

---

# Cross-Domain Transferability of Foliar Nitrogen Prediction in Sugarcane (*Saccharum officinarum*) Through the Integration of UAV and Simulated Spectral Data

---

[Izabelle de Lima e Lima](#)\*, [Marta Laura de Souza Alexandre](#), [Ana Karla da Silva Oliveira](#), [Rodnei Rizzo](#), [Carlos Augusto Alves Cardoso Silva](#), [Peterson Ricardo Florio](#)

Posted Date: 1 April 2026

doi: 10.20944/preprints202604.0003.v1

Keywords: spectral simulation; remote sensing; machine learning



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Cross-Domain Transferability of Foliar Nitrogen Prediction in Sugarcane (*Saccharum officinarum*) Through the Integration of UAV and Simulated Spectral Data

Izabelle de Lima e Lima \*, Marta Laura de Souza Alexandre, Ana Karla da Silva Oliveira, Rodnei Rizzo, Carlos Augusto Alves Cardoso Silva and Peterson Ricardo Fiorio

Department of Biosystems Engineering, "Luiz de Queiroz" College of Agriculture, University of São Paulo, Piracicaba 13418900, SP, Brazil

\* Correspondence: eng.izabellelima@usp.br

## Highlights

### What are the main findings?

- The proposed methodological approach was crucial for the consistency of the results, demonstrating robust predictive performance for nitrogen content, with  $R^2 = 0.75$  (PLSR) and  $0.76$  (RF) for UAV data, and  $R^2 = 0.75$  (PLSR) and  $0.74$  (RF) for simulated data.
- Even with a reduced spectral range, the simulated data retained high predictive power, with indices such as ChlRe ( $R^2 = 0.74$ ) and NDVI ( $R^2 = 0.71$ ), indicating that information relevant for N estimation was preserved.

### What are the implications of the main findings?

- This approach validates the use of spectral simulation as a strategy for transferring information from the hyperspectral domain to multispectral sensors without any significant loss of performance.
- The results underscore the potential for application in operational settings with limited data, enhancing the feasibility of nutritional monitoring via UAV using more generalizable models.

## Abstract

Remotely Piloted Aircraft (RPAs) equipped with multispectral sensors have emerged as promising tools for estimating foliar nitrogen content (FNC). In this context, this study applied a methodological approach aimed at simulating UAV multispectral data using hyperspectral leaf data obtained in a controlled environment, with the objective of evaluating its predictive potential and its transferability to field data collected by UAVs for FNN estimation. To this end, spectral bands and indices equivalent to those of UAV-mounted sensors were simulated based on hyperspectral data acquired by a benchtop sensor, and subsequently used in modeling via Partial Least Squares Regression (PLSR) and Random Forest (RF). The results showed similar performance across the levels, with  $R^2$  values of  $0.75$  and  $0.76$  for PLSR and RF on the UAV data, and  $0.75$  and  $0.74$  for PLSR and RF on the simulated data, respectively. The RF model also performed well in cross-domain validation, with  $R^2 = 0.70$  when calibrated with simulated data and applied to UAV data. Furthermore, the simulated data maintained high predictive power even with a reduced sample size. It is concluded that spectral simulation constitutes a viable strategy for expanding the applicability of nutritional monitoring using multispectral sensors.

**Keywords:** spectral simulation; remote sensing; machine learning

## 1. Introduction

Nitrogen (N) plays a central role in the development of sugarcane (*Saccharum officinarum* L.), directly influencing vegetative growth and yield efficiency [1,2]. However, efficient N management is complex, as its availability in the soil is conditioned by dynamic processes of absorption, leaching, and volatilization [2]. Within the plant, nitrogen availability is influenced by physiological and environmental gradients, including the light gradient across the canopy, with a tendency for leaf nitrogen content (LNC) to decrease from upper to lower leaves [3]. Although laboratory analyses of leaf tissue remain the gold standard for nutritional diagnosis, such methods are destructive, costly, and have low scalability, which limits their application in large areas and in frequent monitoring programs.

Agriculture has increasingly incorporated monitoring technologies capable of supporting strategic decisions at the operational level. In this context, remote sensing (RS) offers a non-destructive and scalable alternative for estimating TFN in sugarcane [3,4]. Remotely Piloted Aircraft (UAVs), equipped with multispectral sensors, stand out as agile tools for acquiring reflectance data at different wavelengths, enabling spatial mapping of crop nutritional status [4]. Additionally, laboratory and field hyperspectral sensors, by recording hundreds of narrow bands, also offer high spectral resolution, increasing sensitivity to capture subtle variations in reflectance associated with NSP at different acquisition levels [5].

A significant strategy that has been gaining traction is the use of hyperspectral data to simulate multispectral bands equivalent to widely used sensors, such as those on the Sentinel-2A and PlanetScope satellites and the DJI Phantom 4 Multispectral [6,7]. This simulation, typically performed via spectral convolution, allows for the reproduction of spectral responses from operational sensors and enables the use of spectral correction and standardization techniques across instruments, such as direct standardization and channel-based standardization, with the potential to reduce discrepancies between spectra and enhance comparability across scales and acquisition platforms [7,8]. In addition to facilitating methodological transfer between the laboratory and the field, this approach reduces the dimensionality of the hyperspectral dataset, mitigating spectral redundancy effects inherent in adjacent bands [9].

In practical applications, the literature has emphasized the use of spectral indices (SI) derived from combinations of specific bands, capable of summarizing optical properties related to both pigment absorption in the visible region and structural scattering in the near-infrared (NIR), including bands in the RedEdge region, such as the NDVI, VARI, and Clre [10–12]. These indices have demonstrated a consistent association with plant biochemical and physiological parameters, particularly chlorophyll and N [13,14], and are widely used as sensitive indicators of variations in plant physiological status [4,15].

At the same time, recent advances in machine learning (ML) methods, such as Random Forest (RF) and Partial Least Squares Regression (PLSR), have expanded the ability to model nonlinear relationships between spectral variables and nutritional attributes [16]. Fiorio et al. [17] reported robust performance of partial least squares regression (PLSR) models integrated with spectral reflectance, with  $R^2$  values exceeding 0.80 in predicting foliar N. Similarly, Kumarasiri et al. [18] observed consistent performance of PLSR and Random Forest (RF) with multispectral UAV images, achieving  $R^2$  values of up to 0.72 and 0.66 in the validation set, respectively.

Despite these advances, evidence regarding the equivalence and transferability between simulated leaf data from hyperspectral sensors under controlled conditions and data obtained by multispectral sensors in flight remains limited, especially when evaluated within the same experimental design and under validation strategies that test generalizability across domains (laboratory and field). This methodological gap hinders the consolidation of integrated approaches that use the greater sensitivity of hyperspectral data to calibrate, in a reproducible manner, multispectral sensors applied to canopy-scale monitoring.

In this context, the present study aims to evaluate the consistency between spectral bands and indices obtained from simulated multispectral bands based on hyperspectral data acquired by a benchtop sensor and those obtained by an airborne multispectral sensor. Additionally, we seek to

compare the performance of the PLSR and RF models in estimating foliar nitrogen content (FNC), considering intra-domain and inter-domain validations, as well as the spatial distribution of nitrogen in the experimental area.

## 2. Materials and Methods

The methodology adopted in this study was structured into eight main stages: (i) setting up and conducting the field experiment; (ii) acquisition of multispectral data using UAVs; (iii) acquisition of hyperspectral and laboratory data (Leaf +1); (iv) preprocessing and extraction of spectral signatures in QGIS; (v) standardization and preprocessing of hyperspectral data; (vi) simulation of multispectral bands from hyperspectral leaf data; (vii) calculation of spectral indices and predictive modeling of leaf nitrogen content; and (viii) prediction and spatialization of nitrogen in the experimental area.

This methodological workflow enabled the integration of reflectance data obtained by a multispectral sensor mounted on a UAV, benchtop hyperspectral data, and laboratory analyses, allowing for the evaluation of the potential of spectral simulation regarding cross-domain calibration and the generalization of predictive nitrogen models in sugarcane. The study steps are presented in Figure 1.

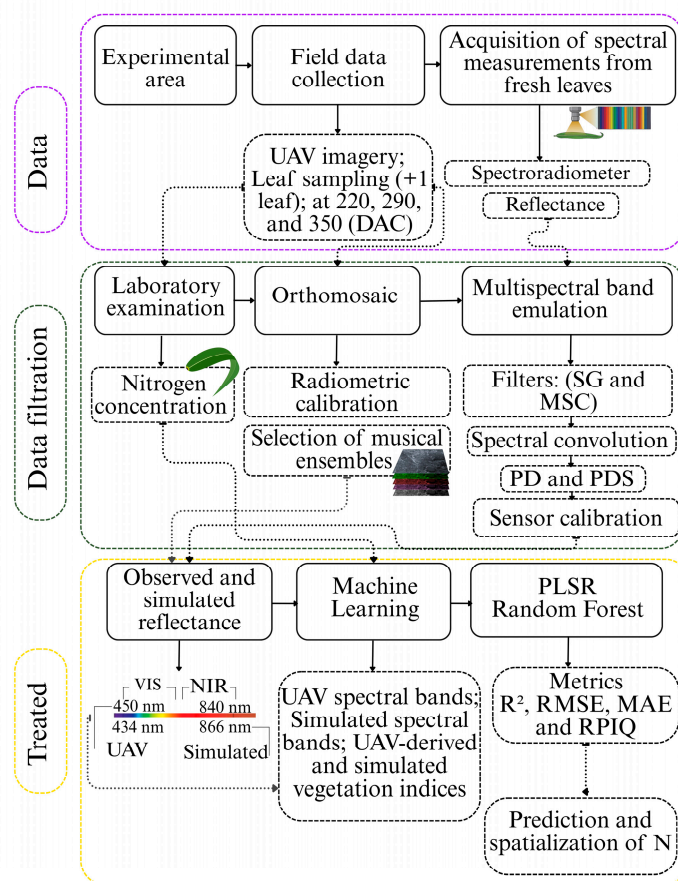
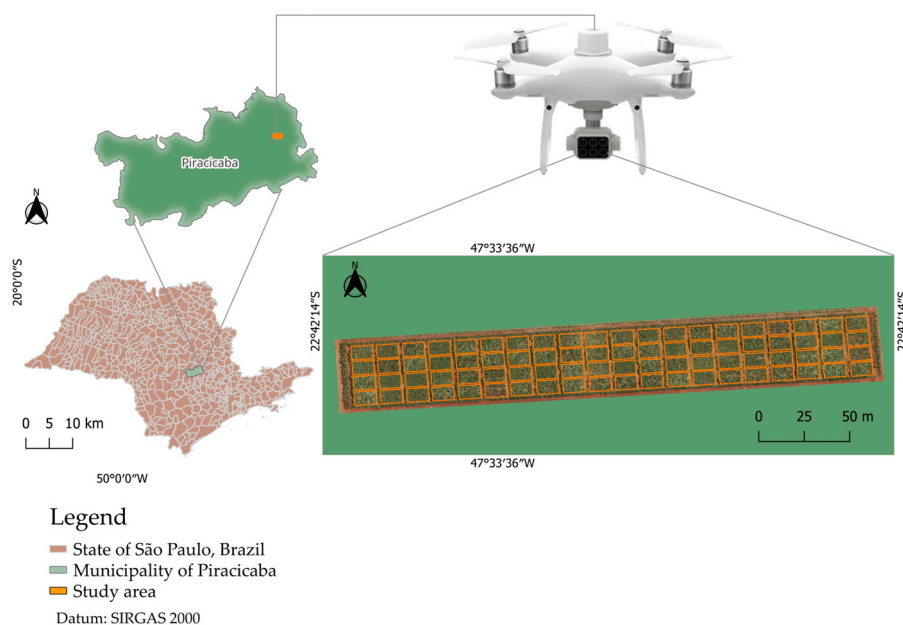


Figure 1. Methodological Workflow.

### 2.1. Study Area and Experimental Conditions

The experiment was conducted at a research site located at the Sugarcane Technology Center (CTC) in the municipality of Piracicaba, state of São Paulo, at geographic coordinates 22°42'15" south latitude and 47°33'34" west longitude, as shown in Figure 2. The soil in the experimental area is classified as Red-Yellow Argisol (PVA), the region's climate is humid subtropical (Cwa), according to the Köppen-Geiger classification, featuring hot and rainy summers and dry winters, with an average annual precipitation of less than 1,400 mm [19].



**Figure 2.** Location map of the experimental area, situated in the municipality of Piracicaba, São Paulo, Brazil, within the research facilities of the CTC.

## 2.2. Field Experimental Design

The experiment was conducted using a factorial design with four blocks, five sugarcane varieties (CT01, CT02, CT03, CT04, and CT05), and four treatments, totaling 80 plots, in their third year of growth. Four sources of N variation were used: control (T1 – no application), conventional (T2 – 1.0 kg t<sup>-1</sup>), optimized 1 (T3 – 1.2 kg t<sup>-1</sup>), and optimized 2 (T4 – 1.2 kg t<sup>-1</sup>). Each plot consisted of six rows, 12 m long and 6 m wide.

## 2.3. Acquisition of Multispectral Images Using a UAV

The images were acquired using a multispectral sensor mounted on the DJI Phantom 4 Multispectral, P4M RTK unmanned aerial vehicle (UAV) (DJI Technology, Shenzhen, China). The system consists of five multispectral cameras covering the blue (B), green (G), red (R), red edge (RE), NIR, and RGB bands (Table 1). Additionally, the equipment features an RTK (Real-Time Kinematic) module, which ensures greater accuracy in the georeferencing of the collected data.

**Table 1.** Characteristics of the multispectral sensor mounted on the UAV, including central wavelength (nm) and bandwidth (nm).

Band ( $\lambda$ )	Central Wavelength (nm)	Bandwidth (nm)
Blue (B)	450	$\pm 16$
Green (G)	560	$\pm 16$
Red (R)	650	$\pm 16$
Red-edge (RE)	730	$\pm 16$
Near-infrared (NIR)	840	$\pm 26$
RGB	20 megapixels	

Three flights were conducted 220, 290, and 350 days after cutting (DAC), between 10:00 a.m. and 11:00 a.m., in accordance with the established flight plan. The images were acquired at an altitude of 50 m, resulting in a ground sample distance (GSD) of 3.0 cm, with 80% forward overlap and 70% side overlap. For georeferencing, in addition to RTK correction, ground control points with known

coordinates were used, obtained using a Trimble R4 RTK receiver (Trimble Inc., Sunnyvale, California, USA). The radiometric consistency of the reflectance data was ensured through the use of a calibration plate, which was recorded by the sensor immediately before and after each flight.

#### 2.4. Acquisition of Leaf Samples

Leaf sampling was conducted by collecting six fresh sugarcane leaves, selecting the +1 leaf, which corresponds to the first fully expanded leaf [17]. Leaves were chosen at random within each plot, excluding those at the edges. After collection, the leaves were placed in plastic bags and transported in coolers with ice to the laboratory, where hyperspectral measurements were performed. This technique was adopted to preserve the turgidity and spectral properties of the leaves [20,21]. The measurements were taken on the middle third of each leaf. After obtaining the spectra, the leaves were sent for laboratory leaf analysis to determine TFN. The collections were made during the 2024/2025 growing season, when the crop was in its third year of cultivation (replanted stand). Leaf samples were collected on the same day as the drone flights, corresponding to 220, 290, and 350 days after cutting (DAC), in order to standardize multispectral and hyperspectral sampling.

#### 2.5. Acquisition of Laboratory Hyperspectral Data

The hyperspectral spectra were obtained using the FieldSpec® 3 benchtop spectroradiometer [22]. The equipment operates in the 350–2500 nm range, divided into the visible (VIS) and near-infrared (NIR) regions with a 1.4 nm sampling interval, and the short-wave infrared (SWIR) region with a 2 nm interval.

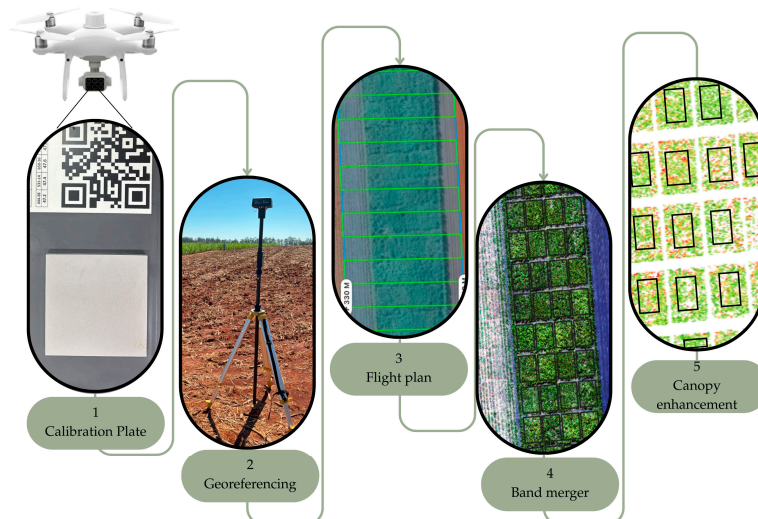
Measurements were conducted in a controlled environment using the Leaf Clip® accessory [22], which ensures orthogonal geometry and constant light intensity, minimizing variations caused by sample positioning or the angle of incidence. To ensure data consistency, radiometric calibration was performed every 15 minutes, using the Lambertian surface integrated into the Leaf Clip® as a reference standard, following ASD protocols and the methodology adopted by Fiorio et al. [17].

#### 2.6. Spectral Data Preprocessing

##### 2.6.1. Calibration and Extraction of UAV Multispectral Data

A total of 3,870 aerial photographs were acquired during the three data collection campaigns. The files were downloaded and processed using photogrammetry software based on the Structure from Motion (SfM) technique, Agisoft Metashape Professional Edition 2.0.1 [23]. The processing involved three-dimensional reconstruction and the generation of an orthomosaic of the study area. To ensure the spectral consistency of the images, radiometric correction was performed by inserting the photograph of the calibration plate recorded before and after each flight. This procedure allowed for the standardization of reflectance values and resulted in the generation of an orthomosaic for each of the five corrected spectral bands, corresponding to the multispectral sensor mounted on the UAV.

To isolate the spectral signature of the sugarcane canopy and reduce ground noise, in accordance with the procedures established by Picado et al. [24], the bands were imported into Quantum GIS [25] version 3.40. Initially, the Normalized Difference Vegetation Index (NDVI) was calculated to highlight vegetation response and minimize interference from exposed soil and shading. For this calculation, a threshold of 0.65 was adopted, with pixels having higher values considered active vegetation. This cutoff value, as reported by Wang et al. [26], proved effective in highlighting the active canopy and enhancing the accuracy of the spectral response. The spectral bands were cropped using the mask generated by the NDVI in QGIS. Next, a buffer was applied to delineate only the usable area of the plot, eliminating the borders. Finally, zonal statistics were performed, which allowed us to obtain, through tabulation, the median reflectance values corresponding to each spectral band in each plot. As illustrated in Figure 3.

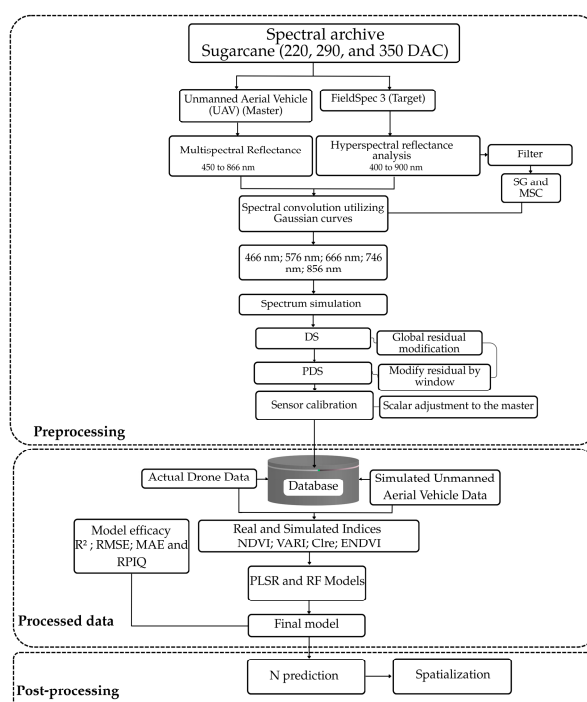


**Figure 3.** Methodological schedule for data extraction. .

### 2.6.2. Standardization of Hyperspectral Data

The spectra were initially recorded in radiance levels (raw values) and converted to reflectance using ViewSpec Pro software (version 5.7). To represent each plot, the median of the leaf readings was used. In the next step, the reflectance spectra underwent a quality control and standardization process prior to band simulation. In this process, responses with evident noise or values outside the interquartile range of each plot were eliminated. Additionally, spectral cropping was performed to exclude unstable regions, such as the range below 400 nm, characterized by a low signal-to-noise ratio, and the sections adjacent to the detector junctions (~1000 and ~2500 nm), which do not correspond to the wavelengths parameterized by the Phantom 4 Multispectral sensor.

### 2.7. Methodological Sequence of the Simulation and Modeling Technique



**Figure 4.** Flowchart of the spectral calibration process between hyperspectral (FieldSpec) and multispectral (Phantom 4) data. The hyperspectral data undergoes preprocessing involving Savitzky-Golay (SG) filtering and

Multiplicative Signal Correction (MSC), followed by spectral convolution using Gaussian curves, DS and PDS were applied, along with cross-sensor calibration and scalar correction to the master, with evaluation based on performance metrics ( $R^2$ , RMSE, MAE, RPIQ).

### 2.7.1. Spectral Convolution

For each plot, reflectance was represented by the median of the readings, generating a representative hyperspectral signature for each plot. Next, the spectra were cropped to the 400–900 nm range, with a continuous spectral resolution of 1.4 nm, in order to reduce interference outside the region of interest, standardize the analyzed spectral domain, and enable more consistent simulations across sensors. Subsequently, spectral preprocessing procedures were applied to minimize instrumental noise and light interference, using Multiplicative Signal Correction (MSC) and the Savitzky-Golay (SG) filter [27], implemented in RStudio [28] (version 4.5.0). MSC was employed to correct additive and multiplicative variations associated with light scattering among the samples, while the SG filter was used to smooth the spectral signatures, preserving the overall behavior of the reflectance curve. The MSC formulas are presented in Equations (1), (2), and (3).

$$\text{Spectra}_{\text{avg}} = \frac{\sum_1^n \text{Spectra}_i}{n} \quad (\text{Eq. 1})$$

$$\text{Spectra}_i = K_i \times \text{Spectra}_{\text{avg}} + b_i \quad (\text{Eq. 2})$$

$$\text{Spectra}_{\text{MSC},i} = \frac{\text{Spectra}_i - b_i}{k_i} \quad (\text{Eq. 3})$$

Where:

**Spectra<sub>avg</sub>** corresponds to the average of all leaf spectra;

**n** represents the total number of spectral data points;

**Spectra<sub>i</sub>** refers to each individual leaf spectrum;

**k<sub>i</sub>** and **b<sub>i</sub>** are correction coefficients obtained by linear regression from Spectra<sub>avg</sub>;

**Spectra<sub>MSC,i</sub>** corresponds to the spectrum corrected using the MSC method.

Next, spectral convolution was applied to reproduce the response of a multispectral sensor from the hyperspectral data, a procedure widely used in cross-validation and time-series integration [29]. For the DJI Phantom 4 Multispectral (P4M), the simulated reflectance of each band was calculated as a weighted average of the FieldSpec® 3 reflectance using spectral response functions (Equation 4), approximated by Gaussian curves, parameterized by the central wavelengths and full width at half maximum (FWHM) values provided by the manufacturer (Blue  $450 \pm 16$  nm; Green  $560 \pm 16$  nm; Red  $650 \pm 16$  nm; RedEdge  $730 \pm 16$  nm; NIR  $840 \pm 26$  nm).

It should be noted that the spectral response functions (SRFs) of the full P4M bands are not published as detailed normalized curves by the manufacturer; therefore, the Gaussian approximation constitutes a methodological limitation. To mitigate potential biases resulting from this simplification and to align instruments, we applied Direct Standardization (DS) and Window-Based Direct Standardization (PDS) after the simulation.

$$R_b = \frac{\sum_{i=1}^n R(\lambda_i) \text{SRF}_b(\lambda_i) \Delta\lambda_i}{\sum_{i=1}^n \text{SRF}_b(\lambda_i) \Delta\lambda_i} \quad (\text{Eq. 4})$$

Where:

$R_b$  is the simulated reflectance of band;

$b$ ,  $\lambda_i$  are the wavelengths measured by the FieldSpec;

$\Delta\lambda_i$  the spectral resolution (1.4 nm), and the denominator ensures normalization, and

$\text{SRF}_b(\lambda_i)$  corresponds to the spectral response function of band  $b$ .

### 2.7.2. Transfer Calibration Using (DS/PDS)

Cross-calibration between spectral domains was performed using Direct Standardization (DS) and Window Direct Standardization (PDS), classical calibration transfer techniques [30,31]. The DS method is based on establishing a global spectral transfer matrix derived from the mathematical relationship between the spectra of a set of standard samples measured by the master instrument and the spectra measured by the target instrument.

On the other hand, PDS is better at capturing variations dependent on the local wavelength, i.e., by windows. From these coefficients, a transfer matrix is constructed and applied to the spectra of the target instrument in order to maximize similarity with the spectra of the reference (master) instrument.

In the context of Direct Partial Standardization (DPS), the window width ( $\omega$ ) serves as a central hyperparameter, responsible for balancing the trade-off between local accuracy associated with narrow windows and numerical robustness favored by wider windows; it is therefore defined empirically [32]. In this work, the methodological strategy consisted of integrating DS and PDS, adopting a window width corresponding to five wavelengths, aligned with the number of evaluated bands.

This choice allowed for consistent local corrections, reduced the dimensionality in each window, and contributed to greater numerical stability. Thus, the spectral segmentation of the PDS maintained the flexibility inherent in the DS, while minimizing the number of hyperparameters to be adjusted [33].

In this study, the multispectral sensor mounted on the UAV was defined as the master, since it serves as the spectral reference and the target calibration scale. The actual bands captured by the UAV represent the comparison standard, such that all transformations aim to align the simulated spectra with this domain. Thus, the quality of the standardization process was evaluated based on the correspondence between the corrected data and the original measurements of the master.

The hyperspectral spectrum obtained in the laboratory, after convolution with the estimated spectral response functions (SRFs), was defined as the target, as it represents the source to be adjusted to the master. The sequential application of DS, PDS, and scalar adjustment aims to transform the simulated spectra, reducing instrumental discrepancies and maximizing similarity with the actual bands of the UAV. Thus, the (hyperspectral) target corresponds to the corrected spectral domain, whose performance is continuously evaluated relative to the master (UAV).

### 2.7.3. Sensor-to-Sensor Calibration

After applying the standardization techniques, sensor-to-sensor calibration was performed as a complementary step to adjust individual scales for each band, using the same anchor samples (plots) employed in the calibration process. This procedure aimed to definitively align the simulated values with the scale of the UAV multispectral sensor, correcting residual differences in both offset and scale. As a result, the transformed spectra showed greater quantitative agreement with the actual UAV measurements.

### 2.7.4. Nitrogen-Sensitive Vegetation Indices

The spectral bands extracted from the sugarcane orthomosaic were used to calculate spectral indices (SI), defined as transformations obtained by combining two or more multispectral bands. In this study, the selection of these indices was guided by their higher correlation with foliar nitrogen content (FNC), as they can capture variations related to pigmentation, photosynthetic activity, and canopy structure [34,35].

To ensure comparability between the datasets, the spectral indices were calculated based on the spectral bands extracted from the buffer of each experimental plot. For the simulated data, the same spatial and spectral representation was adopted; that is, the values were organized to correspond to the same plots and the same set of bands considered in the observed data. This ensured standardization of the unit, allowing for consistent comparisons. Furthermore, the same indices were calculated for both domains (Table 2), which ensured methodological uniformity and greater robustness in the evaluation of TNF.

**Table 2.** List of vegetation indices used in the modeling, along with their respective equations and references. The blue, green, red, red-edge, and near-infrared bands are represented by B, G, R, RE, and NIR, respectively.

ID	Vegetation Index	Formula	References
----	------------------	---------	------------

1	Normalized Difference Vegetation Index (NDVI)	$(\text{NIR} - \text{R}) / (\text{NIR} + \text{R})$	[36]
2	Visible Atmospheric Resistance Index (VARI)	$(\text{G} - \text{R}) / (\text{G} + \text{R} - \text{B})$	[37]
3	Chlorophyll Index - Red-Edge (ChlRe)	$(\text{NIR}) / (\text{RED}) - 1$	[38]
4	Improved Normalized Difference Vegetation Index (ENDVI)	$(\text{NIR} - \text{G}) - (2 \times \text{B}) / (\text{NIR} - \text{G}) + (2 \times \text{B})$	[39]

## 2.8. Spearman Correlation

To investigate the relationship between the bands obtained by UAV, the simulated bands, and the spectral indices derived from both sources, we used Spearman's correlation coefficient, a widely used nonparametric technique for measuring the strength and direction of monotonic associations between variables. This coefficient is particularly suitable when the data do not meet the assumptions of normality or when the relationship between variables is not strictly linear, and it is calculated based on the ranking of the observed values. Its values range from  $-1$  to  $1$ , where positive coefficients indicate direct monotonic associations, negative coefficients indicate inverse associations, and absolute values closer to  $1$  represent stronger relationships between the analyzed variables [40,41].

## 2.9. Modeling of Nitrogen Content Using Machine Learning

### 2.9.1. Linear Regression

Linear regression is one of the simplest and most widely used statistical techniques, designed to model the relationship between a dependent variable (IE) and an independent variable (TFN). This technique assumes that there is a linear relationship between the variables and is commonly tested using the  $p$ -value associated with the model's slope coefficient, where values of  $p < 0.001$  indicate that this relationship is statistically significant [42]. Linear regression provides a simple and effective basis for estimating foliar nutrients based on spectral data, allowing for the quantification of the strength and direction of the association between (IE) and TNF.

### 2.9.2. Partial Least Squares Regression (PLSR)

The Partial Least Squares Regression (PLSR) machine learning technique was used with the NIPALS algorithm. This technique was used to estimate nutritional content based on the reflectance of multispectral (canopy) and simulated (foliar) data. PLSR is widely used in predictive analyses with spectral data, as it projects the original variables into a reduced-dimensional space composed of latent variables (or factors). During the calibration process, the model simultaneously integrated information from correlated independent variables (spectral bands and indices) and dependent variables (N content), seeking to identify the minimum number of factors necessary to explain the variability in the responses. Furthermore, PLSR stands out for its ability to mitigate multicollinearity among independent variables. This characteristic is especially relevant in hyperspectral data, where the strong correlation between adjacent bands compromises the application of conventional methods. In addition to overcoming this challenge, PLSR extracts latent components that condense the variability of the spectra and simultaneously explain multiple variables of interest, resulting in robust and stable models, as has already been demonstrated in various studies [4,43].

Excessive use of latent components may lead the model to interpret noise as relevant signal, resulting in inferior predictive performance, with increasing overfitting effects as more factors are included than necessary [44]. Thus, a well-balanced model should combine quality and predictive power, making it suitable for use in remote sensing studies focused on plant nutrition. To determine

the optimal number of factors, K-fold cross-validation was applied with the aim of minimizing the “Root Mean PRESS” statistic. In addition, a variable importance in projection (VIP) analysis was performed, calculated based on the loading weights and the variance explained by the selected components, allowing for the identification of the variables that contributed most to nitrogen prediction in the model’s latent space.

### 2.9.3. Random Forest (RF)

The Random Forest (RF) model was applied; it is widely used in high-dimensional data processing and is one of the most commonly adopted methodologies for handling hyperspectral data [35]. RF is an algorithm developed by Breiman [45], frequently used independently for regression and classification problems. This method employs a bootstrap aggregation or bagging process, in which each tree is trained on a subsample of the initial dataset. Thus, each tree contributes an individual prediction. The combination of these predictions resulted in the final estimate [46]. During the prediction process, the model used the average or majority of the predictions, ensuring a balance in the error of each tree. This improved the model’s generalization ability [47].

The RF model operates based on the definition of several main parameters: the number of trees (ntree), the number of predictor variables considered for each tree (mtry), and the minimum size of terminal nodes (nodesize) [48,49]. In this study, some of these parameters were manually defined as follows: number of trees (ntree = 1000), number of variables considered at each split (mtry = 11), minimum terminal node size (min.node.size = 20), and split criterion based on variance reduction (splitrule = "variance"). These hyperparameters were defined according to the criteria suggested by Chen et al. [44] in order to avoid unsatisfactory model performance.

### 2.9.4. Performance Indicators

The quality of the models for predicting TNF in sugarcane was assessed by comparing the observed values with those predicted by the models. For these analyses, validation metrics such as the coefficient of determination ( $R^2$ ), root mean square error (RMSE), mean absolute error (MAE), and the ratio of performance to interquartile range (RPIQ) were used, whose mathematical expressions are presented in Equations 5, 6, 7, and 8. According to Rodrigues et al. [50],  $R^2$  classifies prediction models as: poor ( $\leq 0.50$ ), moderate (0.50–0.65), good (0.65–0.80), very good (0.80–0.90), and excellent ( $\geq 0.90$ ).

$$R^2 = \frac{[\sum (\gamma_p - \underline{\gamma}_p) \cdot (\gamma_o - \underline{\gamma}_o)]^2}{[\sum (\gamma_p - \underline{\gamma}_p)^2 \cdot (\gamma_o - \underline{\gamma}_o)^2]} \quad (\text{Eq. 5})$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (\text{Eq. 6})$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |X_i - \hat{X}_i| \quad (\text{Eq. 7})$$

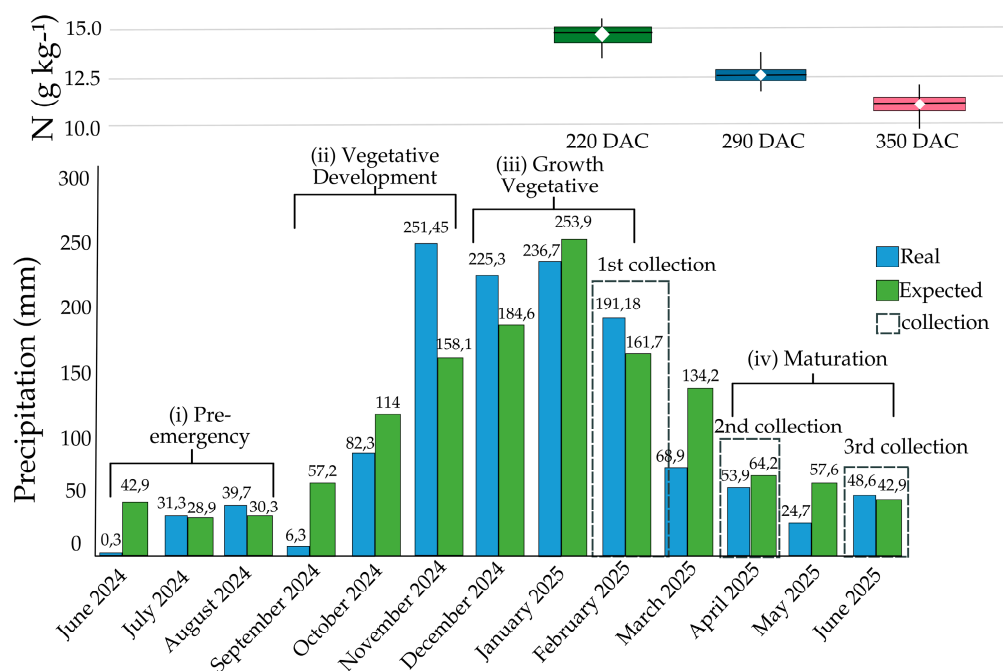
$$\text{RPIQ} = \frac{\text{IQR}_{\text{obs}}}{\text{RMSE}} \quad (\text{Eq. 8})$$

To ensure the robustness and generalization capability of the models, two complementary validation strategies were adopted. The first employed k-fold cross-validation ( $k = 10$ ), in which the dataset of 240 observations was divided into ten balanced subsets; in each iteration, nine subsets were used for training and one for validation, ensuring that all data were tested at least once. The second strategy applied the hold-out method, splitting the data into 70% for training and 30% for independent testing. This combined approach enables a comprehensive evaluation of predictive performance while mitigating the risk of bias associated with data partitioning.

### 3. Results

#### 3.1. Leaf Nitrogen Content (LNC) and Precipitation Throughout the Season

Figure 5 shows precipitation throughout the 2024/2025 sugarcane growing season, along with nitrogen data for the respective samples collected at 220, 290, and 350 days after planting. The phenological cycle of sugarcane consists of four stages: (i) pre-emergence from 0 to 90 days, (ii) vegetative development from 90 to 180 days, (iii) vegetative growth from 180 to 270 days, and (iv) maturation from 270 to 360 days. Among these stages, vegetative development is the one with the highest nutrient concentrations in the plantation [24]. A water deficit was observed during the initial months of the 2024 harvest season, from June to October, corresponding to a phase of extreme importance for the crop's vegetative development.



**Figure 5.** Dynamics of leaf nitrogen content and precipitation (expected vs. actual) in sugarcane throughout the crop cycle, evaluated at 220, 290, and 350 DAC.

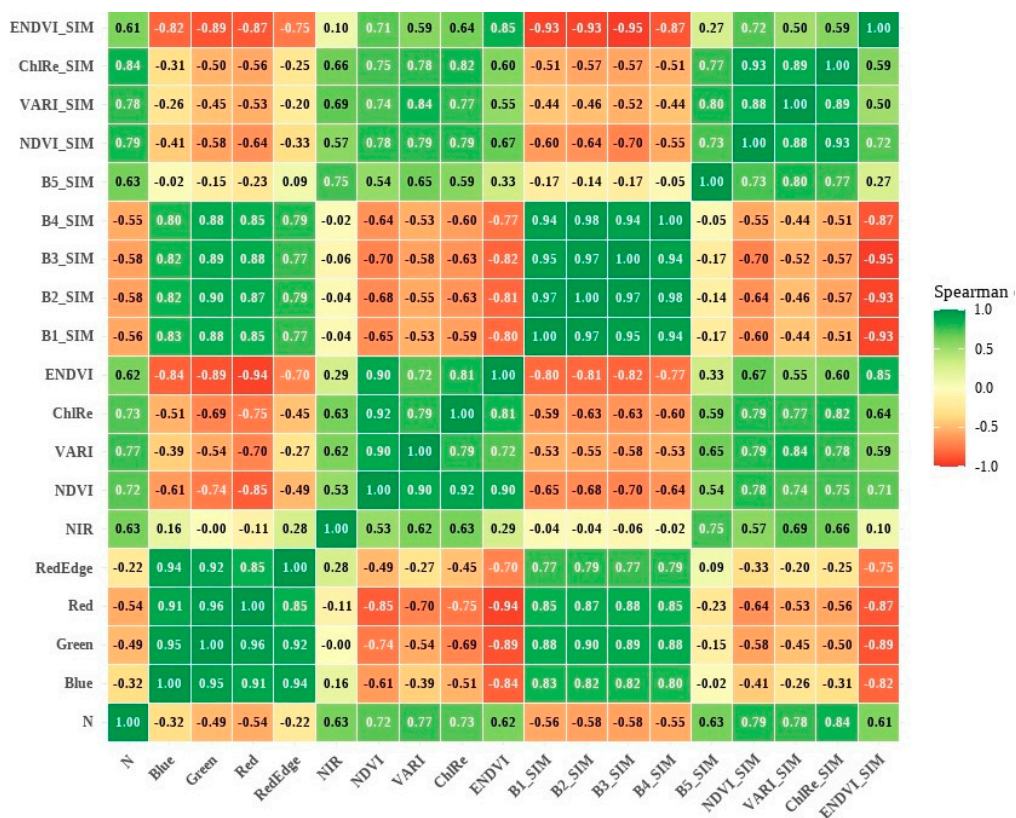
Precipitation patterns throughout the sugarcane growing cycle were consistent with the crop's physiological demands at each phenological stage. Although a slight water deficit was observed at the beginning of the growing season, the crop showed satisfactory development in later stages. This behavior is consistent with the variation in foliar nitrogen content across sampling periods, since, during vegetative growth, increased water availability promotes photosynthetic activity, canopy expansion, and N uptake [17].

In contrast, the reduction in leaf N content at 290 and 350 DAC, at the onset of maturation, may reflect both physiological changes inherent to the progression of the growth cycle and the internal redistribution of the nutrient, with a decrease in its concentration in the leaves and greater allocation of assimilates to the stems [51].

#### 3.2. Spearman Correlation Between Original and Simulated Bands and Indices in Relation to LNC

In this study, Spearman's correlation coefficient ( $r$ ) was used to assess the association between N content and the reflectance values of the original and simulated bands, as well as the spectral indices (SI) calculated from these bands. The analysis focused on the Blue, Green, Red, RedEdge, and NIR spectral regions due to their greater interaction with chlorophyll and pigments directly involved in the photosynthetic process and, consequently, closely related to TNF. Thus, the DAC dataset was

considered, which enabled a more robust assessment of the relationship between the variables under different conditions, as illustrated in Figure 6.



**Figure 6.** Spearman correlation matrix between nitrogen content and original and simulated vegetation indices. “SIM” refers to simulated data, indices, and bands.

The spectral bands and (IE) showed consistent associations with foliar nitrogen (N) content, with negative correlations in the visible spectrum and positive correlations in the near-infrared. The spectral bands and (SI) showed associations in different directions, with negative correlations in the visible spectrum and positive correlations in the near-infrared. In the bands obtained by UAV, for example, the “*r*” values were  $-0.32$ ,  $-0.49$ ,  $-0.54$ , and  $-0.22$  for Blue, Green, Red, and RedEdge, respectively, while in the NIR,  $r = 0.63$ .

The simulated bands reproduced a similar correlation pattern, but with more pronounced magnitudes (B1SIM  $r = -0.56$ ; B2SIM  $r = -0.58$ ; B3SIM  $r = -0.58$ ; B4SIM  $r = -0.55$ ; B5SIM  $r = 0.63$ ), indicating consistency between the simulation and the original data. This result indicates a stronger association between chlorophyll content and higher photosynthetic activity, which increases absorption in the VIS region, particularly in the Blue, Green, and RedEdge regions. In contrast, the signal in the NIR is more influenced by multiple scattering linked to the internal structure of the leaf and canopy architecture, which may support the positive correlation observed in this spectral region.

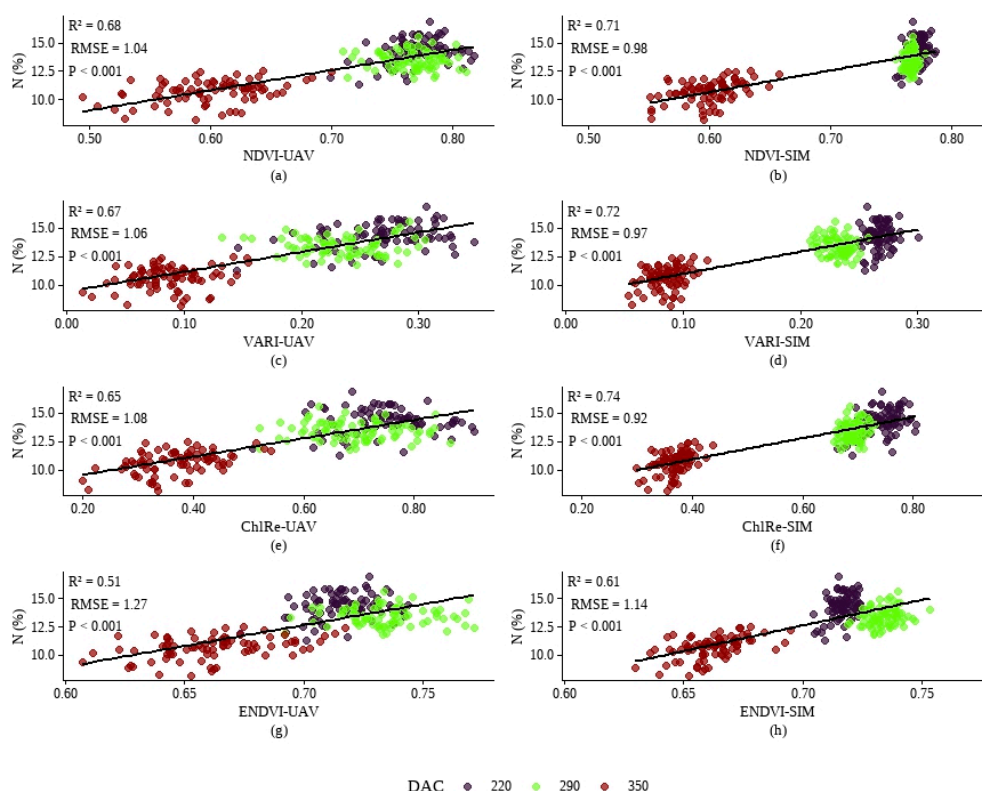
The (SI) indices showed a similar correlation with TNF in the assessed domains, where for NDVI,  $r = 0.72$  and simulated  $r = 0.79$ , for VARI,  $r = 0.77$  and simulated  $r = 0.78$ ; for ChlRe,  $r = 0.73$  and simulated  $r = 0.84$ ; and for ENDVI ( $r = 0.62$ ; simulated  $r = 0.61$ ). Furthermore, the indices derived from the simulated bands generally showed slightly higher correlations than those obtained with the original data, which is consistent with the origin of the bands: since they are constructed from narrow-band spectral information, they tend to better preserve subtle spectral features and, consequently, respond to small variations related to chlorophyll and N.

Finally, the comparison between indices calculated from actual UAV data and those derived from simulated data revealed the following correlations: NDVI vs. simulated ( $r = 0.78$ ), VARI vs. simulated ( $r = 0.84$ ), ChlRe vs. simulated ( $r = 0.82$ ), and ENDVI vs. simulated ( $r = 0.85$ ). Taken together,

these results indicate high agreement between the indices obtained from the bands measured by the UAV and those generated from the simulated bands, suggesting that the simulation adequately preserved the spectral behavior necessary to reproduce the patterns captured by the indices and, consequently, reinforcing the consistency and usefulness of both datasets.

### 3.3. Exploratory Estimation of Nitrogen Concentration in Sugarcane Using Vegetation Indices from UAV Data and Derived from the Simulated Dataset

Figure 7 shows the linear regression between N concentration (%) and the spectral indices (SI) used in this study (NDVI, VARI, ChlRe, and ENDVI), calculated from actual UAV data and their corresponding simulated (SIM) values in relation to their phenological stage at the DACs. These SIs summarize differences in reflectance in specific regions of the spectrum and, therefore, serve as compact descriptors of the biophysical and biochemical state of the vegetation.



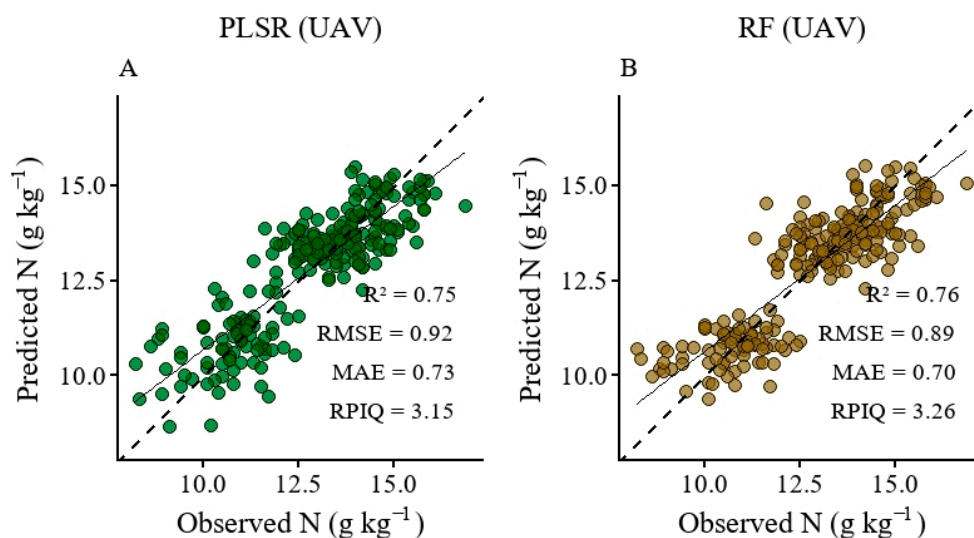
**Figure 7.** Linear regressions between nitrogen content (N, %) and the NDVI, VARI, ChlRe, and ENDVI indices, comparing values obtained directly by UAV and their corresponding simulated values (SIM). Each panel shows the linear fit, including  $R^2$ , RMSE, and significance (P), and letter codes (a–h).

The indices calculated from the canopy's spectral response were statistically significant ( $P < 0.001$ ) for TNF estimation, with performance ranging from "good" to "moderate." The NDVI, for example, showed  $R^2 = 0.68$  and  $RMSE = 1.04$ , followed by the other indices (VARI,  $R^2 = 0.67$ ,  $RMSE = 1.06$ ; ChlRe,  $R^2 = 0.65$ ,  $RMSE = 1.08$ ; ENDVI,  $R^2 = 0.51$ ,  $RMSE = 1.27$ ). The EIs obtained from the spectral response of the simulated data reproduced the trend observed in the EIs derived from UAV data, but with more clearly defined associations regarding leaf N content and the different DACs. This behavior can be explained by the nature of the simulated data, whose origin favors a more sensitive spectral response and one less subject to canopy interference, allowing for the capture of variations related to the leaf's nutritional status with greater discrimination. In this dataset, NDVI ( $R^2 = 0.71$  and  $RMSE = 0.98$ ), VARI ( $R^2 = 0.72$  and  $RMSE = 0.97$ ), and ChlRe ( $R^2 = 0.74$  and  $RMSE = 0.92$ ) performed satisfactorily, while ENDVI showed lower performance ( $R^2 = 0.61$ ;  $RMSE = 1.14$ ).

### 3.4. Modeling of Nitrogen Content Using a UAV-Derived Spectra and Simulated Data

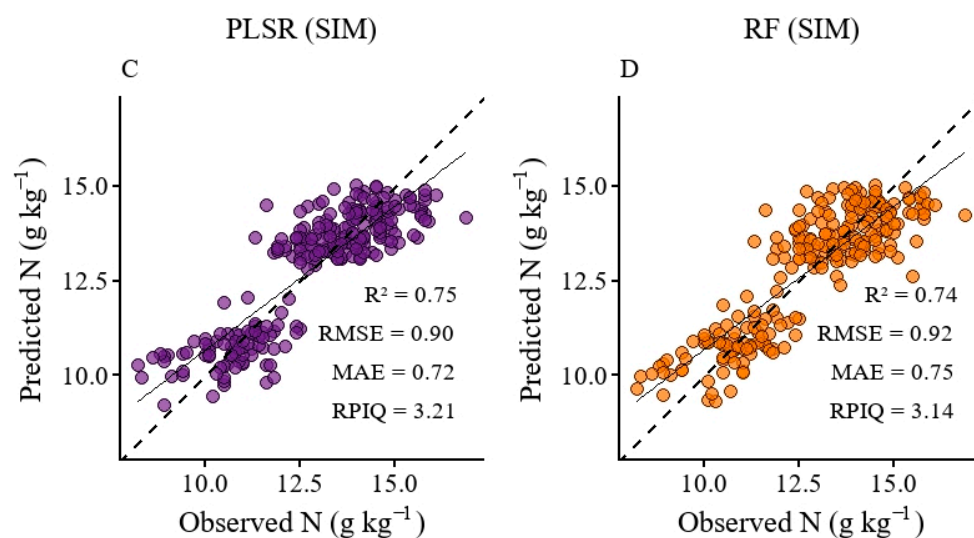
The sugarcane TNF prediction model was developed using original multispectral data and simulated spectral data obtained during the stump development phase. For modeling, two widely established algorithms in spectral analysis were employed: Partial Least Squares Regression (PLSR) [52] and Random Forest (RF) [53].

In this context, both models demonstrated good predictive capability for the UAV data (Figure 8B). The PLSR model showed “good” performance, with an  $R^2$  of 0.75, an RMSE of 0.92  $\text{g kg}^{-1}$ , an MAE of 0.73, and an RPIQ of 3.15. Similarly, the RF model also showed satisfactory performance, with an  $R^2$  of 0.76, an RMSE of 0.89  $\text{g kg}^{-1}$ , an MAE of 0.70, and an RPIQ of 3.26.



**Figure 8.** Performance of predictive nitrogen models using UAV data. (A) PLSR model (UAV), (B) RF model (UAV).

Similar results were also observed for the models calibrated using the simulated data, where PLSR yielded  $R^2 = 0.75$ , RMSE = 0.90  $\text{g kg}^{-1}$ , MAE = 0.72, and RPIQ = 3.21, while RF generated  $R^2 = 0.74$ , RMSE = 0.92  $\text{g kg}^{-1}$ , MAE = 0.75, and RPIQ = 3.14 (Figures 9C and 9B).

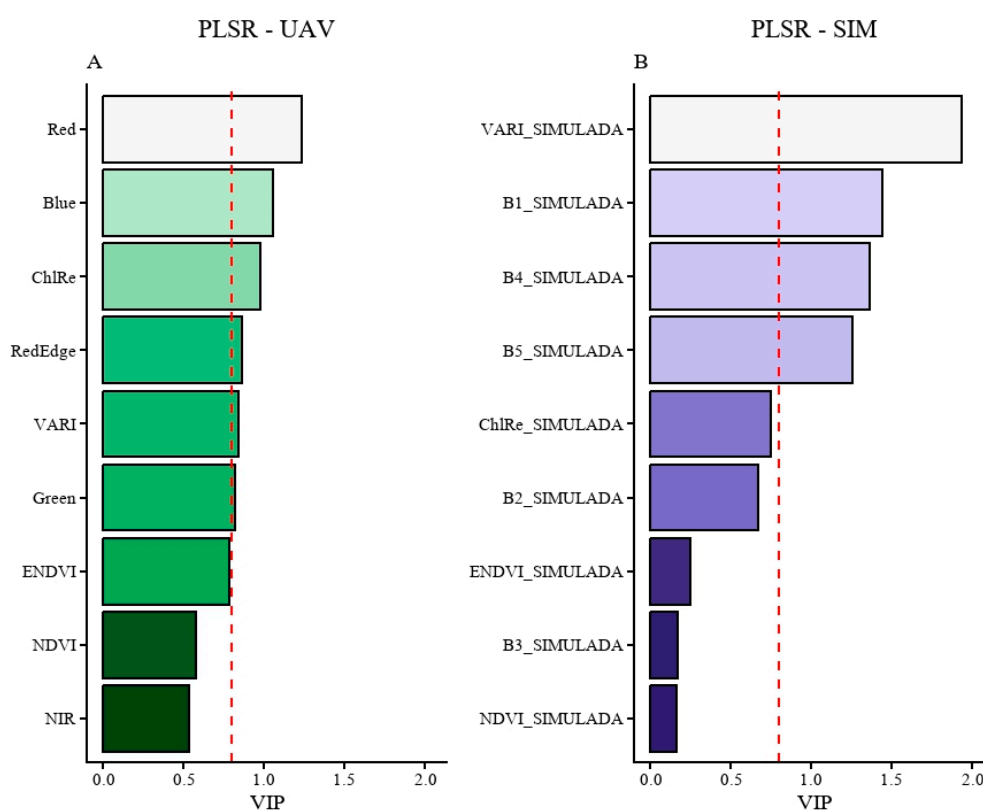


**Figure 9.** Performance of predictive nitrogen models using simulated data. (C) PLSR model (SIM), (D) RF model (SIM).

Overall, the performance of the models calibrated using data collected from the crop canopy via UAV and data simulated from leaf spectral curves showed similar values across all indicators and performed well, indicating that the methodology employed was able to satisfactorily reproduce the spectral behavior observed in the real data.

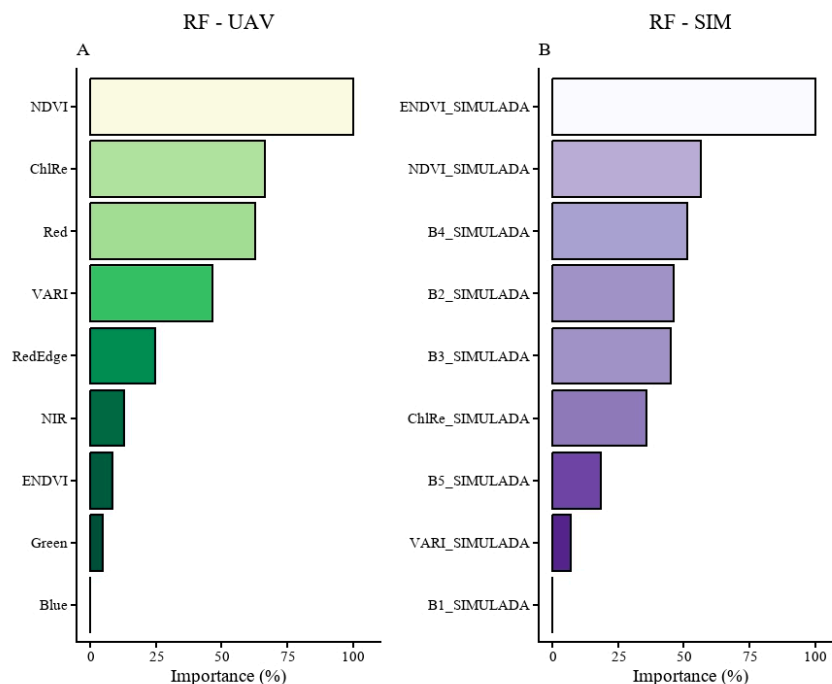
### 3.5. Variable Importance in Projection

Analysis of the Variables of Importance in Projection (VIP) values for the PLSR model (Figure 10A and B) revealed the variables that contributed most to the prediction of TFN. In the original canopy data, the Red, Blue, and RedEdge bands stood out in that order of importance, in addition to the ChlRe and VARI indices. For the simulated data, the most relevant variables include the Blue, RedEdge, and NIR bands, as well as the VARI index, in that order. In both datasets, these variables exceeded the heuristic threshold of 0.8, indicating an above-average contribution to explaining TFN variability.



**Figure 10.** Variable importance analysis (VIPs) for the PLSR models in predicting nitrogen content. (A) PLSR model using UAV data. (B) PLSR model using simulated (SIM) data.

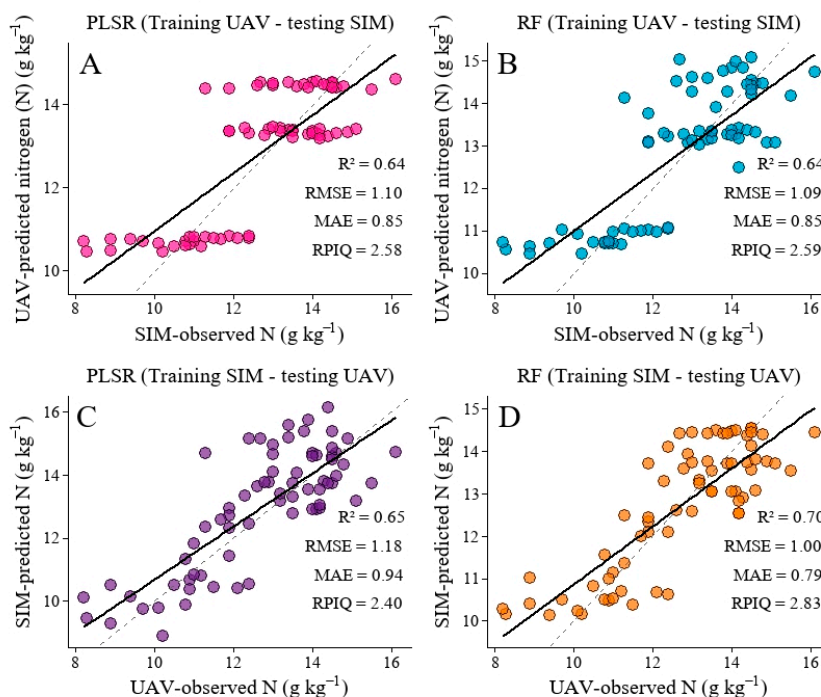
In contrast, the importance of variables by permutation was estimated by randomizing the values of each predictor, followed by an assessment of the impact of this change on the model's overall performance, as shown in (Figure 11 A and B). In the Random Forest (RF) model, this approach revealed a more concentrated distribution of predictive relevance across a smaller number of variables. In the UAV dataset, the NDVI, Clre, and VARI indices stood out, in addition to the Red band. In the simulated dataset, however, the greatest contributions were observed for the ENDVI\_SIM and NDVI\_SIM indices, as well as for the RedEdge and Green bands (B4 and B3\_SIM), as shown in Figures 9C and 9D. This behavior is consistent with the structural differences between the evaluated models.



**Figure 11.** Permutation-based variable importance in the RF model for nitrogen prediction. (A) RF model (UAV), (B) RF model (SIM).

### 3.6. Independent Validation

In this step, the hold-out validation technique was employed, in which 70% of the data was used for training and 30% for validation on an independent dataset. This procedure was used to analyze the transferability of the models between the UAV (original) and simulated domains in predicting TFN (Figure 12), comparing the PLSR and Random Forest (RF) algorithms.



**Figure 12.** Independent validation across domains (original UAV and simulated bands) for predicting leaf nitrogen content (PLSR and RF).

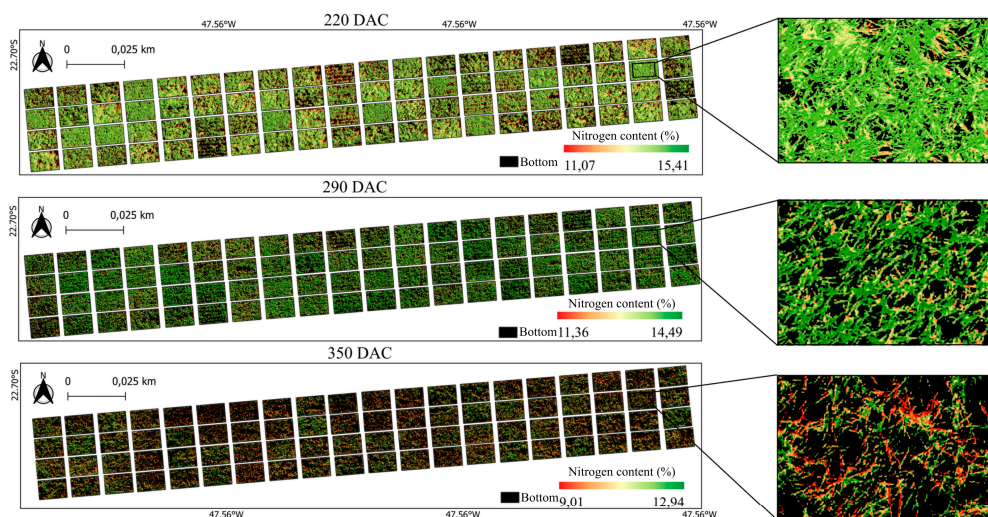
Panels A and B represent the models calibrated with the UAV bands and spectral indices and subsequently applied to the simulated bands and spectral indices. In turn, panels C and D present the reverse arrangement, in which the models were trained in the simulated domain and evaluated in the UAV domain, while maintaining the validation criterion.

The transferability of TNF models between domains (UAV and SIM), comparing PLSR and Random Forest in both transfer directions, yielded “moderate” results in scenarios (A, B, and C), with  $R^2 > 0.64$ , demonstrating that there is a stable relationship between the domains, though with a loss of accuracy associated with the change in spectral representation. Notably, the RF trained on SIM data and tested on UAV data showed the best performance in terms of explanation and error ( $R^2 = 0.70$ ; RMSE = 0.99 g kg<sup>-1</sup>; MAE = 0.78; RPIQ = 2.85), demonstrating greater robustness of the nonlinear model.

### 3.7. Spatialization and Mapping of Leaf Nitrogen Content

When comparing different modeling strategies for estimating TNF, it was found that the validated Random Forest (RF) model exhibited the best predictive performance, even when applied to a reduced dataset. Due to its robustness and flexibility in handling high-dimensional spectral data and spatial variability, the RF model was selected for the spatialization stage. Previous studies show that RF provides more accurate predictions than traditional linear methods, in addition to allowing the integration of multiple spectral bands and environmental variables [54].

In addition to demonstrating statistical stability, the model proved capable of spatializing foliar N content, which strengthens its applicability in pixel-by-pixel prediction of multispectral images. Using the global model, applied to each scenario (220, 290, and 350 DAC), it was possible to generate a map of the spatial distribution of N content in the experimental area, considering different phenological stages, as illustrated in Figure 13.



**Figure 13.** Spatial distribution of nitrogen content predicted by the Random Forest model, validated with simulated data.

As shown in Figure 23, the estimated TFN in the leaf is represented by a color gradient, in which red tones indicate lower concentrations and green tones indicate higher concentrations, following the concentration scale. It is noted that N values decrease gradually throughout the growth stages, stem elongation, and the maturation stage; for example, at 220 DAC, the leaf N content ranges from 11.07 to 15.41 (%), while at 350 DAC, the leaf N content ranges from 9.01 to 12.94 (%). Furthermore, the spatial pattern obtained indicates good model performance, which maintained consistent predictions even when using simulated data, suggesting the ability to generalize across domains. This behavior

is particularly relevant when the model is applied to high-resolution images, in which intra-plot variability tends to be more evident and, therefore, requires greater stability in the estimates.

## 4. Discussion

### 4.1. Analysis of the Correlations Between Original and Simulated Spectral Bands and Indices, and of the Linear Regression Applied to Spectral Indices for Predicting TFN

The relationships between spectral bands and spectral indices obtained at the canopy scale using a multispectral sensor mounted on a UAV, and those generated from simulated leaf data derived from hyperspectral data, were analyzed across different phenological stages. This approach is relevant, since spectral bands form the basis for the construction of mathematical indices, and their relevance stems from their association with specific absorption characteristics of photosynthetic pigments, especially chlorophyll [55].

The negative correlations observed for the Blue, Green, Red, and RedEdge bands, as well as the positive correlation for the NIR band (Figure 8), were verified in both datasets and reflect the spectral behavior of the sugarcane canopy, in which higher chlorophyll and foliar nitrogen contents are associated with greater absorption in the visible region and higher reflectance in the near-infrared [35]. However, since the sensor used is multispectral and operates with relatively broad bands (466, 576, 666, and 746 nm), the spectral response obtained by UAV tends to integrate and smooth out narrower spectral features. Therefore, in studies using hyperspectral data on sugarcane, the regions most sensitive to nitrogen are typically identified more specifically in the Green band, near 550 nm, and in the RedEdge region, between 680 and 750 nm [15,17].

Linear regression of spectral indices derived from the multispectral sensor showed similar performance in estimating TNF. Among these, the NDVI stood out, with the highest coefficient of determination ( $R^2 = 0.68$ ) and the lowest error ( $RMSE = 1.04 \text{ g kg}^{-1}$ ). NDVI combines red and near-infrared (Red and NIR) bands and is sensitive to “green” intensity, canopy vigor, and leaf nitrogen status [56–58]. Authors such as Kumarasiri et al. [18] also reported good performance of the NDVI in predicting TNF ( $R^2 = 0.77$ ), corroborating this finding; Li et al. [59] showed that incorporating ES from multispectral UAV data improves predictive models. Taken together, these results reinforce the potential of EIs, associated with N in the canopy.

Among the simulated indices, ChlRe showed the best fit for TNF ( $R^2 = 0.74$ ;  $RMSE = 0.92 \text{ g kg}^{-1}$ ), followed by VARI ( $R^2 = 0.72$ ;  $RMSE = 0.97 \text{ g kg}^{-1}$ ) and NDVI ( $R^2 = 0.71$ ;  $RMSE = 0.98 \text{ g kg}^{-1}$ ). These findings corroborate those of Martins et al. [60], in which the authors highlight the high sensitivity of indices derived from hyperspectral data for sugarcane in predicting TNF. Complementarily, in a study conducted by Hassani et al. [61], evaluating corn and sorghum with re-calibrated spectral data, the NDVI derived from simulated UAV data yielded  $R^2 = 0.72$  and  $RMSE = 0.22 \text{ g kg}^{-1}$  for sorghum, while for corn, ChlRe derived from ASD data yielded  $R^2 = 0.85$  and  $RMSE = 0.18 \text{ g kg}^{-1}$ , and NDVI derived from Landsat 8 OLI data yielded  $R^2 = 0.81$  and  $RMSE = 0.21 \text{ g kg}^{-1}$ . Taken together, these results reinforce that the use of indices such as ChlRe, VARI, and NDVI constitutes robust alternatives for nutritional prediction, especially when applied to simulated data.

### 4.2. Model Performance in Nitrogen Prediction Using a UAV and Simulated Data (PLSR and Random Forest)

The leaf is the primary light-absorbing organ and plays a central role in nitrogen assimilation; thus, its composition and structure directly influence the crop's spectral response [62,63]. This influence is most evident in the bands associated with chlorophyll absorption, which tend to respond to variations in photosynthetic activity and, consequently, in the plant's nutritional status [56]. In the present study, this relationship was initially observed in the correlation analysis (Figure 6) and subsequently confirmed by the VIP values of the PLSR model (Figure 10). In the original canopy data, the Red ( $r = -0.54$ ) and Blue ( $r = -0.32$ ) bands stood out both for the magnitude of the correlation and for their contribution to the model. Similarly, in the simulated data, this pattern was most evident for

the Blue ( $r = -0.56$ ) and RedEdge ( $r = -0.55$ ) bands. This convergence between correlation and VIP reinforces the relevance of the visible and RedEdge regions for predicting foliar nitrogen content, in addition to indicating that, even under different acquisition conditions, the N-related spectral response was preserved [17,64].

This behavior helps explain the similar performance obtained by the models in both domains. PLSR yielded  $R^2 = 0.75$ , RMSE of  $0.92 \text{ g kg}^{-1}$ , MAE of  $0.73$ , and RPIQ of  $3.15$  for the UAV data (Figure A), and  $R^2 = 0.75$ , RMSE of  $0.90 \text{ g kg}^{-1}$ , MAE of  $0.72$ , and RPIQ of  $3.21$  for the simulated data (Figure C). The similarity between these metrics indicates that spectral simulation preserved the information necessary for nitrogen prediction without significant loss of performance [65]. These results corroborate previous studies that also reported good performance in estimating foliar nitrogen from spectral data. Li et al. [59], for example, obtained  $R^2 = 0.79$  and RMSE =  $0.11 \text{ g kg}^{-1}$  using PLSR on sugarcane based on LiDAR images. Lu et al. [66] observed similar results in winter wheat, with  $R^2 = 0.65$  and RMSE =  $0.13 \text{ g kg}^{-1}$ . Hassani et al. [61], however, reported even better performance using simulated UAV data derived from a field hyperspectral sensor, with  $R^2 = 0.81$  and RMSE =  $0.21 \text{ g kg}^{-1}$ . Taken together, these findings reinforce the potential of integrating different platforms and sensors for estimating foliar nitrogen in the same crop.

The Random Forest model showed similar predictive performance across both datasets, with slightly better results for the UAV data, where it achieved  $R^2 = 0.76$ , RMSE =  $0.89 \text{ g kg}^{-1}$ , MAE =  $0.70$ , and RPIQ =  $3.26$ . For the simulated data, the model achieved  $R^2 = 0.74$ , RMSE =  $0.92 \text{ g kg}^{-1}$ , MAE =  $0.75$ , and RPIQ =  $3.14$ . This slight advantage of the UAV data may be related to the RF's ability to model nonlinear relationships between spectral variables and the plants' nutritional status, more efficiently capturing the variability observed directly in the field [67]. The variable importance analysis reinforces this interpretation by indicating that the model concentrated its predictive power on a reduced set of key predictors. In the RF-UAV (Figure 11A), the NDVI and ChlRe indices stood out most prominently, followed by the Red band and the VARI index, highlighting the relevance of regions associated with chlorophyll, vegetative vigor, and visible light absorption for estimating foliar nitrogen content [17]. In RF-SIM (Figure 11B), the greatest contribution was observed for ENDVI\_SIMULATED and NDVI\_SIMULATED, followed by the B4\_SIMULATED, B2\_SIMULATED, and B3\_SIMULATED bands, indicating that, even after spectral simulation, the model retained sensitivity to the spectral bands most closely related to the crop's nutritional status.

These findings are consistent with recent literature on sugarcane. Li et al. [59] demonstrated the feasibility of estimating canopy N concentration using multispectral UAV imagery, finding a strong correlation between observed and predicted values in the best-performing model ( $R^2 = 0.79$ ; RMSE =  $0.11$ ). Furthermore, Picado et al. [24], although using multiple linear regression, also confirmed the potential of remote sensing with the MicaSense RedEdge-P camera for mapping foliar nutrients in sugarcane, achieving 75% accuracy for nitrogen. In contrast, Soltanikazemi et al. [68] reported more moderate RF performance ( $R^2 = 0.59$ ; RMSE =  $0.08 \text{ g kg}^{-1}$ ), while Hassani et al. [61], working with simulated UAS data for corn, observed lower agreement in validation ( $R^2 = 0.38 \pm 0.18$ ; RMSE =  $0.40 \pm 0.08$ ). This contrast reinforces the consistency of the results obtained in the present study and highlights the potential of spectral simulation as a promising strategy for nitrogen prediction, especially given the still limited number of studies using simulated data applied to sugarcane.

#### 4.3. Validation

When comparing the performance of the models using hold-out validation (Figure 12), it was found that cross-domain transfer depended on the direction of training and testing. In PLSR modeling, scenario (A), calibrated with UAV data and validated in the simulated domain, yielded  $R^2 = 0.64$ , RMSE =  $1.10 \text{ g kg}^{-1}$ , MAE =  $0.85$ , and RPIQ =  $2.58$ , while scenario (C), fitted with simulated data and evaluated in the UAV domain, resulted in  $R^2 = 0.65$ , RMSE =  $1.18 \text{ g kg}^{-1}$ , MAE =  $0.94$ , and RPIQ =  $2.40$ . Although performance was similar between the two transfer directions, there was variation in the error metrics, indicating that the model's generalization is sensitive to the domain used in

calibration. This can be explained by the fact that PLSR requires large datasets and, when dealing with smaller or unknown datasets, may have limitations, functioning as a “black box” [69].

This finding is consistent with recent studies that highlight limitations in the transferability of PLSR models across different contexts, including spatial, temporal, and phenological variations [70,71]. Furthermore, studies using imaging spectroscopy have also shown that phenology and differences between domains can critically affect the robustness and applicability of these models outside the conditions under which they were calibrated [72,73].

In RF modeling, scenario (B), calibrated with UAV data and validated in the simulated domain, yielded  $R^2 = 0.64$ ,  $RMSE = 1.09 \text{ g kg}^{-1}$ ,  $MAE = 0.85$ , and  $RPIQ = 2.59$ , while scenario (D), adjusted with simulated data and validated in the UAV domain, resulted in  $R^2 = 0.70$ ,  $RMSE = 0.99 \text{ g kg}^{-1}$ ,  $MAE = 0.78$ , and  $RPIQ = 2.85$ . The significance of this finding lies in the incorporation of an alternative representation of the dataset through the simulation of UAV bands, contributing to the model calibration process that enhanced its generalization capacity with respect to real data.

Studies conducted by Pan et al. [74] reinforce the view that models built using simulated data derived from hyperspectral reflectance and data actually acquired by UAVs demonstrate that combining these domains can yield more stable models with better performance across different data sources. Similarly, Chen et al. [7], when comparing hyperspectral data obtained by UAVs with simulated multispectral data, including sensors onboard UAVs and satellites, for TNF estimation, observed that simulation constitutes a promising strategy for expanding the operational feasibility of applications, although it also influences predictive performance depending on the configuration of the training and test domains.

In comparison, the models trained on the simulated dataset exhibited better data distribution and, consequently, greater predictive power for TNF, possibly due to the higher sensitivity of the source dataset, as previously reported in studies using hyperspectral sensors [15,17,75]. Furthermore, it is estimated that approximately 75% of foliar nitrogen is allocated to chloroplasts [76,77]. In this context, measurements obtained with fewer external interferences, such as variations in lighting, shading, and leaf angle [78]. This combination tends to represent the relationship between the nutrient, the spectral response, and the aerial image more consistently.

#### 4.4. Spatial Distribution of Nitrogen Content in Sugarcane

The correlation between the simulated spectral responses and TNF demonstrates that the simulated data preserve the spectral sensitivity inherited from the source dataset (hyperspectral). According to Silva et al. [15], hyperspectral data exhibit greater subtlety in the spectral response associated with N. This preserved spectral sensitivity not only enhances the predictive performance of the models but also expands their potential for spatial-scale application, enabling a shift from point-specific estimates of nitrogen content to its spatial distribution across the cultivated area [79–81].

The spatial distribution of nutrients is essential for guiding the application of variable-rate fertilizers [68]. Among nutrients, nitrogen monitoring is particularly challenging due to its high dynamics in the soil and rapid losses through volatilization and leaching [82]. In sugarcane, this relationship has been extensively explored through indirect measures of chlorophyll, such as SPAD readings and optical sensors, since chlorophyll tends to track variations in the plants' nitrogen nutritional status [17,83].

The spatial analysis in this study showed that the behavior of the sugarcane canopy throughout the phenological stages is consistent with the so-called “nitrogen dilution effect,” according to which the concentration of N in tissues tends to decrease as the plant accumulates dry matter and progresses toward senescence [84].

In this context, assessments conducted at 220 DAC tend to represent a physiologically more active canopy, whereas at 290 and 350 DAC, they reflect a reduction in TNF associated with physiological changes in the plant. Thus, this relationship depends not only on N content but also on how canopy reflectance is determined by the interaction between the vegetation's biochemical and

structural attributes such as pigments, leaf moisture, leaf angle, and internal leaf structure which can control a significant portion of radiation absorption and the process of light transmission and scattering [85].

The behavior of the sugarcane canopy throughout the phenological stages is consistent with the so-called nitrogen “dilution effect,” according to which the concentration of N in plant tissues tends to decrease as the plant accumulates dry matter and advances through the cycle toward senescence [84].

## 5. Conclusions

This study provided valuable information for simulating data derived from hyperspectral spectra, to support the calibration and application of multispectral sensors mounted on UAVs for estimating TNF in sugarcane. The EIs derived from the simulated bands maintained an association with TNF for the respective samples, differing slightly from the EIs calculated directly from the UAV, in addition to showing high agreement among themselves, indicating that the simulation procedure preserves the spectral behavior necessary to reproduce patterns observed in flight.

The application of machine learning techniques, particularly RF, demonstrated predictive capability in both domains (UAV:  $R^2 = 0.76$ ; Simulated:  $R^2 = 0.74$ ) and consistent performance in independent validation, notably in the transfer scenario where the RF trained on the simulated data and tested on the UAV achieved the best result ( $R^2 = 0.70$ ). Taken together, these findings indicate sufficient compatibility between spectral representations from different sources to support robust models, reinforcing the potential of integrating spectral simulation and machine learning as a strategy to expand the applicability of multispectral sensors in the field and guide future research involving multispectral and hyperspectral sensors coupled with UAVs.

**Author Contributions:** Conceptualization, I.d.L.e.L., M.L.d.S.A., A.K.d.S.O., R.R., C.A.A.C.S., and P.R.F.; methodology, I.d.L.e.L.; M.L.d.S.A., A.K.d.S.O., C.A.A.C.S., software, I.d.L.e.L., M.L.d.S.A.; validation, I.d.L.e.L., M.L.d.S.A., C.A.A.C.S., and R.R.; formal analysis, I.d.L.e.L., A.K.d.S.O., M.L.d.S.A.; investigation, R.R.; resources, P.R.F.; data curation, M.L.d.S.A., R.R and P.R.F.; writing—original draft preparation, I.d.L.e.L.; writing—review and editing, I.d.L.e.L., M.L.d.S.A.; visualization, R.R.; supervision, P.R.F.; project administration, P.R.F.; funding acquisition, P.R.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Luiz de Queiroz Agricultural Studies Foundation—FEALQ, Brazil. Additional support was provided by the São Paulo Research Foundation (FAPESP, grant 2024/10366-7) and by the Coordination for the Improvement of Higher Education Personnel (CAPES; Proc., 88887.027867/2024-00, 88887.146948/2025-00, 88887.993154/2024-00 and 88887.993148/2024-00).

**Data Availability Statement:** The data will be made available upon request.

**Acknowledgments:** To the Luiz de Queiroz Agricultural Studies Foundation (FEALQ) for funding the publication of this work. To DMLAB (Dinardo Miranda Agricultural Analysis Laboratory) for performing the chemical analyses. To the São Paulo Research Foundation (FAPESP) for their support.

**Conflicts of Interest:** The authors declare they are not aware of any conflicts of financial interest or personal relationships that could have influenced the work reported in this article.

## References

1. Barros, P. P. Da S.; Fiorio, P. R.; Demattê, J. A. De M.; Martins, J. A.; Montezano, Z. F.; Dias, F. L. F. Estimation of leaf nitrogen levels in sugarcane using hyperspectral models. *Ciência Rural*, v. 52, n. 7, p. e20200630, 2022. <https://doi.org/10.1590/0103-8478cr20200630>.
2. Giordano, M.; Giordano, G.; Esposito, D.; Cozzolino, E. The fate of nitrogen from soil to plants: influence of agricultural practices in modern agriculture. *Agriculture*, v. 11, n. 10, p. 944, 2021. <https://doi.org/10.3390/agriculture11100944>.

3. Li, X.; Ba, Y.; Zhang, M.; Nong, M.; Yang, C.; Zhang, S. Sugarcane nitrogen concentration and irrigation level prediction based on UAV multispectral imagery. *Sensors*, v. 22, n. 7, art. 2711, 2022. <https://doi.org/10.3390/s22072711>.
4. Reyes-Trujillo, A.; Daza-Torres, M. C.; Galindez-Jamioy, C. A.; Rosero-García, E. E.; Muñoz-Arboleda, F.; Solarte-Rodríguez, E. Estimating canopy nitrogen concentration of sugarcane crop using in situ spectroscopy. *Heliyon*, v. 7, n. 3, e06566, 2021. <https://doi.org/10.1016/j.heliyon.2021.e06566>.
5. Silva, A. L. Da; Oliveira, G. A. De; Souza, J. P. de; et al. Nitrogen use efficiency in agriculture: advances and challenges. *Ciência Rural*, v. 51, n. 6, e20200630, 2021. <https://doi.org/10.1590/0103-8478cr20200630>.
6. Segl, K.; Richter, R.; Küster, T.; Kaufmann, H. End-to-end sensor simulation for spectral band selection and optimization with application to the Sentinel-2 mission. *Applied Optics*, v. 51, n. 4, p. 439-449, 2012. <https://doi.org/10.1364/AO.51.000439>.
7. Chen, X.; Miao, Y.; Kusnierek, K.; Li, F.; Wang, C.; Shi, B.; Wu, F.; Chang, Q.; Yu, K. Potential of multi-source multispectral vs. hyperspectral remote sensing for winter wheat nitrogen monitoring. *Remote Sensing*, v. 17, n. 15, art. 2666, 2025. <https://doi.org/10.3390/rs17152666>.
8. Almeida, A. C.; Lima, J. C.; Sousa, F. D. et al. Prediction of nitrogen status in sugarcane fields using remote sensing data. *Computers and Electronics in Agriculture*, v. 175, p. 105685, 2020. <https://doi.org/10.1016/j.compag.2020.105685>.
9. Matése, A.; Prince Czarnecki, J. M.; Samiappan, S.; Moorhead, R. J. Are unmanned aerial vehicle-based hyperspectral imaging and machine learning advancing crop science? *Trends in Plant Science*, v. 29, n. 2, p. 196-209, 2024. <https://doi.org/10.1016/j.tplants.2023.09.001>.
10. Atherton, J.; Zhang, C.; Oivukkamäki, J.; Kulmala, L.; Xu, S.; Hakala, T.; Honkavaara, E.; Macarthur, A.; Porcar-Castell, A. What does the NDVI really tell us about crops? Insight from proximal spectral field sensors. *EarthArXiv*, 2020. Preprint. <https://doi.org/10.31223/osf.io/eayph>.
11. Zhang, L.; Wang, A.; Zhang, H.; Zhu, Q.; Zhang, H.; Sun, W.; Niu, Y. Estimating leaf chlorophyll content of winter wheat from UAV multispectral images using machine learning algorithms under different species, growth stages, and nitrogen stress conditions. *Agriculture*, v. 14, n. 7, art. 1064, 2024. <https://doi.org/10.3390/agriculture14071064>.
12. Zou, M.; Liu, Y.; Fu, M.; Li, C.; Zhou, Z.; Meng, H.; Xing, E.; Ren, Y. Combining spectral and texture feature of UAV image with plant height to improve LAI estimation of winter wheat at jointing stage. *Frontiers in Plant Science*, v. 14, art. 1272049, 2024. <https://doi.org/10.3389/fpls.2023.1272049>.
13. Yang, X.; Deng, X.; Li, Z.; Xu, J.; Zhou, Y. Monitoring sugarcane growth and nitrogen status using UAV-based multispectral imagery and machine learning. *UAVs*, v. 6, n. 9, p. 230, 2022. <https://doi.org/10.3390/UAVs6090230>.
14. Li, S.; Wang, C.; Wang, L.; Deng, F. Estimating the nitrogen content of sugarcane based on UAV hyperspectral images combined with machine learning algorithms. *Sensors*, v. 22, n. 7, p. 2711, 2022. <https://doi.org/10.3390/s22072711>.
15. Silva, C. A. A. C.; Fiorio, P. R.; Rizzo, R.; Rossetto, R.; Vitti, A. C.; Dias, F. L. F. Detection of nutritional stress in sugarcane by VIS-NIR-SWIR reflectance spectroscopy. *Ciência Rural*, v. 53, n. 12, p. e20220543, 2023. <https://doi.org/10.1002/fes3.70195>.
16. Martins, J. A.; Fiorio, P. R.; Silva, C. A. A. C. et al. Aplicação de índices vegetativos para estimativa de nitrogênio foliar em cana-de-açúcar usando dados hiperespectrais. *Sugar Tech*, v. 26, p. 160-170, 2024. <https://doi.org/10.1007/s12355-023-01329-1>.
17. Fiorio, P. R.; Silva, C. A. A. C.; Rizzo, R.; Demattê, J. A. M.; Luciano, A. C. Dos S.; Silva, M. A. da. Prediction of leaf nitrogen in sugarcane (*Saccharum* spp.) by vis-NIR-SWIR spectroradiometry. *Heliyon*, e26819, 2024. <https://doi.org/10.1016/j.heliyon.2024.e26819>.
18. Kumarasiri, D. S.; Wijetunga, C.; Gunathilaka, L. D. Use of UAV imagery to predict leaf nitrogen content of sugarcane cultivated under organic fertilizer application. *Tropical Agricultural Research*, v. 35, n. 1, p. 67-79, 2024. <https://tar.sljol.info/articles/8700>.
19. Alvares, C. A. et al. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, v. 22, n. 6, p. 711-728, 1 dez. 2013.

20. Lee, M. A.; Huang, Y.; Yao, H.; Thomson, S. J.; Bruce, L. M. Determining the effects of storage on cotton and soybean leaf samples for hyperspectral analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 7, n. 6, p. 2562-2570, 2014. <https://doi.org/10.1109/JSTARS.2014.2330521>.
21. Tavares, T. R.; Fiorio, P. R.; Seixas, H. T.; Garcia, A. C.; Barros, P. P. Da S. Effects of storage on vis-NIR-SWIR reflectance spectra of Mombasa grass leaf samples. *Ciência Rural*, v. 50, n. 3, 2020. <https://doi.org/10.1590/0103-8478cr20190587>.
22. ASD - Analytical Spectral Devices. (2010). *FieldSpec® 3 User Manual*. <Http://Www.Asdi.Com/>.
23. AGISOFT LLC. *Agisoft Metashape Professional Edition, version 2.0.1*. St. Petersburg, Russia: Agisoft LLC, 2022
24. Picado, E. F.; Romero, K. F.; Heenkenda, M. K. Mapping spatial variability of sugarcane foliar nitrogen, phosphorus, potassium and chlorophyll concentrations using remote sensing. *Geomatics*, v. 5, n. 1, art. 3, 2025. <https://doi.org/10.3390/geomatics5010003>.
25. QGIS DEVELOPMENT TEAM. *QGIS Geographic Information System. Version 3.40*. [S.l.]: QGIS Association, 2024. Disponível em: [qgis.org](http://qgis.org).
26. Wang, Y.; Zhang, Z.; Huang, J.; Zhu, D.; Liu, L.; Huang, W.; Wang, C.; Zhao, J.; Wu, J.; Zhang, D. Development of a UAV-based vegetation index for monitoring crop growth at various spatial resolutions. *Sensors*, v. 22, n. 9, p. 3503, 2022. <https://doi.org/10.3390/s22093503>.
27. Savitzky, A.; Golay, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, v. 36, n. 8, p. 1627-1639, 1964. <https://doi.org/10.1021/ac60214a047>.
28. R Core Team. *R: A language and environment for statistical computing*. 2024. Disponível em: <https://www.R-project.org/>.
29. Burggraaff, O. Biases from incorrect reflectance convolution. *Optics Express*, v. 28, n. 9, p. 13801-13816, 2020. <https://doi.org/10.1364/OE.389713>.
30. Wang, Y.; Kowalski, B. R. Calibration transfer and measurement stability of near-infrared spectrometers. *Applied Spectroscopy*, v. 46, n. 5, p. 764-771, 1992. <https://doi.org/10.1366/0003702924124808>.
31. Wang, Y.; Veltkamp, D. J.; Kowalski, B. R. Multivariate instrument standardization. *Analytical Chemistry*, v. 63, n. 23, p. 2750-2756, 1991. <https://doi.org/10.1021/ac00023a016>.
32. Mishra, P. A tutorial on standard normal variate and its variants: mathematical derivations and applications in calibration transfer. *TrAC, Trends in Analytical Chemistry*, v. 143, p. 116331, 2021. <https://doi.org/10.1016/j.trac.2021.116331>.
33. Ji, W.; Viscarra Rossel, R. A.; Shi, Z. Accounting for the effects of different instruments and protocols on mid-infrared spectral soil measurements. *European Journal of Soil Science*, v. 66, n. 5, p. 822-839, 2015. <https://doi.org/10.1111/ejss.12271>.
34. Xue, J.; Su, B. Significant remote sensing vegetation indices: a review of developments and applications. *Journal of Sensors*, 2017, art. 1353691, 17 p. <https://doi.org/10.1155/2017/1353691>.
35. Zhang, Y.; Wang, J.; Li, H.; Xu, X.; Chen, Q. UAV hyperspectral and optimized multispectral sensing for accurate estimation of rice nitrogen status. *Remote Sensing*, v. 17, n. 15, p. 2666, 2025. <https://doi.org/10.3390/rs17152666>.
36. Rouse, J.; Haas, R.; Schell, J.; Deering, D.; Harlan, J. Monitoring vegetation systems in the Great Plains with ERTS. In: *THIRD EARTH RESOURCES TECHNOLOGY SATELLITE-1 SYMPOSIUM*. [S.l.]: [s.n.], 1973. p. 309-317.
37. Gitelson, A. A.; Kaufman, Y. J.; Stark, R.; Rundquist, D. Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment*, v. 80, n. 1, p. 76-87, 2002. [https://doi.org/10.1016/S0034-4257\(01\)00289-9](https://doi.org/10.1016/S0034-4257(01)00289-9).
38. Gitelson, A. A., Gritz, Y., & 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(3), 271–282. <https://doi.org/10.1078/0176-1617-00887>
39. Maxmax. *ENDVI*. 2015. <http://www.maxmax.com/endvi.htm>

40. Schober, Patrick; Boer, Christa; Schwarte, Lothar A. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, v. 126, n. 5, p. 1763-1768, maio 2018. <https://doi.org/10.1213/ANE.0000000000002864>.
41. Wei, Hsiang-En; Grafton, Miles; Bretherton, Michael; Irwin, Matthew; Sandoval, Eduardo. Evaluation of Point Hyperspectral Reflectance and Multivariate Regression Models for Grapevine Water Status Estimation. *Remote Sensing*, v. 13, n. 16, art. 3198, 2021. <https://doi.org/10.3390/rs13163198>.
42. *Frontiers In Remote Sensing*. Remote sensing estimation of nitrogen content in crops using vegetation indices: a review. *Frontiers in Remote Sensing*, 2024. <https://doi.org/10.3389/frsen.2024.1414540>.
43. Xue, J.; Liu, J.; Wang, X. et al. Estimation of leaf nitrogen content in sugarcane using UAV-based hyperspectral imagery and machine learning algorithms. *Remote Sensing*, v. 14, n. 11, p. 2576, 2022. DOI: <https://doi.org/10.3390/rs14112576>.
44. Chen, H.; Tan, C.; Lin, Z.; Wu, T. Classification and quantitation of milk powder by near-infrared spectroscopy and mutual information-based variable selection and partial least squares. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, v. 189, p. 183-189, 2018. <https://doi.org/10.1016/j.saa.2017.08.034>.
45. Breiman, L. Random forests. *Machine Learning*, v. 45, n. 1, p. 5-32, 2001. <https://doi.org/10.1023/A:1010933404324>.
46. Liu, Y.; Wang, Y.; Zhang, J. A New Machine Learning Algorithm: Random Forest. In: Liu, B.; Ma, M.; Chang, J. (eds.). *Information Computing and Applications. ICICA 2012. Lecture Notes in Computer Science*, v. 7473. Berlin; Heidelberg: Springer, 2012. [https://doi.org/10.1007/978-3-642-34062-8\\_32](https://doi.org/10.1007/978-3-642-34062-8_32).
47. Tang, K. et al. Identifying the influencing factors of soil nitrous acid emissions using random forest model. *Atmospheric Environment*, v. 318, p. 120875, 2024. <https://doi.org/10.1016/j.atmosenv.2024.120875>.
48. Liu, M. et al. Multivariable integration method for estimating sea surface salinity in coastal waters from in situ data and remotely sensed data using random forest algorithm. *Computers & Geosciences*, v. 75, p. 44-56, 2015. <https://doi.org/10.1016/j.cageo.2014.10.016>.
49. Friedman, Jerome H.; Meulman, Jacqueline J. Multiple additive regression trees with application in epidemiology. *Statistics in Medicine*, v. 22, n. 9, p. 1365-1381, 2003. <https://doi.org/10.1002/sim.1501>.
50. Rodrigues, M.; Nanni, M. R.; Cezar, E.; Santos, G. L. A. A. Dos; Reis, A. S.; Oliveira, K. M. De; Oliveira, R. B. de. Vis-NIR spectroscopy: from leaf dry mass production estimate to the prediction of macro- and micronutrients in soybean crops. *Journal of Applied Remote Sensing*, v. 14, 2020. <https://doi.org/10.1117/1.JRS.14.044505>.
51. Leite, J. M.; Ciampitti, I. A.; Mariano, E.; Vieira-Megda, M. X.; Trivelin, P. C. O. Nutrient partitioning and stoichiometry in unburnt sugarcane ratoon at varying yield levels. *Frontiers in Plant Science*, v. 7, art. 466, 2016. <https://doi.org/10.3389/fpls.2016.00466>.
52. Hassan, Muhammad Adeel; Yang, Mengjiao; Rasheed, Awais; Yang, Guijun; Reynolds, Matthew; Xia, Xianchun; Xiao, Yonggui; He, Zhonghu. A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Science*, v. 282, p. 95-103, 2019. <https://doi.org/10.1016/j.plantsci.2018.10.022>.
53. Zhang, C.; Yang, G.; Li, H.; Tang F.; Liu C.; Zhang Y.; (2018). Remote sensing inversion of leaf area index of winter wheat based on random forest algorithm. *Scientia Agricultura Sin.*51 (5), 855–867. <https://doi.org/10.3864/j.issn.0578-1752.2018.05.005>
54. Osco, L. P.; Ramos, A. P. M.; Pereira, D. R.; Moriya, É. A. S.; Imai, N. N.; Matsubara, E. T.; Estrabis, N.; Souza, M. De; Marcato Junior, J.; Creste, J. E. Predicting canopy nitrogen content in citrus-trees using random forest algorithm associated to spectral vegetation indices from UAV-imagery. *Remote Sensing*, v. 11, n. 24, p. 2925, 2020. <https://doi.org/10.3390/rs11242925>.
55. Imran, H. A.; Gianelle, D.; Rocchini, D.; Dalponte, M.; Martín, M. P.; Sakowska, K.; Wohlfahrt, G.; Vescovo, L. VIS-NIR, RedEdge and NIR-shoulder based normalized vegetation indices response to co-varying leaf and canopy structural traits in heterogeneous grasslands. *Remote Sensing, Basel*, v. 12, n. 14, p. 2254, 2020. <https://doi.org/10.3390/rs12142254>.
56. Narmilan, A.; González-Barrios, P.; Mondal, S.; Ninomiya, S.; Peña, R. J.; Dreisigacker, S. Hyperspectral vegetation indices for the prediction of stomatal conductance and chlorophyll fluorescence in spring wheat

- under different water regimes. *Remote Sensing*, v. 14, n. 5, art. 1140, 2022. <https://doi.org/10.3390/rs14051140>.
57. Gutman, G.; Skakun, S.; Gitelson, A. Revisiting the use of red and near-infrared reflectances in vegetation studies and numerical climate models. *Science of Remote Sensing*, v. 4, art. 100025, 2021. <https://doi.org/10.1016/j.srs.2021.100025>.
  58. Miphokasap, P.; Honda, K.; Vaiphasa, C.; Souris, M.; Nagai, M. Estimating canopy nitrogen concentration in sugarcane using field imaging spectroscopy. *Remote Sensing*, v. 4, n. 6, p. 1651-1670, 2012. <https://doi.org/10.3390/rs4061651>.
  59. Li, X.; Ba, Y.; Zhang, M.; Nong, M.; Yang, C.; Zhang, S. Sugarcane nitrogen concentration and irrigation level prediction based on UAV multispectral imagery. *Sensors*, v. 22, n. 7, art. 2711, 2022. <https://doi.org/10.3390/s22072711>.
  60. MARTINS, J. A.; FIORIO, P. R.; SILVA, C. A. A. C. et al. Aplicação de índices vegetativos para estimativa de nitrogênio foliar em cana-de-açúcar usando dados hiperespectrais. *Sugar Tech*, v. 26, p. 160-170, 2024. <https://doi.org/10.1007/s12355-023-01329-1>.
  61. Hassani, K.; Gholizadeh, H.; Taghvaeian, S.; Natalie, V.; Carpenter, J.; Jacob, J. Avaliando o impacto da resolução espacial do sensoriamento remoto baseado em UAS e da resolução espectral do sensoriamento proximal na precisão da recuperação de nitrogênio em culturas. *International Journal of Remote Sensing*, v. 44, n. 14, p. 4441-4464, 2023. <https://doi.org/10.1080/01431161.2023.2237162>.
  62. Lv, Z.; Zhao, W.; Kong, S.; Li, L.; Lin, S. Overview of molecular mechanisms of plant leaf development: a systematic review. *Frontiers in Plant Science*, v. 14, e1293424, 2023. <https://doi.org/10.3389/fpls.2023.1293424>.
  63. Yoneyama, T.; Suzuki, A. Light-independent nitrogen assimilation in plant leaves: nitrate incorporation into glutamine, glutamate, aspartate, and asparagine traced by <sup>15</sup>N. *Plants*, v. 9, n. 10, art. 1303, 2020. <https://doi.org/10.3390/plants9101303>.
  64. Miphokasap, P.; Wannasiri, W. Estimations of nitrogen concentration in sugarcane using hyperspectral imagery. *Sustainability*, v. 10, n. 4, art. 1266, 2018. <https://doi.org/10.3390/su10041266>.
  65. Biriukova, K.; Celesti, M.; Evdokimov, A.; Pacheco-Labrador, J.; Julitta, T.; Migliavacca, M.; Giardino, C.; Miglietta, F.; Colombo, R.; Panigada, C.; Rossini, M. Effects of varying solar-view geometry and canopy structure on solar-induced chlorophyll fluorescence and PRI. *International Journal of Applied Earth Observation and Geoinformation*, v. 89, art. 102069, 2020. <https://doi.org/10.1016/j.jag.2020.102069>.
  66. Lu, N.; Wang, W.; Zhang, Q.; Li, D.; Yao, X.; Tian, Y.; Zhu, Y.; Cao, W.; Baret, F.; Liu, S.; Cheng, T. Estimation of nitrogen nutrition status in winter wheat from unmanned aerial vehicle based multi-angular multispectral imagery. *Frontiers in Plant Science*, v. 10, art. 1601, 2019. <https://doi.org/10.3389/fpls.2019.01601>.
  67. Chatraei Azizabadi, E.; El-Shetehy, M.; Cheng, X.; Youssef, A.; Badreldin, N. In-Season Potato Nitrogen Prediction Using Multispectral UAV Data and Machine Learning. *Remote Sensing*, v. 17, n. 11, art. 1860, 2025. <https://doi.org/10.3390/rs17111860>.
  68. Soltanikazemi, M.; Minaei, S.; Shafizadeh-Moghadam, H.; Mahdavian, A. Field-scale estimation of sugarcane leaf nitrogen content using vegetation indices and spectral bands of Sentinel-2: application of random forest and support vector regression. *Computers and Electronics in Agriculture*, v. 200, art. 107130, 2022. <https://doi.org/10.1016/j.compag.2022.107130>.
  69. Huang, Y.; Chen, W.; Tan, W.; Deng, Y.; Yang, C.; Zhu, X.; Shen, J.; Liu, N. Transfer learning for enhancing the generality of leaf spectroscopic models in estimating crop foliar nutrients across growth stages. *International Journal of Applied Earth Observation and Geoinformation*, v. 139, art. 104481, 2025. <https://doi.org/10.1016/j.jag.2025.104481>.
  70. Ji, F.; Li, F.; Hao, D.; Shiklomanov, A. N.; Yang, X.; Townsend, P. A.; Dashti, H.; Nakaji, T.; Kovach, K. R.; Liu, H.; Luo, M.; Chen, M. Unveiling the transferability of PLSR models for leaf trait estimation: lessons from a comprehensive analysis with a novel global dataset. *New Phytologist*, v. 243, p. 111-131, 2024. <https://doi.org/10.1111/nph.19807>.

71. Helsen, K.; Bassi, L.; Feilhauer, H. Evaluating different methods for retrieving intraspecific leaf trait variation from hyperspectral leaf reflectance. *Ecological Indicators*, v. 130, p. 108111, 2021. <https://doi.org/10.1016/j.ecolind.2021.108111>.
72. Schiefer, F.; Schmidlein, S.; Kattenborn, T. The retrieval of plant functional traits from canopy spectra through RTM-inversions and statistical models are both critically affected by plant phenology. *Ecological Indicators*, v. 121, art. 107062, 2021. <https://doi.org/10.1016/j.ecolind.2020.107062>.
73. Chlus, A.; Townsend, P. A. Characterizing seasonal variation in foliar biochemistry with airborne imaging spectroscopy. *Remote Sensing of Environment*, v. 275, p. 113023, 2022. <https://doi.org/10.1016/j.rse.2022.113023>.
74. Pan, Y.; Li, J.; Zhang, J.; He, J.; Zhang, Z.; Yao, X.; Cheng, T.; Zhu, Y.; Cao, W.; Tian, Y. Estimating leaf nitrogen accumulation considering vertical heterogeneity using multiangular unmanned aerial vehicle remote sensing in wheat. *Plant Phenomics*, v. 6, art. 0276, 2024. <https://doi.org/10.34133/plantphenomics.0276>.
75. Silva, C. A. A. C.; Rizzo, R.; Oliveira, A. K. Da S.; Castro, M. P. P.; Alexandre, M. L. De S.; Lima, I. De L. E.; Demattê, J. A. M.; Fiorio, P. R. Interspecies prediction of nitrogen content in processed plant samples using spectroscopic modeling and transfer learning. *Food and Energy Security*, v. 15, e70195, 2026. <https://doi.org/10.1002/fes3.70195>.
76. Alharbi, K.; Haroun, S.A.; Kazamel, A.M.; Abbas, M.A.; Ahmaida, S.M.; Alkahtani, M.; Alhusnain, L.; Attia, K.A.; Abdelaal, K.; Gamel, R.M.E. Physiological Studies and Ultrastructure of *Vigna sinensis* L. and *Helianthus annuus* L. under Varying Levels of Nitrogen Supply. *Plants* 2022, 11, 1884. <https://doi.org/10.3390/plants11141884>.
77. Mu, X.; Chen, Y. A resposta fisiológica da fotossíntese à deficiência de nitrogênio. *Plant Physiology and Biochemistry*, v. 158, p. 76-82, 2021. <https://doi.org/10.1016/j.plaphy.2020.11.019>.
78. Alexandre, M.L.d.S.; e Lima, I.d.L.; Nilsson, M.S.; Rizzo, R.; Silva, C.A.A.C.; Fiorio, P.R. Sugarcane (*Saccharum officinarum*) Productivity Estimation Using Multispectral Sensors in UAVs, Biometric Variables, and Vegetation Indices. *Agronomy* 2025, 15, 2149. <https://doi.org/10.3390/agronomy15092149>.
79. Yu, K. Q.; Zhao, Y. R.; Li, X. L.; Shao, Y. N.; Liu, F. et al. Hyperspectral imaging for mapping of total nitrogen spatial distribution in pepper plant. *PLOS ONE*, v. 9, n. 12, e116205, 2014. <https://doi.org/10.1371/journal.pone.0116205>.
80. Pechanec, V.; Mráz, A.; Rozkošný, L.; Vyvlečka, P. Usage of airborne hyperspectral imaging data for identifying spatial variability of soil nitrogen content. *ISPRS International Journal of Geo-Information*, v. 10, n. 6, art. 355, 2021. DOI: <https://doi.org/10.3390/ijgi10060355>.
81. Verrelst, J.; Rivera-Caicedo, J. P.; Reyes-Muñoz, P.; Morata, M.; Amin, E.; Tagliabue, G.; Panigada, C.; Hank, T.; Berger, K. Mapping landscape canopy nitrogen content from space using PRISMA data. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 178, p. 382-395, 2021. <https://doi.org/10.1016/j.isprsjprs.2021.06.017>.
82. Bassi, D., Menossi, M. & Mattiello, L. Nitrogen supply influences the establishment of photosynthesis along the sugarcane leaf. *Scientific Reports*, v. 8, art. 2327, 2018. <https://doi.org/10.1038/s41598-018-20653-1>.
83. Amaral, L. R. Do; Molin, J. P. Sensor óptico no auxílio à recomendação de adubação nitrogenada em cana-de-açúcar. *Pesquisa Agropecuária Brasileira*, v. 46, n. 12, p. 1633–1642, 2011. <https://doi.org/10.1590/S1678-3921.pab2011.v46.9781>.
84. Santana, A. C. A. et al. Critical nitrogen dilution curves and productivity assessments for plant cane. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v. 24, n. 4, p. 244–251, 2020. <https://doi.org/10.1590/1807-1929/agriambi.v24n4p244-251>.
85. Ollinger, S. V. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist*, v. 189, n. 2, p. 375-394, 2011. <https://doi.org/10.1111/j.1469-8137.2010.03536.x>.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.