Session 6

Applications of Unmanned Aerial Systems

73. Machine learning for estimating leaf area index in almond orchards using high-resolution multispectral and point-cloud data

C.Minuesa[1], M. Quintanilla-Albornoz[1], A. Pelechá[1], J. Gené-Mola[1], L. Aparicio[2] and J.Bellvert[1]
[1]*Efficient Use of Water in Agriculture Program, Institute of Agrifood Research and Technology (IRTA), Fruitcentre, Parc AgroBiotech, 25003 Lleida, Spain; * cesar.minuesa@irta.cat*
[2]*Fruit Production Program, Institute of Agrifood Research and Technology (IRTA), Fruitcentre, Parc AgroBiotech, 25003 Lleida, Spain*

Abstract

In irrigation management, the leaf area index (LAI) is crucial to understanding the crop energy flux distribution. This study proposed an indirect method to estimate LAI by combining UAV-derived multispectral and point-cloud data with machine learning (ML) algorithms. Fractional intercepted photosynthetically active radiation (fiPAR) measurements from two almond orchards with varying ages and water regimes were used to develop models. Random forest (RF) and artificial neural network (ANN) were evaluated. Results showed ANN outperformed RF, achieving higher accuracy (R^2=0.94, RMSE=0.07 m²/m²) on the test dataset and better generalization across the orchards.

Keywords: almond trees, LAI, machine learning, UAV

Introduction

Monitoring the leaf area index (LAI) is essential for modelling key physiological processes in plant canopies. LAI quantifies the leaf surface area per unit ground area (m²/m²; Watson, 1947), providing insights into light interception and distribution within the canopy. This distribution affects photosynthetic activity, canopy temperature, and water vapor exchange, which are critical for estimating canopy transpiration. In the energy-balance model, LAI plays a central role in intercepting short- and long-wave radiation and in determining canopy aerodynamic resistance.

Current approaches for measuring LAI include both direct and indirect methods (Gower *et al.*, 1999), each with specific applications and limitations. Direct techniques, such as destructive sampling, provide precise LAI measurements that are essential for calibrating indirect assessments but are labour-intensive and impractical for extensive or temporal studies. In contrast, indirect methods, such as those using LAI meters (e.g., AccuPAR, LAI-2000) and hemispherical photography, infer LAI by assessing sunlight occlusion, allowing for efficient and automated data collection. However, while indirect techniques facilitate broader spatial sampling, they remain constrained by the effort required to cover large areas (Jonckheere *et al.*, 2004).

To overcome the spatial and temporal limitations of traditional direct and indirect methods, advanced technologies such as remote sensing (RS) and LiDAR (Sanz *et al.*, 2018) have emerged as powerful alternatives. Significant progress in RS technology has improved LAI estimation by leveraging canopy reflectance data. Studies have shown that LAI can be estimated using spaceborne sensors, including AVHRR (Franch *et al.*, 2017), MODIS (Huang *et al.*, 2015), Landsat (Gao *et al.*, 2012), WorldView-2 (Psomiadis *et al.*, 2017), and Sentinel-2 (Sandonís-Pozo *et al.*, 2024; Weiss and Baret, 2016), as well as unmanned aerial vehicles (UAV) (Srinet *et al.*, 2019). RS approaches are generally categorized into (i) physically-based methods, which utilize radiative transfer models (Sahoo *et al.*, 2023), and (ii) empirical methods, which rely on linear or non-linear regression with vegetation indices (VIs) (Nguy-Robertson *et al.*, 2014). Point cloud variables have also been integrated into empirical methods, particularly in UAV-based remote sensing (Gao *et al.*, 2022; Quintanilla-Albornoz *et al*., 2023). The limited spatial resolution of satellites restricts their effectiveness in orchard plots, where finer detail is essential to capture canopy-level variation, making UAV platforms preferable. Despite this, efforts to estimate LAI in woody crops such as almond orchards using UAV platforms are scarce, with only few studies to date addressing this need (Bellvert *et al*., 2021; Quintanilla-Albornoz *et al*., 2023).

Machine learning (ML) algorithms offer promising advancements for data-intensive research in empirical modelling approaches. Various algorithms have been applied to estimate LAI using UAV platforms, including support vector regression (SVR), extreme gradient boosting (XGBoost), and random forest regression (RF) (Zhang *et al*., 2021; Gao *et al*., 2022). Comparative studies on ML algorithm performance for LAI estimation have found that RF generally achieves superior results (Gao *et al*., 2022; Chiu and Wang, 2024). Deep learning, a subset of ML, has also gained attention for its capacity to process large datasets and detect complex patterns (Kamilaris and Prenafeta-Boldú., 2018). Although not as widely adopted as traditional ML algorithms, artificial neural networks (ANN) are increasingly common in LAI estimation with UAV data (Xu *et al*., 2023; Cheng *et al*., 2024).

Despite significant advances in LAI estimation methods, current approaches face limitations when applied to woody crops like almond orchards. The low spatial resolution of satellite-based methods and the scarcity of UAV-based studies in heterogeneous canopy structures highlight the need for more precise and adaptable approaches. This study aimed to address these gaps by evaluating the performance of RF and ANN as empirical modelling algorithms for LAI estimation using UAV data. Accurate LAI predictions are crucial for improving crop transpiration assessments through energy balance models, which support effective irrigation management in almond orchards. To test the generalizability of these models, their performance was validated in two distinct orchard plots.

Materials and methods

Study area

The study was conducted in two almond orchards (Figure 1): Plot 1, a commercial orchard in Castelldans (CDL), Spain (41.499419° N, 0.80294° W, WGS 84), and Plot 2, located at the IRTA experimental station in Les Borges Blanques (BRB), Spain (41.509504° N, 0.853615° W, WGS 84). In Plot 1, trees were spaced 5 m apart with 6 m between rows, while Plot 2 featured trees spaced 2 m apart with 4.5 m between rows. Both orchards, irrigated with drip systems, were studied during the 2023 growing season.

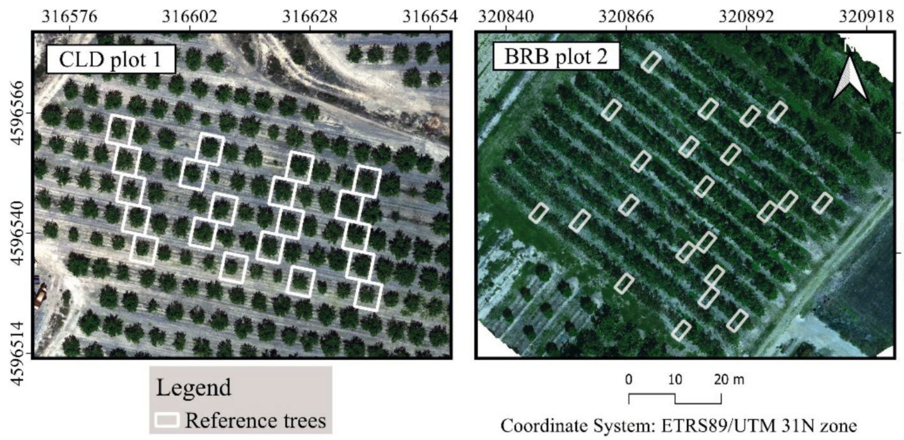

Figure 1. Locations of the case study area.

LAI ground truth acquisition

The ground truth for LAI was obtained using canopy-intercepted photosynthetically active radiation (fIPAR) measurements and subsequent calculations. fIPAR was measured on clear days concurrent with image acquisition, between 11:00 and 14:00 h (local time), using a portable ceptometer (AccuPAR LP80, Decagon Devices, Pullman, WA, USA). Incident PAR was recorded both above and below the canopy for each reference tree (Figure 1). A total of 150 readings were collected across 8 UAV flights, with 100 reading from Plot 1 and 50 readings from Plot 2, ensuring comprehensive coverage of tree spacing. Of these, 20 measurements were taken below the canopy for each reference tree. The ceptometer was positioned horizontally at ground level, perpendicular to the rows. Instantaneous fIPAR ($fiPAR_i$) was calculated as the ratio of average below-canopy PAR to incident PAR recorded in open, sunlit areas.

LAI was subsequently derived from $fiPAR_i$ measurements using the Norman-Jarvis model (Norman and Jarvis, 1974). Leaf absorptivity in the PAR region was set at 0.86071, and a spherical canopy structure was assumed (χ=1) based on almond tree characteristics (Zarate-Valdez *et al.*, 2012). The sunbeam fraction of incident radiation (fb) was calculated following Spitters *et al.* (1986), and the extinction coefficient (K) was determined using Beer's law, incorporating the sun's zenith angle. Canopy transmittance (τ) was computed as $1-fiPAR_i$. The resulting LAI values served as the target variable for machine learning models designed to estimate LAI based on additional features.

Image data processing and feature extraction

The image acquisition campaign involved five UAV flights over plot 1 on May 18, June 16, July 17, August 2, and August 22, and three flights over plot 2 on May 29, July 17, and August 22 in 2023. Each flight took place near solar noon (14:00 h local time) under clear sky conditions, using a Dronehexa XL (DRONETOOLS, Seville, Spain) equipped with a Micasense RedEdge-MX multispectral camera (Micasense, Seattle, WA, USA). This camera captured five spectral bands at wavelengths of 475±20 nm, 560±20 nm, 668±10 nm, 717±10 nm and 840±40 nm. Flights were conducted at 50 meters a.g.l., achieving a spatial resolution of 0.03 m.

All images underwent radiometric, atmospheric, and geometric corrections. In situ spectral measurements of ground targets were acquired using a FieldSpec 4 Standard-Res Spectroradiometer (Malvern Panalytical, UK) with a wavelength range of 350–2500 nm and optical resolution of 3–10 mm. Calibration was performed using a white Spectralon panel and a dark reference. Image mosaicking and digital elevation model (DEM) generation from multispectral point clouds were conducted with Agisoft Metashape Professional, and final corrections were applied in QGIS 3.36.

To extract model features from UAV multispectral and point cloud data, a canopy layer was delineated by selecting pixels with height values exceeding 1.5 m and a soil-adjusted vegetation index (SAVI) above 0.2 (Quintanilla-Albornoz *et al.*, 2023). From this canopy layer and the DEM, structural features were derived, including fractional canopy cover (f_c), maximum canopy height ($h_{c,max}$), and mean canopy height ($h_{c,mean}$). Additionally, mean, and standard deviation values were calculated for multiple VIs, specifically the normalized difference vegetation index (NDVI), modified triangular vegetation index (MTVI2), and SAVI.

Machine learning algorithms

RF and ANN were applied in this study due to the strong performance of RF for LAI estimation (Chiu and Wang, 2024; Gao *et al.*, 2022) and the growing use of ANN for capturing complex patterns (Cheng *et al.*, 2024; Xu *et al.*, 2023).

RF is an ensemble method using bootstrap aggregation to build a collection of decorrelated decision trees and average their predictions. To reduce overfitting, an hyperparameter optimization process was carried out to define the number of trees (10–50) and minimum samples per split (2–10), using the validation dataset. This model was implemented using the scikit-learn Python library.

For the ANN, the ReLU activation function and the 'adam' optimizer were used. Hyperparameters, including the number of hidden layers (1–3) and neurons per layer (60–100), were optimized using the same validation set. Shallow networks were prioritized due to the limited training data (105 samples). Early stopping was applied by analysing training and validation curves to prevent overfitting. ANN development was conducted using the Keras Python library.
The dataset was split into 70% training, 15% validation, and 15% test sets. The performance of both methods, RF and ANN, was evaluated by comparing the LAI estimated by the machine learning models based on UAV-derived data with the LAI measured from fiPAR measurements using the Norman-Jarvis model, considered as the ground truth. The workflow is depicted in Figure 2. The following metrics were used for this evaluation: coefficient of determination (R^2), root mean square error (RMSE), relative RMSE (RRMSE), mean absolute error (MAE), and the Nash-Sutcliffe efficiency (NASH).
The dataset and developed code have been made publicly available for reproduction of results at https://github.com/UEA-Irta/UAV-biophysical-estimation.git.

Results and discussion

Machine learning models

Figure 3 shows the model performance in the test set for each ML algorithm. Model selection was based on performance metrics on validation set, with parameters tuned for optimal results during the hyperparameter optimization process applied in both RF and ANN models. The final RF model was configured with 45 trees in the forest (n_estimators=45) and a minimum of five samples per leaf (min_samples_leaf=5). For the neural network, a two-layer architecture was employed, with one hidden layer consisting of 85 neurons (hidden_layers=1, neurons=85). This configuration yielded the best performance in the validation set at epoch 173, providing a basis for comparison with the RF model. The ANN outperforms overall RF model, with higher R^2 (0.94 vs. 0.92) and lower relative root mean square error (RRMSE=0.07 vs. 0.10 m^2/m^2). In addition, Figure 3 shows that the RF model slightly overestimates LAI values below 0.8 m^2/m^2 and underestimates values above this threshold, suggesting a systematic bias.

Figure 4 compares the LAI maps obtained with the final RF and ANN models on 18 May 2023 in plot 1 (CLD) and 29 May 2023 in plot 2 (BRB). In both fields, LAI estimations using ANN displayed greater variability compared to RF, justified by the high variability of tree sizes in the study site. Additionally, RF estimations showed lower LAI values compared to ANN in the second orchard

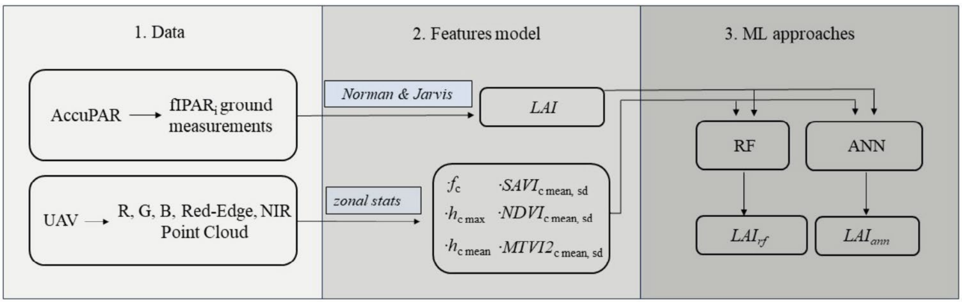

Figure 2. A three-step experiment for training ML algorithms for LAI estimation. Step 1 collects data from an AccuPAR ceptometer ($fPAR_i$ measurements) and UAV (multispectral and point cloud data). Step 2 estimates initial LAI with the Norman and Jarvis model using $fPAR_i$ measurements and extracts model features. Step 3 applies RF and ANN to predict LAI values.

(BRB), which had taller and denser trees, corroborating the underestimation of LAI by RF for high-LAI samples observed in the test set (Figure 3).
In terms of computational complexity and speed, RF estimated the LAI in both plots in 0.01 s, while ANN processed these plots in 0.19 s. This demonstrates that RF is faster, although both methods are efficient in computational time.

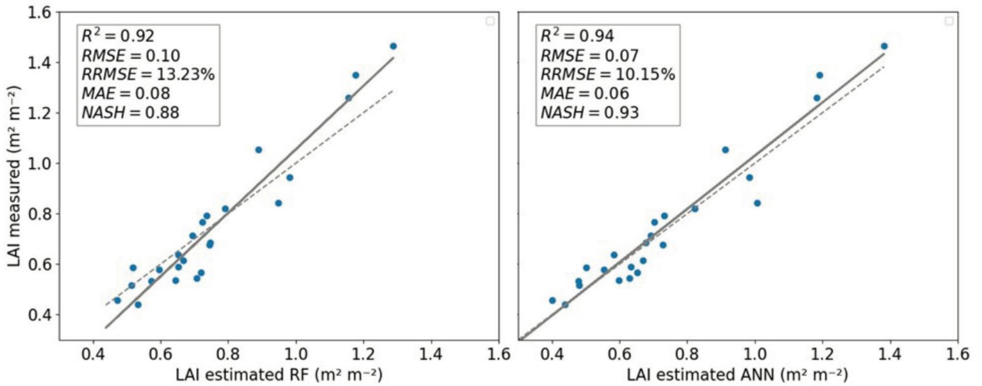

Figure 3. Comparison between the performance of the two final models (RF and ANN) in the test set.

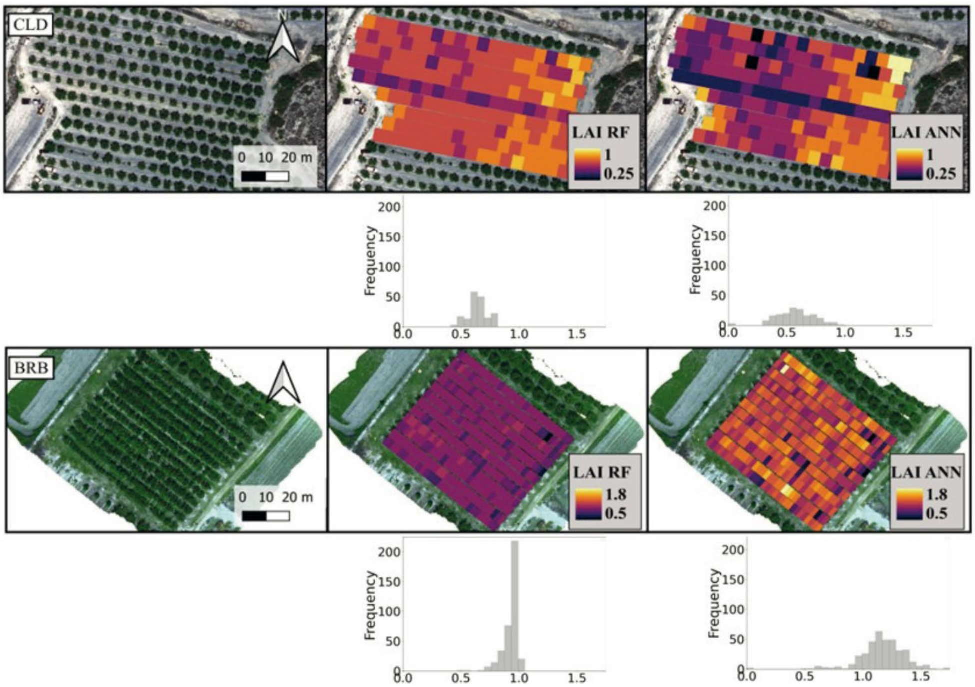

Figure 4. Differences in the prediction of LAI data by both models in plot 1 (CLD) on 18/05/2023 and plot 2 (BRB) on 29 May 2023.

Conclusion

In this study, an empirical model leveraging high-resolution multispectral and point-cloud data from UAVs, combined with machine learning techniques, was shown to be an effective solution for estimating LAI -an essential parameter for quantifying crop water needs and support other management decisions in almond orchards. ANN outperformed RF on the test set and demonstrated better generalization across both plots when applied to new data. Therefore, the results suggest that ANN algorithms can enhance compared to traditional machine learning techniques, highlighting their potential for accurate LAI estimation in complex orchard environments. Reliable LAI predictions play a key role in refining crop transpiration analysis within energy balance models, contributing to optimized irrigation strategies in almond orchards. Future work will extend this study to other fruit orchard varieties and cultivars, as well as expand the dataset to train deeper neural networks capable of further improving performance.

Acknowledgement

This study has been supported by the project ET4DROUGHT (No. PID2021-127345OR-C31) funded by the Ministry of Science and Innovation (MICINN-AEI) of Spain.

References

Bellvert, J., Nieto, H., & Pelechá, A. (2021) Remote sensing energy balance model for the assessment of crop evapotranspiration and water status in an almond rootstock collection. Frontiers in Plant Science, 12, 1.

Cheng, Q., Ding, F., Xu, H., Guo, S., Li, Z., & Chen, Z. (2024). Quantifying corn leaf area index using machine learning and UAV multispectral imaging. Precision Agriculture, 25(8), 1777–1799.

Chiu, M.S., & Wang, J. (2024). Evaluation of machine learning regression techniques for estimating winter wheat biomass using biophysical, biochemical, and UAV multispectral data. Drones, 8(7), 287.

Franch, B., Vermote, E., Roger, J.-C., Murphy, E., Becker-Reshef, I., Justice, C., …, & Devadiga, S. (2017). A 30+ year AVHRR land surface reflectance climate data record and its application to wheat yield monitoring. Remote Sensing, 9(3), 296.

Gao, F., Anderson, M.C., Kustas, W.P., & Wang, Y. (2012). Simple method for retrieving leaf area index from Landsat using MODIS leaf area index products as reference. Journal of Applied Remote Sensing, 6(1), 063554.

Gao, R., Torres-Rua, A., Aboutalebi, M., White, W., Anderson, M., Kustas, W., …, & Gowing, I. (2022). LAI estimation across California vineyards using sUAS multi-seasonal multi-spectral, thermal, and elevation information and machine learning. Irrigation Science, 40, 731–759.

Gower, S.T., Kucharik, C.J., & Norman, J.M. (1999). Direct and indirect estimation of leaf area index, f(APAR), and net primary production of terrestrial ecosystems. Remote Sensing of Environment, 70(1), 29–51.

Huang, J., Ma, H., Su, W., Zhang, X., Huang, Y., Fan, J., & Wu, W. (2015). Jointly assimilating MODIS LAI and ET products into the SWAP model for winter wheat yield estimation. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(9), 4060–4071.

Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M., & Baret, F. (2004). Review of methods for in situ leaf area index determination part I. Theories, sensors, and hemispherical photography. Agricultural and Forest Meteorology, 121(1), 19–35.

Kamilaris, A., & Prenafeta-Boldú, F.X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70–90.

Nguy-Robertson, A.L., Peng, Y., Gitelson, A.A., Arkebauer, T.J., Pimstein, I., Herrmann, A., Karnieli, A., Rundquist, D.C., & Bonfil, D.J. (2014). Estimating green LAI in four crops: potential of determining optimal spectral bands for a universal algorithm. Agricultural and Forest Meteorology, 192, 140–148.

Norman, J.M., & Jarvis, P.G. (1974). Photosynthesis in Sitka spruce (*Picea sitchensis* (Bong.) Carr.): III. Measurements of canopy structure and interception of radiation. Journal of Applied Ecology, 11(2), 375–398.

Psomiadis, E., Dercas, N., Dalezios, N.R., & Spyropoulos, N.V. (2017). Evaluation and cross-comparison of vegetation indices for crop monitoring from Sentinel-2 and WorldView-2 images. In Remote Sensing for Agriculture, Ecosystems, and Hydrology XIX. International Society for Optics and Photonics, Bellingham, WA, USA. Vol. 10421, pp. 239–251.

Quintanilla-Albornoz, M., Miarnau, X., Pelechá, A., Casadesús, J., García-Tejera, O., & Bellvert, J. (2023). Evaluation of transpiration in different almond production systems with two-source energy balance models from UAV thermal and multispectral imagery. Irrigation Science, 43, 29–49. https://doi.org/10.1007/s00271-023-00888-1

Sahoo, R.N., Gakhar, S., Rejith, R.G., Verrelst, J., Ranjan, R., Kondraju, T., …, & Chinnusamy, V. (2023). Optimizing the retrieval of wheat crop traits from UAV-borne hyperspectral image with radiative transfer modelling using Gaussian process regression. Remote Sensing, 15, 5496.

Sandonís-Pozo, L., Oger, B., Tisseyre, B., Llorens, J., Escolà, A., Pascual, M., & Martínez-Casasnovas, J.A. (2024). Leafiness-LiDAR index and NDVI for identification of temporal patterns in super-intensive almond orchards as response to different management strategies. European Journal of Agronomy, 159, 127278.

Spitters, C.J.T., Toussaint, H.A.J.M., & Goudriaan, J. (1986). Separating the diffuse and direct component of global radiation and its implications for modeling canopy photosynthesis Part I. Components of incoming radiation. Agricultural and Forest Meteorology, 38(1–3), 217–229.

Sanz, R., Llorens, J., Escolà, A., Arnó, J., Planas, S., Román, C., & Rosell-Polo, J.R. (2018). LIDAR and non-LIDAR-based canopy parameters to estimate the leaf area in fruit trees and vineyard. Agricultural and Forest Meteorology, 260–261, 229–239.

Srinet, R., Nandy, S., & Patel, N.R. (2019). Estimating leaf area index and light extinction coefficient using random forest regression algorithm in a tropical moist deciduous forest, India. Ecological Informatics, 52, 94–102.

Watson, D.J. (1947). Comparative physiological studies on the growth of field crops: II. The effect of varying nutrient supply on net assimilation rate and leaf area. Annals of Botany, 11, 375–407.

Weiss, M., & Baret, F. (2016). S2ToolBox Level 2 products: LAI, FAPAR, FCOVER—Version 1.1. Sentin. ToolBox Level2 Prod. 53. Available online at https://step.esa.int/docs/extra/ATBD_S2ToolBox_L2B_V1.1.pdf

Xu, W., Yang, F., Ma, G., Wu, J., Wu, J., & Lan, Y. (2023). Multiscale inversion of leaf area index in citrus tree by merging UAV LiDAR with multispectral remote sensing data. Agronomy, 13(11), 2747.

Zarate-Valdez, J.L., Whiting, M.L., Lampinen, B.D., Metcalf, S., Ustin, S.L., & Brown, P.H. (2012). Prediction of leaf area index in almonds by vegetation indexes. Computers and Electronics in Agriculture, 85, 24–32.

Zhang, J., Cheng, T., Guo, W., Xu, X., Qiao, H., Xie, Y., & Ma, X. (2021). Leaf area index estimation model for UAV image hyperspectral data based on wavelength variable selection and machine learning methods. Plant Methods, 17(49).