nursery stock. For the 2012 survey, DEFRA randomly selected 2,900 farms (48%) from the roughly 60,000 English farms that fit the minimum size criteria. Under UK law, response to the farming practices survey is voluntary. The DEFRA farming practices survey was implemented by paper mail. The farm categories shown reflect the farm's primary enterprise.
- The Australian data in Table 1 was collected in 2012 by the Commonwealth Science and Industrial Research Organization (CSIRO) for the Grain Research and Development Corporation (GRDC) (Llewellyn and Ousman, 2014). CSIRO is an Australian government entity operating under the Ministry of Industry and Science. The survey was done in the major grain growing regions of Australia in the states of New South Wales, Victoria, South Australia and Western Australia. Respondents were randomly selected from a comprehensive database of grain growers with over 500 ha of grain. Survey participation was voluntary and the interview was done by telephone. Responses were obtained from 573 farmers.
- Denmark Statistics has collected data annually since 2017 on precision agriculture (Denmark Statistics, 2017 & 2018). A stratified sample is drawn from all farms with cultivated area registered with the government for administrative and regulatory purposes. The data from the 2017 sample is listed in Table 1 because it included data on both GNSS guidance and VRT. The results in 2018 were similar, but GNSS guidance had risen to 23% of farms. The 2017 data are based on responses from 6281 farms from a total of 33,580 farms on the government list.

Key differences between the US, UK, Australian and Danish surveys include size and type of farm. The US survey included some very small farms due to the USDA's definition of a farm, as previously discussed. Depending on the crop, price and yield, $1000 of production might represent a farm of less than one hectare. The UK arable farms must have at least 20 ha and the Australian respondents were selected from a list of farms with over 500 ha of grain. The Danish survey drew from a list of all farms registered with the government including horticultural producers and very small farms. In all three countries, the farm businesses surveyed may have included livestock and other enterprises. The US survey asked specifically for the technology used for the focus crop. By selecting from a list of farmers with over 500 ha of grain, the Australian survey may have chosen a sample more specialized in grain production. Because of the different types of farms included in the UK sample, the 'Cereals' and 'Other Crops' data may be the most comparable to the USDA ARMS data and the Australian information.

Table 1. PA adoption for government studies using random sampling methods.[1]

	Data collect. year	GNSS guidance[2]	Yield monitor	Yield map[3]	VRT any[4]	VRT fert.	GNSS soil map[5]	EC[6]
United States		percentage of area using the technology						
Maize	2016	59%	68%	45%	NA	29%	22%	NA
Rice	2013	53%	58%	NA	21%	21%	12%	NA
Peanuts	2013	49%	NA	NA	18%	18%	24%	NA
Soybean	2012	45%	63%	NA	22%	NA	19%	NA
Winter Wheat	2009	35%	36%	6%	14%	11%	6%	NA
United Kingdom		percentage of holdings using the technology						
All Farms	2012	22%	NA	11%	16%	NA	20%	NA
Cereals	2012	46%	NA	25%	31%	NA	38%	NA
Other Crops	2012	40%	NA	18%	23%	NA	31%	NA
Australia		percentage of growers who have used the practice						
All Respondents	2012	77%	NA	33%	15%	49%	NA	15%
New South Wales	2012	76%	NA	26%	11%	34%	NA	15%
South Australia	2012	79%	NA	30%	15%	52%	NA	13%
Victoria	2012	74%	NA	33%	16%	50%	NA	16%
W. Australia	2012	79%	NA	43%	17%	55%	NA	19%
Denmark		percentage of farms with cultivated land						
All Farms	2017	16%	NA	NA	7%	NA	NA	NA

[1] NA = not available. Sources: US: USDA ERS (2018), except maize is USDA/ERS (2019); UK: UK DEFRA (2013), Australia: Llewellyn and Ousman (2014), Denmark: Denmark Statistics (2017).
[2] For USDA ARMS survey, the GNSS guidance prior to 2013 is 'guidance or autosteer.' For UK in 2012, the GNSS question was general. It asked about GPS and gave autosteer and guidance as examples. Australia asked specifically about autosteer use. The Danish survey asked specifically about 'RTK GPS' where RTK is Real Time Kinematic.
[3] For USDA ARMS survey, 'yield map created.' UK DEFRA asked only if used 'yield mapping' without explaining. For Australia, Llewellyn and Ousman report data on 'yield map adopters.'
[4] 'VRT Any' indicates variable rate technology for any purpose including fertiliser, seed and plant protection chemicals.
[5] For USDA ARMS Survey through 2013 'GPS device used to create soil properties map.' For UK DEFRA survey, asked only if used 'soil mapping'. The Australian survey asked about 'soil testing' but this apparently was not necessarily GNSS linked.
[6] For USDA ARMS survey, 'soil properties map based on electrical conductivity.' In Australia, electrical conductivity (EC) or Gamma. UK and Denmark surveys did not ask about EC.

- GNSS guidance: The US, UK, Australian and Danish surveys show the widespread adoption of GNSS guidance by mechanised grain farmers. Even though the Australian survey is six years ago, their guidance adoption percentages are highest (Table 1). This is consistent with the fact that the first auto guidance systems were introduced in Australia (in 1997 at least 3 years before introduction in North America) (Quick, 2007) and the relatively large size of the farms in the Australian sample. The US survey estimates show the growth of GNSS guidance use with the winter wheat survey in 2009, showing 35% of winter wheat area farmed with either lightbar or autosteer guidance, to the 2016 maize survey with 59% of maize area farmed with guidance technologies. The 2012 DEFRA survey shows that 46% of cereal crop farms used some type of guidance, which is similar to the level of use in the US data at that time. The GNSS guidance estimates for Denmark are exclusively for RTK GPS guidance and consequently omit the light bar and other 'manual' guidance technologies.
- Yield mapping: The yield monitor and yield map columns for the USDA ARMS data show that many US farmers have a yield monitor, but do not bother to make yield maps. For example, the 2016 maize data indicates that 68% of the area was harvested with a combine equipped with a yield monitor, but only 45% of the area was yield mapped. Yield maps are usually considered the first step in making management use of the yield monitor information, but possession of yield maps does not guarantee that the maps were used in management. Industry sources indicate that most new combines in these four countries come with yield monitors as standard equipment. The percentage making yield maps in the three countries is comparable: US, 45% of maize area in 2016, UK, 25% of cereal farmers in 2012, and Australia, 33% of grain farmers in 2012. The Danish report did not include yield monitor adoption estimates.
- VRT: In the four countries in Table 1, use of 'VRT for any purpose' only exceeded 30% for cereal farms in the UK. The surveys generally estimate it at 20% or less across these four countries. At 7% the Danish estimate is among the lowest in this group. Oddly, the percentage of farms in the Australian data using variable rate fertiliser is higher than using VRT for any purpose. This is because of a quirk in the Australian data that includes manual variable rate in the VRT fertiliser estimates, but manual is not included in the 'VRT for any purpose' estimates. The USDA estimates illustrate that, in most cases, computer-controlled variable rate is fertiliser. The percentages in the VRT fertiliser column are the same or only slightly less than is the total VRT for any purpose column.
- GNSS soil mapping: In many countries site-specific soil maps are commonly the foundation for VRT fertiliser recommendations, but the data indicates that only a modest proportion of farms have such maps. The highest percentage among this group is 38% for cereal farms in the UK. The Australian and Danish reports did not include GNSS soil mapping.
- EC: In the 2012 Australian survey, an overall 15% of farms report doing soil electrical conductivity (EC) mapping. A question about soil EC maps was asked in some earlier USDA ARMS surveys. In the 2010 maize survey (not shown) and the 2002 soybean survey only 1% of area was managed with soil EC maps. The UK DEFRA and Danish surveys did not ask about EC.

PA adoption data from European studies that use other methodologies is summarized in Table 2. Because of differences in sampling, technology definitions and the questions asked, comparison among studies can provide only a rough impression of the trends. For more information on the methodology used in each study, readers should consult the original documents.

Some of the studies in Table 2 use almost standard random sample techniques and should have reliably representative data. The Lawson *et al.* (2011) data for Germany, Denmark and Finland seems to have been collected with random sampling methods. In Denmark and Finland, questionnaires were mailed to a random sample of farms registered with their governments. In Denmark, the list was from arable, dairy and swine farms, omitting horticultural operations unlike the Denmark Statistics surveys (2017, 2018). In Germany, farms were randomly selected from a list of farms receiving over

Table 2. PA adoption rates for European farmers in other studies.

Location	First author	Collection year	Method	Source of list	GNSS guidance	Yield mapping	VRT fertiliser	GNSS soil test mapping	EC	Remote sensing images
					percent of farmers					
Germany	Lawson	2008	post	EU farm support >40,000€	37%	12%	8%	20%	1%	11%
Denmark	Lawson	2008	post	gov't registered farmers	9%	3%	4%	6%	1%	1%
Finland	Lawson	2008	post & web	gov't registered farmers	1%	1%	3%	1%	0%	0%
Italy	Cavallo	2010	in person	farm show participants	~46%	NA	NA	NA	NA	NA
UK	Pickthall	2016	web	author contacts	20%	13%	14%	15%	NA	NA

€40,000 in EU support. The PA adoption estimated by Lawson *et al.* (2011) fits with the patterns seen in Table 1. GNSS guidance adoption was highest with modest interest in the other PA technologies. Cavallo *et al.* (2014) interviewed more than 300 Italian visitors at the International Exposition of Agricultural Machinery (EIMA) in Bologna in 2010. They found that about 46% used GNSS guidance. This is larger than the GNSS guidance adoption rate reported by Lawson *et al.* (2008) in northern Europe, but consistent with the fact that larger than average farms were over represented. Pickthall and Trivett (2017) conducted a web based survey of UK farmers and found PA adoption levels very similar to those in the 2012 DEFRA survey (Table 1) in spite of the fact that the respondents tended to be younger than most UK farmers. The survey was sent via email and social media to contacts known to the authors, with 77 responses. The mean age of respondents was 34 years, while the average age of UK farmers was 59 years.

Because it is not a survey, data from the Digital Agriculture Adoption Observatory in France is not in Table 2. It is a partnership involving two universities, eight research institutes and 28 companies (cooperatives, manufacturers or service providers in digital agriculture) that gathers data on farm use of PA technology reported by the companies. This data fits the pattern of modest adoption of VRT, EC and remote sensing. It indicates that about 4% of French crop farmers used VRT in 2018, mostly for fertiliser, about 10% of field crop area was managed with satellite, aircraft or drone images in 2017, and less than 1% of French agricultural land was mapped by either electrical conductivity or electrical resistance sensing in 2017 (Digital Agriculture Adoption Observatory 2017, 2018a, 2018b).

Conclusions

From the confusing mass of data collected from studies that used different sampling methods and analysed in a variety of ways emerges a picture of rapid adoption of GNSS guidance, and slower adoption of VRT of all kinds. Adoption of information gathering technologies like yield mapping and remote sensing which are used to inform VRT decisions has also been relatively slow. These patterns hold in Europe, as well as in the rest of the world.

The data show that GNSS guidance adoption in Europe is slightly lower than in the Americas and Australia. This is probably because, in Europe, farms tend to be smaller, many of them are grazing-based livestock farms and, in some countries, a substantial proportion are small horticultural farms. In contrast, variable rate fertiliser adoption in parts of Europe is at the upper end of the worldwide range. For example, the 31% adoption of VRT fertiliser on UK cereal farms is one of the highest VRT adoption estimates based on a representative sample.

References

Cavallo, E., Ferrari, E., Bollani, L., and Coccia, M., 2014. Attitudes and behaviour of adopters of technological innovations in agricultural tractors: A case study in Italian agricultural system. Agricultural Systems 130 44-54. doi:10.1016/j.agsy.2014.05.012

Denmark Statistics, 2017. Satellite technology is gaining ground among young farmers. (in Danish) https://www.dst.dk/da/Statistik/nyt/NytHtml?cid=29269 (last accessed 24/1/19)

Denmark Statistics, 2018. Advanced technology occupies Danish fields. (in Danish) https://www.dst.dk/da/Statistik/nyt/NytHtml?cid=30775 (last accessed 24/1/19)

Digital Agriculture Adoption Observatory, 2017. Adoption of electrical resistance and electrical conductivity to characterize agricultural soils. (in French) http://agrotic.org/observatoire/2017/12/06/usages-de-la-cartographie-pour-caracteriser-les-sols-agricoles (last accessed 24/1/19)

Digital Agriculture Adoption Observatory, 2018a. Adoption of intraparcel variable rate. (in French) http://agrotic.org/observatoire/2018/12/05/usages-de-la-modulation-intra-parcellaire (last accessed 24/1/19)

Digital Agriculture Adoption Observatory, 2018b. Adoption of remote sensing in agriculture – 2017 Season. (in French) http://agrotic.org/observatoire/2018/04/25/usages-de-la-teledetection-en-agriculture-campagne-2017 (Last accessed 24/1/19)

Italian Ministry of Agriculture, Food and Forestry, 2015. Guidelines for the development of precision agriculture in Italy. (In Italian) https://www.politicheagricole.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/12069 (last accessed 24/1/19)

Lawson, L. G., Pedersen, S. M., Sørensen, C. G., Pesonen, L., Fountas, S., and Werner, A., *et al.,* 2011. A four nation survey of farm information management and advanced farming systems: A descriptive analysis of survey responses. Computers and Electronics in Agriculture 77 7-20. doi:10.1016/j.compag.2011.03.002

Lewellyn, R. and Ousman, J., 2014. Adoption of precision agriculture-related practices: Status, opportunities and the role of farm advisors. Commonwealth Scientific and Industrial Research Organisation (CSIRO). https://grdc.com.au/resources-and-publications/all-publications/publications/2014/12/adoption-of-precision-agriculture-related-practices (last accessed 24/1/19)

Lowenberg-DeBoer, J., 2018. The economics of precision agriculture. In: Precision Agriculture for Sustainability, J. Stafford (Ed.), Cambridge, UK: Burleigh Dodds Science Publishing Ltd. doi:10.19103/AS.2017.0032.19

Pickthall, T. and Trivett, E., 2017. An investigation into the barriers that prevent the adoption of precision farming technologies in combinable cropping in the UK. In: Aspects of Applied Biology 135: Precision systems in agricultural and horticultural production, I. Grove and R. Kennedy (Eds.), p. 29-37. https://www.aab.org.uk/product-page/aspects-135-precision-systems-in-agricultural-and-horticultural-production (last accessed 24/1/19)

Quick, G., 2007. Remarkable Australian Farm Machines: Ingenuity on the Land. Rosenberg Publishing. Kenthurst, NSW, Australia.

United Kingdom Department for Environment, Food and Rural Affairs (DEFRA)., 2013. Farm Practices Survey Autumn 2012 – England. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/181719/defra-stats-foodfarm-environ-fps-statsrelease-autumn2012edition-130328.pdf. (last accessed 24/1/19)

United Kingdom Parliamentary Office of Science and Technology (POST), 2015. POSTnote 505. Precision farming. https://researchbriefings.parliament.uk/ResearchBriefing/Summary/POST-PN-0505. (last accessed 24/1/19)

United States Department of Agriculture (USDA) Economic Research Service (ERS), 2018. 'ARMS farm financial and crop production practices' data product, https://data.ers.usda.gov/reports.aspx?ID=17883 (last accessed 24/1/19)

United States Department of Agriculture (USDA) Economic Research Service (ERS), 2019. Personal communication.

Collection of agricultural data and evaluation of their usefulness for farm management in Australian cropping and red meat industries

A. Zhang[1*], J. Trindall[2] and E. Hobman[1]
[1]*Commonwealth Scientific and Industrial Research Organisation, Ecosciences Precinct, 41 Boggo Road, Dutton Park Qld 4102, Australia; airong.zhang@csiro.au*
[2]*Cotton Research and Development Corporation, 2 Lloyd St, Narrabri NSW 2390, Australia*

Abstract

Computer-assisted telephone interviews (CATI) were conducted with 786 Australian primary producers in cropping and red meat industries to examine on-farm data collection practices and producers' perceptions of the usefulness of such data for making farm management decisions. The results revealed that the collection of agricultural data was not undertaken extensively by producers, suggesting that the practice was at an early stage of adoption in the industry. Data collection varied significantly across data types, as well as across producers and industry sectors. Producers who collected data tended to perceive that data would bring greater benefits on a range of variables (e.g. business profit, risk management), as compared to their counterparts who did not collect data at all.

Keywords: agricultural data, collection rate, data usefulness, farm management

Introduction

Advances in digital technologies offer the potential for transformational change in agriculture, providing a precision management approach to improve productivity and profitability while minimising environmental impact. Rapid developments in computing power, sensing technologies, robotics, big data, the Internet of Things, and cloud computing are creating opportunities for more data-driven approaches to farm management (Griffith *et al.*, 2013; Keogh and Henry, 2016; Wolfert *et al.*, 2017). The application of digital agriculture tools and big data analytics can increase farm productivity through input efficiencies and increased output. To realise the potential benefits of digital agriculture and big data, it is essential that primary producers collect on-farm data to inform their decision-making (Bramley and Ouzman, 2019; Miller *et al.,* 2017; Robertson *et al.*, 2012).
Funded by Australian Government Department of Agricultural and Water Resources' Rural Research & Development for Profit program, all 15 Research and Development Corporations (RDCs) in Australia have undertaken a detailed analysis of what digital agriculture means for Australia, and what needs to be done in the future for Australian producers to capture the potential value from agriculture data (Trindall *et al.*, 2017; Zhang *et al.*, 2017). As part of this effort, the present research conducted a cross-industry producer survey in Australia to enhance the understanding of effective data technology adoption and data sharing pathways.
The purpose of the project component reported here was to provide a benchmark of the state of data collection by Australian primary producers in the cropping and red meat industries. Such understanding can inform targeted strategies and policies to enable producers to increase their uptake of precision agriculture technologies and better utilise agricultural data for productivity, profitability and sustainability.

DOI 10.3920/978-90-8686-888-9_107, © Wageningen Academic Publishers 2019

Materials and methods

Respondents and procedure

A specialised agricultural research survey company was engaged to conduct the data collection. The survey was conducted via computer-assisted telephone interviewing (CATI) during the period of 7th March to 18th April in 2017. The sampling specifications for each industry were defined in consultation with the participating Research and Development Corporations (RDCs). Potential respondents were drawn from the research company's database. In addition, various RDCs publicized the survey in their newsletters and invited their members to participate. To be eligible for the survey, respondents had to be the primary decision-maker on the farm.

In total, 260 producers (M_{age}=54.8, SD_{age}=12.6; 10% female) from the cropping industries and 526 producers (M_{age}=58.4, SD_{age}=12.5; 14% female) from the red meat industries participated in the survey. For respondents who had multiple components to their business (i.e. it spanned more than one industry; for example, beef and grain), they were asked to indicate the major component of their business. The survey items were answered in relation to the major component of their business. For example, if the respondent had both beef and grain, and indicated beef as the major component to their business, the respondent would be classified as beef/grain mixed, and all answers would refer to their beef component. On the other hand, if the respondent indicated grain as the major component, the respondent would be classified as grain/beef mixed, and all answers would refer to their grain component. This classification principle applied to all other mixed combinations (i.e. beef/sheep mixed, sheep/grain mixed, and grain – grain/beef/sheep). Table 1 presents the number of respondents from each industry, as well as their average farm size and business intensity.

Measures

Data type. Cropping and red meat industries collected different sets of data due to their divergent business operation needs. The key data types were identified in consultation with industry experts. Respondents from cropping industries were asked whether they collected the following agricultural data: yield mapping, soil mapping, crop sensing (e.g. NDVI), weed pressure mapping, soil moisture sensing, on-farm weather station, and irrigation use.

Table 1. Average farm size and business intensity by industry.

Industry	Number of respondents	Average farm size (ha)	Business intensity		
			unit	Average	Median
Red meat industry (n=526)					
Beef	126	11,723	number of beef cattle	1,868	775
Beef/grain mixed	64	6,779	number of beef cattle	1,516	430
Beef/sheep mixed	94	6,792	number of beef cattle	662	270
			number of sheep	4,046	3,000
Sheep meat (lamb)	59	8,726	number of sheep	21,154	2,800
Sheep/grain mixed	94	3,466	number of sheep	4,287	3,000
Sheep wool	89	7,281	number of sheep	4,954	3,000
Cropping industry (n=260)					
Grain	77	3,936	area planted to grain	2,464	2,020
Grain – grain/beef/sheep	73	4,569	area planted to grain	2,516	2,023
Cotton	30	6,866	area planted to cotton	3,595	403
Rice	15	2,424	area planted to rice	171	100
Sugarcane	65	335	area planted to sugarcane	159	80

Respondents from red meat industries were asked whether they collected the following agricultural data: soil mapping, pasture/vegetation mapping, individual animal or herd production, individual animal or herd feeding data, on-farm weather station data, animal breeding data, veterinary medicine record data, and water use/quality data.

Evaluation of data usefulness was measured by asking respondents to indicate how useful each data type that they collected was in helping them make farm management decisions on a 5-point scale (1 = no use at all, 5 = extremely useful). Thus, an evaluation score for each data type was collected.

General benefits of data on four aspects – farm management, business profit, operation efficiency and risk management – was measured with four items. After respondents listed the agricultural data that they were collecting and evaluated each data type, they were asked to indicate their overall rating of the data they had collected. For the participants who had not collected any data, they were asked to indicate their overall rating of the data if they had collected them. The survey questions are presented in Table 2.

Table 2. Survey questions on general benefits of data.

Respondents who indicated they collect data	Respondents who indicated that they do not collect any data
Farm management 'Overall, how useful was the data you had collected in helping you make farm management decisions?' (1 = no use at all, 5 = extremely useful)	*Farm management* 'Overall, how useful would be the data in helping you make farm management decisions if you had them?' (1 = no use at all, 5 = extremely useful)
Business profit 'Overall, how much contribution had the data made to your farm business profit?' (1 = no positive contribution at all, 5 = increased profit greatly)	*Business profit* 'Overall, how much contribution would the data make to your farm business profit if they had them?' (1 = no positive contribution at all, 5 = increased profit greatly)
Operation efficiency 'Overall, how much contribution had the data made to the efficient running of your farm?' (1 = no positive contribution at all, 5 = increased efficiency greatly)	*Operation efficiency* 'Overall, how much contribution of the data make to the efficient running of your farm if they had them?' (1 = no positive contribution at all, 5 = increased efficiency greatly)
Risk management 'Overall, how much contribution would the previously mentioned data you currently collect make to the risk management of your farm operations?' (1 = no positive contribution at all, 5 = improved risk management greatly)	*Risk management* 'Overall, how much contribution would the previously mentioned data make to risk management of your farm operations if you had them?' (1 = no positive contribution at all, 5 = improved risk management greatly)

Results

Data collection rates and reported usefulness of collected data for farm management decisions

Cropping industries

Table 3 presents the percentage of respondents from each cropping industry that collected certain types of data. On average, data collection rates were reasonably low across all data types. Yield mapping data (57%) was the most collected data across industries, followed by soil mapping data (43%), on-farm weather station data (31%), soil moisture sensor data (23%), crop sensing data (20%), irrigation use (18%) and weed pressure mapping data (12%). Apart from the cotton industry, there were still a significant proportion of producers (23%) who did not collect any agricultural data. This proportion of 'non-collectors' ranged from 20% (rice industry producers) to 27% (grain/beef/sheep producers).

There was wide variation in data collection rates for all data types across the cropping industries. Respondents from the cotton industry were the most active in collecting various types of data, while those from sugarcane industry lagged behind in collection rates across nearly all types of agricultural data.

Table 4 presents the findings on the perceived usefulness of each type of data (that they had collected) in helping the producer make farm management decisions. As indicated by the average scores, all types of data were regarded as quite useful for informing farm management decisions. In particular, irrigation use data was rated as the most useful, followed by weed pressure mapping data and soil moisture sensor data, soil mapping data, on-farm weather station data, yield mapping data, and crop sensing data. There were some variations in reported usefulness of each data type across industries. For example, irrigation use data was highly regarded as useful by respondents from the cotton industry (M=4.44, SD=0.51, out of 5), but was rated as moderately useful by data collectors from the sugarcane industry (M=3.40, SD=1.17, out of 5).

Red meat industries

Table 5 presents the percentage of respondents for red meat industries that currently collected major types of data. Similar to cropping industries, data collection rates on average were reasonably low across all data types. Veterinary medicine records (63%) were the most collected data type, followed by animal breeding data (54%), individual animal or herd production data (49%), on-farm weather station data (37%), pasture/vegetation mapping data (27%), soil mapping data (25%), individual animal or heard feeding data (23%), and water use/quality data (22%). On average, 14%

Table 3. Data collection rates by the cropping industries.

Industry	Yield mapping	Soil mapping	Crop sensing (e.g. NDVI)	Weed pressure mapping	Soil moisture sensor	On-farm weather station	Irrigation use	None of the listed data is collected
Cotton (n=30)	73%	53%	33%	13%	67%	47%	60%	7%
Grain (n=77)	61%	46%	25%	12%	22%	36%	4%	23%
Grain/Beef/ Sheep (n=73)	49%	38%	12%	11%	19%	26%	7%	27%
Rice (n=15)	60%	47%	20%	20%	20%	33%	67%	20%
Sugarcane (n=65)	51%	40%	17%	11%	8%	22%	15%	25%
Total (n=260)	57%	43%	20%	12%	23%	31%	18%	23%

Table 4. Mean and standard deviation of reported usefulness of collected data for farm management decisions by data collectors in cropping industriesM (SD).[1]

Industry	Yield mapping	Soil mapping	Crop sensing (e.g. NDVI)	Weed pressure mapping	Soil moisture sensor	On-farm weather station	Irrigation use
Cotton (n=30)	3.91 (0.92)	4.31 (0.79)	3.80 (1.14)	4.25 (0.96)	4.30 (0.87)	4.50 (0.94)	4.44 (0.51)
Grain (n=77)	3.60 (1.26)	3.83 (0.95)	3.53 (0.96)	3.56 (1.13)	3.76 (1.20)	3.64 (1.19)	4.00 (1.00)
Grain/beef/ sheep (n=73)	3.47 (1.06)	3.82 (0.82)	3.22 (1.20)	4.63 (0.52)	4.00 (1.11)	3.68 (0.89)	4.60 (0.89)
Rice (n=15)	3.78 (0.97)	4.14 (0.90)	3.67 (0.58)	4.00 (1.00)	4.33 (0.58)	3.20 (1.30)	4.40 (0.84)
Sugarcane (n=65)	3.85 (1.03)	3.77 (1.24)	3.91 (1.22)	4.00 (1.16)	4.20 (1.10)	3.79 (1.25)	3.40 (1.17)
Total	3.68 (1.10)	3.90 (0.98)	3.62 (1.07)	4.06 (1.00)	4.07 (1.03)	3.80 (1.13)	4.20 (0.91)

[1] 1 = no use at all; 5 = extremely useful.

of respondents did not collect any agricultural data whatsoever. The proportion of 'non-collectors' varied across industries, ranging from 9% (beef/grain mixed) to 25% (sheep/grain mixed).
There was also variation in data collection rates for all data types across the red meat industries. In particular, respondents from beef/sheep mixed appeared to be more active in collecting various data, while sheep/grain mixed collected less, comparatively.
Table 6 presents the findings on the perceived usefulness of each type of data (that they had collected) in helping them make farm management decisions. Similar to that observed in the cropping industries, data collectors in the red meat industries rated all types of data as being quite useful for informing farm management decisions. On average, animal breeding data was rated as the most useful, followed by individual animal or herd production data, individual animal or herd feeding data, soil mapping data, pasture/vegetation mapping data, water use/quality data, veterinary medicine record, and on-farm weather station data. There were some variations in reported usefulness of each data type across industries. For example, water use/quality data was rated very useful by data collectors from beef/grain mixed (M=4.13, SD=1.06, our of 5), but was regarded as moderately useful by data collectors from sheep wool (M=3.24, SD=1.11, out of 5).

Overall evaluation of general benefits of data by collectors and non-collectors
Figure 1 and 2 present the overall evaluation of general benefits of data in farm management, business profit, operation efficiency and risk management by respondents who collected one or more types of data (i.e. collectors) and those who did not collect any data (i.e. non-collectors). As displayed in Figure 1 and 2, data collectors from both cropping and red meat industries highly evaluated the data they had collected in helping them with farm management, business profit, operation efficiency and risk management. In contrast, non-collectors viewed the benefits of data less favourably, with average ratings below the mid-point of the scale, across all four benefits (Note: 1= no use/contribution at all).

Discussion

The present study revealed that the collection rates of agricultural data were still low for both the cropping and red meat industries. In addition, there were variations in the collection rates of various agricultural data. For example, collection rates ranged from the highest (yield – 57%) to the lowest (weed pressure – 12%) in cropping industries; and the collection rates ranged from the highest of veterinary medicine record (63%) to the lowest of water use/quality data (22%) in red meat industries.

Table 5. Data collection rates by the red meat industries.

Industry	Soil mapping	Pasture/ vegetation mapping	Individual animal or herd production	Individual animal or herd feeding	On-farm weather station	Animal breeding	Veterinary medicine record	Water use/ quality	None of the listed data is collected
Beef (n=126)	23%	33%	54%	23%	26%	50%	63%	25%	14%
Beef/grain mixed (n=64)	28%	30%	52%	25%	45%	41%	67%	23%	9%
Beef/sheep mixed (n=94)	29%	29%	52%	23%	42%	61%	70%	19%	10%
Lamb (n=59)	20%	24%	44%	29%	49%	54%	63%	17%	19%
Sheep/grain mixed (n=94)	27%	19%	31%	16%	31%	47%	45%	23%	25%
Sheep wool (n=89)	23%	24%	56%	27%	38%	67%	72%	19%	10%
Total (n=526)	25%	27%	49%	23%	37%	54%	63%	22%	14%

Table 6. Mean and standard deviation of reported usefulness of collected data for farm management decisions by data collectors in red meat industries.[1]

Industry	Soil mapping	Pasture/ vegetation mapping	Individual animal or herd production	Individual animal or herd feeding	On-farm weather station	Animal breeding	Veterinary medicine record	Water use/ quality
Beef (n=126)	4.10 (0.94)	3.95 (1.06)	3.93 (1.15)	4.00 (1.17)	3.42 (1.09)	4.27 (0.81)	3.77 (1.20)	3.68 (1.33)
Beef/grain mixed (n=64)	3.61 (0.92)	3.58 (0.90)	3.67 (1.16)	3.50 (1.27)	3.69 (1.07)	4.12 (0.71)	3.58 (1.18)	4.13 (1.06)
Beef/sheep mixed (n=94)	3.52 (1.31)	3.78 (1.01)	3.98 (1.05)	3.77 (0.92)	3.51 (1.19)	4.02 (0.96)	3.65 (1.17)	3.78 (1.11)
Lamb (n=59)	3.92 (0.90)	3.86 (0.77)	4.15 (0.78)	4.00 (0.87)	3.14 (1.09)	4.02 (0.96)	3.70 (1.05)	3.30 (0.83)
Sheep/grain mixed (n=94)	4.12 (1.01)	3.33 (1.09)	3.86 (0.95)	3.80 (0.86)	3.52 (1.35)	4.19 (0.86)	3.45 (1.17)	3.77 (0.97)
Sheep wool (n=89)	3.50 (1.24)	3.67 (1.02)	4.04 (0.97)	3.79 (0.93)	3.44 (1.21)	4.27 (0.89)	3.48 (1.27)	3.24 (0.97)
Total	3.81 (1.10)	3.74 (1.01)	3.94 (1.04)	3.83 (1.01)	3.46 (1.17)	4.15 (0.89)	3.62 (1.18)	3.67 (1.11)

[1] 1 = no use at all; 5 = extremely useful.

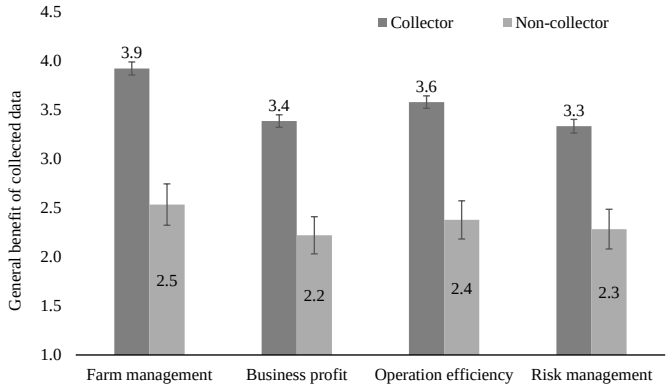


Figure 1. Cropping industry – overall evaluation of data benefit by collectors and non-collectors. Note: Farm management (1 = no use at all, 5 = extremely useful); business profit (1 = no positive contribution at all, 5 = increased profit greatly); Operation efficiency (1 = no positive contribution at all, 5 = increased efficiency greatly); Risk management (1 = no positive contribution at all, 5 = improved risk management greatly).

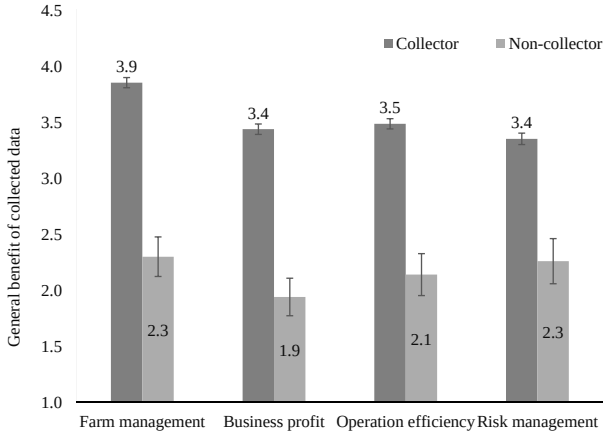


Figure 2. Red meat industry – overall evaluation of data benefit by collectors and non-collectors. Note: Measurement scale as in Figure 1.

These results indicate that there is significant room for improvement in data collection among producers. Many producers may not be fully capitalising on the precision agriculture technologies available to collect data (more) easily.

While data collection rates were relatively low across the board, the value of these data was highly recognised by the producers. While this finding is unsurprising, it does suggest that a concerted effort among all stakeholders to help identify the potential value of agricultural data could be an influential factor in encouraging more widespread data collection. It may be worthwhile learning from the experiences of cotton producers who appear to be on a relatively more advanced adoption pathway. Additionally, further investigation is required to identify the underlying barriers to collecting agricultural data. For example, even if producers perceive the benefits of data, there may exist practical barriers (e.g. lack of time and resources) preventing them from taking action. This is an area that requires further attention.

The findings highlight the importance of communicating and demonstrating the concrete benefits that digital technologies can bring. A concerted effort among all stakeholders is needed to help identify the potential value of on-farm agricultural data applications, and to identify the barriers that may limit data collection practices (including precision agriculture technology adoption). The findings have implications for governments, RDCs and research institutions who are seeking to develop targeted strategies and policies to facilitate the utilisation of agricultural data. The behaviour of producers is critical for securing the large productivity gains that digital agriculture has promised to deliver for the industry.

Conclusions

Collection of agricultural data is essential for precisions agriculture, providing a data-driven management approach to improve productivity and profitability. However, the present research shows that, even though the value of the agricultural data was highly regarded by producers, the data collection rates were still low across cropping and red meat industries in Australia. This broad and in-depth benchmark has great implications for key stakeholders including governments, RDCs and research institutions to develop targeted strategies and policies which will enable producers to better utilise agricultural data for productivity and profitability and capitalise on the opportunities created by digital agriculture and big data.

Acknowledgements

This project is supported by funding from the Australian Government Department of Agriculture and Water Resources as part of its Rural R&D for Profit programme. We acknowledge and thank the 'Precision to Decision' project supporters. We also thank all producers who took part in the survey: Cotton Research and Development Corporation, Meat and Livestock Australia, Dairy Australia, Grains Research and Development Corporation, Sugar Research Australia, AgriFutures Australia, Australian Wool Innovation, Australian Meat Processing Corporation, Australian Livestock Export Corporation Ltd.

References

Bramley, R. G. V. & Ouzman, J., 2019. Farmer attitudes to the use of sensors and automation in fertilizer decision-making: nitrogen fertilization in the Australian grains sector. Precision Agriculture, 20 (1), 157-175.

Griffith, C., Heydon, G., Lamb, D., Lefort, L., Taylor, K., & Trotter, M., 2013. Smart Farming: Leveraging the impact of broadband and the digital economy. New England, Australia: CSIRO and University of New England.

Keogh, M., & Henry, M., 2016. The implications of digital agriculture and big data for Australian agriculture. Research Report, Australian Farm Institute, Sydney, Australia.

Miller, N. J., Griffin, T. W., Bergold, J., Ciampitti, I. A., & Sharda, A., 2017. Farmers' adoption path of precision agriculture technology. In Taylor, J. A., Cammarano, D., Prashar, A. & Hamilton, A. (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences 8 (2) 708-712).

Robertson, M. J., Llewellyn, R. S., Mandel, R., Lawes, R., Bramley, R. G. V., Swift, L., *et al.,* 2012. Adoption of variable rate technology in the Australian grains industry: status, issues and prospects. Precision Agriculture, 13, 181-199.

Trindall, J., Rainbow, R., & Leonard, E., 2018. Enabling digital agriculture in Australia. Farm Policy Journal, 15 (1), 1-6.

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M., 2017. Big Data in Smart Farming – A review. Agricultural Systems, 153, 69-80.

Zhang, A., Jakku, E., Llewellyn, R., & Baker, E. I., 2018. Surveying the needs and drivers for digital agriculture in Australia. Farm Policy Journal 15 (1), 25-39.

Section 16
Small farms/small holders and precision agriculture

Evaluating the potential benefits of field-specific nitrogen management of spring maize in northeast China

X. Wang[1], Y. Miao[1,2*], R. Dong[1], Y. Guan[1], D.J. Mulla[2]
[1]*College of Resources and Environmental Sciences, China Agricultural University, Beijing 100193 China, P.R.*
[2]*Precision Agriculture Center, Department of Soil, Water, and Climate, University of Minnesota, St. Paul, MN 55108 USA; ymiao@umn.edu*

Abstract

The objective of this study was to evaluate the potential benefits of field-specific N management (FSNM) in rain-fed maize farming systems of Northeast China. Two farmer's fields (black and sandy soils) were selected to determine the field-specific optimum N application rates. The results indicated that the optimum N rate was 231 and 224 kg/ha for black soil and sandy soil fields, respectively. The FSNM increased the N use efficiency (NUE) by 7-41 and 5-34% relative to farmer N management (FMNM) for black and sandy soil fields, respectively, while maintaining the corresponding N surplus at about 6 and 59 kg/ha.

Keywords: optimum N application rate, field-specific management, soil type

Introduction

China is the world's largest producer, consumer and importer of chemical fertilizers (Miao *et al.,* 2011; Zhang *et al.,* 2013). Overuse of nitrogen (N) has resulted in very serious environmental problems, threatening the sustainability of Chinese agriculture and economy (Norse and Ju, 2015; Strokal *et al.,* 2016). There is an urgent need to improve N management in China (Miao *et al.,* 2011; West *et al.,* 2014; Zhang *et al.,* 2015). The first step taken by Chinese scientists to improve N management is to develop regional optimum N rates (RONR) to avoid significant over- or under-application problems (Cui *et al.,* 2013; Cao *et al.,* 2017). Such strategies can be easily adopted by farmers and don't involve any additional costs. However, due to the significant field-to-field variability of indigenous soil N supply and crop N demand, this fixed rate and time strategy will unavoidably result in sub-optimal N management in different fields or regions of a field (Cao *et al.,* 2012; Cui *et al.,* 2013). There is a growing interest in China to adopt precision N management (PNM) strategies for different scales of farming systems to further improve N use efficiency (NUE) by matching N supply with crop N requirement in both space and time. It is necessary to evaluate the potential benefits of different PNM strategies first.

Northeast China, with an arable land area of 18.2 Mha, is the largest maize production region in China, accounting for more than 30% of the nation's total maize production (Sun *et al.,* 2016). The typical farmer practice in this region is to apply about 300 kg/ha N before planting with a plant density of 55,000-60,000 plants/ha (Chen *et al.,* 2014; Cui *et al.,* 2018). The regional best management practice (RBMP) recommends about 200 kg/ha N fertilizer and 70,000 plants/ha planting density, which can improve grain yield and farmer's income while reducing the reactive N losses and greenhouse gas emissions compared with farmer's management practices in this region (Cui *et al.,* 2018). The average field size in this region is about 1-2 ha, field-specific management will be a practical PNM strategy for this scale of farming.

So far, little has been reported about the potential benefits of field-specific N management (FSNM) in China. This study took the medium scale spring maize (*Zea mays* L.) farming system in Northeast China as a case study. The objectives of this study were to: (1) determine the field-specific optimum

DOI 10.3920/978-90-8686-888-9_108, © Wageningen Academic Publishers 2019

N application rates; (2) evaluate the potential benefits of FSNM for the spring maize production in Northeast China.

Materials and methods

Site description

This study site is located in Lishu County (43°02'-43°46'N, 123°45'-124°53'E), Jilin Province. Two farmer's fields with contrasting soil types were studied, a black soil equivalent to a typic Haploboroll and a sandy soil equivalent to a typic Cryopsamment according to the USDA Soil Taxonomy (Keys to Soil Taxonomy). This study site is located in the North Temperate Zone, which is characterized by a semi-humid continental monsoon climate and is a major region for spring maize production in Northeast China. The mean annual temperature is 6.6 °C, the annual sunshine is 2,656 h, the frost-free period is 142 days each year, and the annual cumulative temperature (>10 °C) is 3,056 °C. The annual average precipitation is 556 mm, with about 84% occurring during the crop-growing season from May to September.

Field experiments

The same field experiment was conducted from 2015 to 2017 in the black soil and sandy soil fields, covering 1 ha in each field. The experiment used a two-factor randomized complete block design with three replicates involving six N rates (from 0 to 300 kg N/ha for maize with an increment of 60 kg N/ha) and three planting densities (D1: 55,000 plants/ha, D2: 70,000 plants/ha, D3: 85,000 plants/ha) in each field. For every N treatment, one-third of the N fertilizer and all the phosphorus (90 kg P_2O_5/ha) and potassium (90 kg K_2O/ha) fertilizers were blended into the top 0.2 m soil as basal fertilizers before planting, and the remaining N fertilizer was applied as side-dressing at the V8 growth stage. Because natural precipitation was enough to meet the growth demand for the maize growth on the black soil field, no irrigation was applied during the study period. However, maize at the sandy soil field always suffered from drought stress and irrigation of about 50 mm of water was applied once before the anthesis growth stage each year. All plots were kept free of weeds, insects and diseases with chemicals based on standard practices. Each experiment used the same local maize variety (Liangyu 66).

Sample collection and data calculation

At maize harvest stage, three plant samples were randomly selected from each plot and split into stalks, leaves and grains. All plant samples were dried in the oven at 105 °C for 1 h and then at 85 °C to a constant weight to determine the dry biomass. They were ground into fine powder, respectively, to determine N concentration by the Kjeldahl digestion method (Nelson and Sommers, 1962). After sampling, grain yield was determined by harvesting the middle 20-m^2 area of each plot, standardized to 14% grain moisture content. After that, the N use efficiency (NUE) – partial factor productivity (PFP), agronomy efficiency (AE), and recovery efficiency (RE) were calculated using the equations:

$$PFP\ (kg/kg) = Y_N / N_F \quad (1)$$

$$AE\ (kg/kg) = (Y_N - Y_0) / N_F \quad (2)$$

$$RE\ (\%) = 100 \times (PNU_N - PNU_0) / N_F \quad (3)$$

where Y_N and Y_0 are the grain yield in N fertilizer application plots and N_0 (control) plots, respectively, and PNU_N and PNU_0 are the plant N uptake in N application plots and N0 plots, respectively, and N_F is N from the applied fertilizer rate. The N surplus (N_S) was defined as N fertilizer application rate (N_F) minus plant N uptake (PNU).

The economic income, defined as marginal return to N (E, $/ha), was calculated according to the formula:

$$E = Y \times P_Y - N \times P_N \quad (4)$$

where Y is the grain yield (kg/ha), P_Y is the grain price ($/kg), N is the N fertilizer application rate (kg/ha), P_N is the N fertilizer price ($/kg). The prices of maize grain and N fertilizer were 0.31 and 0.62 $/kg, respectively, in China (Cui *et al.*, 2018).

Statistical analysis

Quadratic plus plateau (Q+P) model was selected to describe maize grain yield and marginal return response to N rate. The PROC NLIN procedure in SAS (Version 8.0, SAS Institute Inc., Cary, NC, USA) was used for the analysis. Analysis of variance (ANOVA) was conducted using the general linear model procedure in SPSS (Version 19.0, SPSS Inc., Chicago, IL, USA).

Results

The differences in field-specific optimum nitrogen rates

In this study, the lowest N application rates that would obtain the maximum grain yield or marginal return were defined as agronomic optimum N rate (AONR) or economic optimum N rate (EONR), respectively. The N application rate which leads to the soil N being in balance (N surplus was 0 kg/ha) was defined as the N balance optimum N rate (NONR). The grain yield, marginal return, N surplus, and NUE were calculated for the different N application rates (Farmer N rate (FMNR), RONR, AONR, EONR, and NONR) for each field, respectively (Figure 1).

Compared with the AONR and EONR, the NONR (231 kg/ha) produced similar yield and marginal return in the black soil field, while reducing N fertilizer by 10-14% and keeping the soil N in balance (N surplus=0 kg/ha). Thus, NONR was used for this field as the optimum N rate (Figure 1A). For the sandy soil field (Figure 1B), although the NONR kept the soil N in balance (N surplus=0 kg/ha), it significantly decreased the grain yield by 12-14% and marginal return by about 7% compared with AONR and EONR. The EONR (224 kg/ha), with a 20% reduction in N application rate, produced similar grain yield and marginal return as the AONR, and was used as the optimum N rate for the sandy soil field.

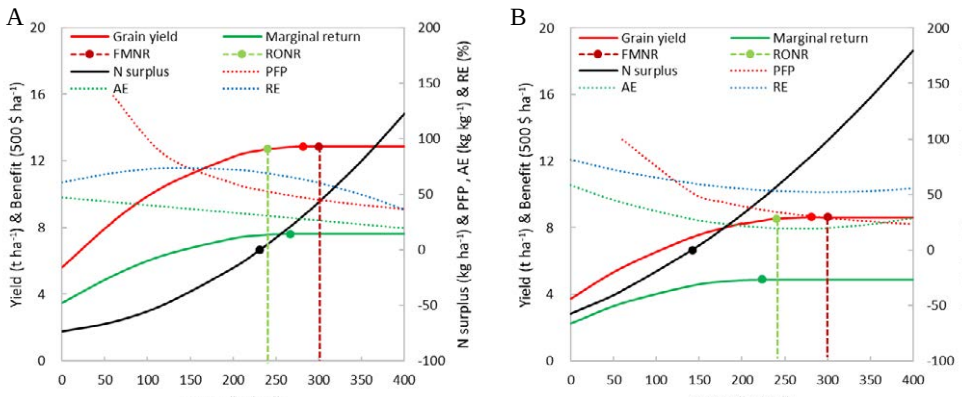


Figure 1. Comparison of different optimum N application rates based on different targets for different soil types (A) black soil; (B) sandy soil. (solid dots indicate the highest point, or N surplus=0 kg/ha). FMNR is farmer N rate. RONR is regional optimum N rate. PFP is partial factor productivity. AE is agronomic efficiency, and RE is recovery efficiency.

The effects of field-specific optimum N application rates

The comparison of FSNM with the farmer N management (FMNM) and the regional optimum N management (RONM) for two soil fields across three years are illustrated in Table 1. In the black soil field, the FSNM decreased the N application rates by about 23% and increased the NUE by about 7-41% compared with FMNM. It decreased the N application rates by about 4% and increased the NUE by about 3% compared with RONM. In addition, the FSNM produced similar grain yield and marginal return as FMNM and RONM, while reducing N surplus to 6 kg/ha. In the sandy soil field, the FSNM decreased the N application rates by about 25% and increased the NUE by about 5-34% compared with FMNM, respectively, while decreasing the N application rates by about 7% and increasing the NUE by about 9-11% compared with RONM, respectively. The FSNM produced similar grain yield and marginal return as FMNM and RONM and decreased the N surplus by 48 and 8% compared with FMNM and RONM, respectively.

Discussion

During the recent decades, to increase farmers' income, N application rates have been optimized to maximize economic returns by considering both grain yield price and N application cost (Xia and Yan, 2011; Calvo *et al.,* 2014). The economic optimum is reached when the marginal cost of the N fertilization corresponds to the marginal revenue, that is, when the returns above the N fertilizer cost are maximized (Hernandez and Mulla, 2008). If the goal is to optimize economic return, the optimum N rate would be decreased in comparison with AONR by 5 and 20% in the black soil and sandy soil fields, respectively. Therefore, the lower N application rate of EONR than AONR resulted in a significant decrease in N surplus and increase in NUE, especially for the sandy soil field (Figure 1). This would allow producing similar grain yield, and economic returns as with AONR with less N fertilizer and fewer negative environmental impacts. In China, the EONR rates appear superior to AONR as an N recommendation strategy for farmers.

Currently, high input and low use efficiency of resources in crop (or food) production has led to the deterioration of environment (Zhang *et al.,* 2018). From the environmental perspective, assessing and controlling active N (Nr) releases should become one of the primary focuses of crop production (Leach *et al.,* 2012; Cui *et al.,* 2013). The N surplus can be used as an index of the efficiency of N management in cropping systems (Van Groenigen *et al.,* 2003; Cui *et al.,* 2013; Zhang *et al.,* 2015). In this study, both the black and sandy soil fields had a significant increase in N surplus with increases in N application rate. When considering the soil N balance (N surplus was 0 kg/ha), the optimum N application rate was significantly decreased compared with AONR by about 18 and 49% in the black and sandy soil fields, respectively (Figure 1). Therefore, basing N fertilizer applications on

Table 1. Comparison of the average N rate (kg/ha), grain yield (t/ha), marginal return ($/ha), N surplus (kg/ha), partial factor productivity (PFP) (kg/kg), agronomic efficiency (AE) (kg/kg), and recovery efficiency (RE) (%) from different N managements for two soil types across three years.

Field	N management	N rate	Grain yield	Marginal return	N surplus	PFP	AE	RE
Black soil	FMNR	300.0	12.21	3,598.6	41.5	40.7	20.7	0.62
	RONR	240.0	12.83	3,827.6	11.4	53.4	28.4	0.66
	FS-NONR	231.1	12.76	3,813.8	5.9	55.2	29.3	0.67
Sandy soil	FMNR	300.0	8.03	2,303.6	104.6	26.8	14.9	0.46
	RONR	240.0	7.90	2,300.3	63.7	32.9	18.1	0.50
	FS-EONR	223.5	8.03	2,351.0	58.5	35.9	20.0	0.48

NONR after simultaneously considering grain yield, economic benefit and environmental impact will contribute to both food security and sustainable development. It should be noted, however, that for the sandy soil field, applying the NONR would result in lower yield and marginal returns. Since environmental benefits are not currently reflected in monetary terms, farmers would not adopt the NONR approach unless the government provides them with financial incentives.

The commonly adopted practical RONM strategy had good potential to improve N management in China (Cui *et al.,* 2013; Wu *et al.,* 2014). However, N fertilizer is applied at a fixed rate and timing without accounting for spatial and temporal variability in soil N supply and crop N demand. The FSNM strategy, which considered differences in soil type, could decrease the N application rates by about 24% and increase the NUE by about 5-41% compared with FMNM respectively. It decreased the N application rates by about 5% and increased the NUE by about 2-11% compared with RONM. Grain yield and marginal return were not reduced. This strategy, however, does not consider year-to-year weather variability, which can significantly influence the optimum N rates (Miao *et al.,* 2006). Future studies are needed to use proximal and remote sensing technologies to diagnose maize N status and guide in-season FSNM (Li *et al.,* 2014; Xia *et al.,* 2016). Crop growth modeling can also be used to determine field (soil)-specific N rates based on long-term simulations or guide in-season FSNM using current season and historical weather data (Miao *et al.,* 2006; Puntel *et al.,* 2018).

Conclusions

With development of agriculture in China, concepts of high grain yield and high economic income with the aim of minimizing environmental cost has become widely accepted. Therefore, the trade-offs among agronomic, economic and environmental factors need to be understood to determine the optimum N application rate during crop production. Due to the spatial variation in soil type, the optimum N application rates, aiming at different production purposes, should be adjusted on a site-specific basis. In summary, the findings suggest that the determination of optimum N application rate should consider the trade-offs between agronomy, economy, and environment for specific soil types. Compared with farmer N management, the field-specific N management strategy showed good potential to significantly increase NUE and reduce N surplus, but needs to move toward field-specific in-season N management to further improve NUE relative to regional optimum N to support spring maize production in Northeast China.

Acknowledgements

This project was funded by the National Key Research and Development Program of China (2016 YFD0200600, 2016YFD0200602), National Basic Research Program (2015CB150405), the UK Biotechnology and Biological Sciences Research Council (BB/P004555/1), and Norwegian Ministry of Foreign Affairs (SINOGRAIN II, CHN-17/0019).

References

Calvo, N., Rozas, H., Echeverría, H., Diovisalvi, N. (2014). Using canopy indices to quantify the economic optimum nitrogen rate in spring wheat. Agronomy Journal, 107 (2) 459-465.

Cao, Q., Cui, Z., Chen, X., Khosla, R., Dao, T., Miao, Y. (2012). Quantifying spatial variability of indigenous nitrogen supply for precision nitrogen management in small scale farming. Precision Agriculture, 13, 45-61.

Cao, Q., Miao, Y., Feng, G., Gao, X., Liu, B., Liu, Y., *et al.* (2017). Improving nitrogen use efficiency with minimal environmental risks using an active canopy sensor in a wheat-maize cropping system. Field Crops Research, 214, 365-372.

Chen, X., Cui, Z., Fan, M., Vitousek, P., Zhao, M., Ma, W., *et al.* (2014). Producing more grain with lower environmental costs. Nature, 514(7523), 486.

Cui, Z., Yue, S., Wang, G., Meng, Q., Wu, L., Yang, Z., *et al.* (2013). Closing the yield gap could reduce projected greenhouse gas emissions: a case study of maize production in China. Global Change Biology, 19(8), 2467-2477.

Cui, Z., Zhang, H., Chen, X., Zhang, C., Ma, W., Huang, C., *et al.* (2018). Pursuing sustainable productivity with millions of smallholder farmers. Nature, 555(7696).

Hernandez J., Mulla, D. (2008). Estimating uncertainty of economically optimum fertilizer rates. Agronomy Journal, 100(5), 1221-1229.

Leach, A., Galloway, J., Bleeker, A., Erisman, J., Kohn, R., Kitzes, J. (2012). A nitrogen footprint model to help consumers understand their role in nitrogen losses to the environment. Environmental Development, 1 (1), 40-66.

Li, F., Miao, Y., Feng, G., Yuan, F., Yue, S., Gao, X., *et al.* (2014). Improving estimation of summer maize nitrogen status with red edge-based spectral vegetation indices. Field Crops Research, 157, 111-123.

Miao, Y., Mulla, D., Batchelor, W., Paz, J., Robert, P., Wiebers, M. (2006). Evaluating management zone optimal nitrogen rates with a crop growth model. Agronomy Journal, 98(3), 545-553

Miao, Y., Stewart, B., Zhang, F. (2011). Long-term experiments for sustainable nutrient management in China. A review. Agronomy for Sustainable Development, 31(2), 397-414.

Nelson, D., Sommers, L. (1962). Determination of total nitrogen in plant material. Agronomy Journal, 65(1), 423-425.

Norse, D., Ju, X. (2015). Environmental costs of China's food security. Agriculture Ecosystems & Environment. 209, 5-14.

Puntel, L.A., Sawyer, J. E., Barker, D. W., Thorburn, P. J., Castellano, M. J., Moore, K. J., *et al.* (2018). A system modeling approach to forecast corn economic optimum nitrogen rate. Frontiers in Plant Science, 9, 436. doi: 10.3389/fpls.2018.00436.

Strokal, M., Ma, L., Bai, Z., Luan, S., Kroeze, C., Oenema, O., *et al.* (2016). Alarming nutrient pollution of Chinese rivers as a result of agricultural transitions. Environmental Research Letters, 11, 024014.

Sun, Z., Li, Z., Lu, X., Bu, Q., Ma, X., Wang, Y. (2016). Modeling soil type effects to improve rainfed corn yields in Northeast China. Agronomy Journal, 108, 498-508. doi:10.2134/agronj2015.0232.

Van Groenigen, K., Six, J., Harris, D., Blum, H., Van Kessel, C. (2003). Soil ^{13}C–^{15}N dynamics in an N_2-fixing clover system under long-term exposure to elevated atmospheric CO_2. Global Change Biology, 9(12), 1751-1762.

West, P., Gerber, J., Engstrom, P., Mueller, N., Brauman, K., Carlson, K., *et al.* (2014). Leverage points for improving global food security and the environment. Science, 345, 325-328.

Wu, L., Chen, X., Cui, Z., Zhang, W., Zhang, F. (2014). Establishing a regional nitrogen management approach to mitigate greenhouse gas emission intensity from intensive smallholder maize production. PLoS One, 9, e98481.

Xia, T., Miao, Y., Wu, D., Shao, H., Khosla, R., Mi, G. (2016). Active optical sensing of spring maize for in-season diagnosis of nitrogen status based on nitrogen nutrition index. Remote Sensing, 8 (7), 605.

Xia, Y., Yan, X. (2011). Comparison of statistical models for predicting cost effective nitrogen rate at rice-wheat cropping systems. Soil Science & Plant Nutrition, 57 (2), 320-330.

Zhang, W., Dou, Z., He, P., Ju, X., Powlson, D., Chadwick, D., *et al.* (2013). New technologies reduce greenhouse gas emissions from nitrogenous fertilizer in China. Proceedings of the National Academy of Sciences of the United States of America, 110(21), 8375-8380.

Zhang, X, Davidson, E., Mauzerall, D., Searchinger, T., Dumas, P., Shen, Y. (2015). Managing nitrogen for sustainable development. Nature, 528(7580), 51-59.

Zhang, Y., Wang, H., Lei, Q., Luo, J., Lindsey, S., Zhang, J., *et al.* (2018). Optimizing the nitrogen application rate for maize and wheat based on yield and environment on the Northern China Plain. Science of the Total Environment, 550 (7677), 469.

Combining crop modelling and remote sensing to create yield maps for management zone delineation in small scale farming systems

H. Zha[1], D. Cammarano[2*], L. Wilson[3], Y. Li[1], W.D. Batchelor[4] and Y. Miao[1,5]
[1]*China Agricultural University, 2 Yuanmingyuan W. Rd, 100193, Beijing, China, P.R.*
[2]*James Hutton Institute, Errol Road, Invergowrie, DD25BQ, United Kingdom; davide.cammarano@hutton.ac.uk*
[3]RSK ADAS Ltd., 172 Chester Road, WA60AR, Helsby, United Kingdom
[4]*Auburn University, 200 Tom Corley Building, 36849 Auburn, USA*
[5]*University of Minnesota, 1991 Upper Buford Circle, 55108 St Paul, USA*

Abstract

The objective of this study was to develop a management zone delineation strategy using reconstruction of yield maps with crop modelling and remote sensing data. Wangzhuang Village in Quzhou County (China), with an average farm size of 0.3 to 0.5 ha, was selected for this study. Satellite images, detailed soil nutrient samples from fields in the village and satellite-derived soil brightness images were available. Their integration allowed for the delineation of management zones at the village level.

Keywords: management zones, remote sensing, crop model, fertilization

Introduction

Small scale farms represent about 80% of the farming area of China and given the challenge of producing more food on the same amount (or less) of land without polluting the environment, such areas are faced with tough challenges. Farmers manage their fields by experience and need science-based evidence to make the system more efficient. The mismanagement of nitrogen (N) fertilizer is a known problem in the small holder farming systems of China (Miao *et al.*, 2011). To optimize agronomic inputs such as N, an approach that matches N application rates to the soil and crop needs is required (Nawar *et al.*, 2017). In many larger-scale arable systems around the world, the fields are divided into management zones (MZs); sub-regions within the field that have similar combinations of yield-limiting factors and are managed accordingly (Miao *et al.*, 2018). Jin *et al.* (2017) combined a crop simulation model with remotely sensed data to map yield heterogeneity on smallholder farms in East Africa. However, the small size of single farms in the North China Plain (NCP) does not allow for cost-effective on-farm management using farm scale MZs. One solution would be to consider the whole village as a single farm unit and divide it into several MZs, since fertilizer management practices at the Village scale are relatively uniform because of the co-operative structure (Zhang *et al.*, 2016). Science and Technology Backyards (STB) were established originally in Quzhou County in 2009 by China Agricultural University to carry out specific research and extension services aimed at transferring research technology to smallholder farmers (Zhang *et al.*, 2016). Using the network of research and extension facilities provided by the STB, the overall aim of this study was to define MZs at the village scale by integrating historical satellite images, constructed yield estimations using crop modelling, and soil information derived from remote sensing.

DOI 10.3920/978-90-8686-888-9_109, © Wageningen Academic Publishers 2019

Materials and methods

Site description

The village chosen for the study was Wangzhuang village in Quzhou County in the NCP (30°51'55"N; 115° 02' 06' E). The cropped area of the village covered approximately 177 ha and comprised a total of 540 fields, with an average farm size of 0.33 ha. The main cropping system of the village was winter wheat (sown in October) and summer maize (sown in mid-June). The soil was classified as silty-loam.

Overall methodological approach

Satellite data for the village were obtained from the Landsat satellites (30 m pixels) for 2008-2018 using the USGS explorer website (USGS, 2018). Atmospherically-corrected cloud-free scenes from the end of April/beginning of May were selected to coincide with flowering in winter wheat. The green normalized difference vegetation index (GNDVI) was calculated for each scene. Soil brightness (SOB) data at 3 m spatial resolution were purchased from Courtyard Agriculture Ltd. and were derived from optical satellite images from the RapidEye satellite at a time in the season when the land was un-cropped (bare soil). Most of the fields were contiguous, with 'boundaries' being narrow ridges of bare soil. There was little other vegetation or paths/roads present. Therefore, most of the pixels corresponded to cropland (Figure 1).

Soil N was measured for the top 200 mm within each field prior sowing of wheat in 2015. Soil organic carbon (OC) within the top 200 mm was sampled on a regular grid in 2015, prior sowing of wheat, for a total of 40 points across the village. Spatial soil texture and OC information were also available in digital format (He *et al.*, 2009; Li and Hu, personal communication).

The CSM-CERES-Wheat v4.7 crop simulation model distributed with the Decision Support System for Agrotechnology Transfer (DSSAT, Jones *et al.*, 2003) was used in this study to simulate wheat yields. The crop model was previously evaluated on experimental data on N responses reported by Cao *et al.* (2017). The input data for the model were soil texture, OC and N, daily weather data (maximum and minimum temperature, solar radiation and precipitation). The soil physical properties were obtained from the work of He *et al.* (2009) and the chemical properties were obtained from the STB soil survey that comprised a total of 540 samples, one per field. The model was set up to run for 30 years of weather data to generate 30 m-spatial maps. To represent the range of management conditions and initial soil water conditions it was run for each field in a factorial combination with 2 levels of initial soil water content (high and low) before sowing; 3 sowing dates (1 Sep, 1 Oct, 1 Nov); and 2 N fertilization levels (a total of 100 kg N/ha and a total of 300 kg N/ha reflecting the farmers' practice). Irrigation amount was applied following the common farmers'



Figure 1. Overview of the fields in Wangzhuang village, China. The fields were separated only by a thin bare patch of soil. The distribution of the wheat plants highlights how planting happened at the same time across the fields.

practices as discussed in Cao *et al.* (2017). Therefore, for each field 360 simulated yield values were generated. For each year and each field, the simulated yield values were averaged to obtain a 'mean' yield value representing the range of management/initial conditions for a given year. The practice of simulating factorial yield combinations, studying the gaussian distribution of the simulated outputs and taking the mean value as indicative of the yield expression under average conditions has been applied to the simulation of food security of small holders farming systems of Sub-Saharan Africa and South East Asia (Rosenzweig and Hillel, 2015).

Spatial and temporal variability in GNDVI and simulated yield were quantified following the method of Basso *et al.* (2012). The spatial variability of GNDVI from 2008 to 2018, and simulated yield over 30 years, was calculated from the relative percentage difference of GNDVI/ simulated yield at each 30 m pixel or field from the average GNDVI/ simulated yield obtained for the whole cropped area of the village, according to the following equation:

$$\bar{\sigma}_{si}^2 = \frac{1}{n}\sum_{k=1}^{n}\left[\frac{y_{i,k}-\bar{y}_k}{\bar{y}_k} \times 100\right] \quad (1)$$

Where n is the total number of available years, $k = 1\ldots,n$ is each year, $\bar{\sigma}_{si}^2$ is the average percentage difference at location i, $\bar{y}_k$ is the average of the variable obtained for the whole village at year k, $y_{i,k}$ the variable monitored at location i at year k.

The temporal variability of patterns in the GNDVI and simulated yield data was calculated as the temporal variance (variable value monitored at each location minus the whole village mean). The variance calculation was computed by:

$$\bar{\sigma}_{ti}^2 = \frac{1}{n}\sum_{k=1}^{n}\left(y_{i,k} - \bar{y}_{i,n}\right)^2 \quad (2)$$

Where $\bar{\sigma}_{ti}^2$ is the temporal variance value at location i, $\bar{y}_{i,k}$ is the value of the variable monitored at location i at year k, $y_{i,n}$ is the average variable value at location i over the n years.

The GNDVI variability metrics and the soil brightness images were summarised at the field scale for consistency with the representation of simulated yield and soil chemical properties.

For site-specific agronomic management of any input there is the need to develop MZs that will be subject to a unique combination of potential yield-limiting factors (Fraisse *et al.*, 2001). Management zones will be most effective if the variation (of the factors under consideration) within them is minimised, and there are a manageable number of zones. The partitioning around medoids (PAM) method was used (Kaufman and Rousseeuw, 1990), which is a more robust alternative to k-means clustering when noisy data are expected, as is the case in this study due to unmeasured variation in management. The term medoid refers to an observation within a cluster for which the sum of the distances between it and all the other members of the cluster is a minimum. PAM requires *a-priori* information of the number of clusters, but it does more computation than the more commonly used k-means clustering to ensure that the medoids it finds are truly representative of the observations within a given cluster.

Field-scale variables for (1) spatial and temporal variability in GNDVI; (2) spatial and temporal variability in simulated yield; (3) mean soil brightness and (4) spatial coordinates of field centroids were normalised and used as variables with the PAM method to define clusters.

Relative variance (RV) of measured soil properties at the field scale was used to evaluate the accuracy of this approach for delineating MZs (Miao *et al.,* 2018), given by;

$$RV = 1 - S_w^2/S_T^2 \quad (3)$$

Where σS_W^2 is the total within-zone variance of the soil property and S_T^2 is the total village level variance of the corresponding property. RV approximates the amount of variability explained by the MZ delineation and can be interpreted similarly to the R^2 value in a regression analysis in terms

of the proportion of variation explained. RV was calculated on a per-field basis for measured total N and OC because they best characterise areas that would benefit from differential management.

Results

The spatial pattern of the LandSat GNDVI image and the soil brightness averaged to field scale are shown in Figure 2A and 2B, respectively. The spatial and temporal variability of the GNDVI time-series are shown in Figure 3A and 3B, respectively.
The spatial distributions of measured soil N and OC throughout the village are shown in Figure 4A and 4B respectively. Soil N varied between 0.68 and 1.37% (Figure 4A). The soil OC varied between

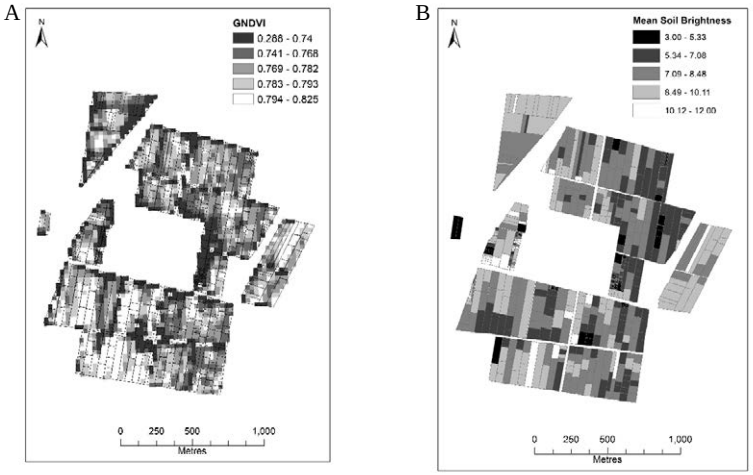


Figure 2. GNDVI map from the LandSat 8 satellite (30 m pixel resolution) taken at flowering in 2016, overlaid with field boundaries (A) and soil brightness for 2017 averaged to field-scale (B).

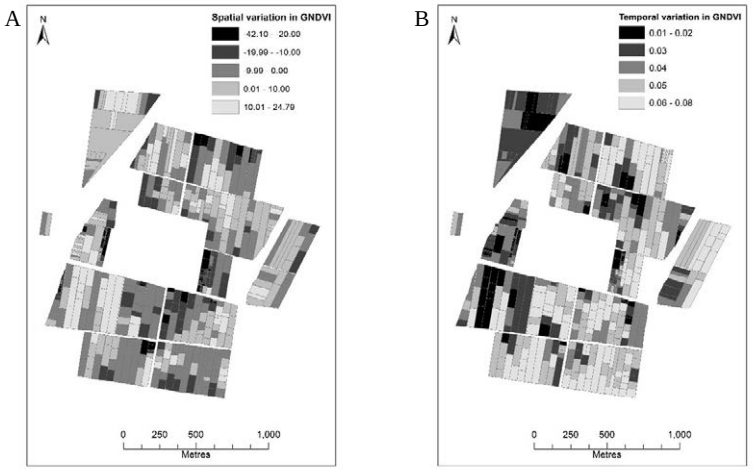


Figure 3. Spatial variability (A) and temporal variance (B) of 10-year time series of GNDVI.

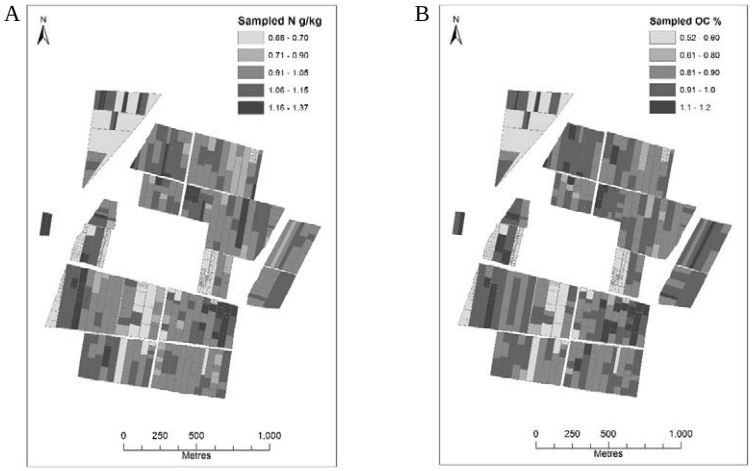


Figure 4. Measured soil total N (g/kg) (A) and soil organic carbon (%)(B).

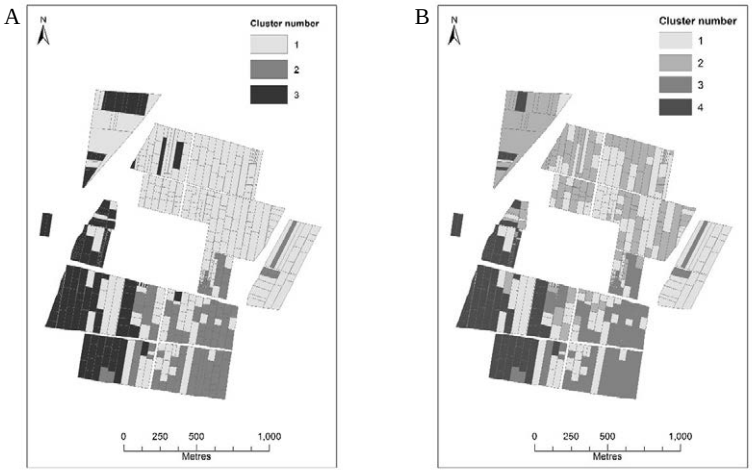


Figure 5. Clusters obtained using remote sensing approach specifying a three-cluster solution (A) and a four-cluster solution (B).

0.52 and 1.2% and its spatial variability matched the soil N content (Figure 4B). The results of using soil brightness, variability in GDNVI derived from remote sensing and variability in simulated yield maps for clustering the village into several MZs are shown Figure 5A and 5B, specifying three and four clusters, respectively. The three-cluster solution divided the village into relatively spatially-coherent zones, whereas the four-cluster solution classified fields in the original cluster 1 into two separate clusters with less spatial coherence.

The RV of measured soil variables total N and OC within these clusters compared to the village mean is shown in Table 1. The combination of remote sensing data and modelling was a poor predictor of measured N and OC in this study. However the four-cluster solution explained more of the variability in measured soil properties than the three-cluster solution.

Table 1. Relative variance of measured soil variables total N and organic carbon at the field scale and the spatial and temporal variability in GNDVI and simulated yield for each of the tested clustering solutions.

Clustering approach	Relative variance (RV)	
	Total N	Organic carbon
Remote + model 3 zones	1.0	1.9
Remote + model 4 zones	4.7	6.5

Discussion

Overall, the integration of variables derived from remote sensing and crop modelling allowed for a discrimination of zones within the field, albeit with poor explanation of the variability in measured N and OC. This is likely to be due to the very narrow range in these measured parameters within the village, and unmeasured variability in field management and inputs. Satellite images of crops provide an indirect tool to obtain spatial information of crop growing conditions for a given year and therefore are a good tool to quantify spatial field variability; with a long enough time-series also enabling a quantification of temporal stability in this spatial variation.

The subdivision of the field into 4 zones improved the explanation of the variability in measured soil parameters; however, there is a trade-off between number of zones and site-specific management. The spatial coherence of the zones also needs to be considered, since it is likely to be more economically and practically efficient to zone the village into 'blocks' rather than by individual field. It has been found in the literature that 3 zones are a common number adopted in commercial precision agriculture solutions regardless of the size of the field (Basso *et al.*, 2012). More zones could have been defined; but this would translate into more management recommendations and organisation of farmers into more co-operative 'clusters', which would add to the complexity and time commitment for co-operative leads.

The use of generated yield maps made with CSM helped to consider the impact of weather variability on crop growth and yield. The model requires observed and spatially-defined soil information and historical weather data. CSM can generate both spatial and temporal information of crop growing conditions. The satellite data do not consider the variability, or the direct impact of soil N and OC and a discrimination of zones based only on this information might carry some limitation. CSM, on the other hand uses as input soil N and OC; and, those two parameters interact with the climatic conditions to influence daily crop growth.

Results of this study can be considered as a preliminary method based on integration of different data to delineate MZs that could then be refined as uniform management is encouraged within these zones.

Conclusions

Spatial and temporal data of simulated crop growth integrated with crop growth and soil variation measurements from satellite can be used to delineate zones at the village level. In conclusion, satellite data, linked with other spatial information (e.g. local soil sample data) and crop growth models could be used to delineate management zones at village level to help optimise inputs.

Acknowledgements

This project was funded by the ACTNN+ Network, the National Basic Research Program (2015CB150405), and Norwegian Ministry of Foreign Afairs (SINOGRAIN II, CHN-17/0019). We thank Prof. Hu Kelin and Li Baoguo from China Agricultural University and Dr Yong He from the Chinese Academy of Agricultural Science for providing the spatial soil texture information of the Village. We thank the anonymous reviewers for the feedback provided.

References

Basso, B., Fiorentino, C., Cammarano, D., Cafiero, G., Dardanelli, J., 2012. Analysis of rainfall distribution on spatial and temporal patterns of wheat yield in Mediterranean environment. European Journal of Agronomy 41, 52-65.

Cao, Q., Miao, Y., Feng, G., Gao, X., Liu, B., Liu, Y. *et al.*, 2017. Improving nitrogen use efficiency with minimal environmental risks using an active canopy sensor in a wheat-maize cropping system. Field Crop Research 214, 365-372.

Fraisse, C., Sudduth, K.A., Kitchen, N.R., 2001. Delineation of Site-Specific Management Zones by Unsupervised Classification of Topographic Attributes and Soil Electrical Conductivity. Transactions of the ASAE 44, 155-166.

He, Y., Chen, D., Li, B.G., Huang, Y.F., Hu, K.L. Li, Y. *et al.*, 2009. Sequential indicator simulation and indicator kriging estimation of 3-dimensional soil textures. Australian Journal of Soil Research 47, 622-631.

Jin, Z., Azzari, G., Burke, M., Aston, S., Lobell, D.B., 2017. Mapping smallholder yield heterogeneity at multiple scales in Eastern Africa. Remote Sensing 9 (9) 931. doi: 10.3390/rs909031.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L. A. *et al.*, 2003. The DSSAT cropping system model. European Journal of Agronomy 18, 235-265.

Kaufman L, Rousseeuw PJ. Partitioning Around Medoids (Program PAM). In: Finding Groups in Data: An Introduction to Cluster Analysis. Hoboken, NY, USA: Wiley. p. 1990.

Miao, Y., Mulla, D., Robert, P.C., 2018. An integrated approach to site-specific management zone delineation. Frontiers of Agricultural Science and Engineering 5, 432-441.

Miao, Y., Stewart, B.A., Zhang, F., 2011. Long-term experiments for sustainable nutrient management in China. A review. Agronomy for Sustainable Development 31, 387-414.

Nawar, S., Corstanje, R., Halcro, G., Mulla, D., Mouazen, A.M., 2017. Delineation of Soil Management Zones for Variable Rate Fertilization: A review. Advances in Agronomy 143,175-245.

Rosenzweig, C., Hillel, D., 2015. Handbook of Climate Change and Agroecosystems: The Agricultural Model Intercomparison and Improvement Project (AgMIP) Integrated Crop and Economic Assessment. World Scientific Press. https://doi.org/10.1142/p970

United States Geological Survey (USGS), 2018. Earth Explorer (https://earthexplorer.usgs.gov/) verified Nov 2018.

Zhang, W., Cao, G., Li, X., Zhang, H., Wang, C., Liu, Q., *et al.*, 2016. Closing yield gaps in China by empowering smallholder farmers. Nature 537, 671-674.

Section 17
On farm experimentation with site-specific technologies

Operating Farmer Innovation Groups (FIGs) for testing yield enhancing ideas using on-farm experimentation

C. Clarke[1*], D. Brightman[2], S. Roques[1], S. Kendall[3], D. Kindred[1] and R. Sylvester-Bradley[1]
[1]*ADAS, Battle Gate Road, Boxworth, CB23 4NN Cambridge, United Kingdom; christina.clarke@adas.co.uk*
[2]*Gaydon Hill Farm, Gaydon, Warwick, CV35 0HQ Warwickshire, United Kingdom*
[3]*ADAS, Meden Vale, Nottinghamshire, NG20 9PD Mansfield, United Kingdom*

Abstract

Tramline comparisons within agricultural trials can now be analysed statistically, allowing joint science generation by researchers and farmers. Through the Yield Enhancement Network (YEN), multiple Farmer Innovation Groups (FIGs) were set up to both facilitate farmer to farmer knowledge exchange and to replicate tramline trials across farms, to enable implementation of innovative ideas. The FIG presented here is testing the hypothesis that yield enhancement can be achieved in wheat by improving crop nitrogen uptake through a spring potash application. Experience and learnings from the running of this FIG are discussed here along with trial results from the first season.

Keywords: on-farm experimentation (OFE), agronomics, farmer innovation group (FIG), farmer-centric research, tramline trials, yield mapping

Introduction

There is a current disconnect in agricultural research between researchers endeavouring to answer fundamental scientific questions and farmers, together with their advisors, who are undertaking the practice of science at the ground level. To improve progress in crop productivity and sustainable intensification, a more holistic approach is needed which better integrates scientific findings with practice. On-farm experimentation (OFE), undertaken by farmers and facilitated by researchers, can be used as a tool to bridge this gap.

Knowledge from farmers at the local scale is pivotal in making gains in agricultural research due to the variability that exists within fields, farms and regions in terms of soil, topography and microclimate (Steiner, 1998). With the recent advances in precision agriculture, OFE has become more accessible, taking advantage of yield mapping capabilities, global navigation satellite systems (GNSS) and satellite imagery (Piepho and Spilke, 2011; Griffin *et al.*, 2008). These technologies support site-specific management and allow 'innovation in context', proposed by Macmillan and Benton (2014) as the key to achieving improvement in crop productivity.

On-farm trials are currently being used not just to develop new practices but to facilitate joint knowledge generation and exchange by researchers, industry representatives, farmers and advisors. The need to form partnerships between farmers and researchers has been recognised by the European Commission. 'Thematic Networks' and Agricultural Knowledge and Innovation Systems (AKIS) have been formed under the EU's Horizon 2020 programme, with the aim to create networks that facilitate this generation and exchange of knowledge between farmers, advisors and researchers (EIP-AGRI, 2016, 2018).

OFE is taking place within the Yield Enhancement Network (YEN) which originated in 2012 to identify arable innovators and assist them in testing yield enhancing ideas (Sylvester-Bradley and Kindred, 2014). It is run as a yield competition based on both highest yield achieved and highest percentage of biophysical potential achieved; the calculations are based on availability of light and water, to encourage engagement of both higher and lower yielding farms. Interaction within YEN

DOI 10.3920/978-90-8686-888-9_110, © Wageningen Academic Publishers 2019

consists of a large cohort of farmers (270 cereal and 70 oilseed rape farmers took part in 2018) as well as industry representatives and academics.

A programme of experimentation has been set up within the YEN to test yield enhancing ideas through the formation of multiple Farmer Innovation Groups (FIGs). These FIGs are farmer-led and facilitated by a researcher, creating a network of farmers trialling the same idea and exchanging knowledge to guide the decision making process of implementing these innovative ideas on farm. Ingram (2010) demonstrated the value of farmer to farmer interaction and the role it played in improving learnings from experimentation and adaptation of reduced tillage practices. A community of reduced tillage farmers, which extended to researchers and agronomists, allowed farmers to reflect on their problems and learn from each other's experiences. Generally, farmers base the majority of their decisions on tradition and fashion modified by local observations (Sylvester-Bradley, 1991). These FIGs facilitate the sharing of explanatory concepts between farmers and scientists which is important in the progression of reasoned thinking on how yields can feasibly be enhanced. The testing of these yield enhancing ideas is made possible through recently developed methodologies encompassed within agronomics (Sylvester-Bradley *et al.*, 2017). This includes statistical analysis of combine yield map data from tramline trial comparisons which takes into account the spatial variability inherent within fields and ascribes a level of confidence to treatment effects which can be interpreted by both the farmer and researcher (Marchant *et al.*, 2019).

This paper discusses the process behind the formation of FIGs, from the origin of the yield enhancing idea through to the setting up of tramline trial comparisons, and the analysis and discussion of the results using the case study of a spring potash FIG set up in early 2018. This particular FIG is testing the hypothesis that a spring application of plant available potassium will improve crop nitrogen uptake in wheat due to the interaction of these nutrients within the plant (Johnston and Milford, 2012). The importance of the facilitator role and interaction between the researcher(s) and farmer(s) will be examined.

Methods

Formation of a farmer innovation group

The establishment of FIGs within YEN started initially with the formulation of yield enhancing ideas produced by YEN farmers. 'Ideas Labs' were held in January 2017 and June 2018 with the objective of encouraging farmers to undertake trials testing their own yield enhancing ideas. The format of these meetings consisted of an introduction in how to set up a scientifically sound trial with the first step in this process consisting of asking good questions and taking useful measurements. The YEN encourages farmers to think about yield as the capture and conversion of resources, principally light and water, therefore discussions were guided around these concepts. Farmers worked in groups, facilitated by a researcher and/or industry representative to produce yield enhancing ideas which could be tested using tramline trials. These ideas were formed into the YEN ideas treasury (Figure 1). FIGs were then formed by lead farmers who were enthused by the discussions and wanted to take an idea forward to test on their farm. The role of the lead farmer is significant in keeping with the farmer-centric nature of the trials; managing decisions such as treatments, participant recruitment to ensure trial replication, and group liaison. During the initial stages of setting up the FIG, the facilitator had an important role in validating the idea to be tested; relevant literature and the current state of knowledge were summarised, to ensure that the trial was novel and the right questions were being asked. This then led to choosing appropriate metrics and designing tests to answer the question.

The spring potash FIG was formed by a farmer who suspected that despite high potassium (K) indices associated with the heavy clay soils on their farm, there was not sufficient levels to support rapid nitrogen (N) uptake in the spring. Therefore, the trial was setup to study the potential benefit of a spring potash application, to coincide with the main N application. The metrics were yield

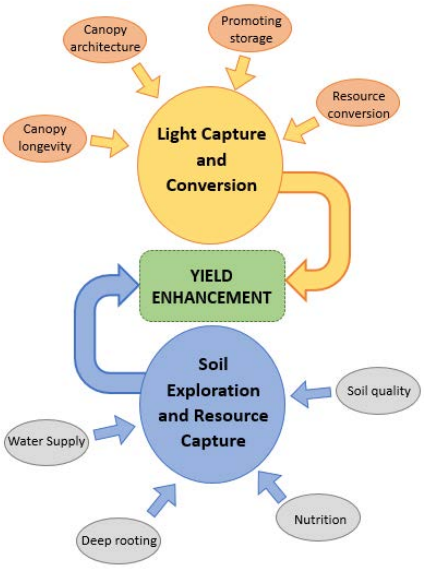


Figure 1. Diagrammatic representation of the YEN 'Ideas Treasury' concepts, which inter-relates yield enhancing concepts originating from farmers.

and crop N and K offtake. The lead farmer enlisted additional farmers into the FIG who were also farming on potassium rich clay soils and were interested in the idea.

Protocol for tramline trial set up and harvesting

In order to undertake a fair test to study treatment effects using tramline comparisons, it is essential that within-field spatial variability and logistics are taken into account when choosing the right field. This includes history of management differences within the field, avoiding fields which contain telegraph poles or trees and is of a sufficient size. Field size should be a minimum of 250 m in length to allow for cutting of headlands and end of combine runs in the data analysis and tramline number is sufficient to allow for the correct number of treatments to be applied. The number of treatments should take into account the ability to replicate the trial; most fields within the UK have around six to eight tramlines. These stipulations were communicated to the participants of the FIG by the facilitator prior to the farmers identifying their trial fields.

The spring potash FIG treatments consisted of 0, 100 and 200 kg/ha of Muriate of Potash (MoP, 60% K_2O) applied using the farmers' spinning disc spreaders, therefore requiring twelve tramlines in total to replicate treatments, due to the need to have a discard tramline in between treated tramlines. This was achieved in three of the four trials but the fourth farmer could not replicate treatments within fields so chose to replicate across fields; this was expected to reduce precision due to the variability introduced between the two fields.

The harvesting procedure was a critical part of the trial planning to ensure accurate yield map data or to maximise the precision of treatment comparisons made using a weighbridge. The facilitator visited each of the tramline trials to discuss with the farmer the method of harvesting. A harvest procedure was discussed with each FIG participant that included: cutting at least two full combine header widths within each treated tramline, calibrating the combine yield monitor and recording how the trial was harvested e.g. to identify which cuts were not full or crossed over treatment boundaries. The trial visit took place after the spring potash treatments were applied, providing an opportunity to confirm trial design with the farmer to aid later data analysis. The longitude and

latitude of each tramline wheeling at both ends of the field were recorded, using a GNSS device, to enable treatment location on the yield map at the data analysis stage.

Agronomics data analysis and reporting results

Georeferenced yield data were taken directly from each FIG participant's combine harvester and cleaned to ensure the data were normally distributed for the analysis. This included discarding data from headlands, ends of combine runs, runs which spanned the boundary between treatments and where the header was not full and extreme values outside the expected yield range (Figure 2). The data then underwent 'offset correction' to find the optimal shift up or down the combine run to correct for the lag in time for grain to reach the yield monitor sensor. The modified dataset was then statistically analysed using Spatial Discontinuity Analysis (SDA; Marchant *et al.*, 2019). This method uses a linear mixed effect model dividing spatial variation of yield into fixed (treatment effects) and random effects (underlying variation or residual variation). Then a model of underlying variation was applied to the data to account for spatial variation across rows and along rows, and for the effect of the treatment. The analysis resulted in a cleaned yield map, arithmetic mean treatment yields, confidence limits and a modelled treatment effect for each site.

To best communicate the statistical analysis of the yield map data, short reports were produced for each FIG participant which summarised: the background of the idea, presented NDVI images of the trial field at significant time points in the season (sourced from satellite imagery: www.datafarming.com.au) to demonstrate potential within-field variation or visible treatment effects and then results were reported both in tabulated form and interpreted. Confidence intervals at 95% were stated for each treatment difference to indicate how researchers define the statistical significance of the result. However, because farmers and scientists have contrasting standards of proof, with farmers not relying on statistical differences to define treatment effects, the probability was also calculated on a 0-1 scale with 0 indicating a certain disadvantage, 1 indicating a certain advantage, and 0.5 indicating no evidence of a difference either way.

Results

Functioning of spring potash FIG

A number of different learnings could be taken from running the spring potash FIG in its first season with most depending on successful communication. Instructions, protocols and requests needed to be clear and concise, sent through an agreed communication method (email, social media, mobile group chats, etc.) and ensured each FIG participant implemented treatments with acceptable

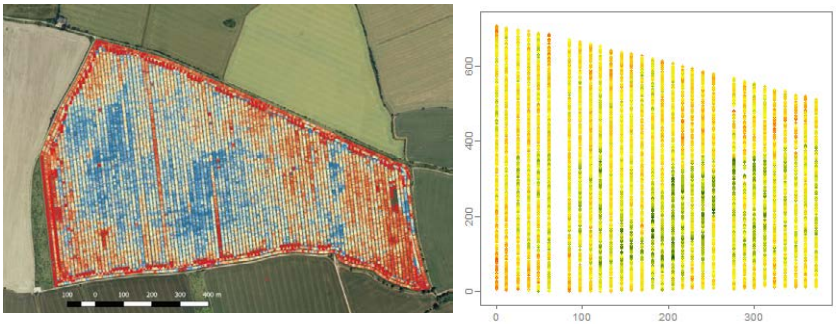


Figure 2. Agronomics analysis. Left: Example of raw yield map data. Right: Example of cleaned yield data, following removal of headlands, anomalous combine runs and both ends of the field. Both images scale is in metres.

consistency between fields. Actions of the farmers determined the success of the FIG findings overall so participants needed to be engaged throughout the season, communicating with the facilitator and other participants to discuss trial details and progress, and exchanging knowledge of outcomes. In the first year the spring potash FIG successfully implemented three replicated tramline trials and a fourth trial replicated across fields. Treatments were consistent throughout the trials and treatment location and yield map data were communicated in a timely fashion. As is the nature of farmer trials, there was an element of risk introduced which is the reliance on the farmer and their equipment to successfully undertake a scientific experiment. The farmer's contractors are also included in this as they need to be aware of the correct protocols for treatment application and harvesting. Within the spring potash FIG, quality of the yield map data was an issue. One of the FIG participants failed to communicate that they did not have access to a yield mapping combine but yield was measured from each treated tramline using a weighbridge and treatment replication allowed for statistical analysis of the data. Additionally, a second FIG participant had implemented a new yield mapping system on their farm which failed to record yields across the whole field trial, resulting in a dataset which was missing a large proportion of the treated area. Despite these circumstances, a sufficient dataset was produced from the first season to take learnings both from the functioning of the FIG and to help answer the initial question of the potential yield benefit of a spring application of potash.

Spring potash FIG trial results

There was no significant effect of spring potash on yield in four of the five trials as a result of applying potash at 100 and 200 kg/ha in the form of MoP (Table 1). A significant treatment effect did occur in the trial where no yield map was obtained and yield was measured, in each replicated treated tramline, using a weighbridge. Yield increases due to MoP were consistent with around 0.25 t/ha for each 100 kg/ha MoP applied. The confidence intervals associated with the treatment differences, stated in Table 1, are indicative of the trial precision. The data from FIG participant 1 was associated with the yield map missing a proportion of the treated tramlines and participant 4a and 4b are un-replicated trials.

Discussion

The formation of Farmer Innovation Groups (FIGs) originating within the Yield Enhancement Network began in early 2018 and therefore the results and learnings presented in this paper are from a single growing season. The 2017-18 growing season in the UK was characterised by abnormally warm and dry conditions from May to July (Met Office, 2018). It is likely these drought conditions reduced crop responses to nutrient applications, exaggerated the effects of soil variation on yield and increased within-field variability, reducing precision of the tramline trials conducted. The confidence intervals achieved from these trials are higher than the average error associated with small plot trials which can 'prove' treatment differences between 0.3 and 0.5 t/ha. The spring potash FIG will be continued into a second season to improve confidence in the results obtained, with four of the five trials indicating a nil yield response to the application of MoP but the only significant treatment effect showing a consistent yield increase related to the quantity of MoP applied.

The main learnings from the spring potash FIG include: (1) The same concepts and metrics should be agreed between the researcher and farmers at the initial stages of FIG formation. This was successfully achieved in the spring potash FIG in which a hypothesis and the treatments to test the hypothesis were agreed soon after the group was formed and communicated effectively between participants, resulting in the successful set up of tramline trials that were consistent across farms. (2) FIG size should take into account potential trial errors or loss of data and allow for results to be compared across sites to gain a broader picture of outcomes (Piepho and Spilke, 2011). Although treatment effects will be site dependant, general patterns can still be drawn across trials which can benefit farmers and advisors outside of the FIG itself. (3) An able and enthusiastic lead farmer substantially improves

Table 1. Results from each spring potash FIG trial using yield maps analysed by agronomics method, unless otherwise stated.

FIG participant	Treatment (kg/ha MoP)	Mean yield (t/ha, 15% mc)	Yield advantage over untreated (t/ha ± 95% CI)[1]	Probability of yield advantage[2]
1	Untreated	10.43		
	100		-0.55±1.192	0.18
	200		-0.60±1.107	0.14
2	Untreated	10.26		
	100		-0.33±0.562	0.12
	200		-0.26±0.553	0.18
3[3]	Untreated	9.23		
	100	9.46	+0.23±0.122*	
	200	9.72	+0.49±0.122*	
4a	Untreated	9.75		
	100		-0.60±0.798	0.07
	200		-0.66±0.799	0.05
4b	Untreated	10.75		
	100		-0.05±1.294	0.47
	200		-0.18±1.279	0.39

[1] 95% confidence interval. * Indicates where difference is significant at the 95% confidence level.
[2] Probability on a 0-1 scale, where 0 indicates a certain disadvantage, 1 indicates a certain advantage, and 0.5 indicates no evidence of a difference either way.
[3] Statistical analysis of tramline yield averages measured using weighbridge, not agronomics. Analysed using ANOVA (P<0.01).

co-ordination and communication within the FIG. The farmer lead within the spring potash FIG was pivotal in treatment design, recruiting participants and communicating ideas and progress to both the facilitator and involved farmers which is important in both keeping with the farmer-centric nature of the research but also indicated the benefits of farmer-to-farmer knowledge exchange.

Spring potash FIG participants commended the pre-harvest trial visit to discuss the harvest procedure and geo reference tramline wheelings, which enabled farmers to harvest the field as a whole rather than in subsections. This visit also allowed discussion of other trial components and could be timed before or after treatment application. The spring potash FIG was unsuccessful in organising a meeting after harvest with all farmers in attendance which would be of great value to gage opinions on success of trials and to discuss the trial results. When trials are continued over multiple seasons, it is inevitable that participants may lose interest, particularly when results are not positive. The farmer-centric nature of this research aims to reduce this phenomenon. In the case of the spring potash FIG, three of the five participants agreed to repeat the trials in a second season. This is difficult to avoid but may be prevented by keeping communication concise, keeping cost for the farmer minimal and promoting benefits of undertaking scientifically sound on-farm trials.

The fundamental advantage of undertaking OFE for the farmer is the fact that they are able to see the results first hand. The value of these results are more significant than those from a traditional scientific experiment as they are relevant to the farmer's land (Cook *et al.*, 2013). Additionally, tramline trials are undertaken at scales that farmers can appreciate, leading to the farmer being able to make informed decisions from the trial results and allowing them to bring practical considerations into

the trial interpretation, such as economics. The statistical analysis of yield map data taken directly from the combine allowed scientific robustness to be applied to these on-farm tests, and accounted for the large within-field underlying variability, which a farmer is likely to ignore. Through using this scientific rigour in analysing OFE data, more farmer-centric research can be undertaken which better links scientific experiments and farmer activities. Maat (2011) suggests that it is this 'experimental gap', created by the separating of activities between farmers and researchers, which needs to be addressed in order to bring actual yields closer to potential yields. For OFE and agronomics to become common practice within agricultural research, these methods need to be accepted by the scientific community, in conjunction with improved engagement between researchers and farmers.

Acknowledgements

We are grateful for funding from the European Agricultural Fund for Rural Development (Grant No. 107433) under the EIP-Agri scheme.

References

Cook, S., Cock, J., Oberthur, T., Fisher, M. 2013. On-farm experimentation. Better Crops, International Plant Nutrition Institute 97: 17-20.

EIP-AGRI. 2018. Agricultural Knowledge and Innovations Systems. https://ec.europa.eu/eip/agriculture/en/publications/eip-agri-brochure-agricultural-knowledge-and (accessed December 2018)

EIP-AGRI. 2016. Thematic Networks., https://ec.europa.eu/eip/agriculture/en/my-eip-agri/thematic-networks (accessed December 2018)

Griffin, T., Dobbins, C., Vyn, T., Florax, R., Lowenberg-DeBoer, J. 2008. Spatial analysis of yield monitor data: case studies of on-farm trials and farm management decision making. Precision Agriculture 9: 269-283.

Ingram, J. 2010. Technical and social dimensions of farmer learning: An analysis of the emergence of reduced tillage systems in England. Journal of Sustainable Agriculture 34: 183-201.

Johnston, A. E. and Milford, G. F. J. 2012. Potassium and nitrogen interactions in crops. Potash Development Association.https://www.pda.org.uk/technical-potash-notes/potassium-and-nitrogen-interactions-in-crops (accessed December 2018)

Maat, H. 2011. The history and future of agricultural experiment. NJAS – Wageningen Journal of Life Sciences 57: 187-195.

Macmillan, T., Benton, T. 2014. Engage Farmers in Research. Nature 509: 25-27.

Marchant, B., Rudolph, S., Roques, S., Kindred, D., Gillingham, V., Welham, S., *et al.,* 2019. Establishing the precision and robustness of farmers' crop experiments. Field Crops Research 230: 31-45.

Met Office. 2018. Summer 2018.https://www.metoffice.gov.uk/climate/uk/summaries/2018/summer (accessed December 2018)

Piepho, H and Spilke, J. 2011. Statistical aspects of on-farm experimentation. Crop and Pasture Science 62: 721-735.

Steiner, K. 1998. Using farmer's knowledge of soils in making research results more relevant to field practice: Experiences from Rwanda. Agriculture, Ecosystems and Environment 69: 191-200.

Sylvester-Bradley, R., Kindred, D., Marchant, B., Rudolph, S., Roques, S., Calatayud, A., *et al.,* 2017. Agronōmics: transforming crop science through digital technologies. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences 8: 728-733.

Sylvester-Bradley, R. and Kindred, D. 2014. The Yield Enhancement Network: Philosophy and results from the first season. Aspects of Applied Biology, Agronomic Decision Making in an Uncertain Climate 125: 53-62.

Sylvester-Bradley R. 1991. Modelling and mechanisms for the development of agriculture. Aspects of Applied Biology, The Art and Craft of Modelling in Applied Biology 26: 55-67.

Improving yield mapping accuracy using remote sensing

R. Gonçalves Trevisan[1*], L.S. Shiratsuchi[2], D.S. Bullock[3] and N.F. Martin[1]
[1]*University of Illinois at Urbana Champaign (UIUC), Crop Sciences, W201 Turner Hall, 1102 S. Goodwin Ave, Urbana, IL 61801, USA; rodrigo7@illinois.edu*
[2]*Louisiana State University Agricultural Center, School of Plant, Environmental and Soil Sciences, 104 Sturgis Hall, Baton Rouge, LA 70803, USA*
[3]*UIUC, Agricultural and Consumer Economics, 339 Mumford Hall, 1301 W. Gregory, Urbana, IL 61801, USA*

Abstract

The objective of this work was to investigate the use of remotely sensed vegetation indices to improve the quality of yield maps. The method was applied to the yield data of twelve cornfields from the Data Intensive Farm Management project. The results revealed the need to time shift the yield values up to three seconds to better match the sensor readings with the geographic coordinates. The residuals of the yield prediction model were used to identify points with unlikely yield values for that location, as an alternative to traditional approaches using local spatial statistics, without any assumption of spatial dependence or stationarity. The temporal and spatial distribution of the standardized coefficients for each experimental unit highlighted the presence of trends in the data. At least five out of the twelve fields presented trends that could have been induced by data collection.

Keywords: on-farm precision experimentation, normalized difference vegetation index, data filtering, error correction

Introduction

On-Farm precision experimentation (OFPE) is becoming an important resource to understand the spatial variation of crop response to management practices and thus improve agronomic decisions. High-quality yield data are fundamental for obtaining real insights from this type of experiment. Previous research studies have demonstrated that yield data quality affects the outcome of OFPE data analysis (Griffin *et al.,* 2008). Thus, there is great interest in finding new methods to improve yield monitor information quality. Most publications aiming to establish guidelines to improve yield data quality are focused on removing erroneous observations. The majority of data cleaning methodologies include global frequency distribution-based filtering rules, complemented by local or spatial methods (Leroux *et al.,* 2018). However, even when carefully followed, these methods leave sources of uncertainty unaccounted. Temporal drift and other artificial trends can be present in the yield data when sensors are not properly maintained, cleaned or calibrated. These problems are harder to detect when multiple machines are used to harvest the same field, because they generate data with different levels of noise and bias. The spatial resolution of yield data is of greater importance in OFPE scenarios, impacting the minimum size of the plots and the costs of trial implementation. A better understanding of the data quality can guide the decision to what level the data should be aggregated prior to further analysis.

The presence of distinct treatments in neighboring plots poses an extra complexity not accounted for in conventional techniques for filtering spatial data. Conventional methods assume that the spatial dependence of the yield data distribution can be used to estimate the yield value at one point based on its neighbors' yield values, and that the difference between the observed and the estimated yield values may be used as a criterion for defining outlying points (Lyle *et al.,* 2014). However, this assumption is generally invalid for OFPEs, which are designed to vary seed and nitrogen rates in neighboring plots, meaning that neighboring yield value samples may be drawn from different

DOI 10.3920/978-90-8686-888-9_111, © Wageningen Academic Publishers 2019

frequency distributions. Incorporating this information is necessary to avoid the removal of points at which yield was measured accurately, and therefore avoid the introduction of biases in the analysis. The use of auxiliary sources of information may be an important strategy to deal with the particularities of OFPE data outlined above. Remotely sensed information with high temporal and spatial resolution is becoming increasingly available, and has been successfully used to predict yields at different scales (Khanal *et al.,* 2018; Peralta *et al.,* 2016). Therefore, the objective of this work was to investigate the use of remotely sensed vegetation indices to improve the quality of yield maps.

Materials and methods

Yield data from twelve cornfields were recorded during the 2017 harvest season by combines equipped with different yield monitoring systems. There were seven rainfed fields, six of them in in Illinois (Fields 03, 05, 08, 09, 10 and 11) and one in Ohio (Field 6), and five irrigated fields, four of them in Nebraska (Fields 02, 04, 07 and 12) and one in Kansas (Field 01). The number of variables available for each field varied according to the yield monitor model and software used to extract the data. The minimum variables consisted of only longitude, latitude and yield, while for some datasets information on moisture content, harvest time and machine orientation was also available. In order to replicate the same procedure on all fields, only the yield and coordinates data were used. The field trials were conducted as part of the Data-Intensive Farm Management (DIFM) project, in which OFPE is used to generate data to improve understanding of crop responses to seeding and nitrogen fertilizer application rates. The fields where chosen to represent the range of yield variation commonly observed in these trials (Figure 1). Each field had on average 40 ha and 200 experimental units, represented by 85 m long and 18 m wide polygons. This information was incorporated using the polygon identification as a predictor factor, without relating the experimental units with the treatment applied. This method was chosen instead of using the treatment values and the model residuals for subsequent analysis to fully separate the data cleaning from the model evaluation steps, since it could introduce biases.

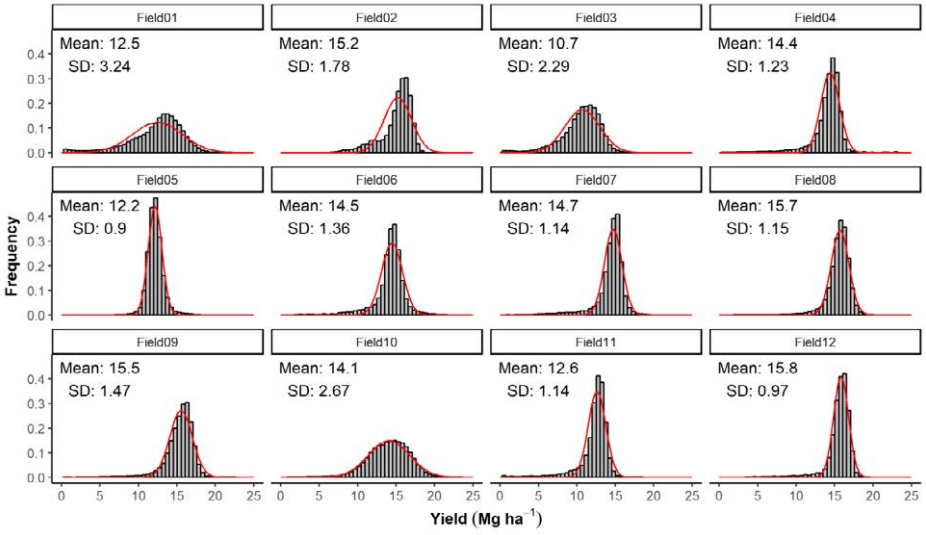


Figure 1. Histograms of corn yield distribution (Mg/ha) for the 12 fields under the Data Intensive Farm Management project. The red line represents the expected frequency for normally distributed data with the mean and standard deviation (SD) shown in each histogram.

Remote sensed data consisted of Planet Labs PSSE4 imagery with RGB and NIR bands at three-meter spatial resolution (Planet Team, 2017). Cloud-free images were chosen from consecutive fifteen-day periods, beginning thirty days after sowing and ending fifteen days before harvest. Thus images were acquired on eight to ten dates per field. The image products were already corrected to surface reflectance (Houborg and McCabe, 2018), and the values were used directly to calculate the normalized difference vegetation index (NDVI) for each date. Principal component analysis was applied to NDVI data for the whole season in order to generate six orthogonal variables containing more than 90% of the variance in the NDVI. The values of the pixel touching each yield point were transferred to the yield dataset.

Yield data were first submitted to exploratory descriptive statistics and basic error removal, with the exclusion of all the points with values differing from the average by more than three times the interquartile range. Multiple linear regression models were fitted to predict yield variability, considering the six principal components (PCs) and plot identification as predictors, with no interactions. Yield values were also shifted across time, from ten seconds behind to ten seconds ahead of the registered time position. The records in the dataset were in chronological order and shifting yield values in time was done by shifting the index of the record. The time shift that resulted in the highest R-squared was considered as the correct time delay for each field. The final multiple linear regression model used in subsequent analysis was fit to the yield corrected for time delay. The autocorrelation of sequential points in time was used to identify the level of noise in the yield registered by the yield monitoring sensors. The same method was applied to the predicted yield for comparison.

The coefficient estimates for each experimental unit were standardized to have zero mean and unit variance. These coefficients were fitted in the yield prediction model to each polygon and represent the portion of yield variance not explained by the PCs due to the different treatments or changes in yield not related to changes in the NDVI. The standardized coefficients were compared to the relative time of harvest, defined as the average relative file position of the yield points over that plot, to investigate the presence of trends or drifts over time. The same values were also mapped to explore the spatial distribution of the plot coefficients. Finally, the standardized residuals from the yield prediction model were used as indicators of the probability of a point having been erroneous. All procedures described and figures presented in this paper were developed using the R programing language (R Core Team, 2018).

Results and discussion

The histograms of the yield distribution in each field as recorded by the yield monitor (Figure 1) show that the data follow a normal distribution in most situations. Values are concentrated around the mean, and there are more values below the average than would be expected from sampling the normal distribution. This reflects the biological meaning of these values, since many factors can reduce the yields, but increases in yield are limited by the crop's maximum biological potential. The datasets represent a range of high yielding fields, varying from 10.7 Mg/ha, slightly lower than the national average of 11.1 Mg/ha in the 2017 season (United States Department of Agriculture – USDA, 2017), to 15.8 Mg/ha. The standard deviation of yield varied from less than 1.0 Mg/ha to more than 3.0 Mg/ha, illustrating differences in field growing conditions.

The correlation between PCs and yield was usually higher for the first PC (Table 1), which explains most of the variability. The main exception was field 6, in which the second and third components had the highest absolute value of correlation. The ability of the time series of NDVI images to predict the yield varied among the fields. The best results were observed in field 2, with 75% of the variance explained by the model. The worst performance was observed in field 10, with only 16% of the variance explained by the model. These results are important because the utility of the satellite images to improve the accuracy of the yield maps is directly related to their ability to predict yield.

Nevertheless, the low performance of the yield prediction can also be caused by noise in the yield data, and by the inability of the model to capture the underlying relations. Further exploration of the field 10 data revealed that the problem was very likely to be with the yield data itself (Figure 2B). Consistent results of the optimum time shift were obtained by representing the R-squared of the yield prediction as function of the time shift applied to the yield. All fields presented the highest R-squared between zero and three seconds, with decreasing values when large time shifts were applied. Only three of the fields did not increase the R-squared when the yield values where shifted in time, meaning that the right delay was already been used in the monitor configuration. The data in four fields should be shifted by one second, in other four fields by two seconds, and in one field by three seconds in order to match the yield values with the coordinates (Figure 2A).

Fields 4 and 10 presented noisy data with negative temporal autocorrelation (Figure 2B). The accuracy of the data with these characteristics is questionable. It is likely that a sensor malfunction was the reason for such behavior. Discarding the data, especially for field 10 could be the more reasonable decision. For field 4, aggregating the data every few seconds could be a viable strategy. If the data is aggregated to every experimental unit, the short scale noise introduced by the sensor may not be a problem. The autocorrelation in the yield prediction was usually higher, evidencing that the NDVI variation is smoother than the variability registered by the yield monitor.

The standardized residuals of the yield prediction model were used to identify points with unlikely yield values for that location, as an alternative to the traditional approach of using local statistics without any assumption of spatial dependence or stationarity. The traditional criterion of excluding points outside the range of three times the standard deviation can be applied.

In the two fields chosen as example (Figure 3), most of the errors highlighted with the more intense green and red colors are due to subtle changes in the machine speed or the wrong cutting width. There is no tendency of eliminating points close to the border of the experimental units with different treatments or to consider as erroneous the data points with low yields due to low nitrogen rates, for example. Although the proposed methodology requires the use of additional information, it can be considered simpler than most of the current procedures recommended for yield data filtering (Lyle *et al.,* 2014), avoiding the use of many arbitrarily chosen steps and parameters.

Table 1. Correlation coefficient between each principal component of the NDVI images and the yield, and the coefficient of determination of the yield prediction using all PCs and the experimental unit polygon identification.

Field	PC1	PC2	PC3	PC4	PC5	PC6	R^2
1	**0.65**	-0.20	-0.08	0.08	0.04	-0.04	**0.73**
2	**0.74**	-0.12	0.08	0.17	-0.12	0.02	**0.75**
3	**0.38**	0.05	0.03	-0.03	0.00	0.08	**0.61**
4	**0.52**	0.27	0.32	-0.05	-0.37	-0.23	**0.45**
5	**0.40**	0.22	0.05	0.01	0.00	0.08	**0.41**
6	0.07	-0.29	**-0.33**	-0.04	0.04	-0.04	**0.29**
7	**0.48**	-0.11	-0.16	0.22	0.02	-0.05	**0.50**
8	**0.45**	0.34	0.11	-0.01	0.00	0.00	**0.48**
9	0.45	**0.50**	-0.11	0.00	0.05	-0.05	**0.58**
10	**0.18**	-0.17	0.17	0.00	-0.03	0.00	**0.16**
11	0.16	**-0.23**	-0.16	0.01	-0.02	0.08	**0.28**
12	**0.50**	0.31	-0.15	0.05	-0.01	-0.03	**0.59**

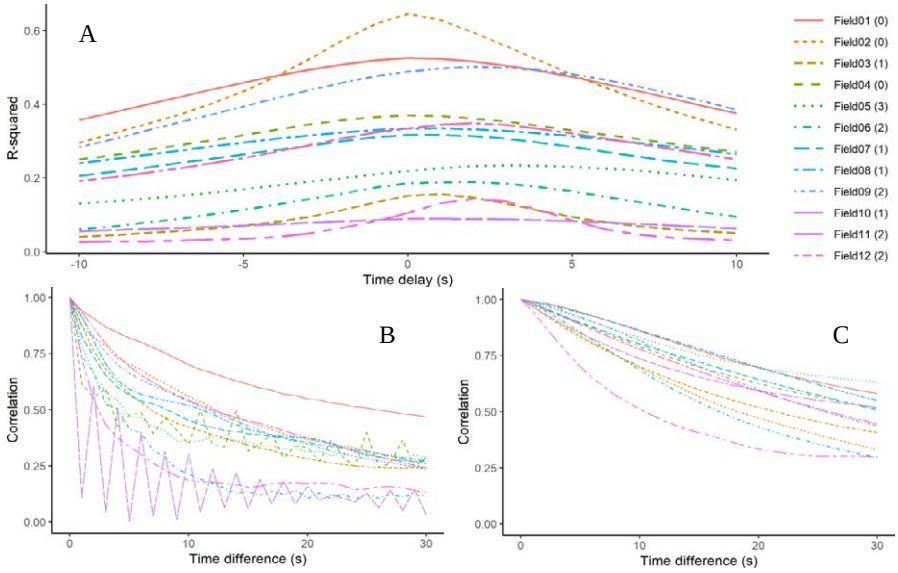


Figure 2. (A) Yield prediction performance as function of changing the delay of the yield record. The number inside the parenthesis in the legend represents the time with the highest R-squared. (B) Temporal autocorrelation of the yield data observations in each field. (C) Temporal autocorrelation of the predicted yield in each field.

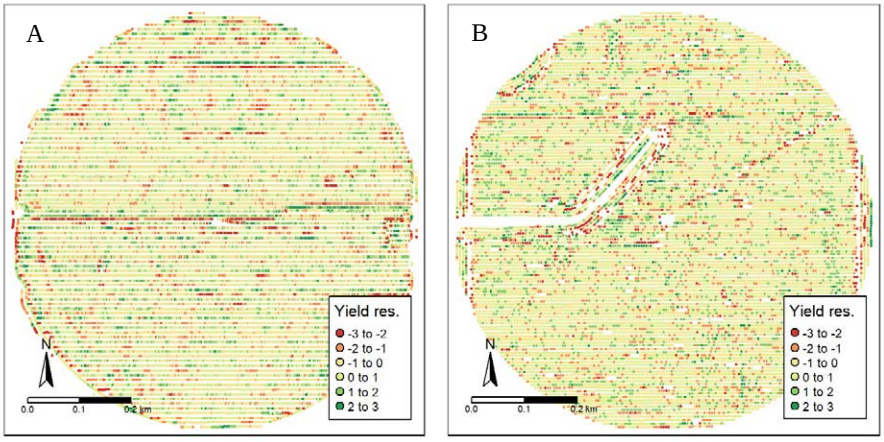


Figure 3. Spatial distribution of the standardized yield residuals in fields 1 (A) and 2 (B).

The inclusion of other layers in the yield prediction model was also considered, especially the use of variables related to field topography. The use of the satellite images was preferred due to the difficulty of obtaining accurate elevation data for all fields and the possible correlation between the slope of the field and the errors in the yield monitor.

The independence of the yield estimation errors and the yield monitor errors was one of the assumptions to identify the presence of artificial trends in the yield data. The trends are defined as

the presence of a significant change in the estimated average of the yield in each polygon representing one experimental unit.

Visualizing the standardized estimated coefficient of each polygon as function of the time of harvest can reveal such trends. Because the rates are randomly assigned to the plots, these coefficients should not present any correlation with time. This behavior can be observed in fields 2 and 4 (Figure 4), for example. Various levels of trends can be observed in other fields, from small deviation to the straight line in fields 6 and 12 to large deviations in fields 1 and 5. The spatial distribution of the estimated coefficients for each experimental unit can be used to further investigate these trends (Figure 5). The method used cannot distinguish between trends introduced by data collection and trends related to other management factors that may affect yield without affecting NDVI. Independent of the factor causing the trend, it is important to acknowledge the existence of these trends and to account for it when analyzing the yield response, either by excluding part of the data or including grouping factors.

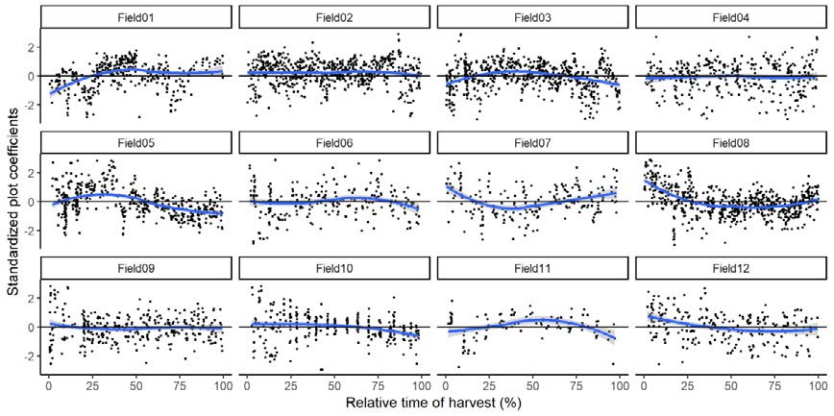


Figure 4. Standardized coefficients for each experimental unit compared to the relative time of harvest. The line represents a smoothed moving average.

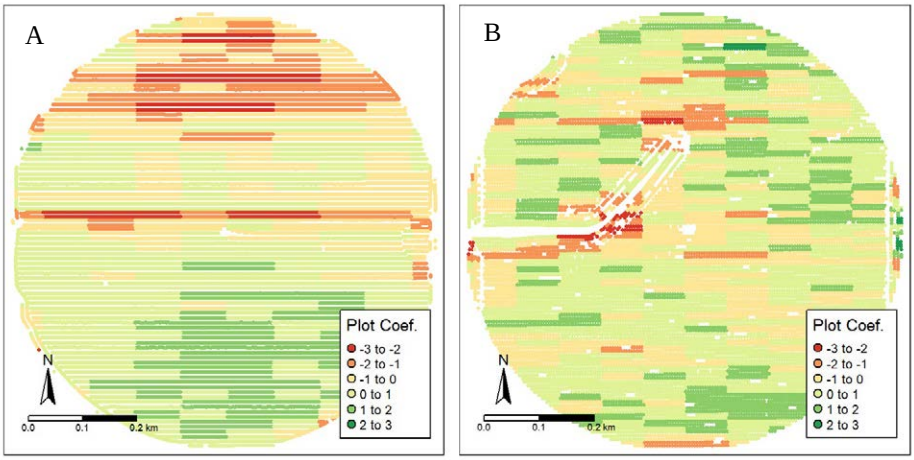


Figure 5. Spatial distribution of the standardized coefficients for each experimental unit in fields 1 (A) and 2 (B).

Differences similar to the pattern observed in field 2 (Figure 5A) are expected due to the different treatments in the field and the fact that some factors that affected yield may not have affected NDVI, especially late in the season, when the NDVI gets saturated. Differences such as the ones observed in field 1 (Figure 5B) are unlikely to be caused by the treatment effects in yield. The main stripe of negative values in the middle of the field is clearly an error and should be eliminated. Besides that, there are still three different clusters of errors present in the data. The group of negative values in the north of the field seems to be related to the harvest, since they are mostly aligned with the harvest direction. The group of values in the bottom of the field may be related to irrigation, since it shows a radial distribution.

The use of an independent sensor to calibrate the load cell at the impact plate, such as a weighing system for the grain tank could contribute to the validation of this methodology. The use of simulation can also be an alternative to validate the proposed methodology (Leroux *et al.,* 2017), but the delays in yield record and the presence of trends should be included in the simulated data.

Conclusions

The proposed method allowed the identification of the most likely time shift to be applied to correct the yield monitor values. This value varied from zero seconds, meaning that the value set in the monitor configuration was right, to three seconds, meaning that the value set in the yield monitor configuration should be increased by three to better match the sensor readings with the coordinates. The residuals of the yield prediction model were used to identify points with unlikely yield values for that location, based only on the global frequency distribution of the values. This method can be seen as an alternative to the traditional approaches using local spatial statistics, without any assumption of spatial dependence or stationarity.

The temporal and spatial distribution of the standardized coefficients for each experimental unit highlighted the presence of trends in the data. Although the method cannot explain the reason for the trend, it can help to improve the quality of subsequent analysis by accounting for the trends in the model definition.

Acknowledgements

This research was funded in part by a USDA National Institute of Food and Agriculture Food Security Program grant, award number 2016-68004-24769. We also acknowledge Planet Labs for providing API tools and images, as part of Planet's Education and Research Program.

References

Griffin, T. W., Dobbins, C. L., Vyn, T. J., Florax, R. J. G. M., & Lowenberg-DeBoer, J. M. (2008). Spatial analysis of yield monitor data: case studies of on-farm trials and farm management decision making. Precision Agriculture, 9 (5), 269-283. doi:10.1007/s11119-008-9072-2

Houborg, R., & McCabe, M. F. (2018). A Cubesat enabled Spatio-Temporal Enhancement Method (CESTEM) utilizing Planet, Landsat and MODIS data. Remote Sensing of Environment, 209, 211-226. doi:10.1016/j.rse.2018.02.067

Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N., & Shearer, S. (2018). Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. Computers and Electronics in Agriculture, 153, 213-225. doi:10.1016/j.compag.2018.07.016

Leroux, C., Jones, H., Clenet, A., Dreux, B., Becu, M., & Tisseyre, B. (2017). Simulating yield datasets: an opportunity to improve data filtering algorithms. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences, 8(02), 600-605. doi:10.1017/S2040470017000899

Leroux, C., Jones, H., Clenet, A., Dreux, B., Becu, M., & Tisseyre, B. (2018). A general method to filter out defective spatial observations from yield mapping datasets. Precision Agriculture, 19 (5), 1-20. doi:10.1007/s11119-017-9555-0

Lyle, G., Bryan, B. A., & Ostendorf, B. (2014). Post-processing methods to eliminate erroneous grain yield measurements: review and directions for future development. Precision Agriculture, 15 (4), 377-402. doi:10.1007/s11119-013-9336-3

Peralta, N. R., Assefa, Y., Du, J., Barden, C. J., & Ciampitti, I. A. (2016). Mid-Season High-Resolution Satellite Imagery for Forecasting Site-Specific Corn Yield. Remote Sensing, 8 (10), 1-16. doi:10.3390/rs8100848

Planet Team. (2017). Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA, USA. https://api.planet.com

R Core Team. (2018). R: A Language and Environment for Statistical Computing. Vienna, Austria. https://www.r-project.org

United States Department of Agriculture – USDA. (2017). Quick Stats 2.0. Washington DC.: U.S. Department of Agriculture, National Agricultural Statistics Service. https://quickstats.nass.usda.gov

Model validation of the dynamics of a no-till seeding assembly with a magneto-rheological damper system

G.M. Sharipov*, D.S. Paraforos and H.W. Griepentrog
University of Hohenheim, Institute of Agricultural Engineering 440d, Garbenstr. 9, 70599, Stuttgart, Germany; galibjon.sharipov@uni-hohenheim.de

Abstract

To overcome challenges in achieving a consistent seeding depth in no-till seeding, modelling and simulation of the dynamics of a no-till seeding assembly equipped with a magneto-rheological (MR) damper were carried out. This was followed by the implementation of the MR damping system on the seeding assembly. In order to verify the correctness of the simulation model, the simulation outputs, i.e. pitch angles and impact forces, were correlated with the measured ones based on criteria such as root-mean-squared errors and correlation coefficients. The correlation analyses depicted a good agreement, with an RMS error percentage below 25% and the correlation coefficient above 0.6, between all the simulated and measured data sets.

Keywords: impact forces, pitch angles, seeder dynamics, simulation model

Introduction

In no-till seeding, correct seeding depth is a very important factor that affects germination and emergence. Both the germination and emergence ultimately affect optimal crop growth and, in general, yields (Özmerzi *et al.,* 2002). Inconsistency in the seeding depth mainly occurs due to the dynamic responses of the seeder to harsh soil conditions. The variability of coulter dynamic responses to heterogeneous soil conditions is due to forces and vertical displacements that are not fully regulated during seeding (Al-Kheer *et al.,* 2011). The motion behaviour of the coulter assembly can be described by the forces resulting from the interaction between the coulter and the soil (Hasimu and Chen, 2014). The coulter assembly mainly consists of the coulter arm and tine, where the draught and vertical forces arise, and the packer wheel, which is affected by the profile impact forces. By taking into account these forces while controlling the assembly vertical movement, the performance could be significantly optimised for a proper dynamic behaviour that results in better seeding depth. Development of a simulation model of the seeder dynamics with a damping system would be the first step towards implementation on the seeding assembly.
By introducing a mathematical model of a semi-mounted seeding implement and the simulation of its dynamic response to surface undulations, Lawrance (1969) revealed that the vertical excessive oscillations of the seeding component, described by the motion behaviour of the seed depositing apparatus, occur due to the dynamic response of the machine to soil undulations. To develop an automatic depth control system of a seeder, its dynamic response was modelled and simulated under laboratory condition by Weatherly and Bowers (1997). By introducing a simulation model of an oscillating sub-soiler dynamic behaviour, Shahgoli *et al.* (2010) demonstrated the effect of the field profile and stiffness and damping of sub-soiler tyres on the vertical motion behaviour of the machine. Of the many techniques for optimising the dynamic response of no-till seeders and controlling the seed placement mechanism (Nielsen *et al.*, 2018), one of the common methods is to regulate the applied downforce on the entire seeding assembly (Rui *et al.*, 2016); however, in this method, advanced control cannot be achieved due to the non-linearity in compressing/depressing component (spring or rubber) (Gratton *et al.,* 2003) and the complexity of the assembly design.

DOI 10.3920/978-90-8686-888-9_112, © Wageningen Academic Publishers 2019

In spite of all the above-mentioned studies, there has been very limited research on developing a control model of the vertical movement of a no-till direct seeder with a semi-active MR damper using the corresponding actual surface profile and the measured reaction forces. The MR damper system consists of a hydraulic cylinder, magnetic coils and MR fluid. The MR fluid is a viscous liquid combined with magnetically fast-responsive particles in micron size. So, the flow motion of the MR liquid can be adjusted by regulating the strength of the magnetic field applied to the magnetic coil. By modelling and simulating the dynamic motion behaviour of the seeding assembly with the MR damping system, Sharipov *et al.* (2017a) demonstrated that the dynamic response of the seeder to harsh soil conditions, in terms of the reduction in amplitudes of the developed forces and vertical displacements, could be significantly improved. The performance of the MR damper was captured using up-to-date sensors such as robotic total stations, inertial measurement units and laser pointers, whose accuracy has been assessed using an industrial robotic arm (Paraforos *et al.,* 2017). Consequently, the performance of the georeferenced seeding depth with the corresponding seeder dynamics was evaluated (Sharipov *et al.,* 2018). This evaluation necessitated further validation of the simulation model by utilizing the measured inputs such as vertical forces and surface profile resulting from a real-life test.

The aim of the project was to test the simulation model of the vertical dynamics of the seeding assembly with the MR damper to verify its correctness using the vertical forces and field profile from the measured data. To evaluate, the simulation outputs, i.e. pitch angles and impact forces, were compared with the measured ones, based on the correlation criteria such as root-mean-squared (RMS) error and correlation coefficient.

Material and methods

Instrumentation and data acquisition

A machine prototype was developed comprising a metal frame that carried a Green Drill 200 seed dosing mechanism and two ConTeC Schare no-till seeding assemblies (both from AMAZONEN-Werke H. Dreyer GmbH & Co. KG, Hasbergen, Germany) (Figure 1A). The seed dosing mechanism was equipped with a seed-metering wheel for wheat (*Triticum aestivum* L.). For configuring the seeding rate and the speed of the seed shaft, a dedicated in-cab control terminal was employed. The two seeding assemblies with a 0.25 m inter-row distance were attached to a rotating square rod controlled by a hydraulic cylinder. The hydraulic cylinder provided a down pressure with a value of 6.89 MPa for the two seeding assemblies. To reduce the effect of the vertical and profile impact forces that are the main reason for excessive vertical movements of the seeding assembly, an RD-8040-1 MR damper (LORD, Baltimore, USA) was implemented on one seeding assembly between the wheel shank and the packer wheel. This was carried out by using an extra shank and a bearing system as indicated in Figure 1B. More detailed information about the performance of the MR damper can be found in Sharipov *et al.* (2018).

All necessary sensors were fixed on the seeder mainframe and the two seeding assemblies (Figure 1A and 1B) to determine the traversed field surface profile of both seeding assemblies and to measure the developed machine dynamics, such as forces and tilting information. An SPS930 total station (Trimble, Sunnyvale, CA, USA) tracking a Trimble MT900 machine control prism fixed on the seeder mainframe was employed to acquire the in-field absolute geo-referenced position of the seeder mainframe. A DT50 laser range finder (SICK AG, Waldkirch, Germany) detected the displacement of the seeding assembly with the MR damper relative to the main frame. Three VN-100 inertial measurement units (IMUs) (VectorNav, Dallas, TX, USA), with one on the seeder mainframe and the other two fixed to the seeding assemblies, were used to gather real-time tilt information (roll-pitch-yaw) of the seeder mainframe and the seeding assemblies, respectively. Furthermore, the total station was employed to fuse the data from both IMU and laser range finder in order to improve the vertical displacement and pitch estimation. To determine the vertical and profile impact forces, the

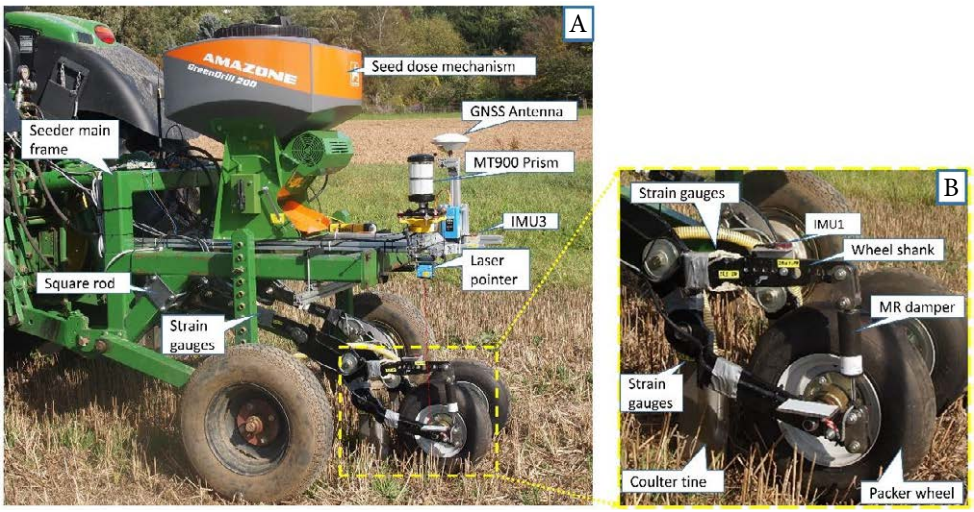


Figure 1. (A) The machine prototype and sensor-frame with the sensors for measuring machine dynamics, and (B) the MR damper.

strains resulting from the loads on the coulters and the packer wheels were recorded using six linear 350 Ω DY41-1.5 strain gauges (HBM GmbH, Darmstadt, Germany) in a full-bridge configuration, attached at the critical points on the coulter assemblies.

Mathematical modelling and simulation

The seeding assembly comprised a coulter, a wheel shank and a packer wheel. To develop a mathematical model of the seeding assembly with the semi-active MR damper system, a mass-spring-damper model was engaged. The semi-active MR damper was assumed to be located between the wheel shank and the packer wheel, as represented in Figure 2. The model considered the mass of the packer wheel as an unsprung mass m_w [kg], and the mass of the coulter together with the wheel shank as a sprung mass m_c [kg]. In the model, the deformation and damping force of the packer wheel could be characterised by the displacements and the velocity of the unsprung mass, respectively. Consequently, the profile impact forces F_{sp} [N] were expressed by the following expression:

$$F_{sp} = K_w(x_p - x_w) + C_w(\dot{x_p} - \dot{x_w}) \tag{1}$$

Where x_w, $\dot{x}_w$ are the vertical displacement and velocity of packer wheel, in [m] and [m/s], respectively; x_p is the vertical surface profile that the packer wheel followed; K_w is the stiffness of the tyre [N/m] and C_w is the damping coefficient of the tyre [Ns/m]. More detailed information about the stiffness and damping coefficient of the wheel tire can be found in Sharipov *et al.* (2017b). In this model, the damping force F_{mr} [N] resulting from the semi-active suspension system could be the input force to the seeding assembly. Therefore, the profile impact force F_{sp} was substituted for F_{mr}. As a result, the equations for motion behaviour of the coulter assembly, including a semi-active MR damper system, were yielded as follows:

$$\left(m_w + \frac{m_c}{4}\right) L_a^2 \ddot{\theta} = F_{mr} L_a - ((m_w + m_c) g L_c + F_v L_v)\theta - M_f \tag{2}$$

$$m_w \ddot{x}_w + C_s\left(\dot{x}_w - L_a\dot{\theta}\right) + K_s(x_w - L_a\theta) + C_w\dot{x}_w + K_w x_w = F_{sp} - F_{mr} \tag{3}$$

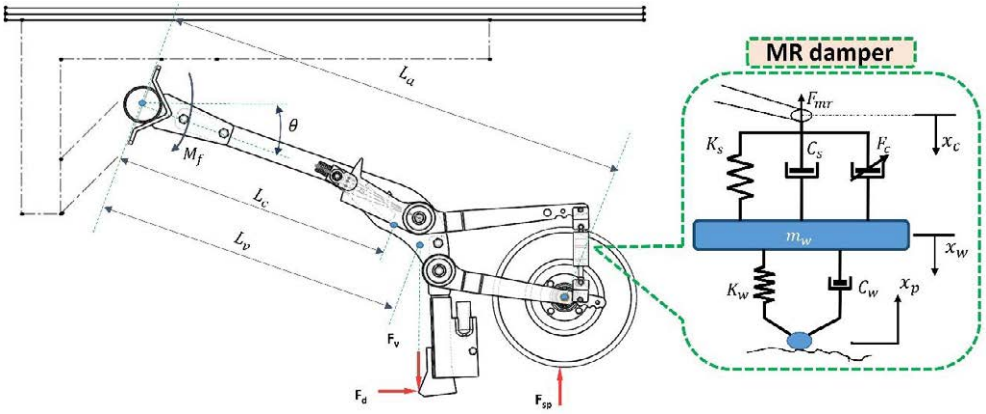


Figure 2. Representative view of the seeding assembly with the semi-active MR damper system.

where θ is the pitch angle of the seeding assembly motion; K_s and C_s are the stiffness and damping coefficient of the semi-active suspension passive components, in [N/m] and [Ns/m], respectively; m_c and m_w are the mass in [kg] of the coulter assembly and the packer wheel, respectively; L_c is the distance between the coulter fixing point on the main frame and the centre of mass [m]; and L_v is the distance between the coulter fixing point on the main frame and the point that the vertical forces are acting on the coulter [m].

Bouc-Wen model

In the model (Figure 2), F_c is the control force of the MR damper. However, different hysteresis models could be applied for the damper model to address the hysteresis and nonlinear behaviour of the MR damper. The Bouc-Wen model (Figure 3) was chosen to investigate the non-linear behaviour of the MR damper due to its superior performance over other models (Sharipov *et al.,* 2017a). The damping force in the Bouc-Wen model is described by the following expressions:

$$F_{mr} = C_s(u)\,\dot{x}_c + K_s(u)\,x_c + \alpha\, z + f_0 \tag{4}$$

where the internal state variable is governed by:

$$\dot{z} = \gamma\, z\, |x_c|\, |z|^{n-1} - \beta\, x_c\, |z|^n + A\, x_c \tag{5}$$

where γ, β and A are the control parameters for the internal state variable that can vary from a sinusoidal to a quasi-rectangular function of time; f_0 is the force due to the accumulator that considers the pre-yield stress of the damper; α is the parameter that defines the influence of the model on the active control force; $K_s(u)$ and $C_s(u)$ are the stiffness and damping coefficients, respectively, with respect to the control voltage u. The active control force is equal to the damping force coming from the model.

Performance evaluation

The mathematical Equations 1-5 for the semi-active MR damper models were implemented and simulated in MATLAB and Simulink. The measured field surface profiles and the vertical forces of Figure 4 were used as inputs to the simulation model to simulate the pitch angle of the seeding assembly and the damped forces of the MR damper with the supplied current of 0.5 A (Figure 4A and 4B) and 0.7 A (Figure 4C and 4D). From our performed simulations, it was apparent that the performance efficiency of the MR damper based on the Bouc-Wen model depended on the hysteresis

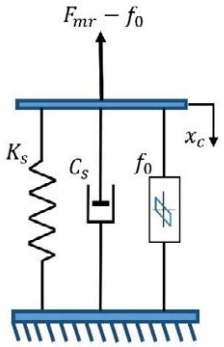


Figure 3. Schematic view of the Bouc-Wen model for the MR damping system.

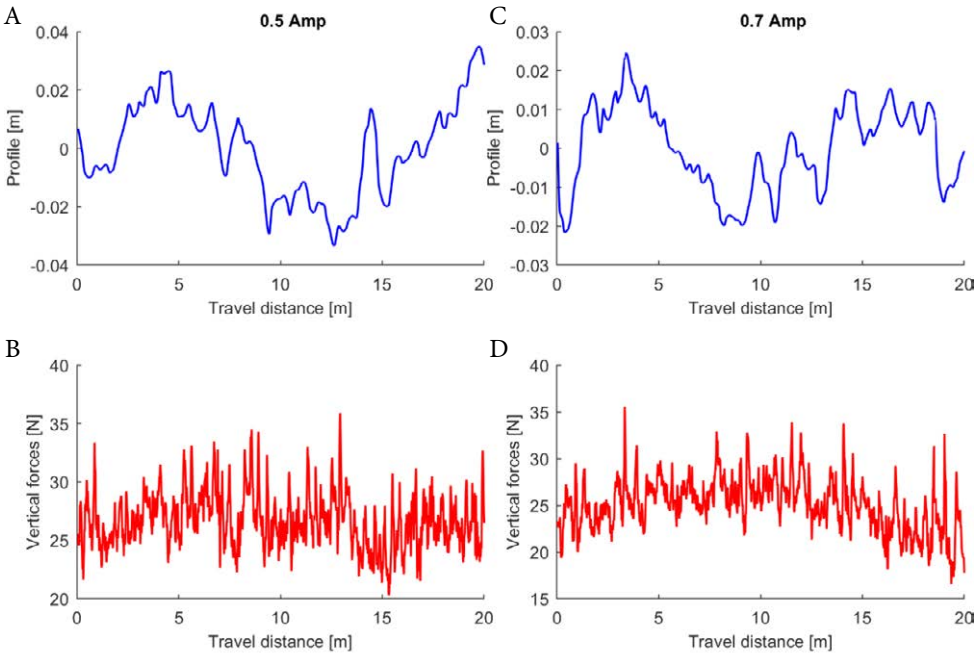


Figure 4. The measured (A and C) field surface profile and (B and D) vertical forces, when the MR damper was supplied with 0.5 A and 0.7 A.

loop parameters. Therefore, all numerical values of the hysteresis model (Bouc-Wen) parameters were found by trial and error estimation. This was performed by locating some rational values of the hysteresis loop until to have those loop parameters made the model fit. The real-life test of the MR damper-implemented seeding assembly performance was investigated with six different current input levels 0 A, 0.1 A, 0.3 A, 0.5 A, 0.7 A and 1 A applied on the coil of the MR damper. A constant operation speed of 10 km/h was set for all trials. Analysis of the MR damper performance was carried out at input currents of 0.5 and 0.7 A as the best performance was achieved at these currents (Sharipov *et al.,* 2018)

In order to verify the accuracy of the simulation model, outputs such as the simulated pitch angles and damped forces were compared with those measured ones in the fields. The comparison analyses were carried out by calculating RMS error percentage and correlation coefficient between the simulated and measured data. RMS errors described the absolute fitting of the simulated data to the measured one. RMS error values of less than 25% were considered acceptable. Subsequently, the fitting accuracy, in terms of similarity relationship, of the simulated data to the measured data was also described with a correlation coefficient. Values higher than 0.6 could be considered as practically acceptable (Sharipov *et al.,* 2017b). These thresholds for the RMS error and correlation coefficient were considered as acceptable due to a ripple effect of the lateral forces ignorance, sinkage of the packer wheel and the hysteresis loop parameters estimation in the modelling.

Results and discussion

Figure 5 demonstrates the pitch angles and damped forces resulting from both simulation and real-life test of the seeding assembly with the MR damper supplied with 0.5 A and 0.7 A. The analyses of the correlation between the simulated and measured pitch angles from the MR damper-implemented seeding assembly, when it was supplied with 0.5 A (Figure 5A), indicated an RMS error value of 14% and correlation coefficient of 0.92, respectively. At a current of 0.7 A, the figures were 18% and 0.86 (Figure 5C). For the damped forces resulting from the simulation and real-life test of the seeding assembly with the MR damper supplied with 0.5 A (Figure 5B), the RMS error and correlation coefficient were equal to 22% and 0.63, respectively. Patterns were similar for the simulated and measured damped force at 0.7 A (Figure 5D) with RMS error of 24% and correlation coefficient of 0.75. These analyses revealed that all the RMS errors and correlation coefficients were in the range of acceptable values. In addition, the pattern as the correlation coefficient was high when the RMS

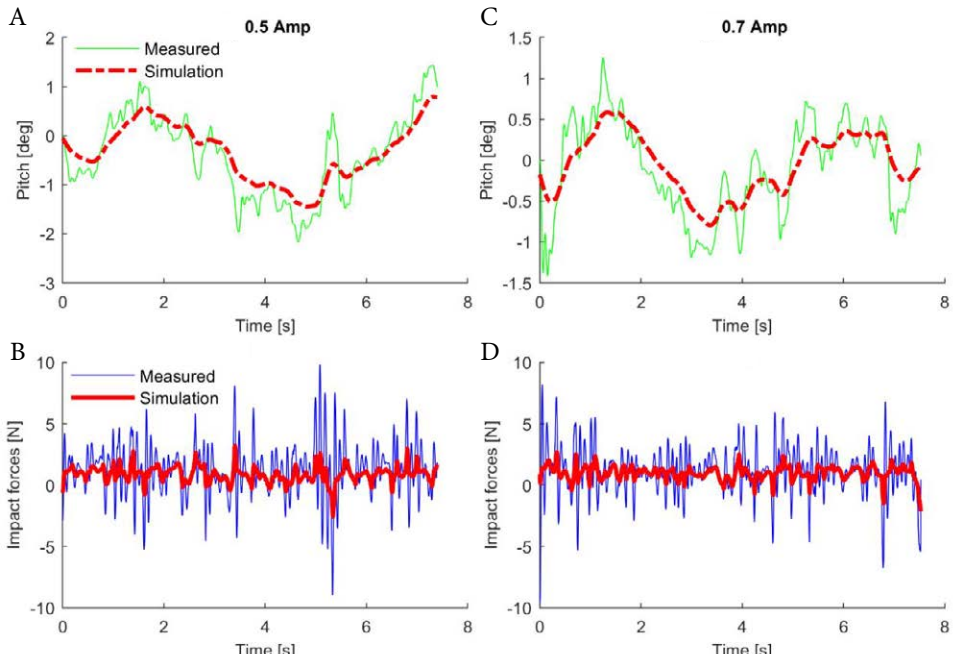


Figure 5. The measured and simulated (A) pitch angle of the aggregate and (B) profile impact forces with MR damper control current of 0.5 A and 0.7 A.

indicated a low error, confirmed the correctness of the correlation. Taking into account all the above numbers, it could be stated that the simulation model of the MR damper-implemented seeding assembly is capable of producing the same results as the measured ones.

Conclusions

A simulation model of the vertical motion dynamics of the seeding assembly with an MR damper was developed and its dynamic responses, in terms of the profile impact forces and pitch angle of the assembly vertical motion, were simulated using the measured surface profile and vertical forces. The accuracy of the simulation model was verified by comparing the RMS error percentage and the correlation coefficient of the simulation outputs with that of the measured ones. The comparison analyses depicted a good agreement, with the RMS error percentage below 25% and the correlation coefficient above 0.6, between the simulated and measured datasets for both 0.5 A and 0.7 A supply current. The model is capable of producing similar results to the measured ones. Thus, it could be used as a basis to find the optimal settings for the MR damper but also further develop and optimise the entire seeding assembly in eliminating seeding depth variation.

Acknowledgements

The financial support of GA nr 213-2723/001-001-EM Action2 TIMUR (Training of Individuals through Mobility to EU from the Uzbek Republic) project is gratefully acknowledged. This work was financially supported by the German Federal Ministry of Food and Agriculture (BMEL, grant number 2817ERA09H). The authors are very thankful to Dr. R. Resch and Ch. Gall from AMAZONEN-WERKE H. Dreyer GmbH & Co.KG (Osnabrück, Germany) for providing the seed dose mechanism.

References

Al-Kheer, A. A., Kharmanda, M. G., Hami, A. El, & Mouazen, A. M. (2011). Estimating the variability of tillage forces on a chisel plough shank by modeling the variability of tillage system parameters. Computers and Electronics in Agriculture, 78(1), 61-70. doi:10.1016/j.compag.2011.06.001

Gratton, J., Chen., Y., & Tessier, S. (2003). Design of a spring-loaded downforce system for a no-till seed opener. Canadian Biosystems Engineering, 45(2), 29-35.

Hasimu, A., & Chen, Y. (2014). Soil disturbance and draft force of selected seed openers. Soil and Tillage Research, 140, 48-54. doi:10.1016/j.still.2014.02.011

Lawrance, N. S. (1969). A method of Analyzing Dynamic Responses of A Semi-mounted Farm Implement. Thesis, The Ohio State University, Columbus, OH, USA.

Nielsen, S. K., Munkholm, L. J., Lamandé, M., Nørremark, M., Edwards, G. T. C., & Green, O. (2018). Seed drill depth control system for precision seeding. Computers and Electronics in Agriculture, 144 174-180. doi:10.1016/j.compag.2017.12.008

Özmerzi, A., Karayel, D., & Topakci, M. (2002). Effect of Sowing Depth on Precision Seeder Uniformity. Biosystems Engineering, 82(2), 227-230. doi:10.1006/bioe.2002.0057

Paraforos, D. S., Reutemann, M., Sharipov, G., Werner, R., & Griepentrog, H. W. (2017). Total station data assessment using an industrial robotic arm for dynamic 3D in-field positioning with sub-centimetre accuracy. Computers and Electronics in Agriculture, 136, 166-175 doi:10.1016/j.compag.2017.03.009

Rui, Z., Tao, C., Dandan, H., Dongxing, Z., Kehong, L., Xiaowei, Y., *et al.* (2016). Design of depth-control planting unit with single-side gauge wheel for no-till maize precision planter. International Journal Agriculture and Biological Engineering, Open Access at http://www.ijabe.org, 9(6). doi:10.3965/j.ijabe.20160906.2394

Shahgoli, G., Fielke, J., Saunders, C., & Desbiolles, J. (2010). Simulation of the dynamic behaviour of a tractor-oscillating subsoiler system. Biosystems Engineering, 106(2), 147-155. doi:10.1016/j.biosystemseng.2010.03.002

Sharipov, G. M., Paraforos, D. S., & Griepentrog, H. W. (2017a). Modelling and simulation of the dynamic performance of a no-till seeding assembly with a semi-active damper. Computers and Electronics in Agriculture, 139, 187-197. doi:10.1016/j.compag.2017.05.010

Sharipov, G. M., Paraforos, D. S., & Griepentrog, H. W. (2017b). Modelling and simulation of a no-till seeder vertical motion dynamics for precise seeding depth. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences 8 455-460 doi:10.1017/S2040470017000590

Sharipov, G. M., Paraforos, D. S., & Griepentrog, H. W. (2018). Implementation of a magnetorheological damper on a no-till seeding assembly for optimising seeding depth. Computers and Electronics in Agriculture, 150, 465-475. doi:10.1016/j.compag.2018.05.024

Weatherly, E. T., & Bowers, C. G. (1997). Automatic depth control of a seed planter based on soil drying front sensing. Transactions of ASAE, 40(2), 295-305.

Variation across scales indicates that best progress in crop yields should come from farmer-centric research

R. Sylvester-Bradley[1*], S. Clarke[2], D. Kindred[1], S. Roques[1], P. Berry[3] and S. Welham[4]
[1]*ADAS Boxworth, Cambridge, CB23 4NN, United Kingdom; roger.sylvester-bradley@adas.co.uk*
[2]*ADAS Gleadthorpe, Meden Vale, Mansfield, NG20 9PD, United Kingdom*
[3]*ADAS High Mowthorpe, Duggleby, Malton, YO17 8BP, United Kingdom*
[4]*Stats4Biol Consultancy, Welwyn Garden City, AL8 6EB, United Kingdom*

Abstract

Four multi-site datasets with 188 to 14,155 wheat yields from the UK showed two- to five-fold variation. This was partitioned using REML into predictable and unpredictable effects, with the latter being ostensibly associated with season (year), and exceeding predictable variation. Of the predictable factors, 'farm' accounted for most variation in each case. Effects of variety choice and other husbandry factors were small. Some of the farm × season variation must have arisen from inter- and intra-field variation as well as from interactions with season; however, the extent of farm-scale variation suggests that progress in yield enhancement must come through research at farm-scale, to understand farmers' attitudes and designs, as well as their soils, environments and husbandry.

Keywords: yield variation, spatial scale, farmer-centric research, on-farm experiments

Introduction

Progress in crop productivity at national (UK), regional (Europe) and indeed global scales falls far short of the rate required to sustain the global population trajectory (>2%/year; Ray *et al.*, 2013). Yet estimates of potential crop yields (Foulkes *et al.*, 2009) and farmers' best yields (Kindred and Sylvester-Bradley, 2019) both indicate ample scope for yield enhancement. Much store is placed by electorates, governments and crop technologists in digitally-facilitated 'precision agriculture' to effect crop yield enhancement, hence food security (NASEM, 2018; Gove, 2018). However, the precision of husbandry decisions is not, of itself, an obvious precursor of higher yields; evidence that fine variations in husbandry elicit significant yield enhancements is unconvincing (e.g. Kindred *et al.*, 2018). So how could precision agriculture bring a resumption of crop yield enhancement in the many global regions where yields have stagnated? The hypothesis explored here is that most controllable yield variation occurs at farm- and field-scale. If this is shown to be true, the main value in precision technologies might be in using them to analyse, understand and prove farm-specific differences; it would seem that digitally-enabled field-scale experiments (Sylvester-Bradley *et al.*, 2018) could reveal to each farm its bespoke best practices and system designs.

Several initiatives to enable and develop 'farmer-centric research' are now emerging (Cook *et al.*, 2013; EU, 2017; Macmillan and Benton, 2014; Sylvester-Bradley *et al.*, 2018), driven by the assumption that development of crop husbandry to increase crop yields must make fastest progress if research is 'up-scaled' such that the scale of yield comparisons becomes similar to that at which most variation arises, particularly anthropogenic variation. Perhaps inevitably, farmers themselves appear to be responsible for making the most influential decisions on crop yields. However, in designing their individual farming systems and imposing their design decisions across whole farms, most farmers confound multiple husbandry choices. So simple analyses of associations between husbandry and crop performance at the farm scale can prove problematic and inconclusive; experimentation is needed. It is thus important to assess the scale at which agronomic research and experimentation should focus. This paper reports an initial attempt at estimating how the evidently large variability

DOI 10.3920/978-90-8686-888-9_113, © Wageningen Academic Publishers 2019

in crop yields is partitioned between husbandry factors and across spatial scales. For this initial exercise, the data used were those that describe yields of wheat in the UK and were easily available to the authors. Analyses embrace temporal variation, but this is largely used to provide replication of spatial effects; temporal trends are not considered here.

Data sources

Four separate datasets were collated describing annual yields of winter wheat in the UK during the past 17 growing seasons (identified by their harvest years). Spatial scales of individual yield values in these four datasets ranged from 50 m^2 to whole tramlines, whole fields or many farms (Table 1). None of the datasets was originally created to study spatial variation; all datasets included missing information and combinations of spatial units and levels of husbandry factors were often incompletely represented. The datasets, with their strengths and weaknesses, were as follows:

1. RL (Recommended List; AHDB, 2018): Approximately 17,500 yields, each being an average of three replicate plots totalling ~50 m^2 of a common variety from each of 477 trials conducted in all harvest years between 2002 and 2018. The trials were sited on ~28 well-managed farms per year in wheat-growing regions throughout the UK. Trials were managed by the AHDB or their agents according to a common protocol each comparing 26-54 (mean 37) varieties to inform variety recommendations and choice. The standard protocol for and data from these trials are publicly available via the AHDB website (2015; 2018). The main strengths of this dataset were in its size, hence likely capacity to show multiple effects, its full representation of all UK regions, its standard methodology, and the extent of genetic variation that it addressed. Shortcomings of this dataset were that some soil and husbandry details were not recorded in the early years, varieties changed gradually over seasons, and some farms differed between seasons. To meet the capacity of the analytical software, varieties with less than 60 yield values were omitted from the analysis, reducing the dataset size to 14,155 yields.
2. LearN: ~580 yields were obtained from farmers' tramline trials, set up to compare effects of different rates of fertiliser nitrogen (N) (Kindred *et al.*, 2018). Yields here were from each of four tramlines per field (~0.2 ha), two treated with the farmer's standard rate of N (ranging from 174 to 352 kg/ha), one treated with 60 kg/ha less and the other treated with 60 kg/ha more. Three tramline trials were conducted on each of about 16 farms in each year over the four harvest

Table 1. Structure of multi-site datasets describing field yields of wheat in the UK. Note some yield values had missing explanatory data.

Dataset	RL	LearN	YEN	Defra census
Spatial unit	>45 m^2	>0.2 ha	>2 ha	multi-farm
Harvest years, AD	2002-2018	2014-2017	2013-2018	1999-2017
Regions[a]	7	6	10[b]	11[c]
Counties	28	12	40	NA
Farms	168	17	123[c]	NA
Varieties	118d[d]	17[e]	52e[e]	NA
No. yield values	14,155	575	587	209

[a] AHDB, 2015.
[b] As AHDB but incl. 3 non-UK.
[c] Defra, 2017.
[d] With >60 yield values.
[e] With >1 yield value.

years from 2014 to 2017. Farms were selected from 55 candidates, avoiding those with much organic manure use, to provide groups representing land of different soil types across England. The main strengths of this dataset were its inclusion of the same farms over seasons, inclusion of more than one field per farm and season, hence the ability to test for farm and field effects, and intra-field variation, and effects of N fertiliser. Shortcomings of this dataset were the omission of the AHDB's West region, partial confounding of variety choice and manure use with farm, and determination of yields using monitors on harvesters (which may not have been well calibrated).

3. YEN: 587 yields were collated from farmers' entries in the Yield Enhancement Network's competitions in harvests 2013 through to 2018 (Sylvester-Bradley and Kindred, 2014). Each yield was from an area exceeding 2 ha and often from a whole field; mean area per entry was 15 ha. Fields and husbandry treatments were chosen by the farmers, according to their own experience and advice, usually to represent the best yields from their farm, but sometimes to represent typical yields. Strengths of this dataset were its wide geographical range, representation of real farming conditions, and trustworthy yields (all were from a weigh-bridge and witnessed). Shortcomings of this dataset were in the proportion of husbandry data missing, and the limited number of fields and seasons per farm (Table 1).
4. Defra: 188 average regional yields were derived from the government census of UK agriculture (Defra, 2018) for harvests between 1999 and 2017; data are routinely collated from ~30,000 farms according to a standard methodology (Defra, 2017) but the smallest production unit of published data is a region; individual farm yields are not published. The majority (>98% by area) of wheat grown in the UK is sown in the autumn but these national and regional yields include small areas of spring-sown wheat. This dataset described all UK farms, whereas the other datasets were restricted to farms and farmers engaged with the research community. This dataset was included to compare regional with seasonal variation; clearly it lacks any fine-scale information, so could only be used for this one purpose.

These datasets were analysed using the REML directive in Genstat (Payne *et al.*, 2017) to assess the predominant factors contributing to the variation in yield. Each factor in the analysis was included as a random, rather than a fixed term. The term 'region' was consistent between RL, LearN and YEN datasets, except that YEN data included three additional non-UK regions (Ireland; the Netherlands, and Other); regions in the Defra dataset employed similar sized but slightly different divisions to AHDB regions (Defra, 2017).

Results

All seasons represented in these datasets post-dated the start of 'yield stagnation' in the UK (Knight *et al.*, 2012) so, although significant seasonal variation was evident, with lowest yields in 2012 and highest in 2015, Defra and RL data showed no significant linear trends through successive seasons (Figure 1), and the other datasets were too short to detect linear trends. Mean yield in the Defra dataset was 7.6 t/ha, compared to 10.4 to 11.4 t/ha in datasets which arose from trials and research-centric farms (Table 2).

Yield variation in the Defra dataset was significantly less than in the other datasets (Table 2) because of the much larger crop area represented by each value, and the much more numerous source data from which these values were derived. Variation in the other datasets was larger and similar; the range was largest in the RL dataset probably due to the much greater quantity of data. At least a 3-fold range in yield occurred within each of these datasets, with maxima being more than double the national average of 7.6 t/ha and similar to the world record wheat yield of 16.8 t/ha harvested in New Zealand in 2017 (Guinness World Records, 2017). This degree of variation appeared to offer ample scope for analysis and attribution to different scales and causes.

Inclusion of the Defra dataset, whilst adding nothing on farm and husbandry effects, allowed a precise assessment of regional effects, with yields in the east and north (8.0 t/ha) exceeding those

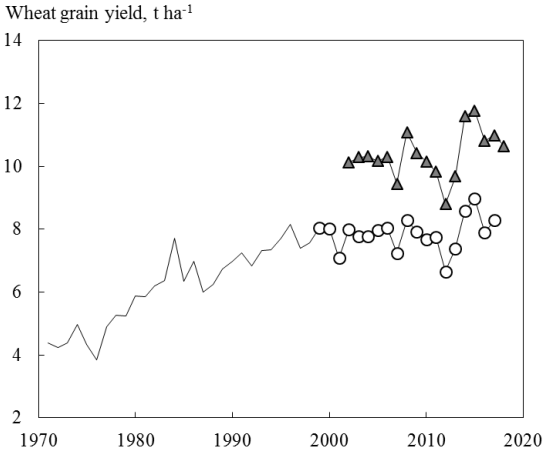


Figure 1. Seasonal yield variation in the Defra (open circles) and RL (closed triangles) datasets. The phase of increasing yields from 1971 is shown for the Defra dataset (no symbol) but was not included in the analyses here.

Table 2. Summary of variation in wheat yields within UK multi-site datasets.

	RL	LearN	YEN	Defra census
Mean yield, t/ha	10.40	11.43	10.74	7.57
SD, t/ha	1.82	1.77	1.88	0.84
Max. yield, t/ha	18.62	16.07	16.50	9.33
Min. yield, t/ha	3.45	6.56	5.01	4.97

in the west and south by 0.9 t/ha; regional effects were not so clear in the more detailed datasets (Table 3) and although county-level effects may have masked regional effects, county comparisons were not included in analyses here because there were insufficient farms within counties to support separate assessments.

Variation in the three detailed datasets was initially partitioned between factors considered to be unpredictable (i.e. variation involving season, plus residual variation) and factors that are explicable and predictable, or might become so through further research (Table 3). In all datasets, and particularly in the RL, unpredictable variation exceeded predictable variation. Given that there was no systematic replication within any dataset, residual variation was generally small, being comprised of interactions other than those explicitly identified, so season and its interactions with farm and region were largely responsible for the unpredictable yield variation. The origin of this must be not only real interactions with season but also intra-farm spatial variation at field and sub-field scales, especially in the RL where yield values arose from only ~50 m^2 land.

The RL dataset was much the largest and so offered most scope to assess the importance of explanatory factors, but in all three datasets 'farm' contributed most to the predictable variation (Table 3). In the RL dataset, farm differences contributed 11% of the total variance. Of the farms represented in four or more seasons within the RL dataset, the range of mean yields was from 8.5 to 12.5 t/ha, and the average SD for these sites was 1.48 t/ha. Thus, whilst the farm effect was not

Table 3. Proportions (*V*) of total variance attributed by REML analysis to a range of predictable spatial, agronomic and genetic factors, and to unpredictable (seasonal and residual) effects in four datasets describing UK wheat yields. SE: standard error. NA: factors not included in the analysis.

Dataset	AHDB RL		LearN		YEN		Defra	
	V	SE	*V*	SE	*V*	SE	*V*	SE
Predictable								
Region	0.01	0.02	0.00		0.01	0.04	0.49	0.22
Farm (= site)	0.11	0.04	0.22	0.13	0.24	0.07	NA	
Field	NA		0.12	0.06	NA		NA	
Soil type	0.00	0.01	0.00		0.11	0.08	NA	
Previous crop	0.09	0.05	0.02	0.03	0.06	0.04	NA	
Manure use	NA		0.04	0.05	0.01	0.01	NA	
Variety	0.04	0.01	0.03	0.04	0.00		NA	
Variety × Region	0.00	0.00	NA		NA		NA	
Variety × Farm	0.01	0.00	NA		NA		NA	
Fertiliser N[1]	NA		0.02	0.02	0.00		NA	
Unpredictable								
Season (= year)	0.11	0.05	0.22	0.20	0.19	0.13	0.33	0.11
Season × Region	0.05	0.03	0.00		0.00		0.18	0.02
Season × Farm	0.56	0.05	0.17	0.07	0.09	0.06	NA	
Season × Other	0.01	0.00	0.10	0.05	0.01	0.01	NA	
Residual	0.01	0.00	0.06	0.01	0.28	0.13	NA	
Total	1.00		1.00		1.00		1.00	

[1] N policy for LearN (low, normal, high); 50 kg/ha N groups for YEN.

immediately evident when comparing yields from small parts of individual fields, it nevertheless proved to account for the largest proportion of predictable variance.

Previous crop was significant in the RL dataset, and more important than soil type whereas soil type was more important in the YEN dataset. Variety was a significant factor in the RL & LearN datasets but accounted for only ~4% of total yield variance. Variety was not significant in YEN.

The LearN dataset provided replication across tramlines within fields and fields within farms and hence enabled a more precise assessment of effects at farm and sub-farm spatial scales. Both proved important, with the field effect being more certain. Surprisingly, in the LearN dataset, although there were effects of fertiliser N in all tramline trials (Kindred *et al.*, 2018), the contribution of N fertiliser treatments to total variance was very small compared to the effects of farm and the large but uncertain effects of season.

As with the RL and the LearN datasets, the effect of farm in the YEN dataset proved as large and relatively more consistent than any other factor. There were 123 farms with repeat entries; only 19 of these were responsible for more than 3 entries, and the definition of a farm in the YEN database was more uncertain than in the other datasets. Husbandry factors did not account for significant proportions of the variance although analysing factors as fixed effects showed significant associations of yield with soil type, previous crop, fertiliser P and N, fungicide use, frequency of plant growth regulator use, use of slurry and form of fertiliser N (Kindred and Sylvester-Bradley, 2019).

Discussion

There is no need for an exhaustive study to reveal how much crop yields vary; farmers, all those that supply them, and all those that they supply, are eminently aware of the extreme variation in crop yields, even in the UK where serious droughts are infrequent. It is an indictment of the farming and crop science industries that much of this variation still remains unrecorded, unanalysed and unexplained. The apparent dependence of yields on seasonal weather, and the focus of research at finer scales, have contrived to thwart the clear assessment, analysis and attribution of yield variation between predictable and unpredictable factors and between different spatial scales, as well as between wide ranging husbandry practices.

All datasets examined included good representation across regions, and husbandry effects were well represented, yet it is evident that a farm effect was very influential. This was despite farming factors such as soil type, rotation, manure use, varietal choice and fertiliser N use being extracted separately. Of course, intra-field yield variation is also often large, and it is a moot point which of farm-scale or field-scale variation is likely to be the more tractable through research investment. Research into farm effects tends to be problematic because experimentation is difficult, whilst research into variable rates of nutrients or seeds has shown how difficult it is to counteract inherent intra-field variation with single inputs (e.g. Kindred *et al.*, 2017). Since yield is a multi-faceted variable, it is likely that future research must go beyond simple conventional considerations of crop genetics, establishment, nutrition and protection. The findings here indicate that farmers' systems, designs and attitudes must be examined for their influences on yield, as well as research at different spatial scales, so that the combinations of factors that account for most of the variation can be identified. Precision Farming technologies are developing apace and, with networks of researchers and farmers working to a common cause, many of these can now be harnessed to facilitate research across scales (Sylvester-Bradley *et al.*, 2018; Marchant *et al.*, 2018).

Much of the yield variation in all datasets examined in this exploratory study originated at scales of a field or larger. Much larger datasets are now available from farm harvesters so, if calibration and analytical issues can be overcome (e.g. Muhammed *et al.*, 2016) and allow robust inter-farm comparisons, further data analyses are justified to validate the primary scales and patterns of yield variation. Clearly, in further analyses, it will be important to recognise that farm differences may be as much due to differences between farmers, their skills, attitudes, motivations and behaviours, as they are due to physical farm differences (e.g. machinery, soils & weather). If the findings here are confirmed, great potential for yield progress should arise through better understanding of farm or farmer-mediated effects on crop productivity.

Conclusions

Given the urgency of increasing food production and the apparent extent of field, farm and farmer effects on crop productivity, there is a strong case for transferring substantial investments into farmer-centric research. It should be noted that research using conventional small plots is unable to test hypotheses that might explain inter-farm, inter-field and inter-management zone yield effects. It will therefore be important to continue with development of organisations, methodologies and technologies that enable research at larger scales. The strong influence of 'farm' indicates that analysis and testing of social and behavioural issues as well as physical contrasts should be included. Farmer-centric research has great potential and is required urgently.

Acknowledgements

We are grateful to the AHDB for publishing cleaned data from the Recommended List trials, for funding the LearN project and for sponsoring the YEN. We are grateful to all 17 sponsors of

the Cereal YEN for permitting us to publish this analysis. (http://www.yen.adas.co.uk/Sponsors/CEREALYEN.aspx)

References

AHDB (2015) Recommended List Cereal trials protocol 2017-21. 51 pp. https://cereals.ahdb.org.uk/media/1095071/Protocol-CER-17-21-AHDB-RL-Cereal-trials-H17-21-October-2015-.docx (last accessed Dec 2018)

AHDB (2018) Recommended List and Harvest Results archive. at https://cereals.ahdb.org.uk/varieties/current-trials-and-harvest-results/archive.aspx. (last accessed Dec 2018)

Cook, S., Cock, J., Oberthür, T. & Fisher, M. (2013) On-farm experimentation. Better Crops with Plant Food 97, 17-20.

Defra (2017). June survey of agriculture and horticulture in England – methodology. https://www.gov.uk/guidance/structure-of-the-agricultural-industry-survey-notes-and-guidance (last accessed Dec 2018)

Defra (2018). Farming Statistics. https://www.gov.uk/government/statistics/farming-statistics-final-crop-areas-yields-livestock-populations-and-agricultural-workforce-at-1-june-2017-uk (last accessed Dec 2018)

EU (2017). A strategic approach to EU agricultural research innovation. Final outcome of the European Conference: 'Designing the path' 26-28 January 2016, Brussels. https://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/agri_strategypaper_web_1.pdf (last accessed Dec 2018)

Foulkes, M.J., Reynolds, M.P. & Sylvester-Bradley, R. (2009). Genetic Improvement of Grain Crops: Yield Potential. 5 In V. Sadras, & D. Calderini (Eds.) Crop Physiology: Applications for Genetic Improvement and Agronomy. San Diego, USA: Academic Press. pp 355-38

Gove, M. (2018). Farming for the next generation. https://www.gov.uk/government/speeches/farming-for-the-next-generation (last accessed Dec 2018)

Guinness World Records (2017). Highest wheat yield. http://www.guinnessworldrecords.com/world-records/highest-wheat-yield (last accessed Dec 2018)

Kindred D., Clarke S., Sylvester-Bradley R., Hatley D., Roques S., Morris N., *et al.* (2018). Using farm experience to improve N management for wheat (LearN). AHDB Report No. 596. pp. 82 https://cereals.ahdb.org.uk/media/1420839/pr596-final-project-report.pdf. (last accessed Dec 2018)

Kindred D., Sylvester-Bradley R., Milne A., Marchant B., Hatley D., Kendall S., *et al.* (2017). Spatial variation in nitrogen requirements of cereals, and their interpretation. In JA Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of 11th European Conference on Precision Agriculture. Advances in Animal Biosciences 8, 303-307.

Kindred, D.R. & Sylvester-Bradley, R. (2019). https://www.adas.uk/News/attention-to-detail-is-this-the-key-to-high-yields-in-cereals (last accessed Feb 2019)

Knight, S., Kightley, S., Bingham, I., Hoad, S., Lang, B., Philpott, H., *et al.* (2012). Desk study to evaluate contributory causes of the current 'yield plateau' in wheat and oilseed rape. Stoneleigh, UK: AHDB. Report 502, pp. 225.

Marchant, B., Rudolph, S., Roques, S., Kindred, D., Gillingham, V., Welham, S., *et al.* (2018). Establishing the precision and robustness of farmers' crop experiments. Field Crops Research 230, 31-45.

McCullagh, P. & Clifford, D. (2006). Evidence for conformal invariance of crop yields. Proceedings of the Royal Society of London, Series A 462, 2119-2143.

Macmillan, T. & Benton, T.G. (2014). Engage farmers in research. Nature 509, 25-27.

Muhammed, S., Milne, A., Marchant, B., Griffin, S. & Whitmore, A. (2016). Exploiting yield maps and soil management zones. Report 565. Stoneleigh, UK: AHDB.

NASEM (2018). Science breakthroughs to advance food and agricultural research by 2030. Washington DC, USA: The National Academies Press. doi.org/10.17226/25059.

Payne, R., Welham, S. & Harding, S. (2017). A Guide to REML in Genstat® (19th Edition). VSN International, 101 pp.

Ray, D.K., Mueller, N.D., West, P.C. & Foley, J.A. (2013). Yield trends are insufficient to double global crop production by 2050. PLoS ONE 8(6): e66428.

Sylvester-Bradley, R., Kindred, D. & Berry, P. (2018). Agronōmics: eliciting food security from big data, big ideas and small farms. In Proceedings of the 14th International Conference on Precision Agriculture. https://www.ispag.org/proceedings/?action=abstract&id=4830 (last Accessed Dec 2018)

Site-specific treatment responses in on-farm precision experimentation

R.G. Trevisan[1*], D.S. Bullock[2] and N.F. Martin[1]
[1]*University of Illinois at Urbana Champaign (UIUC), Crop Sciences, W201 Turner Hall, 1102 S. Goodwin Ave, Urbana, IL 61801, USA; rodrigo7@illinois.edu*
[2]*UIUC, Agricultural, and Consumer Economics, 339 Mumford Hall, 1301 W. Gregory, Urbana, IL 61801, USA*

Abstract

There is a large body of research on characterizing field variability of soil characteristics and their impact on crop yields. In contrast, site-specific information about crop responses to agronomic treatments is less frequent. On-farm precision experimentation (OFPE) brings important information to understand the spatial variation of crop response to agronomic practices and thus to improve agronomic decisions. The objective of this work was to investigate the spatial variability of corn yield responses to nitrogen and seed rates using OFPE in four fields in the US Midwest. Geographically-weighted regression was applied to generate local regression coefficients, which were used to delineate response zones in each field. The results showed the existence of great potential to adjust the rates of these inputs according to the response of each zone identified by the proposed method. The results of this study can be applied to re-evaluate expectations on variable rate prescriptions guided largely by soil and variability.

Keywords: *Zea mays* L., nitrogen, chessboard design, geographically-weighted regression, yield response functions

Introduction

In recent decades, site-specific technologies including yield monitors, remote sensing imaging and variable rate input application have become increasingly available to farmers. Many farmers have access to spatial information, and even access to software tools to process this information to aid the decision-making process. However, the process is usually aided by a knowledge base largely supported by agronomic trials, which involves changing management practices and subsequent monitoring of the effect on the system output (Pringle *et al.,* 2004). Traditional agronomic experimentation is usually conducted in small plots that by design are managed to generate broad inferences from small field areas. This scenario was conceived in a time when all experimental steps were labor-intensive, restricting the feasible extension of the plots. For decisions interacting with environment and management, the results obtained in this experiment may not align with farmers' field conditions. To improve the adoption of site-specific management, it is necessary to obtain additional information about how site-specific treatments impact crop responses (Bullock and Bullock, 1994). Recent advances in GNSS-based technologies have enabled agronomic experiments to be applied on-farm and in much larger areas than was previously possible. Such experiments generate large amounts of site-specific response data that can be used to understand the spatial variation of optimum rates. This involves approximating site-specific yield response functions, which requires methods of spatial analysis (Bullock and Lowenberg-DeBoer, 2007). Site-specific prescriptions of inputs often make the assumption that crop responses are expressed by a linear function of yield or some soil parameter, given that there is insufficient information about site-specific responses to farm management. The Data Intensive Farm Management (DIFM) project has developed a platform to carry out many on-farm experiments and better understand site-specific yield responses.

DOI 10.3920/978-90-8686-888-9_114, © Wageningen Academic Publishers 2019

The most prominent field design applied at the DIFM is the chessboard arrangement covering the entire field. This arrangement generated an unprecedented amount of information, requiring appropriate statistical methods to extract meaningful information. Geographically-weighted regression (GWR) is a statistical methodology that can be used to highlight local differences in the relationship between spatial variables. The use of this data analysis technique is particularly suitable for OFPE field arrangements since the treatments are randomly assigned to entire fields. In GWR, the parameters of the models are allowed to vary in space and can be mapped and interpreted as a spatial variable (Fotheringham, 1997; Plant, 2012). It can be used as an effective method to evaluate the assumption of stationarity in situations of significant variation of regression coefficients. Therefore, the objective of this study was to apply GWR to evaluate spatial variability of corn response to nitrogen and seed rates in on-farm precision experiments.

Materials and methods

Field experiments

The datasets used in this research were generated from corn (*Zea maiz* L.) on-farm trials conducted under the DIFM project. The fields were chosen to represent typical corn production systems in the Corn Belt, with two rainfed fields from Illinois and two central pivot irrigated fields from Nebraska, in the 2017 growing season (Table 1). Each of these fields had completely randomized designs (CRD) with at least four levels of corn seed rates (SR) and four levels of nitrogen rates (NR). Each field had on average 40 ha and 200 experimental units. All treatment levels were implemented in the field using variable rate-enabled seed planters and nitrogen applicators. Each experimental unit was 85 m long and about 18 m wide, representing two times the machinery swath width. Other farming practices were kept constant throughout the field and were conducted by the farmers in accordance with standard protocols for the region.

Data collection and processing

Yield data was collected during harvest using combine yield mapping systems. Digital elevation maps (DEM) and soil electrical conductivity (EC) were used to characterize field variability. DEM maps were created by interpolating the elevation obtained by the real-time kinematic global navigation satellite system receiver (RTK GNSS). Other topographical indices were derived from the DEM to represent slope, curvature and terrain wetness index. EC data were collected previous to planting using a Veris U3 sensor system (Veris Technologies, Inc. Salina, KS, USA). The yield point co-ordinates were used to sample the information from the interpolated EC and DEM maps. The length of the experimental unit was further divided to produce almost squared polygons (Figure 1A). These were used as the observational units, and all data were averaged to these polygons with approximately 400 m^2 (Figure 1B).

Table 1. Description of the four cornfields used for the on-farm precision experimentation.[1]

Field	Location	Irrig.	Seed rate seeds/ha (×1000)	N rate (kg/ha)	ECS (mS/m)	Elev. (m)	Yield (Mg/ha)
1	Effingham – IL	No	66-96	78-129	20 (4)	180	11.2 (1.7)
2	Moultrie – IL	No	76-96	96-147	28 (5)	210	11.9 (0.4)
3	Hamilton – NE	Yes	69-89	24-99	40 (7)	572	14.9 (0.6)
4	Clay – NE	Yes	69-91	0-90	49 (6)	522	15.9 (0.8)

[1] Irrig: Irrigation; ECS: Soil electrical conductivity (0-0.30 m); Elev: Elevation; Numbers inside parenthesis are the standard deviations.

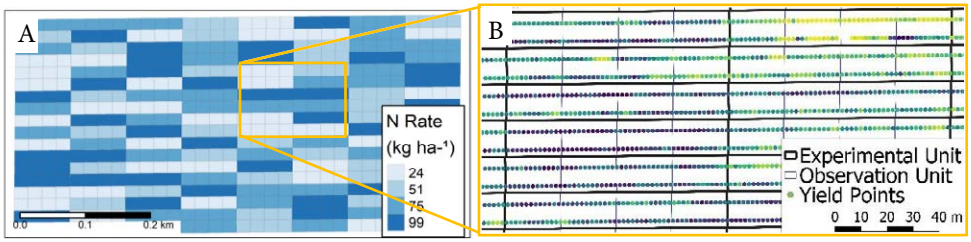


Figure 1. Field trial arrangement of the on-farm precision experiment in Field 3 at Hamilton-NE (A), and subsection of the field illustrating yield monitor data and the observational units (B).

Methodology of analysis

Based on ordinary least squares, GWR modifies the estimates of intercept (β_0) and regression coefficients (β_1) as a function of the geographic position (x_i,y_i) (Equation 1).

$$Y_i = \beta_0(x_i, y_i) + \beta_1(x_i, y_i) + \varepsilon_i, \qquad i = 1, \ldots, n. \tag{1}$$

Following matrix notation regression coefficients are estimated as:

$$b(x_i, y_i) = (X'G(x_i, y_i)X)^{-1} X'G(x_i, y_i)Y, \tag{2}$$

Where Y is the vector whose components are the Y_i and X is a matrix containing the values of the explanatory variables X_i. In this notation, the GWR model incorporates local variation in the estimated coefficients $b(x_i, y_i)$ by incorporating a geographic weighting matrix G(x_i, y_i). This diagonal matrix incorporates the distance d_{ij} between predictors X_i at (x_i, y_i) and dependent observations Y_j at location (x_j, y_j). As d_{ij} increases, the explanatory variable is expected to decrease in influence over the response variable (Plant, 2012). This decay is captured by elements g_{jj} of the matrix G (Equation 3).

$$g_{ij}(x_i, y_i) = exp\left(-\left[\frac{d_{ij}}{h}\right]^2\right) \tag{3}$$

For this study, a Gaussian adaptive spatial kernel was used, with a bandwidth h set to include for each location several neighbors corresponding to 5% of the total number of data points. Following these specifications, GWR analysis was conducted using the R package spgwr (Bivand and Yu, 2017). The regression model had yield as the dependent variable and the NR and SR effects as the independent variables, plus an intercept. Therefore, for each co-ordinate in the center of the Gaussian kernel, three model parameters were estimated, as well as statistical metrics such as the standard error and the R-squared.

The resulting coefficients of NR and SR were standardized between 0 and 1, by subtracting the minimum and dividing by the maximum value. These coefficients were labeled as seed rate response index (SRI) and nitrogen rate response index (NRI). As the index increases, so does the responsiveness to treatments. These indices were submitted to unsupervised k-means clustering to create three response zones for each field to better illustrate the differences. The indices were also compared with auxiliary variables representing the spatial variability of soil and terrain parameters, using correlation analysis.

Results and discussion

Overall yield values were higher in the irrigated fields in Nebraska, averaging 15.4 Mg/ha, than the non-irrigated fields in Illinois, averaging 11.4 Mg/ha (Table 1). The coefficient of yield coefficient of variation (CV) was smaller than 10% for most fields, except for Field 1 in which the CV was 15%. The low CVs are probably due to the aggregation of the data in the polygons, which tends to smooth the variations, and to the high average yield observed for the fields, which denotes favorable weather conditions.

Although the overall variation is not large, it is possible to observe the presence of spatial variation in the yield in the fields. The main distinct area in Field 3 (Figure 2A), is the low yielding region in the northwest of the field. Comparing the yield with the results of the GWR analysis in this area for the SR (Figure 2B), most of this area can be considered as highly responsive to the increment in seed rates, indicating a negative correlation between yield and the optimum SR. However, the east side of the field presents the highest yields and also the highest responses to SR. The ability of using the results from the GWR analysis to make this type of observation, which would not be possible without mapping the regression coefficients, is one of the motivations for recommending its application in the analysis of OFPE data.

Similar results regarding the spatial variation of the response index were observed in the four fields used in this study. These results reiterate the importance of incorporating the spatial variability as one of the bases for agronomic experimentation. The presence of high responses in regions with contrasting yield levels can be seen as a good indication that yield potential or target yield alone are not good predictors of the optimum rates. Yield maps of previous years are commonly used as the deciding factor to delineate management zones or to make variable rate prescriptions. However, there is no guarantee that the optimum rate in two given regions with similar yields is in fact the same.

Unsupervised k-means clustering was used to generate yield response zones as an alternative to the traditional yield-based management zone delineation. There is no reason to assume the same zones are optimum for different inputs, therefore distinct response zones were created for SR and for NR in each field. The clusters were labeled as 'High', 'Medium', and 'Low' according to the average response index in each group. The results were consistent across fields in terms of the delineation of meaningful spatial contiguous zones and the different yield response in each zone.

The two fields in Illinois were more responsive to NR than SR. In these fields, the response to NR was positive for the high and medium response zones and no response was observed in the low response zone (Figure 3). In the low response areas of this field, there seems to be no decrease in yield using the lowest N rate at 96 kg/ha in comparison with the higher rates, therefore there exists a potential to apply less N in this region of the field while maintaining the same yields. For the same fields, the response to SR was positive only in the high response zone, no response was observed in the medium response zone and negative response was observed in the low response zone. In this

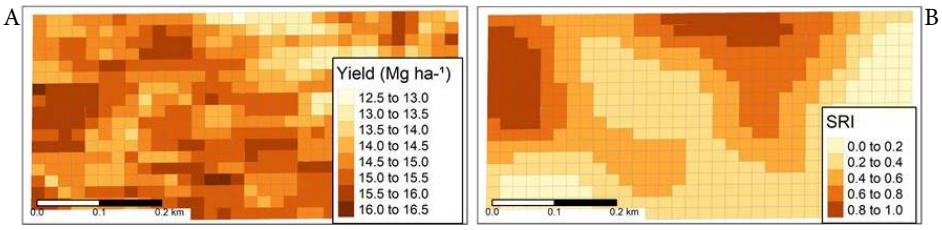


Figure 2. Corn yield (A), and the seed rate response index (B) in Field 3 at Hamilton-NE, for the 2017 crop season.

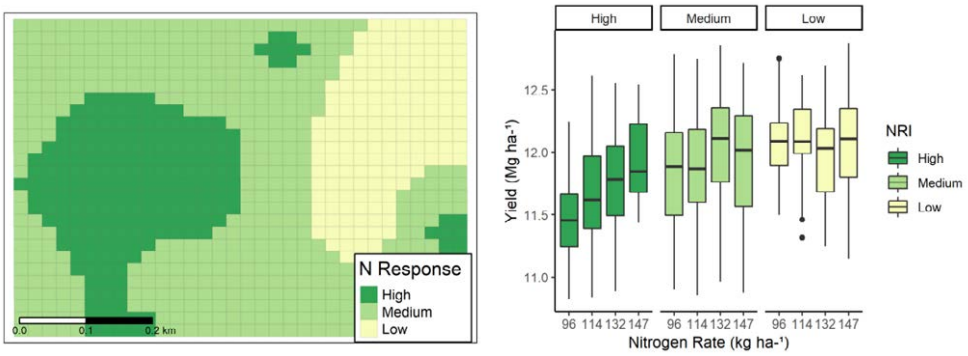


Figure 3. Zones of yield response to nitrogen (NRI) and corresponding box-plots showing the different yield response to nitrogen rate treatments, for the Field 2 at Moultrie, IL.

scenario, the average rate was already optimized for the average response of the field, but there is still potential to increase the yields using the same amount of SR, by redistributing it inside the field according to the response zones.

The two irrigated fields in Nebraska were more responsive to SR than NR. In these fields, the response to SR was positive for the high and medium response zones and no response was observed in the low response zone (Figure 4). In the high response zones, the yield is increased by more than 1,500 kg/ha when the SR is increased from 69k to 89k, in the medium response zone the yield increase for the same variation is still more than 1000 kg/ha while, in the low response areas, of this field the increase is within the variation represented in the boxplot. Therefore, there is a great potential to adjust the seed rates according to the response of each zone. The response to NR in the same fields was positive only in the high response zone, no response was observed in the medium response zone and negative response was observed in the low response zone.

Correlations between the NRI and SRI and different field variables are shown in Table 2. In general, the field characteristics impacting SRI are different than the ones impacting NRI. Elevation had a dominant role impacting SRI with correlations of 0.76 and -0.60 for Field 1 and Field 3, respectively. For most of the fields, the correlation between electrical conductivity and topographic characteristics

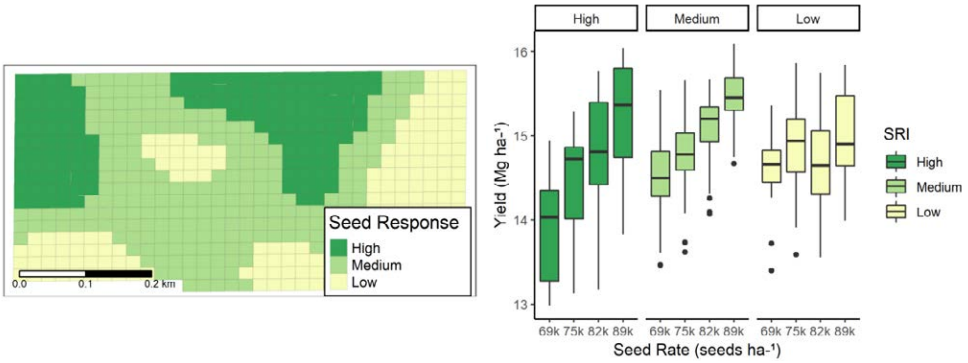


Figure 4. Zones of yield response to seed rate (SRI) and corresponding box-plots showing the different yield response to seed rate treatments, for the Field 3 at Hamilton, NE.

Table 2. Correlation coefficient between the crop response to seed and nitrogen rate and field characteristics.[1]

Field	Variable	ECS	ECD	Elevation	Slope	Curvature	TWI
1	SRI	0.08	0.17	-0.60	0.19	0.00	0.17
	NRI	-0.26	-0.22	-0.02	0.20	-0.04	0.08
2	SRI	0.41	0.30	-0.38	-0.36	0.26	0.11
	NRI	-0.11	-0.01	0.19	0.12	-0.11	0.02
3	SRI	-0.17	-0.20	0.18	-0.05	-0.28	-0.08
	NRI	-0.13	-0.10	0.26	-0.09	-0.38	-0.14
4	SRI	-0.30	-0.28	0.76	0.58	-0.38	0.01
	NRI	0.16	0.08	-0.44	-0.29	0.27	-0.17

[1] ECD = soil electrical conductivity (0-0.90 m); ECS = soil electrical conductivity (0-0.30 m); NRI = nitrogen rate responsiveness index; SRI = seed rate responsiveness index; TWI = terrain wetness index.

were less than 0.30. These low correlations imply that there might be interactions between these field characteristics within the field, impacting these different yield responses to NRI and SRI treatments. Depending on availability, more variables could be added to this list, including the yields and remotely sensed data in previous years. Multivariate statistical methods and other methods of data analysis can also be employed to explore the additive contribution of different factors. Future work will also focus on developing econometrics models to identify optimal treatment rates based upon input prices and site-specific yield responses to NR and SR in fields. In addition, as more fields are added to the DIFM project, relationships between site-specific responses and site-specific characteristics will be further investigated.

Conclusions

The DIFM project has developed a platform to carry out many on-farm experiments and better understand site-specific yield responses. The use of GWR identified differential yield responses to seed rate and nitrogen in cornfields in different producing regions in US Midwest. Therefore, there is great potential to adjust the rates of these inputs according to the response of each zone identified by the proposed method.

The presence of high responses in regions with contrasting yield levels can be seen as a good indication that yield potential or target yield alone might not be good predictors of the optimum rates. These results should be taken into account when setting expectations on variable rate prescriptions guided largely by soil and yield variability.

Acknowledgements

This research was funded in part by a USDA National Institute of Food and Agriculture Food Security Program grant, award number 2016-68004-24769.

References

Bivand, R., & Yu, D. (2017). spgwr: Geographically Weighted Regression. https://cran.r-project.org/package=spgwr (last accessed December 2018)

Bullock, D. S., & Bullock, D. G. (1994). Calculation of optimal nitrogen fertilizer rates. Agronomy Journal, 86(5), 921-923.

Bullock, D. S., & Lowenberg-DeBoer, J. (2007). Using spatial analysis to study the values of variable rate technology and information. Journal of Agricultural Economics, 58 (3), 517-535. doi:10.1111/j.1477-9552.2007.00116.x

Fotheringham, A. S. (1997). Trends in quantitative methods I: stressing the local. Progress in Human Geography, 21 (1), 88-96.

Plant, R. E. (2012). Spatial data analysis in ecology and agriculture using R. Boca Raton, FL, USA: CRC Press.

Pringle, M. J., Cook, S. E., & McBratney, A. B. (2004). Field-scale experiments for site-specific crop management. Part I: Design considerations. Precision Agriculture, 5(6), 617-624. doi:10.1007/s11119-004-6346-1

Section 18
Software and mobile apps for precision agriculture

Easy water stress detection system for vineyard irrigation management

G. Brunel*, L. Pichon, J. Taylor and B. Tisseyre
ITAP, Irstea, Montpellier SupAgro, Univ. of Montpellier, Montpellier, France; guilhem.brunel@supagro.fr

Abstract

The monitoring and management of vine water stress is a major issue for many vineyards. A method based on the observation of shoot tip growth makes it possible to determine the water stress of plants without the drawbacks of standard instrumental methods. However, this 'shoot tip' method is rarely used in the field, as observations and index calculations need to be done manually. Therefore, while the data collection can be relatively quick, data interpretation is not well supported. This article presents a mobile application, 'Apex', developed to answer this challenge. The Apex application allows a simple and fast recording of observations and a pooling of spatialized observations. The developed application makes it possible to envisage the collaborative water stress monitoring with crowdsourcing observations.

Keywords: vine water status, shoot tip index, smartphone application, *Vitis vinifera*

Introduction

Many authors have demonstrated the significant role of water status in vine behaviour and berry composition (Cifre *et al.,* 2005). Monitoring vine water status is of importance for yield and quality management purposes. The vine water status naturally changes over time according to the climatic characteristics of the season. In addition, many authors have shown that it presents a significant spatial variability at different scales: at the within field level (Tisseyre *et al.,* 2005), at the vineyard level (Taylor *et al.,* 2011) and at the regional scale (Baralon *et al.,* 2012). Characterizing the temporal and the spatial variability of vine water status is therefore of paramount importance at these different scales to provide the wine industry with relevant decision support tools: either for irrigation scheduling or for other management practices in the case of non-irrigated vineyards. Several standard instrumental methods have been proposed to measure plant water status (Acevedo-Opazo *et al.,* 2008). However, these are not easy to obtain, since they are manual techniques requiring heavy materials and specific skills. These constraints make systematic spatio-temporal vine water status measurements difficult to carry out and time-consuming. In order to overcome these constraints, simple alternative approaches have been proposed (Acevedo-Opazo *et al.,* 2008). Among them, the shoot tip index (STI) proposed by Trambouze *et al.* (2009) and Rodríguez-Lovelle *et al.* (2009) is of interest as it is based on simple observations of vine shoots. Although more imprecise than standard methods, it has the advantage of being easily achievable without expensive equipment and by inexperienced operators. This method is therefore interesting to monitor temporal and spatial changes in vine water status when several operators may be involved (workers, growers, advisors) in collecting observations.

However, the STI method has some limitations that limits its dissemination: (1) it requires the observers to count the Apex (end of shoots) of approximately 50 shoots and to class them into three different classes (full growth, moderate growth, growth stopped); (2) the STI needs to be calculated based on the Apex observations; and then (3) the observed STI value needs to be recorded and transferred to a database with additional information, like the time and the location of observation. Although simple, these different operations can be cumbersome because they require information to be recorded and a calculation to be made in the field. The aim of this article is to present a mobile

DOI 10.3920/978-90-8686-888-9_115, © Wageningen Academic Publishers 2019

application, called ApeX, whose purpose is to tackle this issue by providing a simple tool for: (1) collecting Apex notations; (2) calculating STI; (3) geo-referencing and timing STI observations; and (4) sending STI observation to a server.

Materials and methods

The specifications of ApeX were defined by involving the end-users in development and the developments followed the precepts of UX Design (Hekkert, 2011). This step made it possible to highlight the characteristics that defined the technical choices of the application. The main ones are summarized hereafter:

- Dissemination: The application must be widely deployable to different organisations and stakeholders (growers, advisors, etc.) whose smartphone (model, operating system, etc.) may be different.
- Accessible: Users are not necessarily familiar with professional applications. The use must be simple and user-friendly while maintaining the data integrity.
- Offline mode: Network coverage is variable according to the areas of use, the application must operate in offline mode as well as in connected mode.

The developments followed the precepts of UX Design, which refers to a person's emotions and attitudes about using a product or service. User experience is dynamic over time due to changing circumstances of usage and changes to individual systems, as well as the wider usage context. User needs can be described in three ways: Why, What and How. 'Why' corresponds to the motivations, values and vision of the context. 'What' describes the features and functionalities necessary to achieve the 'Why'. And finally, 'How' corresponds to the way of handling and aesthetics. As shown in Figure 1, the application design process follows an iterative cycle consisting of four phases based on Design-thinking methods:

- First phase consists of discovering user needs, requirements and constraints.
- Second phase consists of imagining all the scenarios and solutions to meet the needs. At the end of this phase, all participants make a pragmatic choice on the new features to be added.
- Third phase consists of offering the user a visual set that is both operational and functional. The aim is to combine comfort of use without losing efficiency.
- Fourth phase consists of quickly implementing the chosen characteristics and performing tests in real conditions by the end user.

Unlike a traditional approach where a finalized solution is produced after the third phase, here the cycle will be renewed after a test period. This is needed because the user experience is dynamic and

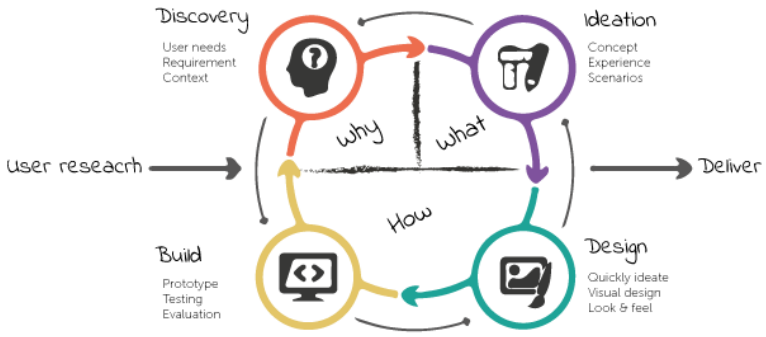


Figure 1. This diagram shows the iterative design process to discover, define, design and develop a digital solution that meets user needs.

constantly changing over time. Four experts were integrated into the ApeX development cycles (6 complete cycles), these users were able to test and validate the various updates.

ApeX was developed as a hybrid app. A hybrid app is one that combines elements of both native and web applications. Hybrid applications execute within wrappers that are targeted to each platform, and rely on standards-compliant application programming interface (API) bindings to access each device's capabilities, such as embedded sensors, data, network status, like a native application. A hybrid app approach was chosen to facilitate the dissemination of the application across more than one platform, without having to re-implement it with each platform's language and tool set.

The development of the application was programmed using Ionic framework based on Apache Cordova (Griffith, 2017). Apache is an open-source mobile development framework, which allows the use of standard web technologies – HTML5, CSS3, and JavaScript – for cross-platform development. AngularJS is a JavaScript-based open-source front-end web application framework to address many of the challenges encountered in developing applications. The JavaScript components complement Apache Cordova. It aims to simplify both the development and the testing of such applications by providing a framework for client-side model-view-controller (MVC) and model-view-view model (MVVM) architectures, along with components commonly used in rich internet applications. The application uses Typescript, a superset of JavaScript, which allows access to a number of extra features, such as type of declarations and interfaces. JavaScript code runs on a NodeJS server, which is a platform built on Chrome's V8 JavaScript engine. It differs from other platforms in that it uses a non-blocking approach to perform I/O asynchronously. This workspace is coupled with Gulp, which offers a toolkit for automating time-consuming tasks in the development workflow (Figure 2). All data created and inserted into the application is synchronized with the ApeX cloud when the smartphone has a 4G or Wi-Fi connection. To operate in areas without a network, all data is stored in a local database supported by SQLite and are automatically synchronized on a MySQL server database when the network allows it. Communication between the smartphone and the server is provided by a JavaScript API. The application uses the smartphone's GNSS to locate observations. The web application was designed and developed to quickly and efficiently visualize the collected data. The web application was developed using HTML5 mark-up language, CSS3 styling and WordPress API for a responsive web design. It automatically adjusts to different screen sizes, allowing it to be easily viewed on any device or at any screen resolution.

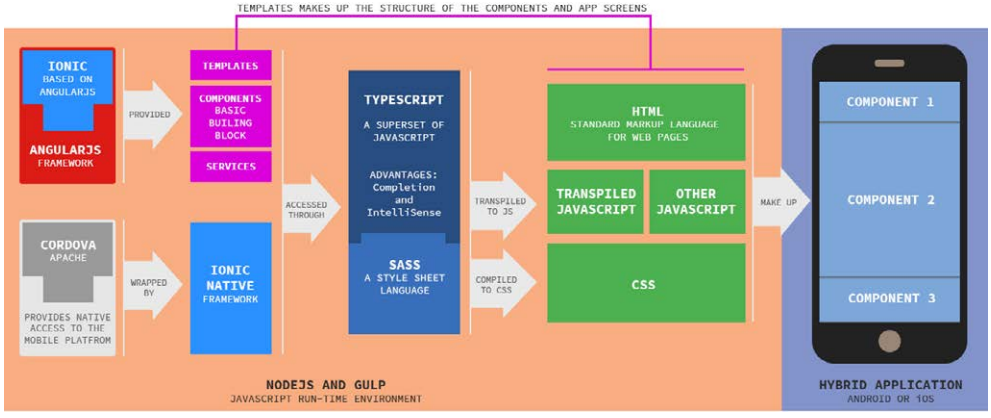


Figure 2. Technologies used for the development of ApeX.

The ApeX app was tested by twenty users in the summer of 2018 in the Occitanie and Provence-Alpes-Côte d'Azur regions of southern France. One session represents fifty observations of shoot tips in a vineyard field, which is necessary for the STI to be calculated to estimate vine water stress. More than 1500 sessions distributed over 270 different vine fields were collected over the season.

Results

Application interface

The calculation page is divided into three areas (Figure 3A): (1) a list to choose the field that the observation is performed in. This list is ordered according to the user's GNSS position, (2) three buttons for shoot tip selection. This area takes up 70% of the screen for comfort of use. Each button has a different vibration to indicate user choice, (3) a dynamic counter at the bottom of the screen to allow the user to know the number of observations made. A long vibration of the phone indicates the achievement of the fifty observations necessary to calculate the STI. The STI button is only clickable after fifty observations to compute STI.

The home page (Figure 3B) navigates to the calculation pages and allow visualization of the results. For each field, the user can access: (1) the current growth rate of the vine (vigour); (2) the growth dynamics based on past and current growing STI rates; and (3) a scale range of the current STI (water stress). To improve the quality of the user experience, other interfaces have been developed (instructions, contact form, field characteristics, etc…).

Spatial distribution of observations

More than 1500 sessions (each of 50 shoot observations), from 22 operators from 7 different companies (extension services, advisers and research institutes) were collected with the ApeX app during 2018. Figure 4 shows the sessions collected at two different scales: the within-field scale (Figure 4A) and the regional scale (Figure 4B).

At the within field scale, (Figure 4A), the 50 observations were plotted to provide more detailed information on the monitoring that has been carried out. Some observations were positioned outside

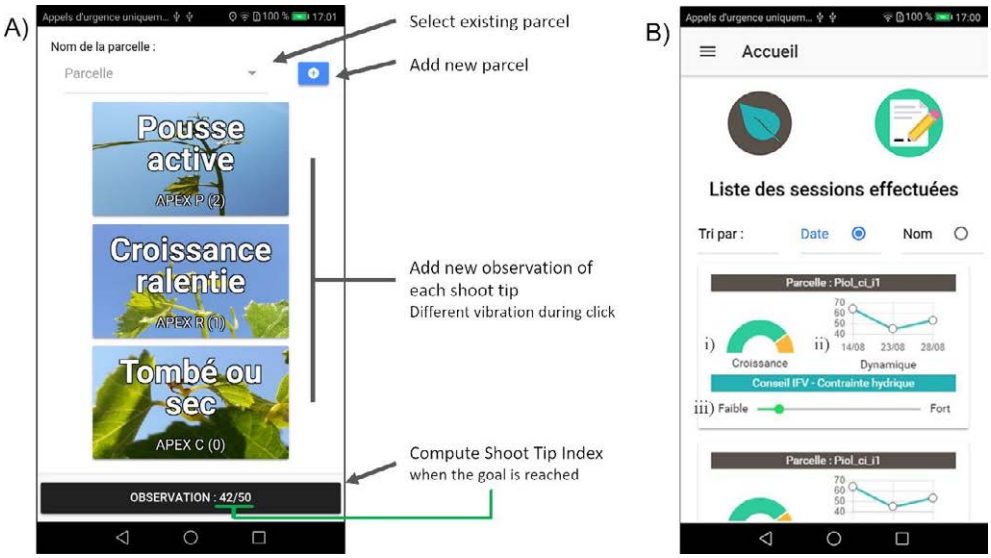


Figure 3. Screenshot of both main interfaces of ApeX.

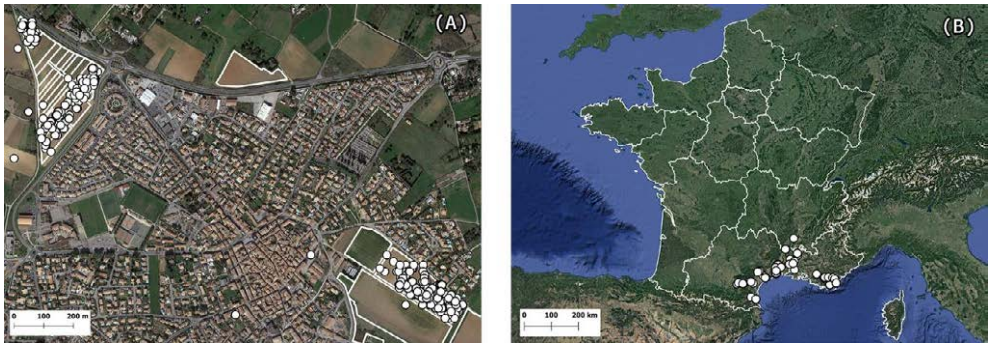


Figure 4. Observations collected with Apex during 2018 in southern France, (A) example at the within-field level, (B) spatial distribution of measurements at the regional level.

the fields. This result can be explained by the low quality of the smartphone's geolocation sensor but also by poor operating conditions (fast switching on of the smartphone, locations with poor positioning signal quality, etc.). All together, these problems resulted in spatial outliers. It highlighted the need to identify and to correct outliers before the app is made available to professionals.
At the regional scale, the data were collected mainly around the Mediterranean basin (Figure 4B). This spatial distribution of the observations was mostly driven by the location of partners who were interested in using the ApeX application and for whom the monitoring of water stress is a major issue.

Temporal distribution of observations
Figure 5A shows the number of sessions collected during the 2018 season for each two-week interval. The first sessions were made at the beginning of June with about fifty sessions over the first two weeks. The number of sessions then increased sharply to ~400 over the same interval at the beginning of July. It then decreased during August, with only a few sessions during September and October. ApeX was predominantly used to decide the date of the first irrigation during the dry summer period when vine water stress reached critical values. In the Mediterranean region, this decision is usually taken in July. The increased and decreased number of observations corresponded respectively to the start and end of a period of risk for the vine. Most users who started using the application kept on using it during the season. This indicated that it met users' expectations and on the other hand that its usability was satisfactory to them.
Figure 5B shows an example of the evolution of the STI indicator over the season in a vineyard. The STI was close to 0 on June 19th, corresponding to significant vegetative growth and no water restriction. The STI increased steadily during June, as vine growth decreased with the occurrence of a significant water restriction. The STI then stabilized at around 60, indicating moderate to weak growth between mid-July and early September. Slight variations in the STI were observed during this period. This information, combined with agronomic expertise, could be used to identify periods of water stress as a decision support vineyard management.

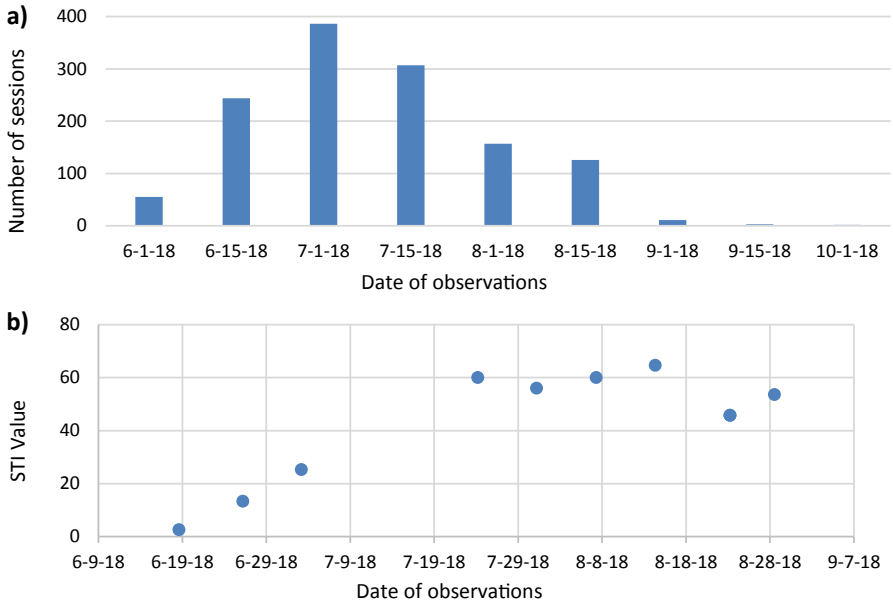


Figure 5. Use of ApeX over the season and examples of results (June – October) (A) the count of sessions made each fortnight and (B) the average recorded STI at approximately weekly intervals in a field.

Discussion

The development of ApeX followed the MVC pattern, in which the user interface (presentation), functional process logic (decision rules code), computer data storage and data access were developed and maintained as independent modules. This application architecture provided a flexible model and a reusable application. By segregating an application into tiers, it was possible to refine the existing model or add new ones based on shoot tip analysis, without having to rebuild the entire application. The specifications of ApeX were realized by involving the end-users and the developments followed the precepts of UX Design. This reduced development times by directly integrating feedback and requests for improvement. A suggestion form was added in the first version so that the user could feedback directly during use. Consequently, some changes were made, for example, vibrations during clicks were added to allow the user to feel actions without having to check the screen at each observation. This basic feature was widely appreciated, as it considerably increased user comfort. Users also felt involved in the application, which explains the success and the large amount of data collected during the first test season.

The results of this study showed that the ApeX application allows the collection of observations on vine growth in a simple and rapid way. These observations were georeferenced and shared at different scales. This work assumed that the ApeX method has been validated by comparison with reference methods in previous work (Trambouze *et al.,* 2009). New experiments will have to be conducted to validate whether the use of the ApeX application improves or alters the relationship with reference measurements. In this work, the use of the application was restricted to a limited number of users, but the number of sessions collected was still relatively large. When the ApeX app is freely released to the wider public, it should collect a very large number of sessions at very different scales. This wider use will be an opportunity to develop a tool to help monitor water stress

at a large scale. However, it does raise the question of the quality of the data that will be produced by the application. This aspect is indeed a major challenge in the use of crowdsourced data (Minet *et al.,* 2017). It will be necessary to set up mechanisms to characterize and improve the quality of the data collected. Finally, it seems relevant to study how this new data source, that provides data in large quantities of low to medium quality, can complement and enrich the reference data.

Conclusions

The ApeX application was developed to monitor vine water status using a simple and robust method that can potentially be used by stakeholders with very different skills. The application is free of charge. It is a novel decision support tool that can be an interesting alternative for small growers who cannot afford alternative commercial services. The version presented in this work is a first approach. However, after a first season of use, the results showed that users have adopted the application throughout the monitoring season. This indicated that it meets a need and that the interface is appropriate.
The app can be used at the field or the within-field scale for temporal or spatio-temporal monitoring of vine water status. It can also be used at a regional scale by aggregating data, although one limitation remains the quality and imprecision of the information obtained with this type of approach.

Acknowledgements

We thank the '*Institut Français de la Vigne et du Vin*' for their involvement throughout the development of the application. We also thank the Chambers of Agriculture who participated in the test campaign during the summer of 2018. The Occitanie region for its financial support in the frame of the crowd-viti project (repere project).

References

Acevedo-Opazo C., Tisseyre B., Ojeda H., Ortega-Farias S. and Guillaume S., 2008. Is it possible to assess the spatial variability of vine water status? Journal International Des Sciences De La Vigne et Du Vin 42, 203-220.

Baralon K., Payam J-C., Salancon E. and Tisseyre, B., 2012. SPIDER: spatial extrapolation of the vine water status at the whole denomination scale from a reference site. Journal International Des Sciences De La Vigne et Du Vin (3), 167-175.

Cifre J., Bota J., Escalona J.M., Medrano H. and Flexas J., 2005. Physiological tools for irrigation scheduling in grapevine (Vitis vinifera L.). An open gate to improve water-use efficiency? Agriculture, Ecosystems & Environment 106, 159-170.

Griffith C. 2017. Mobile App Development with Ionic Revised Edition Cross Platform Apps with Ionic Angular & Cordova. Sebastopol, CA, USA: O'Reilly Media.

Hekkert P., van Dijk M. B. 2011. Vision in design: A guidebook for innovators. Amsterdam, the Netherlands: BIS Publishers.

Minet J., Curnel Y., Gobin A., Goffart J.P., Mélard F., Tychon B., *et al.,* 2017. Crowdsourcing for agricultural applications: A review of uses and opportunities for a farm sourcing approach. Computers and Electronics in Agriculture. 142, Part A, 126-138.

Rodriguez-Lovelle B., Trambouze W., and Jacquet O., 2009. Evaluation de l'état de croissance végétative de la vigne par la « méthode des ApeX » (Evaluation of the vegetative growth status of the vine by the 'ApeX method'). Le Progrès agricole et viticole, 126 (4), 77-88.

Taylor J.A., Acevedo-Opazo C., Pellegrino A., Ojeda H. and Tisseyre B., 2011. A comment on inter-field spatial extrapolation of vine (Vitis vinifera L.) water status. Journal International Des Sciences De La Vigne et Du Vin (2), 121-124.

Tisseyre B., Ojeda H., Carillo N., Deis L. and Heywang M., 2005. Precision viticulture and water status, mapping the pre-dawn water potential to define within vineyard zones. In: H. R. Shultz (Ed.). Proceedings of 14th GESCO congress, Geisenheim, Germany: Groupe d'Etudes des systèmes de Conduite de la Vigne, pp 23-27.

Trambouze W., Rodriguez-Lovelle B. and Jacquet O., 2009. The shoot tip method: a practical tool to evaluate the vine growth status. In: Viticulture & Enology, XVIth international Giesco Symposium. Davis, California, pp 112-120.

A model for precision irrigation scheduling of soybeans for the South-eastern U.S.

V. Liakos[1*], W. Porter[1], J. Kichler[2], A. Sawyer[3], D. Pavlou[1], A. Orfanou[1] and G. Vellidis[1]
[1]*Deprtment of Crop & Soil Science, University of Georgia, Tifton, GA, USA; vliakos@uga.edu*
[2]*Colquitt County Extension, University of Georgia, Moultrie, GA, USA*
[3]*Wilcox County Extension, University of Georgia, Rochelle, USA*

Abstract

This paper describes an evapo-transpiration-based model which will be utilized by a smartphone app for scheduling irrigation in soybeans in 2019. The model calculates the daily soil water balance by using weather, soil and irrigation data. The evaluation of the model was made in 2018. A 12 ha soybean field in Georgia was selected for the experiment. The field was divided in half. The one half was irrigated uniformly based on the farmer's irrigation strategy while the other half received irrigation uniformly when the model recommended it. Irrigation management zones (IMZs) were delineated using Sentinel 2 satellite images from the previous two years. The combination of satellite images with irrigation management zones will give the opportunity to study the soil moisture variability among the IMZs in order to apply variable rate next year.

Keywords: efficiency, remote sensing, variability, water deficit, soil water tension

Introduction

The major oilseed crops in U.S are soybeans, sunflower seeds, cottonseeds, canola, rapeseeds and peanuts. Soybeans are dominant, accounting for about 90% of U.S oilseed production (and one third world production).
In the U.S, the irrigated area dedicated to soybean have increased steadily over the past two decades because irrigation serves both to reduce risk of crop loss and to build resiliency and yield stability. Additionally, as competition for water is increasing rapidly in areas normally associated with plentiful water resources, there is evidence of water withdrawals exceeding recharge and deterioration of water quality in many areas worldwide (Siebert *et al.*, 2010; Scanlon *et al.*, 2012). As a result, the organizations representing growers are investing in the development of irrigation scheduling tools which improve irrigation water use efficiency. In response, a significant amount of research has been conducted on this topic.
Crop water needs are a function of their phenological stage. For example, McGuckin *et al.* (1987) optimized irrigation scheduling using accumulated heat units and not a chronological frame. Evapo-transpiration (ET) is an important factor in estimating daily plant water use. Several irrigation scheduling tools have been developed which use estimated crop ET (ETc) to develop irrigation recommendations. These models typically use a crop coefficient (Kc) to calculate ETc from a reference ET (ETo) as shown in Equation 1.

$$ET_c = ETo \times Kc \quad (1)$$

The crop coefficient (Kc) is widely used to estimate crop water use and to schedule irrigation. Kc changes during the life cycle of the plant. For annual crops, it typically begins with small values after emergence and increases to 1.0 or above when the crop has the greatest water demand. Kc decreases as crops reach maturity and begin to senesce. The seminal work on this approach is known as FAO Irrigation and Drainage Paper 56 (Allen *et al.*, 1998).

DOI 10.3920/978-90-8686-888-9_116, © Wageningen Academic Publishers 2019

Although models which use only ETc to estimate irrigation requirements have been used extensively because they are simple, they do not consider the moisture available in the soil profile and do not calculate a soil water balance. Consequently, they do not take into account soil water stored in the soil profile which sometimes leads to over-application of irrigation water. Incorporating soil water balance increases the number of parameters needed as well as the complexity of the model; in general, the more parameters used, the more accurate the model. However, for an irrigation scheduling model to be adopted, the model must balance accuracy with ease of use because crop consultants, farmers or other professionals making daily irrigation decisions require sophisticated yet easy-to-use and easily accessible tools.

The increasing availability of online environmental measurements required for ET-based irrigation schedules (i.e. temperature, solar radiation, relative humidity, wind speed and rainfall) published by local, state and regional weather station networks has facilitated the development of a wide variety of web-based irrigation scheduling tools. The University of Florida's PeanutFARM (peanutfarm.org; Rowland, 2016) is a web-based irrigation scheduling tool for southeastern peanut production. PeanutFARM uses cumulative adjusted growing degree days (aGDD) (Rowland *et al.*, 2006) and ET from weather station networks to estimate crop water use and provides daily irrigation recommendations. Washington State University's AgWeatherNet (http://weather.wsu.edu) uses meteorological data from nine weather station networks in the western United States and Canada to develop irrigation scheduling recommendations for 56 agronomic and horticultural crops. Additional tools are described by Singels and Smith (2006).

Recent technological advances that allow for widespread internet access through handheld devices such as tablets and smartphones provide a platform on which to deliver sophisticated yet easy-to-use ET-based irrigation scheduling tools. The functionality of a smartphone app differs from a web tool in that apps are with the user at all times since they reside on the smartphone, are readily accessible, and engage the user through notifications (Migliaccio *et al.*, 2015, 2016). Some apps use notifications, similar to text messages, to prompt users to respond to critical events and eliminate the need to interact with the tool on a daily basis.

This paper describes a model for irrigation scheduling in soybeans. This model will be used by a smartphone app which will be released in 2019 and will advise farmers when is the appropriate time to irrigate soybeans. The App will not recommend irrigation rates for each irrigation event. The authors' objectives were to develop a novel ET-based irrigation scheduling model for soybeans that requires minimal inputs and will help farmers irrigate soybeans efficiently.

Material and methods

The soybean model is an ET-based soil water balance model. It uses meteorological data, soil parameters, crop phenology, crop coefficients and irrigation applications to estimate root zone soil water deficits (RZSWD) in terms of amount of water and percent of total deficit. The model does not deliver specific irrigation recommendation (irrigation rates) but the model results can be utilized to make appropriate irrigation decisions.

ET and Kc

The soybean model downloads from ftp servers meteorological data from Georgia Automated Environmental Monitoring Network (GAEMN) and Florida Automated Weather Network (FAWN) to calculate reference ET (ETo) using the Penman-Monteith equation (Allen *et al.*, 1998). The model also uses crop coefficient values (Kc) to estimate crop ET (ETc). However, no Kc curves are available for soybeans in southeastern USA (there are available for other regions and can be imported into the model). Thus a prototype Kc curve developed using as reference the Kc curve recommended by the Food and Agriculture Organization of the United Nations (FAO, 2019). Additionally, soil water tension data were used from field experiments which were carried out in 2016 and 2017 at south

Georgia, USA to retroactively calibrate the FAO Kc curve so that 40% RZSWD coincide with the soil water tension of 50 kPa at the root zone. Field experiments took place also in 2018 to calibrate and validate the Kc curve for conditions in southern Georgia and northern Florida. Changes in the model's Kc values are driven by accumulated heat units commonly referred to as growing degree days (GDDs). GDDs are calculated using Equation 2.

$$GDD = [(T_{max} + T_{min)} / 2] - T_{base} \quad (2)$$

The GDD calculation for soybeans considers that T_{base} is equal to 15.5 °C. Any temperature below T_{base} is equal to T_{base} and consequently the GDD is equal to zero. Additionally, daily average air temperature higher than 37.7 °C is considered equal to 37.7 °C because growth is limited above this temperature. The specific GDDs required for each phenological stage of soybeans are based on previous experiments where soybean growth stages were monitored and associated with GDDs in south Georgia (Figure 1).

Soil water balance

The model calculates ETc to estimate the daily crop water use. Variables such as ETc, measured precipitation, irrigation and soil water holding capacity (Figure 2) are used to estimate the plant available soil water (there is no need to use soil moisture sensors). In the model, it is assumed that there is no water deficit during the planting date. Furthermore, the initial rooting zone depth is 152 mm because the planting depth of soybeans is 152 mm and increases by 7.6 mm/day until it reaches a maximum depth of 762 mm. At emergence, the soil profile from 0 mm to 762 mm is assumed to be at 85% of maximum plant available soil water holding capacity. The daily plant-available soil water is calculated by subtracting the previous day's ETc from the previous day's plant available soil water and adding any precipitation or irrigation events. The model uses an effectiveness factor of 85% for all sprinkler irrigation events to account for evaporation and drift before the water droplets reach the soil. The model assumes that 90% of measured precipitation reaches the soil to account for canopy interception and other possible losses. All these parameters are used to calculate root zone soil water deficit (RZSWD) in mm and % RZSWD.

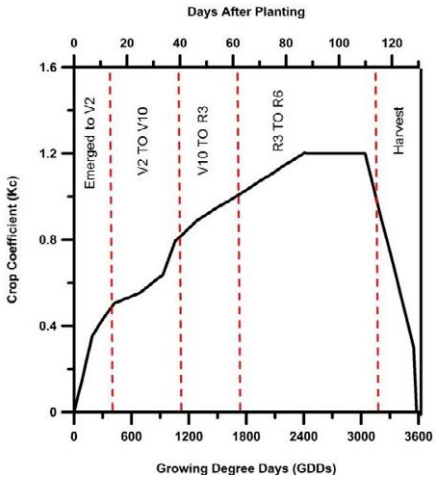


Figure 1. The crop coefficient curve developed for Georgia. The red dotted lines show the different growth stages for soybeans.

DAP	Julian Day	Avg Temp (F)	Base Temp (F)	Daily Heat Units	Accumulated Heat Units (GDD)	Kc	ETo (in/day)	ETc (Kc*ETo avg) (in/day)	Rooting Depth (in)	Available Soil Water (in)	Irrigation Applied (in)	Effective Irrigation (in)	Rain (in)	Effective Rain (in)	Irrigation + Rain (in)	Root Zone Water Deficit (in)	Root Zone Water Deficit (%)
0	136	72.2	60.00	0.00	0	0	0	0.00	6.00	0.92	-	-	0	-	-	0.00	0%
1	137	74.7	60.00	####	15	0.044	0.1495	0.01	6.30	1.01	-	-	0	-	-	0.00	0%
2	138	73.7	60.00	####	28	0.088	0.1567	0.01	6.80	1.05	-	-	0	-	-	0.01	1%
3	139	77.5	60.00	####	46	0.132	0.1462	0.02	7.30	1.13	-	-	0	-	-	0.02	2%
4	140	73.0	60.00	####	59	0.176	0.1879	0.03	7.80	1.21	-	-	0	-	-	0.04	3%
5	141	71.2	60.00	####	70	0.22	0.1632	0.04	8.30	1.28	-	-	0	-	-	0.07	6%
6	142	73.1	60.00	####	83	0.264	0.1619	0.04	8.80	1.36	-	-	0	-	-	0.11	8%
7	143	74.0	60.00	####	97	0.308	0.1681	0.05	9.30	1.44	-	-	0	-	-	0.15	11%

Figure 2. The excel spreadsheet shows the variables which are required to run the model. Users have to import data such as average daily temperature, Kc, ETo, soil water holding capacity, rooting depth, precipitation and irrigation (yellow columns).

On-farm irrigation experiment and model validation

The experiment was conducted during the 2018 growing season in a 12 ha soybean field (31 03' 58.57" N, 83 40' 34.00" W) located in Colquitt County of Georgia. The field was planted with maturity group VI soybeans on June 19th 2018. Before planting the field was divided into IMZs by using data from Sentinel-2 satellite images such as normalized difference vegetation index (NDVI) with 10 m spatial resolution and normalized difference water index (NDWI) with 20 m spatial resolution. Both variables were selected for the IMZs delineation because they are related to soil texture and crop growth. Figure 3 shows the delineated IMZs; the colors of the IMZs are related to different types of zones. The IMZs were delineated by using the EZZone software (https://ezzone.pythonanywhere.com). This software is a free online resource for delineating agricultural management zones from univariate georeferenced data. After the IMZ delineation, the field was divided into two parts (Figure 3). The western part was irrigated uniformly based on soybean model and the eastern part was irrigated uniformly using the grower's standard irrigation practice. The farmer was irrigating

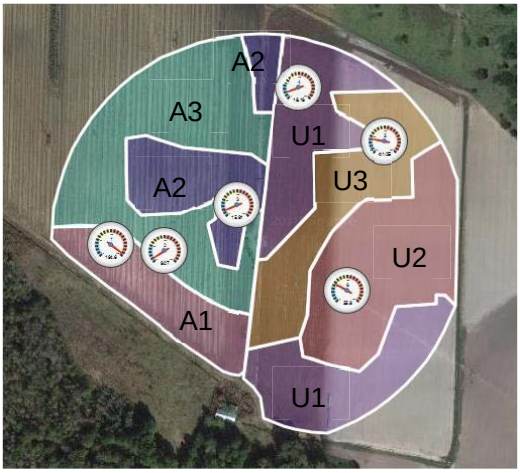


Figure 3. The experimental field. The polygons within the field present the delineated irrigation management zones (IMZs) based on the NDVI and NDWI. The gauges show the exact location of UGA SSA sensor nodes installed in the field. Gages show the weighted average soil water tension (SWT) of the three sensor readings. Average SWT = (0.5 × SWT at 200 mm) + (0.3 × SWT at 400 mm) + (0.2 × SWT at 610 mm).

the soybeans after identifying water stressed plants. The field was instrumented with the University of Georgia Smart Sensor Array (UGA SSA) which consists of smart sensor nodes (Figure 4) and a base station. The UGA SSA is a fully wireless sensing system which measures soil water tension in soybeans at the depth of 200, 400, and 610 mm using Watermark™ sensors (Vellidis *et al.*, 2013). At the current experiment, three nodes were installed at the eastern part and three at the western part of field. The sensor readings were used only to monitor the soil moisture in each zone and to avoid damaging the plants in case of a problem with the water deficit calculations. The use of the UGA SSA was beneficial for research and the farmer because it demonstrated the within-field soil moisture variability.

Results and discussion

Soil moisture data using the irrigation scheduling from the Soybean Model vs Grower's Method

Figure 5 presents the soil moisture as recorded by the sensors in node 2 which was located in the western part of the field where irrigation was applied based on data from the soybean model. Although, on three occasions, the soil water tension (SWT) of the shallow sensors exceeded the desired limit of 70 kPa, in general this graph represents an ideal soil moisture condition for most of the growing season.

Figure 6 presents the soil moisture readings in node 6 in the eastern part of the field irrigated using the grower's standard method. In Figure 6, soil moisture tension at up to 400 mm depth ranged from 0 kPa to 30 kPa indicating a wet soil profile probably resulting in leaching of nutrients and may have generated runoff from irrigation. However, both irrigation treatments received the same amount of irrigation. The high soil moisture variability at the different locations even if both treatments received the same amount of irrigation (5 irrigation events, 102 mm total irrigation, farmer was irrigating two to three days after irrigating the soybean model treatment) indicates that the field variability is high and the use of variable rate irrigation should be beneficial.

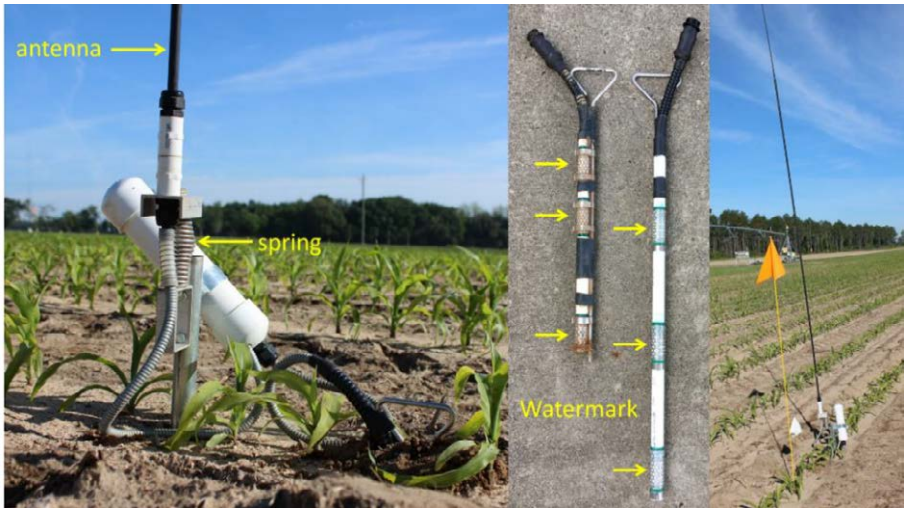


Figure 4. (Left) A UGA SSA node installed in soybeans. The electronics are housed in the white PVC container. The spring allows the antenna to bend when farm vehicles pass overhead. (Middle) The UGA SSA sensor probe integrates three Watermark sensors and can be customized to any length. (Right) Antennae can vary in length to accommodate the crop.

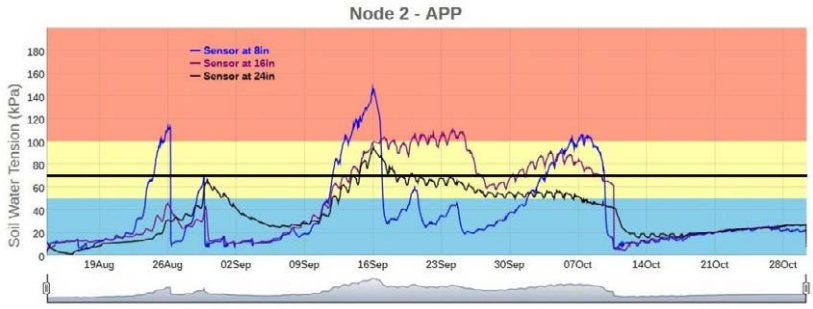


Figure 5. Soil water tension variability during the growing season. The blue line shows the soil water tension at 200 mm, the red at 400 mm and the black at 610 mm.



Figure 6. Soil water tension variability during the growing season. The blue line shows the soil water tension at 200 mm, the red at 400 mm and the black at 610 mm.

Model yield vs Farmer yield

The crop was harvested on 16th November 2018. The harvester was equipped with a yield monitor. Figure 7 presents the yield variability for both irrigation treatments. The yield range of the red points is from 326 to 2,555 kg/ha, the yellow points from 2,555 to 3,564 kg/ha, the green points from 3,564 to 4,304 kg/ha, the light blue points from 4,304 to 4,573 kg/ha and the dark blue points from 4,573 to 6,214 kg/ha. The lowest average yield (2,649 kg/ha) was found in farmer treatment and the highest (4,155 kg/ha) in the app treatment (Table1). Thus, the average yield in the model treatment area was 3,635 kg/ha which was 436 kg/ha higher than the yield in the farmer treatment area. As the amount of irrigation in both treatments was the same, it is obvious that the field variability is high and the variable rate application of inputs should be beneficial.

Conclusions

The use of the soybean model which calculates the daily soil water balance using weather data and soil type showed promise as an alternative to existing decision support tools. Furthermore, the integration of the soybean model with the UGA SSA system was very helpful for the evaluation of the soybean model. The use of satellite images and vegetation indices proved to be very helpful to delineate IMZs. The combination of IMZs with the real time soil moisture data enabled the authors to supervise the soil moisture condition of the field in real time. The results showed that the yield with the models treatment was 672.5 kg/ha higher than the yield with the farmer treatment even though both treatments received the same amount of irrigation. This indicates that the field variability is

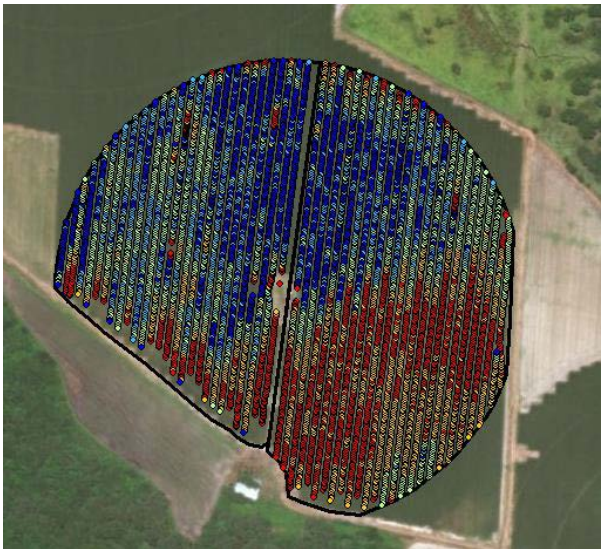


Figure 7. Yield variability of soybeans as recorded from the yield monitor.

Table 1. Descriptive statistics of the yield in each IMZ.

Zones[1]	Avg yield (kg/ha)	Min yield (kg/ha)	Max yield (kg/ha)	SD yield (%)
A1	2,765	368	6,058	9.42
A2	4,155	364	6,214	4.92
A3	3,985	327	6,020	5.96
U1	3,377	357	5,444	8.82
U2	2,649	372	5,036	5.97
U3	3,569	326	6,009	7.86

[1] See Figure 3.

high and the benefits of using variable rate irrigation in the next year should be beneficial. In 2019, the Soybean App will be released and a VRI system will be installed at the field. Thus yield data and more satellites images will be utilized to delineate new IMZs. Moreover the node location will be identified based on the soil moisture data from 2018 and the new IMZs. The Soybean App is not designed to send irrigation recommendations to VRI systems but users will be able to register several zones for one field to receive irrigation recommendation for each zone individually.

References

Allen, R.G., Pereira, L.S., Raes, D., Smith, M. 1998. Crop Evapotranspiration – Guidelines for Computing Crop Water Requirements. FAO Irrigation and drainage paper 56. Rome, Italy: Food and Agriculture Organization of the United Nations. ISBN 92-5-104219-5.

Food and Agriculture Organization of the United Nations. 2019. ETc- Single crop coefficient (Kc). Chapter 6. http://www.fao.org/docrep/x0490e/x0490e0b.htm [accessed February 2019].

McGuckin, J.T., Mapel, C., Lansford, R. and Sammis, T. 1987. Optimal control of irrigation scheduling using random time frame. American Journal of Agricultural Economics, 69, 123-133.

Migliaccio, K.W., K.T. Morgan, C. Fraisse, G. Vellidis, L. Zotarelli, C. Fraisse, B. A. *et al.,* 2015. Smartphone apps for irrigation scheduling. Transactions of the ASABE. 59 (1) 291-301 doi:10.13031/trans.59.11158.

Migliaccio, K.W., K.T. Morgan, C. Fraisse, G. Vellidis, L. Zotarelli, C. Fraisse, B. A. *et al.,* 2016. Performance evaluation of urban turf irrigation smartphone app. Computers and Electronics in Agriculture 118, 136-142 doi:10.1016/j.compag.2015.08.015.

Scanlon, C.C., Faunt, L., Longuevergne, R.C., Reedy, W.M. Alley, V.L., McGuire, P.B., *et al.,* 2012. Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley, Proceedings of the National Academy of Sciences, 109 (24) 9320-9325.

Siebert, S., Burke, J., Faures, J.M., Frenken, K., Hoogeveen, J., Döll, P., *et al.,* 2010. Groundwater use for irrigation – a global inventory. Hydrology and Earth System Sciences 14 1863-1880.

Singels, A. and Smith, M. T. 2006. Provision of irrigation scheduling advice to small-scale sugarcane farmers using a web-based crop model and cellular technology: a South African case study. Irrigation and Drainage 55, 363-372.

Rowland, D. L., Sorensen, R. B., Butts, C. L., & Faircloth, W. H. 2006. Determination of maturity and degree day indices and their success in predicting peanut maturity. Peanut Science 33 (2), 125-136.

Rowland, D. L. 2016. PeanutFarm irrigation and harvest tool is now mobile friendly. University of Florida Extension. http://nwdistrict.ifas.ufl.edu/phag/2016/05/20/peanutfarm-irrigation-and-harvest-tool-is-now-mobile-friendly [accessed February 2019].

Vellidis, G., Tucker, M., Perry, C., Reckford, D, Butts, C., Henry, H., *et al.,* 2013. A soil moisture sensor-based variable rate irrigation scheduling system. In: J.V. Stafford (Ed.), Precision Agriculture 2013 – Proceedings of the 9th European Conference on Precision Agriculture. Wageningen, the Netherlands: Wageningen Academic Publishers pp. 713-720. doi: 10.3920/978-90-8686-778-3.

Automated irrigation scheduling for drip-irrigated plum trees.

S. Millán[1*], C. Campillo[1], J. Casadesús[2], M.J. Moñino[1], A. Vivas[1] and M.H. Prieto[1]
[1]*Centro de Investigaciones Científicas y Tecnológicas de Extremadura (CICYTEX), Finca La Orden, Junta de Extremadura. 06187 Guadajira (Badajoz), Spain; sandra.millan@juntaex.es*
[2]*Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Centre UdL IRTA, Av. Alcalde Rovira Roure, 191, 25198 Lleida, Spain*

Abstract

During the past few years, due to a large growth in water demand, agricultural irrigation of fields has become a concern. For this reason, suitable water irrigation management plays an important role in agricultural irrigation systems, since efficient water management and regulated deficit irrigation strategies (RDI) can help to increase water savings and farmers' profitability while reducing environmental impact. These benefits are easier to apply in commercial farms due to advances in technological areas, which offer the opportunity of integrating a variety of sensors and accumulated knowledge into a decision support system (DSS). IRRIX, is a web platform hosted in the cloud and not requiring human intervention, that allows automated irrigation, resulting from an algorithm derived from a combination of water balance with a feedback adjustment mechanism using soil moisture sensors. Therefore, the main objective of this work was to analyse the IRRIX automatic irrigation system, for the automatic management of irrigation in the cultivation of plums and its suitability to form part of RDI strategies. The results show that automatic irrigation is a viable alternative and can facilitate the use of irrigation strategies aimed at a more efficient use of water.

Keywords: feedback, regulated deficit irrigation, moisture sensor, automated irrigation

Introduction

Water is a scarce resource (Romero *et al.*, 2012), so that rational use of it should be encouraged. Problems derived from lack of water will likely increase if long-term predictions on global climate change are correct. Meteorological records suggest significant increases in temperature and decreases in annual precipitation, which will entail a reduction of the available water resources for the XXI century (Turral *et al.*, 2011). In this context, proper management plays an important role in agricultural irrigation systems, since an efficient water management system can help to decrease water consumption, improving farm profitability and reducing environmental impact. Therefore, the use of technological tools in the agricultural sector should allow improved and more rational use of water by optimizing irrigation.

At the scientific level, the processes that determine the water demand of crops and the various methodologies that have been developed to irrigate efficiently are well known. The water balance method is the most widely used system to determine the irrigation needs of a crop, where the water inputs to the soil-plant system must be balanced with the expected outputs (Allen *et al.*, 1998). The most relevant component of the water balance used to determine irrigation is the evapotranspiration of the crop, ETc, or water requirement of the crop that tests both the evaporation in the plot and the transpiration of the plants, and is usually estimated as the product of the evapotranspiration of a reference crop, ETo, and a crop coefficient, Kc, in the form $ETc = ETo \times Kc$. In this relationship, ETo represents the demand imposed by the meteorological conditions while Kc integrates the physical and biophysical differences between the reference crop and the crop which is to be estimated for evapotranspiration (Allen *et al.*, 1998). The water balance method presents the disadvantage that any inaccurate parameterization will produce a systematic error that will be accumulated throughout the crop cycle, due to deviations that occur between the estimated values and the actual consumption.

DOI 10.3920/978-90-8686-888-9_117, © Wageningen Academic Publishers 2019

Irrigation control based on feedback from moisture sensors is an alternative to the water balance system, where the main advantage is the ability to adjust irrigation to the needs of a specific plot (Casadesús *et al.*, 2012). However, irrigation control based on only sensors presents some weaknesses such as the risk deriving from breakdowns, variability in the measurements (Nolz and Kammerer, 2017) and difficulty of interpreting the sensors, especially for drip irrigation, where soil water distribution is heterogeneous.
In general, the methods of water balance and irrigation control using sensors present advantages and complementary disadvantages, thus, combining both approaches seems the best future way to improve the efficiency of irrigation water in agricultural systems: i.e. determine the irrigation dose from a water balance model using sensors to adjust that model to the real situation of each plot. The use of information and communication technologies (ICTs) can help to realise this objective. In addition, these tools are designed for agricultural systems with high added value, such as woody crops, but these tools have even greater potential in agricultural systems that require very controlled irrigation management such as regulated deficit irrigation (RDI). RDI consists of applying a volume of irrigation lower than theoretical needs in the periods of the crops cycle in which the plants are less sensitive to water deficit (Fereres and Soriano, 2007).
The main objective of this work was to demonstrate the reliability of a web application for automated irrigation scheduling in an early maturing Japanese plum crop 'Red Beaut'.

Materials and methods

The experiment was conducted during one year (2017) in a 1.0 ha orchard of an early maturing Japanese plum (*Prunus salicina Lindl. 'Red Beaut'*), located in 'Finca La Orden' experimental farm, Badajoz (latitude 38°51'19.06'N, longitude 6°40'18.90'W, datum WGS84), Spain. Plum trees were planted in spring 2005 at a spacing of 6×4 m, in an east-west row orientation. Trees were irrigated daily using a drip system with a single lateral line per tree row located close to the base of the tree, with four pressure compensating in-line emitters with 4 l h^{-1} discharge rates, spaced at 0.5 m.

Automatic irrigation system.
The automatic irrigation system was composed of two components: a) sensors installed in the field and b) IRRIX, a web platform whose algorithm results from a combination of water balance with soil moisture sensor feedback adjustment mechanisms.

a. Sensor installed in the field: To monitor the soil moisture, 15 10HS capacitive moisture sensors (Decagon Devices Inc., Pullman, WA, USA) were installed at different depths (six sensors at 300 mm depth and located under the dropper; six sensors at 600 mm depth and located under the dropper; three sensors at a depth of 300 mm and located between droppers). An air temperature sensor, a solenoid valve (Rain BirdEurope SCN, Aix-en-Provence, France), a water meter (MTK, Zenner, Madrid, Spain), a solar panel, a voltage regulator (BlueSolar PWM-Pro, Vitron Energy Blue Power, Almere, the Netherlands) and a relay (SMD-CD16AC, Campbell Scientific Inc., Logan, UT, USA.), were also installed. All sensors were connected to a datalogger (CR1000, Campbell Scientific Inc., Logan, UT, USA.) which stored the data once every 5 minutes. The data was downloaded to an IRRIX server four times a day.
b. IRRIX is a web platform hosted in the cloud that carries out the following daily tasks:

1. Data collection of sensors installed in the field. IRRIX downloads sensor data at periodic intervals throughout the day and at the user's request. The evapotranspiration of the reference crop (ETo) and rainfall can be obtained from weather sensors connected directly to these tools or online through an external meteorological network (REDAREX, http://redarexplus.gobex.es). ETo was estimated daily from the air temperature sensor, using the Hargreaves equation (Hargreaves and Allen, 2003).

2. Analyse all data and calculate irrigation water volumes. IRRIX analyses all incoming data to detect anomalies and detect if any important event has occurred in the system (irrigation, rain). Then, IRRIX analyses the set of data to determine the irrigation dose and adjust the irrigation dose with the information provided by the humidity sensors. To achieve this, this tool integrates a water balance model.
3. Irrigation scheduling. IRRIX sends the updated irrigation doses to the datalogger. Then, this device will order the activation of the rest of the equipment (solenoid valve or pumps, etc.) to apply the required irrigation doses.
4. Interact with users. IRRIX is an autonomous system whose main objective is to free the user from work. The main function of the user is to check that the system has worked correctly and the irrigation campaign has passed as expected. Logically, if there is any anomaly in the system the user must resolve this anomaly.

In addition, before starting the irrigation campaign, the user must input to IRRIX a seasonal plan that foresees roughly how the water will be distributed throughout the irrigation campaign. For this, a set of curves must be defined with the automated control system being positioned between those curves, such that the system must be between a maximum and a minimum of accumulated irrigation.

Irrigation treatments

Three irrigation treatments were applied during the 2017 season: Control, RDI and Automatic. The Control treatment was irrigated to replace the ETc, minus effective rainfall. RDI received 40% of the water given to the Control after harvest until onset of leaf fall. The automatic treatment was carried out by means of the IRRIX application, and the seasonal plan that was introduced in to the IRRIX, being the same as the RDI treatment.

Measurements

Irrigation volumes were recorded daily by digital water meters (CZ2000-3M, Contazara, Zaragoza, Spain) installed on each replicate. The water use efficiency (WUE, kg/m^3) was calculated as the radio between yield and the irrigation applied in each treatment.

Results and discussion

Figure 1 shows a screenshot of the plot that is managed through IRRIX. This screen shows what has happened in the irrigation plot in the last 7 days, so that the technician or farmer, by visualizing the traffic lights, interprets whether the system is working correctly or if there is a fault in the system.

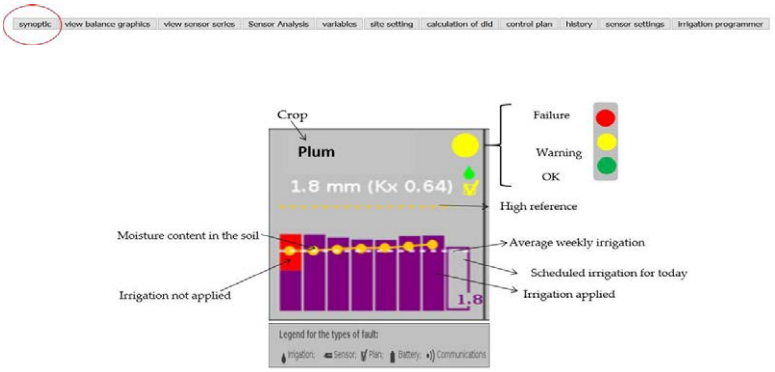


Figure 1. Screenshot of the 'synoptic' tab of IRRIX in the year 2017.

The system can detect if an anomaly has occurred in any of the sensors, in the irrigation, in the seasonal plan, system batteries or if any communication error occurs.
Figure 2 shows a screenshot of the 'control plan' tab of IRRIX, where it is possible to observe the accumulated irrigation during the entire irrigation campaign of 2017 with its corresponding upper and lower limits. The automatic treatment was placed between the upper and lower limits, so this treatment followed the plan established.
Figure 3 shows an example of the sensor display during the irrigation campaign. Figure 3a shows the daily superimposed ETo with the irrigation application measured by the water meter in order to compare them. However, Figure 3b presents the daily pattern of each of the sensors installed in the plot during the complete irrigation campaign in 2017. It can be observed that each sensor shows its own daily pattern, which makes it difficult for users to interpret them. The variability of the sensors increases in the case of localized irrigation, since the difficulty in the use of soil moisture sensors comes from the complex spatial distribution of water, characterized by the formation of humid bulbs around the emitters, with the rest of the soil being little or unaffected by irrigation (Choi *et al.*, 2016). In this sense, it is not necessary that irrigators interpret this data, since IRRIX interprets sensors data.
Figure 4 shows the comparison between the ETc calculated using the recommended Kc and the ETo obtained with the Penman-Monteith formula from the data obtained from a meteorological station located on the plot belonging to the REDAREX network in the year 2017. Also, this figure presents adjusted ETc by IRRIX with the information provided by the sensors (Hargreaves equation) and the adopted irrigation strategy. However, as can be seen, the adjusted ETc was inferior to the theoretical, even in the period without stress, when taking into account the reserve of water in the soil and the advisable stress.
Figure 5 shows the water applied to each irrigation treatment in the year 2017. The reduction in irrigation water in the RDI treatment with respect to Control was on average 38% while, for the automatic treatment, it was 41%. It is possible to observe that the automatic treatment reduced the water applied in pre-harvest when 100% of the water needs of the crop were applied. The algorithm of Casadesús *et al.* (2012) suggests that the water balance model allowed a rapid response to address the changes that could occur, while the feedback mechanism could adapt the amount of water to the requirements of each plot.
Figure 6 shows the WUE calculated for each one of different treatments. The WUE increased notably in the treatments with less water applied: 1.23 and 1.51 kg of plums were produced per cubic meter of water provided in RDI and Automatic respectively, compared to only 0.50 kg in the Control.

Conclusions

IRRIX can handle irrigation almost unassisted throughout the whole campaign, in the sense that the irrigator need only be present if they are notified of any anomaly in the irrigation system. In addition, the system can be configured to carry out various irrigation strategies, in this case, for the

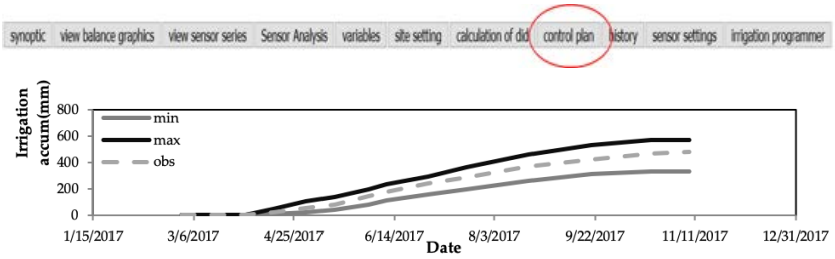


Figure 2. Screenshot of the 'control plan' tab of IRRIX in the year 2017.

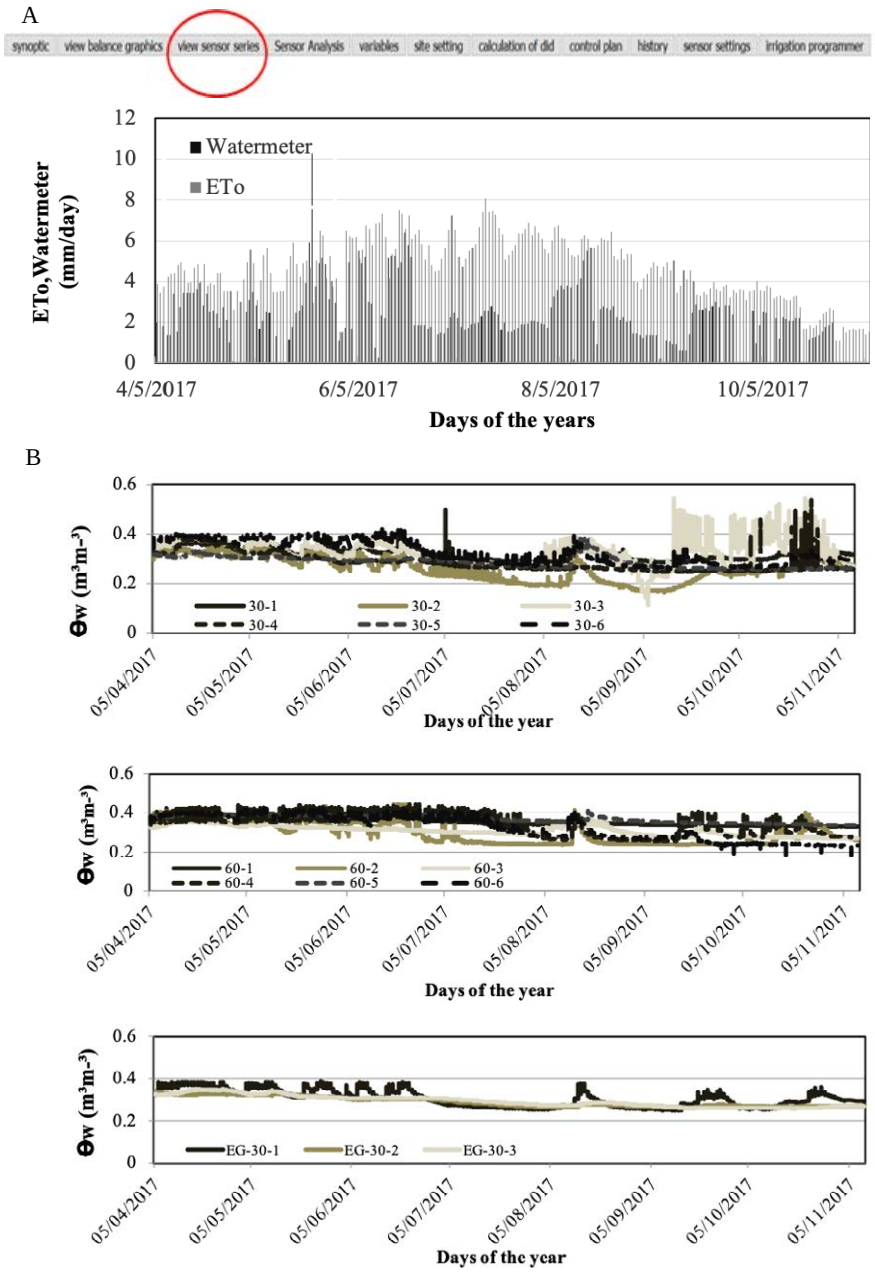


Figure 3. Information showed the 'view sensor series' tab of IRRIX in the year 2017. (A) Daily accumulaion of the ETo and the irrigation applied; (B) values of the 15 capacitive sensors installed in the irrigation sector to different depths.

implementation of deficit irrigation strategies. In addition, with IRRIX, a water saving is produced increasing the production.

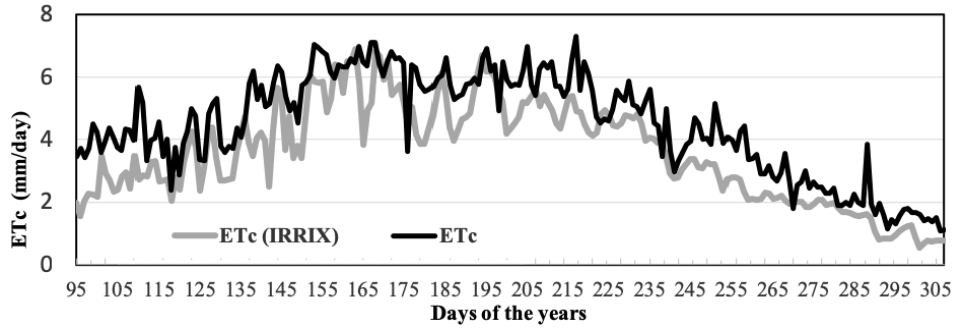


Figure 4. Seasonal evolution of the 'theoretical' ETc and ETc (IRRIX) in the year 2017.

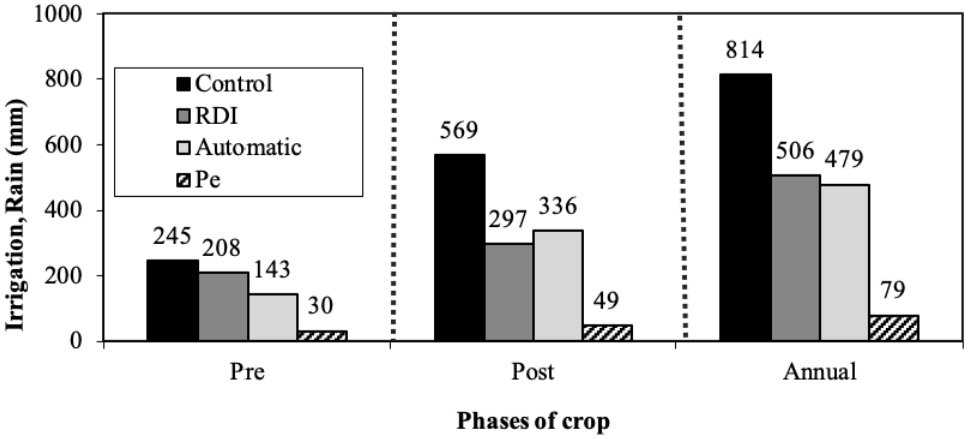


Figure 5. Total irrigation water applied under three irrigation treatments in 2017. 'Pre' is the period from fruit set to harvest of the current year; 'Post' is the period from harvest to onset leaf fall of the current year; 'Annual' is the period from onset of leaf falls of the previous year to onset of leaf fall of the current year.

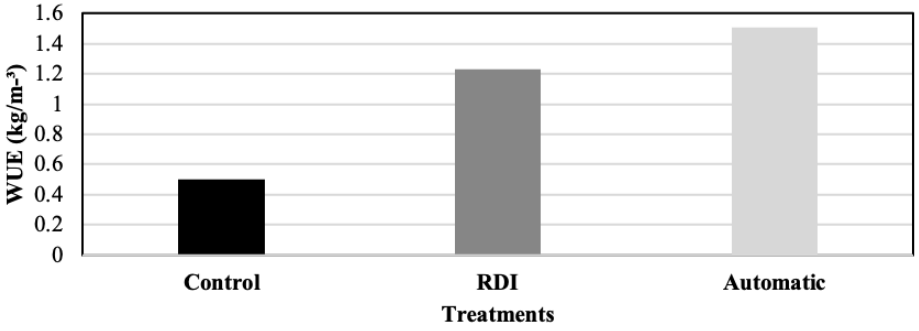


Figure 6. WUE in different treatments of irrigation.

Acknowledgements

This project was funded INIA (RTA2013-00045-C04 project), CCESAGROS projects and Junta de Extremadura.

References

Allen RG, Pereira LS, Raes D, Martín M, 1998. Crop evapotranspiration. Guidelines for computing crop water requirements. Irrigation and Drainage Paper 56, Rome, Italy: FAO

Casadesús J, Mata M, Marsal J, Girona J, 2012. A general algorithm for automated scheduling of drip irrigation in tree crops. Computers and Electronics in Agriculture 83, 11-20.

Choi, K. Y., Choi, E. Y., Kim, I. S., & Lee, Y. B. (2016). Improving water and fertilizer use efficiency during the production of strawberry in coir substrate hydroponics using a FDR sensor-automated irrigation system. Horticulture, Environment, and Biotechnology, 57(5), 431-439.

Fereres E and Soriano MA, 2007. Deficit irrigation for reducing agricultural water use. Journal of Experimental Botany 58, 147-159.

Hargreaves, G. H., & Allen, R. G. (2003). History and evaluation of Hargreaves evapotranspiration equation. Journal of Irrigation and Drainage Engineering, 129(1), 53-63.

Nolz, R., & Kammerer, G. (2017). Evaluating a sensor setup with respect to near-surface soil water monitoring and determination of in-situ water retention functions. Journal of hydrology, 549, 301-312.

Romero, R., Muriel, J. L., García, I., & de La Peña, D. M. (2012). Research on automatic irrigation control: State of the art and recent results. Agricultural water management, 114, 59-66.

Turral, H., Burke, J., Faures, J.M., 2011. Climate Change, Water and Food Security. FAO Water Report 36. Rome, Italy: FAO.

Drone dataflow – a MATLAB toolbox for extracting plots from images captured by a UAV

A.K. Mortensen[1*], M.S. Laursen[2], R.N. Jørgensen[2] and R. Gislum[1]
[1]*Department of Agroecology, Aarhus University, Forsøgsvej 1, 4200 Slagelse, Denmark; anmo@agro.au.dk*
[2]*Department of Engineering, Aarhus University, Finlandsgade 22, 8200 Aarhus N, Denmark*

Abstract

The use of unmanned aerial vehicles (UAVs) with mounted cameras has seen a rapid increase in popularity within precision agriculture research for surveying fields. Typically, the collected images are stitched into georeferenced maps using photogrammetry software. This paper proposes a MATLAB toolbox for easing the workflow. The toolbox consists of two modules: (1) detection of ground control points in images collected by UAVs; and (2) extraction of regions-of-interest from georeferenced maps generated from the images. Use of the toolbox is exemplified on selected plots from a plot experiment throughout a growing season.

Keywords: drone, software, maps, unmanned aerial vehicle

Introduction

Surveying fields using unmanned aerial vehicles (UAVs) has become increasingly popular in recent years, especially in precision agricultural research. The UAV allows the researcher to scan the entire field quickly and easily with a camera mounted on the UAV. Though, the entire field is scanned, the researcher is often only interested in specific regions-of-interest (ROIs) within the field and being able to extract these regions from the collected images. In a plot experiment, the ROIs will typically coincide with the individual plots. Here it is important to relate the generated image data reliably to the effect of the treatment of each individual plot. In a commercially run field, the ROIs will typically be areas where physical samples have been collected with the purpose of either calibrating a model to the rest of the field (Mortensen *et al.*, 2015) or creating a generalized model across fields. The typical workflow, when using images collected by a UAV to describe ROIs, is as follows. First, ground control points (GCPs) are distributed around the field and their position is recorded. This is followed by scanning the field using one or more cameras mounted on the UAV. Next, the GCPs are identified in the images. The position of the GCPs in the real world and the images are used by photogrammetry software to stitch the images into georeferenced maps with a high level of accuracy. Finally, the ROIs are extracted from the maps, so they can be compared to the physical measurements. Solutions exist for each of the individual steps in the described workflow. E.g. coded targets for GCPs (Ahn *et al.*, 2001; Moriyama, 2008; Calvet *et al.*, 2016), photogrammetry software such as Pix4D (Pix4D, Lausanne, Switzerland), Photoscan (Agisoft, St. Petersburg, Russia) and OpenDronemap (www.opendronemap.org), and geographic information systems (GIS) applications such as ArcGIS (Esri, Redlands, CA, USA) or QGIS (www.qgis.org) may be used for ROI extraction. These solutions, however, tend to focus on only their specific step or sub problem of the entire workflow.
This paper proposes a framework for processing images collected using a UAV and extracting ROIs from maps generated from these images. The framework is MATLAB toolbox, which consists of two main modules for detecting GCPs and extracting ROIs, respectively. The framework and modules are described in detail in the following section. This is followed by an example use from a plot experiment. In the final sections, the toolbox is discussed and the work concluded.

DOI 10.3920/978-90-8686-888-9_118, © Wageningen Academic Publishers 2019

Materials and methods

The Drone Dataflow MATLAB toolbox consisted of two main modules (Figure 1): (1) the GCP detection module and (2) the ROI extraction module. The GCP detection module detected GCPs in images captured by a UAV. The ROI extraction module extracted ROIs from georeferenced maps generated by photogrammetry software.
The toolbox assumed that an approximate 3D position (latitude, longitude and altitude) of all images, GCPs and ROIs were known with respect to a common world co-ordinate system. It further assumed that generated maps were georeferenced in the same world co-ordinate system.
The toolbox requires MATLAB R2018a or newer, as well as the following MATLAB toolboxes: Optimization Toolbox, Mapping Toolbox, Image Processing Toolbox, Statistics and Machine Learning Toolbox and Aerospace Toolbox. Code is available at https://github.com/anderskm/drone-dataflow

GCP detection module
The dataflow of the GCP detection module is illustrated in Figure 2. The module took the images collected by the UAV and the world co-ordinates of the GCPs as inputs. The images had to be geotagged with respect to the world co-ordinate system. The output of the module was the image co-ordinates of all the detected GCPs.
First, the EXIF tags of the images were used to determine the camera model from a set of predefined camera models. The predefined camera models supported both perspective and fisheye camera models and were customisable by the user.

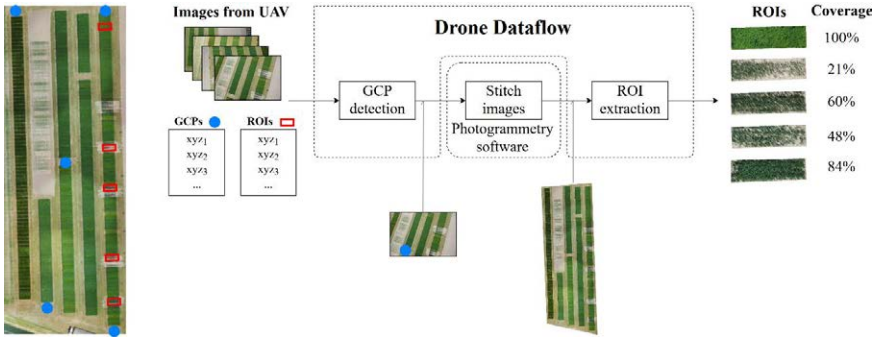


Figure 1. Overview of the Drone Dataflow in relation to the overall workflow of extracting ROIs from images collected using a UAV. The left field is 255×83 m. The extracted ROIs are 2.3×7.9 m.

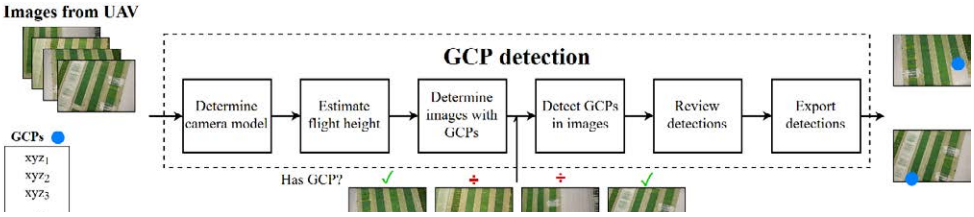


Figure 2. Block diagram of the GCP detection module. Geotagged images and a list with GPC co-ordinates are input into the module. The module outputs the image co-ordinates of the detected GCPs.

Next, the flight height of each image was estimated. The flight height of a given image was estimated as the difference in altitude between the ground and the image (Figure 3). The altitude of the ground was, however, unknown. To estimate the altitude of the ground, a plane was fitted to the world co-ordinates of the GCPs using RANSAC (Fischler and Bolles, 1981). The flight height of each image was then the height difference in altitude between the image and the ground plane at the image xy-co-ordinate.

The camera model and flight height of a given image were then used to evaluate if the image might contain a detectable GCP. The evaluation was performed by projecting each of the image corners onto the estimated ground plane (Figure 3). If the orientation (yaw, pitch and roll) of the camera was available from the EXIF tags, these were also used in the projection. If the orientation was not available, yaw was estimated from the image positions, and pitch and roll were assumed to be zero. For a given image, the yaw was estimated as the orientation (with respect to North) of the vector in the xy-plane pointing from the position of the current image to the next image. The image projection onto the ground plane created a four-sided polygon on the ground plane. For each GCP, it was checked as to whether it was located inside the polygon. The evaluation only considered the xy-co-ordinates of the GCP and polygon to account for the GCPs, which were not located exactly on the ground plane. If it was determined that a given image might contain a GCP, the image was flagged for further processing (Figure 2 and 3).

Next, each flagged image was processed to detect the GCPs present in it. Along with the image, image position and orientation, flight height, camera model and GNSS positions were passed to the GCP detection algorithm. Each GCP detection included the image co-ordinates, where it was detected in the image, and an ID of the detected GCP. The GCP detection algorithm is interchangeable as long as it accepts the above-mentioned inputs and outputs. The examples presented in this paper uses the N-fold edge marker as GCPs (Midtiby, 2018)[4].

The detected GCPs were reviewed by the user through a review GUI. The review process included verifying that the detections were GCPs, had the correct ID and were assigned to the correct corresponding real world GCP. Finally, the GCPs were exported to a format readable by photogrammetry software.

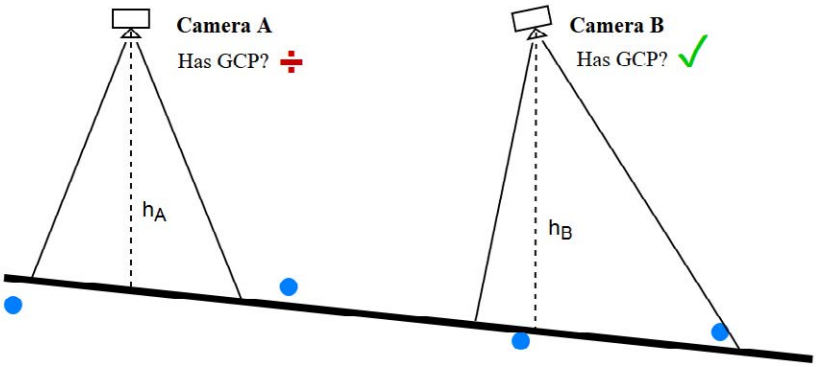


Figure 3. Estimation of individual camera height. The circles indicate ground control points. The thick line is the fitted ground plane. The dashed lines are the height h_A and h_B of the two cameras A and B, respectively. The thin solid lines are the field of view of each camera.

[4] https://github.com/henrikmidtiby/MarkerLocator.

ROI extraction module

The ROI extraction module extracted ROIs from a georeferenced map of interest. The module also extracted user specified features from each ROI. Figure 4 shows a block diagram of the ROI extraction module.

A georeferenced map of interest and a list of ROIs were input to the module. Optionally, a reference map could also be input to the module. Each ROI was handled individually. For each ROI, a neighbourhood around the ROI was extracted. The neighbourhood was found by extending the bounding box of the ROI by 50 pixels on all sides. If a reference map is used, a similar neighbourhood was extracted from the reference map. The pixels of the map of interest and the reference map were not necessarily perfectly aligned due to variations in the image data used to generate the two maps. Therefore, the neighbourhood extracted from the reference map was transformed using bilinear interpolation into the same co-ordinate system as the map of interest. This allowed the two maps to be subtracted from each other to create a difference map:

$$N_\Delta = N_M - T_{R\to M}(N_R) \tag{1}$$

where N_M is the neighbourhood extracted from the map of interest, N_R is the neighbourhood extracted from the reference map, and $T_{R\to M}$ is the transformation from the reference map to the map of interest. N_Δ is the resulting difference neighbourhood, when the reference map neighbourhood has been subtracted from the map of interest neighbourhood. If a reference map was not provided to the ROI extraction module, the neighbourhood extracted from the map of interest was simply forwarded to the next step without any alteration.

From the extracted neighbourhood, the smallest axis aligned bounding box around the ROI was extracted. A weighted mask with the same size as the bounding box was also extracted. Each pixel

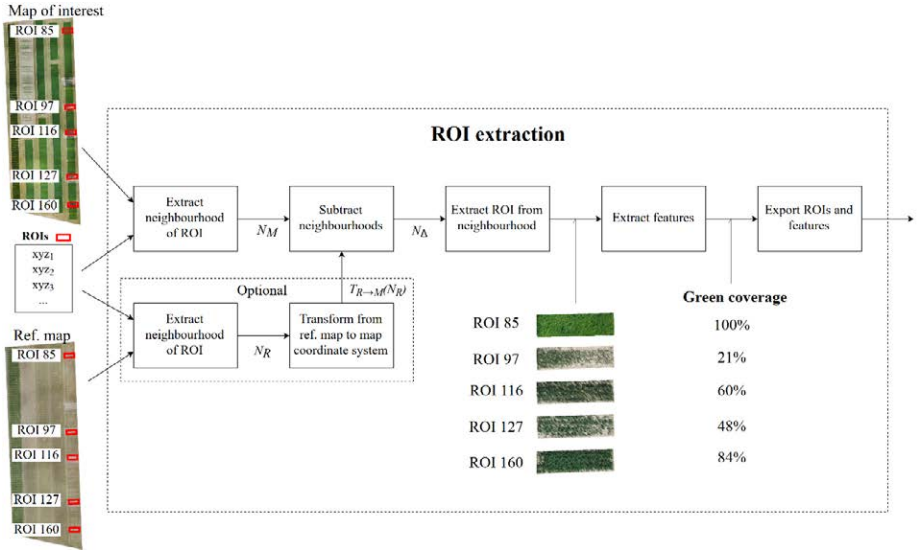


Figure 4. Block diagram of the ROI extraction module. A georeferenced map of interest, a list of ROIs and, optionally, a georeferenced reference map were input into the module. The module extracted the ROIs and user specified features from the map. If a reference map was specified, it was subtracted from the map before the ROIs were extracted. The variables refer to the variables in Equation 1.

in the mask was weighted with respect to the relative area of the pixel located inside the ROI. E.g. if 50% of the pixel area was located inside the ROI, the weight was 0.5.
After extracting a ROI, features were extracted from the ROI. The user-defined features could be low level such as weighted mean intensity or more high level, such as percentage of green coverage. When extracting the features, the weighted mask was used to determine the relative importance of each pixel in the ROI.
In the last step, the extracted ROIs and features were exported. The ROIs were exported as georeferenced maps. Each ROI map was cropped to the smallest bounding box and all pixels outside the ROI were set to be fully transparent. The features were exported to an excel-spreadsheet to ease the later data analysis.

Example use
The plot experiment (latitude: 55.32851° N, longitude: 11.38731° E) shown in Figure 1 was scanned approximately weekly by an eBee (senseFly SA, Cheseaux-sur-Lausanne, Switzerland) with the S.O.D.A. (senseFly SA, Cheseaux-sur-Lausanne, Switzerland) RGB camera mounted. The weekly scans were carried out from April to July in 2018. Prior to scanning the field, five GCPs were placed in and around the field (Figure 1). The GNSS co-ordinates of the GCPs as well as the four corners of each plot were recorded. Five plots with perennial ryegrass (*Lolium perenne* L.) were selected as ROIs based on their visual variation throughout the growing season (Figure 5).
For each weekly flight, the images were processed using Drone Dataflow and Pix4D as photogrammetry software. First, the GCP detection module in Drone Dataflow was used to detect the GCPs in the images. Next, the images were stitched into georeferenced orthomosaics as well as a digital surface model (DSM) using Pix4D. The ROIs were extracted from both maps using the ROI extraction module in Drone Dataflow. From the orthomosaic, the green coverage was extracted as a feature from each ROI. The green coverage was determined as the ratio between number of green plant pixels with respect to the total number of pixels within a given ROI. A pixel was regarded as a green plant pixel, if the excess green (ExG) exceeded the excess red (ExR) (Meyer and Neto, 2008):

$$ExG = \frac{2g-r-b}{r+b+g} \quad (2)$$

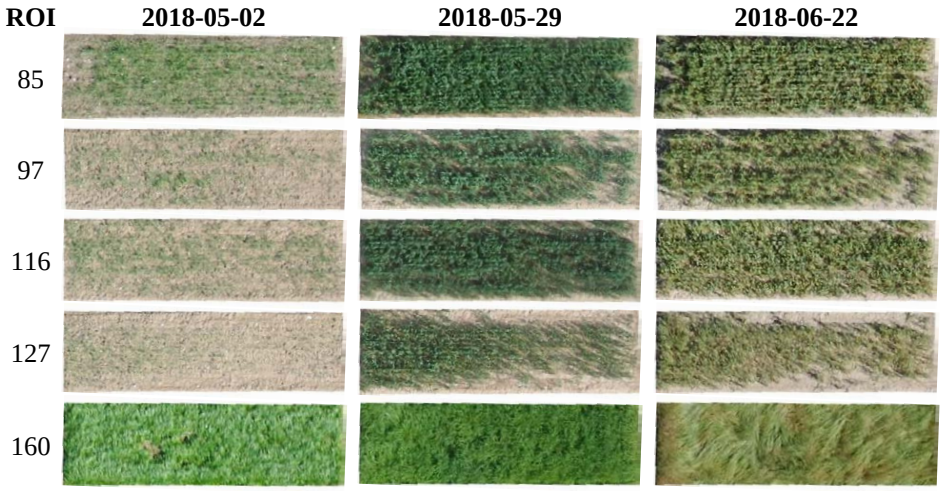


Figure 5. Extracted ROI at three different dates. The ROIs are each approximately 2.3×7.9 m. The ROI ids corresponds to the ROIs in Figure 6.

$$ExR = \frac{1.4r - g - b}{r + b + g} \quad (3)$$

where r, g and b is the intensity of the red, green and blue channel of a given pixel. *ExG* and *ExR* are the excess green and the excess red of a given pixel, respectively.

From the DSM, the maximum height of each ROI was extracted as a feature. The DSM from the first flight was used as a reference map to get a relative height with respect to the first flight instead of a height with respect to the world co-ordinate system.

The estimated green coverage and relative height over time for each of the ROIs are shown in Figure 6 and 7, respectively. For all ROIs, an increase in green coverage over time was experienced until the beginning of June, where the crop began to senesce. A visually good correspondence was also observed between the extracted ROIs (Figure 5) and the estimated green coverage (Figure 6). The green coverage of ROI 160 quickly saturated at 1 and the growth of the crop could not be tracked by the coverage after this point. The height of all the ROIs increased with time, except ROI 160, which quickly dropped in the middle of June indicating something had gone wrong. Studying the extracted ROIs (Figure 5), it was evident, that the drop in height was likely due to lodging.

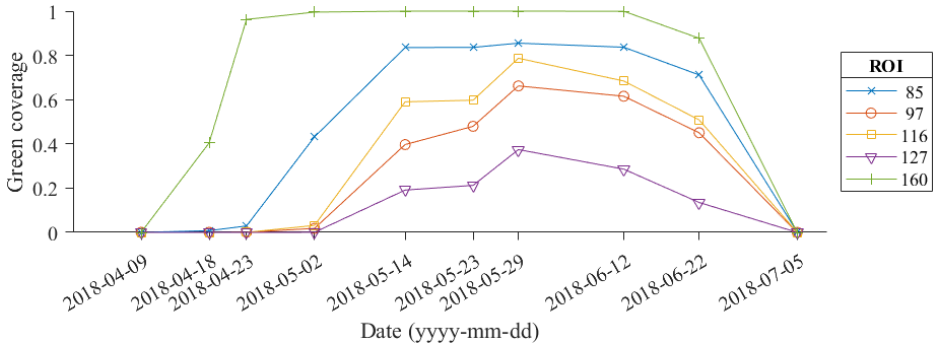


Figure 6. Green coverage estimates of five ROIs during the 2018 growing season. The ROIs correspond to selected plots with perennial ryegrass (Figure 5). Green coverage was determined using Equation 2 and 3.

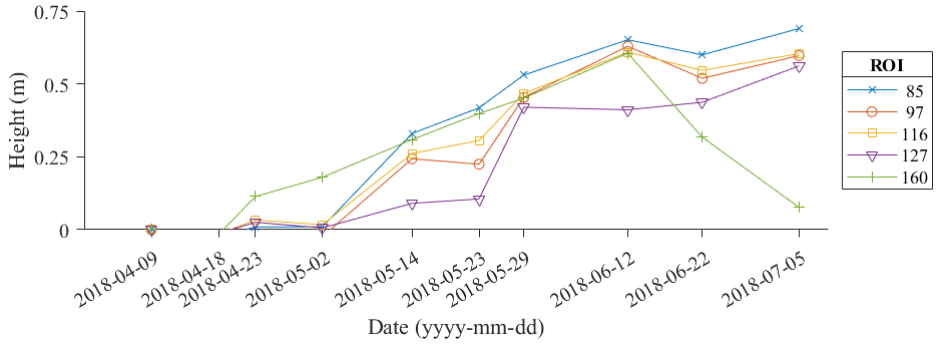


Figure 7. Crop height estimates of five ROIs during 2018 growing season. The ROIs correspond to selected plots with perennial ryegrass (Figure 5). The heights were calculated using April 9 as reference zero height.

Discussion

The photogrammetry software Photoscan comes with its own built-in detection of coded targets (or GCPs). The detection is, however, limited to one type of GCPs, which potentially locks the user to both the specific GCPs and the Photoscan.
Newer UAVs such as the eBee RTK (senseFly SA, Cheseaux-sur-Lausanne, Switzerland) come with an RTK GNSS, which provides images geotagged with a very high accuracy and precision. The maps generated from these images without using GCPs have an accuracy comparable to maps generated using GCPs. These new high accuracy UAVs could make the GCP detection module irrelevant, but GCPs can be used as a redundancy.
The toolbox is implemented in MATLAB, which is a commercial product. In line with the free and open source nature of the toolbox, it would be preferable with an underlying non-commercial solution. E.g. the programming language python (www.python.org).

Future work

Future work will focus on better integration with the photogrammetry software. E.g. in the current version of Drone Dataflow, the detected GCPs must be imported into the photogrammetry software manually. In the future, the detected GCPs could be written directly to the project files of the photogrammetry software. This will help streamline the workflow even further.

Conclusions

The proposed MATLAB toolbox streamlines the dataflow involved in extracting ROIs from images collected using UAVs. This both saves time for the end-user and ensures consistency in the data processing. The modular building blocks of the toolbox make it easy to extend it to other types of GCPs, photogrammetry software and ROI-features.

References

Ahn, S. J., Rauh, W., & Kim, S. I. (2001). Circular coded target for automation of optical 3D-measurement and camera calibration. International Journal of Pattern Recognition and Artificial Intelligence, 15 (06), 905-919

Calvet, L., Gurdjos, P., Griwodz, C., & Gasparini, S. (2016). Detection and accurate localization of circular fiducials under highly challenging conditions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 562-570).

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 24 (6), 381-395.

Meyer, G. E., & Neto, J. C. (2008). Verification of color vegetation indices for automated crop imaging applications. Computers and Electronics in Agriculture, 63 (2), 282-293.

Midtiby, H. Skov. (2018). Robust detection of n-fold edges using convolution with a complex kernel. In preparation. https://github.com/henrikmidtiby/MarkerLocator (last accessed March 2019)

Moriyama, T., Kochi, N., Yamada, M., & Fukaya, N. (2008). Automatic target-identification with the color-coded-targets. In The International Archives of Photogrammetry and Remote Sensing, Beijing, XXI Congress, WG (Vol. 1, pp. 39-44).

Mortensen, A. K., Gislum, R., Larsen, R., & Jørgensen, R. N. (2015). Estimation of above-ground dry matter and nitrogen uptake in catch crops using images acquired from an octocopter. In J.V. Stafford (Ed.) Precision agriculture'15 Proceedings of the 10th European Conference on Precision Agriculture. (pp. 227-234). Wageningen, the Netherlands: Wageningen Academic Publishers.

Section 19
Decision support for precision agriculture

A multi-disciplinary approach for the precision management of lodging risk

P. Berry[1*], A. Blackburn[2], M. Sterling[3], Y. Miao[4], D. Hatley[5], D. Gullick[2], G. Joseph[3], D. Whyatt[2], D. Soper[3], J. Murray[2] and C. Baker[3]
[1]*ADAS High Mowthorpe, Duggleby, Malton, North Yorkshire, United Kingdom; pete.berry@adas.co.uk*
[2]*Lancaster Environment Centre, Lancaster University, Lancaster, United Kingdom*
[3]*Department of Civil Engineering, University of Birmingham, United Kingdom*
[4]*Department of Soil Water & Climate, University of Minnesota, USA*
[5]*ADAS Gleadthorpe, Meden Vale, Mansfield, Nottinghamshire, United Kingdom*

Abstract

This paper describes how wind engineering, geospatial science and crop agronomy disciplines can be combined to provide a decision support system for farmers to manage lodging risk, using the Crop Failure Assessment due to Lodging Losses (CROPFALL) framework. CROPFALL calculates lodging risk using information about topography, land cover, soil type and meteorological data, combined with a mechanistic model of lodging, and crop parameters. The risk of lodging is calculated at regional, farm, field and sub-field scales. The impact of crop management changes on lodging risk is simulated, demonstrating that modest crop management changes can achieve large reductions in lodging risk.

Keywords: lodging, risk prediction, wind engineering, remote sensing, agronomy

Introduction

Lodging is the permanent displacement of plant stems from their vertical position and usually only occurs after flowering (Figure 1). This can reduce crop yield by up to 80% and causes several knock-on effects including reduced grain quality, greater drying costs and slower harvest (Berry and Spink, 2012). It is a problem that limits crop productivity globally. Lodging is a complex phenomenon influenced by many factors: wind, rainfall, topography, soil type and crop characteristics (Berry *et al.*, 2003). The effective control of lodging using crop and soil management requires lodging prediction before the crop is sown to enable strategic control options to be employed (e.g. variety choice, seed rate) and at an early stage of crop growth to enable tactical (within-season) control measures (e.g. nitrogen fertiliser, plant growth regulator chemicals). The multi-faceted challenge of



Figure 1. Example of lodging in wheat.

DOI 10.3920/978-90-8686-888-9_119, © Wageningen Academic Publishers 2019

lodging control requires a multi-disciplinary approach. Therefore, this paper describes how wind engineering, geospatial science and agronomy disciplines can be combined to provide a decision support system for farmers to manage lodging risk with the Crop Failure Assessment due to Lodging Losses (CROPFALL) framework.

Outline of CROPFALL lodging risk management system

An outline of the CROPFALL lodging risk decision support system is described in Figure 2. At the core of CROPFALL is a mechanistic plant lodging model which explains the interaction of crops with wind and rain (Baker *et al.*, 2014). The lodging model calculates the wind induced base bending moment of a shoot from measured plant characteristics (crop height at centre of gravity, natural frequency and ear area, for which the measurement methodology is described in Berry *et al.* (2004) and wind speed. The base bending moment is then compared to the failure moments (strength at the point of failure) of the stem base (typically 100 to 300 Newton millimetres (Nmm) in cereals) and of the anchorage system (typically 200 to 600 Nmm in cereals). Stem lodging is assumed to occur when the base bending moment of a single shoot exceeds the failure moment of the stem base. Root lodging is assumed when the base bending moment of all the shoots belonging to a single plant exceeds the failure moment of the anchorage system. The failure moment of the anchorage system is determined by the spread and depth of the root plate and soil strength, which in turn is determined by the soil type, moisture content and daily rainfall. The lodging model calculates the minimum wind speed for stem lodging (stem failure wind speed) or root lodging (root failure wind speed). Figure 3 illustrates how different combinations of rainfall and wind speed result in different likelihoods of stem and root lodging.

Using data from a network of meteorological stations, wind speed, wind direction and rainfall, spatially continuous representation of these values is spatially interpolated using 3 dimensional kriging. The wind speed and direction data are further refined by using a process-based wind flow model which accounts for the effects of topography and surface roughness. This increases the spatial resolution of the wind data such that within field variations can be represented. In this way CROPFALL produces a map describing how wind speed, wind direction and rainfall varies spatially (Figure 4). Such outputs can be derived from any particular storm event that is captured in the meteorological data archive, for a statistical representation of long term conditions or extremes, or

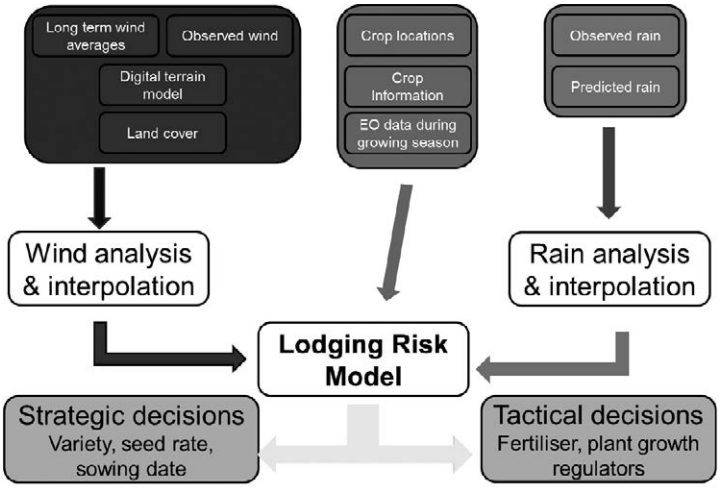


Figure 2. Outline of the CROPFALL decision support system for predicting lodging risk.

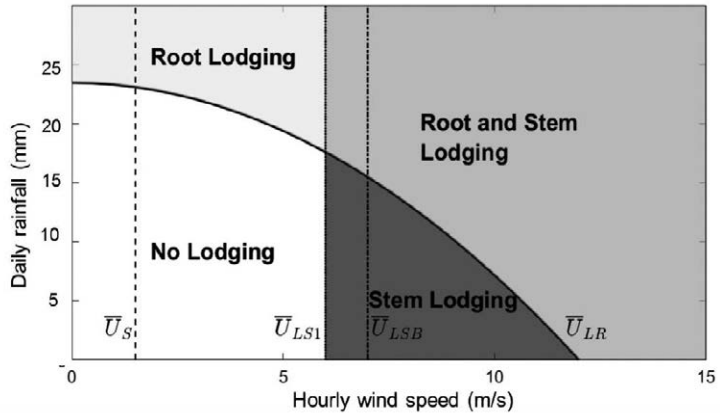


Figure 3. Lodging areas in the mean wind velocity/daily rainfall plane. U_S root lodging wind velocity when the soil is saturated; U_{LS1} stem lodging at 1st stem node wind velocity; U_{LSB} stem lodging at base wind velocity; U_{LR} root lodging wind velocity (adapted from Baker *et al.*, 2014).

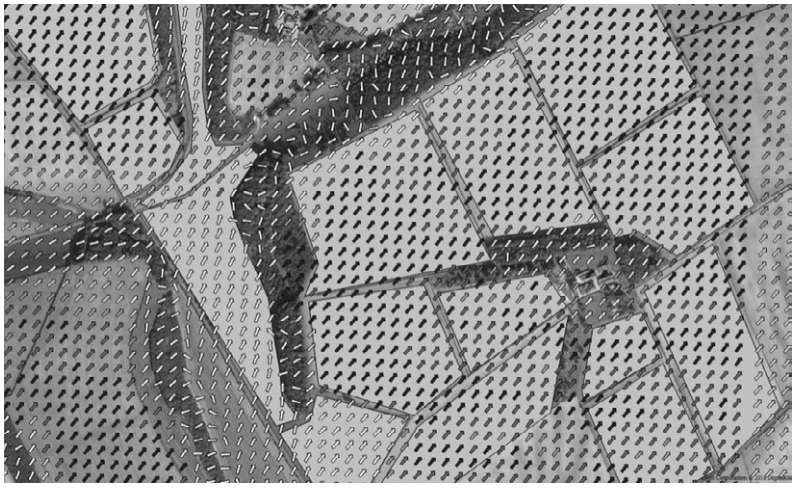


Figure 4. Example of wind speed and direction output from CROPFALL. White arrows represent low wind speeds and dark arrows represent high wind speeds. Shaded polygons represent field areas for a particular farm.

for hypothetic scenarios that are generated by manipulating the source data. On the other hand, earth observation (EO) technologies are used to update the lodging risk during the growing season based on canopy size measured in terms of green area index (GAI). A larger canopy size at the start of stem extension is linked with the development of weaker stems (Figure 5).

CROPFALL is used to calculate the risk of stem and root lodging before the crop is sown to enable the farmer to make strategic crop management decisions to optimise choice of crop species, crop variety, seed rate and sowing date. An updated lodging risk map is generated during the early part of the crop growing season (using EO data) to enable the farmer to make tactical crop management

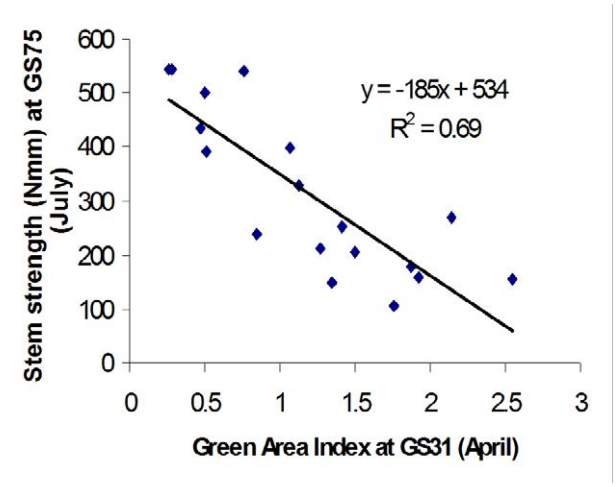


Figure 5. Winter wheat: the relationship between green area index measured at the start of stem extension (Growth Stage (GS) 31 (Tottman, 1987)) measured using a moving belt leaf area meter (Li-Cor Model 3100, Delta-T Devices, Burwell, Cambridge, UK) and the stem strength measured in the summer at mid-grain filling (GS75) using a three-point breaking test. Y = -185x + 534, R^2=0.69; *P*<0.001).

decisions to optimise nitrogen fertiliser and plant growth regulator chemicals to minimise the risk of lodging. CROPFALL uses information describing how changes to crop management affect the risk of stem and root lodging (Table 1) to adjust the lodging risk.

Results and discussion

CROPFALL predicts the risk of lodging at a range of scales from regional (Figure 6) to farm, field and sub-field scales (Figure 7). Spatial variation in lodging risk is driven by variation in topography, wind speeds, rainfall and soil type. The regional scale output indicates that the pattern of lodging risk is quite heterogeneous. The field scale output shows that substantial variation in lodging risk

Table 1. Effect of crop management on the wind speed required to cause stem or root lodging in winter wheat (Berry, 2018).

Factor	Increase in stem failure wind speed m/s	Increase in root failure wind speed m/s
Variety (1 point greater lodging resistance score)[1]	1.0	1.0
Delayed sowing (per week)	0.5	0.5
Less plants established (per 100 plants/m)	0.8	1.8
Less soil residual N (per 50 kg N/ha)	2.3	1.3
Plant growth regulator	1.4	1.4
Delayed and less fertiliser	1.4	0.8
Spring rolling (pre-GS30)	0	0.8

[1] Lodging resistance score described in AHDB recommended variety list (www.ahdb.co.uk).

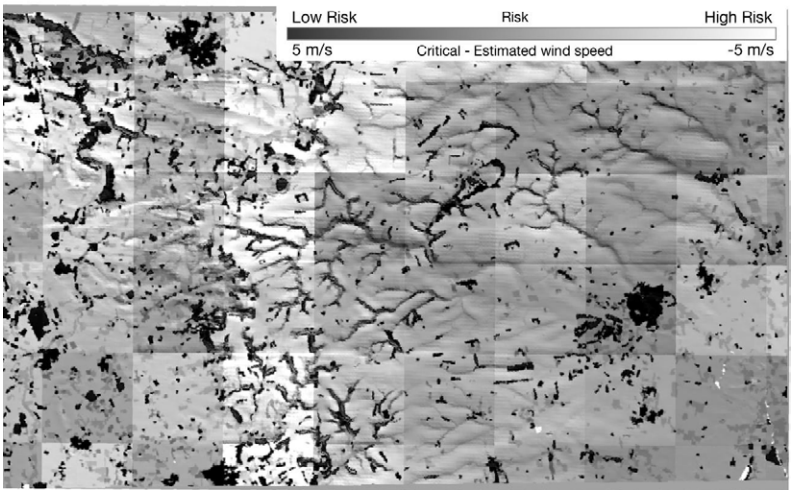


Figure 6. Predicted lodging risk of winter wheat for a 40×30 km region. The lowest risk regions depicted are expected to withstand winds of 5 m/s or more than expected for the region, whilst the highest risk areas are expected to experience lodging 5 m/s below the expected wind speeds.

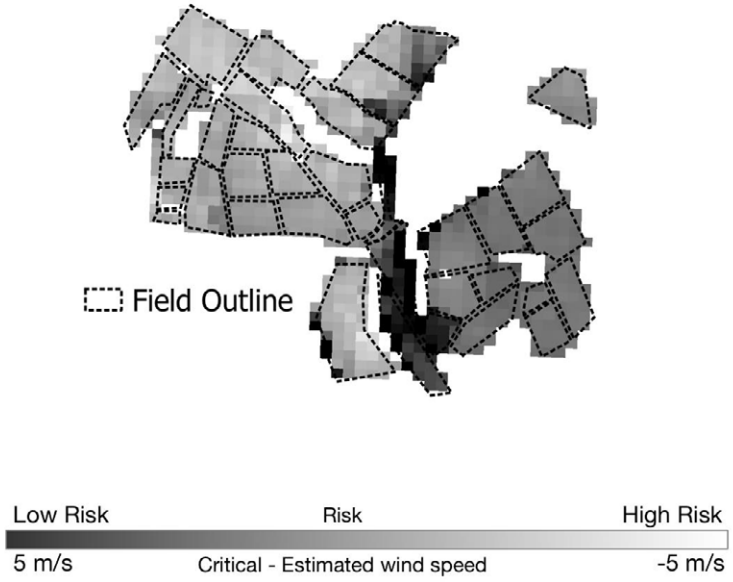


Figure 7. Lodging risk prediction for winter wheat at field and sub-field scale; Less risky areas are depicted as darker regions, whilst higher risk areas are shown lighter in colour.

should be expected both between, and, within fields. Risk is considered as the critical wind speed expected to cause lodging minus the expected windspeed at each location. Lower and negative values are considered at the highest risk, whilst higher scores are considered at the lowest risk. The

range of these values changes from site to site, which is presented qualitatively on a continuous scale of low to high risk.

CROPFALL has been used to calculate the risk of lodging for winter wheat grown on several fields in a region of the UK using a lodging susceptible variety with a Lodging Resistance score of 5 out of 9 (where 9 is highly resistant) and a lodging resistant variety with a Lodging Resistance score of 8 out of 9. The concept of the lodging resistance score is used in the UK AHDB recommended variety lists (www.ahdb.co.uk). The effect of this variety change is to predominantly high to predominantly medium risk. This illustrates that strategic decisions by the farmer using just one crop management decision can make large changes to the risk of lodging. CROPFALL also has the capacity to predict the impact of altering different combinations of crop management factors, such as seed rate, nitrogen fertiliser rate and plant growth regulator chemicals, to enable farmers and agronomists to develop the optimum crop management strategy when attempting to minimise lodging risk.

Figure 8A illustrates that there is commonly substantial variation in lodging risk within fields, with some fields containing regions with low, medium and high lodging risk. This information can be used to develop variable rate application maps for crop inputs, such as nitrogen (N) and plant growth regulators (PGR), to ensure that the regions of the field with highest lodging risk have the lowest N rate and highest PGR rate.

Conclusions

This paper demonstrates how diverse disciplines (wind engineering, geospatial science and crop agronomy) can be combined to produce a decision support system for farmers to manage lodging risk at a range of scales. Initial sensitivity testing indicates that commercially practical changes in crop management can be used to potentially minimise the lodging risk, even in environments which are conducive to lodging (high wind and rainfall). Ongoing work is further testing the CROPFALL framework approach under a range of different environmental conditions.

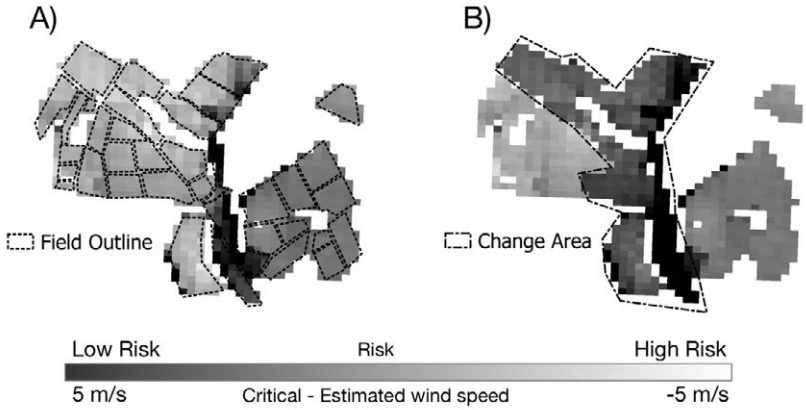


Figure 8. Lodging risk prediction for winter wheat. (A) Winter wheat variety with lodging resistance score (LRS) of 5 out of 9, field outlines shown. (B) Similar fields but with Winter wheat variety with LRS of 8 out of 9 planted in central fields resulting in a lower lodging risk.

Acknowledgements

This work was made possible through funding from the UK Biotechnology and Biological Sciences Research Council, UK (Grant Ref: BB/P004555/1).

References

Baker, C.J., Sterling, M. and Berry, P.M. 2014. A generalised model of lodging. Journal of Theoretical Biology 363, 1-12.

Berry, P.M. 2018. Lodging Resistance in Cereals. In: Meyers R. (Ed.) Encyclopedia of Sustainability Science and Technology. New York, NY, USA: Springer.

Berry, P.M., Sterling, M., Baker, C.J., Spink, J.H. and Sparkes, D.L. 2003. A calibrated model of wheat lodging compared with field measurements. Agricultural and Forest Meteorology. 119, 167-180.

Berry, P.M., Sterling, M., Spink, J.H. Baker, C.J., Sylvester-Bradley, R., Mooney, S., *et al.,* 2004. Understanding and reducing lodging in cereals. Advances in Agronomy 84, 215-269.

Berry, P.M. and Spink, J. 2012. Predicting yield losses caused by lodging in winter wheat. Field Crops Research 137, 19-26.

Tottman, D.R., 1987. Decimal code for the growth stages of cereals. Annals of Applied Biology 110, 683-687.

Extending the CERES-Beet model to simulate leaf disease in sugar beet

E. Memic[1*], S. Graeff[1] and W.D. Batchelor[2]
[1]*University of Hohenheim, Stuttgart, Germany; emir.memic@uni-hohenehim.de*
[2]*Biosystems Engineering Department, Auburn University, Auburn, AL, USA*

Abstract

In this study, techniques have been developed to use the CERES-Beet model to evaluate *Cercospora* effects on sugar beet yield and sugar content. Foliar disease effects on sugar beet growth and yield were reflected through model variables associated with crop growth and biomass accumulation and evaluated with observed values. Data collection was carried out in field experiments with different levels of disease. Model performance was evaluated based on the statistics averaged across one calibration and six evaluation experiments. Model statistics across seven scenarios were: LAI (RMSE=0.60, d-stat=0.87), tops weight dry matter (DM) (RMSE=673 kg/ha, d-stat=0.86), tuber (DM) (RMSE=2,538 kg/ha, d-stat=0.97), sugar yield (RMSE=1,960 kg/ha, d-stat=0.96), suggesting that the technique can be used as a suitable tool to simulate the impact of *Cercospora*.

Keywords: sugar beet disease, *Cercospora beticola*, CERES-Beet

Introduction

One of the most fundamental problems of various crop growth models is their inability to properly simulate foliar disease effects on crop growth and yield estimates. In the case of sugar beet (*Beta vulgaris*), leaf disease effects on sugar beet yield and sugar content are not accounted for in the simulation process and so miscalculation of field inputs like fungicides or fertilizers may result. The model calibration process itself and, subsequently, the evaluation process will lead to yield overestimations, if crop diseases are not properly accounted for. One of the earliest efforts to simulate *Cercospora* leaf spot was done by Bourgeois (1989) based on the PNUTGRO peanut crop growth model (Boote *et al.*, 1986). Similar efforts for evaluating disease effects on photosynthesis and yield estimates were done by Nokes and Young (1991).
Economic consequences of foliar diseases in sugar beet are evident and quantifiable in the context of extractable sugar losses. The leaf disease *Cercospora beticola* has a significant influence on sugar beet yield, and is the most important foliar disease of sugar beet in Germany. Economically less significant leaf diseases in Germany are: *Ramularia beticola*, *Uromyces betae*, and *Phoma betae* (Wolf and Verreet, 2002), and are normally not treated with specific fungicides. These three diseases appear later in the growing season, and are less invasive than *Cercospora*. Field scouting of *Cercospora* leaf spot is possible. However, from the moment when the disease is visible to the human eye, it is already difficult to control with fungicides without significant yield losses. Therefore, previously developed disease models are set up as forecast models taking favourable weather conditions for disease outbreak into account. One of the most significant attempts of predicting and quantifying *Cercospora* was conducted by the CERCOPRI model developed by Rossi and Battailani (1991). Rossberg *et al.* (2000) adapted CERCOPRI in order to predict early *Cercospora* epidemics and to define fungicide applications resulting in CERCBET 1. The CERCBET 1 model was modified and evolved into CERCBET 3 (Racca *et al.*, 2004). Other significant *Cercospora* forecasting models were: the leaf spot model for sugar beet (Windels *et al.*, 1998), the integrated pest management system in Germany (Wolf and Verreet, 2002), and the integrated surveillance of leaf disease in sugar beet (Mittler *et al.*, 2004). All of these model development efforts were conducted in order to reduce

DOI 10.3920/978-90-8686-888-9_120, © Wageningen Academic Publishers 2019

time consuming scouting and reduce over-application of fungicides. Chemicals used to prevent leaf disease outbreaks cause extreme environmental pollution and cannot be used without limitation, and disregard of economic costs.

Current crop disease models rely on a variety of weather models, in-field scouting and disease thresholds. Considering that the occurrence of plant disease depends on specific environmental factors and changing microclimates within a plant stand, diseases often exhibit a patchy distribution in the field. Hence the over-application of fungicides is often the case, due to a lack of techniques that can identify areas differing in disease severity in fields and delineate the appropriate fungicide strategies.

Reliable and timely assessments of plant disease occurrence and spread are the basis for planning targeted plant protection activities in the field. Visual plant disease estimations by human raters are the most common way to collect this data. In combination with advanced sensor technologies, crop modelling techniques might be as useful tool for site-specific pest management programs in sustainable crop production as they take site-specific differences in crop growth, microclimate, soil, etc. that impact disease outbreak and development into account. In combination with appropriate sensor technologies relevant input data for crop growth models can be collected.

The objective of this study was to develop a technique to simulate the impact of leaf disease on sugar beet yield and sugar content using the CERES-Beet model to attain accurate yield loss estimates based on the observed disease values. The model approach will be developed on a plot basis to delineate in a first step the impact different disease levels on beet yield and sugar content by setting up the appropriate model algorithms. In a further step, the model can then be combined with site-specific sensor data on disease occurrence to generate spatial fungicide application maps.

Materials and methods

Model development and leaf disease coupling point

Within the CERES-Beet model (CERES-Beet, Leviel, 2000; CSM-CERES-Beet, Anar *et al.*, 2015) foliar disease effects on sugar beet growth and yield are reflected through specific model variables associated with crop growth and biomass accumulation. Currently there are four ways of applying foliar damage through pest coupling points that reduce daily state variables or rate processes: (1) daily absolute damage rate; (2) percent observed damage; (3) daily percent damage rate; and (4) daily absolute damage rate with preference and competition (Batchelor *et al.*, 1993). Minor modifications of the existing pest coupling point 'daily percent damage rate' will enable the integration of foliar disease damage based on observed disease percentages in the model.

The foliar disease *C. beticola* was integrated into the model based on the following Equation 1, where x_{it} is the modified cumulative leaf area state variable after applying daily damage Di on cumulative leaf area state variable Xi on day t.

$$x_{it} = X_{it} - D_{it} \quad (1)$$

For application of leaf damage caused by *Cercospora,* the model variables 'leaf area index' and 'cumulative leaf area' were selected as major coupling points. Leaf area index is calculated from specific leaf area. Based on the calculations in the model and the direct relationship between leaf area and leaf area index, applied damage will affect both cumulative leaf area and leaf area index, as leaf area index damage is subsequently deduced from cumulative leaf area damage. Leaf mass damage proportionally affects leaf N concentration, photosynthesis and subsequently tuber dry matter as well as sugar accumulation in the model. Scouting reports or sensor based disease quantifications are read into the model as percent damage. Percent damage is calculated in Equation 2:

$$D_{it} = \frac{R_{it}}{100} X_{it} \quad (2)$$

where R_{it} is the observed damage percentage applied to the coupling point on day t. The CERES-Beet model simulates individual leaf development (Leviel, 2000), and leaf area is computed based on leaf number and leaf weight. Leaf disease severity is measured in the field and computed as an average per plant. *Cercospora* is commonly observed in the field after row closure occurs. The model can simulate disease effects by entering observed disease severity levels in an input file, which is used to simulate daily damage on leaf area and subsequently, light interception and daily photosynthesis rate. Field experiments were conducted over three years to develop and test the leaf disease subroutines in CERES-Beet. The disease epidemic thresholds were defined as plot-specific percentage of diseased leaves. Values for leaf disease infection progress (%) were converted into leaf disease damage rates (internally in the model) and applied to the selected disease coupling point. As leaf disease creates necrotic lesions, photosynthesis is reduced leading to reductions in final biomass accumulation and yield.

For statistical evaluation of the simulated results the coefficient of determination, the RMSE (root mean square error) and the d-stat (index of agreement) were used. RMSE was used to measure the deviation between measured and simulated values in the same unit (absolute measure of fit). Model performance is evaluated also with the index of agreement (more robust unitless measure), because it is more sensitive than the commonly used correlation estimator between observed and simulated means and variances. Variation range of the index of agreement is 0.0-1.0, and values closer to 1 indicate a better fit.

Field experiment description and data collection

In 2016-2018, field experiments were conducted at three different fields located at Ihinger Hof, Agricultural Research Station of Hohenheim University (30 km away from Stuttgart, latitude: 48.666, longitude: 8.967, elevation: approximately 490 m). Weather data (solar radiation, rainfall, minimum and maximum temperature) for model simulations were collected from a local station 200-600 m away from the fields. According to the World Reference Base (2006), experimental soils can be characterized as vertic Luvisol and vertic Cambisol. Organic carbon was assumed to be on average 3.8% based on historic site measurements. The sugar beet cultivar BTS940, highly susceptible to leaf diseases was planted in all experiments. In 2016, sugar beet was planted on the 120th day of the year (29. April 2016) due to a longer winter period and harvested 177 days after planting. In 2017, sugar beet was planted on the 94th day (4. April 2017) and harvested 184 days after planting. In 2018, planting took place on the 108th day (18. April 2018) and sugar beet was harvested 169 days after planting. In 2016, five different fungicide levels mainly 0, 25, 50, 75 and 100% of the recommended amounts were applied. Fungicide levels 0, 50 and 100% were investigated within this study and labelled as TRT (Treatment) 1, TRT 2 and TRT 3, respectively, in order to differentiate leaf disease ranges and effects of disease within the plots (defined as plot-specific units). Respected repetitions for investigated treatments were averaged in order to get robust model evaluation values across different treatments. All treatment repetitions (all plot-specific units) are going to be thoroughly investigated in the extended studies.

The plots were fully randomised with three repetitions. Observed values used for the simulations are mean values of the three repetitions. Total plot size was 576 m^2 (24×24 m). Plots were evenly divided into sampling and harvesting area. In 2017 and 2018, the treatment consisted of different amounts of *Cercospora inoculum* per plot named as TRT 1 (no inoculum) and TRT 2 (with inoculum), with 4 repetitions in both years. In both years, 1 g/m^2 of inoculum was applied with 5.35×10^6 spores/g in inoculated plots. In 2017, plot size was 96 m^2 (12×8 m) and in 2018 192 m^2 (24×8 m). In 2016, sugar beet plant population emergence rate was counted as 7.5 beets/m^2, in 2017 as 9 beets/m^2 and in 2018 as 8 beets/m^2. In 2016, shortly after sowing, 130 kg N/ha was applied on the field as calcium ammonium nitrate (CAN, 27% N). In 2017, 150 kg/ha and in 2018 140 kg/ha was applied as CAN (27% N) shortly after sowing.

During the growing seasons LAI was measured every two weeks non-destructively using an LAI 2000 (LICOR Inc., Lincoln, NE, USA), by taking one reference measurement above the canopy and four measurements within the canopy. Tops weight (leaf and stem separately) and tuber weight were sampled in two to four week intervals during the growing period (2016-2018). For each tuber sampling date, sugar yield was measured as percent of tuber dry matter (DM).
Leaf disease ratings were conducted after canopy closure (approximately 90% of leaves from one row were touching those in neighbouring rows), starting end of June in each year. Every 2-3 weeks, 10 middle leaves were inspected from 10 plants in 2016, 4 plants in 2017 and from 5 plants in 2018 per plot as part of destructive sampling. Based on the number of the diseased and healthy leaves, plot-specific disease progress (%) was recorded (Table 1) and used as input into the model. Calculated tuber (DM) for corresponding leaf disease percentages is shown in Table 1.
Model calibration was conducted using TRT 3 for 2016 (100% fungicide), and TRT 1 and TRT 2 from 2016 were used in addition to the treatments in years 2017 and 2018 for model evaluation.

Results and discussion

Calibration results are shown as time series graphs (Figure 1). Since leaf disease was introduced as damage on cumulative leaf area per plot, tops and tuber weight (DM) are considered as important indicators of the overall model performance. In Figure 1 (left), LAI (y0 axis) and tops weight (y1 axis) are shown with observed values. In Figure 1 (right) tuber yield (DM) and sugar yield are shown. Resulting model calibration statistics were: LAI (n=14; RMSE=0.45, d-stat=0.96, R^2=0.90), tops weight (DM) (n=7; RMSE=536 kg/ha, d-stat=0.97, R^2=0.95), tuber (DM) (n=9; RMSE=2180 kg/ha, d-stat=0.98, R^2=0.92), sugar yield (kg/ha) (n=9; RMSE=1,691 kg/ha, d-stat=0.97, R^2=0.93).

Table 1. Observed leaf disease (*Cercospora*) values as disease infection progress (%) on day of the year (DOY), tuber dry matter losses (kg/ha) based on the three years of observed data for corresponding treatments.

	DOY	Leaf disease observed values (%)			Tuber loss (DM)(kg/ha)		
2016		TRT 1	TRT 2	TRT 3	TRT 1	TRT 2	TRT 3
	183	0	0	0	0	0	0
	203	3	3	1	30	30	10
	223	13	20	20	237	324	278
	245	22	21	24	557	666	672
	258	47	25	22	981	906	909
	272	48	33	33	1563	1265	1251
2017		TRT 1	TRT 2		TRT 1	TRT 2	
	200	0	0		0	0	
	221	18	46		238	590	
	234	56	74		835	1501	
	263	92	100		3019	3949	
2018		TRT 1	TRT 2		TRT 1	TRT 2	
	214	0	0		0	0	
	227	1	7		9	64	
	234	46	52		252	357	
	241	49	57		906	818	
	255	75	79		1706	1926	
	262	100	100		2261	2467	

The model gave reasonably good estimations of observed data, but future work will be conducted to improve model performance. The model tended to overestimate tuber (DM) early in the crop season, resulting in sugar yield over-estimation as well.

For the evaluation, two treatments from 2016 and all treatments for 2017 and 2018 were used. Results are shown in Figure 2. The 1:1 graphs show measured values on the y-axis and simulated values on the x-axis. Model statistics averaged across 6 evaluation scenarios were LAI (n=62; RMSE=0.74, d-stat=0.78, R^2=0.59), tops weight (DM) (n=34; RMSE=809 kg/ha, d-stat=0.75, R^2=0.62), tuber (DM) (n=38; RMSE=2,895 kg/ha, d-stat=0.96, R^2=0.90), sugar yield (kg/ha) (n=38; RMSE=2,229 kg/ha, d-stat=0.95, R^2=0.90).

The statistics of model evaluation are satisfactory, with over- and under-estimations that are clustered around the 1:1 line. The problem with the evaluation occurred most in the 2018 data, due to serious drought in the region and the inaccuracy of the model to entirely account for those severe drought stresses. Inaccuracies in model simulations decreased in the order of LAI, tops weight (DM) and tuber (DM) yield.

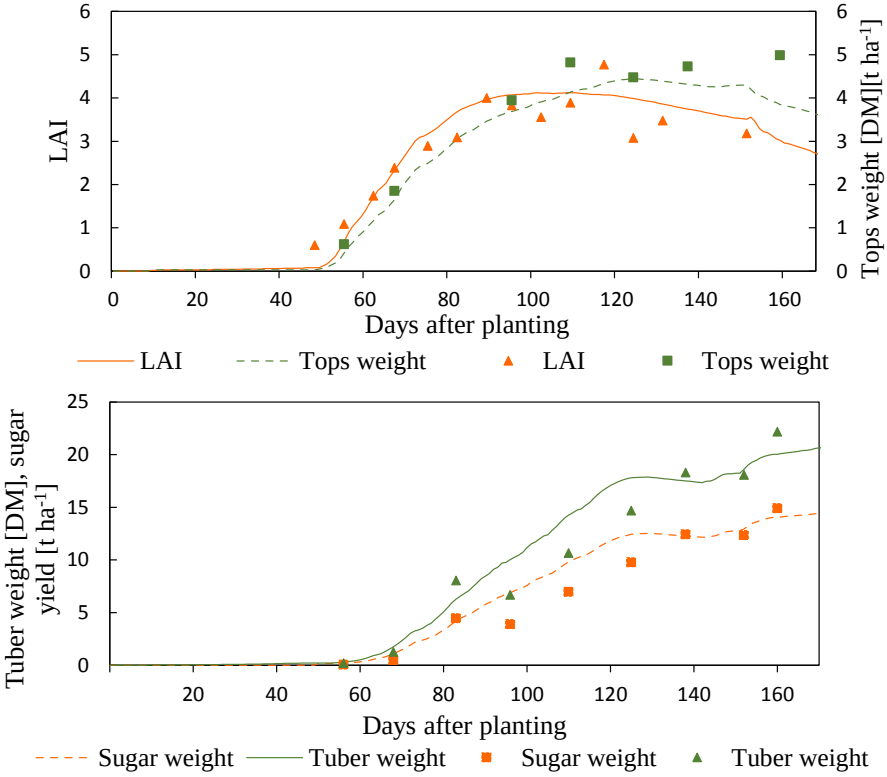


Figure 1. Calibrated tops weight (DM) and LAI (top) and tuber weight (DM) and sugar yield (bottom), all as dry matter is shown in (t/ha).

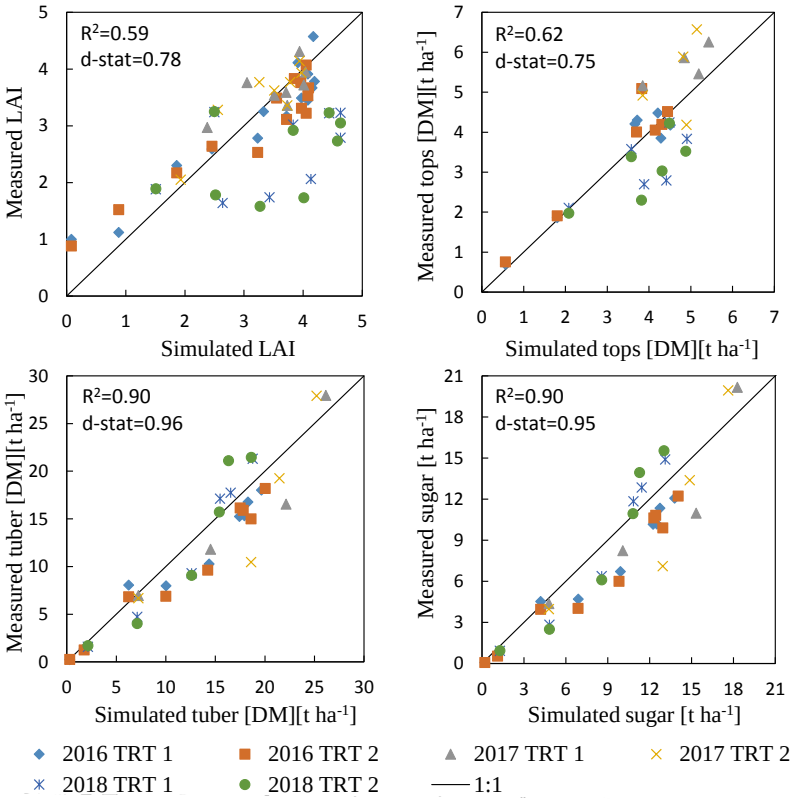


Figure 2. Evaluation of LAI (top left), tops weight (DM) (top right), tuber weight (DM) (bottom left) and sugar yield (bottom right), all dry matter is shown in (t/ha).

Future work

Leaf disease patches (*Cercospora* patches) were observed in the field and recorded in approximate sizes in m^2 per plot. This inspection was considered as a second disease threshold in order to identify the overall disease intensity per plot and thus on an area basis. This intensity will be used in studies to correlate leaf disease progress percent with weather data in order to test the potential of the model to predict diseases.

Conclusions

The goal of the study was to develop a leaf disease subroutine for CERES-Beet that can be used in combination with either manual field scouting or with sensors capable of simulating impact of different leaf disease levels on sugar beet biomass. Observed levels were passed into the model, and were then converted internally into damage rates based on the statistical fit of simulated and measured values. Introduced leaf disease effects played a very important role in attaining better calibration and evaluation results for tuber yield and sugar content during the later sugar beet growing period.

Acknowledgements

The project is supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support program (GZ 313-06.01-28-1-57.054-15), and by the National Institute of Food and Agriculture, U.S. Department of Agriculture, Hatch project (ALA014-1-16016).

References

Anar, M. J., Lin, Z., Teboh, J. and Ostlie, M. 2015. Modelling energy beets using the Decision Support System for Agrotechnology Transfer. ASABE Paper No. RRV15-048. St. Joseph, MI, USA: ASABE.

Batchelor, W.D., Jones, J.W., Boote, K.J. and Pinnschmidt, H.O. 1993. Extending the use of crop models to study pest damage. Transactions of ASAE (36) 1-8.

Boote, K. J., Jones, J. W., Mishoe, J.W. and Wilkerson, G. G. 1986. Modelling growth and yield of groundnut. In Sivakumar, M V K and Virmani, S M and Beckerman, S R (Eds.), Agrometeorology of Groundnut. Proceedings of an International Symposium, Patancheru, India: ICRISAT pp. 243-254.

Bourgeois, G. 1989. Interrelationships between late leafspot disease and florunner peanut: A modelling approach. PhD Dissertation, University of Florida, Gainesville, FL, USA.

Leviel, B. 2000. Evaluation of risks and monitoring of nitrogen fluxes at the crop level on the Romanian and Bulgarian plain. Application to maize, wheat, rapeseed and sugar-beet. Ph. D. dissertation, Institut National Polytechnique de Toulouse, France.

Mittler, S., Petersen, J., Jörg, E. and Racca, P. 2004. Integrierte Bekämpfung von Blattkrankheiten bei Zuckerrüben (Integrated treatment of leaf disease in sugar beet). In Proceedings of the 67th IIRB Congress, Brussels, Belgium: International Institute of Sugar Beet Research, pp 97-106.

Nokes, S. E. and Young, J.H. 1991. Simulation of the temporal spread of leafspot and the effect on peanut growth. Transactions of ASAE 34 (2) 653-662.

Racca, P., Jörg E., Mittler, S. and Petersen, J. 2004. CERCBET 3 – a forecaster for epidemic development of *Cercospora beticola*. In Proceedings of the 67th IIRB Congress, Brussels, Belgium: International Institute of Sugar Beet Research, pp 295-299.

Rossberg, D., Racca P., Joerg E., Kleinhenz, B. 2000. Erste Erfahrung mit dem Modell CERCBET 1 (First application of CERCBEET 1 model). Nachrichtenblatt des Deutschen Pflanzenschutzdienstes 52 (7) 153-159.

Rossi, V. and Battilani, P. 1991. CERCOPRI: a forecasting model for primary infections of cercospora leaf spot of sugar beet. EPPO-Bulletin 21, 527-531.

Windels, C.E., Lamey, H.A., Hilde, D., Widner, J. and Knudsen, T. 1998. A Cercospora leaf spot model for sugar beet: in practice by an industry. Plant Disease 82, 716-726.

Wolf, P. F. J. and Verreet, J. A. 2002. An integrated pest management system in Germany for the control of fungal leaf diseases in sugar beet. The IPM sugar beet model. Plant Disease 86, 336-344.

World Reference Base. 2006. World Reference Base for Soil Resources 2006. World Soil Resources Reports No. 103. Rome, Italy: FAO.

User needs for decision support functionalities in future crop protection software

J. Nikander[1*], S. Hara[2], L. Pesonen[2], M. Jalli[2], P. Erlund[3] and A.I. Nissinen[2]
[1]*Aalto University, Department of Built Environment, PL 14100, 00076 Aalto, Finland; jussi.nikander@aalto.fi*
[2]*Natural Resources Institute Finland, PL 2, 00791 Helsinki, Finland*
[3]*Nylands Svenska Lantbrukssällskap, Elisabetsgatan 21B 9, 00170 Helsinki, Finland*

Abstract

Farmer needs for decision support systems (DSS) were investigated using four focus group interviews, including a total of 15 Finnish farmers. The interviews were analyzed using grounded theory, which produced a selective coding describing the phenomena that influence decision making. The four main coding categories were the effects of profitability, the social network, the plant production-related factors and the operational environment. The results included farmers' opinions on existing DSS as well as the needs that they have to improve the use of DSS software in farm management. A set of preliminary software requirements were drawn from the results.

Keywords: farmer needs, decisions support systems, plant protection, requirements

Introduction

The number of distinct software systems that farmers use is increasing rapidly. The functionality and development of such systems are of significant interest. This can be seen, for example, in the works of Sorensen *et al.* (2010a) and Lawson *et al.* (2011), or Fountas *et al.* (2015), where 140 farm management information systems (FMIS) were reviewed. As the number of systems a farmer uses increases, so does the need for the systems to either fit with existing farming processes or improve them (Sorensen *et al.,* 2010b). New systems that are hard to use, or unsuitable for the existing farming processes, add to the farmer's workload and often provide no benefit.
An increasing number of decision support systems (DSS) are available to help farmers. However, DSS software has not been readily adopted by farmers. Mackrell *et al.* (2009) list common reasons for the poor adoption rate of DSS software, such as complexity of the software, wide prior knowledge requirements for using the software, inconsistencies between the software and farming objectives, insufficient software maintenance, as well as difficult data capture and entry into the software. The problems stem from many sources, including cognitive overload for farmers who need to use many software systems, difficulties in formalizing the farmers' decision making processes, and the strong links between farming activities and specific decisions. Therefore, it is challenging to develop DSS for farming activities, while farmers have interest in systems that would improve farming.
This study investigated farmers' needs for DSS for agricultural pest management. The problems in the work are related to economically efficient but environmentally sound decision making on the use of pesticides. The objective was to study farmers' decision making processes concerning plant protection measures to identify the critical points in the process where DSS could facilitate their work.

Materials and methods

The data used in this work consisted of four focus group interviews, three of which included four farmers and one group with three farmers. The farmers were invited to the groups with

DOI 10.3920/978-90-8686-888-9_121, © Wageningen Academic Publishers 2019

purposeful sampling strategy, where invitations were sent to farmers known to be interested in and knowledgeable about the issues, and ready to voice their opinions on them.

Each interview used the semi-structured group interview approach, where the interviewer had 10 prepared questions to ask from the group, who were given relatively free range in discussing the issues. Additional questions were asked in case the discussion touched topics that were felt relevant by the interviewer. All interviews were recorded and transcribed, and the set of questions was revised after each interview to contain insights gained. After four interviews, the researchers decided that the discussion had sufficiently saturated the topics.

The interview transcripts were analysed using Grounded Theory (Glaser and Strauss, 1999, Strauss and Corbin, 1990). It is a method, where categorization of a data set is derived from the data, instead of using an existing framework. The results of a grounded theory analysis is a coding on the data, which is a hierarchical categorization that describes the phenomena. Data analysis in grounded theory consists of three phases: open coding, axial coding, and selective coding.

In Open coding, the data is examined, and elements in it are named and arranged into preliminary categories. In this part of the work, each data element – lines of the interview transcripts in this work – and its meaning is examined and named. Then, the work moves to axial coding, where the data is grouped into categories, and these categories are turned into different dimensions used to further arrange the data. Finally, in selective coding, the dimensions created in the axial coding are arranged into hierarchies, where most significant categories serve as the core of the categorization, with less significant and more specific elements arranged under these. The result is a hierarchical arrangement of the data that can be used also to categorize further data sets.

The results of this work cover four focus group interviews that had a total of 15 farmers. The interviews were held between July 2017 and February 2018. Each group was asked ten questions, which were slightly revised between interviews. The interviews were held in Finnish, and Table 1 shows the translations of the questions asked in the final interview.

Results

The result of the interview analysis was a selective coding focused on the phenomena that influence farmer decision making on plant protection operations. Based on this coding, user needs for plant protection DSS were derived. The main categorization of the coding is shown in Figure 1. There are four main categories: the effects of profitability, the social network, the plant production-related

Table 1. Interview questions.

1.	Do you use plant protection advisory services, and if you do, what services?
2.	How do you use software for cultivation planning and what do you think about their functionality? What other software do you use?
3.	What kind of information do you need from previous years' plant protection implementation/ realization to make the next year's plant protection plan?
4.	What kind of management practices do you use to manage plant diseases?
5.	Which plant diseases do you take into account in the plant protection plan?
6.	How do you make decision about the timing of plant protection measures?
7.	Where do you get the information to support your plant protection decisions? Do you cross-check the information from other sources?
8.	What kind of observations concerning plant protection issues do you make during the growing season? Do you leave an untreated plot for control purposes?
9.	In what kind of situations do you not make plant protection measures?
10.	How do you estimate whether plant protection measures were a) effective b) profitable?

factors, and the operational environment. These main categories were divided into sub-categories, which can in turn be divided into more categories up to three levels of sub-category hierarchy, as shown in Figure 2. As the main category operational environment includes most of the parameters relevant to current DSS for plant diseases, its subcategories are presented here in detail.

The profitability category contains the means that farmers use to ensure plant protection measures increase profit gained from the harvest. The relevant profitability indicators depend on the type of farm, but in most cases the profitability is based on the expected yield from the crop. The expected yield is compared to the actual yield and yield quality, as well as to the yield due to plant protection measures. An important restriction is also the generally low profitability of farming in Finland and weak economic situation of many farms.

The farmers' social network contains the networks and services the farmers use as the main sources of knowledge when making plant protection decisions. This category consists of contacts with advisory services and other farmers, plant protection product producers and resellers, as well other parties with whom the farmer has entered into contracts. In addition, farmers increasingly also get information from various internet services, where there is less direct human contact.

The production-related factors include the farmers' tacit and explicit expertise and ability to use this expertise to create plant protection plans. The expertise includes understanding the life cycle of the pests, the conditions, where the pests are capable of infecting the crop, and how pest infections affect the plant growth and the yield, as well as how infections can be detected and recognized, how pests build resistance to various active ingredients, and how to consider all this in plant protection planning.

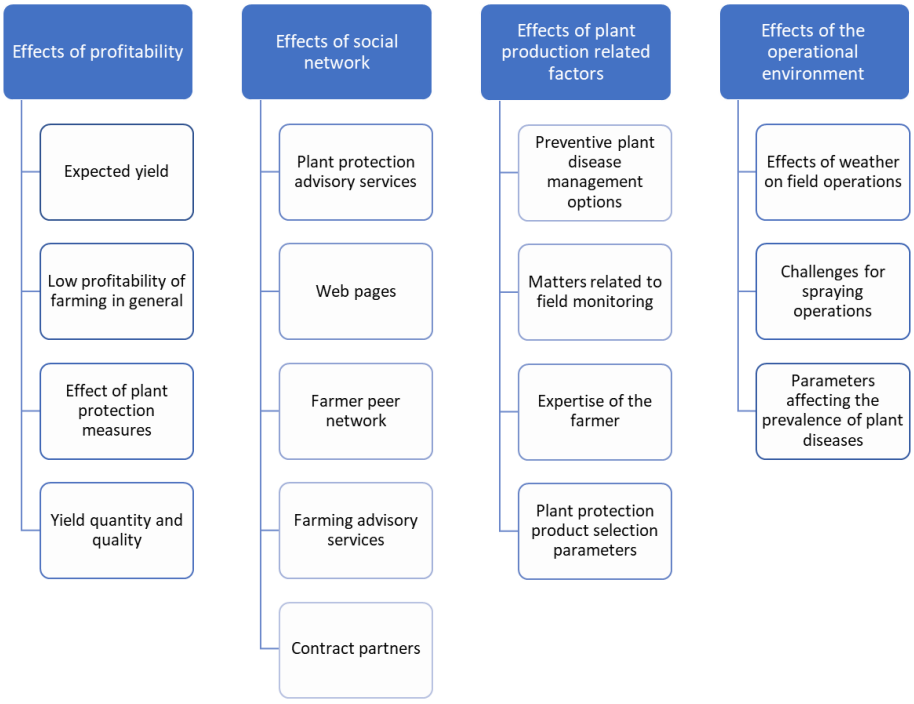


Figure 1. The main categories of selective coding focused on phenomena that influence farmer decision making regarding plant protection operations.

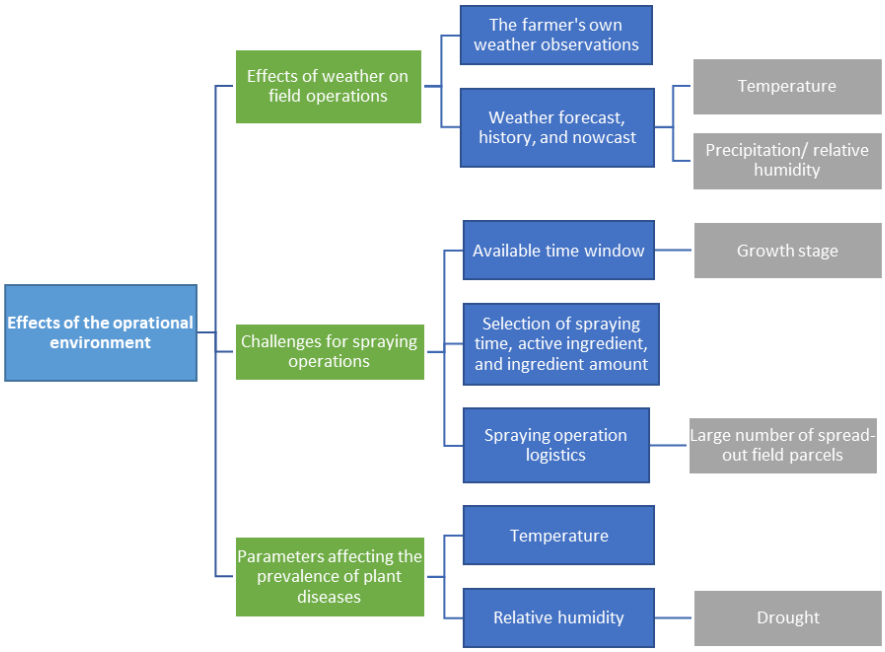


Figure 2. The effects of the working environment on the decision making.

Finally, the operational environment describes how the farmers make plant protection decisions during the growing season and execute plant protection plans. The most important aspect of the decision making in the operational environment is whether there is a need for plant protection, either currently or in the near future. Should there be a need, then the weather history, current weather and weather forecasts influence the decision whether to execute a plant protection operation. Weather can also affect the decision making during operation; high winds or unexpected rain showers can stop the plant protection operation.

The operational environment sets strict limits for executing the plant protection measures at the farm. The short time frame, in which the decision has to be made, was described by one of the farmers: 'The weather is suitable for spraying, vegetation has been growing more quickly than expected, it is 8 pm on Friday and the decision must be made here and now. And no [supporting] information is available.' Many of the interviewed farmers hoped for technical tools that would facilitate making the plant protection decision in a very limited timeframe. Another challenge is that some diseases do not show visible symptoms in the early stages of infection, and the farmer should make plant protection measures when the foliage still looks healthy. Then information about the current disease pressure is crucial. These needs have to be addressed in future DSS systems.

In addition, FMIS need to support automatic documentation of plant protection measures taken. Currently, many farmers print the plant protection plan, take the printed document with them into the tractor, write down the actualized measures into it, and later manually insert this information to the FMIS. However, farmers identified that using a printed copy in tractor as a risk of losing their field notes. Therefore using a tablet or a phone that stores the data directly into a cloud service was suggested as a way to save the notes, even in case of destruction of the device. A farmer, who had used a phone to make field notes, found the small size of the phone screen a challenge to work with the long lists of pesticides. His suggestion was that it should be possible to make a tailored menu of the products that have been brought to the farm to diminish the possibility of selecting a wrong

product from the long menu. It should also be possible to select recently used products without having to search for them again.

Another group of farmers suggested that the sprayer could directly send the data to the FMIS, including amount of active ingredient(s) and water used. Farmers also expressed a hope that weather observations during the spraying operation could automatically be transferred into the FMIS. This functionality would require an on-vehicle weather station. One farmer discussed on-vehicle weather stations: 'Each sprayer could have its own weather station which would continuously record the temperature, wind direction and speed and send the data directly to the computer. The sprayer [weather station] could also receive data from the nearest ground weather station and it could forecast when the rain starts. Then the employee doing the spraying would have time to leave the field [before it rains].'

One critical step in adopting a DSS is their compatibility with FMIS. One of the current DSSs was found quite easy to use, but there was no open interface in the FIMS to get data of field operations directly to DSS. Farmers are not willing to use their time to input the same data into several systems, and are likely to abandon the DSS, if extra work is required. Several farmers felt that to better utilize their own expertise when making the plant protection plans for the coming season, they would need better access to the plant protection history of the field plots. Field history information is required, for example, to fulfil the farming restrictions for the consecutive use of certain active ingredients e.g. in groundwater areas. Furthermore, farmers would like to document their observations about plant protection hot spots, such as patches of wild oats or herbicide resistant weeds. A way to document the issue could be photos added to FMIS. However, this kind of functionality is not available in current systems used in Finland. It was also suggested that when the farmer enters the field, warnings on some special issues should pop up in the app. This would require an application that is location-aware.

The low profitability of plant production sets strict limits to the plant protection measures used. Several farmers mentioned that they will not apply any plant protection measures, if the vegetation is poor and it will not fulfil their expectations of yield level. Economic impact of the plant protection measures was clear for some of the farmers and they were very aware of the costs, for example: 'I have the knowledge of the variable costs by crop species in my head … while I am walking in the field plot and making the decisions.' However, some farmers were not so confident about their own knowledge concerning economic impact of plant protection measures. One farmer suggested to include yield level as a parameter for plant disease models. That allows the DSS to give a recommendation about the active ingredient and to estimate profitability of spraying. Some farmers left untreated control plots to assess the impact of plant protection measures, while other farmers did not. However, both groups assessed the success of the measures by quantity and quality of the yield.

Some farmers made their plant protection decisions without consulting others. When asked if they cross-check the information from other sources, one of them answered: 'what other source could there be, when I have made my own observations?' However, in addition to running their own farm, these farmers usually did additional work, such as selling farming advice or contracting plant protection services to other farmers. There were also farmers who got information to support their decisions from advisors, trusted pesticide vendors, webpages, other farmers, or from the farmers' groups in Facebook or in WhatsApp and/or a newspaper. This group of farmers mainly consisted of part time farmers, or farmers whose main production line was animal husbandry. Therefore, they did not have as much time to invest in developing knowledge of crop protection as the full time cereal farmers. Therefore, the needs for DSS are quite different for farmers with different specializations. The first group seemed to be satisfied with warnings of high plant disease probability and instructions to monitor the crop, while the second group preferred a message that contains the control protocol up to the active ingredient to be used.

Overall, the social aspect of the decision making process was closely linked to the reliability of the decision support information available. One farmer described this: 'When checking the messages

in the mobile phone some messages do not startle you, but if [the message] comes from three advisors or similar sources, then it may [startle you].' The farmers also expressed their concern that the retirement of experienced advisors and trusted pesticide vendors will affect the reliability of the decision support information available.
In the future, the observation network behind disease prognosis should be denser, in order to give predictions that are more reliable. However, a suggestion that farmers would make observations themselves and share them in a network, raised the question about reliability of the information and difficulty of controlling their quality. Therefore, persons making the observations should be trained for the task, which led to the idea about an observation network of experimental farms to validate the predictions available for cereal diseases.

Discussion

The goal of surveying the farmers' needs for decision support functionality was to provide knowledge for further development of existing DSS and FMIS, as well as to influence the development of future systems. In software development user needs typically are tracked starting from very early parts of the project, and can influence the whole development cycle. Classically, user needs are used to create system requirements.

Suggestions for DSS requirements derived from the interviews

The software requirements process itself is complex work that requires expertise. Thus, the following discussion does not present a set of requirements that could directly be used for software development, but provides information that could be exploited by software development projects. Thus, the contents in Table 2 merely summarizes the most important results of the interviews. At most, the Table 2 could be used as a basis for specific requirements for a DSS or FMIS.

Table 2. Farmer needs and corresponding software requirements.

Farmer needs	DSS requirements
Decision support for making plant protection decisions under strict time constraints	DSS should work without human in the loop
Documentation of pest hotspots, and adding field photos to FMIS	Inclusion of geolocated, semi-structured data into the system, photo geolocation
Field-specific warnings upon entering a field	Device geolocation, addition of location-based services to the system
Access to the fields' pest management history	Improved access to field history, user interface improvements for accessing historical data
Automatic documentation of field operations to FMIS	Integration of FMIS and field machinery systems, field data storage and analysis
Weather data integration and storage	Integration of weather data into FMIS, standardization and analysis of weather data
Prediction of economic effects of plant protection measures	Inclusion of economic modelling to DSS
Taking into account the farmer's expertise and circumstances	Different functionality and user interaction for different types of farmers, such as full-time or part-time
Ability to combine observational data from several farms	Vendor-level support for FMIS data integration between different FMIS instances

Some of the requirements listed in Table 2 are already in development, such as the integration of field machinery and FMIS systems in the ISO-11783 standard (ISO, 2015). There are also aspects that are increasingly important as the digitization of agriculture continues, such as the ability to store geolocated data and use it in decision making and user interaction. There are also challenging needs to implement, such as adding economic modelling to DSS even on general level. Similarly, the development of many user interface functionalities may be challenging, as the usability is often at odds with the amount of functionality provided.

Conclusions

Currently, farmers are critical of the decision support provided by existing systems, as well as about FMIS in general. The criticism can, in turn, be used as a basis for the requirements for new systems. In this work, four groups of farmers were interviewed, and the results analysed and turned into a hierarchical description of farmer needs. Several critical points were noted. The need for better integration with farm machinery is currently being standardized. There is a need for geolocated data, including unstructured or semi-structured data such as photos and farmer notes, which would allow the farmers better exploit existing farming history data during farm management, operation planning and execution. Finally, the graphical user interfaces received significant criticism and thus would require further development.

Acknowledgements

This research has been carried out as a part of the SpotIT project (https://www.nibio.no/en/projects/spotit) funded under the C-IPM ERA-NET. The Finnish funding agency for the project is the Ministry of Agriculture and Forestry with funding decision 377/03.01.02/2017. The authors thank all the project partners and interviewed farmers for their help in the project.

References

Fountas, S., Carli, G., Sørensen, C.G., Tsiropoulos, Z., Cavalaris, C., Vatsanidou, A., at al., 2015. Farm management information systems: Current situation and future perspectives. Computers and Electronics in Agriculture, 115, 40-50.

Glaser, B.G. and Strauss, A.L., 1999. Discovery of grounded theory: Strategies for qualitative research. New York, USA: Routledge.

ISO, 2015. Tractors and machinery for agriculture and forestry -- Serial control and communications data network -- Part 10: Task controller and management information system data interchange. Standard 11783-10, International Organization for Standardization

Lawson, L.G., Pedersen, S.M., Sørensen, C.G., Pesonen, L., Fountas, S., Werner, A., *et al.*, 2011. A four nation survey of farm information management and advanced farming systems: a descriptive analysis of survey responses. Computers and Electronics in Agriculture, 77(1), 7-20.

Mackrell, D., Kerr, D. and Von Hellens, L., 2009. A qualitative case study of the adoption and use of an agricultural decision support system in the Australian cotton industry: The socio-technical view. Decision Support Systems, 47 (2), 143-153.

Sørensen, C.G., Fountas, S., Nash, E., Pesonen, L., Bochtis, D., Pedersen, S.M., *et al.*, 2010a. Conceptual model of a future farm management information system. Computers and Electronics in agriculture, 72(1), 37-47.

Sorensen, C.G., Pesonen, L., Fountas, S., Suomi, P., Bochtis, D., Bildsøe, P. *et al.*, 2010b. A user-centric approach for information modelling in arable farming. Computers and Electronics in Agriculture, 73, 44-55.

Strauss, A. and Corbin, J.M., 1990. Basics of qualitative research: Grounded theory procedures and techniques. Thousand Oaks, CA, USA: Sage Publications, Inc.

Automated mixed-scale data fusion for mapping of within-field variation in a national decision support system – the example of pH correction

K. Piikki[1*], M. Söderström[1], H. Stadig[2] and J. Martinsson[3]
[1]*Dept. of Soil and Environment, Swedish University of Agricultural Sciences (SLU), Skara, Sweden; kristin.piikki@slu.se*
[2]*Rural Economy and Agricultural Societies, Skara, Sweden*
[3]*DataVäxt AB, Hyringa Hedåkers Säteri 3, Grästorp, Sweden*

Abstract

Systems in which it is possible to easily combine local and global data for improved local accuracy is an important development for the practical use of digital soil mapping products. In this project, the objective was to further improve this with new map layers in a soil database, combined with interactive functionality for downscaling. Two new raster map layers (buffering capacity and target-pH) were developed and added to the Digital Soil Map of Sweden (DSMS) that previously contained topsoil texture data in a 50×50 m grid. The new map layers were developed using airborne gamma radiation data, a laser-scanned digital elevation model, legacy Quaternary deposit maps and a national soil sample database. In addition, an interactive and automated functionality for downscaling was developed in a freely available decision support system (DSS; Markdata.se) to enable use of the new map layers for within-field variable-rate liming for soil pH correction. In this procedure, local soil data are uploaded to *Markdata.se* and used for local adaptation of DSMS.

Keywords: decision support, digital soil map, variable rate application, scale, soil pH

Introduction

The recommended soil pH for crop production (*target-pH*) depends on soil type; the target pH is relatively low in soils with a low clay content and/or a high content of organic matter (OM), and *vice versa* (Börling *et al.*, 2017). It also depends on the crop species; e.g. if sugar beet (*Beta vulgaris*) are produced, the recommended target-pH is 0.5 pH units higher than for cereals (*ibid.*). The amount of lime required to raise soil pH by one unit (here denominated *buffering capacity*) also depends on soil type, with a larger amount of lime required on soils with a high clay content and/or a high content of OM.

The Digital Soil Map of Sweden (DSMS) is a spatial soil database that covers >90% (~2.4 Mha) of Sweden's arable land. The data, which is stored in raster format at a spatial resolution of 50×50 m, is publicly available under CC-BY license. Before the present study, there were map layers for topsoil clay content, sand content and FAO soil texture class according to FAO ITPS (2015).

Detailed digital soil maps, such as the DSMS, are increasingly being produced for regional, national and even global extents. It would obviously be very cost effective to apply such maps in precision agriculture compared to using maps based on local soil sampling. A number of digital soil mapping products may provide sufficiently high spatial resolution to give the impression that no further soil analyses are required. However, the often large standard error in the predictions at point locations make such data unsuitable for use in precision agriculture. Söderström *et al.* (2016; 2017a) demonstrated that it is often possible to improve the accuracy in a sub-region of a map (e.g. a farm, a field or a country) by use of local soil sample data. Piikki *et al.* (2017) evaluated local adaptation of the DSMS clay content layer on 403 farms. They found that, in some cases, the digital soil map was most accurate. In some cases a map produced solely by the local soil sample data and ordinary

DOI 10.3920/978-90-8686-888-9_122, © Wageningen Academic Publishers 2019

kriging was the most accurate, but most often (in about 50% of the farms) a map produced by a combination of the two adaptation methods used (residual or regression kriging) was the most accurate. They proposed to produce all four maps and chose the locally best one.
This strategy has been implemented in the R package mapsRinteractive (Piikki and Söderström, 2018a) and a Swedish decision support system (DSS) for precision agriculture based on DSMS: *Markdata.se* (Dataväxt AB, Grästorp, Sweden). In the first version of this DSS, the clay content map layer is used for designing a variable-rate application (VRA) file, e.g. for variable seed rate or variable liming rate for soil structure improvement. The user can also decide to upload their own sample data to locally improve the national map and the VRA file. The goals of the present study were:

1. Two new DSMS raster layers of topsoil attributes: buffering capacity and target-pH
2. A new functionality in the DSS *Markdata.se*, where users: (1) can upload local soil pH data, and optionally data on texture and OM; (2) can automatically derive and manually edit lime requirement maps of their own farm; and (3) download VRA files for liming for soil pH correction in various file formats.

Materials and methods

Development of national map layers required for pH liming

The development of national map layers of target-pH and buffering capacity largely followed the procedure to develop already existing soil texture and OM layers in DSMS (summarized below in 1-3). Furthermore, it is based on the same reference soil sample data and the same predictor raster datasets. More details are given by Piikki and Söderström (2018b).

1. Reference data preparation: Target-pH and buffer capacity were calculated for a national soil sample dataset (n=13,500) by applying pedotransfer functions (Equations 1 and 2) on lab results for clay and OM contents. Equations 1 and 2 were derived from tabulated values in Börling *et al.* (2017). The soil samples were taken in a 1-m square grid with a small random displacement of point locations (Paulsson *et al.*, 2015).

$$Target - pH = 6.1 + 0.01 \times clay\,[\%] - 0.033 \times OM[\%] \quad (1)$$

$$Buffering\ capacity\ [unit] = 1.9 + \frac{3.5 \times OM\ [\%] + clay\ [\%]}{3.8} \quad (2)$$

where OM = organic matter content, and unit = tonnes of lime product with 50% CaO per ha needed to raise pH by one unit.

2. Predictor data preparation: Raster datasets of a number of ancillary variables (explained in detail in Piikki and Söderström, 2018b) were prepared:
   a. airborne gamma radiation data (isotopes ^{232}Th, ^{40}K and ^{238}U);
   b. airborne laser-scanned digital elevation data (elevation and topographic position indices in three scales);
   c. legacy Quaternary soil maps (two simplified classifications).
3. Model development and validation: Multivariate adaptive regression splines models (MARSplines; Hastie *et al.*, 2009) were parametrized by the national soil sample dataset. The target-pH and buffering capacity were response variables and extracted data from the raster stack of ancillary variables were used as predictors. First, a random k-fold cross validation (k=10) was performed and validation statistics were computed (details in next section). Then, final models to be used for mapping were parameterized without omitting any samples.
4. Mapping: The final MARSplines models, parameterized using all data, were deployed on the predictor raster stack to make spatially exhaustive predictions of the response variables.

In practice, the entire map area was manually split into 22 sub-regions, for which models were parameterized and deployed individually. The resulting prediction rasters were then merged to

constitute the final map. Mapping and modelling were carried out using ArcGIS (ESRI Inc. Redlands, Ca, USA) and R (R Core Team, 2018) with extension packages earth, maptools, raster, rgdal, rgeos and utils.

Validation of national map layers required for pH liming

A set of validation measures were computed to quantify the accuracy of the produced target-pH and buffering capacity raster layers (Equations 3-5). The mean absolute error (MAE; Janssen and Heuberger, 1995) reflects the mean magnitude of prediction errors irrespective of whether the errors are over-predictions or under-predictions. The mean error (ME; *ibid.*) is the bias; it indicates whether the target variable is generally over-predicted (positive ME) or under-predicted (negative ME). The Nash-Sutcliffe modelling efficiency (E; Nash and Sutcliffe, 1970) is a measure of whether the predictions are better than a mean of the observations. A positive value of E indicates an improvement over the mean of the observed values, while a negative value of E indicates that the predictions are less accurate than the mean, which means that the model performance is very poor. The validation statistics were calculated on the 22 sub-region sets of predicted and observed values obtained by leave-one-out cross validation of the MARSplines modelling in the sub-areas.

$$MAE = \frac{\sum|p-o|}{n} \quad (3)$$

$$ME = \frac{\sum(p-o)}{n} \quad (4)$$

$$E = \frac{\sum(o-p)^2}{\sum(o-\bar{o})^2} \quad (5)$$

where o = observed value; p = predicted value; n = number of observations.

Implementation of functionality to derive VRA files of lime in the DSS

The DSMS raster maps of buffering capacity and target-pH were made available through the web-based DSS *Markdata.se*. The system had to be designed so that it was possible for the user to interactively upload a file with current soil pH data from soil analyses, and optionally also clay content, OM content and magnesium (Mg) content analyses. If clay and OM contents were included, the national map layers of buffering capacity and target-pH should be locally adapted according to Piikki *et al.* (2017), otherwise the DSMS data should be used without modification.

Results

New DSMS map layers

The buffering capacity and target pH map layers, as well as their distributions are presented in Figure 1. The spatial distribution of the maps (in Figure 1A and 1B) reflects the distribution of cropland in southern Sweden.

The accuracy measures from the leave-one-out cross validation of the MARSplines modelling of target-pH and the buffering capacity are presented in Table 1. The positive E values show that the regional modelling was a considerable improvement over using the regional mean observed value for both target-pH and buffering capacity. The relatively large standard deviation indicates a variation in model performance among the mapping regions. The ME values show that there was almost no bias of the target-pH predictions and only a slight positive bias (general over-prediction) in the buffering capacity. The MAEs were considered acceptable for practical application.

Implementation in the web based DSS

The functionality to produce lime requirement maps from a combination of national and local data was implemented in the free-to-use DSS *Markdata.se*. A schematic illustration of the internal

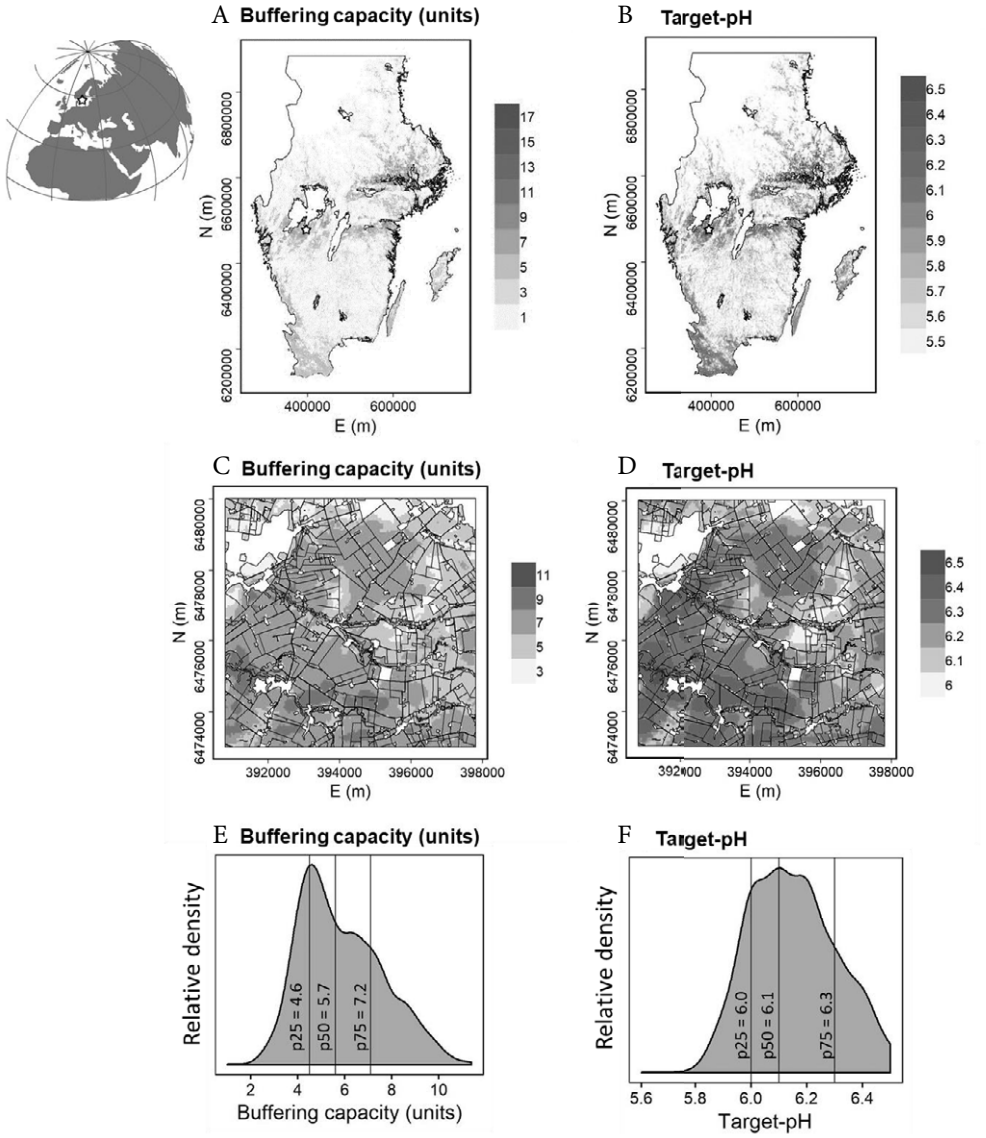


Figure 1. Map layers of (A) buffering capacity and (B) target-pH for the entire extent of the Digital Soil map of Sweden (DSMS) and C) buffering capacity and D) target-pH for a 7×7 km area. The small star on the overview map shows the location of the DSMS area and the small stars in maps (A) and (B) show the locations of maps (C) and (D). The density plots (E) and (F) display distributions and percentiles (p25, p50 and p75) of maps (A) and (B). The unit of the buffering capacity is tonnes of lime product with 50% CaO required per hectare to raise soil pH by one unit. Projected co-ordinate system: Sweref99TM (EPSG: 3006). Field boundaries: Swedish Board of Agriculture (Jönköping, Sweden).

Table 1. Validation statistics. Mean values standard deviation of validation measures computed for the 22 sub-regions for mapping. E = Nash-Sutcliffe modelling efficiency; ME = mean error (bias); MAE = mean absolute error. The unit of the buffering capacity is t lime/ha with 50% CaO required to raise the pH by one unit.

Topsoil attribute	E (relative)	ME (units)	MAE (units)
Target-pH	0.40±0.18	0.00±0.00	0.07±0.01
Buffering capacity (t/ha)	0.31±0.16	0.05±0.15	2.37±0.43

functionality of the system is shown in Figure 2. In essence, the user uploads a text file with georeferenced soil sample data, containing at least soil pH analyses. If magnesium (Mg) data are included in the uploaded file, the user will be notified by a pop-up text message in case the Mg status of the soil is below a certain threshold and, if so, he/she is advised to consider the low Mg status when choosing the lime product. The web interface was programmed as a stepwise guide, where the user successively inputs required data and makes choices (first step shown in Figure 3). In the last step, the lime requirement maps produced can be downloaded in file formats for equipment from various manufacturers. In addition, ESRI shapefiles and text files can be downloaded. Data can also be exported to a cloud-based service and accessed through a smartphone app whereby it can be transferred wireless to the tractor and directly control e.g. a lime spreader (Controlmaster; Dataväxt AB, Grästorp, Sweden). The *Markdata.se* DSS has been visited by approximately 3,000-4,000 unique users per year during the period 2017-2018.

Discussion

To monitor and adjust soil pH where needed is important for nutrient use efficiency in crop production. Sub-optimal soil pH affects availability of nutrients to crop roots and, in severe cases, crops become unresponsive to fertilisers, even if the fertiliser nutrient elements are deficient in the soil. To reach, but not exceed, the target-pH at all locations in a field with non-uniform soil-pH or soil type, variable rate application (VRA) of lime is required.

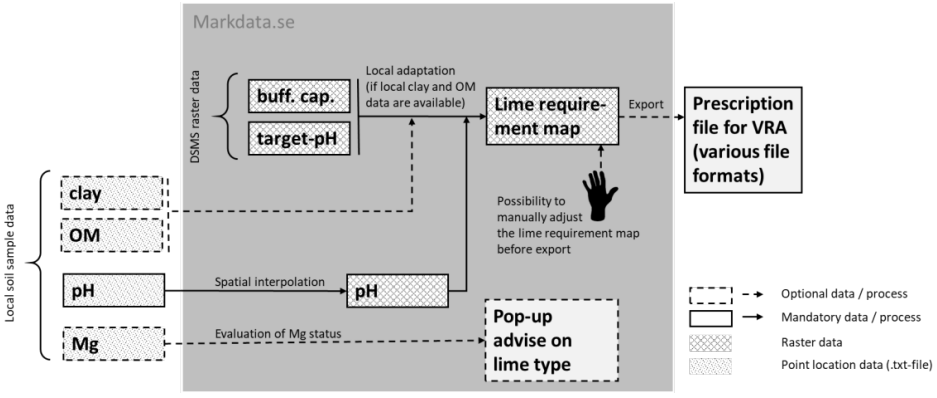


Figure 2. Schematic illustration of data and processes in the DSS *Markdata.se*. OM = organic; buff. cap. = pH buffering capacity. VRA = variable lime rate application.



Figure 3. The start-up page of the decision support system for variable rate application *Markdata.se* (translated to English).

Experience from Swedish crop production is that the practical implementation of VRA in principle follows what is described by e.g. Mulla and Khosla (2015). The interest among farmers increases when it becomes easier to practice VRA of inputs. For example, in 2014, a free web-DSS for VRA of nitrogen (N) was launched (Söderström *et al.*, 2017b). With this system, it was easy to view satellite-based vegetation index maps of crop fields, design VRA and export prescription files; it rapidly gained a large number of users. At the same time, some technological problems for implementing VRA have recently been solved. Many farmers have invested in auto-steering equipment, including high accuracy positioning systems and new user-friendly technologies for wire-less data transfer from DSSs to spreaders and other machinery that are implemented in the *Markdata.se* system that has been launched.

Currently in Sweden, liming is most often carried out by entrepreneurs. According to one of the leading actors in this business, about 40% of their clients vary the lime-rate across the fields, while about 60% apply uniform lime-rates. With a web-based DSS that allows easy within-field mapping of lime requirement for pH correction, the adoption rate of VRA of lime may increase. In the long run, this may have a positive impact on productivity and nutrient-use efficiency in crop production.

Conclusions

New map layers of topsoil target-pH and topsoil buffering capacity were produced in the national Swedish digital soil map DSMS from a set of exhaustive covariates (airborne gamma radiation, a laser-scanned digital elevation model and Quaternary deposit maps) and data from a national soil sampling campaign. The map layers had mean absolute errors of 0.07 pH units, and 2.4 t/ha of lime with 50% CaO, respectively. A new function was implemented in a DSS for precision agriculture *Markdata.se*, where the users can upload their own pH values and, by use of the new DSMS maps, derive lime requirement maps for soil pH correction and download prescription files in various formats.

Acknowledgements

This project was funded by the Geological Survey of Sweden, SGU (contract: 15-03-31), the Swedish Foundation for Agricultural Research (contract: O-15-20-566), the Västra Götaland Region and the Swedish University of Agricultural Sciences (contract: RUN 2018-00141). We acknowledge valuable access to data from the Swedish Board of Agriculture.

References

Börling, K., Kvarmo, P Listh, U., Malgeryd, J. and Stenberg, M. 2017. Rekommendationer för gödsling och kalkning 2018 (Recommendations for fertilisation and, liming 2018). The Swedish Board of Agriculture, JO17:4, (in Swedish) pp. 83-87.

FAO ITPS 2015. Status of the World's Soil Resources (SWSR) – Main Report. Food and Agriculture Organization of the United Nations and Intergovernmental Technical Panel on Soils, Rome, Italy.

Hastie, T., Tibshirani, R. and Friedman, J., 2009. The Elements of Statistical Learning: Data Mining Inference and Prediction, 2nd edition. Springer Series in Statistics, New York, pp. 383-411.

Janssen, P. H. M. and Heuberger, P. S. C. 1995. Calibration of process-oriented models. Ecological Modelling 83 55-66.

Mulla, D. and Khosla, Raj. 2015. Historical Evolution and Recent Advances in Precision Farming. In: Lal, R. and Stewart, B. A. (Eds). Soil-Specific Farming: Precision Agriculture, Boca Raton, FL, USA: CRC Press, Taylor and Francis. pp. 1-36.

Nash, J. E. and Sutcliffe, J. V. 1970. River flow forecasting through conceptual models part I – A discussion of principles. Journal of hydrology 10 282-290.

Paulsson, R., Djodjic, D., Carlsson, Ross C. and Hjerpe, K. 2015. Nationell jordartskartering. Matjordens egenskaper i åkermarken (National mapping of soil types. Topsoil properties in arable land). Swedish Board of Agriculture Report: RA15:19 (in Swedish). 44 pp.

Piikki, K., Söderström, M. and Stadig, H. 2017. Local adaptation of a national digital soil map for use in precision agriculture. In J A Taylor, D Cammarano, A Prashar, A Hamilton (Eds.) Proceedings of the 11th European Conference on Precision Agriculture. Advances in Animal Biosciences 8 430-432.

Piikki, K. and Söderström, M. 2018a. Package 'mapsRinteractive'. Available online: https://cran.r-project.org/web/packages/mapsRinteractive/mapsRinteractive.pdf (Last accessed December 2018).

Piikki, K. and Söderström, M. 2018b. Digital soil mapping of arable land in Sweden–Validation of performance at multiple scales. Geoderma in press, doi:10.1016/j.geoderma.2017.10.049

R Core Team 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org.

Söderström, M., Sohlenius, G., Rodhe, L. and Piikki, K. 2016. Adaptation of regional digital soil mapping for precision agriculture. Precision Agriculture 17 588-607.

Söderström, M., Piikki, K. and Cordingley, J. 2017a. Improved usefulness of continental soil databases for agricultural management through local adaptation. South African Journal of Plant and Soil 34 35-45.

Söderström, M., Piikki, K., Stenberg, M, Stadig, H., Martinsson, J. 2017b. Predicting nitrogen uptake in winter wheat by combining proximal crop measurements with Sentinel-2 and DMC satellite images in a decision support system for farmers. ACTA Agriculturae Scandinavica, Section B–Soil & Plant Science 67 637-650.

Near-real time winter wheat N uptake from a combination of proximal and remote optical measurements: how to refine Sentinel-2 satellite images for use in a precision agriculture decision support system

S. Wolters[1*], M. Söderström[1], K. Piikki[1] and M. Stenberg[2]
[1]*Department of Soil & Environment, Swedish University of Agricultural Sciences (SLU), Skara, Sweden; sandra.wolters@slu.se*
[2]*Swedish Board of Agriculture, Skara, Sweden*

Abstract

The availability of satellite data has facilitated the use of such data in decision support systems (DSS) for precision agriculture. The aim of the present study was to develop and evaluate prediction models for N-uptake in winter wheat (*Triticum aestivum* L.) in Sweden. Ground control measurements were carried out in winter wheat fields on farms using a handheld Yara N-sensor. Sentinel-2 satellite data were extracted for the point locations of the N-uptake measurements. Winter wheat N-uptake could be predicted with a mean absolute error of 10-12 kg N/ha which should be acceptable for implementation in a DSS.

Keywords: nitrogen fertilisation, decision support system, Sentinel-2, N-uptake, variable rate application

Introduction

A continuously growing world population and a changing global perspective on agriculture call for increasing resource use efficiency to meet societal profitability and sustainability demands. By measuring and responding to spatial and temporal variability within fields, efficiency in resource usage can be increased by variable rate practice. Optimal nitrogen (N) fertilisation has the potential to increase nutrient use efficiency and minimize leaching and runoff of N by reduction of fertiliser input in places where enough N is already present for plant uptake (Delin and Stenberg, 2014).
To adapt N supply to crop requirements, in-season N fertilisation is often done two or three times between the developmental stages of tillering and booting. The fine-tuning of fertilisation in winter wheat is of particular importance due to, e.g. the relationship between grain protein content and quality of the end product (human consumption in flour-based staple foods, like bread).
An important challenge is the transfer of optical crop canopy measurements into meaningful information, such as nitrogen content in the above-ground biomass (N-uptake) that can be used in practice by farmers and advisors. Currently, the use of decision support systems (DSS) appears to be a good way to bring the resulting N-uptake information to farmers as decision support for production of variable-rate application (VRA) maps for N fertilisation. These maps can be directly used in the field by use of navigation systems and electronically controllable fertiliser spreaders. A commonly used free DSS in Sweden, developed for this purpose, is CropSAT (Söderström *et al.*, 2017). In some other examples of DSS, vegetation index maps calculated from satellite data are available to facilitate VRA of N site-specifically (Whelan and Taylor, 2013).
The launch of the Sentinel-2 satellites (ESA, EU) has opened up new possibilities to access detailed and timely data for mapping of crop growth within-fields. Prior to this study, possibilities to map N uptake by combining proximal crop measurements with the modified soil-adjusted vegetation index (MSAVI2) index (Qi *et al.*, 1994) derived from a mix of Sentinel-2 and DMC (the UK-DMC-2 satellites, DMCii Ltd, Guildford, UK) satellite images in a DSS for farmers was assessed by Söderström *et al.* (2017). The present study aimed to test how well a number of vegetation indices from Sentinel-2

DOI 10.3920/978-90-8686-888-9_123, © Wageningen Academic Publishers 2019

data can predict N uptake in winter wheat, when calibrated by proximal sensor measurements of N uptake, and validated in a procedure resembling an implementation in a DSS. The study was based on an extensive dataset consisting of more than 700 individual ground measurements retrieved over a period of 3 years.

Materials and methods

Reference data was collected using an optical crop canopy sensor (a handheld version of the Yara N-Sensor, Yara, Oslo, Norway) on a weekly basis by the Swedish Board of Agriculture (Jordbruksverket, Jönköping, Sweden). This was done for a period of three years (2016-2018) at point locations in winter wheat fields on 40-50 farms per year during the period of Zadoks growth stage 22-53 (Zadoks *et al.*, 1974). The farms covered the four main wheat growing districts in Sweden (Figure 1). Fourteen winter wheat cultivars were measured: 'Brons', 'Elvis', 'Festival', 'Frontal', 'Hereford', 'Julius', 'Linus', 'Mariboss', 'Norin', 'Olivin', 'Praktik', 'Reform', 'Stava' and 'Torp'. More than 700 measurements were collected, and the built-in N uptake calibration of the proximal sensor (as is used on the tractor-borne Yara N-Sensors; Reusch, 2003) was used as ground-truth data. Each proximal N uptake measurement was the average of four registrations obtained in four directions, representing an area of approximately 3 m × 3 m. The Sentinel-2 satellites have a revisit time of 5 days at the equator. In the higher latitude regions, where the present study was carried out images are available approximately every 2-3 days. Twelve distinct bands in the range 400-2200 nm are available from Sentinel-2 satellites (Table 1). Spatial resolution varies between 10 and 20 m for ten bands in visible to the short-wave infrared region. For this study, a little more than 80 so called L1C images (containing radiometrically and geometrically corrected top-of-atmosphere reflectance) were used. No additional pre-processing was done on the images other than a manual judgement of the image quality in terms of occurrence of clouds, cloud shadows and haze. The reason for using L1C images in this study was that the DSS CropSAT uses this type of images as input. The Sentinel-2 images were downloaded manually from the Copernicus Open Access Hub (https://scihub.copernicus.eu;

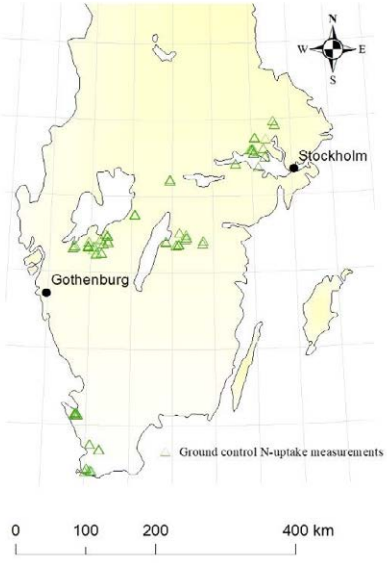


Figure 1. The study area and the locations of ground truth N-uptake measurements over the south parts of Sweden.

Table 1. Sentinel bands with according spectral resolution, central wavelength and bandwidth for Sentinel sensor A (S2A) and central wavelengths and bandwidth for sentinel sensor B (S2B) respectively (Sentinel-2, 2018).

Band	Spatial resolution (m)	S2A & bandwidth (nm)		S2B& bandwidth (nm)	
Band 1 – Coastal aerosol	60	443.9	27	442.3	45
Band 2 – Blue	10	496.6	98	492.1	98
Band 3 – Green	10	560.0	45	559	46
Band 4 – Red	10	664.5	38	665	39
Band 5 – Vegetation red edge	20	703.9	19	703.8	20
Band 6 – Vegetation red edge	20	740.2	18	739.1	18
Band 7 – Vegetation red edge	20	782.5	28	779.7	28
Band 8 – NIR	10	835.1	145	833	133
Band 8A – Narrow NIR	20	864.8	33	864	32
Band 9 – Water vapour	60	945	26	943.2	27
Band 10 – SWIR – Cirrus	60	1373.5	75	1376.9	76
Band 11 – SWIR	20	1613.7	143	1610.4	141
Band 12 – SWIR	10	2202.4	242	2185.7	238

ESA, EU). After downloading, images were again visually inspected for possible haze and clouds with help of a composite image constructed from bands 3, 4 and 8. Clear images were paired (if within ±3 days of acquisition) with the proximal sensor measurements and data were extracted.
Five different vegetation indices were calculated from the Sentinel-2 data. The NDVI (normalized difference vegetation index; Rouse, 1974), MSAVI2, green-red vegetation index (GRVI) (Mohtoka *et al.*, 2010), the leaf chlorophyll index (CI), which was calculated for two different combinations of bands and therefore used as separate indices (Gitelson *et al.*, 2003).
The indices were calculated by Equations 1-5, where ρ is reflectance and the subscript indicates the Sentinel-2 band number:

$$NDVI = \frac{\rho_8 - \rho_4}{\rho_8 + \rho_4} \quad (1)$$

$$MSAVI2 = \frac{1}{2}\left[(2 \times \rho_8 + 1)\sqrt{[(2 \times \rho_8 + 1)^2 - 8 \times (\rho_8 - \rho_4)]}\right] \quad (2)$$

$$GRVI = \frac{\rho_3 - \rho_4}{\rho_3 + \rho_4} \quad (3)$$

$$\text{CI_1= } CI_{red-edge} = \left({}^{\rho_7}/_{\rho_6}\right) - 1 \quad (4)$$

$$\text{CI_2= } CI_{red-edge} = \left({}^{\rho_7}/_{\rho_5}\right) - 1 \quad (5)$$

N uptake data from the proximal sensor measurements were compared with the vegetation index values. To investigate the impact of cultivar on the correlations, the correlation was done for the whole dataset and a subset of the data for solely cultivar Julius, which is the most commonly grown winter wheat cultivar in Sweden.
Linear regression models for prediction of N uptake were parameterised between N uptake and the vegetation index with the strongest correlation to N uptake. To assess the predictive ability of the model, the data was split in an iterative procedure. This procedure aimed at mimicking a real-

world application in which calibration data were only available from other years and earlier in the current season, and only from other locations. This procedure was repeated for all measurements to achieve a full set of predicted and observed values. This was again done for all cultivars, and separately for the cultivar Julius only. All growth stages and a subset consisting of DC 31-45 have also been tested. In summary, the four data subsets were: (1) all cultivars and all growth stages; (2) Cultivar Julius and all growth stages; (3) all cultivars and growth stages DC 31-45, and; (4) Cultivar Julius and growth stages DC 31-45. Validation measures were calculated to determine the fit of the model in comparison with the actual data. The Nash-Sutcliffe modelling efficiency (E) measures the relative improvement in accuracy over the mean of the calibration dataset. The mean error (ME) is a measure of prediction bias. The mean absolute error (MAE) is the average of the absolute prediction errors. The last validation measure is the coefficient of determination (R^2) of a linear regression line between predicted and observed data.

To quickly determine which military grid reference system files were required, the Python programming language with the MGRS package was used (Python software foundation, Wilmington, Delaware, USA). Data was stored in a SQLserver database (Microsoft, Redmond, Washington, USA) and the analysis was done with help of the R programming language (R core team, 2018). ArcMap (ESRI, Redlands, CA, USA) was used for visual judgement of image quality.

Results

In Figure 2, the plotted results show a relatively strong correlation between N-uptake and all of the indices. The strongest correlation was between N uptake and CI_1. NDVI and MSAVI2 showed a tendency of a non-linear relationship between the two variables, likely as an effect of saturation. In this figure, it is also easy to see the larger variance that was found for the GRVI. The difference between the GRVI and the other indices was smaller when the data subset, with only the cultivar Julius, was considered.

The results from the cross-validation of N uptake modelling (Figure 3) show a relatively good fit for all four datasets. The R^2 values were in all cases >0.7, which indicate relatively strong linear relationships between the predicted and the observed values. In this case, the dataset with only the cultivar Julius had the highest value for R^2. Validation statistics, are presented in Table 2. The E value shows the relative improvement in accuracy of the predicted N uptake over the mean of the observed N uptake dataset, and the present E values (E>0.7) confirm good model performance. In terms of E, it was better to use cultivar specific models than models parameterized on a dataset including all 14 cultivars. The MAE was in all datasets >10 kg N/ha. ME is for all datasets close to zero, which indicated only small prediction biases.

Discussion

A translation of N uptake to fertiliser recommendations was not in the scope of the present paper. However, as new prediction models can continuously be parameterized in a satellite-based DSS, this can be linked to fertiliser recommendations quite easily. Interesting to note is that the relatively good predictions of N uptake obtained here were done on L1C-corrected images (that are currently used in the DSS CropSAT), i.e. portraying top-of-atmosphere reflectance. One could even expect improved results if L2A images had been used (bottom-of-atmosphere reflectance), but that remains to be investigated in coming research.

The positional accuracy of the ground-truth measurements were sometimes judged somewhat uncertain, which adds additional difficulties when paired with satellite images. It is recommended to continue refinement of the workflow that has been initiated in this paper, moving towards variable N-rate maps.

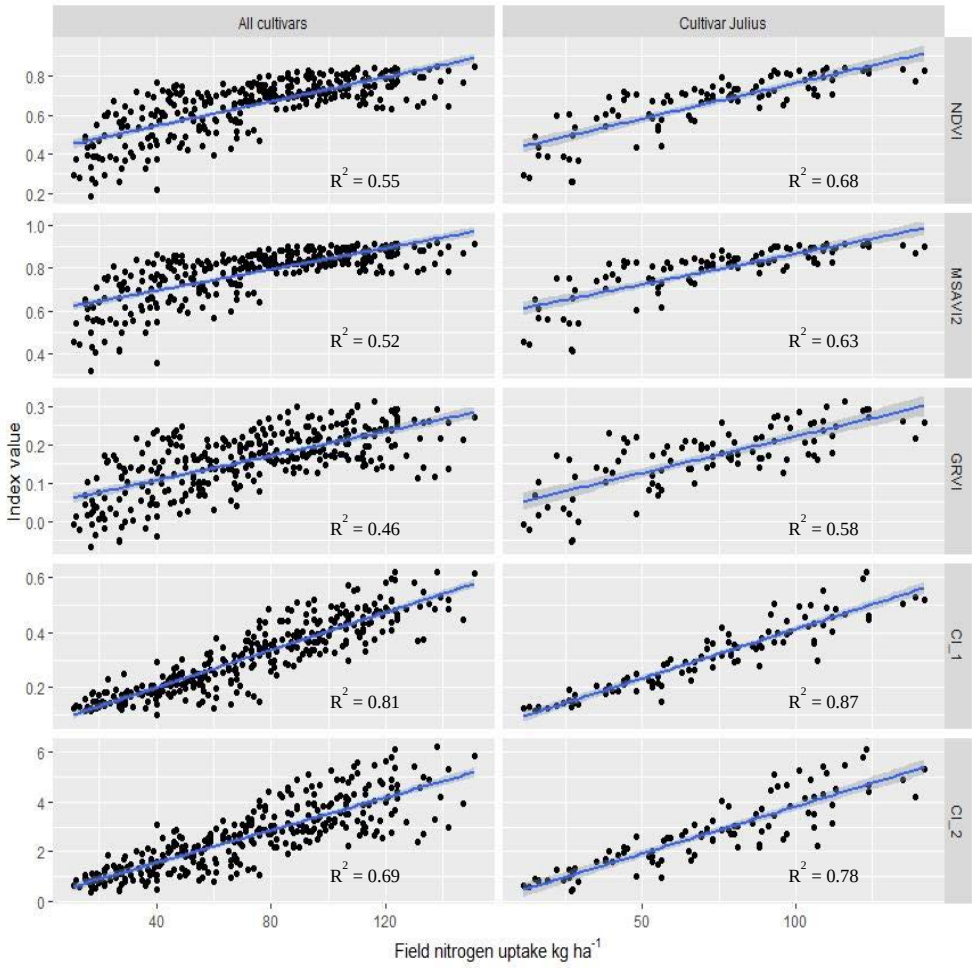


Figure 2. Satellite-based vegetation indices plotted against proximal sensor measured N uptake. The NDVI = normalized difference vegetation index; MSAVI2 = modified soil adjusted vegetation index, GRVI = green-red vegetation index; CI = leaf chlorophyll index. CI was calculated for two different combinations of bands and therefore used as separate indices (CI_1 and CI_2).

Information from optical satellite sensing requires relatively large fields. Lobate or pinnate field shapes, and structures in the field can impinge on this use of satellite data, due to edge effects from deviating vegetation next to the field, as well as the abundance of non-crop elements, such as wind mills. This means that a combination with other methods of N-uptake determination will still be required. Even though maybe not all fields are suitable for satellite N-uptake determination, this will still be the case for many wheat fields in Sweden and elsewhere and can therefore save considerable effort for farmers.

The use of optical remote sensing data is challenging in temperate regions due to substantial cloud coverage and haze. On the other hand, a considerable number of images can potentially be retrieved because of the short revisiting times of the Sentinel-2 satellite constellation at high latitudes.

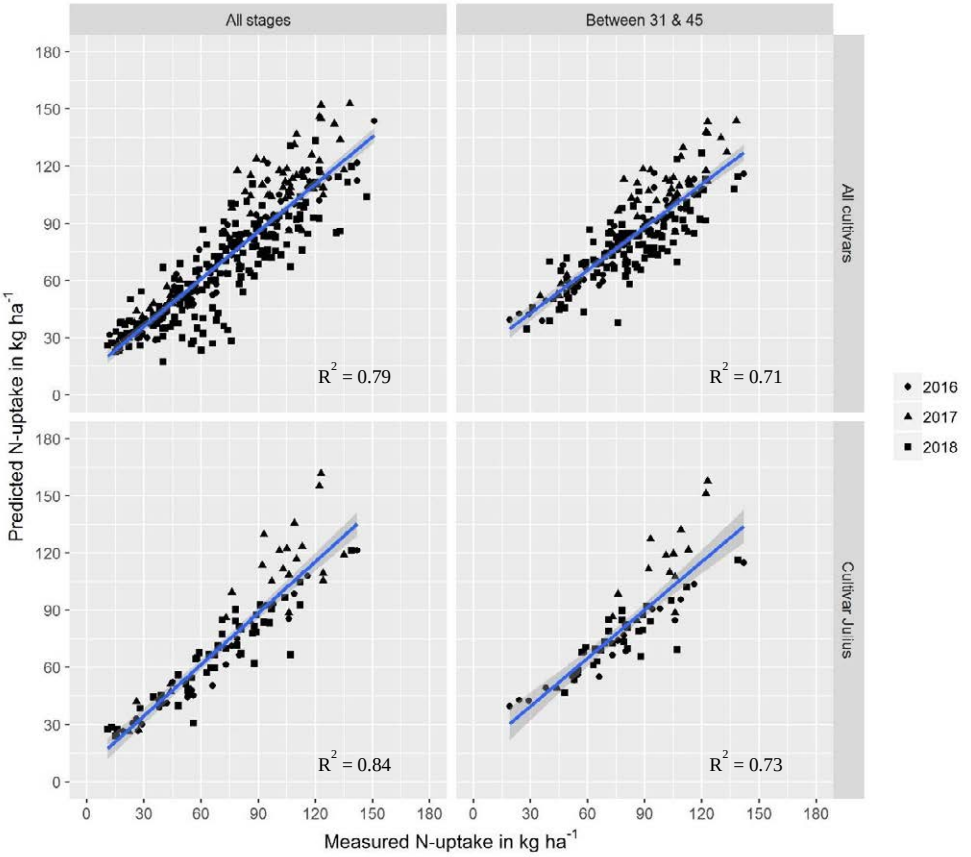


Figure 3. Results from cross-validation – predicted nitrogen (N) uptake plotted against actual N uptake. The four subplots show results from cross-validations for the complete dataset (all cultivars and all growth stages) or different subsets of data (only the most common cultivar (Julius) and/or for a range of growth stages (31-45)). Linear models with 95% confidence intervals were fitted and R-square values are included.

Table 2. Validation statistics of prediction models for N uptake. E = modelling efficiency, ME = mean error, MAE = mean absolute error, Cv. = cultivar, DC = growth stage, n = number of observations.

Statistic	All cultivars All DCs	Cv. Julius All DCs	All cultivars DC 31-45	Cv. Julius DC 31-45
E	0.79	0.83	0.70	0.71
MAE (kg/ha)	12	10	11	11
ME (kg/ha)	1.3	-0.3	0.5	-1.7
R^2	0.79	0.84	0.71	0.73
n	320	97	193	65

Conclusions

Empirical relationships between N uptake in Swedish winter wheat fields and Sentinel-2 data on the same locations were tested. The satellite data were expressed in five commonly used vegetation indices. The chlorophyll index based on bands 6 and 7 showed the strongest correlation with N uptake (R^2=0.81 for all cultivars and R^2=0.87 in a subset including only the most common cultivar: Julius). With the tested strategy, near-real time winter wheat N uptake can be predicted from a combination of proximal and remote optical measurements (10≤MAE<12; kg/ha). The MAE was smaller for the cultivar-specific model but not obviously reduced by restraining the dataset to certain growth stages (DC 31-45). Because Sentinel-2 data is freely available and can be easily accessed via web-request, this has proven to be a promising data source for use in web decision support systems, that can provide the required VRA maps to farmers for the application of N.

Acknowledgements

This project was funded by VGR/SLU (contract: RUN 2018-00141) together with Dataväxt AB, Sweden. Thanks are also due to the Swedish Board of Agriculture for the georeferenced data on N uptake obtained by handheld sensor measurements.

References

Delin, S. and Stenberg, M. 2014. Effect of nitrogen fertilization on nitrate leaching in relation to grain yield response on loamy sand in Sweden. European Journal of Agronomy. 52:291-296.

Gitelson, A.A, Gritz, Y. and Merzlyak, M.N. 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. Journal of Plant Physiology. 160:271-282.

Motohka, T., Nasahara, K., Oguma, H. and Tsuchida, S. 2010. Applicability of Green-Red Vegetation Index for Remote Sensing of Vegetation Phenology. Remote Sensing, 2, 2369-2387.

Qi, J., Chehbouni, A., Huete, A.R. and Kerr, Y.H. 1994. Modified Soil Adjusted Vegetation Index (MSAVI). Remote Sensing of Environment 48:119-126.

Reusch, S. 2003. Optimisation of oblique-view remote measurement of crop N-uptake under changing irradiance conditions. In: Stafford, J. and Werner, A. (Eds.). Precision Agriculture. Proceedings of the 4th European Conference on Precision Agriculture. Wageningen, the Netherlands: Wageningen Academic Publishers. pp. 573-578.

R Core Team. 2018. R: A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. Available from: http://www.R-project.org

Rouse, J.W, Haas, R.H., Scheel, J.A. and Deering, D.W. 1974. Monitoring Vegetation Systems in the Great Plains with ERTS. Proceedings, 3rd Earth Resource Technology Satellite (ERTS) Symposium, vol. 1, pp. 48-62.

Söderström, M., Piikki, K., Stenberg, M, Stadig, H. and Martinsson, J. 2017. Predicting nitrogen uptake in winter wheat by combining proximal crop measurements with Sentinel-2 and DMC satellite images in a decision support system for farmers. ACTA Acta Agriculturae Scandinavica, Section B – Soil & Plant Science. 67, 637-650.

Whelan, B. and Taylor, J. 2013 Precision agriculture for grain production systems. Clayton, Australia: CSIRO publishing.

Zadok, J.C., Chang, T.T. and Konzak, C.F. 1974. A decimal code for the growth stages of cereals. Weed Research 14:415-421.

Section 20
Data mining for precision agriculture

Maize yield prediction based on artificial intelligence using spatio-temporal data

A. Nyéki[1*#], C. Kerepesi[2#], B. Daróczy[2], A. Benczúr[2], G. Milics[1], A.J. Kovács[1] and M. Neményi[1]
[1]*Széchenyi István University, Faculty of Agricultural and Food Sciences, Department of Biosystems and Food Engineering, Vár square 2, 9200 Mosonmagyaróvár, Hungary; nyeki.aniko@sze.hu*
[2]*Institute for Computer Science and Control, Hungarian Academy of Sciences, Kende Str. 13-17, 1111 Budapest, Hungary*
[#]*Joint first authors*

Abstract

The aim of this study was to predict maize yield by artificial intelligence using spatio-temporal training data. Counter-propagation artificial neural networks (CP-ANNs), XY-fused networks (XY-Fs), supervised Kohonen networks (SKNs), extreme gradient boosting (XGBoost) and support-vector machine (SVM) were used for predicting maize yield in 5 vegetation periods. Input variables for modelling were: soil parameters (pH, P_2O_5, K_2O, Zn, Clay content, EC*a*, draught force, Cone index), and micro-relief averages and meteorological parameters for the 63 treatment units. The best performing method (XGBoost) attained 92.1 and 95.3% of accuracy on the training and the test set.

Keywords: maize yield prediction, machine learning, XGBoost, soil variables

Introduction

The current unfavourable effects of agriculture on the biosphere cannot be reduced by the knowledge of traditional experiments (Longchamps *et al.*, 2018). The gap between environmental adverse impacts and the slowly expanding knowledge base is growing.
Nyéki *et al.* (2013, 2017) concluded that increasing input data for plant physiological models can increase accuracy of prediction to only a limited extent. Consequently, there has to be a change in scientific methods, and the potential of big data needs to be better exploited than before, resulting in artificial intelligence (AI).
AI is used by more and more researchers to model a wide range of tasks in agriculture. There are a large numbers of studies in which the relationship between soil properties and yield has been analyzed (Irmak *et al.*, 2006; Miao *et al.*, 2006; Mike-Hegedűs, 2006; Dahikar and Rode, 2014; Pantazi *et al.*, 2016). A growing number of articles are presented in which climate effects are also analyzed by AI (Elgaali and Garcia, 2004; Iizumi *et al.*, 2018).
The gradient boosting machine (GBM) is a very effective state-of-the-art machine learning method (ML) (Friedman, 2001; Chen and Guestrin, 2016). For example, rice yield was forecasted by the XGBoost prediction model from weather information, cultivation data and location information (Maeda *et al.*, 2018). Its prediction accuracy was 74.4%.
Laacouri *et al.* (2018) assessed the corn N status comparing ML and vegetation indices. Hyperspectral images were collected by hexacopter UAV platform. Eight ML algorithms were compared for their accuracy, among other gradient boosting models. The authors concluded that hyperspectral imagery combined with ML improved the assessment of corn N stress status at V5 growth stage (Abendroth *et al.*, 2011), and achieved more than 90% classification accuracy when the entire spectrum was mined. GBM achieved an overall classification accuracy of 89%. SVM, logistic regression (LR), Multi-Layer Perceptron (MLP) and GBM showed promising results.

DOI 10.3920/978-90-8686-888-9_124, © Wageningen Academic Publishers 2019

Folberth *et al.* (2019) explored two ML approaches: XGBoost and random forests. ML was trained on global scale maize simulation of a GGCM (Global Gridded Crop Model). The method provides very high accuracy (R^2>0.96) for predictions of maize yields, evapotranspiration and crop available water. Khanal *et al.* (2018) integrated high resolution remotely sensed data and ML techniques for spatial prediction of soil properties and corn yield. High spatial resolution (<1 m) was used for the capturing of soil properties (SOM, CEC, K, Mg, and pH) and corn yield. The accuracy of seven statistical methods (LM – linear regression, RF – random forest, SVM with linear and radial kernel function, GBM, NN – neural network, and CUB – cubist) was compared for their ability to predict soil properties and corn yield. For pH and corn yield prediction gradient boosting machine and RF model performed better than other models.

These analysis show the XGBoost ML method seems to be beneficial method. However, XGBoost has not been applied for spatio-temporal maize yield prediction.

The aim of the study was to increase the accuracy of maize yield prediction. The above literature methods confirm that the XGBoost gives more accurate yield prediction than ANNs. However, further studies are needed to prove the advantages of XGBoost method. Among these, spatio-temporal studies have to be highlighted, as conclusions can be drawn from several years of experiments in agronomy.

Materials and methods

Study site

Field experiments were carried out in a 23.52 ha experimental field (Figure 1) belonging to Széchenyi István University in the vicinity of Mosonmagyaróvár, Hungary [N47°54'20.00"; E17°15'10.00"]. The experimental field is an alluvial plain of the Leitha River – on which precision agriculture has been applied since 2001.

Crop and soil parameters affecting maize yield

NDVI was calculated from satellite data by Landsat on July 26th, 2006; on July 5th; 2010; on July 29th, 2013 and on August 9th, 2017. The images were collected under high quality atmospheric conditions during the maize flowering period. The soil textures according to USDA indicated the soil types of loam, silty loam and sandy loam.

Instrumentation

In order to find the correct position, a Trimble AgGPS system (Trimble Ltd., USA) was used. The relative elevation (m) was determined by differential global navigation satellite system (DGNSS, Sokkia Radian-IS). Yield measurement was carried out by AGROCOM (ACT260) yield measurement system mounted on a Deutz Fahr grain harvester until 2010. From 2011, an AgLeader (InCommand™ 1200) yield mapping system was used which was mounted on a CLAAS Medion 340 grain harvester. Continuous soil draft measurements were carried out by a self-developed system (Neményi *et al.*, 2006). Site-specific penetrometer measurement was carried out by means of 3T system (Szőllősi, 2003). The soil EC_a was measured by a Veris Soil EC-3100 (Salina, KS, USA) instrument.

Feature table for supervised machine learning

Before supervised machine learning, a feature table was created from the measured parameters. The feature table consisted of 315 (63*5) rows for each treatment unit and year, and also consisted of 48 columns for each parameter (11 of soil parameters, 35 of meteorology parameters and one crop parameter) and the maize yield. In summary, four types of parameters were used: (1) yearly changing soil parameters (measured in 5 years: pH, P_2O_5, K_2O, Zn; measured in 3 years: EC_a in two layers, Veris N3 and N4; measured in 1 year: Cone Index (MPa), pH (water), draught force (kN)); (2) crop parameter (NDVI, measured in 4 years); (3) non-yearly-changing soil parameters (clay content (%),

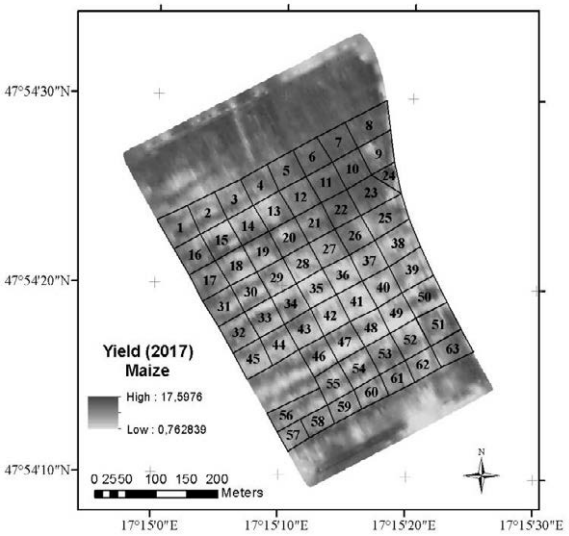


Figure 1. Site-specific maize yield (2017) in study site with 63 treatment units.

relative elevation (m) measured in one year but used in all years); (4) 35 of meteorology features (not changing by treatment units: used five types in each month from April (IV.) to October (X.), for maize vegetation seasons: sum of precipitation, average temperature, average relative precipitation, vaporization, aridity index).

Cross-validation and train-test split

Before training, the feature table was split into two parts, a training set of 50 and a testing set of 13 randomly selected treatment units. Since, five years of measurements were used, 250 training and 65 testing samples were obtained. The best parameters of the machine learning models were selected by 5-fold cross-validation over the training set. The final models were trained on the entire training set and tested on the independent testing set.

Binary classification

Maize yield values varied between 2.46 t/ha and 15.05 t/ha, and the upper limit of the first third of the lowest values of yield was 8.42 t/ha. The maize yield values were divided into two classes: low yield (≤8.42 t/ha) and medium-high yield (>8.42 t/ha). Several types of state-of-the-art supervised machine learning methods were used for classification: counter-propagation artificial neural networks (CP-ANNs), XY-fused networks (XY-Fs), supervised Kohonen networks (SKNs) extreme gradient boosting (XGBoost) and support-vector machine (SVM). CP-ANNs, SKNs and XY-Fs are supervised neural networks derived from hierarchical self-organizing maps (SOMs) (Ballabio *et al.*, 2012), and used for wheat yield prediction in a recent study (Pantazi *et al.*, 2016). XGBoost is an efficient implementation of the gradient boosting machine (Friedman, 2001). XGBoost was applied successfully in a wide variety of classification problems (Chen and Guestrin, 2016).

Evaluation measures for binary classification

Sensitivity, specificity, accuracy and ROC AUC (area under the receiver operating characteristic curve) were calculated to evaluate the classification models. Sensitivity means the proportion of actual positives (low yield samples) and specificity means the proportion of actual negatives

(middle-high yield samples). Accuracy is the proportion of the total number of samples that are correctly identified. ROC curve is defined by the point pairs of true positive rates (sensitivity) and false positive rates (1-specificity) at different threshold settings.

Results and discussion

Three neural networks models (CP-ANN, XY-F, SKN) and a gradient boosting machine (XGBoost) were trained on a feature table of 47 measurements of 63 treatment units in five years, altogether 315 samples. Recall that 50 randomly chosen treatment units were used for training and 13 for testing. The statistics of the training and testing datasets are shown in Table 1. The distributions of features that influence maize yield are shown in Figure 2 and the yield distribution by year in Figure 3.

The distribution of maize yield is shown in Figure 3 in Boxplot.
Table 2 shows the results of 5-fold cross-validation to find the optimal parameters of CP-ANN, XY-F, SKN, XGBoost and SVM. Sensitivity, specificity, accuracy and ROC AUC were calculated to evaluate the classification models. Mean and standard deviation of each evaluation measure were generated by repeating the 5-fold cross-validation ten times. XGBoost attained the highest performance in 3 of the 4 evaluation measures on the calibration set using cross-validations: 92.1% accuracy, 87.9% sensitivity, 96.7% ROC AUC. Results of independent validation are shown in Table 3 for the same models. XGBoost attained the highest performance in 3 of the 4 evaluation measures on the validation set: 95.38% of accuracy, 91.3% of sensitivity, 97.62 of specificity.
Gradient boosted tree algorithms are capable of selecting the most important uncorrelated features by building small decision trees of some of the important features and refining the small models by adding new trees (Kerepesi *et al.*, 2018). The positive and negative influencing factors for maize yield were divided. NDVI was identified as the overall most important positive factor in all years. The following influencing parameters were K_2O and soil electrical conductivity measurements (Veris N3).

Table 1. Summary statistics of the selected soil and crop parameters and the maize yield for the training and test datasets. Min, minimum; Max, maximum; Mean, mean (average); SD, standard deviation; CV coefficient of variation (the ratio of the standard deviation to the mean).

	Years	Training dataset (n=250)					Test dataset (n=65)				
		Min	Max	Mean	SD	CV	Min	Max	Mean	SD	CV
Maize yield (t/ha)	5	3.25	15	9.39	2.62	0.28	2.46	15.05	9.23	2.84	0.31
pH (water)	1	7.62	7.82	7.75	0.04	0.01	7.7	7.93	7.78	0.06	0.01
pH (KCl)	5	7.09	7.83	7.39	0.14	0.02	7.14	7.71	7.4	0.14	0.02
P2O5 (mg/kg)	5	123	415	234.24	55.67	0.24	151	380	250.95	57.84	0.23
K2O (mg/kg)	5	22.6	518	153.05	102.06	0.67	60.1	400	148.6	91.4	0.62
Zn (mg/kg)	5	1.21	4.7	2.91	0.66	0.23	1.74	4.1	2.95	0.57	0.19
Clay content %	5	8.4	21	13.28	3.07	0.23	7.9	18.4	13.14	3.45	0.26
Draught force (kN)	1	1.52	4.72	3.1	0.72	0.23	1.9	6.12	3.4	1.31	0.38
Relative elevation (m)	5	122.18	123.6	122.74	0.34	0	121.96	123.11	122.62	0.36	0
NDVI	4	0.2	0.69	0.45	0.14	0.31	0.22	0.68	0.46	0.14	0.3
VERIS N3	3	5.5	39.12	13.37	6.58	0.49	5.51	26.6	12.82	6.18	0.48
VERIS N4	3	7.16	37.83	17.06	6.81	0.4	6.94	33.26	16.58	7.17	0.43
Cone index (Mpa)	1	30.27	58.74	44.88	6.85	0.15	22.47	58.38	43.88	10.55	0.24

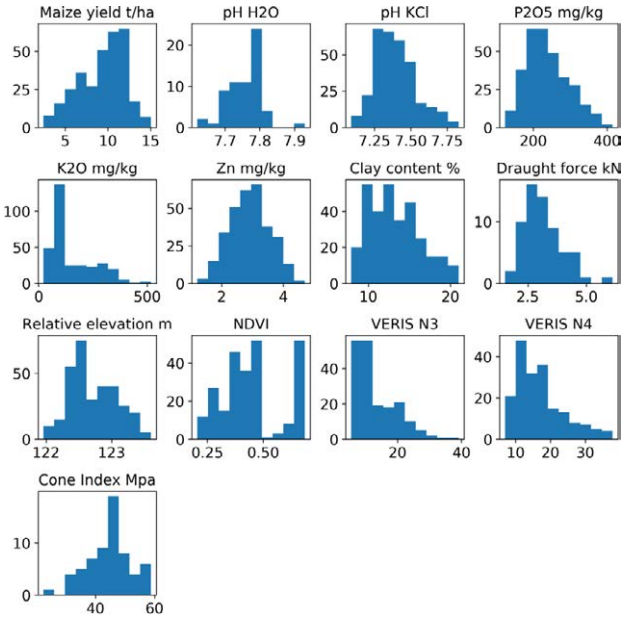


Figure 2. Value distribution of maize yield and its influencing factors. Horizontal-axis: value of the given variable. Vertical axis: frequency.

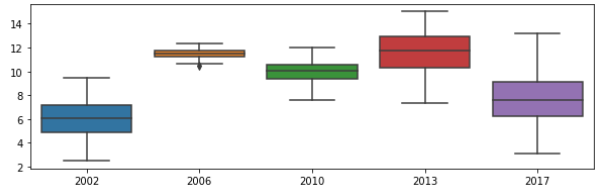


Figure 3. Yield distribution in each year considered in the study.

However, soil draught force had the second largest impact (negative correlation) on yield. The additional influencing variables were EC*a* (Veris N4), relative elevation, P_2O_5, Zn and pH (H_2O), in general, negative for yield in all years.

Conclusions

Five machine learning analysis methods, namely, CP-ANN, XY-F, SKN, XGBoost and SVM were presented. Machine learning to adapt site-specific data on soil and crop with two classifications of maize yield were applied. In this paper, site-specific maize yield in years was presented by the yield prediction model. In this study, CP-ANN, XY-F, SKN, XGBoost and SVM were used for defining the machine learning models between the maize yield and the influencing factors of yield. Meteorological parameters, soil dataset and crop parameter were used. The results showed that the best prediction method (accuracy) was the XGBoost. This prediction accuracy attained on the test set: 95.38% accuracy, 91.3% sensitivity and 97.62% specificity. The results showed that XGBoost was very effective in medium-high maize yield prediction.

Table 2. Results of 5-fold cross-validations (CV) on the training dataset for Counter-propagation artificial neural network (CP-ANN), XY-fused networks (XY-Fs), supervised Kohonen network (SKN), extreme gradient machine (XGBoost) and support-vector machine (SVM). Mean and standard deviation of accuracy, sensitivity, specificity and ROC AUC values were calculated by 10 experiments of 5-fold cross-validation. CP-ANN, XY-F and SKN were trained by 30 by 30 neurons and 50 epochs. XGBoost was trained with 10 trees and maximum depth 2. Linear kernel function with standard and L^2 normalization was used to train SVM.

	No. exps	Accuracy		Sensitivity		Specificity		ROC AUC	
		Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev
CP-ANN	10	0.9076	0.0097	0.8159	0.0272	0.9524	0.0097	NA	NA
XY-F	10	0.916	0.0078	0.8549	0.0272	0.9458	0.0082	NA	NA
SKN	10	0.9144	0.0091	0.8634	0.0138	0.9393	0.0115	NA	NA
XGBoost	10	0.921	0.007	0.879	0.016	0.941	0.007	0.967	0.008
SVM	10	0.9001	0.0065	0.8595	0.0205	0.9205	0.0079	0.9405	0.0060

Table 3. Results of the predictions on the independent test dataset for Counter-propagation artificial neural network (CP-ANN), XY-fused networks (XY-Fs), supervised Kohonen network (SKN), extreme gradient machine (XGBoost) and support-vector machine (SVM), trained on the whole training set with the same parameter settings as described in the legend of Table 2. Accuracy, sensitivity, specificity and ROC AUC values were calculated.

	Accuracy	Sensitivity	Specificity	ROC AUC
CP-ANN	0.8923	0.8261	0.9286	NA
XY-F	0.8615	0.7826	0.9048	NA
SKN	0.8923	0.8261	0.9286	NA
XGBoost	0.9538	0.9130	0.9762	0.9829
SVM	0.9538	0.9565	0.9524	0.9917

Acknowledgements

This work was supported by the EFOP-3.6.3-VEKOP-16-2017-00008 project. The project is co-financed by the European Union at the European Social Fund. The work was supported by the 2018-1.2.1-NKP-00008 "Exploring the Mathematical Foundations of Artificial Intelligence" project.

References

Abendroth L.J., Elmore R.W, Boyer M.J., and Marlay S.K. 2011. Corn growth and development. PMR 1009. Iowa State University Extension and Outreach, Ames, Iowa, USA.

Ballabio, D. and Vasighi, M. 2012. A Matlab toolbox for self organizing maps and supervised neural networks learning strategies. Chemometrics and intelligent laboratory systems 118: 24-32.

Chen, T. and Guestrin, C. 2016. XGBoost: A scalable tree boosting system. In Krisnapruham, B. and Shah, M. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, USA: ACM. pp 785-794.

Dahikar, S. S., Rode, S. V. 2014. Agricultural crop yield prediction using artificial neural network approach. International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering. 2, 683-686.

Elgaali, L., Garcia, L. 2004. Neural Network Modeling of Climate Change Impact on Irrigation Water Supplies in Arkansas River Basin, Hydrology Days, Colorado State University, CO, USA. http://hydrologydays.colostate.edu/Papers_2004/Elgali_paper.pdf

Friedman, J. H. 2001. Greedy function approximation: a gradient boosting machine. Annals of statistics. 29 (5), 1189-1232.

Iizumi, T., Shin, Y., Kim, W., Kim, M., Choi, J. 2018. Global crop yield forecasting using seasonal climate information from a multi-model ensemble. Climate Services. 11, 13-23.

Irmak, A., Jones, J. W., Batchelor, W. D., Irmak, S., Boote, K. J., Paz, J. O. 2006. Artificial neural network model as a data analysis tool in precision farming. Transactions of the ASABE. 49 (6), 2027-2037.

Kerepesi, Cs., Daróczy, B., Sturm, Á., Vellai, T., Benczúr, A. 2018. Prediction and characterization of human ageing-related proteins by using machine learning. Scientific Reports. 8, 1-13.

Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N., Shearer, S. 2018. Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. Computers and Electronics in Agriculture. 153, 213-225.

Laacouri, A., Nigon, T., Mulla, D., Yang, C. 2018. Case study comparing machine learning and vegetation indices for assessing corn nitrogen status in an agricultural field in Minnesota. https://www.ispag.org/proceedings/?action=abstract&id=4974&search=years (last accessed March 2019) Longchamps, L., Tremblay, N., Panneton, B. 2018. Observational studies in agriculture: paradigm shift required. https://www.ispag.org/proceedings/?action=abstract&id=5436&search=years (lasr accessed March 2019)Maeda, Y., Goyodani, T., Nishiuchi, S., Kita, E. 2018. Yield prediction of paddy rice with machine learning. In The 24th International Conference on Parallel and Distributed Processing Techniques and Applications CSREA PRESS. pp 361-365.

Miao, Y., Mulla, D. J., Robert, P. C. 2006. Identifying important factors influencing corn yield and grain quality variability using artificial neural networks. Precision Agriculture 7: 117-135.

Mike-Hegedűs F. 2006. Applying fuzzy logic and neural networks to database evaluation in precision agriculture. PhD thesis in Hungarian. University of West-Hungary, Mosonmagyaróvár.

Neményi, M., Mesterházi, P. Á. and Milics, G. 2006. An Application of Tillage force Mapping as a Cropping Management Tool. Biosystems Engineering. 94: 351-357.

Nyéki, A., Milics, G., Kovács, A. J., Neményi, M. 2013. Improving yield advisory models for precision agriculture with special regards to soil compaction in maize production. In: Stafford, J.V. (Ed.), Precision Agriculture '13. Proceedings of the 9th European Conference on Precision Agriculture. Wageningen, the Netherlands: Wageningen Academic Publishers. pp. 443-451.

Nyéki, A., Milics, G., Kovács, A. J., Neményi, M. 2017: Effect of soil compaction on cereal yield. Review. Cereal Research Communication. 45 (1) 1-22.

Pantazi, X. E., Moshou, D., Alexandridis, T., Whetton, R. L. and Mouaen, A. M. 2016. Wheat yield prediction using machine learning and advanced sensing techniques. Computers and Electronics in Agriculture 121 57-65.

Szőllősi, I. 2003. A 3T SYSTEM készülékkel mért penetrációs ellenállás és nedvességtartalom összefüggése vályog fizikai féleségű talajon. (Correlations between the penetration resistance registered with a 3T SYSTEM instrument and the moisture content of a soil with loam texture.) Agrokémiai és Talajtan. 52 (3-4) 263-274.

Keyword index

A

active
– contour 381
– optical sensors 525
actual map 149
adoption 851, 859
agricultural data 867
agroecology 597
agronomics 893
agronomy 969
anthocyanins 287
apparent electrical conductivity 91, 307
application maps 231
APSIM 531
artificial neural network 657
as-applied maps 53
automated
– irrigation 951
– weed recognition 231
automatic disease identification 141
autonomous
– navigation 767
– robot 331
auto-steering 807

B

banana 413
– crop 299
barley 29, 271
Bayesian Kriging 519
benefit 833
big data 83
biomass 165
biophysical variables 453
BLiSS 67
block cokriging 705
boom section control 53, 239
bootstrapped regression trees 713
breeding 381

C

canopy
– management 187
– temperature 249
capacitance 589
– probe 631
cauliflowers 783
Cercospora beticola 977
cereal breeding program 397
CERES-Beet 977
chessboard design 925
chlorophyll content 299
clustering 623
CNN 381
codes 99
collection rate 867
color 195
– scales 623
colorimeter 571
commercialization 843
competitiveness 597
computational imaging 665
consortium 851
constraint programming 173
convolutional neural networks 141, 713
corn 649, 657
correlation 217
cost 833
cotton 45, 695
crop
– coefficient 437
– damage 405
– growth model 29
– model 91, 133, 469, 883
– sensors 323
– structure 157
– water stress 263, 673, 687
crowdsourcing 735
CubeSats 493

D

2,4-D 209
3D 187
– modelling 165
– digital surface model (DSM) 397
– surface tree model 109
data
– filtering 901
– management 83
– processing 203
– usefulness 867

decision support 985, 993
– system 581, 673, 1001
deep learning 141, 639, 761, 791
deficit irrigation 125
demonstrated value 859
detection 279
differential management 631
digital soil map 993
digitization 83
disease detection 195, 225, 339
drone 365, 959
droplet application 209
drought 249
dry matter 631
durum wheat 23

E

early yield prediction 373
earth observation 469
economic
– value 817
economy 833
eddy covariance 485
EEFlux 431
efficiency 943
electrical conductivity 115
embodied knowledge 843
entry and exit 825
environment 357
error correction 901
evapotranspiration 421, 431, 485

F

factorial cokriging 705
farmer
– centric research 893, 917
– innovation group (FIG) 893
– needs 985
farm management 493, 867
feedback 951
fertilization 883
field-specific management 877
fleet management 775
flux towers 437
fruticulture 109
fungal diseases 255
Fusarium 339
fuzzy delineation 729

G

geographically-weighted regression 925
glufosinate-ammonium 209
glyphosate 209
Google earth engine 445
grain storage 721
grape
– bunch segmentation 791
– detection 373
– quality 181
– yield 181
grapevine downy mildew 141
green
– red vegetation index (GRVI) 37
– snap 405
grid sampling 525
ground-based 2D LiDAR 157
guidance 859

H

hazard rate 843
head blight 225
headland-turning 807
herbicide 279
– mixtures 231
heterogeneity 217
high-density olive 421
horticulture 115
HSV 225
humidity 217
hybrid topological map 767
hyperspectral
– camera 203
– data 705

I

image
– analysis 373, 389, 623
– classification 365
– processing 791
imaging
– spectroscopy 365
impact forces 909
in-field monitoring 287
information intensive 843
integrated weed management 231
intensive olive grove 477

inter-annual variability 29
ion selective electrode 571
'IPMwise' 231
irrigation 133, 421, 485, 649, 695
– prescription 657
– scheduling 649, 695
– uniformity 37
ISO 11783 775

K

kriging 743

L

LAI 469
– estimation 389
Landsat 263, 437
large-scale paddy field 751
laser scanner 115
laser triangulation 315
lateral move irrigation 37
leaf 279
– angle 271
– area 115, 331
– specific weed control 209
– wall area 157
– water potential 263
– wetness 217
licences 99
Linear Imaging Self Scanning 437
load cells 589
local
– effects 217
– Moran's I 721
lodging 969
low-cost 373
– D-GNSS 775

M

machine
– learning 365, 381, 389, 1011
– vision 589
maize 453, 531
– yield prediction 1011
malting barley 581
management
– cell 125
– zones 555, 729, 883
mapping 287
maps 959
market structure 825
measurement 157
– platform 203
melon 381
membership frequency 729
mergers and acquisitions 825
METRIC 431
midday stem water potential 125
missing plants 165
mixed models 299
modelling 45
model sampling 173
moisture sensor 951
morphology 271
multispectral 255, 339, 469
– images 605
– sensor 413
mycotoxin 721
– content correlation 225

N

national survey 851
NDVI 23, 347
nested factors 299
network depth 141
new sensing technologies 287
N fertilisation 563
nitrogen 525, 555, 925
– fertilisation 1001
– fertiliser 61, 293
– fertilization 323
– management 511
– nutrition index 293, 453
– requirement 61
– uptake 23
– use efficiency 531
NNI 453
non-invasive sensors 287
normalised difference vegetation index (NDVI) 173, 477, 901
number of clusters 67
N-uptake 1001

O

object-based image analysis (OBIA) 397
objective model 615

oilseed rape 61
old wheat varieties 461
olive orchard 421, 431
on-farm
– experimentation (OFE) 817, 893
– experiments 917
– precision experimentation 901
optical active sensor
– 347
optimal N splitting 563
optimum N application rate 877

P

parallel tracking system 807
parametric classification 195
pasture 623
PCA 563
permanent pastures 631
pH 539
phenotyping 381, 389
photogrammetry 331
pitch angles 909
Planet-labs 445
plant
– classification 639
– height 397
– phenotyping 249
– protection 985
– protection products 357
plant geometry 315
planting dates 493
plateau functions 519
point cloud 187, 331
precise plant protection 109
precision
– agriculture 833
– irrigation 263, 431
– pesticide application 357
– viticulture 165, 181, 735, 743
precision liming 539
prediction 597
– models 631
privacy 99
probability maps 615
process optimization 775
production
– gap 75
– potential 75
protein content 461, 581
proximal sensing 23, 195, 255, 539

Q

quality degradation index 347

R

radiometry 665
red edge 605
reflectance 293, 605
regulated deficit irrigation 951
remote
– sensing 23, 133, 307, 397, 461, 485, 639, 713, 883, 943, 969
– sensing and sensors 665
requirements 985
RGB
– D sensing 791
– images 389
risk prediction 969
robotics 783

S

sampling optimization 173
satellite 469
– image price 501
– images 477, 493, 581
– performance 501
scale 993
seeder dynamics 909
selective harvesting 783, 799
semantic
– identification 767
– segmentation 639
semivariogram 751
sensor dynamics 203
Sentinel-2 263, 307, 347, 437, 445, 461, 477, 1001
shoot tip index 935
simulation model 909
site-specific weed management 239
smart farming 833
smartphone application 935
soft robotics 799
software 959
soil
– electrical conductivity 681
– fertility 615
– mineral nitrogen 571
– pH 993

– sensors 649
– texture 91
– type 877
– variability 571
– variables 1011
– water
• depletion modeling 687
• tension 657, 943
soil maps 539
sowing dates 493
spatial
– analysis 299, 721
– distance filter 331
– scale 917
– variability 29
species mapping 639
spectral
– analysis 339
– modelling 485
spectrometer 203, 735
spectrometry 279
sprayer 239
spraying deposits 157
spray pattern 239
spring barley 293
stem water potential 445
stereo vision 271
structural equation modelling 751
structure-from-motion 315
sugar
– beet disease 977
– content 735
super-high density olive groves 357
super resolution 665
supervised classifiers 623
support change 705
survival analysis 843

T

technology 851
– adoption 825
temperature 67
temporal
– analysis 67
– and spatial variability 157
– variability 29
texture 195
time 843
– optimisation 807
tomatoes 799
topdressing 563
total soluble solids 287
tramline trials 893
transition areas 729
transparency 99
tree area index 157
trust 99
Tunisia 431

U

unmanned aerial systems 405
unmanned aerial vehicle (UAV) 37, 109, 149, 165, 323, 357, 373, 421, 511, 605, 959
– imagery 413
– sensing 255

V

variability 531, 943
variable
– application 53
– stiffness actuators 783, 799
variable rate 61
– application 149, 511, 993, 1001
– drip irrigation
• 125
– irrigation
• 649, 673, 681, 687
– nitrogen
• 519
– technology
• 859
variogram 721
vegetation
– index 453, 605
– indices 307, 623
vegetative cycle 347
Venus satellite 37
vigour map 149
vine
– architecture 187
– water status 935
vineyard 125
– variability 743
virtual environment 83
visible image 597
viticulture 133
Vitis vinifera 935
VRA phosphorus and nitrogen fertilizing 547

W

water 555
– deficit 943
– flow accumulation 91
– management 681
– needs 115
– stress 249
weed
– classification 761
– control 597
– maps 365
wheat 255, 271, 597
wind
– damage 405
– engineering 969
winter wheat 511, 547, 581, 751
within-field
– variability 501
– yield data 75
WOFOST 469

X

XGBoost 1011

Y

yield 61, 67, 461
– components 563
– estimation 173, 381
– forecasting 713
– mapping 589, 893
– monitor 91, 589
– monitor data 519
– prediction 307
– response functions 925
– variability 45
– variation 917

Z

Zea mays 531, 925
zone sampling 525

Author index

A

Abdelghafour, F. 195
Alchanatis, V. 279, 339, 665
Alonso-Ayuso, M. 323
Andales, A. 555
Andersen, P. 231
Andrade, M. 673, 695
Andrews, W.F.P. 783, 799
Andujar, D. 623
Anken, T. 511, 833
Apolo-Apolo, O.E. 249, 389
Argento, F. 511
Arnó, J. 307
Aru, F. 357
Auzeeri, T. 783

B

Bahat, I. 125, 445, 665
Baker, C. 969
Baragatti, M. 67
Basso, B. 29
Batchelor, W.D. 883, 977
Bauer, P. 673
Bchir, A. 431
Beaulieu, M. 141
Bebronne, R. 255
Beeri, O. 263, 437
Behmann, J. 203
Bellvert, J. 133
Beltarre, G. 563
Benczúr, A. 1011
Ben Dhiab, A. 431
Ben-Gal, A. 125, 445
Ben Meriem, F. 431
Berry, P. 917, 969
Berry, P.M. 61
Berton, A. 165, 373
Bishop, T. 713
Bishop, T.F.A. 45
Blackburn, A. 969
Bohnenkamp, D. 203
Bois, B. 217
Bojer, O.M. 231
Bondesan, L. 649, 657
Börjesson, T. 581
Boschetti, M. 453
Boulent, J. 141
Bourouah, M. 571
Bousetta, W. 431
Bouvry, A. 271
Braham, M. 431
Bramley, R. 743
Bramley, R.G.V. 83
Brandani, G. 461
Brand, O. 665
Brightman, D. 893
Brorsen, B.W. 519
Bruce, M.A. 23
Brunel, G. 935
Bullock, D.S. 817, 901, 925
Busch, S. 783
Buttafuoco, G. 705

C

Cadenat, V. 767
Caffarra, A. 217
Cambel, K. 225
Cammarano, D. 29, 883
Campillo, C. 477, 951
Campos, J. 149
Canavari, M. 833
Carli, G. 833
Carra, M. 157
Casadesús, J. 951
Casa, R. 365, 453
Castrignanò, A. 705
Castro, A.I. De 187
Chen, A. 37
Chenu, K. 493
Cheraiet, A. 157
Cillis, D. 453
Cinat, P. 165, 373
Ćirić, V. 91
Clarke, C. 893
Clarke, S. 917
Codis, S. 157
Cohen, Y. 83, 125, 445, 665
Conforti, M. 705
Cordero, E. 563
Crema, A. 453

D

Da Costa, J.-P. 195
Dafna, A. 381
Damianidis, D. 649, 657
Dandrifosse, S. 271
Daróczy, B. 1011
Dashuta, A. 381
De Castro, A.I. 397
Delpuech, X. 157
De Tejada, V.F. 783
Devy, M. 767
Diago, M.P. 287
Di Gennaro, S.F. 373
Dokoozlian, N. 743
Dong, R. 877
Douzals, J.P. 239
Dufour, J. 767
Dumont, B. 255, 271
Dunkerley, D. 493
Dyrmann, M. 231, 639, 761

E

Edan, Y. 279, 339, 381
Egea, G. 249, 389
Elharar, O. 37
Emmi, L. 767
Erickson, B. 859
Eriksen, J. 639
Erlund, P. 985
Escolà, A. 307
Etzioni, A. 339
Evert, F.K. van 91
Evett, S. 673

F

Fabbri, C. 461
Fajardo, M. 713
Falagas, A. 469
Fantozzi, P. 365
Farkhani, S. 639
Fernandez-Novales, J. 287
Filippi, P. 45, 713
Fontana, F. 29
Ford, C. 743
Foucher, S. 141
Fountas, S. 83, 115, 589, 833
Franzen, D. 525
Fuller, M.P. 783, 799

G

Gabriel, J.L. 323
Gad, S. 279, 339
Gallart, M. 149
Garcia-Cela, E. 721
Garin, G. 217
Gebbers, R. 539, 571
Gée, C. 597
Gennaro, S.F. Di 165
Geraudie, V. 735
Gerhards, R. 225
Gerlach, F. 539
Germain, C. 195
Gertsis, A. 357, 421
Gil, E. 149
Giora, D. 501
Gislum, R. 639, 959
Gobbo, S. 605
Gogumalla, P. 445
Gonçalves Trevisan, R. 901
Graeff, S. 977
Green, B.M. 799
Green, O. 231
Griepentrog, H.W. 331, 775, 807, 909
Griffin, T.W. 843
Grignani, C. 563
Grünzweig, J.M. 125
Guan, Y. 877
Guillemin, J.P. 239
Gullick, D. 969
Guo, H. 83
Gur, A. 381
Gutierrez, S. 287

H

Hackett, R. 293
Hansen, N.C. 687
Harary, I. 279
Hara, S. 985
Harfouche, A. 365
Hathcock, L.A. 405
Hatley, D. 969
Haymann, N. 485
Heiβ, A. 775
Helman, D. 445
Hobart, M. 109
Hobman, E. 867
Höglind, M. 623

Holland, J. 29
Hopkins, B.G. 687
Hørfarter, R. 53
Hübner, R. 807
Hunt, A.G. 531

I

Ingram, B. 721

J

Jalli, M. 985
Jensen, N.-P. 231
Jensen, N.P. 639
Jimenez, A.F. 649, 657
Jiménez-Brenes, F.M. 187
Johnstone, P.R. 531
Jones, G. 239
Jones, H. 729
Jørgensen, R.N. 231, 639, 761, 959
Joseph, G. 969

K

Kalantar, A. 381
Kamman, A. 331
Kaplan, G. 485
Karantzalos, K. 469
Karkee, M. 83
Karstoft, H. 639
Kauser, J. 547
Kayad, A. 181, 501
Kendall, S. 893
Kendall, S.L. 61
Kerepesi, C. 1011
Keresztes, B. 195
Kerry, R. 687, 721
Kersebaum, K.C. 571
Khosla, R. 555, 589
Kichler, J. 943
Kindred, D. 893, 917
Klapp, I. 381, 665
Klein, F.B. 783
Kling, C. 539
Kono, Y. 751
Kosover, E. 279
Koukiasas, N. 209
Kovács, A.J. 1011
Kramer, E. 539

L

Lacerda, L.N. 605
Lachia, N. 851
Lamour, J. 299, 413
Lange, C. 29
Larmure, A. 597
Laura, L. 365
Laurent, C. 67
Laursen, M.S. 639, 761, 959
Léchaudel, M. 299
Leemans, V. 255, 271
Le Moguédec, G. 299
Lenski, I. 445
Leroux, C. 75, 299, 729, 735
Levi Bar Shalom, A. 279
Liakos, V. 943
Liebisch, F. 511
Lienard, A. 157
Li, X. 519
Li, Y. 883
Llop, J. 149
Longchamps, L. 555, 589
López-Granados, F. 187
López-Herrera, J. 323
Lovat, L. 181
Lowenberg-DeBoer, J. 859
Lück, K. 539

M

Machleb, J. 225
Magan, N. 721
Mahlein, A.-K. 203
Maillot, T. 239
Mancini, M. 461
Mao, X. 799
Marani, R. 791
Marín, D. 187
Marinello, F. 83, 181, 421, 501
Marques da Silva, J. 347, 631
Marsal, J. 133
Martínez-Casasnovas, J.A. 83, 307
Martínez-Guanter, J. 249, 389
Martinez-Perez, J. 209
Martin, N.F. 901, 925
Martinsson, J. 993
Mata, M. 133
Matese, A. 165, 373
Matsui, T. 751

McFadden, J.R. 825
Medici, M. 833
Melyon-Delage, R. 217
Memic, E. 977
Mercatoris, B. 255, 271
Merienne, J. 597
Meron, M. 37
Metay, A. 67
Meyer, S. 539
Mey-tal, S. 263, 437
Miao, Y. 877, 883, 969
Michael, Y. 445
Michez, A. 255
Mieno, T. 817
Milella, A. 791
Milics, G. 547, 1011
Millán, S. 477, 951
Mimić, G. 91
Miniotti, E.F. 563
Miranda, C. 187
Mohamed, A. 799
Moñino, M.J. 951
Moral, F.J. 347, 477, 615, 631
Morari, F. 23, 357, 421, 649
Morata, G.T. 649, 657
Moretti, B. 563
Moretto, J. 23
Mortensen, A.K. 639, 959
Moura, D. 413
Mulla, D.J. 431, 877
Murdoch, A.J. 209
Murgia, T. 29
Murray, J. 969

N

Nacson, T. 279, 339
Nagel, A. 539
Nanni, U. 365
Napoli, M. 461
Naud, O. 157, 299, 413
Neményi, M. 1011
Netzer, Y. 125, 445
Nikander, J. 985
Nissinen, A.I. 985
Nutini, F. 453
Nyéki, A. 1011

O

Oger, B. 173
Orfanou, A. 943
Oria, I. 187
Orlandini, S. 461
Orlov-Levin, V. 37
Ortiz, B.V. 649, 657
O'Shaughnessy, S. 673, 695
Ostos-Garrido, F.J. 397
Oz, N. 665

P

Pagay, V. 743
Palombo, A. 365
Pancorbo, J.L. 323
Panneton, B. 589
Paoli, J.N. 239
Papini, S. 665
Paraforos, D.S. 83, 115, 775, 807, 909
Paris, C. 133
Park, E. 519
Pascucci, S. 365
Paulus, S. 315
Pavlou, D. 943
Pedersen, M.F. 833
Pedersen, S.M. 833
Peeters, A. 125
Pelta, R. 263, 437
Peña, J.M. 397
Pérez, J.M. 477
Pérez-Martín, E. 323
Pérez-Ruiz, M. 249, 389
Pesonen, L. 985
Peteinatos, G.G. 225
Petitti, A. 791
Pflanz, M. 109
Philipp, G. 539
Phillippi, E. 555
Pichon, L. 735, 851, 935
Pignatti, S. 365
Piikki, K. 993, 1001
Pilgrim, R.A. 209
Pistón, F. 397
Polese, R. 23
Porter, W. 943
Poulsen, H.V. 53
Prieto, H. 477
Prieto, M.H. 951
Prince Czarnecki, J.M. 405

Q

Quarto, F. 705
Quarto, R. 705
Quemada, M. 323

R

Rabatel, G. 413
Rançon, F. 195
Raz, J. 263
Rebollo, F.J. 477, 615
Rega, M. 217
Reina, G. 791
Reiser, D. 331
Requena, I. 783
Rihan, H. 783
Rocco, M. Di 365
Rodriguez, B.L. 783
Rodriguez-Moreno, F. 91
Rolley-Parnell, E.J. 799
Romani, M. 563
Romboli, Y. 165
Rondepierre, A. 783
Ronga, D. 29
Roques, S. 893, 917
Rozenstein, O. 485
Rueda-Ayala, V. 623
Ruehlmann, J. 539
Rydahl, P. 231

S

Sacco, D. 563
Sadeh, Y. 493
Samiappan, S. 405
Sams, B. 743
Sanchez, L. 743
Sanderson, J. 99
Sandovsky, T. 279, 339
Sanford, S. 209
Santesteban, L.G. 187
Santos, A.F. 605
Sartori, L. 83, 181, 501
Sauerwald, T. 783
Sawyer, A. 943
Scheibe, D. 539
Schimmelpfennig, D.E. 825
Schirrmann, M. 109
Scholasch, T. 67
Schröter, I. 539
Searle, B.P. 531
Serrano, J. 347, 631
Serrano, J.M. 615
Shahidian, S. 347, 631
Sharipov, G.M. 909
Sharp, J.M. 531
Shaw-Sutton, J. 799
Shilo, T. 263, 437
Shiratsuchi, L.S. 901
Silva, R.P. 605
Skovsen, S. 639, 761
Sochen, N. 665
Söderström, M. 581, 993, 1001
Sökefeld, M. 225
Somerville, G.J. 231
Soper, D. 969
Sørensen, C.G. 83
Sozzi, M. 181, 501
Stadig, H. 993
St-Charles, P.-L. 141
Stenberg, B. 83
Stenberg, M. 1001
Sterling, M. 969
Stoelen, M.F. 783, 799
Stone, K. 673
Storer, K. 61
Stougård, K. 53
Sudduth, K. 83
Sui, R. 681
Svedin, J.D. 687
Sylvester-Bradley, R. 893, 917
Szekeres, L. 547

T

Tagarakis, A.C. 91
Tanaka, T.S.T. 751
Tanny, J. 437, 485
Tardaguila, J. 287
Taylor, J. 29, 67, 75, 735, 935
Tenni, D. 563
Théau, J. 141
Thorsted, M.D. 53
Tisseyre, B. 67, 75, 83, 173, 299, 729, 735, 851, 935
Tofannin, A. 605
Tohidloo, G. 833
Tomasi, D. 181
Torres-Sánchez, J. 187, 397

Toscano, P. 373
Trevisan, R.G. 925
Trindall, J. 867
Tsiropoulos, Z. 833
Tsoulias, N. 115

U

Uribeetxebarría, A. 307

V

Vaessen, H.M. 91
Vallverdú, X. 133
Vandini, G. 453
Vázquez Arellano, M. 331
Vellidis, G. 83, 357, 421, 605, 649, 943
Venezia, A. 705
Vér, A. 547
Vergès, A. 157
Vermeulen, P. 255
Villette, S. 239
Vismara, P. 173
Vitti, C. 705
Vivas, A. 951
Vivoli, R. 461
Vogel, S. 539
Vories, E. 695
Vougioukas, S.G. 83

W

Wade, R. 61
Wallor, E. 571
Walter, A. 511
Wang, X. 877
Welham, S. 917
Whelan, B. 713
Whelan, B.M. 45
Whyatt, D. 969
Wilmot, A. 783
Wilson, L. 883
Wiseman, L. 99
Wolters, S. 581, 1001

Y

Yafin, P. 665
Yeager, E.A. 843

Z

Zancanaro, E. 421
Zha, H. 883
Zhang, A. 867
Zhou, P. 799
Zhou, Y. 799
Zhu, X. 493
Zito, S. 217
Zude-Sasse, M. 115