98. Estimating cover crop biomass from optical satellite images for integration in a PrecisionAg service

A. Veloso*, C. Biller, M. Wolde-Mikael, M. Ortolan and A. Jacquin
Airbus Defence and Space, 5 rue des Satellites, BP 14359, F-31031 Toulouse Cedex 4, France ;
** amanda.veloso@airbus.com*

Abstract

Cover crops play a critical role in enhancing soil health, preventing erosion, managing nutrients, and mitigating climate change through carbon sequestration. Accurate estimation of cover crop biomass is essential for optimizing agronomic and environmental benefits, especially for large-scale precision agriculture applications. Despite their potential, estimating biomass accurately and efficiently over large areas remains a significant challenge. This study explores the use of optical multi-satellite imagery to estimate cover crop biomass through an empirical approach based on satellite-derived biophysical parameters and in-situ data. The aim is to integrate these estimations into a precision agriculture service to provide farmers with a turnkey solution that supports better nitrogen management and low carbon practices.

Keywords: biomass estimation, cover crops, nitrogen management, optical remote sensing data, sustainable farming

Introduction

Carbon farming and more specifically cover crops cultivation represents an innovative agricultural approach that enables farmers to capture and sequester carbon in soils (Poeplau and Don, 2015). Their integration into farming systems is increasingly recognized as a critical component of sustainable agricultural practices under policies like the European Union's Common Agricultural Policy (CAP) and similar initiatives worldwide. However, accurately quantifying the biomass of cover crops remains challenging, especially across diverse climatic and cropping systems. Remote sensing technologies, particularly those based on optical satellite imagery, offer a powerful tool to tackle this issue objectively at large scales. Vegetation indices derived from multispectral bands, such as the normalized difference vegetation index (NDVI), have demonstrated potential for estimating biomass, though challenges persist in accurately modeling mixed-species cover crops, especially for high biomass values because of saturation issues. Previous studies have successfully applied remote sensing to estimate cover crop biomass at local scales, leveraging empirical relationships between vegetation indices and biomass. For instance, Goffart *et al.* (2021) validated the utility of Sentinel-2 data for biomass estimation in Belgium, achieving reliable results for monocultures but facing challenges with mixed species. Similarly, do Nascimento *et al.* (2024) highlighted the potential of combining machine learning with dense Sentinel-2 time series for winter cover crop biomass estimation in France. However, the need for scalable, automated and easy to deploy approaches capable of addressing variability in climatic conditions, species diversity, and management practices remains evident. This study aims to build on these advancements by developing a methodology for converting optical satellite imagery into precise biomass evaluations for a range of cover crop systems. The proposed method is based on an asset of firstly estimating biophysical parameters, in particular the Green Leaf Area Index (GLAI), through reflectance model inversion. This process allows to take into account specific plant, soil and satellite sensors characteristics. It results in biophysical estimations that are spatially and temporally consistent over time and less prone to inter-sensor bias. These products are then used to establish an empirical model along with in-situ data gathered from multiple sources. The proposed approach seeks to integrate the estimated biomass data into precision agriculture services, enabling better management of the resources, with a more precise quantification of nutrients available for the subsequent main crops but also providing farmers with tools to quantify soil organic carbon stock changes.

Materials and methods

Field data

A specific field campaign was conducted in 2023 to collect in-situ biomass data. Plots were located across main agronomic production areas in France, representing diverse cover crop species and management practices under contrasting climatic conditions (see Figure 1). Fields were selected based on visual analysis of remote sensing data and in-situ inspection to cover a wide range of biomass values. The main represented species were phacelia, fava beans, radish, vetch, mustard and oat. The dataset was mostly composed of cover crops mixtures rather than single species (see Figure 2). A total of 69 fields were sampled.

Sampling protocol

The elementary sampling unit (ESU) protocol was applied to collect the fresh biomass samples. This protocol was chosen to ensure consistency of spatial representativeness between the satellite and the field data. The first step consisted in identifying large homogeneous areas for sampling (corresponding to 50 m×50 m, ideally 70 m×70 m). The sampling area (ESU) was then defined as a square of 10m x 10m square area (equivalent to an average pixel area). The centre of the ESU had to be at least 40 m from the edge of the plot, up to 60 m in case the plot was surrounded by hedges. Destructive measurements were taken over four to five 1-m^2 sampling plots inside the ESU. The coordinates of the centre of each ESU were recorded with a GPS device (Figure 3). For each sampling plot unit, plant cover species were sorted and fresh biomass weighted separately. The same process was done when weeds and/or crop regrowth were present in the sampling. Photographs were taken along with the measurements at each field to support further analysis and results (Figure 3, right).

An additional dataset was gathered using data from research projects and public institutions collected over past years (2018 to 2023). This data was used to improve the methodology's robustness and to ensure larger representativity of the species, climate, environment and crop practices over the territory. This yielded a total of 113 additional samples of biomass data.

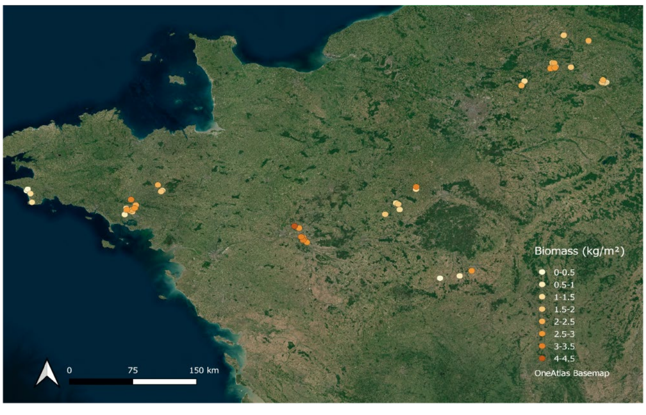

Figure 1. Location of cover crop biomass fields sampled during the in-situ campaign conducted in 2023. Colours of the plot represent their level of measured biomass, in kg/m^2.

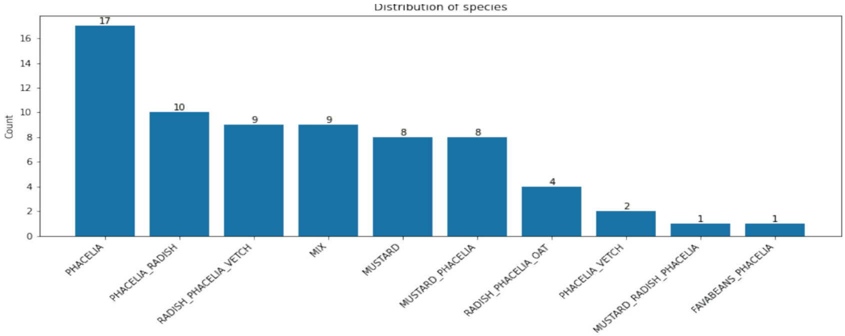

Figure 2. Classification of the different species for the 69 fields and count of fields for each cover crop category.

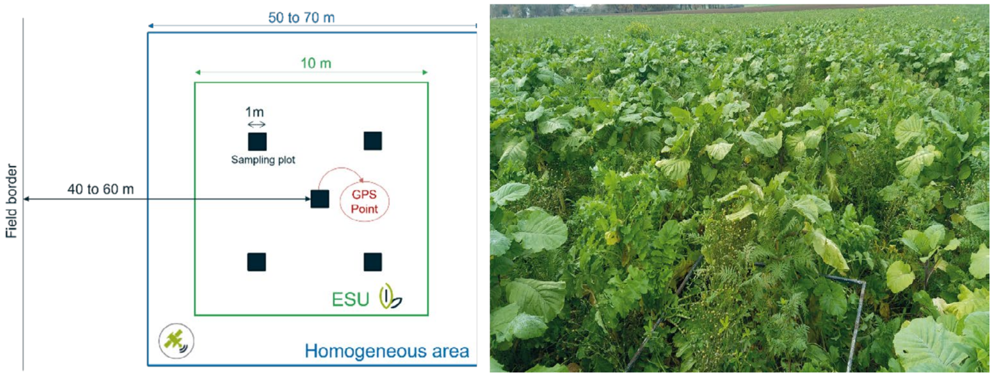

Figure 3. (Left) Schematic illustration of the in situ measurements requirements and the Elementary Sampling Unit (ESU) protocol. (Right) Visual of a sampled plot with a composition of radish, phacelia and vetch, with an average fresh biomass of 3.04 kg/m^2.

Remote sensing data

Data from different optical remote sensing sensors were explored to have acquisitions the closest to the field sampling dates. Images from Sentinel-2, Landsat-8, Pleiades and Pleiades-Neo satellites were considered. Level-1 reflectance products have been used to derive biophysical parameters such as green leaf area index (GLAI), fraction of vegetation cover (FCOVER), chlorophyll content (CHL) and fraction of absorbed photosynthetically active radiation (fAPAR) using the Overland processing chain developed by Airbus DS. This chain is based on physical models, the SAIL and PROSPECT models being core elements of the canopy reflectance model (Jacquemoud & Baret, 1990; Verhoef, 1984), and LOWTRAN (Kneisys *et al.*, 1995) completed with a dedicated cloud mode. The Overland built-in atmospheric model allows autonomous atmospheric correction of reflectances, as well as a masking of clouds and dark shadows (Poilvé, 2010). The inversion of this ensemble of models through minimization techniques allows to estimate a series of biophysical variables.

Images available from sowing date to destruction date were processed. Despite the processing spatial resolution varies according to characteristics of each sensor, data were exported at a common resolution of 5 m. Values of the different biophysical parameters were extracted on sampling plot locations (=ESU centre coordinates) using a 20-m radius buffer to establish the relationships with in situ biomass. Median values were calculated for each sample.

Despite the availability of an extensive temporal series of multi-source images, it is usually a challenge to have the in-situ measurements acquired concurrently with satellite observations (Beriaux *et al.*, 2021). Figure 4 illustrates a temporal series of GLAI derived from different satellite sensors for a phacelia cover crop field.

Model calibration and evaluation

The empirical approach involved testing and establishing relationships between the satellite-based biophysical parameters and the measured fresh aboveground biomass. For each field and corresponding measurement date, a satellite image was manually selected to align as closely as possible with the in-situ sampling date. The selection process adhered to a temporal proximity criterion, allowing images acquired either before or after the ground-based sampling within a specified timeframe, minimizing potential discrepancies caused by temporal misalignment.

An initial analysis was performed to determine the most effective biophysical parameter and/or vegetation index that better correlates with biomass. Linear, polynomial (2 degrees) and exponential regression models were tested, optimizing for the best-fit relationships between satellite-based variable and biomass, looking for the highest coefficient of determination (R^2) and the smallest root mean square error (RMSE). The methodological choice was to prioritize empirical models, given the significant well-known correlation between biophysical parameters, such as Leaf Area Index (LAI), and biomass (Goffart *et al.*, 2021). One notable advantage of this approach is to provide a straightforward and interpretable framework for estimating biomass directly from satellite images, with minimal computational complexity.

Biomass data were split using a train-split method, partitioning the available dataset into two mutually exclusive subsets: a training set (80% of the data) used to build and optimize the model and a test set (20% of data used to assess the model's performance on independent data. The 80:20 splitting integrated specific constraints to ensure representativity of the variability of cover crops species and pedo-climatic environments in both subsets. This process, known as stratified sampling, ensures that the proportion of each class in the training and test sets mirrors the original dataset.

Another important processing step was the outlier's detection. A data point may be classified as an outlier due to anomalous field reference values or atypical biophysical measurements. Outlier detection was mainly accomplished through visual inspection of regression plots and analysis of the associated GLAI and true colours images.

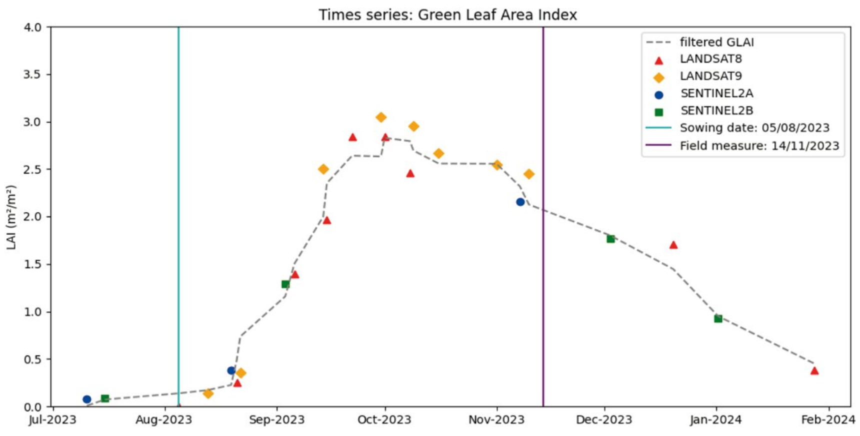

Figure 4. GLAI Time series obtained from reflectance inversion model from different sources (LANDSAT-8 and -9, SENTINEL2-A and -B). The points represent the median values of GLAI for a phacelia field. Sowing date and date of sampling are indicated by vertical bars.

Regression models were evaluated using various statistical metrics as error measures (root mean square error, RMSE, relative root mean square error, RRMSE and the mean absolute error, MAE), bias and the goodness of fit (R^2). The analyses were conducted using Python, leveraging libraries such as pandas, scikit-learn and matplotlib.

From these analyses, a final model was selected for estimating cover crops biomass. Next, the model was applied over the entire dataset at pixel level to generate a cover crop fresh biomass map for each field.

Further applications

Further objective of this study was to integrate the field-specific biomass estimations into a precision agriculture service, enabling this information to be explored for different uses and farmers' needs. Biomass maps were integrated into a commercial product delivered to the farmers allowing them to improve cover crops management. Secondly, biomass information was used as input to quantify i) nitrogen availability from cover crop residues to the next crop and ii) carbon sequestration in the soil induced by the cover crops cultivation. The MERCI method (Constantin *et al.*, 2024) was chosen for addressing these issues. The MERCI decision-support tool allows to estimate nitrogen availability from cover crop residues to the next crop. This output provides farmers liable information for adapting and optimizing their nitrogen management practices, allowing them to reduce the use of mineral artificial fertilizers. MERCI also provides the amount of carbon sequestration in the soil thanks to the cover crops cultivation. This information can support farmers in adopting more sustainable farming practices while also creating opportunities for new revenue streams.

Results and discussion

Analysis of biomass collected data

The distribution of sampled fresh aboveground biomass for each dataset is displayed in Figure 5 (left). It can be observed that biomass values are quite variable depending on the dataset. Median biomass values of samples vary from 0.13 up to 4.16 kg/m^2. This finding is consistent with values from other studies, as reported for winter cover crops in Goffart *et al.* (2021) and do Nascimento *et al.* (2024). In this study, data collection prioritized measurements from plant species most commonly encountered by our stakeholders and not only on range biomass values.

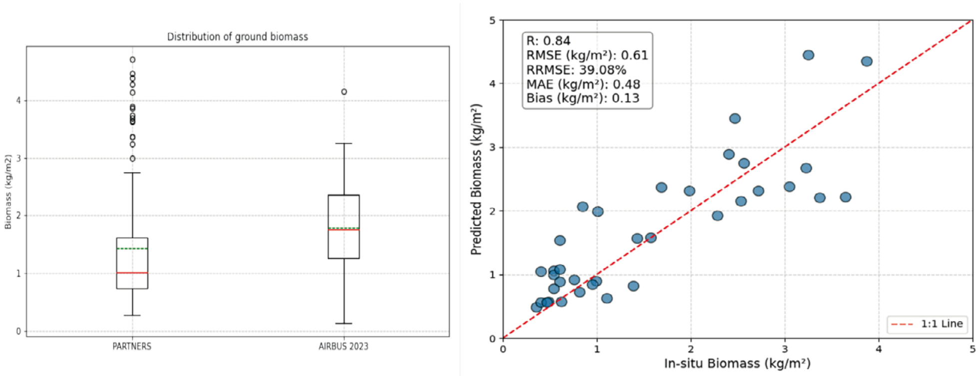

Figure 5. (Left) Distribution of biomass values from both in-situ datasets used in this study: collected by partner institutions and inner 2023 campaign dataset. (Right) Scatter plot showing predicted biomass using the proposed approach versus biomass in-situ measurements (both in kg/m^2) calculated over the validation dataset (n=36). The performance metrics (R, RMSE, RRMSE, MAE and Bias) are displayed in the top-left corner.

Empirical model

From the investigated biophysical parameters, results indicate that the product GLAIxCHL demonstrates better correlations with measured biomass across different cover crop species (Table 1). Still, the GLAI parameter also presents a robust correlation with biomass (Figure 6, right). It is noteworthy that chlorophyll quantification presents greater complexity, since not all remote sensing instruments have the necessary spectral bands required to derive this biophysical parameter. Besides, using a single layer as input of the model (instead of a product as GLAI*CHL) simplifies the analytical and operational process.

The NDVI, widely used for biomass estimation in cover crops due to its sensitivity to vegetation greenness and density, was also tested. However, as expected, the NDVI saturates at high levels of biomass (see Figure 6, left). As vegetation development becomes more important and leaf area index increases, NDVI reaches a threshold where additional increases in biomass no longer result in significant changes in the index. For cover crops, this saturation effect can make it challenging to differentiate between moderate and high biomass levels, limiting its utility for precise quantification of biomass. The GLAI parameter was selected for the subsequent analyses. Few outliers (n=7 samples) were detected and excluded from the calibration dataset to mitigate their impact on the model's performance (Figure 6, right). These points were mainly due to the presence of undetected segments of clouds.

The best-performing model was a polynomial (2 degrees) achieving a R^2 value of 0.65 and a root mean square error (RMSE) of 0.60 kg/m^2 (Figure 5, right). This performance indicates a reliable estimation capability under contrasting field conditions.

One single model was established for all cover crop species. No particular behavior was identified for any of the analyzed cover crops types; however, this could be related with the limited number of samples per cover crop mixture/sole type. The inclusion of additional field data would help improve the model's accuracy, particularly in allowing to distinguish between cover crop types and to increase the range of biomass values, especially for very high values. The selected model was applied to each pixel of all the studied fields, allowing generation of fresh biomass maps (Figure 7), used next for nitrogen and carbon applications.

On average, the mean absolute temporal interval between field sampling measurements and satellite imagery acquisition was 12 days. There is significant variability depending on the datasets and agricultural seasons. This value seems acceptable as the majority of the targeted fields will be winter cover crops. However, for cover crops estimated in spring, the rapid growth rate during this period might necessitate shortening the temporal interval.

Consequently, future work will focus on expanding the dataset to improve model accuracy and scalability, facilitating broader adoption of cover crops as part of climate-smart agriculture strategies. Furthermore, investigating the potential of time series data to better capture the cover crops growth dynamics and to identify the peak of biomass production is a worthwhile path to explore.

Table 1. Correlations between estimations of biophysical parameters and measures of in-situ biomass (n= 182 samples).

	GLAI	FCOVER	NDVI	FAPAR	GLAIxCHL
R^2	0.60	0.55	0.46	0.50	0.63

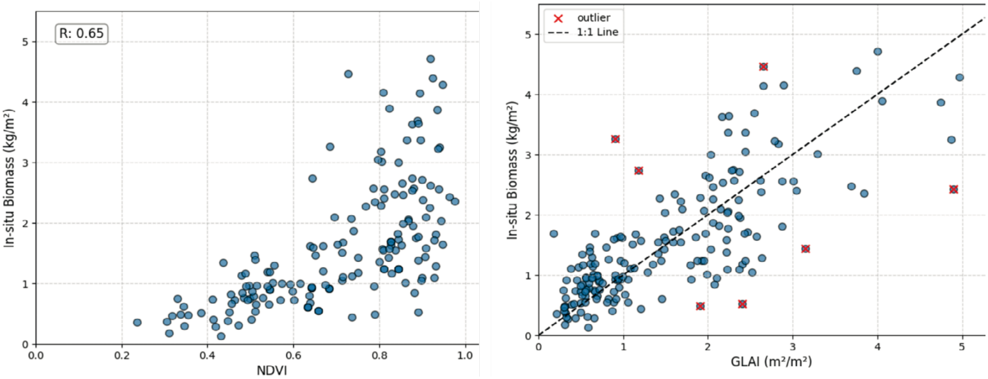

Figure 6. (Left) Scatter plot illustrating the relationship between the normalized difference vegetation index (NDVI) and fresh biomass in-situ measurements, in kg/m^2. (Right) The relationship between green leaf area index (GLAI), in m^2/m^2, and fresh biomass in situ measurements, in kg/m^2. The outliers are indicated by the red crosses.

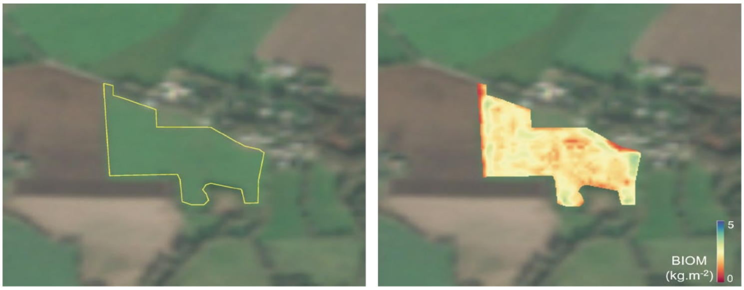

Figure 7. Example of a biomass map calculated using the developed methodology for a phacelia field. Calculated from a Sentinel2 image acquired on 22 November. (Left) The true colour image with the field contour in yellow; (Right) The fresh biomass map (kg^2/m^2).

Conclusions

This research demonstrates the potential of satellite-based GLAI for estimating cover crop biomass through an empirical approach. The integration of this model into an operational precision agriculture service will enable automated generation of biomass estimates at the field scale, providing valuable decision-support tools for farmers. These results represent a significant step providing farmers with precise biomass information to support them towards more sustainable and productive agricultural systems while contributing to climate resilience efforts.

In general, the results showed that the developed model, despite its relatively simplistic empirical approach, was able to correctly predict wet biomass of different types of cover crops (mixtures and distinct range of biomass). This methodology is ready to be deployed at the French territory scale with a turnkey solution for farmers.

The diversity of satellites used for obtaining the GLAI times series and maps, including tasking capabilities and very high resolution, represents a distinctive approach in this domain. Additionally, quality field data collection remains essential to obtain accurate and reliable estimations.

References

Beriaux, E., Jago, A., Lucau-Danila, C., Planchon, V., & Defourny, P. (2021). Sentinel-1 time series for crop identification in the framework of the future CAP monitoring. Remote Sensing, 13(14), 2785.

Constantin, J., Minette, S., Vericel, G., Jordan-Meille, L., & Justes, E. (2024). MERCI: a simple method and decision-support tool to estimate availability of nitrogen from a wide range of cover crops to the next cash crop. Plant and Soil, 494(1), 333–351.

do Nascimento Bendini, H., Fieuzal, R., Carrere, P., Clenet, H., Galvani, A., Allies, A., & Cheschia, É. (2024). Estimating winter cover crop biomass in France using optical sentinel-2 dense image time series and machine learning. Remote Sensing, 16(5), 834.

Goffart, D., Curnel, Y., Planchon, V., Goffart, J. P., & Defourny, P. (2021). Field-scale assessment of Belgian winter cover crops biomass based on Sentinel-2 data. European Journal of Agronomy, 126, 126278.

Jacquemoud, S., & Baret, F. (1990). PROSPECT: a model of leaf optical properties spectra. Remote Sensing of Environment, 34(2), 75–91.

Kneisys, F.X., Abreu, L.W., Anderson, G.P., Chetwynd, J.H., Shettle, E.P., Berk, A., …, & Clough, S.A. (1995). The MODTRAN 2/3 and LOWTRAN 7 model. Ontar, North Andover, MA.

Poeplau, C., & Don, A. (2015). Carbon sequestration in agricultural soils via cultivation of cover crops–A meta-analysis. Agriculture, Ecosystems & Environment, 200, 33–41.

Poilvé, H. (2010). BioPar Methods Compendium of MERIS FR Biophysical Products (report g2-BP-RP-038, EC geoland2 project FP-7-218795).

Verhoef, W. (1984). Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. Remote Sensing of Environment, 16, 125-141.