

Article

Spatio-Temporal Generalization of VIS-NIR-SWIR Spectral Models for Nitrogen Prediction in Sugarcane Leaves

Carlos Augusto Alves Cardoso Silva ^{1,*}, Rodnei Rizzo ², Marcelo Andrade da Silva ² , Matheus Luís Caron ¹ and Peterson Ricardo Fiorio ¹ 

¹ Department of Biosystems Engineering, “Luiz de Queiroz” College of Agriculture, University of São Paulo, Piracicaba 13418900, SP, Brazil

² Department of Exact Science, “Luiz de Queiroz” College of Agriculture, University of São Paulo, Piracicaba 13418900, SP, Brazil

* Correspondence: carlosesalq@usp.br

Abstract: Nitrogen fertilization is a challenging task that usually requires intensive use of resources, such as fertilizers, management and water. This study explored the potential of VIS-NIR-SWIR remote sensing for quantifying leaf nitrogen content (LNC) in sugarcane from different regions and vegetative stages. Conducted in three regions of São Paulo, Brazil (Jaú, Piracicaba and Santa Maria), the research involved three experiments, one per location. The spectral data were obtained at 140, 170, 200, 230 and 260 days after cutting (DAC). From the hyperspectral data, clustering analysis was performed to identify the patterns between the spectral bands for each region where the spectral readings were made, using the Partitioning Around Medoids (PAM) algorithm. Then, the LNC values were used to generate spectral models using Partial Least Squares Regression (PLSR). Subsequently, the generalization of the models was tested with the leave-one-date-out cross-validation (LOOCV) technique. The results showed that although the variation in leaf N was small, the sensor demonstrated the ability to detect these variations. Furthermore, it was possible to determine the influence of N concentrations on the leaf spectra and how this impacted cluster formation. It was observed that the greater the average variation in N content in each cluster, the better defined and denser the groups formed were. The best time to quantify N concentrations was at 140 DAC ($R^2 = 0.90$ and $RMSE = 0.74 \text{ g kg}^{-1}$). From LOOCV, the areas with sandier soil texture presented a lower model performance compared to areas with clayey soil, with $R^2 < 0.54$. The spatial generalization of the models recorded the best performance at 140 DAC ($R^2 = 0.69$, $RMSE = 1.18 \text{ g kg}^{-1}$ and $dr = 0.61$), decreasing in accuracy at the crop-maturation stage (260 DAC), R^2 of 0.05, $RMSE$ of 1.73 g kg^{-1} and dr of 0.38. Although the technique needs further studies to be improved, our results demonstrated potential, which tends to provide support and benefits for the quantification of nutrients in sugarcane in the long term.



Citation: Silva, C.A.A.C.; Rizzo, R.; da Silva, M.A.; Caron, M.L.; Fiorio, P.R. Spatio-Temporal Generalization of VIS-NIR-SWIR Spectral Models for Nitrogen Prediction in Sugarcane Leaves. *Remote Sens.* **2024**, *16*, 4250. <https://doi.org/10.3390/rs16224250>

Academic Editors: Yuhong He and Dongmei Chen

Received: 30 September 2024

Revised: 22 October 2024

Accepted: 30 October 2024

Published: 14 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: digital agriculture; PLS regression; nitrogen sensing; management

1. Introduction

Sugarcane (*Saccharum officinarum* L.) is one of the most harvested crops in the world, largely due to its nutritional value and versatility as a raw material for the food industry [1]. In this context, Brazil holds the position of the world’s leading producer, making the sugar and ethanol sector one of the most important segments of Brazilian agribusiness [2]. Due to its economic importance and high production, the adoption of appropriate agronomic practices is essential, especially for nutritional management, one of the primary factors for successful productivity in sugarcane crops [3,4]. Nevertheless, this is a challenging task that usually requires the intensive use of resources, such as fertilizers, management and water. In this sense, studies have been carried out with the purpose of increasing sugarcane productivity [5,6]; it is essential to adopt appropriate agronomic practices, with emphasis

on nutritional management (e.g., the quantification and variable-rate application of N), one of the key factors for successful productivity [3–5].

Nitrogen, phosphorus and potassium-based fertilization is one of the most widely used agricultural practices in cropping systems [7,8]. Nitrogen (N), for example, plays an important role in crop development, influencing production quantity and quality (sugar yield, vegetative development, fiber quality and leaf production) [9]. In addition to N promoting positive effects on growth, agronomic parameters and sugar content [9,10], it is also one of the primary regulators of several leaf physiological processes, such as photosynthesis [11,12]. Its deficiency reduces the production of chlorophyll, amino acids and energy, which has a direct effect on sugarcane growth and yield [13].

The main technique for acquiring information for the correct application of nutrients is by soil and plant tissue analyses, strategies that sometimes become impractical on a large scale, due to their slow data-acquisition process, high costs, evasiveness and the demand for a large number of samples for an adequate representation of their variability [14]. In this sense, many technological approaches are emerging in favor of crop nutritional monitoring, such as hyperspectral remote sensing [15,16], an alternative with great potential that has rapid data acquisition [17]. This technique addresses different ways to monitor nutrients in plants, including the analysis of canopy and leaves that have been recently removed from the plant for spectral scanning [18,19]. Currently, it is seen as one of the alternatives with great potential [15,20], having as its main advantages the rapid acquisition of data and its non-destructive nature [17], allowing work with numerous samples, which would facilitate the quantification of nutrients on a large scale and spatial variability.

Although it is still little used in the investigation of nutritional aspects, especially in Brazil, this technology has the potential to contribute to obtaining information for the precise application of nutrients in plantations [18,19,21,22]. Nevertheless, few studies in the world test the possibility of transferring their spectral models in independent areas. Pullanagari et al. [23], for example, made N predictions in pastures and tested the models in other areas with edaphoclimatic conditions different from those of the training. The results were inconclusive for the PLSR models (R^2 ranging from 0.45 to 0.48, RMSE between 17% and 19%), but optimistic when testing the one-dimensional convolutional neural network (1D-CNN), with R^2 values ranging from 0.62 to 0.75 and RMSE between 13 and 14.4%. Wang et al. [24] applied their models to the set of data collected in independent areas, having moderate predictive accuracies for N, carbon, leaf mass and equivalent water thickness, with R^2 from 0.48 to 0.55 and NRMSE from 11.6 to 16.8%. This topic has been discussed as one of the main challenges related to nutrient forecasting, which is the transfer of local calibration models to independent areas or varieties, as a measure to assess the accuracy and generalization of predictive models [24,25]. Nevertheless, nitrogen dynamics in sugarcane crops and the spectral response present intrinsic complexities that still need to be studied and that are little addressed in the literature. For example, it is necessary to investigate whether the environment is a determining factor in nitrogen uptake by sugarcane and whether it interferes with the spectral response, and to evaluate the efficiency of models based exclusively on spectral data at different phenological stages and edaphoclimatic conditions different from those used in calibration. Furthermore, it is crucial to investigate whether the performance of the prediction models is equally efficient throughout the entire crop cycle.

To date, few studies have been found that have used sugarcane spectral models and validated them between crop-development stages [15] or between harvests [16] under the Brazilian environmental conditions, and none have been identified comparing regions with different edaphoclimatic conditions. Therefore, this study aims to perform a detailed analysis of the potential of VIS-NIR-SWIR remote sensing for the prediction of nitrogen contents at the leaf level in sugarcane plantations in three areas with different soil and climate conditions throughout the whole vegetative stage of the crop. In other words, we evaluated the performance of the models in phenological stages, climate conditions and production environment different from those considered in the calibration. With

this, the authors hope to understand the influence of external factors on the prediction of N by spectroradiometry, and discuss the methodological limitations that still need to be overcome.

2. Materials and Methods

The method used in this work has nine main phases, as follows: (i) identification of the study areas; (ii) collection of leaf samples; (iii) chemical analysis of LNC; (iv) obtaining the spectral signature; (v) data preprocessing; (vi) identification of patterns in the spectra by clustering; (vii) prediction of LNC by machine learning; (viii) generalization of the models; and (ix) validation. The general workflow can be seen in Figure 1. In the following sections, each of these steps will be described in detail.

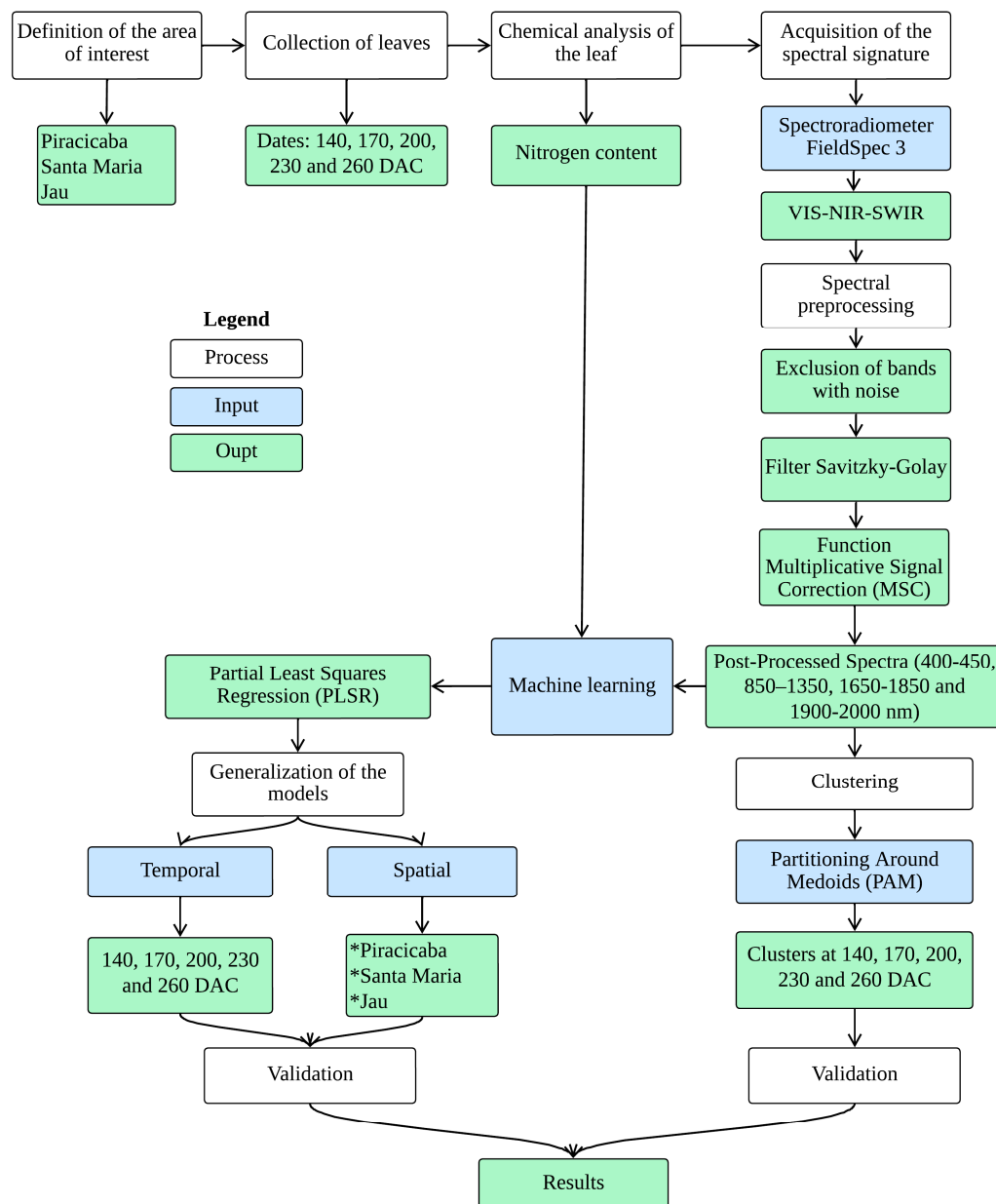


Figure 1. Workflow of the procedure to quantify the N concentrations in sugarcane leaves by the leaf spectral signatures.

2.1. Description of the Experiment

The experiments were installed in the municipalities of Piracicaba, Santa Maria and Jaú, all located in the State of São Paulo, Brazil (Figure 2). According to the Köppen climate classification, the region has a humid subtropical climate (Cwa). The average annual rainfall variation between the municipalities where the experiments were implemented is small, ranging from 1280 mm in the region of Piracicaba [26] to 1344 mm nearby Jaú [27]. The soils in the experimental areas of the municipalities of Piracicaba, Jaú and Santa Maria were classified as Red-Yellow Alfisol (Clayey), Red Oxisol (Sandy Loam) and Quartzarenic Neosol (Sandy Loam), respectively (Figure 2).

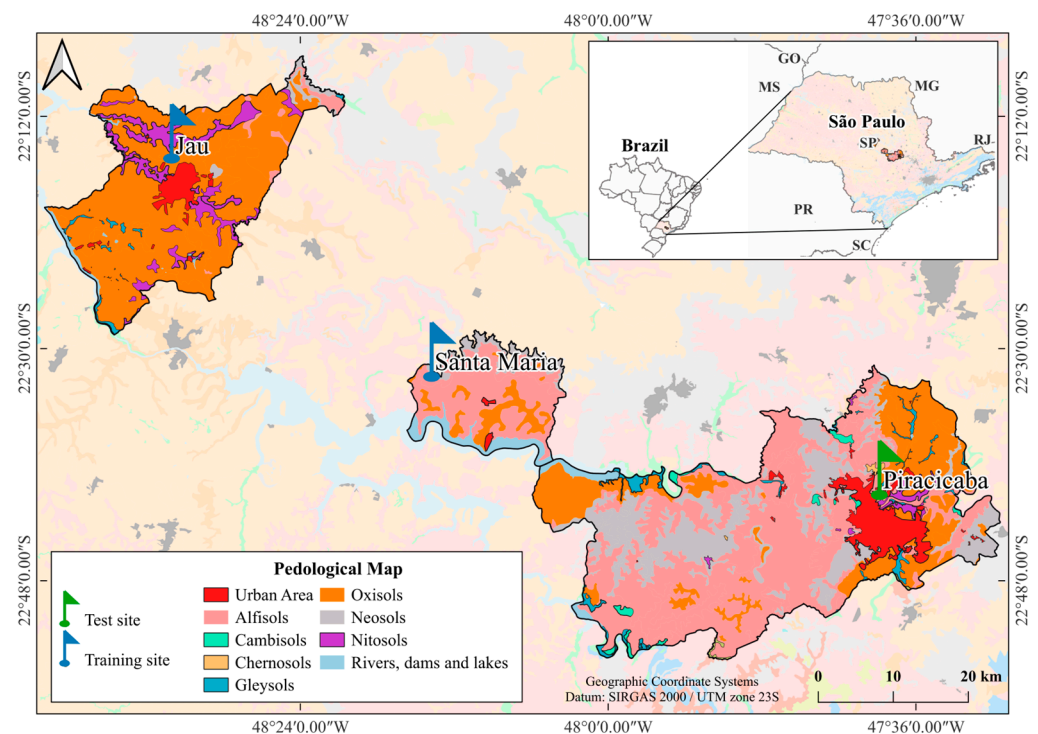


Figure 2. Map with the location of the collection sites, focusing on the soil classes for the municipalities of Piracicaba, Jaú and Santa Maria, characterized by soils of the types Red-Yellow Alfisol (Clayey), Red Oxisol (Sandy Loam) and Quartzarenic Neosol (Sandy Loam), respectively. The training and testing sites of the predictive models are shown. The map was prepared by the authors based on data from Rossi [28].

The experiment was set up in the three study sites using a completely randomized block design with four nitrogen doses (0, 50, 100 and 150 kg ha⁻¹) and six blocks. The variety SP 81 3250 was grown in the three experimental areas, which allowed the comparison of the effect of the environment on the same genetic material. The plots consisted of five rows of sugarcane, spaced 1.5 m apart and 10 m in length; the three central rows were considered for the evaluation area, discarding one meter from the ends in order to avoid border effects.

The initial and annual soil corrections were performed according to the recommendations for the crop, diagnosed by routine soil fertility analyses. Nitrogen doses were applied using ammonium nitrate, distributed over the sugarcane straw in a single dose, at the beginning of each cycle. All other phytosanitary treatments followed the standards of the regional production system adopted by sugarcane producers.

2.2. Collection of Leaf Material and Spectral Measurements in the Laboratory

Field visits were carried out on dates 140, 170, 200, 230 and 260 DAC, totaling five collections. Leaf material was collected for subsequent laboratory analysis by obtaining

20 leaves per plot. The evaluations were performed on the middle third of the first fully expanded leaf from the crop apex [15]. After removal, the leaves were placed in plastic bags and transported in coolers with ice to the geoprocessing laboratory for the spectral readings. This technique was adopted to preserve the turgidity and spectral properties of the leaves [15,29,30].

In the laboratory, spectral measurements were performed using a FieldSpec 3 spectroradiometer. The sensor operates within the 350–2500 nm range, with wavelengths divided into spectral bands: visible/Vis (350–680 nm), near-infrared/NIR (750–1300 nm) and short-wave infrared/SWIR (1300–2500 nm). The spectroradiometer was connected to the Leaf Clip[®] probe via an optical fiber, which is capable of maintaining the same light intensity and orthogonal incidence for all readings, functioning as a fully controlled method. Calibration of the device was performed after reading 20 leaves, using the Lambertian surface embedded in the Leaf Clip[®] as a reference [15].

2.3. Spectra Preprocessing

The spectral data were initially subjected to preprocessing, aiming to correct inconsistencies in the readings caused by external factors such as noise, environmental variations or even light scattering at the time of the readings [15]. Preprocessing occurred in three steps, in the following sequence: (i) removal of the spectral bands that presented large concentrations of noise, namely 400–450 nm, 850–1350 nm, 1650–1850 nm and 1900–2000 nm; (ii) spectra rectification by the function Multiplicative Signal Correction (MSC), which is a mathematical technique (Equations (1)–(3)) that acts mainly reducing the influence of surface scattering of particles in spectral data acquisition [31]; (iii) application of the Savitzky–Golay (SG) filter [32], with 3-point smoothing and second-order polynomial. The following is the expression used to calculate the spectra correction by the MSC method.

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

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

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

where $\text{Spectra}_{\text{avg}}$ is the average of all leaf spectra; n is the total number of spectral data; Spectra_i refers to each individual leaf spectrum; k_i and b_i are correction coefficients obtained by linear regression based on $\text{Spectra}_{\text{avg}}$, and $\text{Spectra}_{\text{MSC},i}$ is the MSC-corrected spectrum.

2.4. Unsupervised Clustering Analysis Partitioning Around Medoids (PAM)

In this study, we evaluated whether the spectral behavior of the leaf samples is influenced by edaphoclimatic conditions in each region. To demonstrate that such environmental conditions are an important factor and must be considered during model calibration and, consequently, in the spectral model transfer processes, a clustering technique was performed based on the spectral behavior of the samples. The aim was to evaluate whether the resulting groups perfectly distinguished each of the experiments and, consequently, the plants grown in different geographic regions. Therefore, the cluster analysis was performed by the application of the unsupervised Partitioning Around Medoids (PAM) algorithm [33], which is based on the K-medoids method, in which data are grouped into k clusters. In this method, the medoid is the element within a cluster whose average distance between it and the rest of the elements within the cluster is the smallest possible. The use of medoids makes the method less susceptible to noise and outliers, when compared to better-known methods, such as k -means, for example, which makes the clustering more accurate [34]. It is widely used in the literature for spectral analyses [34–36]. At the end, the centroid was calculated for each cluster, allowing the dispersion of the spectral data to be assessed.

To evaluate the clustering results, we applied an internal validation that measures the homogeneity of the clusters. Therefore, Silhouette analysis [37] was employed, which

measures how well a point fits into a cluster. The Silhouette coefficient, when close to +1, indicates that the points are very far from the points in the other cluster (meaning the datum is in the correct cluster), and when close to 0, it indicates the points are very close to or even intersecting another cluster [38]. This validation technique assesses the effects of changes in spectral resolution and cluster quality in terms of internal validation and stability [39].

2.5. Machine Learning Prediction Model

Partial Least Squares Regression (PLSR)

The quantification of N concentrations from the leaf spectral signature was performed using the PLSR technique, employing the NIPALS algorithm. PLSR is one of the most widely used techniques in nutrient prediction models that utilize hyperspectral data [15,19,40–42]. This is because PLSR handles correlated independent variables well, in this case, the spectral bands (450–2000 nm) with few observations, reducing them to a set of components and avoiding multicollinearity [43]. During the training stage, PLSR uses the information from the predictor variables (spectra) and predicted variables (N concentrations) to generate new variables, called latent variables (factors). When fitting a model using PLSR, the goal is to find the fewest PLS factors necessary to explain the dependent variables. Too many factors may introduce noise or irrelevant parts into the calibration stage, leading to an unstable model, while too few factors may result in a model with poor performance in both calibration and prediction phases [44]. K-Fold ($k = 10$) cross-validation was used to select the optimal number of factors that minimize the Root Mean PRESS statistic. Additionally, with $k = 10$, we achieve a good balance, allowing the model to be trained on a substantial portion of the data while still reserving a sufficient amount for testing.

2.6. Spatio-Temporal Generalization of the Models

The possibility of determining the ideal period for quantifying LNC in sugarcane using spectral models was tested. Therefore, we used the dates that were common to the three study sites (Piracicaba, Jaú and Santa Maria), such as 140, 170, 200, 230 and 260 DAC, and models for each date were generated. This comparative analysis of the selected dates aimed to identify the moment in which the correlation between spectral data and N concentration is strongest, thus optimizing the process of sugarcane monitoring and management.

The transfer of learning of the models was tested in two situations, both using only spectral data. In the first situation, models were generated to predict LNC completely independently for the same variety, but in regions with different edaphoclimatic conditions. The model was trained with the reflectance data from Jaú and Santa Maria, and tested using information from Piracicaba. The models were tested on the dates 140, 170, 200, 230 and 260 DAC. In other words, the model was tested under conditions completely independent on those used for training, verifying whether the generalization of the models is applicable to sites with different edaphoclimatic conditions.

The second option for the transfer of learning was identified here as Leave-One-Date-Out Cross Validation (LOOCV), in which the possibility of testing the model in vegetative stages that did not participate in the calibration phase was evaluated. Therefore, the model was trained with data from four vegetative stages and tested in with the data that were left out of the training. This process was repeated until the entire cycle was completed between all collections (140, 170, 200, 230 and 260 DAC). Unlike the first validation, LOOCV tests the capacity of generalization of the models between temporal (phenological) stages of crop development.

2.7. Validation of the Models

The PLSR models were calibrated according to the optimal number of factors (FT), obtained through k-fold cross-validation. In the prediction process, the values of the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE) and systematic error (BIAS) were considered, as described in Equations (4)–(7),

respectively. To test agreement, the Willmott index (dr) [45] was used, which reflects the degree to which the measured data are accurately estimated by the predicted data (Equation (8)).

$$R^2 = \frac{[\sum (\gamma_p - \bar{\gamma}_p) \cdot (\gamma_o - \bar{\gamma}_o)]^2}{\sum (\gamma_p - \bar{\gamma}_p)^2 \cdot \sum (\gamma_o - \bar{\gamma}_o)^2} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (5)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |X_i - \hat{X}| \quad (6)$$

$$BIAS = \frac{\sum_{i=1}^{N_s} (\hat{\gamma}_{m,s,i})}{N_s} - \frac{\sum_{i=1}^{N_s} (\gamma_{s,i})}{N_s} \quad (7)$$

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

3. Results

3.1. Leaf N Content for Each Location and Characterization of the Leaf Spectrum

The spectral response of the leaves was categorized for each study location (Jaú, Santa Maria and Piracicaba) into five LNC classes, determined based on a distribution analysis using histograms. The classes were generated from the data collected on dates 140, 170, 200, 230 and 260 DAC, and organized in decreasing order according to the LNC (Figure 3). The bands used in this phase were 450–750 nm, which are between the visible range (450–680 nm) and the red-edge range (680–750 nm). These bands were chosen because they are directly related to the N contents and for scale reasons, to facilitate the visualization of the spectral behavior as a function of the N concentrations.

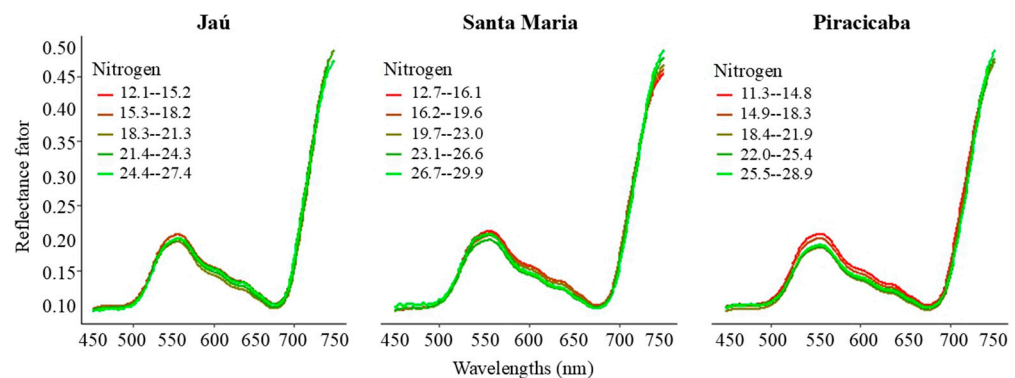


Figure 3. Spectral curves of the sugarcane leaves at the wavelengths of 450–750 nm on dates 140, 170, 200, 230 and 260 DAC, set from the lowest to the highest LNC for the regions of Jaú, Piracicaba and Santa Maria.

Although the variation in N levels was small, the sensor demonstrated the ability to detect these variations in the spectral curves (Figure 3), demonstrating its sensitivity and applicability in identifying small changes in the leaf spectrum. Notably, the leaves with the lowest N concentrations exhibited the highest reflectance factor in the visible range, while those with the highest N levels showed the lowest reflectance, which was more evident in the green band (550 nm). On the other hand, the blue band (450–500 nm) remained stable, regardless of whether the LNC was high or low. Disregarding the collection date or production environment, it was observed that for the plants with higher amplitudes in the spectral response in the visible range (450–680 nm), the red-edge region (680–750 nm)

tended to direct the curves towards the blue bands (Piracicaba). This indicates that the red-edge region is sensitive to variations in LNC in sugarcane crops in situations of low leaf N concentrations. This finding contributes to future studies on the identification of nutritional stress in sugarcane crops.

3.2. Clustering Analysis Using the PAM Technique

Clustering analysis using the PAM method was performed using collections at 140, 170, 200, 230 and 260 DAC, testing the possibility of generating groups from the three different production environments using only leaf spectra (450–1900 nm). Cluster validation was performed using the Silhouette coefficient, which indicated a moderate definition of the clusters, with mean values of 0.40, 0.36, 0.37, 0.45 and 0.38 for the clusters on dates 140, 170, 200, 230 and 260 DAC, respectively. These values suggest that, although the clusters are fairly defined, the points are relatively close to each other or intersect with other clusters, as can be seen in Figure 4.

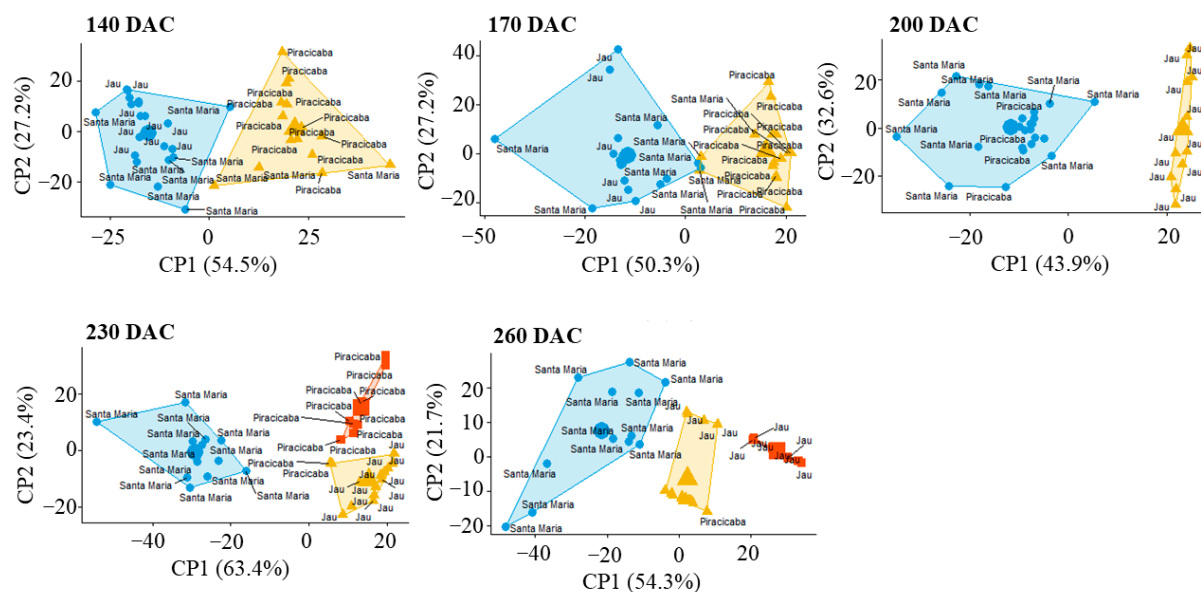


Figure 4. Clustering analysis by the PAM method from the spectral data from dates 140, 170, 200, 260 and 260 DAC.

The formation of clusters followed the same pattern in the first two collections (140 and 170 DAC), with data from Jaú and Santa Maria in the same group, whereas Piracicaba remained in a single group (Figure 4, at 140 and 170 DAC). In the collection at 200 DAC, the groups were inverted, with Piracicaba and Santa Maria in the same cluster. Still, the data from Piracicaba did not show large dispersions, which indicates little variation in the leaf spectral response. Nevertheless, the variation in the spectral response for Santa Maria was greater, confusing the PAM algorithm with the spectral signature of Piracicaba (Figure 4, at 200 DAC). In the last two collections (230 and 260 DAC), concerning the initial phase of maturation, the spectra started to show characteristics that were more distinct by region and more similar within each of them, forming three distinct clusters. At this stage, the dispersions were smaller between the clusters, to the point where few spectral signatures were confused between the groups. Despite the variety and collection dates being the same, the crop may develop at different rates, according to the edaphoclimatic conditions of the region. This justifies better defined clusters at the end of the crop cycle.

After clustering the leaf spectral signatures for each study site (Jaú, Santa Maria and Piracicaba, identified as JAU, STM and PIRA, as shown in Figure 5), boxplots were generated with the N contents based on the clusters. Thus, the distribution of N concentrations was evaluated in each group (Figure 5A). Then, the average reflectance for each cluster was

calculated according to the dates 140, 170, 200, 230 and 260 DAC, allowing the observation of the spectral behavior in the visible range (450–680 nm) of each group (Figure 5B). Therefore, it was possible to determine the influence of the N concentrations on the leaf spectra and how this impacted cluster formation. Overall, the average N contents for the clusters in the collections of 140, 170 and 200 DAC were similar. Additionally, it was observed that the greater the average variation in the LNC of each cluster, the better defined and denser the groups formed were, as in the case of the collections at 230 and 260 DAC (Figures 4 and 5A). Furthermore, as the N variations increased between the clusters, the average spectral curves became more distant, indicating that the N variation influenced the reflectance of the leaves and, consequently, the grouping (Figure 5B).

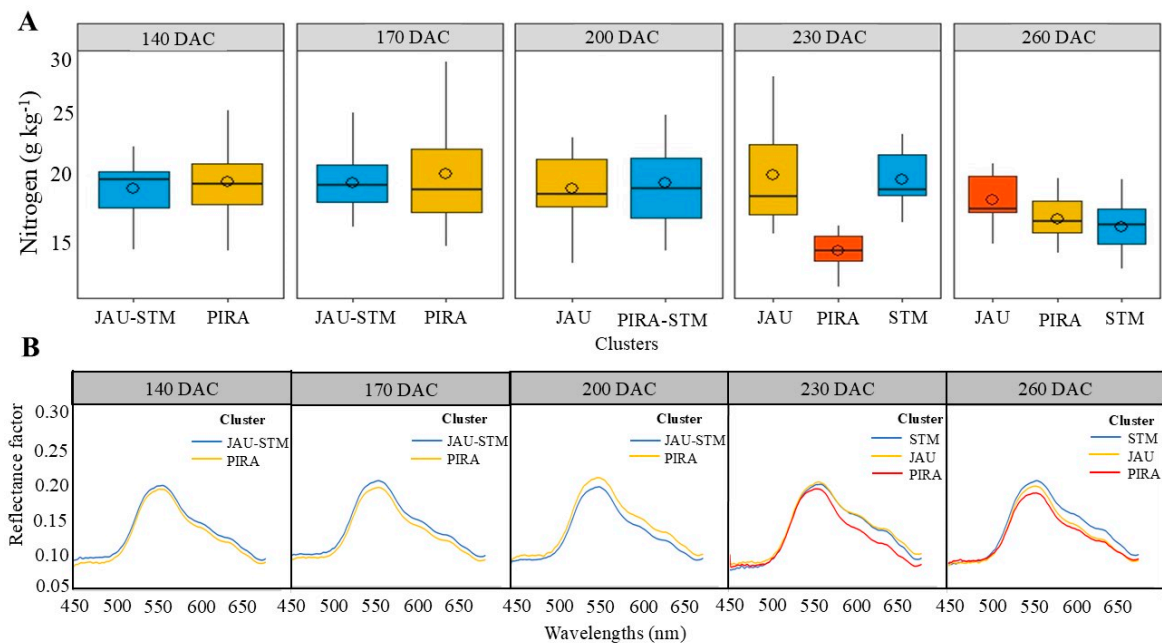


Figure 5. N concentrations based on the clusters formed by the leaf spectra (A), and the mean of the spectral curves for each cluster (B), for the dates 140, 170, 200, 230 and 260 DAC. The acronym JAU-STM refers to the clusters formed from the data from Jaú and Santa Maria; PIRA-STM is Piracicaba and Santa Maria; PIRA is Piracicaba, JAU is Jaú and STM is Santa Maria, according to the clusters shown in Figure 4.

During the phase of exponential development of the crop, with the exception of the collection at 230 DAC, the clusters with the highest N averages presented the lowest reflectance factors in the visible range (450–680 nm), demonstrating the greatest absorption of electromagnetic radiation and, theoretically, the best vegetative development. In this case, the clusters formed from the PIRA spectra recorded the highest N averages and the lowest reflectance factors. On the other hand, at the end of the vegetative cycle of the crop (260 DAC), the average N values for the PIRA clusters decreased, falling below the values for Jaú (JAU), resulting in an increase in the reflectance factor. It is worth noting that the reductions observed in the average N concentrations for Piracicaba (PIRA) were also identified in the spectral curves, maintaining the principle that leaves with higher N contents absorb greater radiation in the visible range (Figure 5A). Therefore, our results denoted a trend that, with higher LNC, there is lower reflectance, regardless of the production environment. In addition, the reduction in N concentrations at the end of the cycle in the sugarcane leaves in the areas of Piracicaba and Santa Maria suggests the crop in these locations started to mature first.

3.3. Prediction of N by Vis-NIR-SWIR Spectra

The prediction of N, in which the models were calibrated and tested at the same location, showed acceptable levels ($R^2 > 0.7$ and $RMSE < 1.59 \text{ g kg}^{-1}$). Nevertheless, the one calibrated with data from all collections (global), which presented R^2 of 0.53, $RMSE = 2.15 \text{ g kg}^{-1}$, MAE of 1.64 g kg^{-1} and $dr = 0.62$, was identified as a model with moderate predictions. Although the R^2 was not considered high, the error in relation to the average N levels considering all locations and during the five collection dates was only 11.7%. Furthermore, it should be considered that in the global validation, the data belonged to three distinct regions and five collection dates per region (140, 170, 200, 230 and 260 DAC). The models generated individually for Jaú, Piracicaba and Santa Maria demonstrated good performances, allowing quantitative predictions. In this study, the R^2 values ranged from 0.70 (Piracicaba) to 0.75 (Jaú and Santa Maria). This was expected, since the models were calibrated and tested for the same location. In this case, the only factor that influenced the performance was the collection dates, considering that the data were obtained at different vegetative stages of the crop. The accuracy parameter (dr) suffered few variations in the three areas, being 0.72, 0.71 and 0.73 for Jaú, Piracicaba and Santa Maria, respectively (Figure 6).

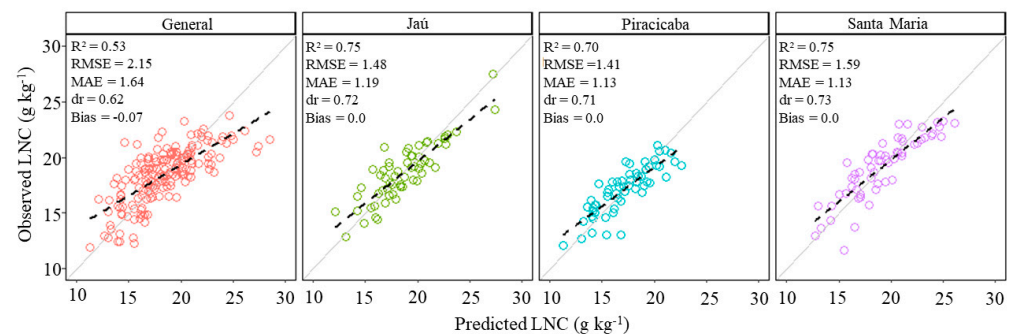


Figure 6. Prediction of LNC by PLSR from the k-fold validation results. The number of factors that best fitted the models were 10, 7, 4 and 7 for the General and the three locations: Jaú, Piracicaba and Santa Maria, respectively.

The highest RMSE was observed for the environment of Santa Maria (1.59 g kg^{-1}), followed by Jaú (1.48 g kg^{-1}) and Piracicaba (1.41 g kg^{-1}). Despite the small variation in the RMSE, in the regions with the greatest dispersion of the spectral data (Figure 4), the highest RMSE was observed (Jaú and Santa Maria). Overall, the models recorded good performances, with a similar error: 8.0% (Jaú), 8.4% (Santa Maria) and 8.2% (Piracicaba) in relation to the average N contents during the five dates (140, 170, 200, 230, 260 DAC). Additionally, MAE, which is a metric that measures the average of the absolute differences between the predicted values and the actual values, varied between 1.64 g kg^{-1} (General) and 1.13 g kg^{-1} (Piracicaba and Santa Maria). In all collection sites, BIAS values were close to zero (global model) or zero (models of Jaú, Piracicaba and Santa Maria), demonstrating a very low bias of the model in relation to the estimated characteristics.

In this study, we aimed to determine the ideal period to quantify the N concentration in sugarcane (Figure 7). The common dates to the three study sites (Piracicaba, Jaú and Santa Maria) were used, specifically at 140, 170, 200, 230 and 260 DAC. The best time was at 140 DAC, with values of $R^2 = 0.90$, $RMSE = 0.74 \text{ g kg}^{-1}$ and $dr = 0.82$. On the other hand, at 260 DAC, beginning of the maturation phase, the quantification of the N contents showed low performance ($R^2 = 0.29$, $RMSE = 1.71 \text{ g kg}^{-1}$ and $dr = 0.44$). On dates 170, 200 and 230 DAC, the R^2 values ranged from 0.42 (170 DAC) to 0.66 (230 DAC), while RMSE ranged from 2.84 g kg^{-1} (170 DAC) to 2.22 g kg^{-1} (230 DAC). It is important to highlight that the models per date were generated from data from three locations with different edaphoclimatic characteristics, which may have influenced the performance.

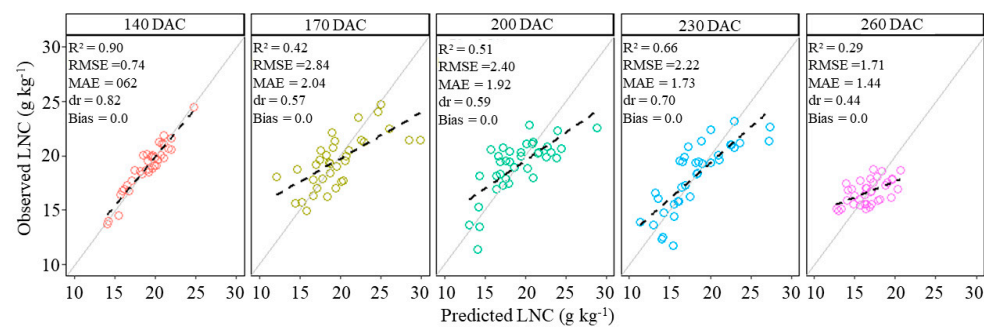


Figure 7. Prediction of LNC by PLSR from the k-fold validation results, where the best date to quantify the leaf N content was evaluated. The data from the three collection sites were used (Jaú, Piracicaba and Santa Maria). The number of factors that best fitted the models was 10, 7, 5, 4 and 7 for the collection dates 140, 170, 200, 260 and 260 DAC, respectively.

3.4. Generalization of the Models

Learning transfer using the LOOCV technique was conducted to investigate the feasibility of testing the models in vegetative stages which did not participate in the calibration phase (Figure 8). Once again, it was observed that the areas with sandy soil texture, Jaú and Santa Maria, presented lower model performance compared to the clayey soil (Piracicaba), with R^2 of 0.06, 0.25 and 0.54, respectively. Despite the low R^2 values, RMSE was only 15.4% (2.87 g kg^{-1}) for Jaú, 10.1% (1.73 g kg^{-1}) for Piracicaba and 14.5% (2.77 g kg^{-1}) for Santa Maria, in relation to the average N contents per site during the five collections. Furthermore, the accuracy indices (dr) demonstrated a more promising performance, reaching 0.34, 0.65 and 0.46 for Jaú, Piracicaba and Santa Maria, respectively. Notably, the area of Piracicaba presented the lowest RMSE, with a value of 1.73 g kg^{-1} , equivalent to 10.1% of error.

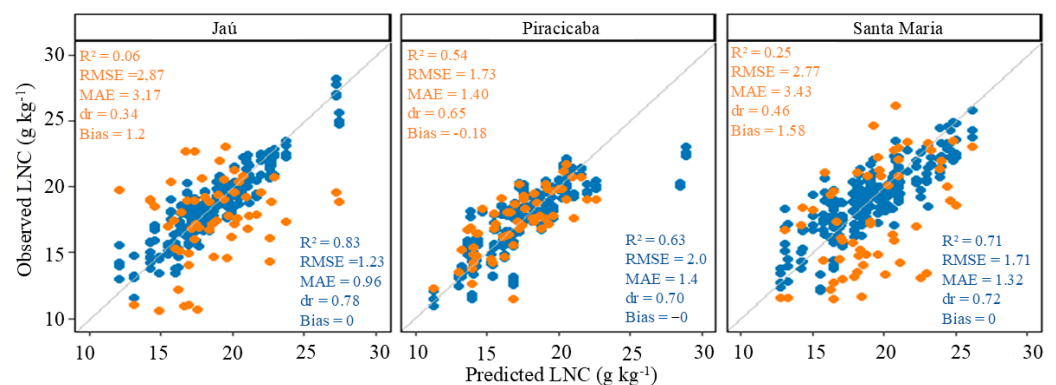


Figure 8. Prediction of LNC by PLSR from the leave-one-date-out cross validation (LOOCV), which evaluated the transferability of the model to vegetative stages which did not participate in the calibration phase (blue values refer to validation and orange, to the test). The number of factors that best fitted the models was 7, 4 and 7 for Jaú, Piracicaba and Santa Maria, respectively.

To evaluate the spatial generalization capacity of the models (Table 1), we tested the prediction potential for the spectral data of one location (Piracicaba) using data from other regions with different edaphoclimatic conditions (Jaú and Santa Maria). The predictive accuracies differed depending on the collection period: 140, 170, 200, 230 and 260 DAC. In this sense, the transfer of learning of the models showed better performance at 140 DAC ($R^2 = 0.69$, $\text{RMSE} = 1.18 \text{ g kg}^{-1}$ and $\text{dr} = 0.61$). Conversely, at the beginning of the maturation phase (260 DAC), the generalization capacity of the model was significantly reduced, with R^2 of 0.05, RMSE of 1.73 g kg^{-1} and dr of 0.38. The collections at 170, 200 and 230 DAC were similar (R^2 ranging from 0.48 to 0.54 and RMSE between 1.02 and 2.56 g kg^{-1}). Although the harvest date occurred in close periods in the three locations (Santa Maria, Jaú

and Piracicaba), vegetative development in the following harvest may have varied between locations. This variation in development, including the onset of maturation at different times, is mainly attributed to differences in edaphoclimatic conditions in each region, and may have directly influenced the performance of the models at the end of the cycle.

Table 1. Statistical parameters of the PLSR models in the training and testing phase in the prediction of N for Piracicaba, trained using data from Jaú and Santa Maria.

	PLSR	140 DAC	170 DAC	200 DAC	230 DAC	260 DAC
Training	Factors	7	7	5	5	2
	R ²	0.93	0.97	0.75	0.6	0.27
	RMSE (g kg ^{−1})	0.62	0.40	1.62	1.99	1.82
	MAE	0.56	0.27	1.35	1.62	1.54
	dr	0.85	0.93	0.73	0.65	0.42
Testing	R ²	0.69	0.49	0.54	0.48	0.05
	RMSE (g kg ^{−1})	1.18	1.75	2.56	1.02	1.73
	MAE	1.47	6.48	4.02	6.74	1.43
	dr	0.61	0.20	0.50	0.14	0.38

4. Discussion

4.1. Influence of Nitrogen on the Leaf Spectral Signature of Sugarcane

Leaf spectral behavior was notably influenced by LNC, regardless of the production environment or collection time (Figure 2). The interaction between the reflectance factor and N concentration was inversely proportional, indicating that the increase in one variable resulted in the decrease in the other, with this relationship being more pronounced at the green wavelengths (550 nm). This characteristic of greater absorption of electromagnetic radiation in the green band in leaves with higher N contents has also been identified in other studies [15,19]. This behavior occurs because VIS spectra are directly related to leaf pigments [15,46], which are responsible for contributing to the physiological functions of the leaves. Chlorophyll, for example, absorbs light energy and transfers it to the photosynthetic apparatus [46]. Green light, in particular, comprises a significant portion of sunlight (14.8%) [47]. It is used in many physiological processes related to plant development, such as growth [48], stomatal opening, flowering [49] and photosynthesis [47]. Compared to the blue (400–500 nm) and red (600–680 nm) bands, which absorb between 80 and 95% of the light, the green band at 550 nm absorbs only about 50% (lettuce) to 90% (broad-leaved evergreen trees) of the radiation [50]. Nevertheless, the absorption of green light can reach greater depth in the leaf and increase photosynthesis by exciting the chloroplasts located in the innermost layers of the mesophyll [48]. As 75% of leaf N is allocated to chloroplasts, and most of it is used for the synthesis of components of the photosynthetic apparatus [51,52], a low N rate can cause negative changes in chloroplasts, since nitrogen is a significant element in both the photosynthesis process and chlorophyll concentration [11,51]. This justifies the greater absorption of radiation in the green range when the N content is higher (Figure 2).

Among the bands from 400 to 750 nm, the blue (450–500 nm) and red (650–700 nm) bands were the only ones that remained stable, with little (red) or almost no (blue) variation in reflectance, regardless of the LNC (Figure 2). The blue and red bands comprise 13.6% and 14% of sunlight, respectively [47]. These spectral bands are notable for their significant energy absorption, with approximately 80 to 95% of the light being absorbed by plants [50]. Pigments are the main source of absorption, especially chlorophylls, with less than 1% of red or blue light being transmitted through the chloroplast [53]. These high absorption properties [48] explain why, despite the variation in the amount of N present in the leaf, the spectral signature remains stable. Furthermore, blue and red light, in addition to being characterized by high energy absorption, influence plant development and physiology [53], such as the regulation of photosynthesis [54], hypocotyl elongation, biomass production and expansion of the leaf area [48,55].

4.2. Environmental Effects on the Spectral Response and Model Performance

N availability is determined by the physical and chemical environment of the soil. Although the response of the nutrients is affected by environmental factors, such as soil temperature and moisture, these factors have a greater impact on N dynamics due to microbial activity [56]. In this study, it was observed that the environment influenced the spectral response of the crop. Although only one variety was used (SP813250), the reflectance for environments with sandy soil texture were similar and with lower average N concentrations (Figure 5A). In this sense, PLSR models were proposed to quantify the N concentrations (g kg^{-1}), using the spectral signatures of fresh leaves as independent variables. Initially, individual predictions were generated for each region (Jaú, Piracicaba and Santa Maria), considering the data from all collections (140, 170, 200, 230 and 260 DAC). Then, a general model was determined using the data from the three areas. Overall, the models proposed per region recorded good performances ($R^2 > 0.70$ and $\text{RMSE} < 1.59 \text{ g kg}^{-1}$), which represents good prediction models, allowing quantitative predictions; on the other hand, the general model had a lower predictive capacity ($R^2 = 0.53$ and $\text{RMSE} = 2.15 \text{ g kg}^{-1}$), with moderate predictions that can be used for evaluation and correlation [40,42]. The results presented in this work are promising for predicting N in fresh leaves and are in agreement with the values found in the literature for sugarcane, $R^2 > 0.70$ and $\text{RMSE} < 1.41 \text{ g kg}^{-1}$ [15]; ryegrass and barley, $R^2 > 0.80$ and $\text{RMSE} < 0.34 \text{ g kg}^{-1}$ [57]; apple trees, $R^2 = 0.6$ [58]; common bean (*Phaseolus vulgaris* L.), $R^2 > 0.63$ and $\text{RMSE} < 4.11 \text{ g kg}^{-1}$ [20].

Subsequently, it was aimed to determine the ideal period to quantify the concentration of N in sugarcane (Figure 7). Therefore, the analyses were tested on dates 140, 170, 200, 230 and 260 DAC, common to the three study sites (Piracicaba, Jaú and Santa Maria). The best date was 140 DAC, with values of $R^2 = 0.90$, $\text{RMSE} = 0.74 \text{ g kg}^{-1}$ and $\text{dr} = 0.82$. On the other hand, at 260 DAC, the beginning of the maturation phase, the quantification of the N contents from the models presented lower performance ($R^2 = 0.29$, $\text{RMSE} = 1.71 \text{ g kg}^{-1}$ and $\text{dr} = 0.44$). Similarly to the results identified in this study, Reyes-Trujillo et al. [43], using the same prediction technique (PLSR), observed that the prediction of N in sugarcane leaves can be performed with greater accuracy when the crop is at the beginning of the vegetative development, such as, for instance, at 90 and 60 days after emergence (DAE), presenting R^2 values of 0.98 and 0.96, respectively.

Although the best times to quantify N are at the beginning of crop development, our results indicate a decrease at 170 and 200 DAC (Figure 7), with improved performance at 230 DAC ($R^2 = 0.66$). At the beginning of the maturation phase (260 DAC), the performance of the model decreases again ($R^2 = 0.29$). Similar results were identified by Reyes-Trujillo et al. [43], with high performance at 60 and 90 DAE, a significant decrease at 120 and 150 DAE ($R^2 = 0.24$), a high performance at the following collection (180 DAE; $R^2 = 0.81$) and a reduction at 210 DAE ($R^2 = 0.46$). One of the factors that may have influenced the performance of the models throughout the sugarcane production cycle was the physiological changes that occur in the plant, affecting the LNC. At the beginning of the development, for example, the plant is in an active growth stage, resulting in greater assimilation of nutrients, including N. It is a crucial nutrient for tiller production, stem and leaf expansion, besides being one of the main components of chlorophyll and the photosynthetic enzymes PEP carboxylase and rubisco [59]. During the sugarcane stalk filling phase, the plant may redistribute N from the leaves to the stalk, which could have affected the LNC concentration, leading to a reduction in foliar N and impacting prediction accuracy. As the sugarcane matures, the demand for N decreases [59], and it is often reported that low N levels in the plant near harvest are essential for increasing sucrose content [60].

4.3. Generalization of the Models in Time and Space

This study used the LOOCV technique to assess the feasibility of testing models in vegetative stages that did not participate in the calibration phase (Figure 8), using only spectral data. It was observed that areas with sandy soil texture (Jaú and Santa Maria) presented lower model performance compared to areas with clayey soil (Piracicaba). In Jaú

and Santa Maria, the values of $R^2 < 0.25$ indicate that the models have weak predictions, capable of distinguishing only high and low values. In contrast, the data from Piracicaba presented a better fit, with $R^2 = 0.56$, suggesting moderate predictions, which may be useful for evaluation and correlation [40,42].

The fall in the performance of the models validated by the LOOCV technique has also been reported in other studies. Sexton et al. [61], for instance, made predictions using PLSR models to quantify N in tobacco leaves, and highlighted a low performance ($R^2 = 0.35$) in the cross-validation (leave-one-out) when all wavelengths (350–2500 nm) were used. Although our results showed low performance by the LOOCV validation, Fiorio et al. [15] also found lower values by LOOCV compared to k-folds, using data from different periods and with higher accuracy ($R^2 = 0.68$ and $RMSE = 1.45 \text{ g kg}^{-1}$), which confirmed the robustness and the possibility of predicting LNC from other collections.

The fall in performance in the transfer of learning of the models between dates may be related to the database used in the calibration phase of the models. This is because the spectral diversity of leaves can be affected by intraspecific differences, such as leaves at different stages [25,62]. Since the LOOCV technique always independently tests the model between vegetative stages, the information provided in the calibration may not have presented all spectral properties of the vegetation, as well as the characteristics of the plants under different levels of nutritional stress and at different temporal (phenological) stages of crop development [24]. Furthermore, sugarcane interacts in a complex way with the environment, and the spectral response can be affected by different agronomic parameters, such as leaf area index, leaf water content, nutritional stress, canopy architecture, among others [63].

To assess the spatial generalization capacity of the models, the prediction potential was tested for the data from one location, using information from the other regions with different edaphoclimatic conditions. In this study, the best performance was identified at 140 DAC ($R^2 = 0.69$, $RMSE = 1.18 \text{ g kg}^{-1}$ and $dr = 0.61$), indicating good prediction models, which allow quantitative predictions [40,42]. Nevertheless, as the crop approaches the maturation phase, the performance of the models decreases, being significantly reduced at the beginning of the maturation phase (260 DAC), with R^2 of 0.05, $RMSE$ of 1.73 g kg^{-1} and dr of 0.38. Although the harvest date occurred in close periods in the three locations (Santa Maria, Jaú and Piracicaba), vegetative development in the following harvest may have varied between locations. In a study of N prediction in pastures from spectral data, a performance of $R^2 = 0.69$ and $RMSE = 18\%$ was observed when PLSR models were applied, being tested in the same location of collection. When tested independently in other environments, the performance was lower (R^2 ranging from 0.45 to 0.48, $RMSE$ between 17% and 19%) [23]. This behavior was also observed in the results presented here. It is believed that as the calibration datasets expand, the performance of the models in terms of accuracy and generalization to other locations will improve [24]. Therefore, the results demonstrated that the spectroradiometry prediction technique is promising and has potential for application in other crop environments, as long as the most representative data sets possible are used [15].

4.4. Limitations and Perspectives

The results showed that the technique of N prediction by spectroradiometry is very promising, but there are still insights that deserve to be investigated to improve the accuracy of the models, especially in the spatial generalization between independent areas. To address this issue, and based on the results of this study, we suggest the use of the most comprehensive database possible, in order to obtain all spectral properties of the vegetation under different levels of nutritional stress, at different phenological stages of sugarcane development and for different edaphoclimatic conditions.

A second option to overcome the challenges faced in the complexity of working with spectral data from fresh leaves would be to explore the potential of predictive models through processed samples (dried and ground). Despite taking longer for data acquisition

due to the time required to process the samples, which involves drying them in an oven, grinding them and then obtaining the spectral response, this method offers other benefits. The main one is the elimination of the water content in the leaf, which directly influences the absorption of radiation, in addition to there being greater consistency in the spectra of the processed leaves, since the spectra of the fresh leaves tend to be more dispersed between repetitions [64].

In future work, we intend to add more independent variables to the spectral models, which may increase their effectiveness. Among these variables are different sugarcane varieties, nitrogen doses and the number of days after cutting. Additionally, we will consider the leaf area index (LAI), chlorophyll content, as well as plant height and biomass. We believe that all of these variables can help explain the variation in LNC.

5. Conclusions

Although the technique of nitrogen quantification by hyperspectral remote sensing needs further studies to be improved, the results presented here demonstrated potential to provide support and benefits in the quantification of nutrients in sugarcane. Moreover, the method of N prediction by VIS-NIR-SWIR spectroradiometry denoted applicability, being sensitive enough to identify small changes in the leaf spectrum, even when there were small nitrogen variations in the leaves.

In this study, we identified that leaves with the highest nitrogen concentrations had the capacity to absorb greater electromagnetic radiation in the visible spectral region (450–680 nm), regardless of the edaphoclimatic conditions. Furthermore, the best time for N prediction and generalization of spectral models in sugarcane leaves is at the beginning of crop development, specifically at 140 DAC.

Author Contributions: Conceptualization, C.A.A.C.S., P.R.F. and R.R.; methodology, C.A.A.C.S. and P.R.F.; software, M.A.d.S., C.A.A.C.S. and R.R.; validation, C.A.A.C.S., P.R.F., M.A.d.S., M.L.C. and R.R.; formal analysis, C.A.A.C.S., P.R.F. and R.R.; investigation, C.A.A.C.S. and P.R.F.; resources, P.R.F.; data curation, C.A.A.C.S. and R.R.; writing—original draft preparation, C.A.A.C.S., P.R.F., M.A.d.S. and M.L.C.; writing—review and editing, C.A.A.C.S., P.R.F., M.A.d.S., M.L.C. and R.R.; supervision, 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 the São Paulo Research Foundation (FAPESP), under grant number 2013/22435-9, and by the Luiz de Queiroz Foundation for Agricultural Studies (FEALQ), which funded the publication of this manuscript.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: To the Luiz de Queiroz Agrarian Studies Foundation—FEALQ, for funding the publication of this work.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Cifuentes, J.; Salazar, V.A.; Cuellar, M.; Castellanos, M.C.; Rodríguez, J.; Cruz, J.C.; Muñoz-Camargo, C. Antioxidant and Neuroprotective Properties of Non-Centrifugal Cane Sugar and Other Sugarcane Derivatives in an In Vitro Induced Parkinson's Model. *Antioxidants* **2021**, *10*, 1040. [\[CrossRef\]](#)
2. Medina, G.S.; Pokorny, B. Agro-industrial development: Lessons from Brazil. *Land Use Policy* **2022**, *120*, 106266. [\[CrossRef\]](#)
3. Shukla, S.K.; Jaiswal, V.P.; Sharma, L.; Dwivedi, A.P.; Nagargade, M. Integration of Bio-products and NPK Fertilizers for Increasing Productivity and Sustainability of Sugarcane-Based System in Subtropical India. *Sugar Tech.* **2023**, *25*, 320–330. [\[CrossRef\]](#)
4. Singh, R.; Gupta, O.P.; Patel, S.K. Energy Use Pattern and Scenario Change in Sugarcane (ratoon) Cultivation for Bhabar Region of Uttarakhand, India. *J. AgriSearch* **2015**, *2*, 119–125.
5. Ismail, M. Response of sugarcane to different doses of Zn at various growth stages. *Pure Appl. Biol.* **2016**, *5*, 311–316. [\[CrossRef\]](#)
6. Shukla, S.K.; Solomon, S.; Sharma, L.; Jaiswal, V.P.; Pathak, A.D.; Singh, P. Green Technologies for Improving Cane Sugar Productivity and Sustaining Soil Fertility in Sugarcane-Based Cropping System. *Sugar Tech.* **2019**, *21*, 186–196. [\[CrossRef\]](#)
7. Lu, C.; Tian, H. Global nitrogen and phosphorus fertilizer use for agriculture production in the past half century: Shifted hot spots and nutrient imbalance. *Earth Syst. Sci. Data* **2017**, *9*, 181–192. [\[CrossRef\]](#)

8. Tian, J.; Tang, M.; Xu, X.; Luo, S.; Condrón, L.M.; Lambers, H.; Cai, K.; Wang, J. Soybean (*Glycine max* (L.) Merrill) intercropping with reduced nitrogen input influences rhizosphere phosphorus dynamics and phosphorus acquisition of sugarcane (*Saccharum officinarum*). *Biol. Fertil. Soils* **2020**, *56*, 1063–1075. [\[CrossRef\]](#)
9. Boschiero, B.N.; Mariano, E.; Torres-Dorante, L.O.; Sattolo, T.M.; Otto, R.; Garcia, P.L.; Dias, C.T.; Trivelin, P.C. Nitrogen fertilizer effects on sugarcane growth, nutritional status, and productivity in tropical acid soils. *Nutr. Cycl. Agroecosyst.* **2020**, *117*, 367–382. [\[CrossRef\]](#)
10. Dinh, H.T.; Watanabe, K.; Takaragawa, H.; Kawamitsu, Y. Effects of Drought Stress at Early Growth Stage on Response of Sugarcane to Different Nitrogen Application. *Sugar Tech.* **2018**, *20*, 420–430. [\[CrossRef\]](#)
11. Bassi, D.; Menossi, M.; Mattiello, L. Nitrogen supply influences photosynthesis establishment along the sugarcane leaf. *Sci. Rep.* **2018**, *8*, 2327. [\[CrossRef\]](#)
12. Dinh, T.H.; Watanabe, K.; Takaragawa, H.; Nakabaru, M.; Kawamitsu, Y. Photosynthetic response and nitrogen use efficiency of sugarcane under drought stress conditions with different nitrogen application levels. *Plant Prod. Sci.* **2017**, *20*, 412–422. [\[CrossRef\]](#)
13. Mailhol, J.; Ruelle, P.; Nemeth, I. Impact of fertilisation practices on nitrogen leaching under irrigation. *Irrig. Sci.* **2001**, *20*, 139–147. [\[CrossRef\]](#)
14. Ranjan, R.; Chopra, U.K.; Sahoo, R.N.; Singh, A.K.; Pradhan, S. Assessment of plant nitrogen stress in wheat (*Triticum aestivum* L.) through hyperspectral indices. *Int. J. Remote Sens.* **2012**, *33*, 6342–6360. [\[CrossRef\]](#)
15. Fiorio, P.R.; Silva, C.A.A.C.; Rizzo, R.; Demattê, J.A.M.; Luciano, A.C.D.S.; da Silva, M.A. Prediction of leaf nitrogen in sugarcane (*Saccharum* spp.) by vis-NIR-SWIR spectroradiometry. *Heliyon* **2024**, *10*, e26819. [\[CrossRef\]](#)
16. Martins, J.A.; Fiorio, P.R.; Silva, C.A.A.C.; Demattê, J.A.M.; da Silva Barros, P.P. Application of Vegetative Indices for Leaf Nitrogen Estimation in Sugarcane Using Hyperspectral Data. *Sugar Tech.* **2023**, *26*, 160–170. [\[CrossRef\]](#)
17. Rodrigues, M.; Cezar, E.; dos Santos, G.L.A.A.; Reis, A.S.; Furlanetto, R.H.; de Oliveira, R.B.; D’Ávila, R.C.; Nanni, M.R. Estimating technological parameters and stem productivity of sugarcane treated with rock powder using a proximal spectroradiometer Vis-NIR-SWIR. *Ind. Crops Prod.* **2022**, *186*, 115278. [\[CrossRef\]](#)
18. Barros, P.P.D.S.; Fiorio, P.R.; Demattê, J.A.D.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. [\[CrossRef\]](#)
19. Silva, C.A.A.C.; Fiorio, P.R.; Rizzo, R.; Rossetto, R.; Vitti, A.C.; Dias, F.L.F.; Oliveira, K.A.D.; Bárbara, M. Detection of nutritional stress in sugarcane by VIS-NIR-SWIR reflectance spectroscopy. *Ciência Rural* **2023**, *53*, e20220543. [\[CrossRef\]](#)
20. Tavares, M.S.; Silva, C.A.A.C.; Regazzo, J.R.; Sardinha, E.J.D.S.; da 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. [\[CrossRef\]](#)
21. Martins, J.A.; Fiorio, P.R.; Barros, P.P.D.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.* **2020**, *43*, e47632. [\[CrossRef\]](#)
22. Nilsson, M.S.; Fiorio, P.R.; Takushi, M.R.H.; Oliveira, A.K.D.S.; Garcia, A.C. Effect of different nitrogen fertilization rates on the spectral response of *Brachiaria brizantha* cv. Marandú leaves. *Eng. Agrícola* **2023**, *43*, e20220008. [\[CrossRef\]](#)
23. Pullanagari, R.R.; Dehghan-Shoar, M.; Yule, I.J.; Bhatia, N. Field spectroscopy of canopy nitrogen concentration in temperate grasslands using a convolutional neural network. *Remote Sens. Environ.* **2021**, *257*, 112353. [\[CrossRef\]](#)
24. Wang, Z.; Chlus, A.; Geygan, R.; Ye, Z.; Zheng, T.; Singh, A.; Couture, J.J.; Cavender-Bares, J.; Kruger, E.L.; Townsend, P.A. Foliar functional traits from imaging spectroscopy across biomes in eastern North America. *New Phytol.* **2020**, *228*, 494–511. [\[CrossRef\]](#)
25. Wan, L.; Zhou, W.; He, Y.; Wanger, T.C.; Cen, H. Combining transfer learning and hyperspectral reflectance analysis to assess leaf nitrogen concentration across different plant species datasets. *Remote Sens. Environ.* **2022**, *269*, 112826. [\[CrossRef\]](#)
26. Alvares, C.A.; Stape, J.L.; Sentelhas, P.C.; de Moraes Gonçalves, J.L.; Sparovek, G. Köppen’s climate classification map for Brazil. *Meteorol. Z.* **2013**, *22*, 711–728. [\[CrossRef\]](#)
27. Pincelli-Souza, R.P.; Bortolheiro, F.P.; Carbonari, C.A.; Velini, E.D.; Silva, M.D.A. Hormetic effect of glyphosate persists during the entire growth period and increases sugarcane yield. *Pest. Manag. Sci.* **2020**, *76*, 2388–2394. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Rossi, M. MAPA PEDOLÓGICO DO ESTADO DE SÃO PAULO: REVISADO E AMPLIADO. Volume 1. São Paulo. 2017. Available online: www.iflorestal.sp.gov.br (accessed on 23 August 2024).
29. 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 J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2562–2570. [\[CrossRef\]](#)
30. Tavares, T.R.; Fiorio, P.R.; Seixas, H.T.; Garcia, A.C.; Barros, P.P.D.S. Effects of storage on vis-NIR-SWIR reflectance spectra of Mombasa grass leaf samples. *Ciência Rural* **2020**, *50*, e20190587. [\[CrossRef\]](#)
31. 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. [\[CrossRef\]](#)
32. Savitzky, A.; Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [\[CrossRef\]](#)
33. Park, H.-S.; Jun, C.-H. A simple and fast algorithm for K-medoids clustering. *Expert. Syst. Appl.* **2009**, *36*, 3336–3341. [\[CrossRef\]](#)
34. Rivera, A.J.; Pérez-Godoy, M.D.; Elizondo, D.; Deka, L.; del Jesus, M.J. Analysis of clustering methods for crop type mapping using satellite imagery. *Neurocomputing* **2022**, *492*, 91–106. [\[CrossRef\]](#)

35. Roth, K.L.; Casas, A.; Huesca, M.; Ustin, S.L.; Alsina, M.M.; Mathews, S.A.; Whiting, M.L. Leaf spectral clusters as potential optical leaf functional types within California ecosystems. *Remote Sens. Environ.* **2016**, *184*, 229–246. [\[CrossRef\]](#)
36. Schirrmann, M.; Hamdorf, A.; Garz, A.; Ustyuzhanin, A.; Dammer, K.-H. Estimating wheat biomass by combining image clustering with crop height. *Comput. Electron. Agric.* **2016**, *121*, 374–384. [\[CrossRef\]](#)
37. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [\[CrossRef\]](#)
38. Shahapure, K.R.; Nicholas, C. Cluster Quality Analysis Using Silhouette Score. In Proceedings of the 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), Sydney, NSW, Australia, 6–9 October 2020; pp. 747–748. [\[CrossRef\]](#)
39. Banas, K.; Banas, A.M.; Heussler, S.P.; Breese, M.B.H. Influence of spectral resolution, spectral range and signal-to-noise ratio of Fourier transform infra-red spectra on identification of high explosive substances. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2018**, *188*, 106–112. [\[CrossRef\]](#)
40. dos Santos, G.L.A.A.; Reis, A.S.; Besen, M.R.; Furlanetto, R.H.; Rodrigues, M.; Crusiol, L.G.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. [\[CrossRef\]](#)
41. Flynn, K.C.; Baath, G.; Lee, T.O.; Gowda, P.; Northup, B. Hyperspectral reflectance and machine learning to monitor legume biomass and nitrogen accumulation. *Comput. Electron. Agric.* **2023**, *211*, 107991. [\[CrossRef\]](#)
42. Rodrigues, M.; Nanni, M.R.; Cezar, E.; dos Santos, G.L.A.A.; Reis, A.S.; de Oliveira, K.M.; de 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. [\[CrossRef\]](#)
43. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
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. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2018**, *189*, 183–189. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Willmott, C.J.; Robeson, S.M.; Matsuura, K. A refined index of model performance. *Int. J. Climatol.* **2012**, *32*, 2088–2094. [\[CrossRef\]](#)
46. 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. [\[CrossRef\]](#)
47. Falcioni, R.; Moriwaki, T.; Bonato, C.M.; de Souza, L.A.; Nanni, M.R.; Antunes, W.C. Distinct growth light and gibberellin regimes alter leaf anatomy and reveal their influence on leaf optical properties. *Environ. Exp. Bot.* **2017**, *140*, 86–95. [\[CrossRef\]](#)
48. Johkan, M.; Shoji, K.; Goto, F.; Hahida, S.; Yoshihara, T. Effect of green light wavelength and intensity on photomorphogenesis and photosynthesis in *Lactuca sativa*. *Environ. Exp. Bot.* **2012**, *75*, 128–133. [\[CrossRef\]](#)
49. Wang, Y.; Folta, K.M. Contributions of green light to plant growth and development. *Am. J. Bot.* **2013**, *100*, 70–78. [\[CrossRef\]](#)
50. Moriwaki, T.; Falcioni, R.; Giacomelli, M.E.; Gibin, M.S.; Sato, F.; Nanni, M.R.; Lima, S.M.; da Cunha Andrade, L.H.; Baesso, M.L.; Antunes, W.C. Chloroplast and outside-chloroplast interference of light inside leaves. *Environ. Exp. Bot.* **2023**, *208*, 105258. [\[CrossRef\]](#)
51. Alharbi, K.; Haroun, S.A.; Kazamel, A.M.; Abbas, M.A.; Ahmida, S.M.; AlKahtani, M.; AlHusnain, L.; Attia, K.A.; Abdelaal, K.; Gamel, R.M. Physiological Studies and Ultrastructure of *Vigna sinensis* L. and *Helianthus annuus* L. under Varying Levels of Nitrogen Supply. *Plants* **2022**, *11*, 1884. [\[CrossRef\]](#)
52. Hak, R.; Rinderle-Zimmer, U.; Lichtenthaler, H.K.; Natr, L. Chlorophyll a fluorescence signatures of nitrogen deficient barley leaves. *Photosynthetica* **1993**, *28*, 151–159.
53. Terashima, I.; Fujita, T.; Inoue, T.; Chow, W.S.; Oguchi, R. Green Light Drives Leaf Photosynthesis More Efficiently than Red Light in Strong White Light: Revisiting the Enigmatic Question of Why Leaves are Green. *Plant Cell Physiol.* **2009**, *50*, 684–697. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Ouzounis, T.; Rosenqvist, E.; Ottosen, C.-O. Spectral Effects of Artificial Light on Plant Physiology and Secondary Metabolism: A Review. *HortScience* **2015**, *50*, 1128–1135. [\[CrossRef\]](#)
55. Johkan, M.; Shoji, K.; Goto, F.; Hashida, S.; Yoshihara, T. Blue Light-emitting Diode Light Irradiation of Seedlings Improves Seedling Quality and Growth after Transplanting in Red Leaf Lettuce. *HortScience* **2010**, *45*, 1809–1814. [\[CrossRef\]](#)
56. Thorburn, P.J.; Meier, E.A.; Probert, M.E. Modelling nitrogen dynamics in sugarcane systems: Recent advances and applications. *Field Crops Res.* **2005**, *92*, 337–351. [\[CrossRef\]](#)
57. Patel, M.K.; Padarian, J.; Western, A.W.; Fitzgerald, G.J.; McBratney, A.B.; Perry, E.M.; Suter, H.; Ryu, D. Retrieving canopy nitrogen concentration and aboveground biomass with deep learning for ryegrass and barley: Comparing models and determining waveband contribution. *Field Crops Res.* **2023**, *294*, 108859. [\[CrossRef\]](#)
58. Ye, X.; Abe, S.; Zhang, S. Estimation and mapping of nitrogen content in apple trees at leaf and canopy levels using hyperspectral imaging. *Precis. Agric.* **2020**, *21*, 198–225. [\[CrossRef\]](#)
59. Singels, A.; Jackson, P.; Inman-Bamber, G. Sugarcane. In *Crop Physiology Case Histories for Major Crops*; Elsevier: Amsterdam, The Netherlands, 2021; pp. 674–713. [\[CrossRef\]](#)
60. van Heerden, P.D.R.; Donaldson, R.A.; Watt, D.A.; Singels, A. Biomass accumulation in sugarcane: Unravelling the factors underpinning reduced growth phenomena. *J. Exp. Bot.* **2010**, *61*, 2877–2887. [\[CrossRef\]](#)

61. Sexton, T.; Sankaran, S.; Cousins, A.B. Predicting photosynthetic capacity in tobacco using shortwave infrared spectral reflectance. *J. Exp. Bot.* **2021**, *72*, 4373–4383. [[CrossRef](#)]
62. Shapira, U.; Herrmann, I.; Karnieli, A.; Bonfil, D.J. Field spectroscopy for weed detection in wheat and chickpea fields. *Int. J. Remote Sens.* **2013**, *34*, 6094–6108. [[CrossRef](#)]
63. Abdel-Rahman, E.M.; Ahmed, F.B.; van den Berg, M. Imaging spectroscopy for estimating sugarcane leaf nitrogen concentration. In Proceedings of the Remote Sensing for Agriculture, Ecosystems, and Hydrology X, Wales, UK, 15–18 September 2008; Neale, C.M.U., Owe, M., D’Urso, G., Eds.; p. 71040V. [[CrossRef](#)]
64. Prananto, J.A.; Minasny, B.; Weaver, T. Near infrared (NIR) spectroscopy as a rapid and cost-effective method for nutrient analysis of plant leaf tissues. *Adv. Agron.* **2020**, *164*, 1–49. [[CrossRef](#)]

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.