agronomy

MDPI

Article

Different Perspectives on the Same Target: Field and Laboratory Spectroscopy for Estimating Nitrogen Content in Sugarcane Leaves

Izabelle de Lima e Lima, Marta Laura de Souza Alexandre, Rodnei Rizzo *, Ana Karla da Silva Oliveira, 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; eng.izabellelima@usp.br (I.d.L.e.L.); martalaura@usp.br (M.L.d.S.A.); anakarla95@usp.br (A.K.d.S.O.); carlosesalq@usp.br (C.A.A.C.S.); fiorio@usp.br (P.R.F.)
* Correspondence: rodnei.rizzo@gmail.com or rizzo@usp.br

Abstract

Proper nitrogen (N) management is essential for increasing the productivity of sugarcane (*Saccharum* spp.) and reducing the economic and environmental impacts associated with excessive fertilizer use. This study compared the performance of two portable spectroradiometers, FieldSpec 3 and HandHeld 2, in estimating foliar nitrogen content based on hyperspectral data in the visible and near-infrared regions, obtained throughout the crop cycle. The experiment was conducted in Piracicaba, São Paulo, Brazil, under four N rates: 0, 60, 120, and 180 kg ha^{-1}. Spectral measurements were taken at the foliar and canopy levels at eight evaluation times, accompanied by laboratory determination of N content. Partial Least Squares Regression (PLSR) and Random Forest (RF) models were fitted using the spectral data and days after cutting (DAC), included as a categorical factor and evaluated using 10-fold internal cross-validation, based on the metrics R^2, RMSE, MAE, and Willmott's refined agreement index (dr). The foliar data performed better with PLSR (R^2 = 0.727; RMSE = 1.381 g kg^{-1}; MAE = 1.109; dr = 0.917) than canopy data (R^2 = 0.591; RMSE = 1.489 g kg^{-1}; MAE = 1.157; dr = 0.866). PLSR also outperformed RF at both acquisition levels. The green (~550 nm) and red edge (~740 nm) regions were the most relevant for N estimation. Under the evaluated conditions, model performance was associated with the spectral acquisition level and conditions, the instrumental configuration, and the modeling strategy employed.

Keywords: canopy reflectance; leaf spectroscopy; leaf nitrogen content; VIS-NIR

Academic Editors: Thanos Balafoutis and Evangelos Anastasiou

Received: 29 May 2026
Revised: 20 July 2026
Accepted: 21 July 2026
Published: 24 July 2026

1. Introduction

Sugarcane (*Saccharum* spp.) stands out among the main global commodities, not only as a source of raw material for sugar production, but also due to its relevance in biomass generation and carbon fixation [1]. In addition, it has high potential for biofuel production [2], which has driven extensive research focused on its nutritional management and productive sustainability.

Among the essential nutrients, nitrogen (N) plays a central role in growth and biomass accumulation, directly influencing sugarcane productivity and technological quality [3]. In contrast, N deficiency reduces the photosynthetic rate and amino acid production, compromising plant growth and productivity [4]. Traditionally, nutritional diagnosis is

performed through laboratory analyses, which, although accurate, are costly and unfeasible on a large scale.

In this context, remote sensing (RS) has advanced as a rapid, non-invasive, and lower-cost alternative, allowing the assessment of parameters such as chlorophyll content and plant nutritional status under field conditions [5]. The interaction process between electromagnetic radiation (EMR) and vegetation leads to the partitioning of incident energy into three fundamental processes: reflection, absorption, and transmission [6]. Among the spectral regions commonly used in RS are the visible (Vis), near-infrared (NIR), and shortwave infrared (SWIR), which can be observed between wavelengths of 350 and 2500 nm and cover approximately 97% of incident solar radiation [7,8].

Spectral behavior, resulting from the interaction between electromagnetic radiation (EMR) and leaves, expresses important leaf attributes, such as internal structure, photosynthetic pigment concentration, and water status [6,9]. Hyperspectral sensors allow these characteristics to be quantified. At the canopy level, measurements allow the monitoring of plants and their structural properties, including canopy density, shape, and plant architecture. Despite the relevance of acquiring canopy spectral data, this information has limitations related to the influence of external factors, such as temperature, humidity, and light intensity [10,11].

In this regard, recent studies have explored the nutritional estimation of sugarcane at leaf and canopy levels using hyperspectral remote sensing. Studies such as Barros et al. [12] demonstrated that canopy reflectance allows the modeling of plant nitrogen concentration, with results of $R^2 = 0.50$ and RMSE = 1.67 g kg^{-1}. However, the quantity and quality of the dataset can directly affect model performance [13]. Miphokasap and Wannasiri [14] demonstrated that canopy hyperspectral data, obtained through field spectroradiometry and orbital imagery (EO-1 Hyperion), allow the estimation of canopy N concentration with high accuracy (R^2cv of 0.78 and RMSEcv of 0.035% N), using multivariate models such as SMLR and SVR. Furthermore, the performance of these models may vary according to climatic conditions, showing that hyperspectral data are highly susceptible to variations in environmental factors.

In addition, accuracy may also be influenced by canopy structural conditions and the crop phenological stage [15]. In this context, the use of spectral preprocessing techniques becomes essential to improve data quality and, consequently, enhance the performance of predictive models. Silva et al. [16], for example, removed noisy spectral regions (400–450; 850–1350; 1650–1850; and 1900–2000 nm) and applied multiplicative scatter correction (MSC) followed by the Savitzky–Golay (SG) filter before calibrating partial least squares regression (PLSR) models to estimate leaf nitrogen content. Such approaches aim to mitigate noise and scattering effects that may compromise spectral interpretation and the predictive capacity of the algorithms.

Fiorio et al. [17] and Silva et al. [16], in their studies using leaf reflectance, showed that data obtained under controlled laboratory conditions resulted in accurate and stable nitrogen models ($R^2 > 0.81$ and RMSE < 1.24 g kg^{-1}; and $R^2 = 0.90$ and RMSE = 0.74 g kg^{-1}), since these measurements are less affected by environmental and spectral interferences. Moreover, understanding spectral variations at the leaf level is essential to more accurately interpret the spectral responses observed at the canopy level. This knowledge provides support for the calibration and correction of data acquired by suborbital sensors [18]. Thus, the combined use of spectra obtained at different levels represents an important functional mechanism, reinforcing the relevance of comparing leaf and canopy spectra.

Despite advances in the application of hyperspectral spectroscopy to estimate nitrogen content in sugarcane, foliar and canopy measurements are still often evaluated independently. Comparing these levels of data acquisition is complex, as the observed differences may stem not only from the spectral characteristics of the instruments but also from the scale of observation, acquisition geometry, lighting conditions, canopy structure, and the contribution of background elements to the recorded signal [15,19]. In a broader context, multisensor integration and fusion approaches have been explored to combine information obtained from different instruments, platforms, and acquisition conditions [20]. However, the application of these strategies requires a prior understanding of the differences associated with acquisition levels and configurations.

It is within this context that the present study is conducted, by performing a direct, multi-temporal comparison between two spectral acquisition configurations for estimating leaf nitrogen content (LNC) in sugarcane. Foliar and canopy measurements were taken at corresponding experimental plots over eight assessment periods and compared within a common spectral band, using equivalent preprocessing and modeling procedures. The contribution of this study, therefore, lies in evaluating, within the same experimental framework, how the level of acquisition and variations occurring throughout the crop cycle affect spectral features, the most informative regions, and the performance of predictive models.

The study was based on the hypothesis that spectra obtained at the leaf level, under controlled laboratory conditions, would provide more accurate and stable predictions of LNC than spectra acquired at the canopy level, under field conditions. Additionally, it was assumed that both acquisition levels would remain sensitive to spectral variations associated with N, especially in the green and red-edge regions. Based on this hypothesis, the objective of this study was to compare the spectral behavior of leaves measured under controlled laboratory conditions with the spectral response of the canopy obtained in the field throughout the sugarcane growing season. Specifically, the study sought to characterize the spectral differences between the acquisition levels, identify the spectral regions most closely related to LNC, and compare the performance of Partial Least Squares Regression (PLSR) and Random Forest (RF) models in predicting this attribute.

2. Materials and Methods

The methodology used in this study consisted of eight steps (Figure 1): (i) definition of the experimental area; (ii) acquisition of the spectral signature of the samples at the canopy level; (iii) acquisition of the spectral signature of the leaves using the laboratory sensor; (iv) leaf chemical analysis to determine leaf nitrogen content; (v) preprocessing of the acquired data; (vi) characterization of the spectral signatures as a function of leaf nitrogen content (LNC); (vii) prediction of nitrogen content using Partial Least Squares Regression (PLSR) and Random Forest (RF) models; (viii) validation of the results.

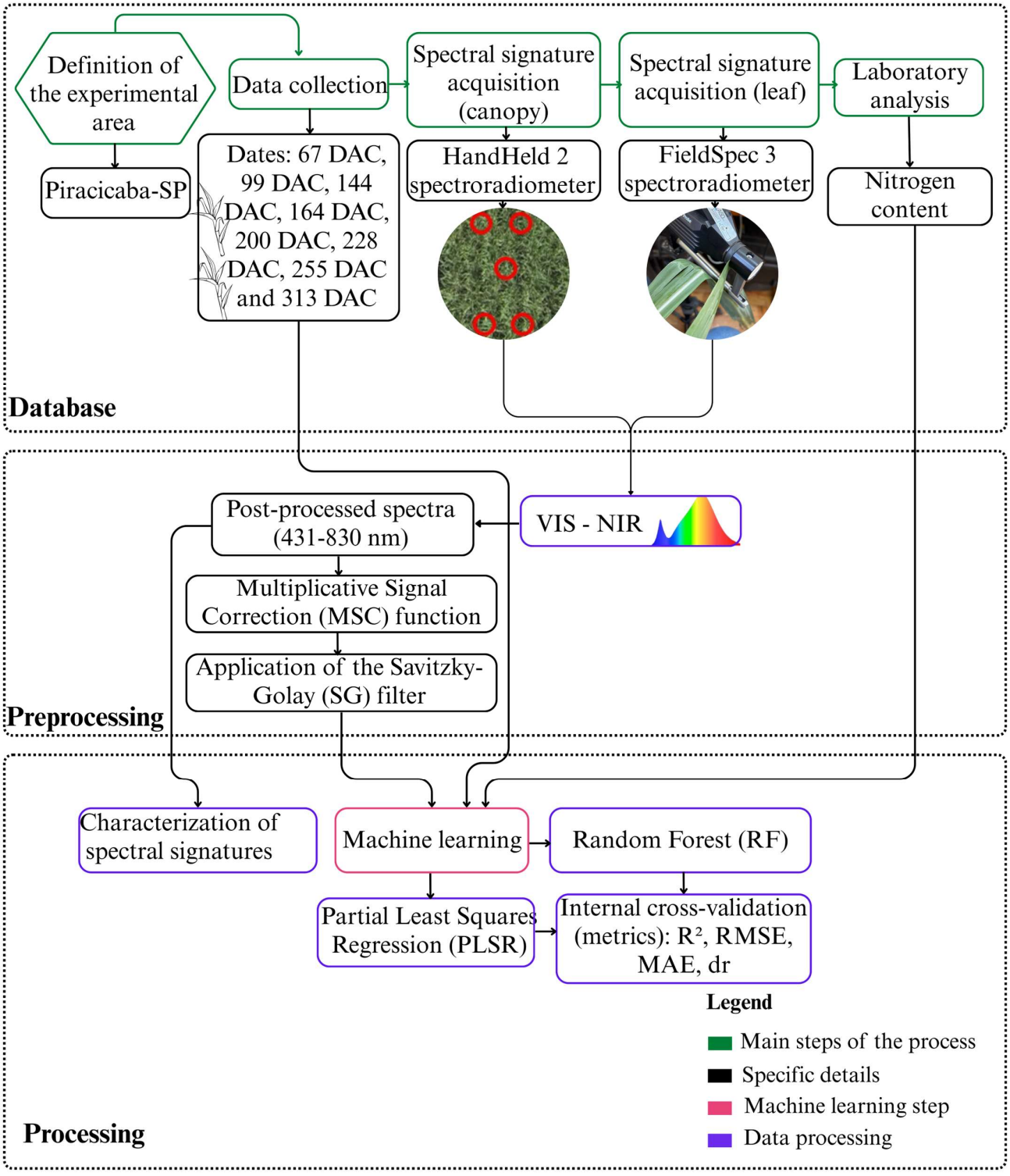

Figure 1. Flowchart representing the methodological steps.

2.1. Description of the Experiment

The experiment was conducted in the city of Piracicaba, located in the state of São Paulo, Brazil, specifically between the coordinates 22°41′02″ and 22°41′04″ S latitude and 47°38′44″ and 47°38′42″ W longitude (Figure 2). According to the Köppen climate classification, the region is characterized as having a humid subtropical climate (CWa), with an average annual precipitation of approximately 1280 mm [21]. The predominant soil in the study area is Red-Yellow Argisol, with a clayey texture.

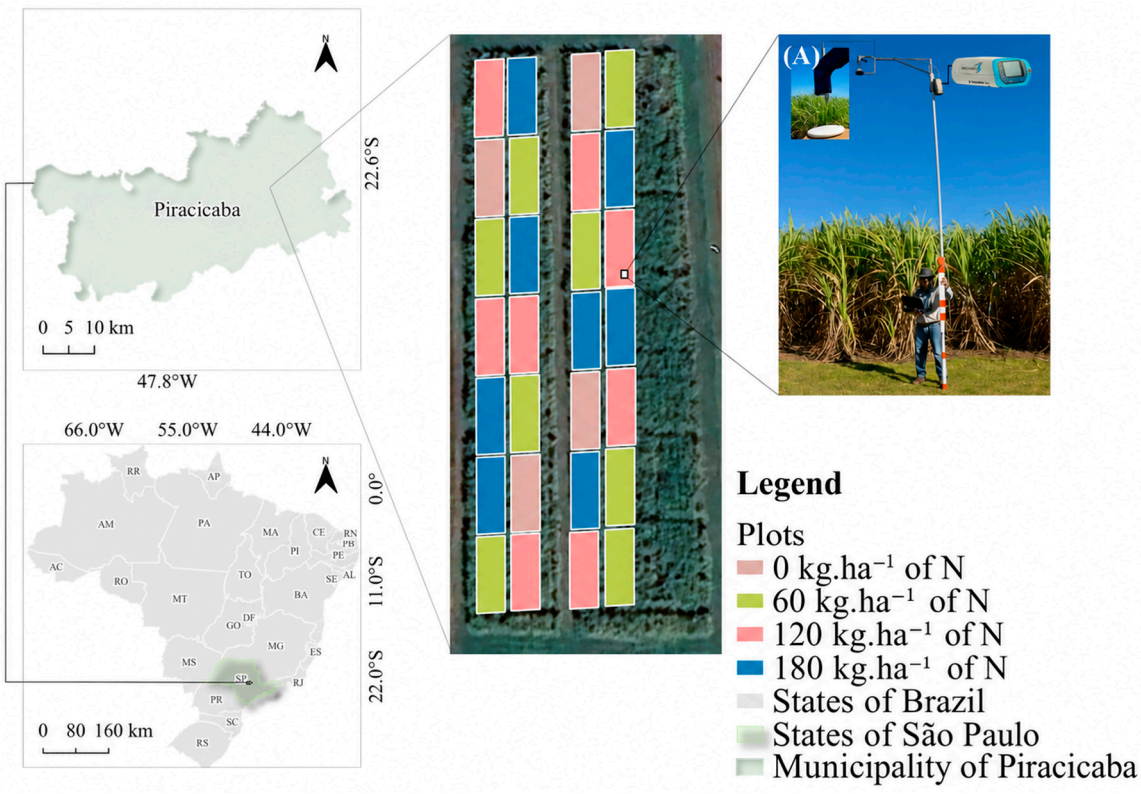

Figure 2. Location map of the experimental area used in the study, located at the APTA Midwest Regional Center in Piracicaba, São Paulo, Brazil. (A) Field sampling using the canopy sensor.

The experiment was conducted using a completely randomized design (CRD), with treatments corresponding to the nitrogen rates applied to the crop (0, 60, 120, and 180 kg ha^{-1}). These rates were part of the fertilizer management strategy used to support crop development and to promote variability in foliar nitrogen content; they were not used as a predictor variable or comparison factor in the estimation models. There were four replicates for the 0 kg ha^{-1} treatment and eight replicates for the other treatments (60, 120, and 180 kg ha^{-1}), totaling 28 experimental plots. Each plot consisted of five rows of sugarcane, each 15 m long and spaced 1.5 m apart. For the evaluations, only the three central rows were considered, disregarding 1.0 m from each end to avoid edge effects. Leaf and canopy samples were collected from the same clump, specifically at 67, 99, 144, 164, 200, 228, 255, and 313 days after cutting (DAC).

2.2. Acquisition of the Canopy Spectral Signature

Spectral reflectance was obtained using a FieldSpec® HandHeld™ 2 spectroradiometer (ASD—Analytical Spectral Devices, Boulder, CO, USA) [22]. This sensor performs spectral measurements in the range of 325 to 1075 nm, with a spectral resolution of 3 nm and a 25° field-of-view angle. Canopy measurements were taken at 1.5 m above the crop, providing a circular field of view of approximately 0.65 m in diameter and an area of 0.34 m^2. A Lambertian panel with known reflectance was used for sensor calibration [22]. Calibration was performed every 20 measurements. For each plot, five canopy measurements were taken at predetermined and georeferenced locations, always before leaf collection for leaf spectral analysis and determination of leaf nitrogen content (LNC).

2.3. Leaf Material Collection and Spectral Signature Acquisition

For laboratory analyses, five plants were selected and georeferenced in each of the 28 plots to ensure sampling accuracy. In each sugarcane stool, two +1 leaves were collected,

totaling 10 leaves per plot. Leaf samples were collected from plants located within the area covered by the field of view projected by the canopy sensor, in order to maintain spatial correspondence between the canopy measurements and the leaf material analyzed. Sampling was conducted on the middle third of the first fully expanded leaf (leaf + 1), counting from the plant's apex. The leaf material was identified, placed in plastic bags, transported in coolers, and sent to the laboratory.

To obtain the leaf spectral signature, a FieldSpec 3 ASD FR Spectroradiometer® [23] (ASD, 2010; ASD—Analytical Spectral Devices Inc., Boulder, CO, USA) was used with the Leaf Clip® attached. The spectroradiometer operates in the spectral range from 350 to 2500 nm, with a resolution of 1.4 nm between 350 and 1050 nm and 2 nm between 1050 and 2500 nm. Periodic calibrations were performed every 20 measurements using the Lambertian panel of the Leaf Clip®. After spectral measurements were taken on fresh leaves, the samples were dried in a forced-air oven at 65 °C until constant weight, subsequently ground, and sent for chemical analysis (Figure 3).

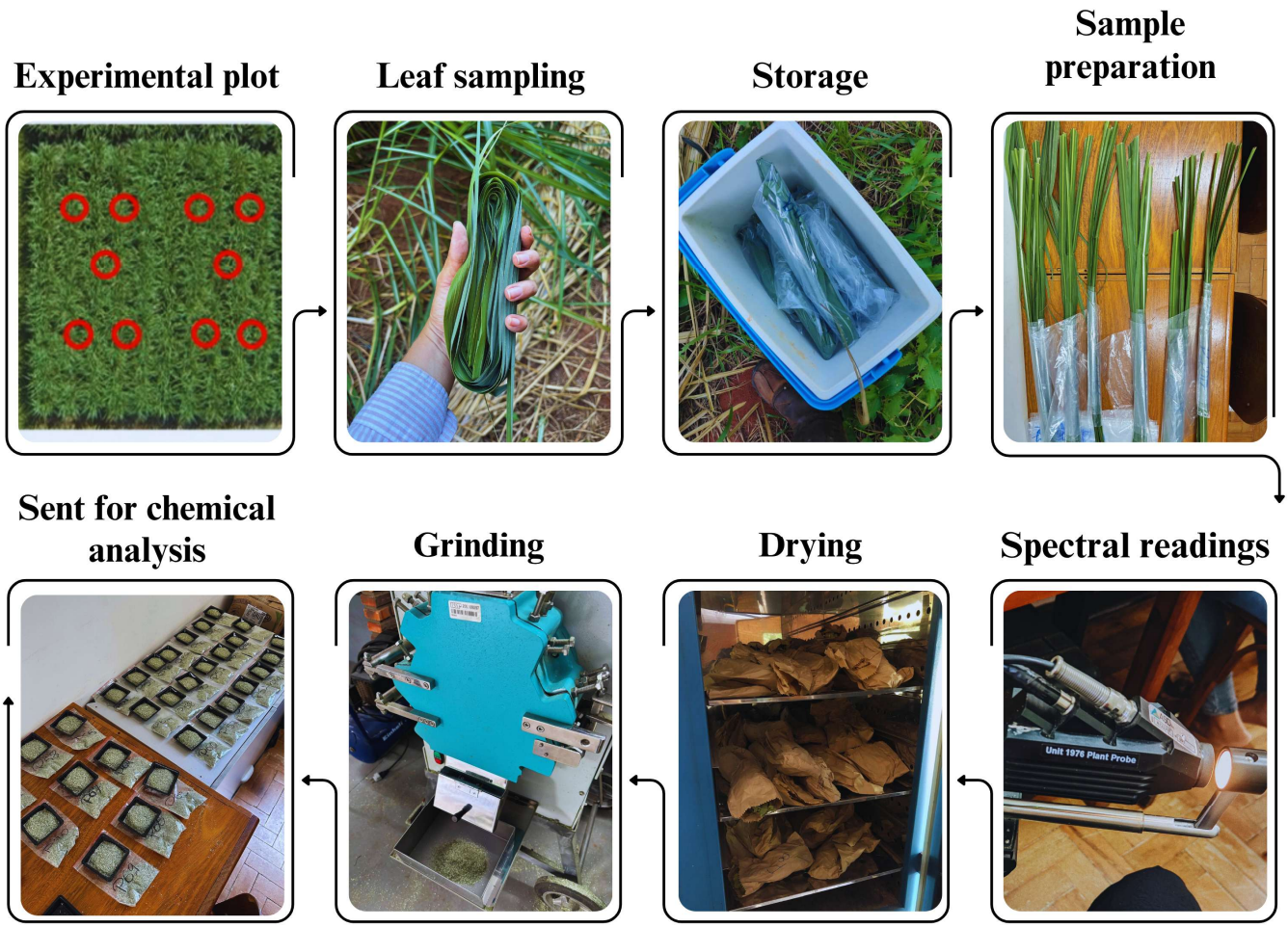

Figure 3. Stages of the experimental process involving leaf sampling and spectral data acquisition.

2.4. Spectral Preprocessing

To reduce distortions associated with instrumental noise, acquisition conditions, and radiation scattering effects [17], the same preprocessing workflow was adopted for both foliar and canopy data. However, all steps were performed separately for each sensor, without combining the spectra from the two acquisition levels.

Initially, the five spectral readings obtained at each experimental unit were organized by plot and assessment period. For each wavelength, the arithmetic mean of the read-

ings corresponding to the same plot and collection date was calculated, yielding a single representative spectrum per experimental unit and assessment period.

Next, the datasets were harmonized by selecting the spectral range common to both sensors, which ranges from 431 to 830 nm. For the canopy data, wavelengths between 325 and 430 nm and between 831 and 1075 nm were removed. For the foliar data, the regions between 350 and 430 nm and between 831 and 2500 nm were removed. The exclusion of wavelengths below 431 nm was due to the greater instability observed at the initial ends of the spectra, while the removal of regions above 830 nm was performed to maintain the same spectral range across both acquisition levels. Thus, the comparative analyses and models were developed using the same VIS-NIR spectral bands.

After organizing the readings and selecting the spectral range, Multiplicative Scatter Correction (MSC) was applied [24]. The procedure was performed separately for leaf and canopy data. For each dataset, the average spectrum was calculated using all representative spectra from the respective sensor and used as the reference spectrum, as described in Equation (1):

$$Spectra_{avg}(\lambda) = \frac{1}{n}\sum_{i=1}^{n} Spectra_i(\lambda) \quad (1)$$

where $Spectra_{avg}(\lambda)$ represents the mean spectrum of the respective dataset at wavelength λ, $Spectra_i(\lambda)$ corresponds to the i-th spectrum, and n represents the total number of spectra in the evaluated dataset.

For each individual spectrum, a linear regression was fitted relative to the average spectrum, as described in Equation (2):

$$Spectra_i(\lambda) = k_i\, Spectra_{avg}(\lambda) + b_i \quad (2)$$

where k_i represents the slope coefficient and b_i represents the intercept of the regression fitted for the i-th spectrum. The MSC-corrected spectrum was then obtained, as described in Equation (3):

$$Spectra_{MSC,i}(\lambda) = \frac{Spectra_i(\lambda) - b_i}{k_i} \quad (3)$$

MSC was used to reduce additive and multiplicative differences among spectra, mainly related to radiation scattering effects.

After MSC, the spectra were subjected to Savitzky–Golay smoothing [25] (Savitzky and Golay, 1964). The filter was applied using a three-point window, a second-order polynomial, and a zero-order derivative [26]. Therefore, at this stage, the procedure was used exclusively for spectral smoothing, without derivative calculation.

The complete sequence applied to the data was as follows: (i) organization of the readings by plot and evaluation period; (ii) calculation of the mean representative spectrum for each experimental unit; (iii) restriction of the data to the common spectral range from 431 to 830 nm; (iv) application of MSC; (v) Savitzky–Golay smoothing. All processing steps were performed separately for the leaf- and canopy-level datasets using ParLeS software, version 3.1, as described by Viscarra Rossel [27].

2.5. Spectral Characterization

To better understand and visualize the influence of leaf nitrogen content (LNC) on the spectral curves, two complementary techniques were applied to the post-processed spectra: Continuum Removal (CR) and First Derivative (FD). CR aims to eliminate absorption features that are not related to the target of interest, thereby isolating specific spectral features. According to Clark et al. [28] and Kokaly [29], this technique works as a convex hull fitted to the spectrum, in which each reflectance value within the absorption feature is divided by the corresponding value of the continuum line. In contrast, FD is used to

minimize the influence of background interference and spectral noise [30]. This processing enhances spectral features associated with radiation absorption in specific wavelength regions, facilitating the interpretation of chemical variations related to N content. It should be noted that the removal of the continuum and the first derivative were used only for spectral characterization and were not employed as input variables in the predictive models.

2.6. Pearson Correlation

To investigate the relationship between the Green and Red-Edge spectral regions and nitrogen content, Pearson's correlation coefficient was used. This feature selection technique is widely applied to quantify the degree of association between variables [31]. The coefficient values range from -1 to 1, indicating negative and positive correlations, respectively. Thus, the closer the value is to 1, the stronger the positive correlation between the variables, whereas values close to -1 indicate a strong negative correlation [32].

2.7. Machine Learning Models

2.7.1. Partial Least Squares Regression (PLSR)

PLSR is a multivariate technique based on Partial Least Squares (PLS), designed to automatically select the most relevant variables during the modeling process. The PLSR technique has been widely applied in predictive nutrient models based on hyperspectral data [14,17,33–35].

This technique is effective for handling highly correlated predictor variables, such as spectral bands, particularly in datasets with a large number of predictors relative to the number of observations. PLSR reduces dimensionality by constructing latent variables as linear combinations of the original predictors. These latent variables are derived to maximize the covariance between the predictor matrix and the response variable, thereby reducing the effects of multicollinearity and supporting the estimation of leaf nitrogen content [36]. During model fitting, an excessive number of latent variables may incorporate noise and lead to overfitting, whereas an insufficient number may fail to capture the relevant relationship between the spectral predictors and the response variable [37].

For each spectral acquisition level, leaf and canopy, a PLSR model was fitted independently. Foliar nitrogen content, expressed in g kg^{-1}, was defined as the response variable, while the reflectance values of the spectral bands and days after cutting (DAC), included as a categorical factor, were used as predictor variables. The inclusion of DAC as a factor allowed for the representation of different assessment periods without assuming a linear relationship between time after cutting and foliar nitrogen content.

No Principal Component Analysis (PCA) was applied prior to PLSR. Dimension reduction occurred internally within the model itself, through the construction of latent variables. These variables are linear combinations of the original predictors and are obtained in such a way as to simultaneously represent the variability of the explanatory variables and their covariance with the response variable. This feature allows PLSR to handle the large number of spectral bands and the strong multicollinearity that exists between adjacent wavelengths.

The model was fitted using the plsr function from the pls package in RStudio version 4.5.0 [38]. To internally evaluate model performance and select model complexity, 10-fold cross-validation was employed. Initially, the observations were randomly distributed among the ten folds. At each step, nine folds were used to fit the model, whereas the remaining fold was used to predict the observations not included in that calibration. The procedure continued until each fold had been used once as the validation subset.

The complexity of the PLSR was determined by the number of latent variables. For each number of components evaluated, predictions were generated using cross-validation,

and the sum of the squares of the prediction residuals, known as the Prediction Residual Error Sum of Squares (PRESS), was calculated according to Mevik and Wehrens [39]:

$$PRESS_A = \sum_{i=1}^{n} (y_i \mid \hat{y}_{i,-k,A})^2 \quad (4)$$

where $PRESS_A$ represents the PRESS value obtained using A latent variables, y_i is the observed foliar nitrogen content for the i-th sample, and $\hat{y}_{i,-k,A}$ is the value predicted for that sample by a model with A components, fitted without using the fold to which the observation belonged.

The optimal number of latent variables was defined as the one associated with the lowest PRESS value. The selection was performed using the 'which.min' function, applied to the PRESS values obtained for the different numbers of components evaluated. This procedure resulted in the selection of eight latent variables for the model based on foliar spectra and eight latent variables for the model based on canopy spectra.

After determining the optimal number of latent variables, a variable importance in projection (VIP) analysis was performed. The VIP values summarize the contribution of each spectral variable to the construction of the latent variables and to the explanation of foliar nitrogen content. For spectral interpretation, VIP values greater than 1 were considered indicative of wavelengths with an above-average contribution to the model, allowing for the identification of the most relevant spectral regions for nitrogen prediction.

2.7.2. Random Forest (RF)

RF is an algorithm developed by Breiman [40], commonly used independently for regression and classification problems. RF operates through a bootstrap aggregation, or a bagging process, in which each tree is trained using a subsample of the original dataset. Thus, each tree contributes an individual prediction, and the combination of these predictions produces the final estimate [41]. During the prediction process, the model uses the average or majority of the predictions, ensuring a balance in the error of each tree. This improves the model's generalization capacity [42].

Following the strategy for the RF model, at each level of foliar and canopy spectral acquisition, the models were fitted separately. Prior to RF tuning, the spectral bands were subjected to Principal Component Analysis (PCA), applied separately to the leaf and canopy datasets. PCA was used to reduce dimensionality and redundancy among highly correlated spectral bands. The number of components retained was defined as the smallest number necessary to explain at least 95% of the cumulative spectral variance. Using this criterion, four principal components were retained for the foliar data, explaining 96.17% of the variance, and three components for the canopy data, explaining 96.31%.

Thus, the model retained LNC as the response variable, while the PCA scores derived from the spectral bands and the DAC included as a categorical factor were used as predictor variables. The DAC was included to account for temporal differences between assessment periods and variations associated with crop development, which can simultaneously influence both spectral behavior and LNC, without assuming a linear relationship with time. Previous studies have also shown that the relationship between spectral response and LNC can vary throughout the sugarcane growth cycle [17].

The RF script was configured with 1000 trees (ntree = 1000), 11 variables selected at random in each split (mtry = 11), a minimum size of 20 observations at the terminal nodes (min.node.size = 20), and a splitting criterion based on variance reduction (splitrule = "variance"). The selection of these hyperparameters was based on previous studies [43,44] and on the criteria presented by Chen et al. [37], while their values were defined empirically during preliminary adjustments specific to the foliar and canopy datasets.

2.8. Model Validation

For each acquisition level, model performance was evaluated exclusively through internal 10-fold cross-validation using the 224 observations obtained from a single field experiment. The observations were randomly divided into ten subsets of approximately equal size. In each iteration, nine subsets were used to fit the model, while the remaining subset was used for validation. This procedure continued until each subset had been used once as the validation set. No independent external dataset was used; therefore, the reported metrics represent the internal predictive performance of the models under the evaluated experimental conditions. The predictive accuracy of the models was assessed using the coefficient of determination (R^2), the root mean square error (RMSE), and the mean absolute error (MAE) described in Equations (5)–(7), respectively. The agreement between observed and predicted values was assessed using Willmott's refined agreement index (dr) [45], presented in Equation (8).

$$R^2 = \frac{\left[\sum(\gamma_p - \bar{\gamma}_p) \cdot (\gamma_o - \bar{\gamma}_o)\right]^2}{\sum(\gamma_p - \bar{\gamma}_p)^2 \cdot \sum(\gamma_o - \bar{\gamma}_o)^2} \quad (5)$$

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

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

$$dr = 1 - \frac{\sum_{i=1}^{n}(Y_i - X_i)^2}{\sum_{i=1}^{n}\left(|Y_i - \overline{\overline{X}}| + |X_i - \overline{\overline{X}}|\right)^2} \quad (8)$$

3. Results

3.1. Temporal Variation in Leaf Nitrogen Content

Figure 4 shows the distribution of precipitation in the experimental area and the variation in LNC over the eight evaluation periods, conducted at 67, 99, 144, 164, 200, 228, 255, and 313 days after cutting (DAC). The highest LNC values were observed at 67 DAC, while lower values were recorded, particularly at 99 and 228 DAC. During the remaining evaluation periods, the LNC values varied throughout the crop cycle.

With the exception of the assessment conducted at 67 DAC, the mean LNC values remained below the reference range of 18 to 20 g kg^{-1} recommended for this crop [46]. Differences in rainfall distribution were also observed throughout the experimental period, including above-average rainfall in December. However, since variations in precipitation, crop development, and other environmental factors occurred simultaneously over the course of a single growing season, Figure 4 should be interpreted descriptively. The experimental design of this study does not allow changes in LNC to be attributed solely to precipitation or to the crop's stage of development.

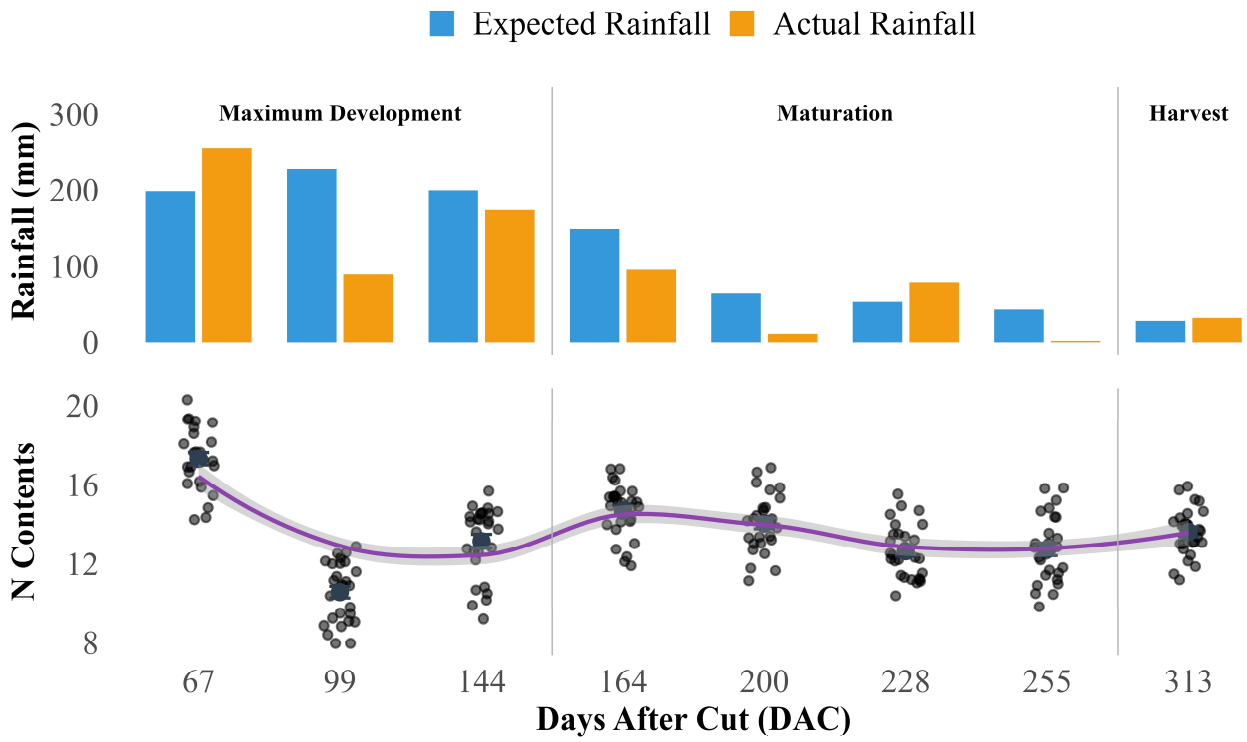

Figure 4. Precipitation distribution and variation in nitrogen content in sugarcane leaves throughout the growth cycle, assessed at 67, 99, 144, 164, 200, 228, 255, and 313 days after harvest. The line represents the smoothed trend of the nitrogen content values in the leaves, and the points correspond to the nitrogen samples.

3.2. Influence of Leaf Nitrogen Content (LNC) on Sugarcane Spectral Behavior at Each Acquisition Level

In this study, two narrow hyperspectral bands were analyzed: the Green band, fixed at 550 nm, and the Red-edge band, fixed at 738 nm, both recognized for their sensitivity to nitrogen content [17]. Pearson correlation analysis was performed to evaluate the relationship between N content and the reflectance values of these bands, following the developmental stages of the sugarcane crop at two hyperspectral acquisition levels, as illustrated in Figure 5.

Pearson correlation coefficient (r) values for the Green and Red-edge bands were observed at both acquisition levels, leaf and canopy. In the analysis of the 67 DACs, the foliar level showed a negative correlation between LNC and the green band ($r = -0.74$) and a positive correlation with the red edge ($r = 0.69$). At the canopy level, correlations of lesser magnitude were observed, with $r = -0.43$ for the green band and $r = 0.47$ for the red edge. In contrast, at the canopy level, during the same sampling date, a negative correlation ($r = -0.43$) was observed for the Green band and a positive correlation ($r = 0.47$) for the Red-edge band. Although the canopy was not yet fully developed, the leaves responded under a reduced canopy N dilution effect. Overall, the leaf sensor showed a better response, with a negative correlation ($r = -0.63$) for the Green band and a positive correlation ($r = 0.64$) for the Red-edge band.

Among the time periods evaluated, strong correlations were observed at 144 DAC at both levels of acquisition. At this stage, the leaf sensor showed a negative correlation ($r = -0.66$) in the Green band and a positive correlation ($r = 0.65$) in the Red-edge band. In contrast, the canopy sensor exhibited even stronger correlations, with ($r = -0.77$) for the Green band and ($r = 0.79$) for the Red-edge band. It is important to highlight the inversely proportional behavior between leaf nitrogen content (LNC) and the Green band in both

sensors; this correlation indicates that as N content increases, spectral values decrease. This is consistent with the crop spectral signature as a function of N values, in which leaves with higher N contents have a lower reflectance factor in the Green band (Figure 6A,B).

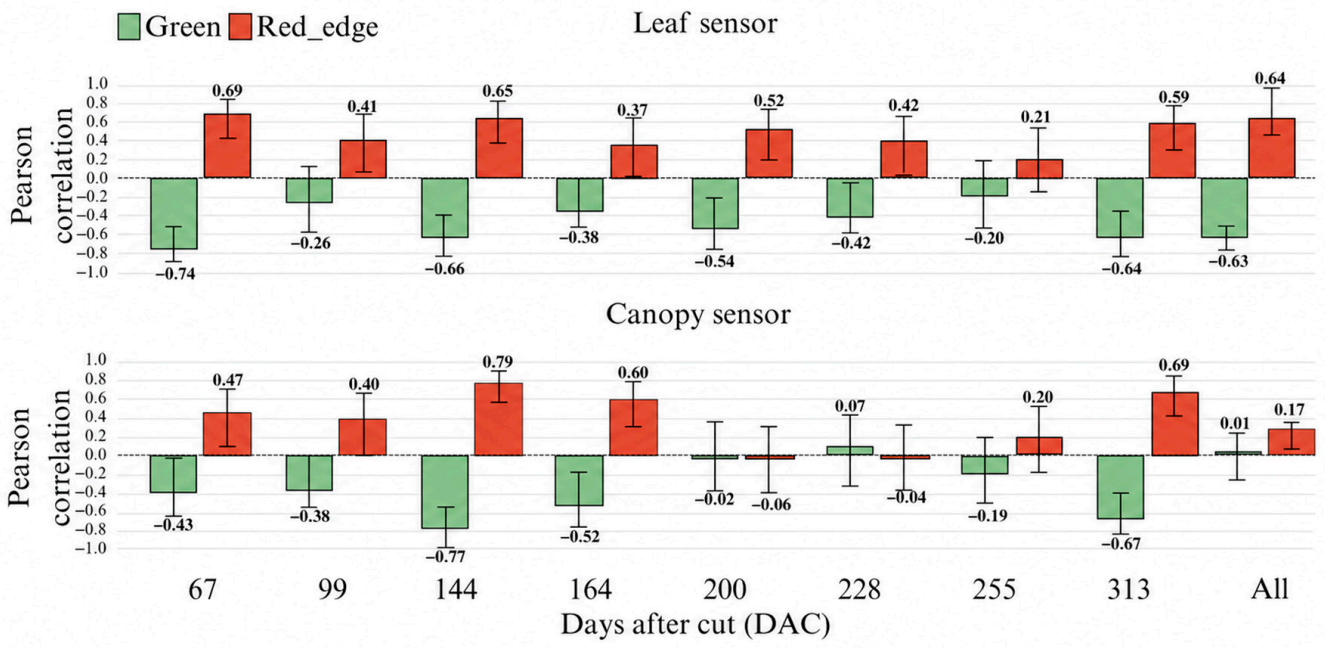

Figure 5. Pearson correlation between N content and the Green and Red-edge bands at different acquisition levels in sugarcane throughout the crop cycle, at different days after cutting (DAC).

Based on the correlation observed at 144 DAC between the two acquisition levels, two spectral techniques were applied to highlight the sensitivity of the spectrum to nitrogen content. Figure 6A,B show the spectrum after Continuum Removal (CR), a technique used to maximize local spectral interpretation by highlighting variations in band depth in the Green region. In contrast, Figure 6C,D illustrate the result of applying the First Derivative (FD), calculated to express the slope of the spectrum in the Red-edge region.

At 144 DAC, the continuum-removed leaf and canopy data showed greater absorption depth in the Green region, associated with increased nitrogen content. When analyzed separately, leaf-level CR stood out for the greater contrast among N levels, with more asymmetric features and broader internal curve shapes, indicating higher sensitivity to N variations. The deeper features centered at wavelengths in the Green region reflect differences in nitrogen availability in the leaf. These differences highlight variations in the spectral features associated with the different N levels observed during this evaluation period. Although this technique helps reduce this bias caused by different levels of spectral acquisition, compared with what was obtained in the laboratory, plant geometry, leaf overlap, shading, light incidence, soil background, and leaf water content may still cause changes in the reflectance signal.

In addition, the application of the first derivative made it possible to more clearly highlight spectral transitions related to nitrogen content, emphasizing variations in the Red-edge region (Figure 6C,D). This technique enhanced the differences in the slope of the spectral curve, making it possible to identify rightward shifts as N content increased. At the leaf level, these variations were visually more pronounced than those observed at the canopy level. At the canopy level, although the same shift pattern was observed, the response showed lower contrast, which may be due to the combined influence of canopy structure, acquisition geometry, and lighting conditions.

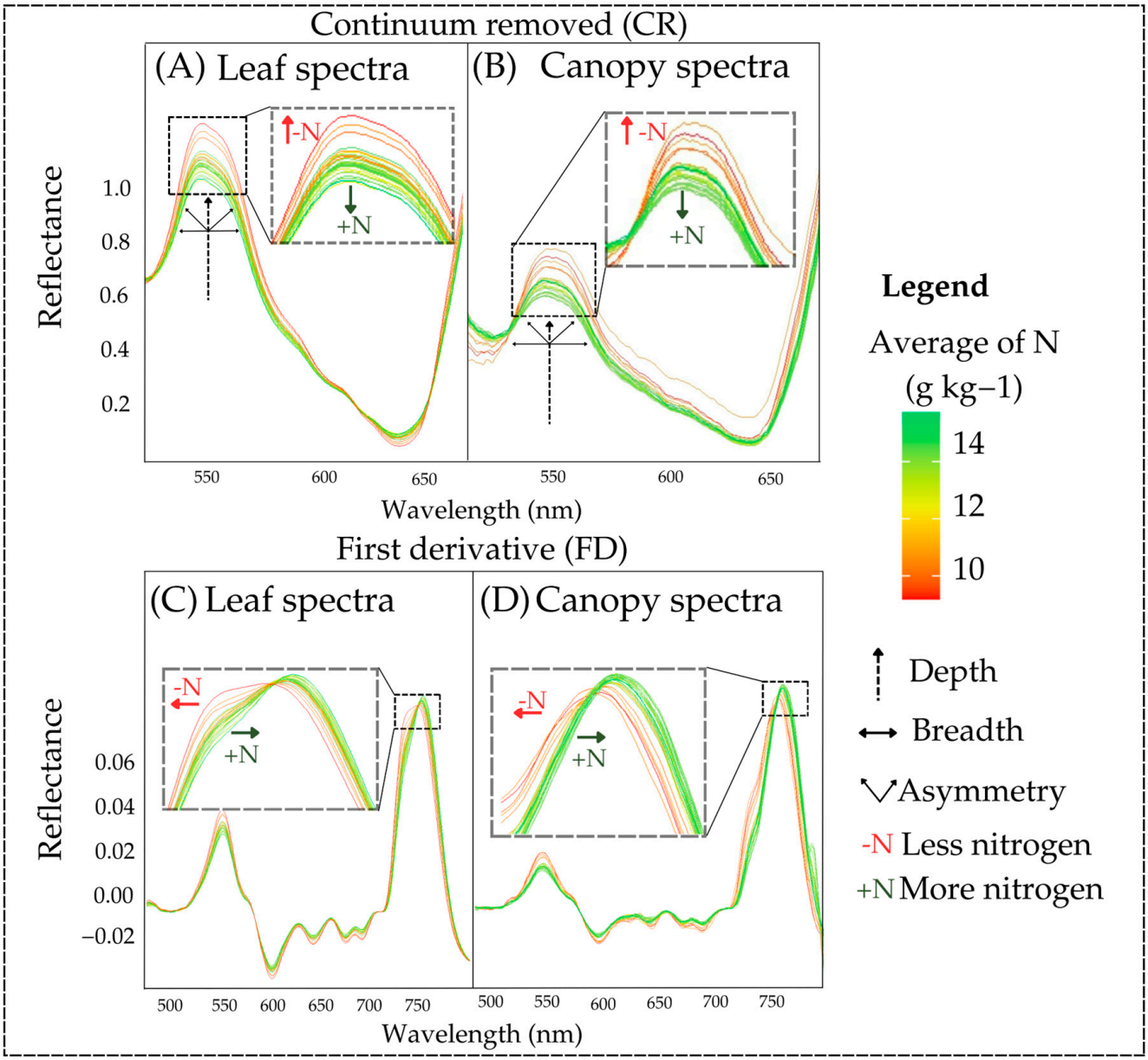

Figure 6. Spectral preprocessing of the data at 144 DAC, representing the continuum-removed leaf spectra (**A**); continuum-removed canopy spectra (**B**); first derivative of the leaf spectra (**C**); and first derivative of the canopy spectra (**D**).

3.3. N Prediction Using VIS and NIR Spectra

Nitrogen prediction based on the VIS-NIR spectra from both sensors was performed using two machine learning models. The datasets used included raw spectra and preprocessed spectra, as described in Section 2.4. When the raw dataset was used, the models from both databases showed moderate performance, highlighting the importance of spectral preprocessing for improving predictive accuracy. The low R^2 values and the high errors observed for the canopy data (R^2 = 0.37–0.51; RMSE = 1.75–2.19 g kg^{-1}; dr < 0.8) indicate that the raw spectra contained noise, which compromised the ability of the models to capture consistent relationships between the spectral variables and nitrogen content (Table 1).

Table 1. Performance of the models Random Forest (RF) and Partial Least Squares Regression (PLSR) on the estimate of nitrogen content in the dataset without preprocessing.

Acquisition Level	Model	R^2	RMSE	MAE	dr
Leaf	RF	0.61	1.75	1.35	0.83
	PLSR	0.65	1.61	1.22	0.89
Canopy	RF	0.51	1.75	1.34	0.79
	PLSR	0.37	2.19	1.37	0.77

For the dataset with the preprocessed leaf data, the model PLSR presented the best performance, with a coefficient of determination (R^2) of 0.727, RMSE of 1.381 g kg^{-1}, MAE of 1.109 and "dr" index of 0.917 (Figure 7A). The model RF, although presenting a slightly inferior performance (R^2 = 0.710, RMSE = 1.422 g kg^{-1}, MAE = 1.113, dr = 0.909), also had good predictive capabilities (Figure 7B).

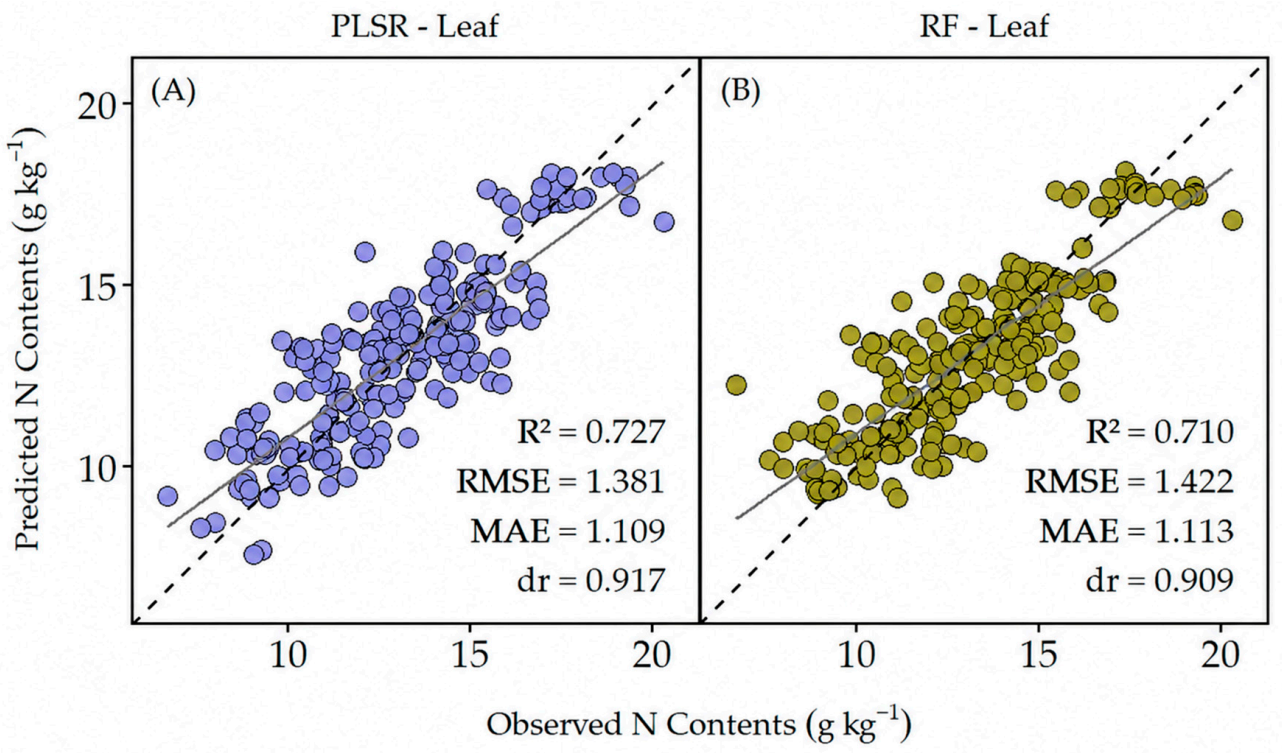

Figure 7. Estimate of sugarcane nitrogen content using the models Partial Least Squares Regression—PLSR (**A**) and Random Forest—RF (**B**) with data preprocessing for the leaf sensor. The solid gray line represents the fitted regression line, while the black dashed line represents the 1:1 reference line between observed and predicted nitrogen contents.

Similar results were also observed for the models calibrated with canopy data, with PLSR reaching R^2 = 0.591, RMSE = 1.489 g kg^{-1}, MAE = 1.157 and dr = 0.866, whereas RF presented R^2 = 0.566, RMSE = 1.559 g kg^{-1}, MAE = 1.254 and dr = 0.813 (Figure 8A,B). The model PLSR, in its essence, derives orthogonal components in order to avoid the multicollinearity that exists in the reflectance data; therefore, the model RF was trained using data derived from the PCA to provide a similar reduction in data dimensionality and multicollinearity.

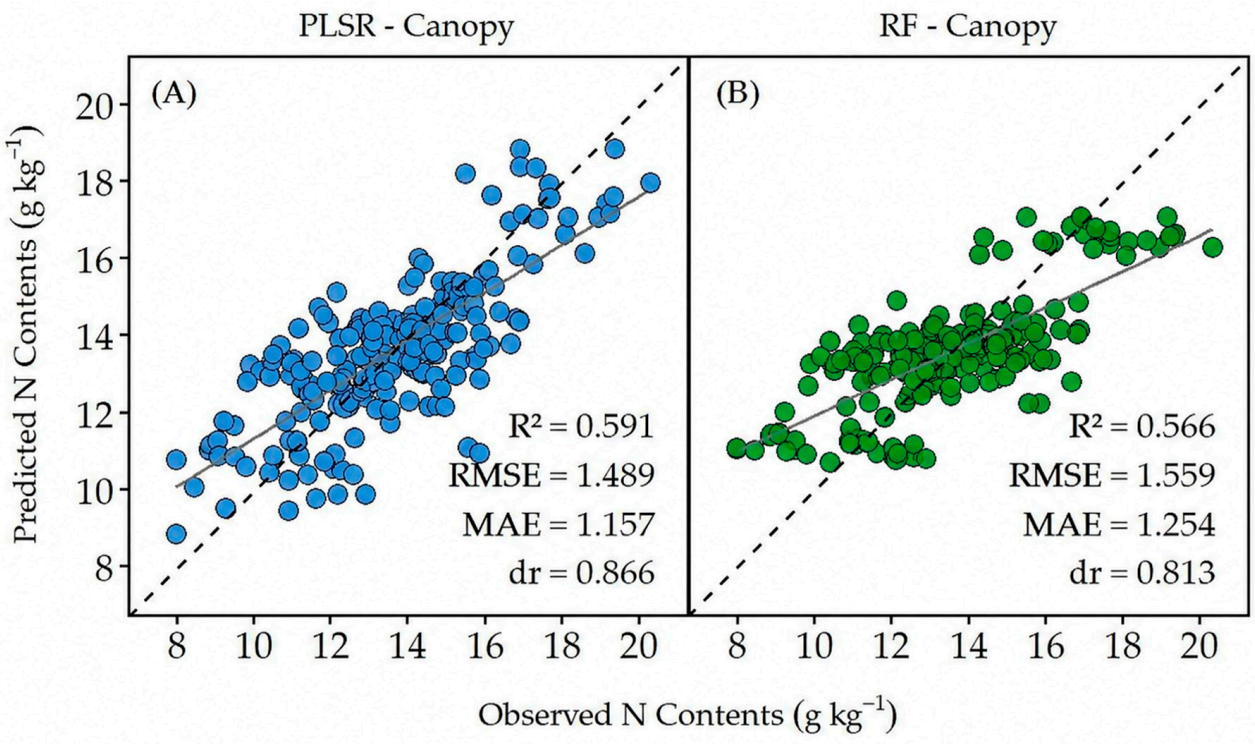

Figure 8. Estimate of the nitrogen content in sugarcane using the models Partial Least Squares Regression—PLSR (**A**) and Random Forest—RF (**B**) with data preprocessing for the canopy sensor. The solid gray line represents the fitted regression line, while the black dashed line represents the 1:1 reference line between observed and predicted nitrogen contents.

Overall, the PLSR model performed slightly better than the RF model, reflecting a better fit to the spectral data evaluated. Agreement was higher in the leaf-level models, while the canopy-level models had lower dr values, ranging from 0.813 to 0.866.

3.4. Variable Importance in Projection

The importance analysis in projection (Vips) of the PLSR model (Figure 9) identified two spectral regions as the most informative for predicting leaf nitrogen (N) content: the Green band (520–550 nm) and the Red-edge band (630–750 nm). For both sensors, in both leaf and canopy measurements, VIPs values above the heuristic threshold of 1 were observed, indicating that these spectral variables contribute more than average to explaining nitrogen variability. Considering that the models were generated from data from all samplings, and that the spectral response varies according to the crop's vegetative stage, differences in Vips values were observed between the sensors. These differences may be associated with the distinct methods of acquiring the spectral signature, which vary by device; for example, the foliar sensor measures a single leaf under controlled conditions. In contrast, the canopy sensor captures the integrated reflectance of the canopy, where the spectral response may include contributions associated with shadows, exposed soil, mulch, and plant geometry (Figure 9B, at 67, 144, and 255 DAC). Furthermore, when comparing the acquisition levels, it is noted that the region near 630 nm shows the greatest discrepancy between leaf and canopy data. These differences in data acquisition methods directly affected model performance (Figures 7 and 8) and, consequently, the VIPs values, particularly in the blue bands (400–500 nm) and red (600–680 nm).

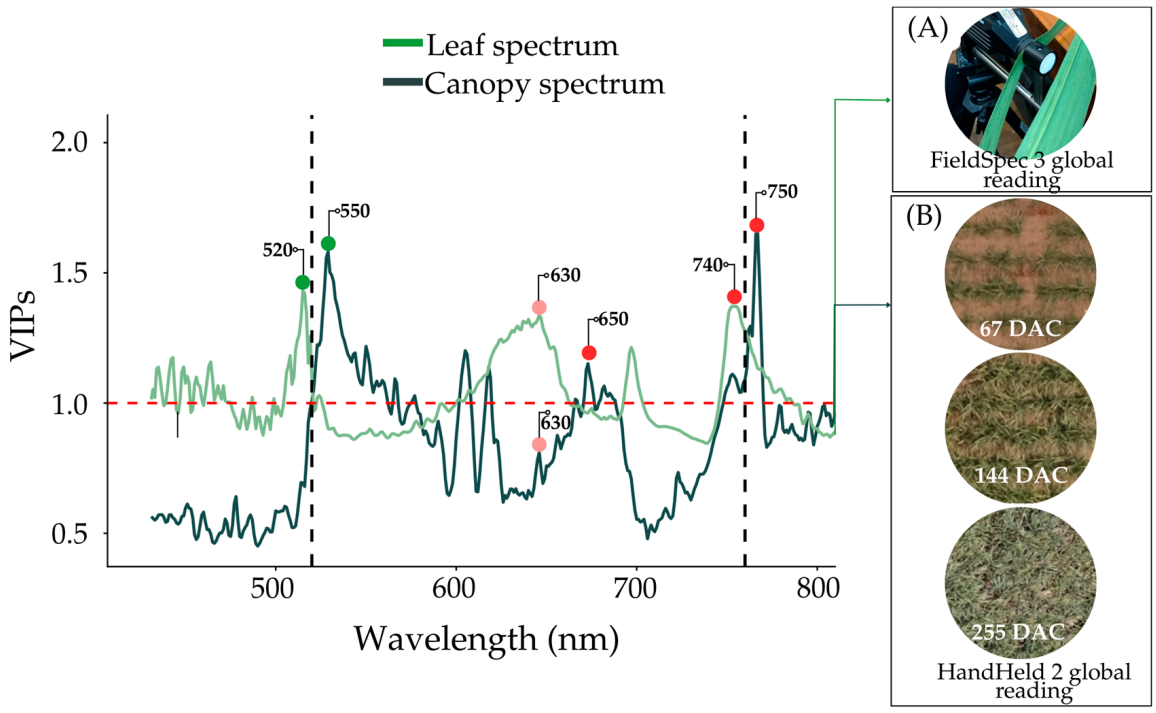

Figure 9. Analysis of the global wavelengths with the greatest importance in the prediction of nitrogen of both sensors. The dotted red line represents the threshold for the selection of the most relevant spectral bands in the PLSR model. (**A**) Leaf sensor; (**B**) illustrative images showing the viewing angle of the canopy sensor under canopy closure.

4. Discussion

4.1. Levels of Acquisition and Spectral Features Related to Nitrogen (N)

The structure and chemical composition of plant tissues determine their optical properties, allowing characteristics such as chlorophyll content, photosynthetic rate, water content, and nitrogen content to be estimated from spectra measured at different acquisition levels [47,48]. However, the geometric levels of acquisition significantly influence the spectral response, since rough surfaces, such as sugarcane leaves, produce diffuse reflections that reach the sensor heterogeneously [49]. This influence is evidenced by the VIPs, CR, and FD, mainly reflecting the effects of the data acquisition environment. As highlighted by Li et al. [19], measurements performed under controlled conditions at the leaf level show minimal environmental interference, whereas canopy-level measurements performed in the field are subject to solar elevation, leaf angle, soil type, and moisture conditions.

Spectral sensitivity to nitrogen (N) content results from optical interactions at the leaf and canopy levels and is strongly influenced by the absorption of incident energy by the plant, a process dependent on chlorophyll [50]. Since chlorophyll contains nitrogen in its structure, variations in this nutrient directly affect its concentration, altering the spectral response. In the distribution of VIPs (Figure 9) above the importance threshold (VIP > 1) [51], both acquisition levels show that, in the visible region, the reflectance peak in the Green region, approximately 520–550 nm, reflects changes in leaf pigments, which are directly related to changes in chlorophyll content and, consequently, nitrogen content [35]. Subsequently, the spectral transition observed at the red edge, between 630 and 740 nm at the leaf level and 650 and 750 nm at the canopy level, follows the increase in reflectance, which is linked to the role of nitrogen in maintaining cellular structure. Studies such as Almeida et al. [52] support this interpretation, highlighting spectral regions for chlorophyll detection and possible nutritional stress, between 535 and 640 nm at the leaf level and 560 and 700 nm at the canopy level.

As demonstrated by the VIPs, the region around 630 nm showed a significant contribution to the leaf-level model. This range, located in the transition toward the red edge, is highly sensitive to structural and biochemical variations in leaves, reflecting interactions associated with pigments and cellular structure [53]. Similar results were reported by Fan et al. [54], who observed a response near 658 nm in corn leaves. At the canopy level, the lower contribution observed around 630 nm may be related to the integrated response of the different components present within the sensor field of view, as discussed by Alexandre et al. [55] and Palla et al. [56]. In addition to vegetation, field measurements may include contributions from soil, straw, shadows, and canopy geometry. However, these components were not quantified separately in this study, making it impossible to determine their individual contribution to the spectral response or to model performance.

The spectral characterization of the curves in relation to nitrogen content using CR subtly enhanced variations in absorption and in the positioning of features associated with N, as described by Sun et al. [57]. A similar behavior was observed in this study, showing that differences in leaf N content are consistently expressed in the Green region. In this spectral interval, greater absorption depth and better distinction among spectral features were observed at 144 DAC, especially between the different acquisition levels. Studies conducted by Silva et al. [16] also identified that leaves with higher nitrogen concentrations show greater capacity to absorb electromagnetic radiation in the visible region of the spectrum at 140 DAC, due to higher photosynthetic activity.

The application of FD confirmed the previously observed correlation between the Red-edge region and nitrogen content (Figure 6C,D), particularly at 144 DAC, a stage characterized by intense vegetative growth according to Martins et al. [58]. This approach highlighted differences in the displacement of the spectral curves, showing that as N content increases, the curve shifts toward the near-infrared region. Although the magnitude of this shift varied between acquisition levels, leaf, and canopy, the response pattern was consistent, indicating greater sensitivity to nutritional variations. This behavior is associated with the absorption of pigments such as chlorophylls and carotenoids [59] and supports previous findings relating N content to the Red-edge region [17,60,61].

4.2. N Prediction: Leaf and Canopy Sensors

The results show that the accuracy of leaf nitrogen prediction in sugarcane was influenced by the acquisition level, preprocessing procedures, and spectral resolution. The superior performance of leaf-level measurements was associated not only with spectral resolution, which may improve the modeling of relationships between spectral responses and nitrogen content [17], but also with the controlled acquisition conditions and the specificity of the measured target. Differences in measurement distance and acquisition geometry may also have influenced the intensity and quality of the reflected signal.

Although the same sets of foliar and canopy data were used in both modeling approaches, it is important to recognize that PLSR and RF handled spectral dimensionality differently. PLSR internally reduces the effects of multicollinearity through latent variables, while RF was trained using components derived from PCA to reduce spectral redundancy and dimensionality. Therefore, the comparison between PLSR and RF should be interpreted with caution, as it reflects different modeling strategies rather than an isolated comparison between algorithms. However, the main objective of this study was to evaluate the influence of the spectral acquisition level on the prediction of foliar N content.

Similarly, other studies have highlighted that the accuracy of nutrient prediction in sugarcane varies according to the spectral acquisition level, whether it be leaf or canopy, and the nutrient evaluated. In studies that directly measured leaves [17,60], the ability to detect multiple nutrients and predict N was very good, especially in the visible and

Red-edge bands. Fiorio et al. [17] reported the highest R^2 values for N prediction using leaf spectra (0.81). On the other hand, studies using canopy measurements [12] also achieved useful N estimates (R^2 = 0.50; RMSE = 1.67 g kg^{-1}). Reyes-Trujillo et al. [61] also reported that spectra obtained from the canopy allowed nitrogen to be modeled with excellent performance (R^2 = 0.98; RMSE = 0.98 g kg^{-1}), especially when sugarcane was still in its early growth stages.

Leaf-level measurements were conducted in the laboratory, directly on individual leaves and under controlled illumination and geometric conditions, whereas canopy acquisitions were performed in the field and integrated effects associated with variability in solar illumination, shadows, leaf orientation and overlap, the presence of straw and exposed soil, and crop phenological stage [12,15,62]. These differences in acquisition conditions and target characteristics may contribute to the greater variability observed in canopy spectra, whereas the standardized measurement environment and the specificity of the foliar target promote greater consistency between reflectance and leaf chemical properties.

The use of spectral preprocessing was important for reducing distortions in the spectral signal. When raw spectra were used, both models showed lower performance, particularly for canopy measurements (R^2 < 0.51; dr < 0.8). This result may reflect the greater variability of the raw canopy spectra; however, the individual sources of this variability were not quantified separately. Nevertheless, PLSR, particularly when combined with preprocessed spectra, achieved the best performance at both acquisition levels, with the highest accuracy observed for leaf-level measurements (R^2 = 0.727; RMSE = 1.381 g kg^{-1}; MAE = 1.109; dr = 0.917).

Preprocessing, by correcting for scattering effects and high-frequency noise, improved the signal-to-noise ratio and, consequently, increased the robustness of the models, especially for PLSR. Similar results were reported by Silva et al. [16] and Reyes-Trujillo et al. [61], who observed high accuracy of the PLSR method in estimating nitrogen concentration from spectral data, both for canopy reflectance and for spectra obtained using a benchtop spectroradiometer. The structural effects of the canopy, combined with spectral noise, underscore the importance of this step, which is essential for improving the accuracy of predictive models [63].

The Variable Importance in Projection (VIP) analysis supported the previous results, showing that wavelengths located in the Green and Red-edge regions were the most relevant for predicting leaf nitrogen concentration for both sensors. A similar result was reported by Martins et al. [58], who observed that variations in leaf nitrogen content were strongly influenced by these spectral regions. Similarly, Fiorio et al. [17] found that the models used to predict nitrogen performed better when restricted to the visible and Red-edge spectra, reaching R^2 values greater than 0.81 and RMSE values lower than 1.24 g kg^{-1}. These results confirm the sensitivity of the visible and Red-edge bands for estimating nitrogen content, since this element is directly associated with the concentration of leaf pigments, particularly chlorophyll [64].

In general, both sensors proved to be sensitive to variations in foliar nitrogen content, albeit with different response magnitudes. The performance results obtained in this study were moderate to satisfactory (R^2 > 0.55), suitable for evaluation and correlation, and served as good quantitative models, especially in the case of the foliar sensor (R^2 = 0.727), highlighting the consistency and potential of reflectance spectroscopy for non-destructive estimates of nitrogen content in sugarcane.

4.3. Limitations and Future Prospects

Although this study provides relevant evidence on the influence of spectral acquisition level and instrument configuration on the estimation of leaf nitrogen content in sugarcane,

some limitations should be acknowledged. The experiment was conducted at a single location, under specific soil, climate, and management conditions, and focused on sugarcane (*Saccharum* spp.). Therefore, the varietal diversity of the crop was not assessed. Because genetic differences among cultivars can influence foliar, structural, and biochemical characteristics and, consequently, spectral responses, the results should be interpreted within the context of the evaluated experimental conditions. Furthermore, although repeated assessments were conducted throughout the crop cycle, the transferability of these models to other regions, growing seasons, soil types, cultivars, and production systems still requires further validation.

In this study, the common spectral range in the VIS–NIR allowed for a comparison between the two sensors. However, the short-wave infrared region may exhibit absorption features associated with proteins, nitrogen-containing compounds, and other biochemical constituents. Thus, future research should evaluate whether preserving or carefully selecting bands in the SWIR region can improve nitrogen prediction, especially in foliar measurements conducted under controlled conditions.

Despite the use of PLSR and RF, two algorithms widely applied to hyperspectral data modeling, future research should include a more comprehensive evaluation of algorithms, systematic optimization of hyperparameters, repeated resampling, uncertainty estimation, and statistical tests to compare model performance. Approaches such as hybrid models, inversion of radiative transfer models, ensemble methods, and deep learning combined with independent validation and strategies such as grouped k-fold or leave-one-plot-out can help increase the robustness of predictions and expand the operational applicability of the models.

In turn, canopy-level measurements, although more suitable for field-scale applications, can be influenced by crop structure, leaf angle, solar radiation, shading, the presence of exposed soil and straw, and plant water status. These uncontrolled sources of canopy-level variability may have contributed to the lower predictive performance; however, their individual effects were not quantified in this study. Therefore, future research should incorporate structural and environmental variables, such as canopy height and closure, vegetation indices, soil background characteristics, straw cover, plant water status, and meteorological data.

Finally, the practical application of the models should be evaluated based on diagnostic ranges and agronomic thresholds for nitrogen sufficiency. This evaluation will determine whether prediction errors are acceptable for supporting nutritional diagnosis and fertilizer recommendations for sugarcane.

5. Conclusions

At 144 DAC, the strongest relationships between spectral responses and leaf nitrogen content were observed at both acquisition levels, with stronger correlations for canopy measurements, reaching $r = -0.77$ at 550 nm and $r = 0.79$ at 738 nm. These relationships indicate variations in spectral features associated with the different N contents observed at this evaluation stage. However, the differences between acquisition levels reflect the combined effects of instrumental configurations and measurement conditions.

Among the evaluated algorithms, PLSR achieved the best performance at both acquisition levels, with R^2 values of 0.727 for leaf data and 0.591 for canopy data. The green region (~550 nm) and the red-edge region (~740 nm) contributed most strongly to nitrogen estimation, highlighting their relevance for characterizing spectral variations associated with the crop's nutritional status.

Overall, the observation scale, acquisition conditions, and data processing were associated with the performance of the spectral models. The results improve the un-

derstanding of the differences between leaf- and canopy-level measurements; however, because they were obtained through internal cross-validation using data from a single experiment, external validation is still required before the models can be applied under other production conditions.

Author Contributions: Conceptualization, I.d.L.e.L., M.L.d.S.A., R.R., A.K.d.S.O., 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. and C.A.A.C.S.; software, I.d.L.e.L. and M.L.d.S.A.; validation, I.d.L.e.L., M.L.d.S.A., R.R. and C.A.A.C.S.; formal analysis, I.d.L.e.L., M.L.d.S.A. and A.K.d.S.O.; 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. and 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: Thanks 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, and 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. Cherubin, M.R.; Carvalho, J.L.N.; Cerri, C.E.P.; Nogueira, L.A.H.; Souza, G.M.; Cantarella, H. Land use and management effects on sustainable sugarcane-derived bioenergy. *Land* **2021**, *10*, 72. [CrossRef]
2. De Souza, A.P.; Grandis, A.; Leite, D.C.C.; Buckeridge, M.S. Sugarcane as a bioenergy source: History, performance, and perspectives for second-generation bioethanol. *BioEnergy Res.* **2014**, *7*, 24–35. [CrossRef]
3. Boschiero, B.N.; Mariano, E.; Torres-Dorante, L.O.; Sattolo, T.M.S.; Otto, R.; Garcia, P.L.; Dias, C.T.S.; Trivelin, P.C.O. Nitrogen fertilizer effects on sugarcane growth, nutritional status, and productivity in tropical acid soils. *Nutr. Cycl. Agroecosyst.* **2020**, *117*, 367–382. [CrossRef]
4. Oliveira, M.P.G.; Bocca, F.F.; Rodrigues, L.H.A. From spreadsheets to sugar content modeling: A data mining approach. *Comput. Electron. Agric.* **2017**, *132*, 14–20. [CrossRef]
5. Li, F.; Miao, Y.; Hennig, S.D.; Gnyp, M.L.; Chen, X.; Jia, L.; Bareth, G. Evaluating hyperspectral vegetation indices for estimating nitrogen concentration of winter wheat at different growth stages. *Precis. Agric.* **2010**, *11*, 335–357. [CrossRef]
6. Noda, H.M.; Muraoka, H.; Nasahara, K.N. Plant ecophysiological processes in spectral profiles: Perspective from a deciduous broadleaf forest. *J. Plant Res.* **2021**, *134*, 737–751. [CrossRef] [PubMed]
7. Abdel-Rahman, E.M.; Ahmed, F.B. The application of remote sensing techniques to sugarcane (*Saccharum* spp. hybrid) production: A review of the literature. *Int. J. Remote Sens.* **2008**, *29*, 3753–3767. [CrossRef]
8. Burnett, A.C.; Anderson, J.; Davidson, K.J.; Ely, K.S.; Lamour, J.; Li, Q.; Morrison, B.D.; Yang, D.; Rogers, A.; Serbin, S.P. A best-practice guide to predicting plant traits from leaf-level hyperspectral data using partial least squares regression. *J. Exp. Bot.* **2021**, *72*, 6175–6189. [CrossRef] [PubMed]
9. Sánchez, M.R.; Silva, C.A.A.C.; Demattê, J.A.M.; Mendonça, F.C.; Silva, M.A.; Romanelli, T.L.; Fiorio, P.R. Predictive modelling of chlorophyll in Mombaça grass leaves by hyperspectral reflectance data and machine learning. *Grass Forage Sci.* **2024**, *79*, 543–556. [CrossRef]
10. Durán, S.M.; Martin, R.E.; Díaz, S.; Maitner, B.S.; Malhi, Y.; Salinas, N.; Shenkin, A.; Silman, M.R.; Wieczynski, D.J.; Asner, G.P.; et al. Informing trait-based ecology by assessing remotely sensed functional diversity across a broad tropical temperature gradient. *Sci. Adv.* **2019**, *5*, eaaw8114. [CrossRef] [PubMed]
11. Zheng, Z.; Zeng, Y.; Schneider, F.D.; Zhao, Y.; Zhao, D.; Schmid, B.; Schaepman, M.E.; Morsdorf, F. Mapping functional diversity using individual tree-based morphological and physiological traits in a subtropical forest. *Remote Sens. Environ.* **2021**, *252*, 112170. [CrossRef]

12. Barros, P.P.S.; Fiorio, P.R.; Demattê, J.A.M.; Martins, J.A.; Montezano, Z.F.; Dias, F.L.F. Estimation of leaf nitrogen levels in sugarcane using hyperspectral models. *Ciência Rural* **2022**, *52*, e20200630. [CrossRef]
13. Inoue, Y.; Guérif, M.; Baret, F.; Skidmore, A.K.; Gitelson, A.; Schlerf, M.; Darvishzadeh, R.; Olioso, A. Simple and robust methods for remote sensing of canopy chlorophyll content: A comparative analysis of hyperspectral data for different types of vegetation. *Plant Cell Environ.* **2016**, *39*, 2609–2623. [CrossRef] [PubMed]
14. Miphokasap, P.; Wannasiri, W. Estimations of nitrogen concentration in sugarcane using hyperspectral imagery. *Sustainability* **2018**, *10*, 1266. [CrossRef]
15. Ollinger, S.V. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytol.* **2011**, *189*, 375–394. [CrossRef] [PubMed]
16. Silva, C.A.A.C.; Fiorio, P.R.; Rizzo, R.; Rossetto, R.; Vitti, A.C.; Dias, F.L.F.; Oliveira, K.A.; Neto, M.B. Spatio-temporal generalization of VIS-NIR-SWIR spectral models for nitrogen prediction in sugarcane leaves. *Remote Sens.* **2024**, *16*, 4250. [CrossRef]
17. Fiorio, P.R.; Silva, C.A.A.C.; Rizzo, R.; Demattê, J.A.M.; Luciano, A.C.S.; da Silva, M.A. Prediction of leaf nitrogen in sugarcane (*Saccharum* spp.) by VIS-NIR-SWIR spectroradiometry. *Heliyon* **2024**, *10*, e26819. [CrossRef] [PubMed]
18. Arroyo-Mora, J.P.; Kalacska, M.; Inamdar, D.; Soffer, R.; Lucanus, O.; Gorman, J.; Naprstek, T.; Schaaf, E.S.; Ifimov, G.; Elmer, K.; et al. Implementation of a UAV-hyperspectral pushbroom imager for ecological monitoring. *Drones* **2019**, *3*, 12. [CrossRef]
19. Li, J.; Wu, H.; Zhang, C.; Zhang, J.; Yang, Y. Hyperspectral characteristics and scale effects of leaf and canopy spectral reflectance. *Agriculture* **2021**, *11*, 1180. [CrossRef]
20. Som-Ard, J.; Suzuki, T.; Uto, K.; Yoneda, K.; Sparovek, G. Multi-sensor integration and variable standardization for crop monitoring: A review of advances and remaining challenges. *Remote Sens.* **2021**, *13*, 712. [CrossRef]
21. Alvares, C.A.; Stape, J.L.; Sentelhas, P.C.; Gonçalves, J.L.M.; Sparovek, G. Köppen's climate classification map for Brazil. *Meteorol. Z.* **2013**, *22*, 711–728. [CrossRef]
22. Analytical Spectral Devices. *FieldSpec HandHeld 2 User Manual; Analytical Spectral Devices: Boulder, CO, USA*; Analytical Spectral Devices: Boulder, CO, USA, 2010; pp. 93p.
23. Analytical Spectral Devices. *FieldSpec 3 User Manual*; Analytical Spectral Devices: Boulder, CO, USA, 2010.
24. Lu, B.; Wang, X.; Liu, N.; He, K.; Wu, K.; Li, H.; Tang, X. Feasibility of NIR spectroscopy detection of moisture content in coco-peat substrate based on the optimization characteristic variables. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2020**, *239*, 118455. [CrossRef]
25. Savitzky, A.; Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [CrossRef]
26. Shen, L.; Gao, M.; Yan, J.; Li, Z.-L.; Leng, P.; Yang, Q.; Duan, S.-B. Hyperspectral estimation of soil organic matter content using different spectral preprocessing techniques and PLSR method. *Remote Sens.* **2020**, *12*, 12026. [CrossRef]
27. Viscarra Rossel, R.A. ParLeS: Software for chemometric analysis of spectroscopic data. *Chemom. Intell. Lab. Syst.* **2008**, *90*, 72–83. [CrossRef]
28. Clark, R.N.; Swayze, G.A.; Livo, K.E.; Kokaly, R.F.; Sutley, S.J.; Dalton, J.B.; McDougal, R.R.; Gent, C.A. Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. *J. Geophys. Res. Planets* **2003**, *108*, 5131. [CrossRef]
29. Kokaly, R.F. Plant phenolics and absorption features in vegetation reflectance spectra near 1.66 μm. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *43*, 55–83. [CrossRef]
30. Lian, S.; Guan, L.; Peng, Z.; Zeng, G.; Li, M.; Xu, Y. Retrieval of leaf chlorophyll content in Gannan navel orange based on fusing hyperspectral vegetation indices using machine-learning algorithms. *Ciência Rural* **2023**, *53*, e20210630. [CrossRef]
31. Lin, H.; Li, Z.; Li, J.; Zhang, F.; Min, M.; Menzel, W.P. Estimate of daytime single-layer cloud base height from Advanced Baseline Imager measurements. *Remote Sens. Environ.* **2022**, *274*, 112970. [CrossRef]
32. Fu, T.; Tang, X.; Cai, Z.; Zuo, Y.; Tang, Y.; Zhao, X. Correlation research of phase angle variation and coating performance by means of Pearson's correlation coefficient. *Prog. Org. Coat.* **2020**, *139*, 105459. [CrossRef]
33. dos Santos, G.L.A.A.; Reis, A.S.; Besen, T.O.; Furlanetto, R.H.; Rodrigues, M.; Crusiol, L.O.T.; de Oliveira, K.M.; Falcioni, R.; de Oliveira, R.B.; Batista, M.A.; et al. Spectral method for macro- and micronutrient prediction in soybean leaves using interval partial least squares regression. *Eur. J. Agron.* **2023**, *143*, 126717. [CrossRef]
34. Flynn, K.C.; Baath, G.; Lee, T.O.; Gowda, P.; Northup, B. Hyperspectral reflectance and machine learning to monitor legume biomass and nitrogen accumulation. *Agronomy* **2023**, *211*, 107991. [CrossRef]
35. Tavares, M.S.; Silva, C.A.A.C.; Regazzo, J.R.; Sardinha, E.J.S.; Silva, T.L.; Fiorio, P.R.; Baesso, M.M. Performance of machine learning models in predicting common bean (*Phaseolus vulgaris* L.) crop nitrogen using NIR spectroscopy. *Agronomy* **2024**, *14*, 1634. [CrossRef]
36. Rodrigues, M.; Nanni, M.R.; Cezar, E.; Dos Santos, G.L.A.A.; Reis, A.S.; Oliveira, K.M.; Oliveira, R.B. Vis-NIR spectroscopy: From leaf dry mass production estimate to the prediction of macro- and micronutrients in soybean crops. *J. Appl. Remote Sens.* **2020**, *14*, 044505. [CrossRef]

37. 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. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2018**, *189*, 183–189. [CrossRef]
38. R Core Team. R: A Language and Environment for Statistical Computing. 2024. Available online: https://www.r-project.org/ (accessed on 5 November 2025).
39. Mevik, B.-H.; Wehrens, R. The pls Package: Principal Component and Partial Least Squares Regression in R. *J. Stat. Softw.* **2007**, *18*, 1–23. [CrossRef]
40. Breiman, L. Random forests. *Mach. Learn.* **2012**, *45*, 5–32. [CrossRef]
41. Liu, Y.; Wang, Y.; Zhang, J. New machine learning algorithm: Random forest. In *Information Computing and Applications, Proceedings of ICICA 2012*; Liu, B., Ma, M., Chang, J., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2021; Volume 7473, pp. 246–252. [CrossRef]
42. Tang, K.; Qin, M.; Han, B.; Shao, D.; Xu, Z.; Sun, H.; Wu, Y. Identifying the influencing factors of soil nitrous acid emissions using random forest model. *Atmos. Environ.* **2024**, *339*, 120875. [CrossRef]
43. Liu, M.; Liu, X.; Liu, D.; Ding, C.; Jiang, J. Multivariable integration method for estimating sea surface salinity in coastal waters from in situ data and remotely sensed data using random forest algorithm. *Comput. Geosci.* **2015**, *75*, 44–56. [CrossRef]
44. Friedman, J.H.; Meulman, J.J. Multiple additive regression trees with application in epidemiology. *Stat. Med.* **2003**, *22*, 1365–1381. [CrossRef]
45. Willmott, C.J.; Robeson, S.M.; Matsuura, K. A refined index of model performance. *Will Int. J. Climatol.* **2012**, *32*, 2088–2094. [CrossRef]
46. Raij, B.; Cantarella, H.; Quaggio, J.A.; Furlani, A.M.C. *Recomendações de Adubação e Calagem para o Estado de São Paulo*, 2nd ed.; Instituto Agronômico: Campinas, Brazil, 1997.
47. Jacquemoud, S.; Ustin, S. *Leaf Optical Properties*; Cambridge University Press: Cambridge, UK, 2019. [CrossRef]
48. Kothari, S.; Serbin, S.P.; Ely, K.S.; Wu, J.; Singh, A.; Rogers, A.; Townsend, P.A. Predicting leaf traits across functional groupsusing reflectance spectroscopy. *New Phytol.* **2023**, *24*, 2350–2365. [CrossRef]
49. Moriya, É.A.S.; Imai, N.N.; Tommaselli, A.M.G. A study on the effects of viewing angle variation in sugarcane radiometric measures. *Bol. Ciênc. Geod.* **2018**, *24*, 85–97. [CrossRef]
50. Wang, Y.; Oliver, T.J.; Croce, R.; Long, S.P. Addition of longer wavelength absorbing chlorophylls into crops could increase their photosynthetic productivity by 26%. *Nat. Commun.* **2025**, *16*, 7933. [CrossRef] [PubMed]
51. Zheng, R.; Chen, Z.; Guan, Z.; Zhao, C.; Cui, H.; Shang, H. Variable importance for projection (VIP) scores for analyzing the contribution of risk factors in severe adverse events to Xiyanping injection. *Chin. Med.* **2023**, *18*, 15. [CrossRef] [PubMed]
52. Almeida, L.M.D.; Ferraz, R.P.; Costa, L.L. Estimation of leaf nitrogen content using hyperspectral reflectance and machine learning algorithms. *ISPRS J. Photogramm. Remote Sens.* **2014**, *97*, 185–193. [CrossRef]
53. Taiz, L.; Zeiger, E.; Møller, I.M.; Murphy, A. *Plant Physiology and Development*, 6th ed.; Sinauer Associates: Sunderland, MA, USA, 2015.
54. Fan, L.; Zhao, J.; Xu, X.; Liang, D.; Yang, G.; Feng, H.; Yang, H.; Wang, Y.; Chen, G.; Wei, P. Hyperspectral-based estimation of leaf nitrogen content in corn using optimal selection of multiple spectral variables. *Sensors* **2019**, *19*, 2898. [CrossRef] [PubMed]
55. Alexandre, M.L.d.S.; Lima, I.d.L.e.; Nilsson, M.S.; Rizzo, R.; Silva, C.A.A.C.; Fiorio, P.R. Sugarcane (*Saccharum officinarum*) productivity estimation using multispectral sensors in RPAs, biometric variables, and vegetation indices. *Agronomy* **2025**, *15*, 2149. [CrossRef]
56. Palla, A. Review of the potential and limitations of visible, near-, and shortwave-infrared reflectance spectroscopy for quantitative soil assessment. *J. Geophys. Res. Atmos.* **2023**, *128*, e2022JD037666. [CrossRef]
57. Sun, H.; Feng, M.; Yang, W.; Bi, R.; Sun, J.; Zhao, C.; Xiao, L.; Wang, C.; Kubar, M.S. Monitoring Leaf Nitrogen Accumulation With Optimized Spectral Index in Winter Wheat Under Different Irrigation Regimes. *Front. Plant Sci.* **2022**, *13*, 913240. [CrossRef] [PubMed]
58. Martins, J.A.; Fiorio, P.R.; Barros, P.P.S.; Demattê, J.A.M.; Molin, J.P.; Cantarella, H.; Neale, C.M.U. Potential use of hyperspectral data to monitor sugarcane nitrogen status. *Acta Sci. Agron.* **2021**, *43*, e47632. [CrossRef]
59. Zhu, L.; Chen, Z.; Wang, J.; Ding, J.; Yu, Y.; Li, J.; Xiao, N.; Jiang, L.; Zheng, Y.; Rimmington, G.M. Monitoring plant response to phenanthrene using the red edge of canopy hyperspectral reflectance. *Mar. Pollut. Bull.* **2014**, *86*, 332–341. [CrossRef] [PubMed]
60. Silva, C.A.A.C.; Fiorio, P.R.; Rizzo, R.; Rossetto, R.; Vitti, A.C.; Dias, F.L.F.; Oliveira, K.A.; Bárbara Neto, M. Detection of nutritional stress in sugarcane by VIS-NIR-SWIR reflectance spectroscopy. *Ciência Rural* **2023**, *53*, e20220543. [CrossRef]
61. Reyes-Trujillo, A.; Daza-Torres, M.C.; Galindez-Jamioy, C.A.; Rosero-García, E.E.; Muñoz-Arboleda, F.; Solarte-Rodriguez, E. Estimating canopy nitrogen concentration of sugarcane crop using in situ spectroscopy. *Heliyon* **2021**, *7*, e06566. [CrossRef] [PubMed]

62. Petibon, F.; Czyż, E.A.; Ghielmetti, G.; Hueni, A.; Kneubühler, M.; Schaepman, M.E.; Schuman, M.C. Uncertainties in measurements of leaf optical properties are small compared to the biological variation within and between individuals of European beech. *Remote Sens. Environ.* **2021**, *264*, 112601. [CrossRef]
63. Rasooli Sharabiani, V.; Soltani Nazarloo, A.; Taghinezhad, E.; Veza, I.; Szumny, A.; Figiel, A. Prediction of winter wheat leaf chlorophyll content based on VIS/NIR spectroscopy using ANN and PLSR. *Food Sci. Nutr.* **2023**, *11*, 2166–2175. [CrossRef] [PubMed]
64. Sims, D.A.; Gamon, J.A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* **2002**, *81*, 337–354. [CrossRef]