Hyperspectral Response to Leaf Nitrogen in Sugarcane: Dynamic Effects of Cultivar, Growth Stage, and Leaf Position with Model Inversion

Guangzai Huang
Guangxi University
Meixue Meng
Guangxi University
Shijian Han
Guangxi University
Zhenli He
University of Florida
Suli Li
lisuli88@163.com

Guangxi University
Yi Gan
Guangxi University
Zhigang Li
Guangxi University

Research Article

Keywords: Sugarcane, Leaf nitrogen, Hyperspectral, Cultivar, Growth stage, Leaf position, Random Forest

Posted Date: February 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8553743/v1

Additional Declarations: No competing interests reported.

Hyperspectral Response to Leaf Nitrogen in Sugarcane: Dynamic Effects of Cultivar, Growth Stage, and Leaf Position with Model Inversion

Guangzai Huang[1,2], Meixue Meng[1,2], Shijian Han[1,2], Zhenli He[3], Suli Li[1,2*], Yi Gan[1,2], Zhigang Li[1,2*]

1 Guangxi Key Laboratory of Sugarcane Biology, College of Agriculture, Guangxi University, 100 Daxue Rd., Nanning 530004, China

2 Key Laboratory of Crop Cultivation and Tillage, College of Agriculture, Guangxi University, 100 Daxue Rd., Nanning 530004, China

3 Department of Soil, Water, and Ecosystem Sciences/Indian River Research and Education Center, Institute of Food and Agricultural Sciences, University of Florida, Fort Pierce, FL 34945, USA

Authors' E-mail:13100458731@163.com; 15578099346@163.com; hanshijian@126.com; zhe@ufl.edu; lisuli88@163.com; GanYi2025@126.com; lizhigangnn@163.com

*Corresponding author: Suli Li (Email: lisuli88@163.com); Zhigang Li (Email: lizhigangnn@163.com), College of Agriculture, Guangxi University, 100 Daxue Rd., Nanning 530004, People's Republic of China

HIGHLIGHTS

1. Quantified the significant interactive effects of cultivar, growth stage, and leaf position on sugarcane leaf nitrogen (N) content and hyperspectral reflectance, demonstrating that a universal spectral model is unsuitable for precision N management.
2. Identified the 2nd leaf as the most stable and optimal sampling unit for spectral-based N monitoring across diverse cultivars and growth stages, providing a standardized protocol to reduce field sampling variability.
3. Determined that the first-derivative transformation combined with Random Forest (FD-RF) outperforms conventional linear models, offering an optimized, cultivar-specific modeling framework for accurate N estimation.
4. Revealed cultivar-specific spectral mechanisms, particularly in the red-edge region, highlighting the need for tailored sensor configurations or indices rather than generic broadband vegetation indices.

IMPACT: This study elucidates the complex, interactive nature of factors influencing the hyperspectral estimation of N in sugarcane and establishes a robust, cultivar-aware modeling framework centered on the 2nd leaf and FD-RF processing.

Keywords: Sugarcane; Leaf nitrogen; Hyperspectral; Cultivar; Growth stage; Leaf position; Random Forest

ABSTRACT

Rapid and non-destructive monitoring of leaf nitrogen (N) content (LNC) is essential for precision N management in sugarcane (*Saccharum officinarum L.*). However, the accuracy of hyperspectral estimation is challenged by the dynamic interactions among cultivar, growth stage, and leaf position. This study systematically investigated the effects of these three factors on LNC and leaf hyperspectral reflectance (400–1000 nm) across six main sugarcane varieties. We identified sensitive spectral bands and developed LNC inversion models using Partial Least Squares Regression (PLSR) and Random Forest (RF). The results revealed highly significant interactive effects ($P < 0.01$) of cultivar, growth stage, and leaf position on both LNC and spectral responses. The 2nd leaf was identified as the most stable and optimal leaf for spectral monitoring, thus is recommended as the standard sampling target for field diagnosis. Correlation analysis showed a consistent negative correlation between LNC and reflectance in the visible region (520–570 nm), a positive correlation in the near-infrared (930–999 nm), and cultivar-specific responses in the red-edge region (700–730 nm). The first derivative (FD) transformation most effectively enhanced spectral features and correlations. The RF algorithm significantly outperformed PLSR, with the "FD + RF" combination yielding the best prediction performance for most varieties (e.g., test set $R^2 = 0.50$ for ROC22). This study elucidates the hyperspectral response mechanisms of sugarcane LNC under multi-factor interactions and provides a robust framework for developing cultivar-specific models, paving the way for precision N diagnosis in sugarcane production.

Context: Although previous studies has resulted in substantial knowledge on crop N monitoring via spectroscopy, systematic investigations into the "leaf N content–spectral characteristics" response mechanisms in sugarcane at the leaf level remain limited.

Aims: This study was designed to address these critical research gaps. The specific objectives were to: (1) quantify the independent and interactive effects of cultivar, growth stage, and leaf position on sugarcane LNC and spectral reflectance; (2) identify the most sensitive spectral bands responsive to LNC changes across different varieties; and (3) construct and evaluate robust estimation models for sugarcane LNC using advanced machine learning algorithms. Our findings are expected to provide a solid theoretical foundation and technical support for developing remote sensing technologies tailored for precision N management in sugarcane fields.

1. INTRODUCTION

Sugarcane (*Saccharum officinarum L.*) stands as a cornerstone of the global agricultural economy, serving as a primary source for sugar and an increasingly vital feedstock for bioenergy production (Heinrichs et al. 2017). Its productivity and profitability are intrinsically linked to effective nutrient management, with nitrogen (N) being the most pivotal element governing growth, yield, and quality (Ali et al. 2025). As a fundamental constituent of amino acids, proteins, nucleic acids, and chlorophyll, N directly regulates key physiological processes, including photosynthesis, carbon and nitrogen metabolism, and the accumulation of dry matter (Liu et al. 2025). However, the efficiency of N fertilizer application in conventional sugarcane cultivation is often suboptimal. Excessive N application not only failed to yield commensurate benefits but also triggered a cascade of detrimental effects, including growth imbalances, increased susceptibility to pests and diseases, and severe environmental pollution such as water eutrophication and greenhouse gas emissions. Consequently, the implementation of precision N management is an indispensable pathway towards achieving the synergistic goals of high yield, superior quality, resource efficiency, and environmental sustainability in the sugarcane industry (Mulla 2013).The prerequisite for this paradigm shift is the ability to perform rapid and accurate diagnosis of the plant's N status.

Traditional methods for diagnosing plant N status predominantly rely on destructive sampling followed by laboratory chemical analysis, such as the Kjeldahl method. While considered the "gold standard" for accuracy, these approaches are inherently labor-intensive, time-consuming, costly, and exhibit significant temporal and spatial lag. These limitations render them impractical for supporting large-scale, high-frequency field decisions, thus failing to meet the real-time information demands of modern precision agriculture (Chowdhury et al. 2024; Li et al. 2025). In this context, hyperspectral remote sensing has emerged as a revolutionary, non-destructive tool for rapid monitoring of crop N. The theoretical foundation lies in the fact that the N status of a plant directly influences the biochemical composition (e.g., chlorophyll, protein content) and physical structure (e.g., mesophyll cell arrangement, palisade tissue thickness) of its leaves. These internal changes subsequently alter the absorption, transmission, and reflection characteristics of the leaf in specific regions of the electromagnetic spectrum, particularly in the visible and near-infrared ranges, forming unique spectral "fingerprints" (Hu et al. 2025; Kanash et al. 2023). By deciphering these spectral features, it becomes possible to indirectly infer the N content of crops.

In recent years, estimating crop N content using hyperspectral technology has become a research hotspot in agricultural remote sensing, yielding significant progress in major staple crops like wheat, maize, and rice. A suite of effective spectral indices and prediction models has been developed, including the Normalized Difference Vegetation Index (NDVI), the Photochemical Reflectance Index (PRI), and parameters derived from the red-edge position (Chegoonian et al. 2025; Meiyan et al. 2023; Yue et al. 2024). For instance, studies on rubber trees have identified a strong correlation between reflectance at 730 nm and leaf N content (Hu et al. 2021). Research on maize has demonstrated that first-derivative transformation of leaf spectra can yield estimation models with higher accuracy (Tang et al. 2025). Furthermore, the integration of spectral features with machine learning algorithms, such as Support Vector Machine Regression, has further enhanced the predictive capability for wheat N content Wang et al. 2024). The results robustly validate the universality and effectiveness of spectral techniques for monitoring N across diverse crop species.

Nevertheless, the successful application of spectral technology to sugarcane N monitoring faces several key scientific challenges. Firstly, the spectral response of crops is profoundly influenced by genotype (cultivar). Different sugarcane cultivars exhibit inherent differences in canopy architecture, leaf morphological and anatomical traits, and nitrogen use efficiency, which may lead to a lack of universality in the "N-spectral" relationship model (Li et al. 2025a). This genetic variability poses a direct obstacle to developing a one-size-fits-all spectral algorithm for variable-rate N application across diverse sugarcane plantations. Secondly, the growth stage is another critical regulatory factor. As sugarcane progresses from the seedling, tillering, elongation, to maturation stages, its growth center, nitrogen demand, and allocation patterns undergo dynamic changes. Consequently, the spectral characteristics of the same leaf at different growth stages and their correlation with N content are likely to be altered (Silva et al. 2024) . This temporal dynamism necessitates growth-stage-specific calibration or adaptive models for accurate season-long monitoring. Moreover, leaf position, serving as a direct indicator of the spatial distribution of N within the plant, is a factor often overlooked in spectral studies. Functional leaves (e.g., the 1st leaf) and older leaves (e.g., the 3rd leaf) differ significantly in N content and physiological function; their combined analysis can introduce substantial noise, reducing the sensitivity and accuracy of estimation models (Li et al. 2022). Inconsistent sampling of leaf position in field practice is therefore a major source of uncertainty that hinders the operationalization of spectral diagnosis.

Although previous studies have resulted in substantial knowledge on crop N monitoring via spectroscopy, systematic investigations into the "leaf N content–spectral characteristics" response mechanisms in sugarcane at the leaf level remain limited. Existing studies have predominantly focused on the canopy scale, which may obscure fine-grained leaf-level responses, or have considered only a single factor, lacking a comprehensive assessment of the interactive effects of cultivar, growth stage, and leaf position. This gap impedes the development of robust, transferable models that can function reliably under real-world field conditions characterized by genetic and temporal diversity. Therefore, this study was designed to address these critical research gaps. The specific objectives were to: (1) quantify the independent and interactive effects of cultivar, growth stage, and leaf position on sugarcane LNC and spectral reflectance; (2) identify the most sensitive spectral bands responsive to LNC changes across different varieties; and (3) construct and evaluate robust estimation models for sugarcane LNC using advanced machine learning algorithms. Our findings are expected to provide a solid theoretical foundation and technical support for developing remote sensing technologies tailored for precision N management in sugarcane fields.

2. MATERIALS AND METHODS

2.1. Experimental Site Description

The field experiment was conducted over the 2023-2024 growing season at the Plant Botanical Garden of the Agricultural College, Guangxi University (E 108.3°, N 22.8°). The region is characterized by a subtropical monsoon climate, featuring year-round warm temperatures and abundant sunshine. During the experimental period, the mean annual temperature was 21.5°C, and the total annual precipitation was approximately 1300 mm (Yang et al. 2024).

2.2. Plant Materials and Experimental Design

Six locally dominant sugarcane cultivars, 'GL07150', 'GL05136', 'ROC22', 'ROC16', 'GT44', and 'GT42', were selected as plant materials for this study. All experimental materials were healthy, one-year-old, newly planted sugarcane, established on March 22, 2024.

A three-factor completely randomized design was employed. The factors and their levels were as follows: **Cultivar (Cultivar):** 'GL07150', 'GL05136', 'ROC22', 'ROC16', 'GT44', 'GT42' (6 levels).
Growth Stage: Seedling stage, Tillering stage, Elongation stage, Maturation stage (4 levels).
Leaf Position: the 1st leaf (the first fully expanded leaf from the apex), the 2nd leaf, the 3rd leaf (the third leaf from the apex) (3 levels), all leaves were fully expanded.

For each cultivar at each growth stage, three biological replicates (plants) were sampled, with each plant providing leaves at the three positions. This resulted in 3 replicates × 3 leaf positions = 9 samples per cultivar per stage. Sampling was conducted four times on: June 22, 2024 (Seedling stage), August 22, 2024 (Tillering stage), October 22, 2024 (Elongation stage), and December 22, 2024 (Maturation stage). During each sampling event, the 1st, 2nd, and 3rd leaves were collected from each plant. This resulted in a total of 6 cultivars × 4 growth stages × 3 replicates × 3 leaf positions = **648 leaf samples**.

2.3. Leaf Hyperspectral Data Acquisition

Hyperspectral reflectance data were acquired using a fiber optic spectrometer (NIR 17S, Shanghai Fuxiang Optical Co., China) with a spectral range of 326–1105 nm and a spectral resolution of 0.8 nm. Measurements were conducted with a leaf clip attachment that held the fiber optic probe at a fixed distance and angle (e.g., 0° viewing angle) relative to the leaf surface, ensuring consistent measurement geometry. Spectral measurements were performed under clear, windless conditions between 10:00 and 14:00 hours local time to minimize the impact of changing solar angles. Immediately after collection, leaf samples were placed in a preservation box and promptly transported to the laboratory for spectral measurement on the NIR 17S reflectance measurement platform. This indoor setup ensured stable and consistent lighting conditions, minimizing the influence of ambient environmental fluctuations(Stamford et al. 2024). For each leaf sample, three points on both sides of the midrib were selected for measurement. The average reflectance value of these six measurements was calculated and recorded as the spectral reflectance for that leaf. A standard white reference panel (Spectralon, Labsphere, USA) was used for calibration before and after each measurement session to correct for instrument dark current and ensure data accuracy Li et al. 2025b).

2.4. Determination of Leaf N Content

Following spectral measurement, leaf samples were immediately collected and placed in sealed bags, stored in a cool box, and rapidly transported to the laboratory. The samples were first deactivated in an oven at 105°C for 30 minutes and then dried at 85°C to a constant weight. The dried samples were ground into a fine powder and passed through a 0.5 mm sieve to prepare homogeneous samples for chemical analysis.

The leaf N concentration was determined using the standard Kjeldahl method (Sparks et al. 1996). Briefly, 0.1 g of the dried powder was accurately weighed and digested with concentrated sulfuric acid using a Kjeldahl digestion system (K-424, BUCHI, Switzerland). The digested samples were then distilled and titrated using a fully automated Kjeldahl nitrogen analyzer (K-370, BUCHI, Switzerland). The nitrogen content was expressed as milligrams per gram of dry mass (mg/g).

2.5. Hyperspectral Data Preprocessing

2.5.1. Savitzky-Golay Smoothing

The raw spectral data often contain high-frequency noise. To enhance the signal-to-noise ratio while preserving the intrinsic shape of the spectral curves, Savitzky-Golay (S-G) smoothing was applied as a primary preprocessing step (Savitzky and Golay 1964). This method is a non-parametric digital filter based on local polynomial least-squares fitting in a moving window. For any data point in the discrete signal sequence, a symmetric window of width 2m+1 (containing 2m+1 neighboring points from -m to +m) is constructed. An nth order polynomial is then fitted to the data within this window:

$$y(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \quad (1)$$

Where x is the relative coordinate within the window (central point x = 0), and a_0, a_1, ..., a_n are the polynomial coefficients. In this study, a second-order polynomial (n=2) with a window size of 11 points (m=5) was employed for smoothing. The SG smoothing window should be chosen based on the noise level and the detail requirements of the signal, using an odd number greater than the polynomial order. The polynomial order should match the complexity of the signal and must be coordinated with the window size to prevent unstable fitting.

2.5.2. Spectral Transformations

Two common spectral transformations were applied to the S-G smoothed data to extract more informative features:

First-Derivative (FD): The first derivative of the reflectance spectrum was calculated to remove baseline drift and amplify subtle spectral features, which is particularly useful for resolving overlapping absorption peaks (Tsai and Philpot, 1998). It was computed based on a differential approximation of the derivative using the formula:

$$R'(\lambda_i) = \frac{R(\lambda_{i+1}) - R(\lambda_{i-1})}{2\Delta\lambda} \tag{2}$$

where, $\rho'(\lambda_i)$ is the first-derivative value at wavelength λ_i, $\rho(\lambda_{i+1})$ and $\rho(\lambda_{i-1})$ are the reflectance values at the subsequent and preceding wavelengths, respectively, and $\Delta\lambda$ is the wavelength interval.

Logarithmic Transformation (Log(1/R)): The reflectance (R) data were transformed to apparent absorbance (A) using the formula $A = \log_{10}(1/R)$. This transformation is grounded in Lambert-Beer law, aiming to establish a linear theoretical foundation between spectral data and biochemical constituent concentrations. It enhances weak absorption features associated with leaf chemical components, such as nitrogen, thereby improving the discernibility of characteristic bands (Jean-Baptiste Féret et al. 2021) .

Given that the reflectance signals in the 323–399 nm and 1001–1105 nm wavelength ranges exhibited severe distortion and low signal-to-noise ratio, these bands were excluded from all subsequent analyses. Only the spectral data within the 400–1000 nm range were utilized.

2.6. Feature Band Selection and Statistical Analysis

2.6.1. Feature Band Selection

Pearson correlation analysis was employed to identify the spectral bands most sensitive to variations in leaf nitrogen content (Hou et al. 2025). The Pearson correlation coefficient (r), which quantifies the strength and direction of a linear relationship between two continuous variables, was calculated between the LNC and the reflectance (and its transformations) at each wavelength across all samples. The coefficient is defined as:

$$r_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{3}$$

Where X_i and Y_i are the spectral reflectance (or transformed value) and LNC of the *i*-th sample, respectively, and $\bar{X}$ and $\bar{Y}$ are their respective sample means. The value of *r* ranges from -1 to 1. For each cultivar and spectral type (raw, FD, Log(1/R)), the wavelengths with the highest absolute correlation coefficients (|r|) with LNC were selected as characteristic bands for subsequent model development.

2.6.2. Data Analysis

Microsoft Excel 2021 (Microsoft Corp., USA) was used for preliminary data organization and calculations. A three-way analysis of variance (ANOVA) was performed using SPSS 26.0 (IBM Corp., USA) to assess the main effects and interaction effects of cultivar, growth stage, and leaf position on both leaf nitrogen content and reflectance values at key characteristic bands. The fixed factors in the model were cultivar, growth stage, and leaf position. When a significant effect was detected ($\alpha = 0.05$), the Least Significant Difference (LSD) post-hoc test was used for multiple comparisons among treatment means. The Python 3.4 (Python Software Foundation, USA) was utilized for all spectral preprocessing steps, including S-G smoothing, derivative calculation, and logarithmic transformation.

2.7. Model Development and Validation

Two modeling approaches were chosen to represent linear and nonlinear paradigms: Partial Least Squares Regression for handling collinear spectral data, and Random Forest for capturing complex, nonlinear relationships. The dataset for each cultivar was partitioned into training and testing sets using a stratified random sampling approach based on the LNC values to ensure representative distribution in both sets. To construct and evaluate models for estimating LNC from spectral data, two distinct algorithms were employed:

Partial Least Squares Regression (PLSR): PLSR is a robust linear technique particularly suited for modeling relationships between a large number of collinear predictor variables (spectral bands) and a response variable (LNC) (Burnett et al. 2021) . Partial least squares regression (PLSR) was performed using the PLS Regression module from the scikit-learn library (version 1.3.0). The model was configured with three components were selected as optimal for capturing the majority of spectral variance related to LNC while preventing overfitting, the maximum number of iterations for the optimization algorithm was set to 2000 (maxiter=2000) (Li et al. 2024).

Random Forest (RF): RF is an ensemble machine learning algorithm that operates by constructing a multitude of decision trees during training and outputting the mean prediction of the individual trees. It is highly effective at capturing non-linear relationships and is relatively resistant to overfitting (Breiman 2001). Random forest regression is implemented using random forest regression module of scikit learn library. The search ranges were: n_estimators = [60, 80, 100], max_depth = [4, 6, 8, None], min_samples_split = [2, 4, 6]. The combination that yielded the lowest cross-validation RMSE was selected (n_estimators=80, max_depth=6, min_samples_split=4). This process ensures a balance between model complexity and generalization ability (Zhang et al. 2024).

2.7.1 Data Splitting and Validation Framework

For each cultivar (n = 108 samples per cultivar), a two-step validation framework was employed to robustly develop and evaluate the models.

Step 1: Model Calibration and Hyperparameter Tuning. The dataset was first split into a temporary training set (83%, 90 samples) and a hold-out test set (17%, 18 samples) using stratified random sampling, ensuring that the distribution of LNC values was representative in both subsets. The hold-out test set was strictly reserved for the final, independent evaluation (Step 2). Only the temporary training set was used for the subsequent model calibration and hyperparameter optimization (e.g., the grid search for RF and component selection for PLSR described above). The performance metrics (RMSE, R^2) of the training set and test set guided the selection of the best hyperparameters and spectral preprocessing method for each cultivar-algorithm combination.

Step 2: Final Model Training and Independent Testing. Using the optimal hyperparameters identified in Step 1, a final model was retrained on the entire temporary training set (90 samples). This final model was then evaluated once on the completely unseen hold-out test set (18 samples). The performance metrics (R^2 and RMSE) reported in Table 3 for the "Testing" set are the results of this single, final evaluation on the independent test set. The metrics for the "Training" set in Table 3 are derived from the final model's performance on the entire dataset used for its training (i.e., the 90 samples).

3. RESULTS

3.1. Effects of Cultivar, Growth Stage, and Leaf Position on Leaf Nitrogen Content

The leaf N content (LNC) of sugarcane was significantly influenced by the independent and interactive effects of cultivar, growth stage, and leaf position. A repeated measures three-way ANOVA revealed that all two-way and three-way interaction effects on LNC were highly significant ($P < 0.01$). Specifically, the F-values were 3.56 for the Cultivar × Leaf Position interaction, 6.07 for the Cultivar × Growth Stage interaction, 4.48 for the Leaf Position × Growth Stage interaction, and 5.83 for the three-way Cultivar × Leaf Position × Growth Stage interaction. This indicates that the differences in LNC were the result of the complex interplay among these three factors.

The dynamic changes in LNC for the six sugarcane varieties across four growth stages and three leaf positions are detailed in Table 1. At the **seedling stage**, cultivar ROC22 exhibited significantly higher N content in the 1st leaf than ROC16, GL07150, GT44, and GT42, but did not differ significantly from GL05136. At the 2nd leaf, GL05136 showed significantly higher N content than ROC16, GL07150, GT44, and GT42, but did not differ significantly from ROC22. At the 3rd leaf, GT42 had significantly higher N content than GL07150, ROC16, and GT44. Within varieties, GL05136 had significantly higher N in the 2nd leaf compared to 1st and 3rd leaves. ROC16 had significantly higher N in the 3rd leaf compared to the 1st leaf, while GT44 showed significantly higher N in the 3rd leaf than in the 2nd leaf, while GT42 showed both the 2nd and 3rd leaves had significantly higher N than the 1st leaf (Table 1).

During the **tillering stage**, varietal differences became more pronounced. At the 1st leaf, GT42 exhibited the highest N content, significantly surpassing the other five varieties. ROC22 also showed significantly higher N content than GL07150 and ROC16 at this leaf position. A similar pattern was observed at the 2nd and 3rd leaf with GT42 leading (Table 1).

In the **elongation stage**, GL05136 consistently demonstrated high N content across leaf positions. At the 1st leaf, GL05136, GT44, and GT42 showed significantly higher N content than GL07150, ROC22, and ROC16. At the 2nd and 3rd leaf, where GL05136 showed significantly higher N content than GL07150, ROC22, ROC16, and GT44. Within varieties, GL05136 had significantly higher N in the 2nd leaf than in the 3rd leaf. Both ROC16 and GT44 exhibited significantly higher N in the 1st leaf than in the 3rd leaf (Table 1).

By the maturation stage, ROC22 generally maintained higher N levels, particularly in the 1st and 3rd leaves. At the 1st leaf, ROC22 exhibited the highest N content, which was significantly greater than that of the other five varieties, while no significant differences were observed among the remaining varieties. At the 3rd leaf, GL07150, GL05136, and GT42 showed significantly higher N content than ROC16 and GT44. Within varieties, GL05136 had significantly higher N in the 3rd leaf compared to the 1st leaf. ROC22 had significantly higher N in the 1st leaf compared to the 2nd leaf, while ROC16 and GT44 showed an opposite trend, with the 1st and 2nd leaves having higher N than the 3rd leaf (Table 1).

In summary, the 2nd leaf often served as a stable indicator, showing high and relatively consistent N content across multiple growth stages and varieties, highlighting its potential as a key leaf for monitoring plant N status. Therefore, from a practical application standpoint, the 2nd leaf is recommended as the standardized sampling unit for spectral-based nitrogen assessment in sugarcane to minimize variability introduced by leaf ontogeny.

Table 1 Leaf nitrogen content across six sugarcane cultivars, four growth stages, and three leaf positions

Growth Stage	Cultivar	1st Leaf ($mg \cdot g^{-1}$)	2nd Leaf ($mg \cdot g^{-1}$)	3rd Leaf ($mg \cdot g^{-1}$)
Seedling	GLO7150	14.83 ± 1.07	**16.11 ± 1.13a**	15.85 ± 1.49
	GL05136	15.94 ± 2.82	**19.95 ± 2.26a**	19.50 ± 2.58
	ROC22	**17.74 ± 2.15**	**18.64 ± 2.59a**	17.55 ± 3.30
	ROC16	13.32 ± 3.26	15.53 ± 2.01	**16.05 ± 1.19a**
	GT44	15.24 ± 2.01	13.53 ± 1.62	**16.59 ± 1.16a**
	GT42	14.78 ± 2.67	17.68 ± 0.99	**19.58 ± 1.43a**
Tillering	GLO7150	13.87 ± 2.93	**15.43 ± 1.89a**	15.11 ± 2.59
	GL05136	15.10 ± 1.09	16.84 ± 2.52	**17.22 ± 1.21a**
	ROC22	16.88 ± 2.46	**16.36 ± 2.78a**	16.08 ± 1.81
	ROC16	14.25 ± 1.90	**15.80 ± 1.11a**	14.81 ± 1.90
	GT44	15.56 ± 1.21	15.51 ± 1.27	**15.75 ± 1.55a**
	GT42	**19.60 ± 2.15a**	19.50 ± 1.94	18.12 ± 1.66
Elongation	GLO7150	12.71 ± 1.13	12.65 ± 0.86	**13.39 ± 3.46a**
	GL05136	16.34 ± 1.50	**17.42 ± 1.01a**	15.38 ± 1.29
	ROC22	13.56 ± 0.89	**14.29 ± 0.59a**	12.87 ± 0.92

Table 1 Leaf nitrogen content across six sugarcane cultivars, four growth stages, and three leaf positions

Growth Stage	Cultivar	1st Leaf (mg·g⁻¹)	2nd Leaf (mg·g⁻¹)	3rd Leaf (mg·g⁻¹)
	ROC16	13.76 ± 3.93	**13.44 ± 1.53a**	11.65 ± 0.77
	GT44	**15.77 ± 2.33a**	14.35 ± 0.89	12.69 ± 2.36
	GT42	15.72 ± 0.72	**16.50 ± 1.44a**	15.22 ± 1.36
Maturity	GLO7150	13.88 ± 2.02	**15.04 ± 1.86a**	14.97 ± 1.50
	GL05136	13.49 ± 1.71	14.72 ± 0.88	**15.87 ± 3.39a**
	ROC22	**17.76 ± 1.58a**	15.90 ± 1.28	17.35 ± 1.83
	ROC16	14.17 ± 2.17	**15.18 ± 1.41a**	11.83 ± 2.16
	GT44	15.11 ± 0.82	**15.25 ± 1.13a**	13.11 ± 1.72
	GT42	14.10 ± 2.02	15.42 ± 1.54	**16.03 ± 0.84a**

Note: Data are presented as mean ± standard deviation (n = 3). Within the same row (i.e., for a given cultivar at a given growth stage), different superscript lowercase letters indicate significant differences among leaf positions ($P < 0.05$, LSD test). Values in bold with a superscript 'a' denote the leaf position with the highest nitrogen content for that specific cultivar at that specific growth stage.

3.2. Spectral Reflectance Characteristics Across Growth Stages, Varieties, and Leaf Positions

The analysis focused on the 400–1000 nm spectral range, as signals outside this range (323–399 nm and 1001–1105 nm) were distorted. The spectral curves for all varieties exhibited the typical pattern of green vegetation across growth stages: a chlorophyll absorption trough in the blue region (around 450 nm), a green reflectance peak (around 550 nm), a strong red absorption trough (around 670 nm), a steep rise in the red-edge region (680–750 nm), and a high-reflectance plateau in the near-infrared (NIR) region (750–1000 nm).

3.2.1. Seedling Stage

Caption for Figure 1: Mean leaf spectral reflectance of six sugarcane varieties at the seedling stage for different leaf positions (1st, 2nd, 3rd). The main panel (a) shows the full spectrum (400–1000 nm), while subpanels (b), (c), and (d) provide detailed views of the green, red, and near-infrared regions, respectively. Shaded areas represent ±1 standard deviation from the mean (3rd). The 2nd leaf position generally exhibits higher NIR reflectance across varieties. At the seedling stage, spectral differences among varieties and leaf positions were already present but relatively modest. In the visible region (400-680 nm), reflectance was generally below 13%, with the maximum difference between varieties and leaf positions being 4.59% at 550.31 nm (GT44 1st leaf vs. GLO7150 3rd leaf). The average slope of the red-edge was approximately $0.36\%\cdot nm^{-1}$(680-750nm). In the NIR plateau (750-1000 nm), reflectance ranged from 33.51% to 43.25%, with an extreme difference of 11.73%. Within a cultivar, the 2nd leaf often showed the highest reflectance in the NIR, followed by the 1st and 3rd leaves (e.g. GT44 and ROC16) (Figure 1).

3.2.2. Tillering Stage

Caption for Figure 2: Mean leaf spectral reflectance at the tillering stage. Spectral differences among varieties and leaf positions become more pronounced compared to the seedling stage, with an overall decrease in visible and NIR reflectance associated with increased chlorophyll content and biomass. During tillering, overall reflectance in the visible and NIR regions decreased compared to the seedling stage, coinciding with increased chlorophyll content. Differences among varieties and leaf positions became more pronounced. The green peak difference increased to 5.23%, and the NIR plateau range widened (32.81%–38.58%), with a maximum difference of 5.77%. The red-edge average slope also $0.36\%\cdot nm^{-1}$. The spectral divergence between cultivar groups, such as the GT series, became more evident (Figure 2).

3.2.3. Elongation Stage

Caption for Figure 3: Mean leaf spectral reflectance at the elongation stage. Reflectance increases in both visible and NIR regions due to leaf maturation and structural development. The spectral distinction between cultivar groups (e.g., GT series) is clear and stable. In the elongation stage, reflectance levels in the visible and NIR regions increased compared to the tillering stage, associated with mature leaf structure. But the difference in the green reflectance peak decreased to 3.33%. The red-edge average slope increased to $0.38\%\cdot nm^{-1}$. The NIR reflectance ranged from 32.80% to 40.31%. The spectral distinction between cultivar groups became clearer and more stable, with the 2nd leaf consistently exhibiting the highest NIR reflectance in most varieties (Figure 3).

3.2.4. Maturation Stage

Caption for Figure 4: Mean leaf spectral reflectance at the maturation stage. Senescence leads to increased visible reflectance and a slight decline in the NIR plateau. The effect of leaf position remains evident, with the 3rd leaf typically showing the highest reflectance. At maturity, reflectance in the visible region increased due to leaf senescence, while the NIR plateau slightly declined. The green peak difference was 5.37%. The red-edge average slope decreased to $0.37\%\cdot nm^{-1}$. The NIR reflectance range was 33.87%–41.99%. The spectral differences between varieties narrowed slightly, but the leaf position effect remained stable, with the 1st leaf typically maintaining the highest reflectance (Figure 4).

Overall, spectral reflectance exhibited characteristic changes associated with growth and senescence: a general decrease from seedling to tillering (increased chlorophyll), an increase during elongation (structural maturation), and a rise in visible reflectance at maturity (senescence). The 2nd leaf consistently showed less spectral variability across these transitions.)

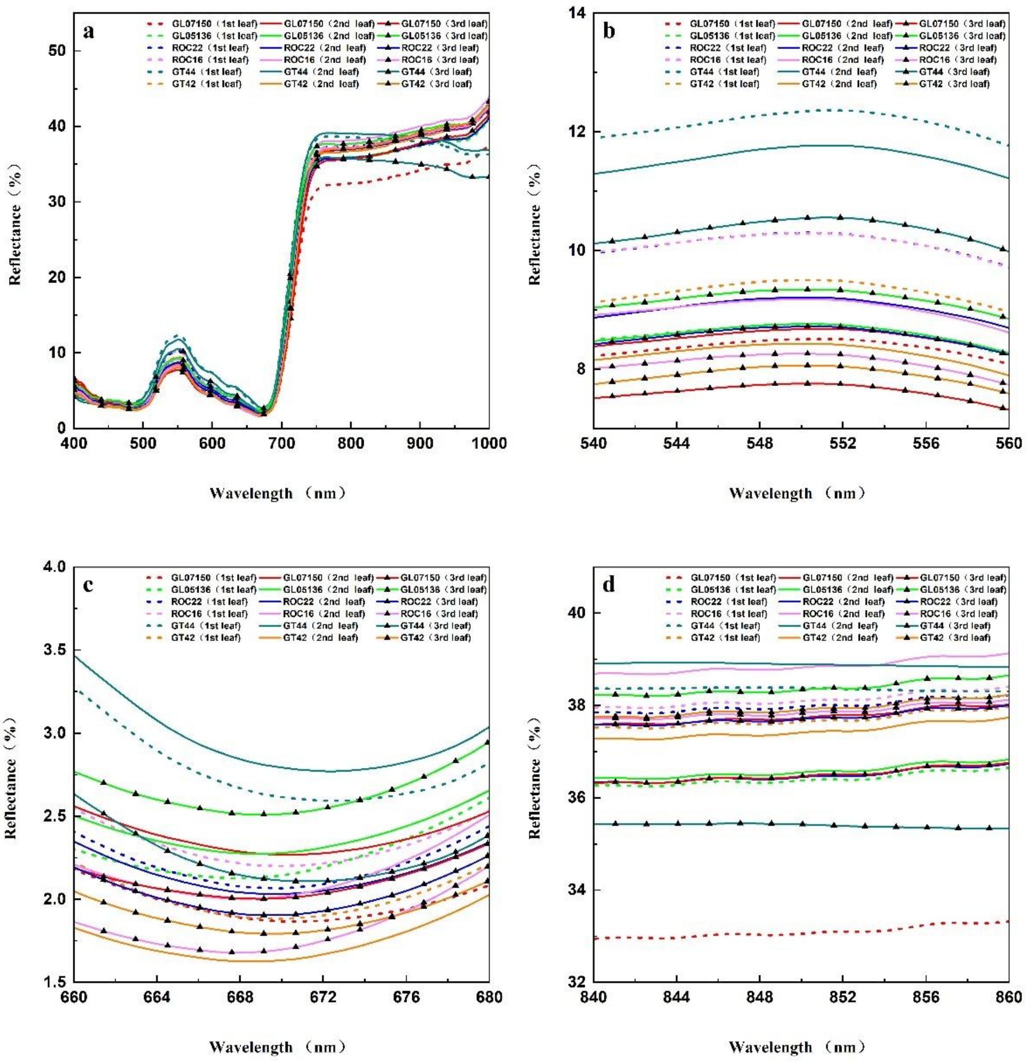

Figure 1 Leaf spectral reflectance of six varieties at the seedling stage. (a) Full spectrum (400–1000 nm): six spectral curves in different colors (one per cultivar), each displayed in three-line styles (representing the three leaf positions). Lightly shaded backgrounds are used to mark the “Visible,” “Red-edge,” and “NIR” regions.

(b) Enlarged view of the green peak (540–560 nm).

(c) Enlarged view of the red absorption trough (660–680 nm).

(d) Enlarged view of the NIR plateau (840–860 nm).

- - The dotted line represents the 1st leaf, — the solid line represents the 2nd leaf, and —▲— the solid line with triangle markers represents the 3rd leaf. Lines represent means (n = 3), and shaded areas indicate ±1 standard deviation.

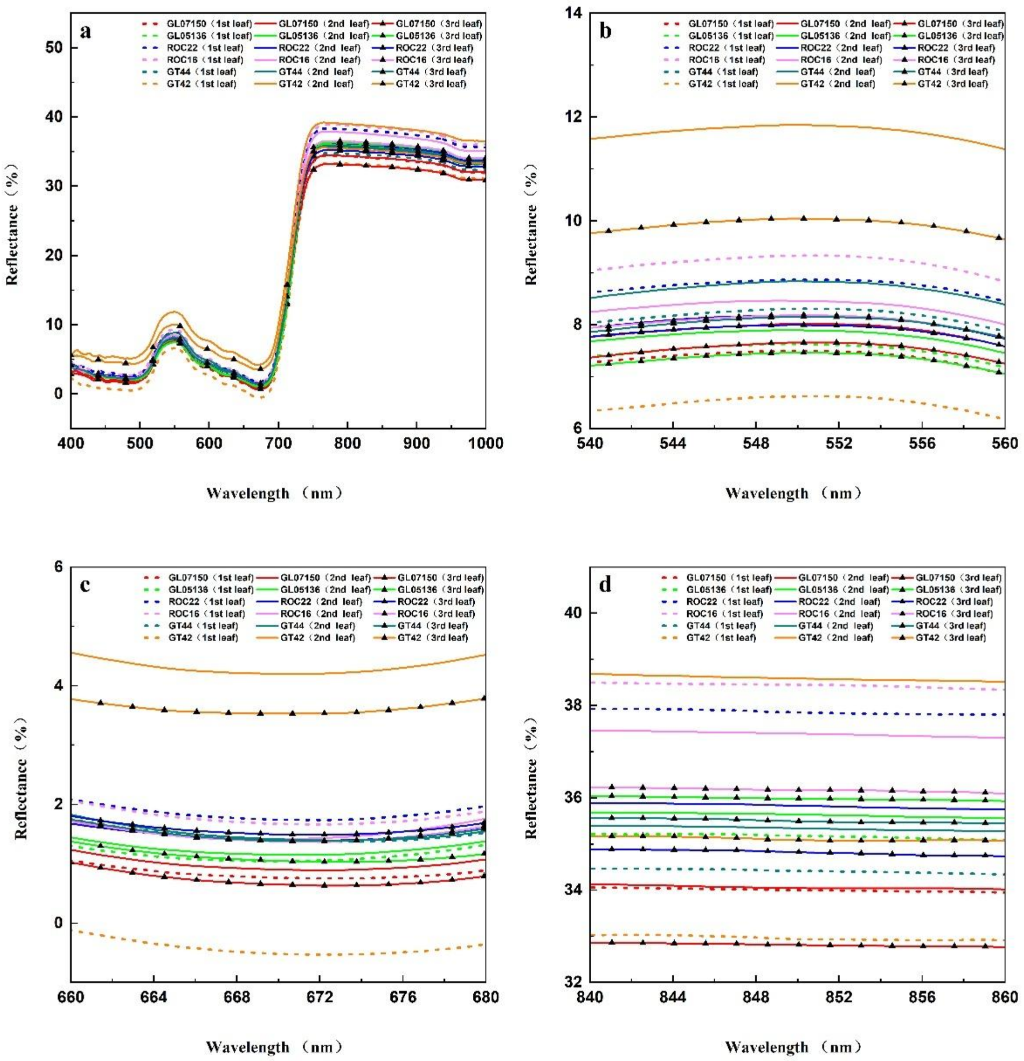

Figure 2 Leaf spectral reflectance of six varieties at the tillering stage. (a) Full spectrum (400–1000 nm): six spectral curves in different colors (one per cultivar), each displayed in three-line styles (representing the three leaf positions). Lightly shaded backgrounds are used to mark the “Visible,” “Red-edge,” and “NIR” regions.

(b) Enlarged view of the green peak (540–560 nm).

(c) Enlarged view of the red absorption trough (660–680 nm).

(d) Enlarged view of the NIR plateau (840–860 nm).

- - The dotted line represents the 1st leaf, — the solid line represents the 2nd leaf, and —▲— the solid line with triangle markers represents the 3rd leaf. Lines represent means (n = 3), and shaded areas indicate ±1 standard deviation.

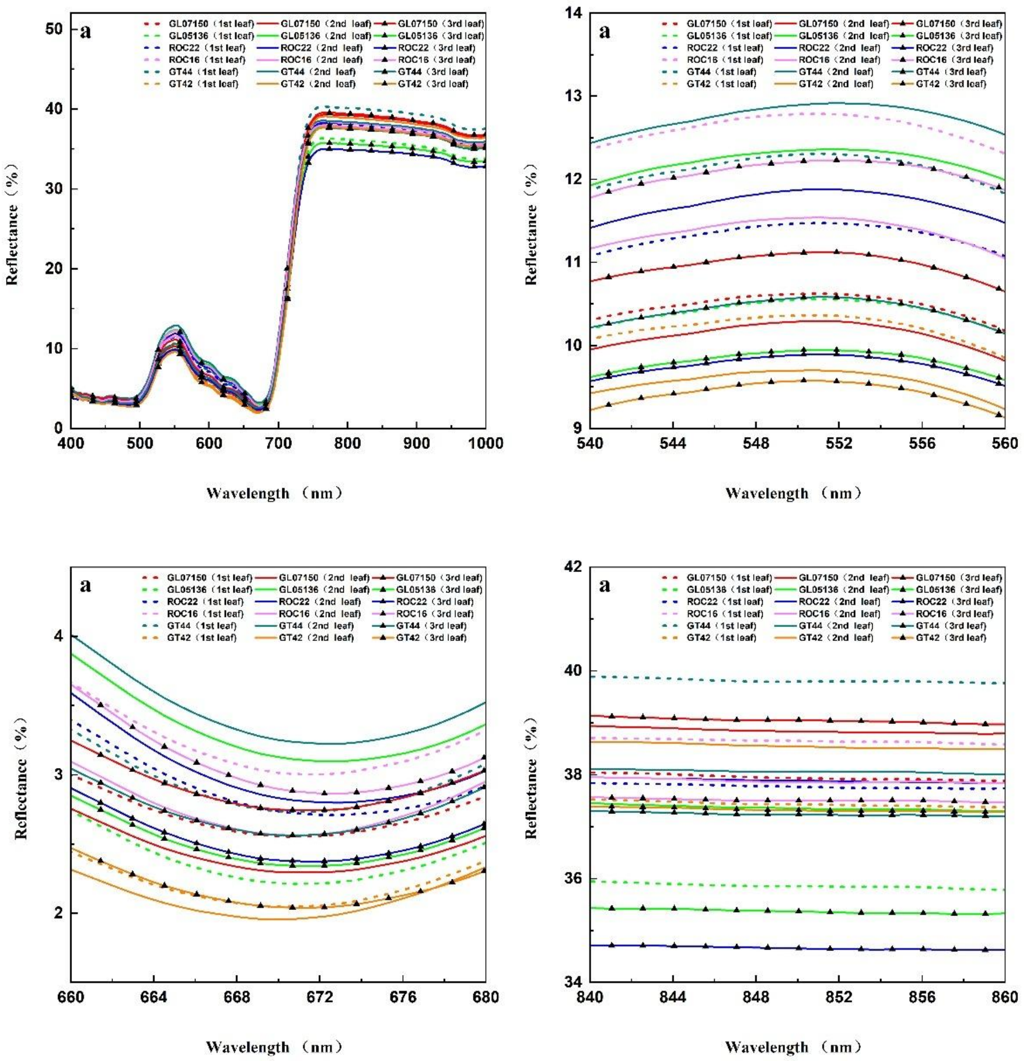

Figure 3 Leaf spectral reflectance of six varieties at the elongation period. (a) Full spectrum (400–1000 nm): six spectral curves in different colors (one per cultivar), each displayed in three line styles (representing the three leaf positions). Lightly shaded backgrounds are used to mark the “Visible,” “Red-edge,” and “NIR” regions.

(b) Enlarged view of the green peak (540–560 nm).

(c) Enlarged view of the red absorption trough (660–680 nm).

(d) Enlarged view of the NIR plateau (840–860 nm).

- - The dotted line represents the 1st leaf, — the solid line represents the 2nd leaf, and —▲— the solid line with triangle markers represents the 3rd leaf. Lines represent means (n = 3), and shaded areas indicate ±1 standard deviation.

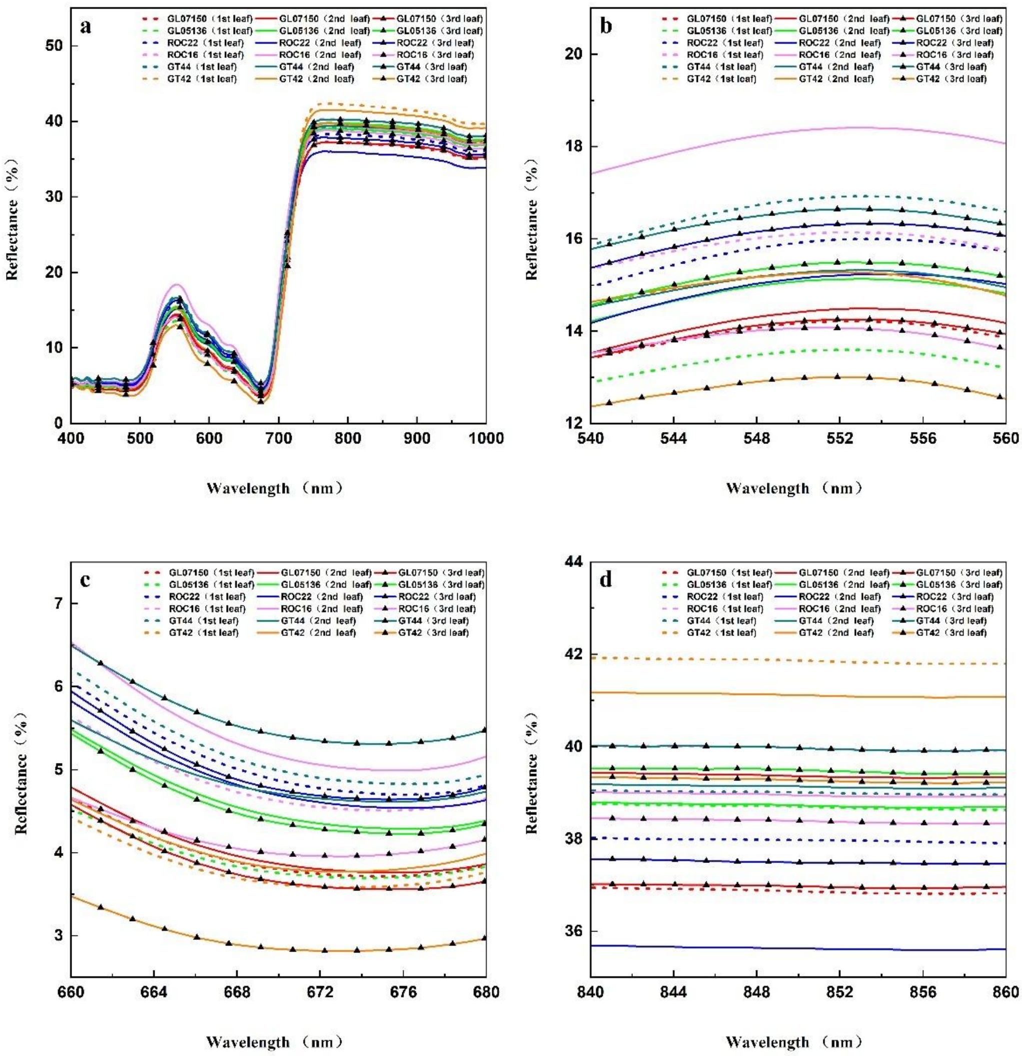

Figure 4 Leaf spectral reflectance of six varieties at the maturation stage. (a) Full spectrum (400–1000 nm): six spectral curves in different colors (one per cultivar), each displayed in three-line styles (representing the three leaf positions). Lightly shaded backgrounds are used to mark the “Visible,” “Red-edge,” and “NIR” regions.

(b) Enlarged view of the green peak (540–560 nm).

(c) Enlarged view of the red absorption trough (660–680 nm).

(d) Enlarged view of the NIR plateau (840–860 nm).

- - The dotted line represents the 1st leaf, — the solid line represents the 2nd leaf, and —▲— the solid line with triangle markers represents the 3rd leaf. Lines represent means (n = 3), and shaded areas indicate ±1 standard deviation.

3.3. Correlation Analysis Between Leaf Nitrogen Content and Spectral Data

The relationship between LNC and spectral reflectance (Original Reflectance, First-Derivative, and Log(1/R)) was analyzed across 108 samples.

3.3.1. Correlation with Original Reflectance (Figure 5a)

Caption for Figure 5a: Pearson correlation coefficients (r) between original spectral reflectance and leaf N content (LNC) for the six sugarcane varieties. GT42 shows the strongest negative correlation in the visible region, while GL05136 shows the strongest positive correlation in the NIR. ROC22 exhibits an atypical positive correlation in the blue region. Visible Region (400–680 nm): A significant negative correlation dominated this region for most varieties. GT42 showed the strongest negative correlation (r = -0.32 at 555 nm). ROC22 was an exception, showing a significant positive correlation in the short-wavelength visible region (r = 0.33 at 410 nm) before transitioning to a weak, non-significant negative correlation.

Red-Edge Region (680–750 nm): This region exhibited significant cultivar-specific differentiation. GT42 maintained a strong negative correlation (r = -0.35 at 709nm), while GL05136's correlation shifted from significantly negative to non-significant negative. Other varieties like GL05136 and ROC22 showed weak or non-significant correlations throughout most of the red edge.

NIR Region (750–1000 nm): A significant positive correlation was observed for most varieties. GL05136 showed the strongest and most stable positive correlation (r = 0.46 at 999 nm). In contrast, GT42 displayed a weak, non-significant positive correlation across the NIR region (Figure 5a).

3.3.2. Correlation with First-Derivative Spectra (Figure 5b)

Caption for Figure **5b**: Pearson correlation coefficients (r) between first-derivative spectra and LNC. The derivative transformation amplifies subtle features, producing sharp correlation peaks, particularly in the red-edge region (e.g., GT42, r = 0.67 at 760 nm), highlighting its sensitivity to LNC variation. The first-derivative transformation amplified subtle spectral features and produced sharper correlation peaks.

Visible Region: The correlations showed sharp "peak-and-trough" fluctuations. Most varieties exhibited strong negative correlation peaks (e.g., ROC22, r = -0.57 at 485nm). GT42 again displayed a unique positive correlation peak (r = 0.52 at 633 nm).

Red-Edge Region: This was the most sensitive region for the first-derivative, with all varieties showing sharp, strong positive correlation peaks. GT42 achieved the highest correlation coefficient in the entire study here (r = 0.67 at 760nm). ROC22 also showed a high correlation (r = 0.55 at 834nm).

NIR Region: The correlations were weaker than in the red-edge but stronger than the original reflectance in the NIR. Multiple weak positive correlation peaks were dispersed across this region (Figure 5b).

3.3.3. Correlation with Log (1/R) Spectra (Figure 5c)

Caption for Figure **5c**: Pearson correlation coefficients (r) between Log (1/R) (apparent absorbance) and LNC. This transformation results in smoother correlation curves, reducing high-frequency noise while maintaining the core negative (visible) and positive (NIR) trends observed in the original reflectance. The logarithmic transformation resulted in smoother correlation curves, reducing noise while retaining key information.

Visible Region: A smooth, significant negative correlation trend was observed. GL07150 showed the strongest correlation (r = -0.32 at 540 nm). ROC22's negative correlation in the green wavelengths was weaker (r = -0.09 at 550 nm) and non-significant.

Red Edge Region: The curves showed a smooth transition from negative to less negative/positive values, without the sharp peaks of the first derivative. GL07150 maintained the strongest negative correlation (r = -0.29 at 704nm).

NIR Region: A stable positive correlation was evident for most varieties, with very smooth curves. GT42 had the strongest correlation (r = 0.47 at 999 nm). GT44 remained the exception with a weak, non-significant correlation (Figure 5c).

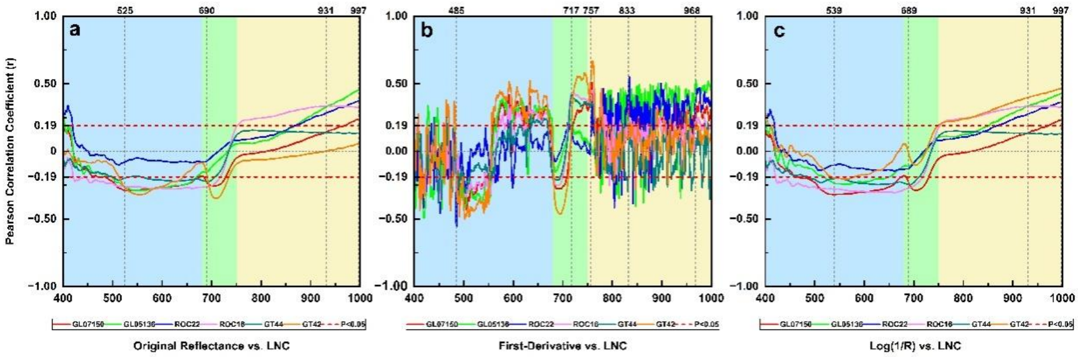

Figure 5 Correlation coefficient curves between the original spectral reflectance, first derivative of leaf spectra, logarithm of leaf spectra and nitrogen for six sugarcane varieties.

(a) Original Reflectance vs. LNC

(b) First-Derivative vs. LNC

(c) Log(1/R) vs. LNC

Note：The shaded background highlights spectral regions: visible (blue), red edge (green), and NIR (yellow). The dashed horizontal line at r = 0 indicates no correlation. The exceptional positive correlation in the visible region for ROC22 underscores the risk of using generic visible band indices (e.g., those based on green reflectance) across all cultivars.)

3.4. Model Inversion of Leaf Nitrogen Content

3.4.1. Characteristic Band Selection

Pearson correlation analysis can objectively identify the most informative spectral bands by quantifying the strength of the linear relationship between each wavelength and the target variable. This method avoids subjective selection bias and directly identifies key spectral regions based on data, providing highly correlated input variables for subsequent modeling. Characteristic bands for each cultivar and spectral type were selected based on the highest absolute Pearson correlation coefficients with LNC. The results showed that:

For **Original Reflectance**, the characteristic bands for GL05136, ROC22, and GT42 were identical and located in the far-NIR (997-1000 nm). ROC16's bands were clustered around 931-934 nm, while GT44's were in the red edge (690-693 nm).

For **First-Derivative** spectra, the selected bands were more dispersed across the spectrum, highlighting the ability of this transformation to identify unique, sensitive wavelengths for different varieties (e.g., ROC22: 485 & 834 nm; GT42: 720-725 nm).

For **Log (1/R)** spectra, the characteristic bands were highly similar or identical to those selected from the Original Reflectance, confirming the retention of core informational content while changing the representation (Table 2).

Table 2: Characteristic wavelengths selected for each cultivar and spectral transformation type, identified as the four bands with the highest absolute Pearson correlation coefficients (|r|) with Leaf Nitrogen Content.

Spectral Type	Cultivar	Characteristic Wavelengths (nm)			
Original Reflectance	GL07150	526.78	527.56	525.99	529.13
	GL05136	999.84	999.11	998.37	997.64
	ROC22	999.84	999.11	998.37	997.64
	ROC16	931.32	932.06	932.80	933.54
	GT44	692.22	690.68	692.98	691.45
	GT42	999.84	999.11	998.37	997.64
First Derivative	GL07150	758.69	757.93	759.45	452.55
	GL05136	968.25	968.99	821.50	993.97
	ROC22	485.02	834.30	485.81	833.55
	ROC16	759.45	760.21	757.93	757.17
	GT44	718.27	717.51	719.04	719.80
	GT42	721.33	725.15	724.39	720.56
Log (1/R)	GL07150	540.91	541.69	540.12	539.34
	GL05136	999.84	999.11	998.37	997.64
	ROC22	999.84	999.11	998.37	997.64
	ROC16	932.06	931.32	932.80	933.54
	GT44	690.68	691.45	692.22	689.91
	GT42	999.84	999.11	998.37	997.64

3.4.2. Model Construction and Validation

The performance of PLSR and Random Forest (RF) models, built using the characteristic bands from the three spectral types, is summarized in Table 3. **Optimal Models:** The combination of **FD spectra with the Random Forest algorithm** yielded the best prediction models for four out of six varieties (GLO7150, GL05136, ROC22, and GT42). For **ROC22**, this model achieved the highest testing set R^2 of **0.50** and an RMSE of 1.74, representing the best generalization performance among all models. For **GL05136**, the FD-RF model showed high training accuracy (R^2 = 0.84) but a marked drop in testing (R^2 = 0.17), suggesting potential overfitting or cultivar-specific spectral instability. For **GT42**, the "First-Derivative + RF" model achieved the lowest testing set RMSE (**1.45**), indicating high prediction accuracy. **Other Models:** For **ROC16** and **GT44**, the best model was the combination of **Log (1/R) spectra with RF**, although their overall performance was weaker (testing set R^2 of 0.18 and 0.08, respectively). **General Trends: Random Forest** consistently outperformed PLSR in most scenarios, demonstrating its superior capability in handling the non-linear relationships between spectra and LNC. **First-Derivative** was the most effective spectral preprocessing method for enhancing model performance, particularly when combined with RF (Table 3).

Models based on **Original Reflectance** generally performed the worst, highlighting the necessity of spectral transformations. Model performance was **cultivar-specific**, underscoring the importance of developing customized or grouped models rather than a single universal model for sugarcane.

Table 3 Performance metrics of the optimal Leaf Nitrogen Content prediction model for each sugarcane cultivar cultivar s of the optimal Leaf Nitrogen Content prediction model for each sugarcane

Cultivar	Optimal Model Combination	Training R^2	Training RMSE	Testing R^2	Testing RMSE	Performance Category
GL07150	FD + RF	0.74	1.12	**0.44**	1.54	Moderate
GL05136	FD + RF	0.84	1.07	0.17	1.84	Moderate
ROC22	FD + RF	0.77	1.23	**0.50**	1.74	Recommended
ROC16	Log (1/R) + RF	0.48	1.72	0.18	2.41	Needs Improvement
GT44	Log (1/R) + RF	0.61	1.12	0.08	2.17	Needs Improvement
GT42	FD + RF	0.83	1.06	0.34	**1.45**	Recommended

Note: RF denotes Random Forest. R^2: Coefficient of Determination; RMSE: Root Mean Square Error (mg/g). Testing set metrics in bold represent the optimal values across all cultivars.

4. DISCUSSION

This research provided a systematic, leaf-level investigation into the dynamic interplay of cultivar, growth stage, and leaf position on the hyperspectral estimation of leaf N content (LNC) in sugarcane. Our findings confirmed that the relationship between LNC and spectral reflectance is dynamically related the interactions among the three key factors: cultivar, growth stage and leaf position of sampling. The results indicated the effect of cultivar is superior to other two factors. The following discussion covered the core findings in the context of plant physiology and remote sensing, as compared with previous reports and elucidated the mechanisms behind model performance variations, and pointed out this study's limitations and future research needs.

4.1. Multi-Factor Interactions: Implications for the Precision Monitoring Paradigm

The most fundamental conclusion of this study is the statistically significant interactive effect of cultivar, growth stage, and leaf position on both LNC and spectral reflectance. This aligns with the results of precision agriculture that genotype-by-environment interactions (G×E) profoundly influence crop phenotypes, including spectral signatures (Fan et al., 2023; Yang et al., 2022). However, much of the existing hyperspectral researches were related to crop N, though acknowledged these factors, only considered to develop universal models that "average out" such variability or studied in isolation (Canicattì et al., 2025; Lin et al., 2022). Our work explicitly quantified the strength and nature of these three-way interactions at the leaf level in sugarcane, a crop lacked with such a detailed coverage. This quantification is crucial because it provided experimental evidence against the use of oversimplified generic models in commercial sugarcane production with genetic and temporal heterogeneity.

The physiological basis for these interactions is well grounded: **Cultivar differences** arise from genetically encoded variations in leaf anatomy (e.g., palisade mesophyll thickness, cellular complexity), biochemical composition (e.g., chlorophyll a/b ratio, epicuticular wax), and nitrogen use efficiency (NUE) (Li et al., 2025; Tross et al., 2024). **Growth stage effects** mirror the shifting source-sink relationships and N translocation patterns as sugarcane progresses from vegetative growth to maturation (Perlo et al., 2020; Wang et al., 2025). **Leaf position effects** directly are related to the spatial gradient of N within the plant, reflecting the distinct physiological roles of young, mature, and senescing leaves (Araki et al., 2025; Ning et al., 2025).

Within this complex interplay, our identification of the **2nd leaf as the most stable and optimal indicator for spectral monitoring** is a key, crop-specific insight. This finding resonates with the concept of a "diagnostic leaf" or "most recently fully expanded leaf" used in other crops like maize and cotton for nutrient status assessment (Chen et al., 2025; Tang et al., 2025). The 2nd2 leaf in sugarcane likely represents a physiological "sweet spot": it is a mature, photosynthetically active source leaf with N status being integrative of plant demand but less transient than the still-developing 1st leaf or the potentially remobilizing 3rd leaf. By experimentally demonstrating its spectral stability across stages and cultivars, this study provided a strong rationale for standardizing field sampling protocols on the 2nd leaf. Adopting this simple, evidence-based protocol can significantly reduce one major source of noise (leaf ontogenetic variation) in large-scale spectral monitoring, thereby improving the consistency and reliability of data collected for precision farming.

4.2. In-Depth Interpretation of Spectral Response Mechanisms: From Phenomenon to Underlying Principles

The observed correlation patterns between LNC and reflectance offer a window into the underlying biophysical mechanisms, while the noted exceptions underscore the complexity introduced by cultivar genetics. Understanding these mechanisms is not merely academic; it guides the rational selection of spectral features and sensors for operational use.

The pervasive **negative correlation in the visible region (400-680 nm)**, particularly around the chlorophyll absorption features (red ~670 nm, blue ~450 nm) and the green reflectance peak (~550 nm), is classic and expected. It is primarily driven by the strong coupling between N and chlorophyll synthesis. Increased N availability typically enhances chlorophyll concentration, leading to greater absorption of photosynthetically active radiation and lower reflectance, especially in the red and blue regions, with a concomitant damping of the green peak (Falcioni et al., 2023). The **notable exception of ROC22**, which showed a weak or even positive correlation in parts of the visible spectrum, is highly instructive, suggesting a potential decoupling of N allocation from chlorophyll production in this specific cultivar. Nitrogen may be preferentially partitioned into non-photosynthetic pools (e.g., soluble proteins, amino acids) or structural components, or its leaf surface properties (e.g., glaucousness) and dominates the reflectance signal in these shorter wavelengths (Li et al., 2022). This cultivar-specific behavior is a critical caution against the blind application of visible-band indices (like simple ratio or NDVI derivatives) across diverse genetic material. It underscores the need for sensor systems or algorithms capable of capturing a broader spectral range to accommodate such genetic divergence.

The **red edge region (680-750 nm)** emerged as the zone of greatest **cultivar-specific differentiation**. The effectiveness of the first-derivative transformation here highlights its power to resolve subtle features. The red edge is phenomenologically a region of rapid change in reflectance, governed by the combined effects of strong chlorophyll absorption in the red and high leaf internal scattering in the NIR. Its exact position (REP) and slope are sensitive to both chlorophyll concentration and leaf mesophyll structure (Li et al., 2025; Spafford et al., 2021). The sharp, high-correlation peaks in the FD spectra for cultivars like GT42 ($r=0.67$ at 760 nm) indicate that the rate of change of reflectance in this region is tightly linked to their LNC. The variation in these red edge responses among cultivars likely is related directly onto their genetic differences in leaf mass per area (LMA), palisade cell development, and chloroplast density. Therefore, the red edge is not just a sensitive region for detection of N but a diagnostic region for identifying *how* different cultivars express their N status optically. This has direct implications for sensor design: broad-band multispectral sensors with fixed red-edge channels may be suboptimal for sugarcane N monitoring across diverse varieties. Hyperspectral sensors or carefully tuned, cultivar-specific narrow-band indices are recommended to capture these genetic nuances for precise diagnosis.

The consistent **positive correlation in the NIR plateau (750-1000 nm)** is largely attributed to the influence of N on leaf internal structure. Nitrogen is a key component of proteins and enzymes involved in building cellular structures. Higher N often correlates with greater leaf thickness and more developed internal air spaces (spongy mesophyll), which enhance multiple scattering of NIR radiation, thereby increasing reflectance (Sukhova et al., 2025). However, this relationship is indirect and can be confounded. Leaf water content, which also strongly absorbs in specific NIR bands (e.g., ~970 nm), and dry matter content are co-varying factors (Maimaitiyiming et al., 2017). The strength of the NIR-LNC correlation may therefore be environment- and cultivar-dependent, explaining why some cultivars showed weaker NIR signals. For field applications, NIR-based indices may be more robust under stable hydration conditions or require correction for leaf water content to improve their reliability for N estimation.

4.3. Model Performance Analysis: Algorithmic Efficacy Meets Biological Complexity

The consistent outperformance of the **Random Forest (RF) algorithm over Partial Least Squares Regression (PLSR)** across most scenarios provided strong evidence that the relationship between the full hyperspectral feature set and LNC in sugarcane is **inherently non-linear**. PLSR is a robust linear method designed for high-dimensional, collinear data, but it assumes linear relationships (Tanzilli et al., 2025). RF, as an ensemble of decision trees, excels at capturing complex, non-linear interactions, thresholds, and hierarchical dependencies within the data without a priori assumptions, making it particularly suited for modeling intricate biological phenomena (Tan et al., 2025). Its built-in feature importance metrics (though not extensively used here) also offer avenues for interpretability. For precision agriculture applications, the choice of algorithm directly impacts prediction accuracy. Our results strongly advocate for the adoption of non-linear, machine-learning approaches like RF over traditional linear models in developing spectral prediction tools for sugarcane.

The **first derivative (FD) transformation** proved to be the most effective spectral preprocessing method. Its success lies in its dual action: it removes additive baseline offsets (e.g., caused by subtle changes in measurement geometry or leaf surface scattering) and amplifies areas of the spectrum where the reflectance curve changes slope (Tsai and Philpot, 1998). The slope changes are often associated with the absorption features of specific biochemicals. By highlighting the shape of the spectrum rather than its absolute magnitude, FD features are more robust to certain types of noise and are particularly adept at isolating features in overlapping regions like the red-edge, which explains the dominance of the "FD+RF" combination. This finding provided a clear preprocessing guideline for future spectral analysis workflows aimed at sugarcane N diagnosis.

A deep dive into the **cultivar-specific model performance** yields valuable methodological and biological insights. The strong performance of the FD-RF model for **ROC22** (Test $R^2 = 0.50$, the highest among all) suggests that for this cultivar, the N status is encoded in a clear and consistent manner within the shape-sensitive features of its spectrum (likely in the red-edge), and this relationship generalizes well to unseen data. Such cultivars are ideal candidates for the immediate development of operational spectral diagnostic tools.

In stark contrast, the case of **GL05136** reveals classic signs of overfitting (Training $R^2 = 0.84$ vs. Test $R^2 = 0.17$). A **methodological explanation** could be that the characteristic bands selected for GL05136, while perfectly correlated with LNC in the training set, were coincidentally aligned with spectral features sensitive to random noise or a confounding factor not represented in the test set. Alternatively, the training set for GL05136 might have an unrepresentative distribution of LNC values. A **physiological hypothesis** is more intriguing: GL05136 may possess a highly plastic or unstable "nitrogen-spectral phenotype." Its spectral response to nitrogen might be strongly moderated by another unmeasured, internally varying factor (e.g., diurnal carbohydrate status, hydraulic condition), making the relationship fundamentally noisier and less predictable. This underscores that model failure can be as informative as success, pointing to cultivars with more complex physiology that may require alternative sensing approaches or larger, more diverse training datasets. For precision management, this implies that a single modeling approach may not suit all cultivars; some may require more sophisticated sensor fusion (e.g., combining reflectance with fluorescence) or the integration of ancillary agronomic data.

The optimal model for **ROC16 and GT44** was "Log (1/R) + RF." The logarithmic transformation converts reflectance to apparent absorbance, which can linearize relationships with constituent concentrations according to the Beer-Lambert law and compress multiplicative scattering effects (Rawal et al., 2024). This preprocessing might be particularly beneficial for cultivars of which leaf surface properties (e.g., thick cuticle, wax) cause significant, variable scattering that obscures absorption features in the raw reflectance.

The overarching conclusion from the modeling exercise is unambiguous: A single, universal hyperspectral model for predicting LNC across all sugarcane cultivars is neither optimal nor realistic. Our results strongly advocate for a cultivar-specific or cultivar-group-specific modeling approach. Future precision agriculture tools for sugarcane should ideally incorporate a two-step process: firstly, cultivar identification (potentially via spectral or image analysis), followed by the application of the appropriate, pre-trained spectral prediction model for that cultivar. This 'precision phenotype' approach is the key to unlocking reliable, large-scale spectral N monitoring.

4.4. From Leaf-Level Insights to Field-Scale Operations: A Pathway for Precision Nitrogen Management

This study, while comprehensive within its controlled leaf-level scope, provides the essential groundwork for operational systems. The transition from controlled leaf measurements to field-scale precision management requires scale translation and system integration.

1. Bridging the Scale Gap: Our leaf-level models establish the fundamental "nitrogen-spectrum" relationships. The next critical step is translating this to the **canopy scale** relevant to UAV, aerial, or satellite sensing. This can be approached through: (i) **Empirical scaling studies** conducting concurrent leaf and canopy hyperspectral measurements across different canopy structures and planting densities. (ii) **Radiative transfer modeling (RTM)** using models like PROSAIL to physically link leaf optical properties (as characterized here) with canopy reflectance, allowing for mechanistic inversion and understanding of background/architecture effects (Féret et al., 2017). (iii) **End-to-end machine learning** trained directly on high-resolution UAV-based hyperspectral imagery of canopies with known LNC (using the 2nd leaf protocol), thereby understanding the complex scaling function directly from data.

2. Towards an Integrated Decision Support System: The goal is informed action. We envision a **hierarchical monitoring and decision framework**:
* **Zone-Level Scouting:** Use satellite or UAV multispectral imagery to identify field zones with potential N stress (e.g., low NDVI).
* **Diagnosis:** Within suspect zones, use portable spectrometers (equipped with leaf clips) to collect spectra from the 2nd leaf, applying the appropriate cultivar-specific FD-RF model developed in this study for accurate LNC diagnosis.
* **Prescription Generation:** Integrate the diagnosed LNC with **agronomic models**, such as critical N dilution curves for sugarcane or crop growth simulation models, to calculate the site-specific N requirement (Lemaire et al., 2008).
* **Precision Application:** Turn the N requirement into variable-rate application equipment to execute spatially optimized fertilization.

3. Limitations and Future Research Directions: To realize this pathway, future work must address:(i) **Multi-environment validation** across different soil types, climates, and management regimes to ensure model robustness. (ii) **Integration with other sensors** (e.g., thermal, LiDAR) to account for confounding factors like water stress and canopy structure. (iii) **Developing low-cost, rapid**

diagnostic tools embedding our key findings (e.g., smartphone-based spectroscopy targeting the 2nd leaf, coupled with cloud-based cultivar-specific model processing). (iv) **Economic and feasibility studies** to evaluate the cost-benefit ratio of implementing such a spectral-based precision N management system compared to conventional practices.

In conclusion, this study elucidates the complex, interactive nature of factors influencing the hyperspectral estimation of N in sugarcane and establishes a robust, cultivar-aware modeling framework centered on the 2nd leaf and FD-RF processing. These findings advance the fundamental understanding of crop spectral responses and provide concrete, actionable technical protocols—optimal sampling strategy, preprocessing method, and algorithm selection—tailored to sugarcane's genetic diversity. By pursuing the outlined research and development pathway, this work lays a solid foundation for translating hyperspectral remote sensing into a tangible, decision-ready technology for sustainable and precise nitrogen management in global sugarcane production.

5. CONCLUSION

In conclusion, this study systematically demonstrated that the hyperspectral estimation of LNC in sugarcane is profoundly governed by the complex interactions of cultivar, growth stage, and leaf position. We identified the 2nd leaf as the most stable and representative indicator for spectral monitoring. Distinct spectral response mechanisms were revealed: a negative correlation in the visible region (520–570 nm) driven by chlorophyll absorption, a positive correlation in the NIR (930–999 nm) linked to leaf structure, and a highly cultivar-specific response in the red-edge region (700–730 nm). For model development, the first-derivative transformation was the most effective preprocessing technique, and the Random Forest algorithm consistently outperformed linear methods, highlighting the non-linear nature of the nitrogen-spectrum relationship. Critically, the optimal model was cultivar-specific, with the "First-Derivative + Random Forest" combination providing the most robust predictions for most cultivars. These findings provide a critical theoretical foundation and a practical, optimized framework for advancing precision N management in sugarcane production via hyperspectral remote sensing.

Beyond these scientific insights, this research delivers clear practical guidance for advancing precision N management in sugarcane. **In practice, we recommend: (1) standardizing field sampling on the 2nd leaf; (2) applying first-derivative preprocessing; (3) building cultivar-specific Random Forest models. This framework can be integrated into spectral-based decision support systems for site-specific N recommendations in sugarcane.** The findings pave the way for developing next-generation, spectral-based decision support tools that can integrate cultivar information, optimize sensor configuration, and ultimately contribute to site-specific N recommendation systems. This progression is essential for improving N use efficiency, reducing environmental impact, and boosting the economic sustainability of sugarcane production.

Implications and Impacts: The findings pave the way for developing next-generation, spectral-based decision support tools that can integrate cultivar information, optimize sensor configuration, and ultimately contribute to site-specific N recommendation systems. And elucidates the complex, interactive nature of factors influencing the hyperspectral estimation of N in sugarcane and establishes a robust, cultivar-aware modeling framework centered on the 2nd leaf and FD-RF processing.

Author Contributions: G.H., Y.G., and M.M.: investigation, data curation. H.C. and S.L.: writing - original draft. S.H., S.L., and L.Z.: conceptualization, methodology. Z.L., S.H., and Z.H.: resources, validation. S.L.: funding acquisition. Z.L. and S.L.: supervision. H.C., Z.H., and Z.L.: writing - review & editing.

Funding:This work was supported by the Science and Technology Major Project of Guangxi (Guike AA22117006, Guike AA20302020-3, Guike AA17204037-4), Science and Technology Major Project of Chongzuo (FA2020006), and the support of Guangxi University and the Guangxi Academy of Science.

Conflict of interest: The authors declare that they have no conflict of interest.

Data Availability Statement: Data is available upon request from the authors.

REFERENCES

Ali, A., Jabeen, N., Farruhbek, R., Chachar, Z., Laghari, A. A., Chachar, S., et al. (2025). Enhancing nitrogen use efficiency in agriculture by integrating agronomic practices and genetic advances. *Frontiers in Plant Science*, *16*. https://doi.org/10.3389/fpls.2025.1543714

Araki, H., Ueno, H., Mizuta, K., Takahashi, T., Shibata, M., & Mano, J. (2025). Timing of nitrogen fertilization and shading affect the transition of nitrogen metabolism in senescing leaves of ripening wheat (Triticum aestivum L.). *Plant Production Science*, *28*(2), 176–185. https://doi.org/10.1080/1343943X.2025.2461776

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Burnett, A. C., Anderson, J., Davidson, K. J., Ely, K. S., Lamour, J., Li, Q., et al. (2021). A best-practice guide to predicting plant traits from leaf-level hyperspectral data using partial least squares regression. *Journal of Experimental Botany*, *72*(18), 6175–6189. https://doi.org/10.1093/jxb/erab295

Canicattì, M., Peng, J., Ciampitti, I., Vallone, M., & Cammarano, D. (2025). Spatiotemporal variability of nitrogen nutrition index in potato fields: A UAV-based machine learning approach using a Bayesian critical nitrogen dilution curve. *Computers and Electronics in Agriculture*, *237*, 110767. https://doi.org/10.1016/j.compag.2025.110767

Chegoonian, A. M., Singh, K. D., Geddes, C. M., Hansen, C., & Wang, H. (2025). Aerial Insights: Advancing Nitrogen Estimation in Field Crops using Multispectral Imaging. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLVIII-M-8–2025*, 7–12. Presented at the 46th Canadian Symposium on Rem ote Sensing (CSRS) "From Mountains to Kitchens: Remote Sensing Innovations for Water, Food & Security" - 16–19 June 2025, Lethbridge, Alberta, Canada. https://doi.org/10.5194/isprs-archives-XLVIII-M-8-2025-7-2025

Chen, S., Qin, S., Wang, Y., Ma, L., Lv, X., Chen, S., et al. (2025). Smartphone-Based Estimation of Cotton Leaf Nitrogen: A Learning Approach with Multi-Color Space Fusion. *Agronomy*, *15*(10). https://doi.org/10.3390/agronomy15102330

Chowdhury, M., Kumar Khura, T., Ahmad Parray, R., Kushwaha, H. L., Upadhyay, P. K., Jha, A., et al. (2024). The use of destructive and nondestructive techniques in concrete nitrogen assessment in plants. *Journal of Plant Nutrition*, *47*(14), 2271–2294. https://doi.org/10.1080/01904167.2024.2342353

Falcioni, R., Antunes, W. C., Demattê, J. A. M., Nanni, M. R., Falcioni, R., Antunes, W. C., et al. (2023). A Novel Method for Estimating Chlorophyll and Carotenoid Concentrations in Leaves: A Two Hyperspectral Sensor Approach. *Sensors*, *23*(8). https://doi.org/10.3390/s23083843

Fan, Y., Feng, H., Yue, J., Liu, Y., Jin, X., Xu, X., et al. (2023). Comparison of Different Dimensional Spectral Indices for Estimating Nitrogen Content of Potato Plants over Multiple Growth Periods. *Remote Sensing*, *15*(3). https://doi.org/10.3390/rs15030602

Féret, J. -B., Gitelson, A. A., Noble, S. D., & Jacquemoud, S. (2017). PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. *Remote Sensing of Environment*, *193*, 204–215. https://doi.org/10.1016/j.rse.2017.03.004

Féret, Jean-Baptiste, Berger, K., de Boissieu, F., & Malenovský, Z. (2021). PROSPECT-PRO for estimating content of nitrogen-containing leaf proteins and other carbon-based constitu ents. *Remote Sensing of Environment*, *252*, 112173. https://doi.org/10.1016/j.rse.2020.112173

Heinrichs, R., Otto, R., Magalhães, A., & Meirelles, G. C. (2017). Importance of Sugarcane in Brazilian and World Bioeconomy. In S. Dabbert, I. Lewandowski, J. Weiss, & A. Pyka (Eds.), *Knowledge-Driven Developments in the Bioeconomy: Technological and Economic Perspectives* (pp. 205–217). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-58374-7_11

Hou, Z., Tan, F., Li, M., Gao, J., Su, C., Jiao, F., et al. (2025). Quantitative Determination of Nitrogen Content in Cucumber Leaves Using Raman Spectroscopy and Multidimensional Feature Selection. *Agronomy*, *15*(8). https://doi.org/10.3390/agronomy15081884

Hu, Wenfeng, Tang, R., Li, C., Zhou, T., Chen, J., & Chen, K. (2021). Fractional order modeling and recognition of nitrogen content level of rubber tree foliage. *Journal of Near Infrared Spectroscopy*, *29*(1), 42–52. https://doi.org/10.1177/0967033520966693

Hu, Wenshi, Zhang, S., Huang, W., Chang, H., Li, Y., Gu, C., et al. (2025). Nitrogen-driven changes in leaf anatomy affect the coupling between leaf mass per area and photosynthesis. *Journal of Experimental Botany*, *76*(12), 3444–3456. https://doi.org/10.1093/jxb/eraf108

Kanash, E. V., Sinyavina, N. G., Rusakov, D. V., Egorova, K. V., Panova, G. G., Chesnokov, Y. V., et al. (2023). Morpho-Physiological, Chlorophyll Fluorescence, and Diffuse Reflectance Spectra Characteristics of Lettuce under the Main Macronutrient Deficiency. *Horticulturae*, *9*(11). https://doi.org/10.3390/horticulturae9111185

Lemaire, G., Jeuffroy, M.-H., & Gastal, F. (2008). Diagnosis tool for plant and crop N status i n vegetative stage: Theory and practices for crop N management. *European Journal of Agronomy*, *28*(4), 614–624. https://doi.org/10.1016/j.eja.2008.01.005

Li, D., Chen, J. M., Yan, Y., Zheng, H., Yao, X., Zhu, Y., et al. (2022). Estimating leaf nitrogen content by coupling a nitrogen allocation model with canopy reflectance. *Remote Sensing of Environment*, *283*, 113314. https://doi.org/10.1016/j.rse.2022.113314

Li, L., Geng, S., Lin, D., Su, G., Zhang, Y., Chang, L., et al. (2022). Accurate modeling of vertical leaf nitrogen distribution in summer maize using *in situ* leaf spectroscopy *via* CWT and PLS-based approaches. *European Journal of Agronomy*, *140*, 126607. https://doi.org/10.1016/j.eja.2022.126607

Li, Q., Song, X., Meng, X., Zhang, J., Zhang, M., Chen, L.-Y., et al. (2025a). Shaping future sugarcane: Ideal plant architecture and breeding strategies. *Molecular Plant*, *18*(5), 725–728. https://doi.org/10.1016/j.molp.2025.04.004

Li, Q., Song, X., Meng, X., Zhang, J., Zhang, M., Chen, L.-Y., et al. (2025b). Shaping future sugarcane: Ideal plant architecture and breeding strategies. *Molecular Plant*, *18*(5), 725–728. https://doi.org/10.1016/j.molp.2025.04.004

Li, Xuan, Zhu, B., Li, S., Liu, L., Song, K., Liu, J., et al. (2025). A Comprehensive Review of Crop Chlorophyll Mapping Using Remote Sensing Approaches: Achievements, Limitations, and Future Perspectives. *Sensors*, *25*(8). https://doi.org/10.3390/s25082345

Li, Xuyang, Shah, I. H., Gong, X., Azam, M., Jinhui, W., Li, P., et al. (2025). Research on High-Accuracy, Lightweight, Superfast Model for Nitrogen Diagnosis and Plant Growth in Lettuce (Lactuca sativa L.). *Horticulturae*, *11*(5). https://doi.org/10.3390/horticulturae11050451

Li, Yafeng, Xu, X., Wu, W., Zhu, Y., Yang, G., Gao, L., et al. (2024). Hyperspectral Estimation of Leaf Nitrogen Content in White Radish Based on Feature Selection and Integrated Learning. *Remote Sensing*, *16*(23). https://doi.org/10.3390/rs16234479

Li, Yumeng, Wang, C., Zhu, J., Wang, Q., Liu, P., Li, Y., et al. (2025). Classification of Nitrogen-Efficient Wheat Varieties Based on UAV Hyperspectral Remote Sensing. *Plants*, *14*(13). https://doi.org/10.3390/plants14131908

Lin, M.-Y., Lynch, V., Ma, D., Maki, H., Jin, J., Tuinstra, M., et al. (2022). Multi-Species Prediction of Physiological Traits with Hyperspectral Modeling. *Plants*, *11*(5). https://doi.org/10.3390/plants11050676

Liu, H., Gao, X., Fan, W., & Fu, X. (2025). Optimizing carbon and nitrogen metabolism in plants: From fundamental principles to practical applications. *Journal of Integrative Plant Biology*, *67*(6), 1447–1466. https://doi.org/10.1111/jipb.13919

Maimaitiyiming, M., Ghulam, A., Bozzolo, A., Wilkins, J. L., Kwasniewski, M. T., Maimaitiyiming, M., et al. (2017). Early Detection of Plant Physiological Responses to Different Levels of Water Stress Using Reflectance Spectroscopy. *Remote Sensing*, *9*(7). https://doi.org/10.3390/rs9070745

Meiyan, S., Jinyu, Z., Xiaohong, Y., Xiaohe, G., Baoguo, L., & Yuntao, M. (2023). A spectral decomposition method for estimating the leaf nitrogen status of maize by UAV-based hyperspectral imaging. *Computers and Electronics in Agriculture*, *212*, 108100. https://doi.org/10.1016/j.compag.2023.108100

Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*, *114*(4), 358–371. https://doi.o rg/10.1016/j.biosystemseng.2012.08.009

Ning, W., Su, L., Shi, D., Ji, M., Ouyang, X., Song, Q., et al. (2025). The near-optimal adjustment of carbon and nitrogen allocations into different organs in early-season rice cultivars with drastically different yield components under nitrogen application. *Frontiers in Plant Science*, *16*. https://doi.org/10.3389/fpls.2025.1537378

Perlo, V., Botha, F. C., Furtado, A., Hodgson-Kratky, K., & Henry, R. J. (2020). Metabolic changes in the developing sugarcane culm associated with high yield and early high sugar content. *Plant Direct*, *4*(11), e00276. https://doi.org/10.1002/pld3.276

Rawal, A., Hartemink, A., Zhang, Y., Wang, Y., Lankau, R. A., & Ruark, M. D. (2024). Visible and near-infrared spectroscopy predicted leaf nitrogen contents of potato varieties under different growth and management conditions. *Precision Agriculture*, *25*(2), 751–770. https://doi.org/10.1007/s11119-023-10091-z

Savitzky, Abraham., & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, *36*(8), 1627–1639. https://doi.org/10.1021/ac60214a047

Silva, C. A. A. C., Rizzo, R., Silva, M. A. da, Caron, M. L., Fiorio, P. R., Silva, C. A. A. C., et al. (2024). Spatio-Temporal Generalization of VIS-NIR-SWIR Spectral Models for Nitrogen Prediction in Sugarcane Leaves. *Remote Sensing*, *16*(22). https://doi.org/10.3390/rs16224250

Spafford, L., le Maire, G., MacDougall, A., de Boissieu, F., & Féret, J.-B. (2021). Spectral subdomains and prior estimation of leaf structure improves PROSPECT inversion on reflectance or transmittance alone. *Remote Sensing of Environment*, *252*, 112176. https://do i.org/10.1016/j.rse.2020.112176

Sparks, D. L., Page, A. L., Helmke, P. A., Loeppert, R. H., Soltanpour, P. N., Tabatabai, M. A., et al. (Eds.). (1996). Methods of Soil Analysis: Part 3 Chemical Methods (1st ed., Vol. 5). Wiley. https://doi.org/10.2136/sssabookser5.3

Stamford, J., Kasznicki, P., & Lawson, T. (2024). Spectral Reflectance Measurements. In S. Covshoff (Ed.), *Photosynthesis : Methods and Protocols* (pp. 333–353). New York, NY: Springer US. https://doi.org/10.1007/978-1-0716-3790-6_17

Sukhova, E., Yudina, L., Zolin, Y., Popova, A., Grebneva, K., Abasheva, K., et al. (2025). Light Scattering of Leaf Surface and Spongy Mesophyll and Concentration of Anthocyanin Influence Typical and Modified Photochemical Reflectance Indices. *Plants*, *14*(21). https://doi.org/10.3390/plants14213255

Tan, J., Jiang, L., Wei, Y., Yao, N., Zhao, G., Yu, Q., et al. (2025). Optimized Random Forest Framework for Integrating Cultivar, Environmental, and Phenological Interactions in Crop Yield Prediction. *Agronomy*, *15*(10). https://doi.org/10.3390/agronomy15102273

Tang, Y., Li, F., Hu, Y., & Yu, K. (2025). NTRI: A novel spectral index for developing a precise nitrogen diagnosis model across pre- and post-anthesis stages of maize plants. *Field Crops Research*, *325*, 109829. https://doi.org/10.1016/j.fcr.2025.109829

Tanzilli, D., Strani, L., Metz, M., Roger, J. M., Lesnoff, M., Ruckebusch, C., et al. (2025). Locally-weighted-RoBoost-PLS: a multivariate calibration approach to simultaneously cope with non-linearities and outliers. *Analytica Chimica Acta*, *1362*, 344167. https://doi.org/10.1016/j.aca.2025.344167

Tross, M. C., Grzybowski, M. W., Jubery, T. Z., Grove, R. J., Nishimwe, A. V., Torres-Rodriguez, J. V., et al. (2024). Data driven discovery and quantification of hyperspectral leaf reflectance phenotypes across a maize diversity panel. *The Plant Phenome Journal*, *7*(1), e20106. https://doi.org/10.1002/ppj2.20106

Tsai, F., & Philpot, W. (1998). Derivative Analysis of Hyperspectral Data. *Remote Sensing of Environment*, *66*(1), 41–51. https://doi.org/10.1016/S0034-4257(98)00032-7

Wang, Wanru, Zhang, L., Liu, S., Chen, M., Yang, X., Wang, W., et al. (2025). Time-Course Transcriptome Analysis Reveals Dynamic Nitrogen Response Mechanisms and Key Regulatory Networks in Sugarcane. *Agronomy*, *15*(9). https://doi.org/10.3390/agronomy15092164

Wang, Wei, Sun, N., Bai, B., Wu, H., Cheng, Y., Geng, H., et al. (2024). Prediction of wheat SPAD using integrated multispectral and support vector machines. *Frontiers in Plant Science*, *15*. https://doi.org/10.3389/fpls.2024.1405068

Yang, H., Ming, B., Nie, C., Xue, B., Xin, J., Lu, X., et al. (2022). Maize Canopy and Leaf Chlorophyll Content Assessment from Leaf Spectral Reflectance: Estimation and Uncertainty Analysis across Growth Stages and Vertical Distribution. *Remote Sensing*, *14*(9). https://doi.org/10.3390/rs14092115

Yang, L., Zhou, Y., Meng, B., Zhan, J., Xi, M., Deng, Y., et al. (2024). High sugarcane yield and large reduction in reactive nitrogen loss can be achieved by lowering nitrogen input. *Agriculture, Ecosystems & Environment*, *369*, 109032. https://doi.org/10.1016/j.agee.2024.109032

Yue, J., Wang, J., Zhang, Z., Li, C., Yang, H., Feng, H., & Guo, W. (2024). Estimating crop leaf area index and chlorophyll content using a deep learning-based hyperspectral analysis method. *Computers and Electronics in Agriculture*, *227*, 109653. https://doi.org/10.1016/j.compag.2024.109653

Zhang, J., Cheng, J., Liu, C., Wu, Q., Xiong, S., Yang, H., et al. (2024). Enhanced Crop Leaf Area Index Estimation via Random Forest Regression: Bayesian Optimization and Feature Selection Approach. *Remote Sensing*, *16*(21). https://doi.org/10.3390/rs16213917