ICPA
17th INTERNATIONAL CONFERENCE ON PRECISION AGRICULTURE
ConBAP 2026
CONGRESSO BRASILEIRO DE AGRICULTURA DE PRECISÃO E DIGITAL
13-16 July 2026 | Porto Alegre, Brazil

Leaf nutrient estimation in soybean from multispectral and multitemporal information using UAV and machine learning

Erney Ricardo Cely Bonilla[1], Claudio Leones Bazzi[1], Kelyn Schenatto[1], Ricardo Sobjak[1], Eduin Torres Avila[1], Marlon Rodrigues[1], Samuel Spricigo[1]

[1]Universidade Tecnológica Federal do Paraná (UTFPR), Medianeira, Paraná, Brasil

A paper from the Proceedings of the

17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture

13-15 July 2026

Porto Alegre, Brazil

Abstract.

Precision agriculture, enabled by remote sensing with Unmanned Aerial Vehicles (UAVs) and Artificial Intelligence, offers solutions for monitoring crop growth and development, supporting decision-making to optimize resources and promote agricultural sustainability. This study evaluated the feasibility of using multispectral imagery captured by UAVs at different phenological stages (V6, V8, and R2) to estimate leaf nutrient concentrations (N, P, K, Ca, Mg, Cu, Zn, and Mn) in soybean crops as an alternative to traditional destructive sampling. Reference values for leaf nutrient content were quantified in the laboratory from 106 sampling points across two commercial soybean fields located in Céu Azul, Paraná, Brazil. Random Forest (RF), Support Vector Regression (SVR), and Artificial Neural Network (ANN) algorithms were trained using 44 vegetation indices and 5 multispectral bands. Five feature extraction strategies were established for each sampling point from the multispectral images: 1) mean-based, 2) mean by subzones, 3) total pixels, 4) pixels with OSAVI index above the median, and 5) pixels with OSAVI index above the third quartile. The models were evaluated using the Coefficient of Determination (R^2) and the Relative Root Mean Square Error (RRMSE), with a 70% training and cross-validation split (kf=5) and a 30% test split. The comparative evaluation of the machine learning algorithms demonstrated that RF achieved a significantly superior performance compared to SVR and ANN for estimating foliar nutrient content in soybean, achieving average R^2 values in the validation sets that were 7.19 to 40.87 units higher. The pixel-based feature extraction strategies generated models with significantly higher performance, reaching an R^2 of 0.39, whereas the other approaches barely reached 0.04. The final RF models, trained on pixel-based data and hyperparameter-optimized, exhibited varying predictive performance across the eight analyzed nutrients. Zinc achieved the most accurate predictions (R^2=0.54, RRMSE=12.65%), while Nitrogen and Phosphorus presented the lowest relative errors (8.43% and 8.67%, respectively). This study highlights the value of multispectral and multitemporal information for soybean nutritional management. It validates a robust machine-learning-based methodology for nutrient estimation, particularly in agronomic studies where field sample sizes are typically limited.

Keywords.

Vegetation Indices, Remote Sensing, Predictive Models, Precision Agriculture

Introduction

The agricultural sector faces a profound transformation to achieve higher levels of efficiency and sustainability, ensuring food security amid challenges such as rapid demographic growth, climate change, and the Sustainable Development Goals (SDGs) (Bertoglio et al., 2021; Martos et al., 2021). Thus, it is necessary to rethink agronomy through a sustainable approach, in which the management of agricultural resources simultaneously preserves ecological resources and promotes economic profitability, utilizing technological tools as means (Bhakta, Phadikar, and Majumder, 2019). In this regard, precision agriculture (PA) is a fundamental strategy to increase crop yields by using production inputs more efficiently, thereby mitigating negative environmental impacts and promoting sustainability (Karunathilake et al., 2023; Rahali et al., 2025).

Remote sensing (RS) is one of the technologies used by PA to acquire georeferenced information over large land areas at high temporal and spatial resolution, without the need for destructive sampling (Omia et al., 2023). Data collected via RS can be used in comprehensive analyses to support decision-making on agricultural productivity (Omia et al., 2023). In particular, multispectral (MS) information captured by RS has proven effective for monitoring foliar nutrient content in crops (De Oliveira et al., 2024), which is fundamental for diagnosing plant nutritional status and correcting potential nutritional deficiencies (Costa et al., 2022).

The integration of RS with UAVs (Unmanned Aerial Vehicles) has enabled the collection of increasingly larger volumes of data (Hassler and Baysal-Gurel, 2019). In this sense, in recent years, artificial intelligence (AI) has enhanced the management and analysis of the vast datasets generated by PA, thereby optimizing predictive modeling and decision-making, while increasing agricultural efficiency and sustainability. Predictive models can assist farmers in planning management practices, thus optimizing resource use while promoting increased crop productivity (Mmbando, 2025).

Given the above, this study proposes practical technological solutions for a more efficient and sustainable management of soybean crops in Southern Brazil. To this end, a robust Machine Learning (ML) modeling methodology was developed to estimate crop nutritional status from multispectral images captured by UAVs. Three algorithms were evaluated—Random Forest (RF), Support Vector Regression (SVR), and Artificial Neural Networks (ANN)—combined with different feature extraction and dataset structuring strategies. This approach emerges as a promising alternative to traditional foliar diagnosis methods (macro and micronutrient analysis), which require destructive sampling, depend on specialized equipment, incur high costs, and use chemical reagents harmful to the environment.

Materials and methods

Experiment mounting and data collection

The experiment was conducted in a commercial soybean plantation located in the municipality of Céu Azul, Paraná State, Brazil, at the following coordinates: 25°06’41” S and 53°49’48” W. The study area is situated in a region with a humid subtropical climate (Cui et al., 2021), with an annual mean temperature of 20°C and an average precipitation of 1,800 mm. The experimental area comprised two distinct plots sown under a no-tillage system: a 16.3-hectare plot (Area A) cultivated with the variety SoyTech 591 I2X (BASF), and a 21.7-hectare plot (Area B) with the variety BMX Zeus IPRO (Brasmax) (Fig. 1). The soil in both areas was classified as a typical Rhodic Ferralsol (Embrapa, 2018).

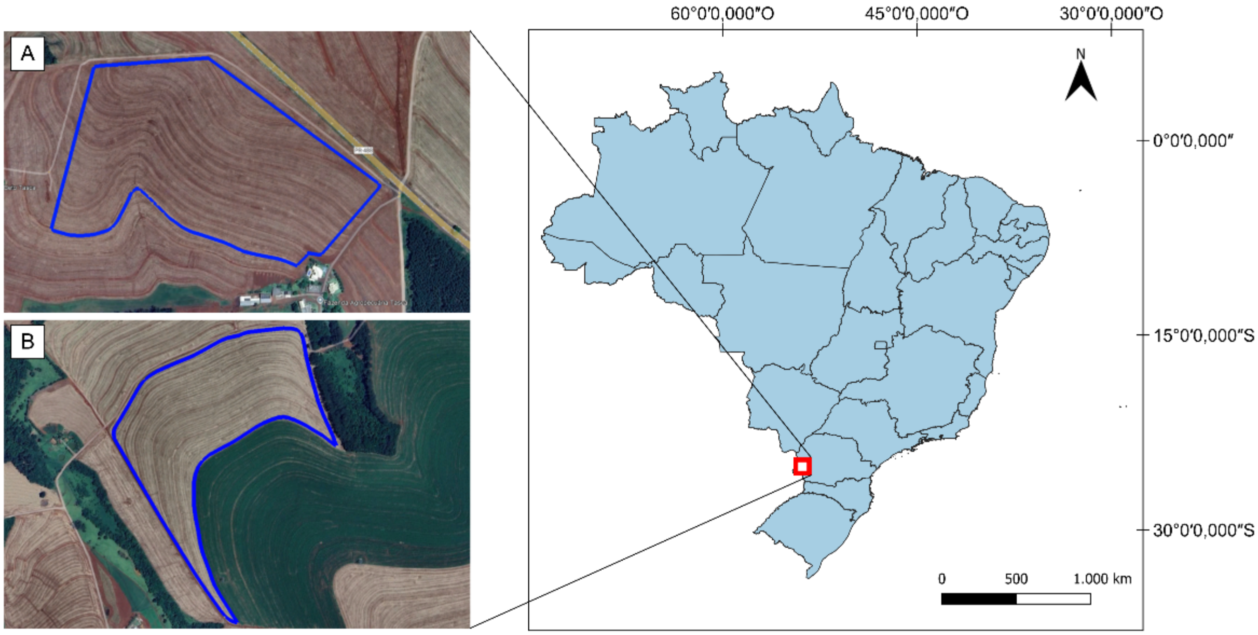

Fig. 1 Location of the sampling areas.

The growing season evaluated was the 2024 second crop (off-season). For the foliar analysis in both cultivation areas, irregular sampling grids were used to establish 106 collection points located along the imaginary centerline between the contour lines present in each area (Fig. 2).

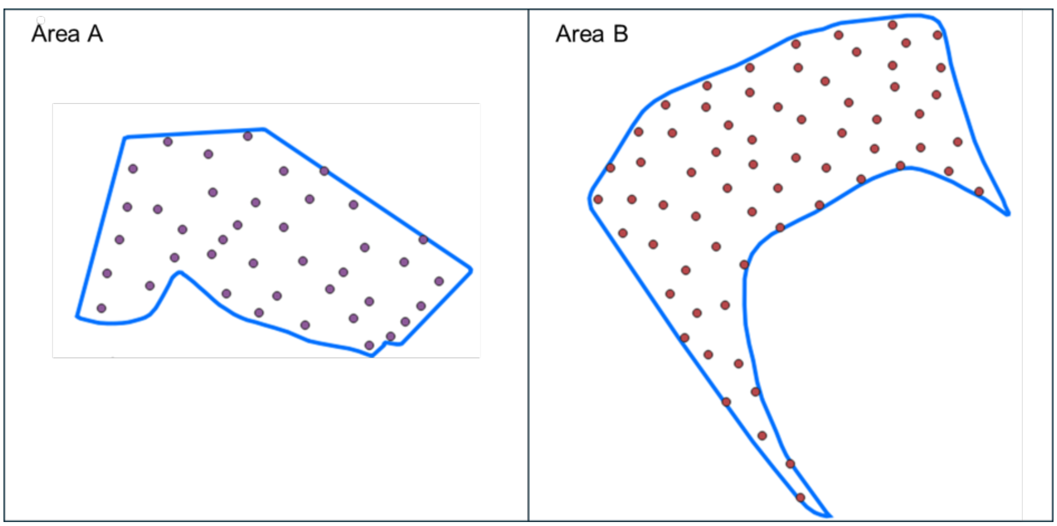

Fig. 2 Sampling grids for analysis of foliage.

Aerial Image Acquisition

Aerial images were acquired using a Nuvem Spectral 2 drone equipped with a Micasense Altum MS sensor capable of imaging in the blue (443 nm to 507 nm), green (533 nm to 587 nm), red (654 nm to 682 nm), red edge (705 nm to 729 nm), and NIR (785 nm to 899 nm) bands.

Flight path planning and flight parameter configuration were performed using the NControl software, provided and recommended by the drone manufacturer. The flights were conducted at an altitude of 120 m with a speed of 12 m/s to obtain images with a pixel size of 5.2 cm and a resolution of 2064 × 1544 in the MS bands. The lateral and vertical image overlap was set to 80%.

A total of four sampling sessions were conducted from crop sowing, during which MS images were acquired at the phenological stages V6: sixth fully developed trifoliate leaf, V8: eighth fully developed trifoliate leaf, and R2: full bloom. The sampling times were defined according to the phenological scale proposed by Fehr et al. (1971).

The dates and corresponding phenological stages for each aerial image sampling of areas A and B during the 2024 second crop are presented in Fig. 3.

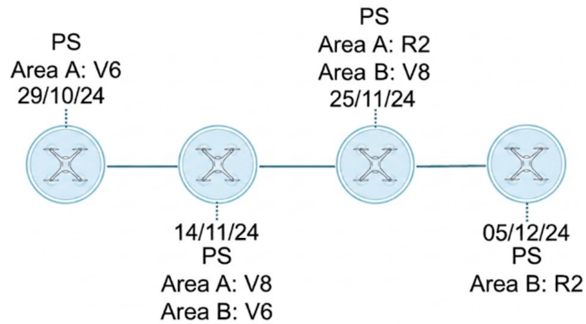

Fig. 3 Dates and phenological stages (PS) corresponding to each aerial image sampling of areas A and B during the 2024 second crop.

Determination of Foliar Nutrient Content

The evaluation of foliar nutrient content was conducted in areas A and B during the 2024 second crop at the R2 phenological stage (full bloom). For this purpose, the third fully developed trifoliate leaf from each plant was collected. At each of the 106 sampling points, 25 trifoliate leaves were collected within a 4 m X 4 m area (Fig. 4).

The 106 composite samples were analyzed at the Plant Tissue and Organic Waste Laboratory (LabSolos) at the UTFPR Pato Branco campus. The procedures described in the Manual of Soil, Plant, and Other Materials Analysis (Tedesco, 1995) were used to determine the contents of N, P, K, Ca, and Mg, while the Manual for Plant Nutritional Status Assessment: Principles and Applications (Malavolta et al., 1989) was followed for the quantification of micronutrients (Cu, Mn, and Zn).

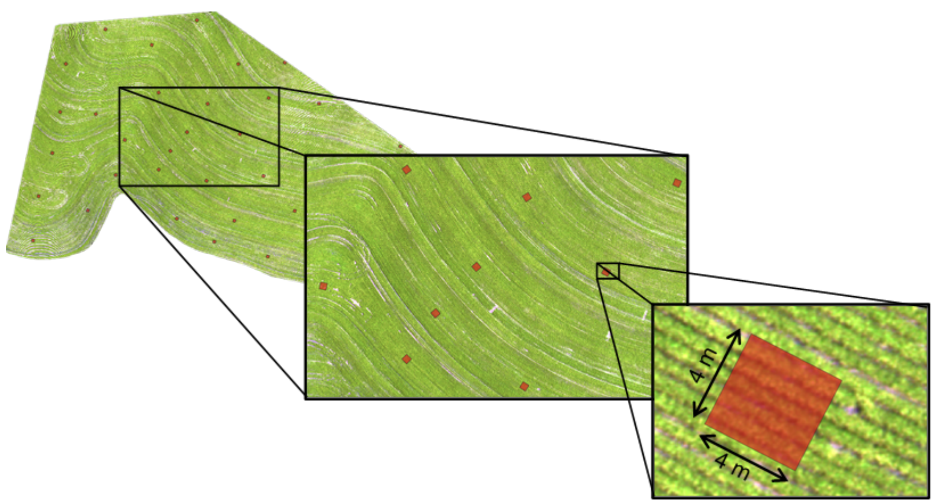

Fig. 4 Region of interest evaluated at each sampling point.

Data Preprocessing

The reflectance maps of the areas at the different phenological stages were generated using Pix4Dmapper 4.6.4 software from the collected multispectral images. The software utilizes algorithms to identify characteristic tie points to align the overlapping images (Xiao et al., 2023). Area boundary clipping, VI calculation, and image feature extraction were performed using QGIS 3.40.11 software.

For the images of all evaluated phenological stages, segmentation was performed by setting a threshold (cutoff value) based on the OSAVI vegetation index. The objective was to exclude soil-related pixels and select only those corresponding to the plants for subsequent analyses.

Calculation of Vegetation Indices (VIs)

The reflectance maps were processed in QGIS 3.40.11 to calculate VIs according to the list and their respective formulas presented in Table 1.

Table 1 Equations for index calculation

Index Name	Abbreviation	Formula
Optimized Soil-Adjusted Vegetation Index	OSAVI	(NIR - Red) / (NIR + Red + 0,16)
Green Normalized Difference Vegetation Index	GNDVI	(NIR - Green) / (NIR + Green)
Normalized Difference Red Edge Index	NDRE	(NIR - RedEdge) / (NIR + RedEdge)
Atmospherically Resistant Vegetation Index	ARVI	(NIR - (2 × Red) + Blue) / (NIR + (2 × Red) + Blue)
Leaf Nitrogen Vegetation Index	LNVI	(NIR / RedEdge) - 1
Modified Simple Ratio	MSR	(NIR / Red - 1)^0.5
Spectral Ratio Index (NIR/Red)	SR1	NIR / Red
Spectral Ratio Index (RedEdge/Red)	SR2	RedEdge / Red
Spectral Ratio Index (RedEdge/Green)	SR3	RedEdge / Green
Spectral Ratio Index (Blue/NIR)	SR4	Blue / NIR
Spectral Ratio Index (Green/NIR)	SR5	Green / NIR
Spectral Ratio Index (Red/NIR)	SR6	Red / NIR
Spectral Ratio Index (RedEdge/NIR)	SR7	RedEdge / NIR
Modified Chlorophyll Absorption in Reflectance Index	MCARI	[(RedEdge - Red) - 0,2 × (RedEdge - Green)] × (RedEdge / Red)
Soil-Adjusted Vegetation Index	SAVI	[(NIR - Red) / (NIR + Red + 0,5)] × (1 + 0,5)
Triangular Vegetation Index	TVI	0,5 × [120 × (NIR - Green) - 200 × (Red - Green)]
Modified Triangular Vegetation Index	MTVI	1,2 × [1,2 × (NIR - Green) - 2,5 × (Red - Green)]
Normalized Difference Vegetation Index	NDVI	(NIR - Red) / (NIR + Red)
Red Edge Position Index / Red Edge Normalized Difference	REP	(RedEdge - Red) / 2
Green Chlorophyll Index	CIGREEN	(RedEdge / Green)-1
False Color Vegetation Index	FCVI	(1.5 × (2×NIR + Blue - 2×Green)) / (2×Green + 2×Blue - 2×NIR + 127.5)
Modified Soil-Adjusted Vegetation Index	MSAVI	0.5 × (2×NIR + 1 - sqrt((2×NIR + 1)^2 - 8×(NIR - Red)))
Enhanced Vegetation Index	EVI	2.5 *((NIR - Red) / (NIR + 6*Red - 7.5*Blue + 1))
Modified Simple Ratio Index	MSR	((NIR/Red)−1)/sqrt((NIR/Red)+1)
Structure Insensitive Pigment Index	SIPI	(NIR - Blue) / (NIR + Blue)
Difference Vegetation Index	DVI	NIR - Red
Normalized Difference Chlorophyll Index	NDCI	(RedEdge - Red) / (RedEdge + Red)
Normalized Red Intensity	NRI	Red / (Red + Green + Blue)
Normalized Green Intensity	NGI	Green / (Red + Green + Blue)
Normalized Blue Intensity	NBI	Blue / (Red + Green + Blue)
Normalized Green-Red Difference	NGMR	(Green - Red) / (Green + Red)
Normalized Green-Blue Difference	NGMB	(Green - Blue) / (Green + Blue)
Normalized Red-Blue Difference	NRMB	(Red - Blue) / (Red + Blue)
Green-Red Difference	GMR	Green - Red
Green-Red Ratio	GDR	Green / Red
Green-Blue Ratio	GDB	Green / Blue
Red-Blue Ratio	RDB	Red / Blue
Normalized NIR Index	NNIR	NIR / (NIR +RedEdge +Green)
Wide Dynamic Range Vegetation Index	WDRVI	((0.1*NIR) - Red) / ((0.1*NIR) + Red)
Chlorophyll Vegetation Index	CVI	(NIR*Red) / Green^2
Green Leaf Index	GLI	(2 * Green - Red - Blue) / (2 * Green + Red + Blue)
Green Atmospherically Resistant Vegetation Index	VARIGREEN	(Green - Red) / (Green + Red + Blue)
Blue Normalized Difference Vegetation Index	BNDVI	(NIR - Blue) / (NIR + Red)
Leaf Chlorophyll Vegetation Index	LCVI	(NIR - RedEdge) / (NIR + Red)

Data Analysis and Algorithms for Foliar Nutrient Estimation

Data processing and analysis to optimize ML models for estimating foliar nutrient levels in the soybean crop were conducted in three steps: preprocessing and feature engineering, selection of the highest-performing models for each nutrient, and hyperparameter optimization.

Feature Extraction Strategies

To verify the influence of the methodology used for dataset construction on the performance of the different algorithms evaluated, five strategies were established for feature extraction at each sampling point from the multispectral images:

- Mean-based: Each observation in the dataset was calculated as the mean of all pixels contained within the area defined for each sampling point, excluding pixels classified as soil (Fig. 5.A)
- Subzone-based: The area of each sampling point was subdivided into sub-areas (0.8 m X 0.8 m squares), and each observation was defined as the mean of the values from each sub-area, excluding pixels corresponding to soil (Fig. 5.B).
- Pixel-based: Each observation was defined as the value corresponding to each pixel within each sampling area, considering only the pixels classified as plants (Fig. 5.C)
- Pixels with OSAVI Index Above the Median (Q2): All pixel values from each sampling area (plants and soil) were extracted, and subsequently, only those with an OSAVI index above the median were selected as observations for the dataset (Fig. 5.D)
- Pixels with OSAVI Index Above the Third Quartile (Q3): All pixel values from each sampling area (plants and soil) were extracted, and subsequently, only those with an OSAVI index above the third quartile (Q3) were selected as observations for the dataset (Fig. 5.E).

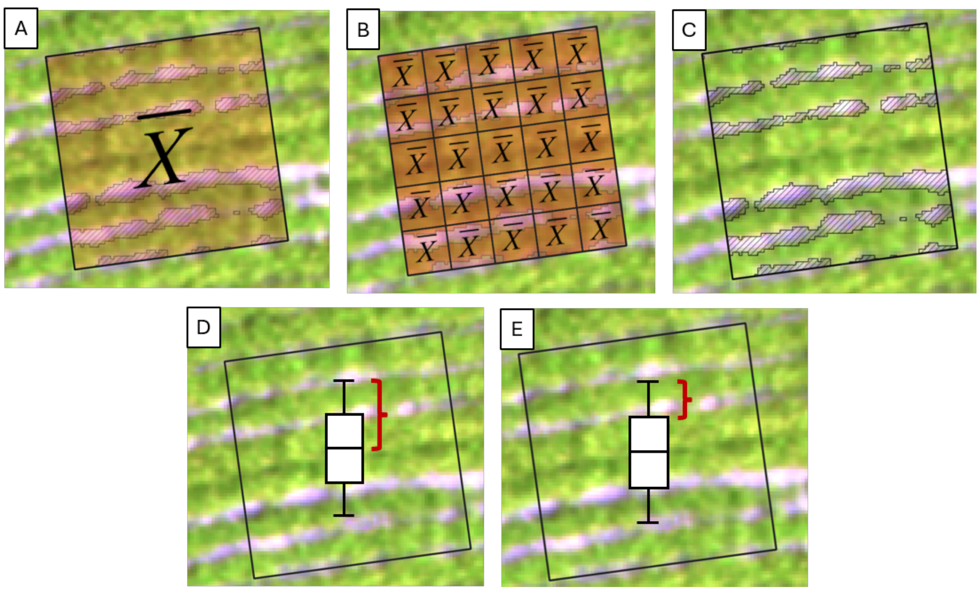

Fig. 5 Feature extraction strategies.

Data processing and predictive modeling

The training and evaluation of the different ML models were performed in a Python 3.12.0 environment. Descriptive data analysis (measures of central tendency, shape, and dispersion) was carried out using the *Pandas* and *NumPy* libraries. For the estimation of foliar nutrient contents, Random Forest (RF), Support Vector Regression (SVR), and Artificial Neural Networks (ANN) algorithms were trained. As input features, 44 vegetation indices and 5 multispectral bands from each of the evaluated phenological stages (V6, V8, and R2) were used. For SVR and ANN, the features were standardized prior to training using the *StandardScaler* class (z-score method) from the *scikit-learn* library. The target variables corresponded to the laboratory-determined foliar nutrient contents (N, P, K, Ca, Mg, Cu, Mn, and Zn). The models were evaluated using the Coefficient of Determination (R^2; Eq. 1), Root Mean Squared Error (RMSE; Eq. 2), and Relative

Root Mean Squared Error (RRMSE; Eq. 3), according to the following equations (Wang et al., 2026):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{1}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \tag{2}$$

$$RRMSE = \frac{RMSE}{\bar{y}} \tag{3}$$

Where y_i represents the concentration of each evaluated nutrient measured in the laboratory for each sample, while $\hat{y}_i$ is the corresponding value predicted by the model for that same point. The variable $\bar{y}$ indicates the arithmetic mean of all actual observed values (quantified in the laboratory) in the sample, and *n* refers to the total number of analyzed observations.

The datasets were partitioned into 70% for training and 30% for testing, with cross-validation using k = 5 folds. To prevent data leakage between the training, validation, and testing sets—particularly in feature extraction approaches that generate multiple features from a single sampling point—the cross-validation was performed using a grouped approach (*GroupKFold*), treating each group as the set of all observations associated with a specific sampling point. The implemented algorithms were *RandomForestRegressor*, *MLPRegressor*, and SVR from the scikit-learn library, utilizing default hyperparameters for the initial performance evaluation stage of the trained models.

Model selection and the highest-performing datasets by nutrient

For selecting and statistically comparing ML algorithms across datasets, the recommendations of Demšar (2006) were followed. Initially, the Friedman test was used as a nonparametric alternative to the repeated-measures ANOVA. This test allows verifying the existence of statistically significant differences in the algorithms' performance based on the average ranking of each model across all datasets. The Friedman statistic was calculated as follows (Eq. 4):

$$\chi_F^2 = \frac{12N}{k(k+1)} * \left(\sum_j R_j^2 - \frac{k(k+1)^2}{4}\right) \tag{4}$$

Where *N* represents the number of evaluated datasets, *k* refers to the total number of algorithms being compared, and R_j is the average rank (mean rank) obtained by the *j-th* algorithm after being evaluated across all scenarios.

From the Friedman statistic, the Iman and Davenport *F*-statistic was calculated, which provides greater statistical power in determining the significance of the differences (Eq. 5):

$$F_F = \frac{(N-1)\chi_F^2}{N(k-1) - \chi_F^2} \tag{5}$$

If the null hypothesis that all algorithms are equivalent was rejected, the Nemenyi post hoc test was used to conduct pairwise comparisons among all evaluated algorithms. This test establishes a Critical Difference (CD); if the difference between the average ranks of two algorithms exceeds this value, they are considered statistically distinct. The significance level established for this test was 5%. The CD was calculated according to the following equation (Eq. 6):

$$CD = q_\alpha * \sqrt{\frac{k(k+1)}{6N}} \tag{6}$$

Where q_a is a statistic that depends on the number of evaluated algorithms (*k*) and the significance level adopted for the test, and *N* represents the number of evaluated datasets.

Hyperparameter Optimization

The hyperparameters of the models trained with the optimal variable set identified in the previous step were optimized using a grid search combined with k-fold cross-validation (k = 5), aiming to maximize the models' R^2. For the RF, the tested hyperparameters were n_estimators: 100 to 1,000 (step of 200), max_depth: None, 10, 20, 30, 40, min_samples_split: 2, 5, 10, 20, min_samples_leaf: 1, 2, 4, 8, and max_features: 'sqrt', 'log2', 'none'.

Results and Discussion

Table 2 presents the mean R^2 values from cross-validation for the three algorithms trained on the five types of generated datasets. Next to each R^2 value, the algorithm's performance rank for the respective dataset is presented in parentheses. When comparing the algorithms' performance across datasets using the Friedman test (Table 2), the F statistic (15.1549) exceeded the critical value (3.114) by nearly 5.

Table 2 Ranking matrix by dataset and regression algorithms for the Friedman test

Dataset	RF	SVR	RN
NxMean	-0,050(2,0)	0,020(1,0)	-7,170(3,0)
PxMean	0,240(1,0)	-0,150(2,0)	-40,630(3,0)
KxMean	0,030(1,5)	0,030(1,5)	-3,330(3,0)
CaxMean	0,130(1,0)	-0,050(2,0)	-4,570(3,0)
MgxMean	0,350(1,0)	-0,010(2,0)	-28,440(3,0)
CuxMean	-0,020(2,0)	0,190(1,0)	-0,600(3,0)
ZnxMean	0,320(1,0)	0,040(2,0)	-2,820(3,0)
MnxMean	-0,540(2,0)	-0,160(1,0)	-2,780(3,0)
NxSubzone	-0,080(2,0)	0,030(1,0)	-0,310(3,0)
PxSubzone	0,400(1,0)	0,040(2,0)	-3,200(3,0)
KxSubzone	0,130(1,0)	0,120(2,0)	0,040(3,0)
CaxSubzone	0,120(1,0)	0,090(2,0)	-0,120(3,0)
MgxSubzone	0,310(1,0)	-0,030(2,0)	-1,360(3,0)
CuxSubzone	0,060(1,0)	0,050(2,0)	-0,320(3,0)
ZnxSubzone	0,280(1,0)	0,220(2,0)	-0,010(3,0)
MnxSubzone	-0,480(2,0)	-0,210(1,0)	-0,580(3,0)
NxPixel	0,040(3,0)	0,100(2,0)	0,120(1,0)
PxPixel	0,380(2,0)	-0,030(3,0)	0,390(1,0)
KxPixel	0,210(1,0)	0,200(2,5)	0,200(2,5)
CaxPixel	0,110(1,5)	0,110(1,5)	0,090(3,0)
MgxPixel	0,260(1,0)	0,160(3,0)	0,220(2,0)
CuxPixel	0,200(1,0)	0,110(3,0)	0,140(2,0)
ZnxPixel	0,310(1,0)	0,220(3,0)	0,300(2,0)
MnxPixel	-0,270(3,0)	-0,220(2,0)	-0,210(1,0)
NxQ2	0,070(2,0)	0,070(2,0)	0,070(2,0)
PxQ2	0,290(1,0)	-0,160(3,0)	0,260(2,0)
KxQ2	0,160(1,0)	0,150(2,5)	0,150(2,5)
CaxQ2	0,140(2,0)	0,120(3,0)	0,160(1,0)
MgxQ2	0,280(1,0)	0,210(3,0)	0,240(2,0)
CuxQ2	0,160(1,0)	0,090(2,5)	0,090(2,5)
ZnxQ2	0,250(1,0)	0,240(2,0)	0,210(3,0)
MnxQ2	0,330(1,0)	-0,290(2,0)	-0,330(3,0)
NxQ3	0,110(1,0)	0,070(2,0)	-0,030(3,0)
PxQ3	0,360(1,0)	-0,090(3,0)	0,180(2,0)
KxQ3	0,240(1,5)	0,240(1,5)	0,010(3,0)
CaxQ3	0,150(1,0)	0,110(2,0)	0,080(3,0)
MgxQ3	0,330(2,0)	0,160(3,0)	0,340(1,0)
CuxQ3	0,330(1,0)	-0,010(3,0)	0,010(2,0)
ZnxQ3	0,130(3,0)	0,230(1,0)	0,140(2,0)
MnxQ3	-0,400(2,0)	-0,290(1,0)	-0,430(3,0)
Average rank	1,438	2,075	2,487

This allowed for firmly rejecting the H_0 of equality in the algorithms' average ranks and, consequently, confirming significant differences in algorithm performance across all evaluated nutrients. The average rank for the RF algorithm (1.438) was the lowest, being 0.64 and 1.05 lower than those of the SVR and ANN algorithms, respectively. The SVR algorithm had an average rank of 2.075, which was 0.412 lower than the NN's average rank (2.487). Therefore, a post hoc test was necessary to determine whether these differences were significant.

The Nemenyi test allowed the establishment of a critical difference of 0.524 for the comparison of average ranks. Fig. 6 shows the results of the test using the diagram proposed by Demšar (2006), in which algorithms with average ranks that do not show significant differences (i.e., whose difference does not exceed the critical value) are connected by a significance bar. In light of this, the RF algorithm had a significantly lower average rank than the other algorithms, demonstrating that it achieved superior performance in estimating the different foliar nutrient contents. On the other hand, SVR and ANN did not show significant differences in performance. When analyzing the R^2 values (Table 2), it can be determined that the greatest performance discrepancies occurred for the smaller datasets, which were structured based on means, as is the case for the *Mean* and *Subzone* datasets, where the maximum performance differences among the algorithms reached up to 40.87, 28.79, and 7.19 R^2 units for the PxMean, MgxMean, and NxMean datasets, respectively, with ANN being the lowest-performing algorithm for these datasets.

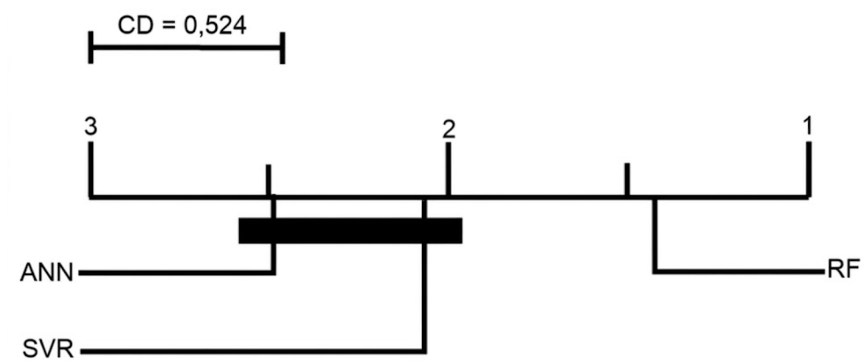

Fig. 6 . Multiple comparison of the algorithms using the Nemenyi test. Algorithms with no significant difference between them (*p > 0.05*) are connected.

RF is recognized for its robustness in identifying nonlinear and complex relationships, which are characteristic of comparisons between spectral data and plant biophysical parameters (Khose and Mailapalli, 2024; Zhang et al., 2025). This robustness stems from its ensemble-based predictions (averaged across multiple decision trees), which enhance its tolerance to noisy data and thereby reduce the risk of overfitting (Khose and Mailapalli, 2024; Han et al., 2025). Furthermore, the algorithm generates models with high stability and interpretability, especially for predicting variables from spectral data (Zhang et al., 2025). Another highlighted attribute of RF is its ability to handle large datasets at high speed and with high efficiency, outperforming other algorithms such as ANNs (Zheng et al., 2018).

With the intention of fully leveraging the potential of the collected data, the present study evaluated four approaches for feature extraction and dataset creation as an alternative to the mean-based methodology widely used in the literature (Jay et al., 2019; Shendryk et al., 2020; Wu et al., 2022; Roth et al., 2023; Khose and Mailapalli, 2024; De Oliveira et al., 2024; Tian et al., 2025; Ge et al., 2025), as detailed in Section Feature Extraction Strategies.

The performance comparison of the algorithms trained on different datasets was performed using the Friedman test (Table 3). The FF value for this comparison (32.822) exceeded the critical value (2.714), indicating significant differences in the performance of the algorithms trained with the different data types. The models trained with the Mean and Subzone datasets consistently maintained the same positions across all nutrients, ranking 5th and 4th, respectively. On the other hand, the models trained on the Pixel, Q2, and Q3 datasets ranked differently across nutrients; however, the Pixel model achieved the best average rank (1.625).

Table 3 Ranking matrix of the feature extraction methods in the Friedman test

Dataset	Mean	Subzone	Pixel	Q2	Q3
N	-2,400(5,0)	-0,120(4,0)	0,090(1,0)	0,070(2,0)	0,050(3,0)
P	-13,510(5,0)	-0,920(4,0)	0,250(1,0)	0,130(3,0)	0,150(2,0)
K	-1,090(5,0)	0,100(4,0)	0,200(1,0)	0,150(3,0)	0,160(2,0)
Ca	-1,500(5,0)	0,030(4,0)	0,100(3,0)	0,140(1,0)	0,110(2,0)
Mg	-9,370(5,0)	-0,360(4,0)	0,210(3,0)	0,240(2,0)	0,280(1,0)
Cu	-0,140(5,0)	-0,070(4,0)	0,150(1,0)	0,110(2,5)	0,110(2,5)
Zn	-0,820(5,0)	0,160(4,0)	0,280(1,0)	0,230(2,0)	0,170(3,0)
Mn	-1,160(5,0)	-0,420(4,0)	-0,230(2,0)	-0,100(1,0)	-0,370(3,0)
Average Rank	5,000	4,000	1,625	2,062	2,312

The Nemenyi test (Fig. 7) demonstrated that all models achieved significantly superior performance compared to those trained with the Mean dataset, except for the Subzone dataset. The latter had an average rank (4) that was significantly lower than that of the model trained with the Pixel dataset (1.625). The models trained with Q2, Q3, and Pixel did not show significant differences in performance.

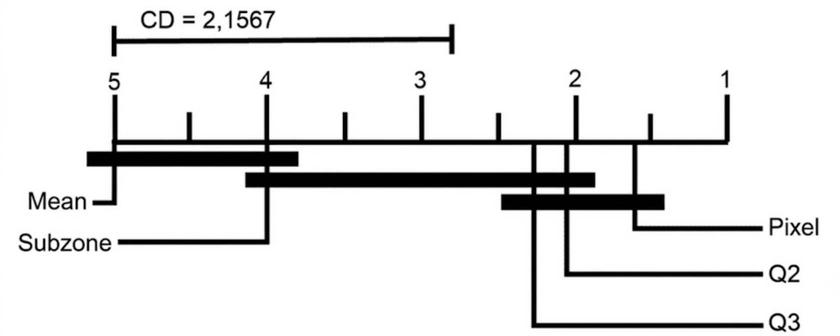

Fig. 7 Multiple comparison of the models trained with different types of datasets using the Nemenyi test. Models with no significant difference between them (*p* > *0.05*) are connected.

It has been discussed that mean-based approaches—namely, those that average all reflectance values within a region of interest, as in the case of the Mean dataset—may lose detailed information and underlying patterns crucial for association with plant biophysical parameters when compared to pixel-based approaches (Tian et al., 2025). However, other authors point out that mean-based approaches minimize the effect of noise present at the pixel level, image misalignment, and plant canopy heterogeneity, thereby improving the reliability of the information (Zhang et al., 2025).

In the present study, the results show that using the pixel-based approach yielded models with better average performance across the three evaluated algorithms (RF, SVR, and NN) than the mean-based datasets (Mean and Subzone) (Fig. 7). This difference becomes more pronounced in algorithms such as NNs. At the same time, the models trained with the Pixel dataset achieved a mean cross-validation accuracy of 0.09-0.39 (excluding Mn), and the models trained with the Mean and Subzone datasets achieved R^2 values ranging from -40.63 to 0.04. This result suggests that the increase in the quantity and detail of information in the Pixel dataset improved the performance of algorithms such as ANNs. Although the latter can, in some cases, achieve accurate predictions on small datasets, larger datasets better leverage the architecture and allow for a more stable convergence of ANN-based models (Benkendorf and Hawkins, 2020).

Since RF was the best-performing algorithm across all datasets, it was used to develop the final models, which were subjected to hyperparameter optimization as described in Section *Hyperparameter Optimization*. The results of the final evaluation using the test set are shown in Fig. 8.

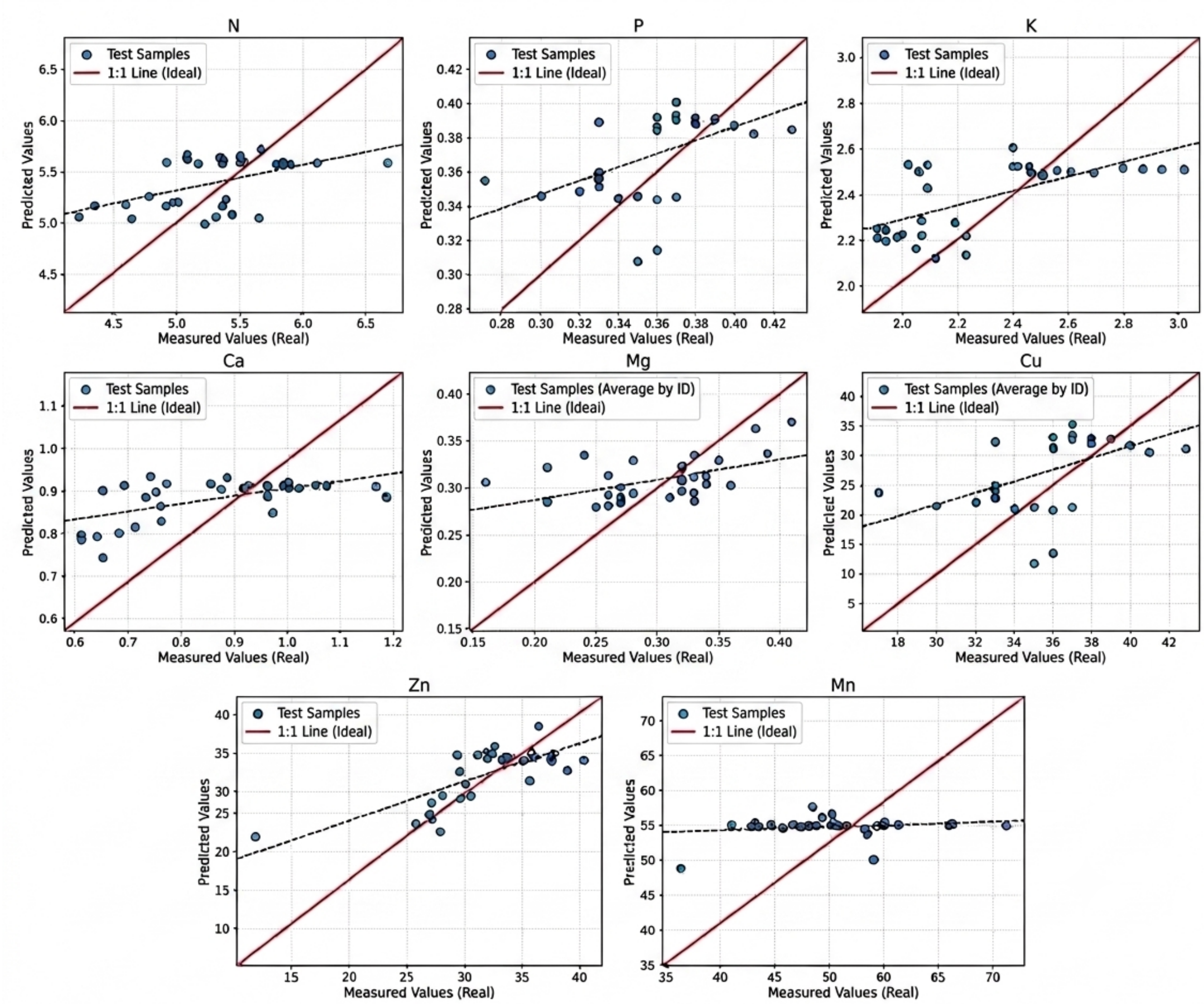

Fig. 8 Final validation of the prediction models using the test set for the different nutrients.

The models evaluated showed variable performance across the analyzed elements, with results ranging from promising to limited. Zinc (Zn) was the best-predicted nutrient, achieving an R^2 of 0.54 and an RRMSE of 12.65%. In comparison, Nitrogen (N) and Phosphorus (P) exhibited the lowest relative errors in the dataset (8.43% and 8.67%, respectively), suggesting that the models captured these elements effectively. These outcomes are encouraging given the inherent complexity of predicting foliar nutrient concentrations from spectral data in a non-destructive manner.

Nevertheless, for most nutrients, R^2 values ranged from 0.21 to 0.35, indicating moderate-to-low predictive capacity and limiting their direct application in precise nutritional diagnosis. Particularly challenging cases include Copper (Cu), with an RRMSE of 28.64%, and Manganese (Mn), with an R^2 of −0.01, suggesting that certain micronutrients present specific difficulties likely associated with their low concentration in foliar tissue and the absence of direct spectral signals in the evaluated regions of the electromagnetic spectrum.

Conclusion

The comparative evaluation of the machine learning algorithms demonstrated that Random Forest (RF) achieved a significantly superior performance compared to Support Vector Regression (SVR) and Artificial Neural Networks (ANN) for estimating foliar nutrient content in soybean, as confirmed by the Friedman and Nemenyi statistical tests. This result is consistent with the recognized robustness of RF in capturing non-linear relationships between spectral data and plant biophysical parameters, particularly when working with noisy datasets.

Regarding feature extraction strategies, the pixel-based approach consistently outperformed the mean-based methodologies, generating models with higher predictive accuracy across all three evaluated algorithms. This finding suggests that preserving spectral variability at the pixel level provides richer information for model training, particularly benefiting algorithms like ANNs, whose architectures are better leveraged with larger, more detailed datasets.

The final RF models, trained on pixel-based data and hyperparameter-optimized, exhibited varying predictive performance across the eight analyzed nutrients. Zinc achieved the most accurate predictions (R^2=0.54, RRMSE=12.65%), while Nitrogen and Phosphorus presented the lowest relative errors (8.43% and 8.67%, respectively), demonstrating the potential of remote sensing-based approaches for monitoring these nutrients. These results establish a solid methodological foundation for future research, where the implementation of complementary feature engineering strategies, such as texture feature extraction or ensemble learning algorithms, could further improve predictive accuracy toward operationally applicable nutritional monitoring tools.

Acknowledgments

The authors express their gratitude to the Federal University of Technology - Parana (UTFPR), the Multi-user Agricultural Technology Laboratory (AGRILAB-MD), the Araucaria Foundation - Paraná, and the CNPq for supporting this work.

References

Benkendorf DJ, Hawkins CP (2020) Effects of sample size and network depth on a deep learning approach to species distribution modeling. Ecological Informatics 60:101137. https://doi.org/10.1016/j.ecoinf.2020.101137

Bertoglio R, Corbo C, Renga FM, Matteucci M (2021) The Digital Agricultural Revolution: A Bibliometric Analysis Literature Review. IEEE Access 9:134762–134782. https://doi.org/10.1109/ACCESS.2021.3115258

Costa L, Kunwar S, Ampatzidis Y, Albrecht U (2022) Determining leaf nutrient concentrations in citrus trees using UAV imagery and machine learning. Precision Agric 23:854–875. https://doi.org/10.1007/s11119-021-09864-1

Cui D, Liang S, Wang D (2021) Observed and projected changes in global climate zones based on Köppen climate classification. WIREs Climate Change 12:e701. https://doi.org/10.1002/wcc.701

De Oliveira PPG, Da Silva-Junior CA, Barron-Zambrano JH, et al (2024) Estimation of macro and micronutrient concentrations in 'valencia' orange leaves at different phenological stages using multispectral image analysis and multiple regression. International Journal of Remote Sensing 45:5506–5529. https://doi.org/10.1080/01431161.2024.2377835

Demšar J (2006) Statistical Comparisons of Classifiers over Multiple Data Sets

Embrapa (2018) Brazilian soil classification system. Embrapa

Fehr WR, Caviness CE, Burmood DT, Pennington JS (1971) Stage of Development Descriptions for Soybeans, *Glycine Max* (L.) Merrill [1]. Crop Science 11:929–931. https://doi.org/10.2135/cropsci1971.0011183X001100060051x

Ge H, Lv G, Qin Y, Shen M (2025) Assessing leaf nitrogen concentration in rice using RGB imaging: a comparative study at leaf, canopy, and plot scales. Front Plant Sci 16:1599177. https://doi.org/10.3389/fpls.2025.1599177

Han Y, Zhang J, Bai Y, et al (2025) Ensemble Learning-Driven and UAV Multispectral Analysis for Estimating the Leaf Nitrogen Content in Winter Wheat

Hassler SC, Baysal-Gurel F (2019) Unmanned Aircraft System (UAS) Technology and Applications in Agriculture. Agronomy 9:618. https://doi.org/10.3390/agronomy9100618

Jay S, Baret F, Dutartre D, et al (2019) Exploiting the centimeter resolution of UAV multispectral imagery to improve remote-sensing estimates of canopy structure and biochemistry in sugar beet crops. Remote Sensing of Environment 231:110898. https://doi.org/10.1016/j.rse.2018.09.011

Karunathilake EMBM, Le AT, Heo S, et al (2023) The Path to Smart Farming: Innovations and Opportunities in Precision Agriculture. Agriculture 13:1593. https://doi.org/10.3390/agriculture13081593

Khose SB, Mailapalli DR (2024) UAV-based multispectral image analytics and machine learning for predicting crop nitrogen in rice. Geocarto International 39:2373867. https://doi.org/10.1080/10106049.2024.2373867

Malavolta E, Vitti GC, Oliveira SA (1989) Avaliação do estado nutricional das plantas ; princípios e aplicações. Potafos

Martos V, Ahmad A, Cartujo P, Ordoñez J (2021) Ensuring Agricultural Sustainability through Remote Sensing in the Era of Agriculture 5.0

Mmbando GS (2025) Harnessing artificial intelligence and remote sensing in climate-smart agriculture: the current strategies needed for enhancing global food security. Cogent Food & Agriculture 11:2454354. https://doi.org/10.1080/23311932.2025.2454354

Omia E, Bae H, Park E, et al (2023) Remote Sensing in Field Crop Monitoring: A Comprehensive Review of Sensor Systems, Data Analyses and Recent Advances. Remote Sensing 15:354. https://doi.org/10.3390/rs15020354

Rahali L, Praticò S, Lanucara S, Modica G (2025) CubeSat constellations: New era for precision agriculture? Computers and Electronics in Agriculture 230:109764. https://doi.org/10.1016/j.compag.2024.109764

Roth RT, Chen K, Scott JR, et al (2023) Prediction of Cereal Rye Cover Crop Biomass and Nutrient Accumulation Using Multi-Temporal Unmanned Aerial Vehicle Based Visible-Spectrum Vegetation Indices. Remote Sensing 15:580. https://doi.org/10.3390/rs15030580

Shendryk Y, Sofonia J, Garrard R, et al (2020) Fine-scale prediction of biomass and leaf nitrogen content in sugarcane using UAV LiDAR and multispectral imaging. International Journal of Applied Earth Observation and Geoinformation 92:102177. https://doi.org/10.1016/j.jag.2020.102177

Tedesco MJ (1995) Análise de solo, plantas e outros materiais., 2nd edn. Departamento de Solos, Universidade Federal do Rio Grande do Sul, Porto Alegre

Tian F, Zhou J, Ransom CJ, et al (2025) Estimating corn leaf chlorophyll content using airborne multispectral imagery and machine learning. Smart Agricultural Technology 10:100719. https://doi.org/10.1016/j.atech.2024.100719

Wu Y, Al-Jumaili S, Al-Jumeily D, Bian H (2022) Prediction of the Nitrogen Content of Rice Leaf Using Multi-Spectral Images Based on Hybrid Radial Basis Function Neural Network and Partial Least-Squares Regression. Sensors 22:8626. https://doi.org/10.3390/s22228626

Zhang X, Hu Y, Li X, et al (2025) Estimation of Rice Leaf Nitrogen Content Using UAV-Based Spectral–Texture Fusion Indices (STFIs) and Two-Stage Feature Selection. Remote Sensing 17:2499. https://doi.org/10.3390/rs17142499

Zheng H, Li W, Jiang J, et al (2018) A Comparative Assessment of Different Modeling Algorithms for Estimating Leaf Nitrogen Content in Winter Wheat Using Multispectral Images from an Unmanned Aerial Vehicle. Remote Sensing 10:2026. https://doi.org/10.3390/rs10122026