



Research article

Estimating canopy nitrogen concentration of sugarcane crop using in situ spectroscopy

Aldemar Reyes-Trujillo^a, Martha C. Daza-Torres^a, Carlos A. Galindez-Jamioy^{b,*}, Esteban E. Rosero-García^c, Fernando Muñoz-Arboleda^d, Efrain Solarte-Rodriguez^b^a School of Environmental & Natural Resources Engineering, Universidad del Valle, Calle 13, No.100-00, Cali, Colombia^b Physics Department, Universidad del Valle, Calle 13, No.100-00, Cali, Colombia^c School of Electrical and Electronic Engineering, Universidad del Valle, Calle 13, No.100-00, Cali, Colombia^d Agronomy Program, CENICANA, Experimental Station, km 26 Cali-Florida, Colombia

ARTICLE INFO

Keywords:

Nitrogen
Sugarcane
Reflectance spectra
Vegetation indices
PLS regression

ABSTRACT

Estimating nitrogen (N) concentration in situ is fundamental for managing the fertilization of the sugarcane crop. The purpose of this work was to develop estimation models that explain how N varies over time as a function of three spectral data transformations in two stages (plant cane and first ratoon) under variable rates of N application. A randomized complete-block experimental design was applied, with four levels of N fertilization: 0, 80, 160, and 240 kg N ha⁻¹. Six sampling events were carried out during the rapid growth stage, where the canopy reflectance spectra with a hyperspectral sensor were measured, and tissue samples for N determination in plant cane and first ratoon were taken, from 60 days after emergence (DAE) and 60 days after harvest (DAH), respectively, until days 210 DAE and 210 DAH. To build the models, partial least squares regression analysis was used and was trained by three transformations of the spectral data: (i) average reflectance spectrum (R), (ii) multiple scatter correction and Savitzky-Golay filter MSC-SG) reflectance spectrum, and (iii) calculated vegetation indices (VIs).

1. Introduction

Worldwide, sugarcane (*Saccharum* spp.) has supplied on average 163.1 million tons of sugar and 88.9 million liters of ethanol in the past decade. Currently, a planted area of approximately 26 million hectares is reported worldwide, and according to CENICANA, (2018), producing one ton of millable cane requires an average of 1.1 kg of nitrogen (N), though this amount varies with the variety. However, the N fertilization of sugarcane is a challenge (Amaral et al., 2014), given that there are few early and reliable alternatives that allow the determination of the availability of this element in the soil (Amaral et al., 2015; Botero et al., 2009), especially in tropical regions, due to N cycle dynamics and soil variability, nor are there adequate methods to determine the availability of N in real time. Aiming to optimize this effort, the concentration of leaf N can be correlated with the spectral response, measured by optical spectroscopy techniques such as the diffuse reflectance of the crop canopy (Amaral et al., 2015; Andrade et al., 2015).

Regarding the application of optical spectroscopy methods to measure reflectance in the sugarcane crop, different studies have found

significant relationships with leaf N concentration under variable rates of N application. Robson et al. (2016) evaluated 15 varieties of sugarcane with satellite sensors and field radiometry at the canopy scale using six normalized difference vegetation indices (NDVIs). Amaral and Molin (2014) evaluated the canopies of four varieties with field radiometry and found significant relationships between the N application rate and three calculated VIs. Abdel-Rahman et al. (2013) worked with one variety of cane and hyperspectral data from a satellite sensor, applying algorithms to simplify their processing and prediction of leaf N concentration. Miphokasap et al. (2012) measured reflectance spectra in the field over the canopy, took the first derivative, calculated VI, and estimated models that significantly explained the N variation in three cane varieties. Portz et al. (2012) found, at three times in physiological development, that a field sensor could identify the variability of N consumption in fields planted with three varieties of sugarcane.

In general, each study reports experiments with different edaphoclimatic conditions, fertilization sources, N application rate and sugarcane varieties. However, sampling numbers have been small, at most three measurement times have been used for both spectral

* Corresponding author.

E-mail address: galindez.carlos@correounivalle.edu.co (C.A. Galindez-Jamioy).

responses and photochemical and biophysical characteristics, and their responses to different N fertilization rates have been calculated only during the vegetative cycle of the crop. This suggests that specific work is required to formulate standard measurement protocols that can determine the spectral response of a particular variety of sugarcane during different stages of its vegetative cycle. Doing so would allow more frequent monitoring, as in the case of experiments with variable N application rates, given the high variability of the forms of N available in the soil and the importance of N in the optimal development of the crop. Additionally, the methods to estimate the leaf N content are destructive, take a long time to generate results, and cause loss of plant material (Botero et al., 2009; Huber et al., 2008).

In the search for solutions to the above shortcomings, remote sensing has proven capable of estimating leaf N content based on spectral signatures. However, when it is done from satellite images, problems have come from the varied climatic conditions and the low resolution that they provide. Image resolutions obtained so far lack the precision required when addressing leaf N, due to the large distance from the leaves to the sensors and because these images can be affected by high cloudiness (Araque and Jiménez, 2009). The purpose of this research was to develop models to estimate the concentration of N in the canopy of the cane variety CC 01–1940 during the rapid growth stage from both lab-determined leaf N and the canopy spectral reflectance obtained in the field.

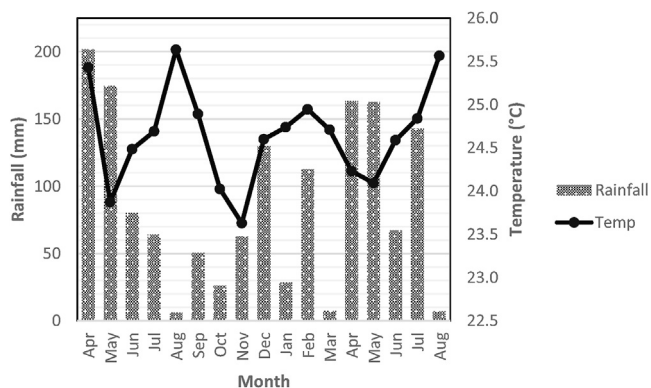


Figure 2. Total monthly mean rainfall and temperature during the two stages of the rapid growth stage of sugarcane variety CC 01–1940 in 2018–2019.

2. Materials and methods

2.1. Study site

The experiment was carried out from April 2018 to August 2019 at the Experimental Station of the Laboratory of Agricultural Waters and Soils (3°22'22.58" N; 76°31'47.57" W) of the Universidad del Valle headquarters, Meléndez, Cali, Colombia (Figure 1). The experimental site

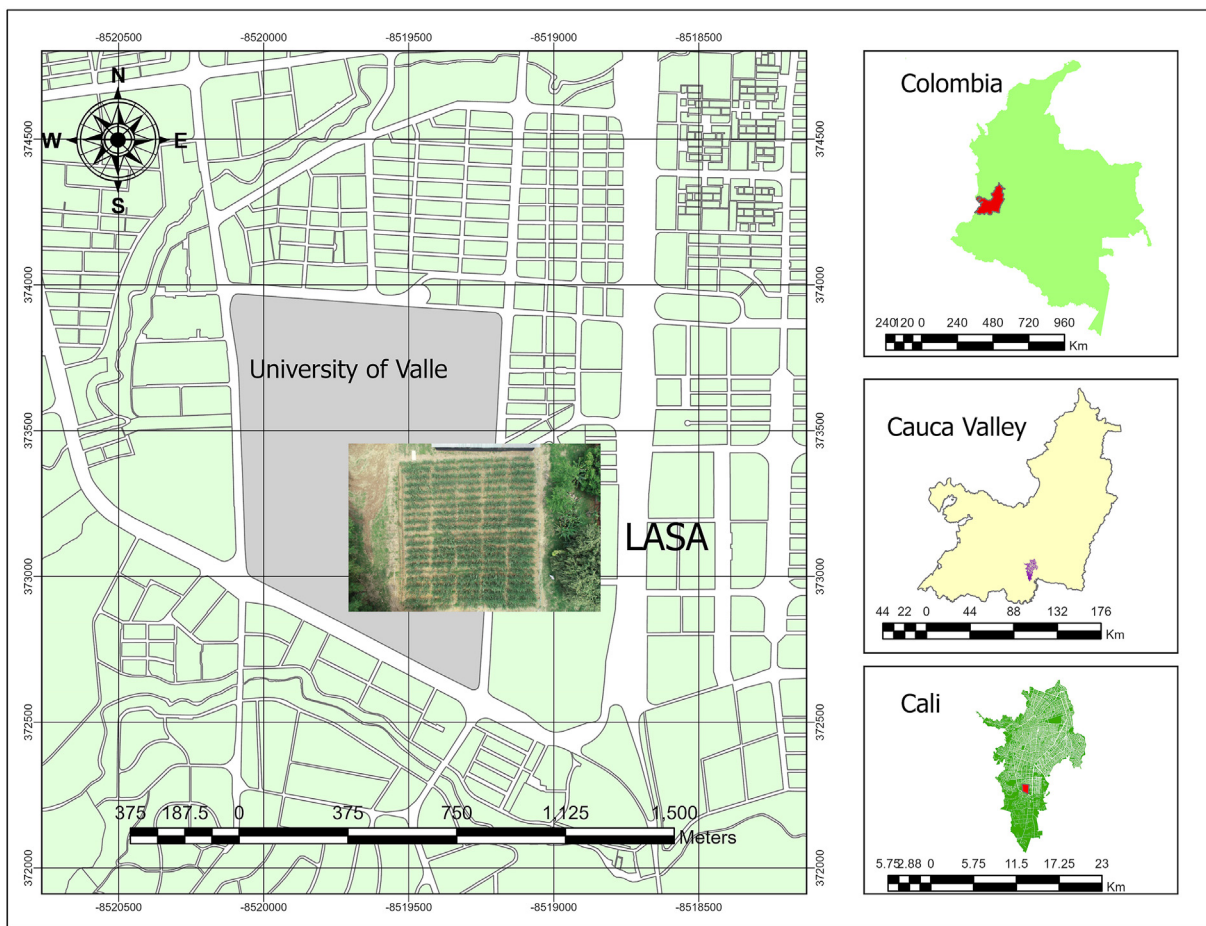


Figure 1. Location of the experiment: Experimental Station of the Agricultural Water and Soils Laboratory, Universidad del Valle.

was located at an elevation of 995 masl in a tropical dry forest climate, characterized by two rainy periods and two dry periods alternating throughout the year. Figure 2 shows the total monthly precipitation and the average temperature during two stages of the rapid growth stage of sugarcane. During the plant cane stage, there was a total precipitation of 575.4 mm and an average temperature of 24.1 °C, and during the first ratoon, there was a total precipitation of 523.8 mm and an average temperature of 24.6 °C. The predominant soil type was Udertic Haplus-toll, belonging to the Arroyo (AY) consociation; it had a slow internal and very slow external drainage and a moderately deep to deep effective root depth.

2.2. Experimental design

For the two stages, a randomized complete-block experimental design was applied with 5 repetitions (blocks), placed according to the organic matter content of the soil to minimize variability inside blocks, applying four different N rates (UAN-32 solution), for a total of 20 experimental units (EUs) each 33 m² in size, consisting of four rows 5 m long and spaced 1.65 m apart, planted with the variety CC 01–1940. To avoid edge effects between treatments, a row was left open between plots. The N fertilizer rates applied in the rapid growth stage were (i) 0 kg N ha⁻¹, (ii) 80 kg N ha⁻¹, (iii) 160 kg N ha⁻¹, and (iv) 240 kg N ha⁻¹. For the application of N as well as the other required primary elements, high-frequency localized fertigation was programmed according to the N application rate of each EU and according to the absorption rate (kg ha⁻¹·month⁻¹) of each proposed element (CENICANA, 2018). Other agricultural practices, including tillage, disease and pest management, and application of herbicides, were carried out according to the recommendations of CENICANA.

The plant cane was planted on April 20, 2018, placing seed at the bottom of the furrow at a rate of 7–10 buds per linear meter. In this stage, six monthly sampling events were conducted from 60 days after emergence (DAE) to 210 DAE. The first ratoon stage began on November 12, 2018, after cutting the plant cane. For this stage, six monthly sampling events were also carried out from 60 days after harvest (DAH) to 210 DAH. In the Cauca River valley, the cane stage is 12–13 months long, but this study only evaluated the rapid growth stage up to 7 months.

2.3. Data collection

2.3.1. Measurement of crop canopy reflectance spectra

To capture the reflectance spectra, a measurement system based on optical spectroscopy in the VIS range was developed. The system consisted of an aerial platform with mechanical and electronic conditioning to transport a STS:VIS microspectrometer (336–822 nm), with an optical resolution of 1.5 nm ± 0.05 nm and 1024 pixels, coupled to a PC and an optical opening control between 1° and 28°, which was located at a vertical distance of approximately 1 m above the canopy of the crop. It also includes a reference panel with Spectralon as the diffusion material and a reflectivity of 99% between 400 and 1500 nm to monitor changes in the reference irradiance and thus avoid saturation, eliminate the effect of changes in lighting conditions, and ensure that the measurements were comparable. The system also ensured independence from atmospheric effects since the aerial platform operated at low altitude.

The system was configured to calculate each reflectance spectrum with an integration time of 200 ms and an average of five readings per sample. In addition, it was programmed to take readings for 30 s over a distance of 5 m, equivalent to the length of the EU. The reflectance spectra were measured between 10:00 h and 15:00 h in both crop cycles and their respective sampling events.

2.3.2. Determination of N

After measuring the reflectance spectra, four plants per EU were randomly selected, in which the first leaf with a visible ligule was identified and cut from the main stem. Then, the middle third of each leaf

was dissected, and the central rib was removed to obtain one leaf tissue sample per EU. These were dried at 60 °C for 24 h and ground. The leaf N concentration was determined by Kjeldahl, performing predigestion for 24 h with H₂SO₄ and catalysts composed of CuSO₄ and K₂SO₄. Next, digestion was performed by subjecting the leaf sample to 250 °C for 30 min and then to 350 °C for 2 h until the samples turned translucent green. Then, distillation was performed with 40% NaOH, and ammoniacal N was collected into a solution of boric acid and indicator. The percentage of N was quantified using H₂SO₄ (0.005 normal) in the titration. The standard error of the lab method for N in the two stages is 0.05%.

Additionally, the dry mass of green leaves per area was quantified to determine the concentration of leaf N per unit area (N_f, in g·m⁻²) and the leaf area index (LAI, in m²·m⁻²) with LAI-2200C equipment in each EU. This was done to determine the concentration of N in the canopy (N_c, in g·m⁻²) (Eq. (1)) based on the fit of the equation used by (Lebourgeois et al., 2012), as follows:

$$N_c = N_f \times LAI \quad (1)$$

2.4. Treatment of reflectance spectra

The transformations of the spectral data were compared to select the best N_c estimation model based on the canopy scale reflectance spectra of the sugarcane crop: (i) average reflectance spectrum per EU (R), (ii) multiple-scatter correction (MSC) and a Savitzky-Golay (SG) filter reflectance spectrum, and (iii) vegetation indices (VIs) calculated from specific spectral bands.

2.4.1. Average reflectance spectrum

The data collection design yielded 10 spectra per row, for a total of 40 per EU and 800 per sampling event. An exploratory analysis was performed to exclude noisy wavelengths of a region (336–449 nm) from the data analysis. Atypical spectra were omitted from the reflectance data between 450 nm and 822 nm since they increase the error in the statistical analysis (Suarez et al., 2016). To equate the spectral data with the N_c values, the average was calculated for each EU.

2.4.2. Multiple-scatter correction and Savitzky-Golay filter

The basis of MSC is the fact that the dependence of light scattering on the wavelength is different from that of chemical-based light absorption. The dispersion in each sample is estimated in relation to that of an “ideal” sample, in this case the average, and the spectrum of each sample is corrected for this.

This version of MSC uses an ordinary least squares regression between the spectrum of sample X and the “ideal” spectrum to calculate the terms *a* and *b* and thus the corrected spectrum X₁ from the following equation (Eq. (2)):

$$X_1 = \frac{(X - a)}{b} \quad (2)$$

The SG filter consists of fitting a set of consecutive data to a polynomial and taking the center point of the adjusted polynomial curve as a new smoothed data point, with the advantage that it smooths the reflectance spectrum while preserving the characteristics of the initial distribution, the relative maxima and minima, and the width between peaks. The filter can be written as (Eq. (3)):

$$x_j^* = \frac{1}{N} \sum_{h=-k}^k C h x_{j+h} \quad (3)$$

where *x_j* is the new value, *N* is a normalization coefficient, *k* is the size of the space on each side of *j*, and *C* and *h* are previously calculated coefficients, which depend on the order selected (Stenvens and Ramirez-López, 2014).

2.4.3. Calculation of vegetation indices

The hyperspectral data obtained allowed VIs to be calculated at specific wavelengths, mostly as normalized differences. A total of 29 VIs were calculated (Table 1) for each of the 800 reflectance spectra per sampling event, and the average of each VI per EU was taken.

2.5. Data analysis

2.5.1. Partial least squares analysis

To determine the ability to estimate the Nc from the spectral data transformations, they were evaluated by partial least squares (PLS). This approach has the ability to treat many correlated independent variables (wavelengths) and relatively few observations, reduce them to a set of components while avoiding multicollinearity, and in particular estimate a set of dependent variables (Nc). In this case, to build the regression model, the matrix for the transformations R and MSC-SG is composed of 779 wavelengths (X) and the variation is made up of 29 VIs, (X) for each value of the dependent variable (Y) in each EU, and is expressed as follows (Eq. (4)):

$$Y = X * B + \epsilon \tag{4}$$

where B is the matrix of regression coefficients and ϵ is the matrix of residuals. The fit of the PLS model was evaluated by selecting an optimal number of components, which was determined by minimizing the root mean square error using leave-one-out cross-validation (RMSEcv), verifying the percentage of variance explained.

Additionally, to evaluate the performance of the model and compare between samplings, the variance explained (r^2), the relative root mean square error (RMSEr), and the residual prediction deviation (RPD) were used, which indicates the behavior of the precision of the fit compared to the average composition of all samples. High r^2 and low RMSEr indicate greater precision and accuracy of a model in estimating Nc (Li et al., 2016b). Simultaneously, the model is considered unacceptable, acceptable, or good depending on the range of r^2 and RPD values, as shown in Table 2 (Wang et al., 2013).

2.5.2. Selection of effective wavelengths

In the transformations proposed to estimate Nc, the relative importance of the wavelengths and the VIs in the PLS regression was evaluated by calculating the variable importance in projection (VIP) defined by Wold et al. (2001). The magnitude of the VIP indicates the contribution

of each wavelength or VI to the value of Nc estimated by the model, independent of the noise present in the information, and works well even when r^2 is low and RMSEr is high. VIP is calculated as (Eq. (5)):

$$VIP_k(a) = K \sum_a W^2 a k \left(\frac{SSY_a}{SSY_t} \right) \tag{5}$$

where $VIP_k(a)$ is the importance of wavelength or the VI in the k^{th} variable based on a model with a factors (PLS components), $W^2 a k$ is the weight of the wavelength or VI variable in the a^{th} factor of the PLS, SSY_a represents the sum of squares explained by Y by the PLS model with a factors, and SSY_t is the total sum of squares of Y explained by all a factors of the PLS model (Li et al., 2016a). According to (Steidle Neto et al., 2017), wavelengths or VIs with VIP values greater than 1 are highly influential in the estimation, between 0.8 and 1 moderately influential, and less than 0.8 insignificant in the estimation of Nc.

2.5.3. Purpose of the model

The objective was to independently evaluate the predictive behavior of the reflectance and VI spectra in each sampling event per stage to take advantage of the potential of the PLS regression and to generate information that would allow us to infer the temporal variability. Therefore, in each sampling event, 20 samples were used (four N treatments, five replicates, for a total of 120 samples per stage). The PLS regression and graphs were drawn using the PLS library run under the R (v.3.6.1) language.

3. Results and discussion

3.1. Evaluation of canopy N

The Nc varied over time in both stages (Figure 3), tending to increase up to 180 DAE, when it reached an average value of 27.39 $g \cdot m^{-2}$ (SD 2.20) in plant cane, while in first ratoon, Nc increased up to 150 DAH, with an average value of 26.29 $g \cdot m^{-2}$ (SD 1.63), and then decreased and stabilized up to 210 days at average values of 17.34 $g \cdot m^{-2}$ (SD 3.87) to 23.88 $g \cdot m^{-2}$ (SD 1.32), respectively. The increase was directly associated with the size and amount of leaf tissue, reflected in a LAI that showed continuous increases between emergence and sample 5. In this stage, the plant directs part of the absorbed element to leaf development. The subsequent decrease in Nc is attributed to N being allocated to development processes

Table 1. Calculated vegetation indices.

Vegetation index	Vegetation index
(1) Clrededge = (750/710) - 1	(16) CVI = (820*668)/(5602)
(2) NDVI _{750/650} = (750 - 650)/(750 + 650)	(17) CVIRE1 = (717*668)/(5602)
(3) NDVIg = (750 - 550)/(750 + 550)	(18) CVIRE2 = (820*717)/(5602)
(4) Clgreen = (820/550)-1	(19) CIRE1 = 717/(668-1)
(5) CCCI = ((820 - 717)/(820 + 717))/((820 - 668)/(820 + 668))	(20) CIRE2 = 820/(717-1)
(6) NDVI = (820 - 668)/(820 + 668)	(21) CIG = 820/(560-1)
(7) NDVIRE = (717-668)/(717 + 668)	(22) CIGRE = 717/(560-1)
(8) NDRE = (820-717)/(820 + 717)	(23) NGRDI = (560-668)/(560 + 668)
(9) GNDVI = (820-560)/(820 + 560)	(24) ENDVI = ((820 + 560)-(2*475))/((820 + 560)+(2*475))
(10) GNDRE = (717-560)/(717 + 560)	(25) ENDRE = ((717 + 560)-(2*475))/((717 + 560)+(2*475))
(11) BNDVI=(820-475)/(820 + 475)	(26) NDRE = ((780-730)/(780 + 730))
(12) BNDRE=(717-475)/(717 + 475)	(27) VOGRE = (740)/(720)
(13) EVI = 2.5*(820-668)/(820 + 6*668-7.5*475 + 1)	(28) DVI = ((780-668)
(14) EVIRE1 = 2.5*(717-668)/(717 + 6*668-7.5*475 + 1)	(29) MNDVI8=((755-730)/(755 + 730))
(15) EVIRE2 = 2.5*(820-717)/(820 + 6*717-7.5*475 + 1)	

(Adapted from: Henrich et al., 2012)

Note: The number in parentheses corresponds to the order assigned in the correlation and regression analysis.

Table 2. Model performance according to R² and RPD.

Statistical methods	Model Performance		
	Unacceptable	Acceptable	Good
r ²	<0.50	0.50–0.75	>0.75
RPD*	<1.40	1.40–2.00	>2.00

* RPD = (SD)/RMSE.

associated with the early maturation stage, where stalk length and thickness increase (CENICANA, 2018).

Although N reached its maximum later in the plant cane than in first ratoon, the increase in both stages was approximately 25 g·m⁻² when comparing the maximum Nc found at 60 DAE and 60 DAH. This was because in the plant cane, the plant must begin to use N for the formation of organs such as roots and the elongation of the first stems, while in first ratoon, the plant has already developed roots and is better adapted to the edaphic conditions; therefore, N is translocated to the upper leaves, as evidenced by the increase in the LAI in this rapid growth stage.

Figure 3 shows that the behavior of the median and the variance, per sample in each stage, changes as a function of the growth stage, with less variation in the first two seasons, according to the interquartile range represented in the box. Therefore, to infer the effect of N on crop development and on the optical properties at the canopy level, a total of 36 models were generated to estimate Nc, including the six sampling events per stage and the three transformations of the spectral data.

3.2. Canopy reflective spectra and influence of the leaf N concentration

The spectral response for the cane variety CC 01–1940 in the two stages evaluated is presented in Figure 4. Reflectance corresponded to the average of 1200 spectra per treatment and a total of 4800 per stage. In plant cane, the reflectance values were lower due to less leaf development in this stage and to its characteristic erectophilic leaf arrangement in the upper canopy. In addition, the N treatments generated differences in the near-infrared (NIR) plateau (750–822 nm), where the reflectance tended to decrease with decreasing N application. The spectral response

in first ratoon had higher reflectance values due to the greater leaf density in this stage, and there was no observed effect of the treatments because of the complex dynamics of the element in the soil type and the N availability, due to factors such as organic matter mineralization and fixation by microorganisms (Franco et al., 2011; Schultz et al., 2012). The effect observed in the NIR plateau is influenced by the canopy structure (LAI, fractional vegetation cover and leaf structure) (Lebourgeois et al., 2012). Plants respond to N supply through changes in chlorophyll concentration, biomass production, and leaf area index (Katsoulas et al., 2016), which alters the canopy reflectance. In the VIS region, the effect is not as evident given the combination of biomass concentration and chlorophyll (Hansen and Schjoerring, 2003).

3.3. Correlation between Nc and reflectance spectra

The spectral transformations were correlated with Nc using the complete dataset. The trend in the first three sampling events in each stage was that Nc had a negative correlation with R in the VIS wavelengths and a positive correlation in the NIR (Fig. 5a, d), in agreement with (Abdel-rahman et al., 2010; Li et al., 2016b), except in sample 2 (90 DAE) in plant cane, which showed indirect and significant relationships across the spectrum. The correlation curve identified regions of more significant activity at 450–500 nm, 590–640 nm, and approximately 700 nm, regions that coincide with the absorption peaks and that are activated by changes in the concentrations and quality of pigments, associated with the supply of N. In later samples, the relationship was positive in both VIS and NIR regions, with significant activity at 470, 481, and 650 nm in first ratoon and at 750–821 nm in plant cane, except in sample 6 (210 DAE) of first ratoon, where the correlation was inverted in the NIR region but was not significant (Figure 5d). Therefore, the first samples presented good regression models, given the statistical significance and differences greater than 0.5 between the maximum and minimum correlation values. In addition, the influence of crop age is evident in its effects on leaf chemistry and canopy architecture, which generate changes in the relationships between variables and the spectral response, a relevant aspect for guiding the monitoring and interpretation of the temporal dynamics of the crop's spectral response.

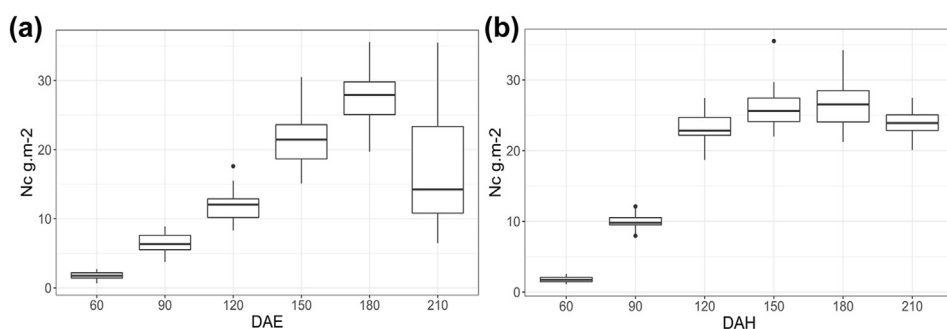


Figure 3. Temporal dynamics of total canopy N content. (a) Plant cane. (b) First ratoon.

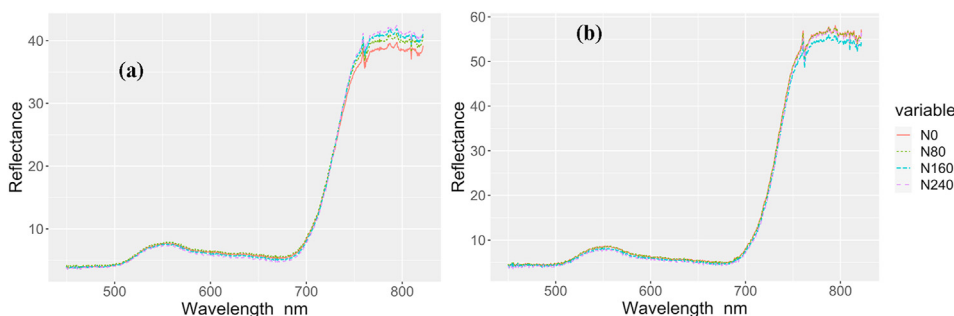


Figure 4. Canopy reflectance spectra per treatment. (a) Plant cane. (b) First ratoon.

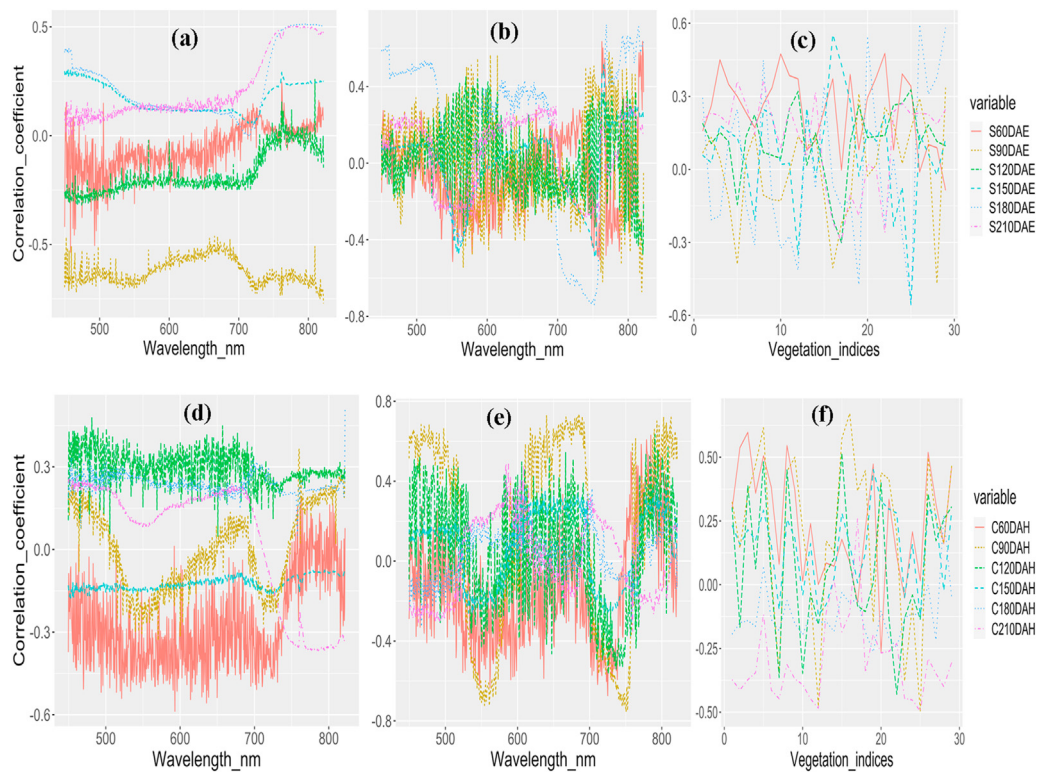


Figure 5. Temporal variability and correlation curve between Nc and reflectance spectra transformations. (a, b, c) Plant cane (d, e, f) First ratoon.

MSC-SG smoothed the correlation graph, facilitating the identification of active areas, without losing the variations of the original data (Barbin et al., 2012) or its relationship with Nc in each sample, as was the aim in this study. The first samples had significant activity in the VIS and NIR regions. For the plant cane stage, the regions 550–600 nm and 750–821 nm stood out, and only the sample from 180 DAE showed significance (Figure 5b). In first ratoon, in the first sample, regions of significant activity occurred at approximately 466, 550, 627, 680, 730, 750, and 804 nm (Figure 5e).

When relating Nc to VI (Fig. 5c, f), it was found that the bioindicators involving the red-edge region were the most significant in both stages evaluated, independent of the sampling event. This transition region is generated by the strong absorption of pigments and the increase in multiple scattering of the canopy (Zhu et al., 2014), the latter when the irradiance is quantified under field conditions and the light source is the sun. In addition, this region is sensitive to the effects of N on the concentrations of pigments, such as chlorophyll and carotenes. By relating this result with those obtained in the previous transformations, the wavelength values in the red edge could be used as significant markers, which better adjust with bioindicators in the different sampling events (VOGRE, NDRE, ENDRE, CIRE2), as in the case of 717 nm–700 nm or 750 nm, which had greater activity in sugarcane.

3.4. Performance of PLS models

The PLS model was proposed with Nc ($\text{g}\cdot\text{m}^{-2}$) as a dependent variable and the transformations of the spectral data (R, MSC-SG, VI) as independent variables. Eighteen models per stage were analyzed to determine the best time to collect data in the field. The optimal number of components (ONC) to include in each model was determined according to the minimum value of the RMSEcv and the r^2 , which is important for omitting excess components, which would lead to an overfitting problem. Tables 3 and 4 show that the ONC varied little with the number of components selected between samples within each transformation, but there was a difference when there were more than six components

between the transformation of VI and the other two in both stages. Given this high variation between samples of each transformation, it was necessary to establish a balance between RMSEcv, the variance explained, and the number of components (Li et al., 2016a).

Once the ONC was selected, the dataset was used to train the model and quantify the Nc estimate. The results of the PLS regression for the plant cane stage showed similar performance in samples 1 and 3 between the R transformation and MSC-SG (Table 3), presenting values of r^2 equal to or greater than 0.92 and RMSEr less than 5%, and the RPD classified the models obtained as those with the best performance. Transformation VI yielded good performance statistics in samples 1 and 2, according to the ranges proposed by (Wang et al., 2013) (Table 3). Additionally, the results show, independent of the transformation method, that the first three sampling times would be the best times to collect field data, in particular sample 1 (60 DAE) in plant cane, which had r^2 values of 0.98, 0.99, and 0.97 for R, SG-MS, and VI, respectively. In first ratoon, the best sampling is at 90 DAH, with r^2 values of 0.98 and 0.99 for R and SG-MS, respectively. The above generally shows, for both samples, that the percentage of variation between the adjusted and the observed data (RMSEr) did not exceed 5%. This is consistent with studies conducted by (Amaral et al., 2015; Portz et al., 2013), who found that the use of canopy reflectance sensors showed better correlations in the early stages of sugarcane crop growth.

In first ratoon (Table 4), in terms of model performance for selecting the best time to estimate the Nc, the R and MSC-SG transformations results that coincided with those of the plant cane stage in the first two time points, with a variation in the predicted values less than 5% and variances explained greater than 95%. Additionally, the performance of the models in sample 5 (180DDC) improved under all transformation methods, which presented r^2 values equal to or greater than 0.75 and RPD >2.0 . In particular, the performance of the VI transformation improved, with an RDP of 4.06 and a variation in the adjusted values of 3.38%, as shown in Table 4. This improvement was attributed to using bioindicators derived from the red-edge region, which mitigated saturation when used in crops with high biomass production, such as

Table 3. Performance of the PLS models between hyperspectral and Nc data in plant cane.

R						MSC-SG					IV				
	DAE	ONC	RMSEcv	Y-r ²	RMSEr (%)	RPD	ONC	RMSEcv	Y-r ²	RMSEr (%)	RPD	ONC	RMSEcv	Y-r ²	RMSEr (%)
60	8	0,40	0,98	3,80	6,65	7	0,39	0,99	2,67	9,50	12	0,36	0,97	4,52	5,57
90	5	0,87	0,85	7,98	2,67	4	1,19	0,63	12,56	1,69	9	1,22	0,81	9,03	2,35
120	7	2,02	0,92	4,49	3,68	7	1,89	0,94	3,86	4,29	3	2,16	0,18	14,66	1,13
150	4	3,20	0,55	11,27	1,53	4	3,40	0,55	11,22	1,53	2	4,61	0,19	15,07	1,14
180	4	2,84	0,74	6,83	2,02	4	2,92	0,73	7,03	1,96	1	3,82	0,15	12,35	1,11
210	1	7,82	0,23	36,51	1,17	1	8,77	0,06	40,42	1,06	1	9,72	0,06	40,44	1,06

DAE: days after emergence. ONC: optimal number of components. RMSE: root mean square error. cv: cross-validation. r²: variance explained. RPD: residual prediction deviation.

Table 4. Performance of the PLS models between hyperspectral and Nc data in first ratoon.

R						MSC-SG					IV				
	DAH	ONC	RMSEcv	Y-r ²	RMSEr (%)	RPD	ONC	RMSEcv	Y-r ²	RMSEr (%)	RPD	ONC	RMSEcv	Y-r ²	RMSEr (%)
60	7	0,35	0,96	4,35	5,32	6	0,34	0,97	3,79	6,11	1	0,38	0,23	19,80	1,17
90	7	0,72	0,98	1,16	9,55	7	0,71	0,99	0,99	11,20	2	1,12	0,35	8,66	1,28
120	2	2,43	0,24	8,54	1,18	1	2,28	0,29	8,29	1,21	2	2,47	0,08	9,43	1,07
150	4	3,41	0,24	9,79	1,17	2	3,35	0,11	10,56	1,09	6	2,89	0,72	5,94	1,93
180	7	4,00	0,81	5,75	2,38	6	3,36	0,75	6,63	2,07	12	3,53	0,94	3,38	4,06
210	4	1,83	0,46	5,66	1,39	1	2,06	0,05	7,51	1,05	9	1,94	0,80	3,42	2,32

DAH: days after harvest. ONC: optimal number of components. RMSE: root mean square error. cv: cross-validation. r²: variance explained. RPD: residual prediction deviation.

sugarcane. These bioindicators are sensitive to wide ranges of N content, photosynthetic pigments, leaf area, and biomass (Darvishzadeh et al., 2008; le Maire et al., 2008; Mutanga and Skidmore, 2007; Xiuhua et al., 2015).

Building PLS models considering Nc and canopy-scale reflectance will improve model performance due to the strong influence of N fertilization on the development of the leaf component of the canopy, and in turn on the spectral response, as found by Lebourgeois et al. (2012), Lofton et al. (2012), who even quantified it from VIs of specific bands. In addition, having well-fitting models in the early stages is essential for managing fertilization operations at appropriate times, both for the fertilizer's effect on the vegetative development of the crop and the efficiency of the fertilization.

We also observed that the models obtained with the R method offered good performance in samples 1, 2, and 3 and acceptable performance in samples 4 and 5. Similar results regarding the use of raw spectral data have been reported by (Stevens and Ramirez-López, 2014; Suarez et al., 2016). This implies that the transformations that reduce noise associated with variations in the intensity of solar radiation, high cloud cover, in the angle of incidence, and topographic effects (Tewari et al., 2008) do not significantly improve the performance of PLS models in the early stages of sugarcane crop development.

3.5. Interpretation of the PLS model

The contribution of each wavelength to the model performance was analyzed by observing the behavior of the regression coefficients of the models (Barbin et al., 2012). Figure 6 shows the results for the models with better performance, and in both stages, activity was observed in the blue region between 455 and 470 nm, characterized by the high absorption of radiation associated with concentrated pigments and providing a measure of the stress caused by their degradation and the consequent detection of N (Basyouni and Dunn, 2013; Katsoulas et al., 2016). Another region of high contribution to the model is in the red-edge region, with wavelengths at approximately 725, 750, 760, 780, and 797 nm, which is a representative region of photosynthetic capacity

and physiological stress because of the maximum sensitivity of absorption and reflectance in this region, allowing an accurate evaluation of crop biochemistry (Cordon, 2009; Zhao et al., 2007). In the NIR, the contribution of 820 nm was significant in both stages. In general, the coefficients of sample 1 (60 DAE) in both stages showed the same active zones but with less contributions in the magnitudes of their coefficients, which we attributed to the lower concentrations of pigments and N in this first stage of physiological development in the plant. The contribution of the coefficients coincided with the active regions identified in the correlation analysis, confirming that variations in the amplitude of the correlation graph (Figure 5) indicate the probability of good model performance. For VI, the greatest contributing index was VOGRE (data not shown), which is reported to be highly correlated with the leaf concentrations of N and chlorophyll (Bulcock and Jewitt, 2010). To calculate this index, the red-edge region and NIR wavelengths are used, which in the case of sugarcane can be adjusted due to the high influence of these values found in these regions.

3.6. Identification of effective wavelengths by VIP

The VIP graph (Figure 7) identifies the wavelengths with the greatest importance in the evaluation of Nc. According to their magnitudes, the blue- and red-edge regions had wavelengths with local maxima similar to those identified from the contributions of the coefficients. In both stages, it was common for the red-edge region to present a higher density of local maxima (719, 722, 725, 730, 758, and 760 nm), with some sights in the blue (450, 459, and 462 nm), green (505 and 534 nm), red (671 nm), and NIR (approximately 809 and 820 nm). Some of these bands were identified as making high contributions according to the regression coefficients of the best-performing models. An interesting result is the high influence found by the blue region, which is not commonly associated with plant health or physiological status (Robles et al., 2010; Suarez et al., 2016), indicating the potential value of including it in the adjustment of bioindicators can monitor environmental stress by N. In general, the VIP of the samples with good-performing models were represented by the same active regions and with similar contributions in magnitude,

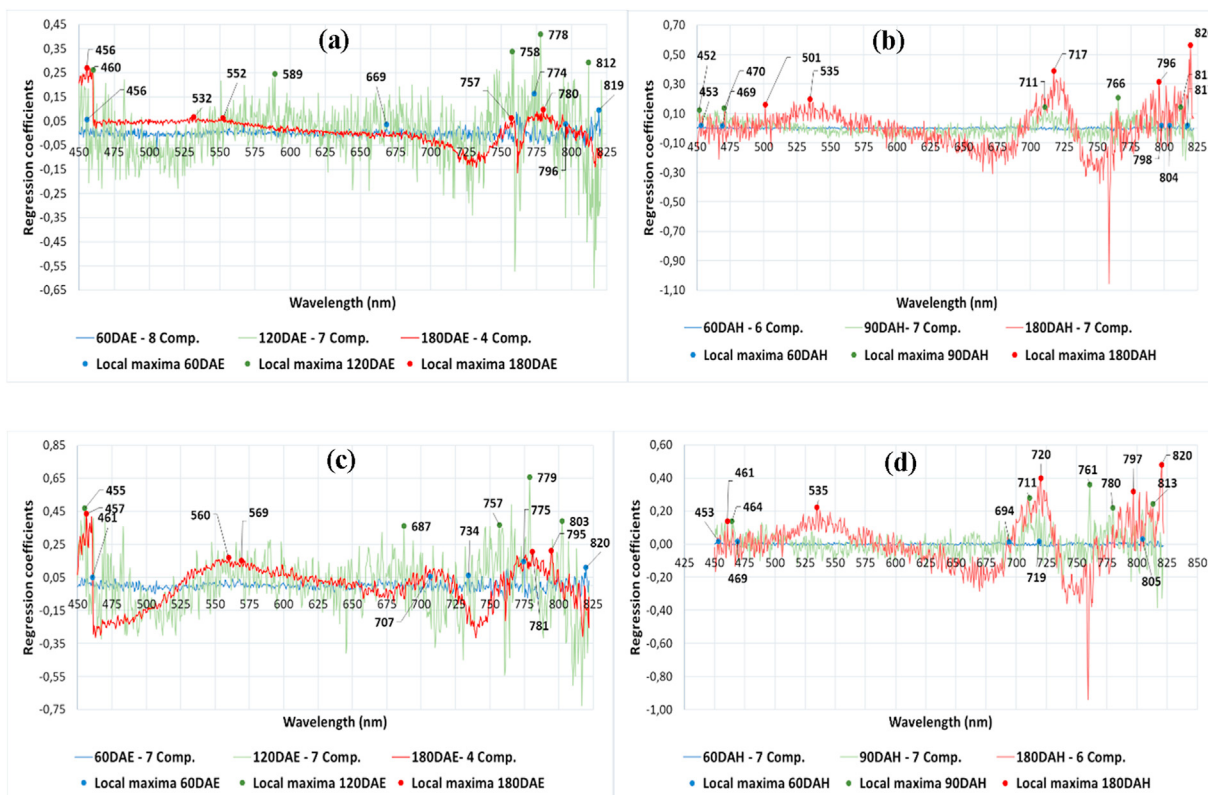


Figure 6. Regression coefficients for each wavelength of the best-performing PLS models. Method R for (a) plant cane and (b) first ratoon. Method MSC-SG for (c) plant cane and (d) first ratoon.

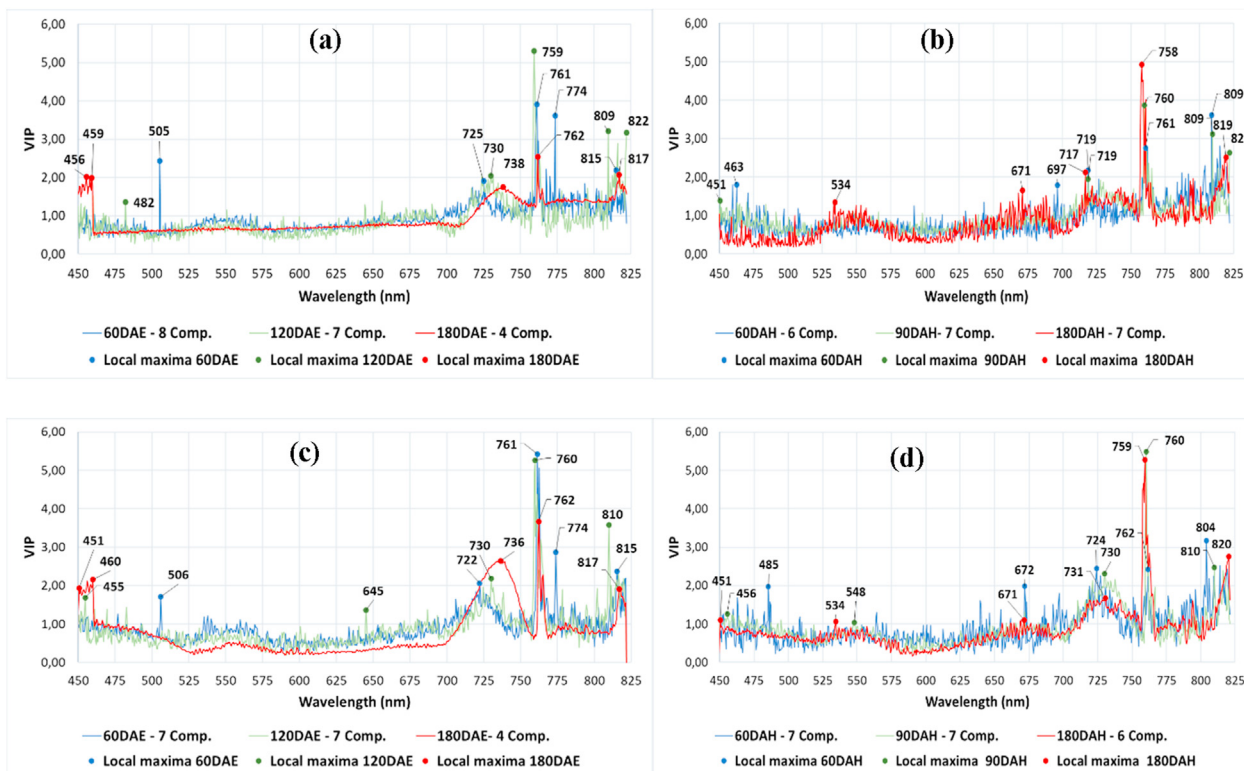


Figure 7. VIP for each wavelength of the best-performing PLS models. Method R for (a) plant cane and (b) first ratoon. Method MSC-SG for (c) plant cane and (d) first ratoon.

which can be attributed to the similar concentrations of Nc in the stages studied.

With respect to the vegetation indices, those calculated with the spectral band of 717 nm presented highly influential values. Additionally, the DVI presented the highest values of 2.82 and 2.54 for the plant cane and first ratoon stages, respectively. The results are in accordance with Lepine et al. (2016), who found strong correlations between DVI and N concentration in the canopy.

4. Conclusions

The Nc varied over time in both stages, increasing until 180 DAE for the plant cane stage. In first ratoon, Nc increased until 150 DAE. The different timing was due to physiological aspects of the plant. In relation to the spectral response and the influence of N application rate, the reflectance values were lower in the plant cane than in first ratoon, mainly due to the effect of the canopy architecture of sugarcane. The variable application of N generated differences in the NIR plateau (750–822 nm), where the reflectance was directly proportional to N application rate.

The results of this study show that the Nc concentration can be accurately estimated from crude canopy reflectance spectra when the sugarcane crop is young, with high values of r^2 ($r^2_{(60DAE)} = 0.98$, $r^2_{(120DAE)} = 0.92$, $r^2_{(60DAH)} = 0.96$, $r^2_{(90DAH)} = 0.98$), RPD ($RPD_{(60DAE)} = 6.65$, $RPD_{(120DAE)} = 3.68$, $RPD_{(60DAH)} = 5.32$, $RPD_{(90DAH)} = 9.55$) and low values of RMSEr ($RMSEr_{(60DAE)} = 3.80$, $RMSEr_{(120DAE)} = 4.49$, $RMSEr_{(60DAH)} = 4.35$, and $RMSEr_{(90DAH)} = 1.16$). This indicates that the reflectance spectra well represented the variations related to the N concentration of the leaf component of the canopy. These models allowed us to identify the effective wavelengths around different regions of the spectrum (450, 462, 505, 534, 671, 809, and 820 nm) and higher-density wavelengths of the red-edge region (719, 722, 725, 730, 758, and 760 nm), as well as VIs that consider this region in their calculation, such as VOGRE and DVI.

From the perspective of agronomic management, good-performing models for the first stages are critical when deep-root fertilization is the main method of improving soil fertility. If, however, fertigation is used, there would be no limitations to using spectral information to guide the targeted application of fertilizers. The results that relate the canopy spectral response to the Nc concentration show that it is possible to use this response to cover the large commercial areas of sugarcane in the Cauca River Valley and accurately diagnose the N nutrition status of the crop.

Declarations

Author contribution statement

Aldemar Reyes Trujillo: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Martha Constanza Daza Torres; Efrain Solarte Rodriguez: Analyzed and interpreted the data.

Carlos Augusto Galindez: Analyzed and interpreted the data; Wrote the paper.

Esteban Emilio Rosero García: Contributed reagents, materials, analysis tools or data.

Fernando Muñoz Arboleda: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Funding statement

This work was supported by the Universidad del Valle for funding the research project “Autonomous aerial system for mapping the N content of

a crop using spectral microsensors” and Universidad del Valle and CENICAÑA (CI-2961).

Data availability statement

The data that has been used is confidential.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

- Abdel-rahman, E.M., Ahmed, F.B., Van Den Berg, M., 2010. Estimation of sugarcane leaf nitrogen concentration using in situ spectroscopy. *Int. J. Appl. Earth Obs. Geoinf.* 12, 52–57.
- Abdel-Rahman, E.M., Ahmed, F.B., Ismail, R., 2013. Random forest regression and spectral band selection for estimating sugarcane leaf nitrogen concentration using EO-1 Hyperion hyperspectral data. *Int. J. Rem. Sens.* 34 (2), 712–728.
- Amaral, L., Molin, J., 2014. The effectiveness of three vegetation indices obtained from a canopy sensor in identifying sugarcane response to nitrogen. *Agron. J.* 106 (1), 273–280.
- Amaral, L., Molin, J., Portz, G., Finazzi, F., Cortinovet, L., 2014. Comparison of crop canopy reflectance sensors used to identify sugarcane biomass and nitrogen status. *Precis. Agric.* 16 (1), 15–28.
- Amaral, L., Molin, J., Schepers, J., 2015. Algorithm for variable-rate nitrogen application in sugarcane based on active crop canopy sensor. *Agron. J.* 107 (4), 1513–1523.
- Andrade, H., Amaral, L., Molin, J., Cantarella, H., 2015. Sugarcane response to nitrogen rates, measured by a canopy reflectance sensor, 1, 840–848.
- Araque, T., Jiménez, A., 2009. Caracterización de firma espectral a partir de sensores remotos para el manejo de sanidad vegetal en el cultivo de palma de aceite. *Revista Palmas* 30 (3 SE-Artículos), 63–79. <https://publicaciones.fedepalma.org/index.php/palmas/article/view/1455>.
- Barbin, D.F., ElMasry, G., Sun, D.-W., Allen, P., 2012. Predicting quality and sensory attributes of pork using near-infrared hyperspectral imaging. *Anal. Chim. Acta* 719, 30–42.
- Basyoumi, R., Dunn, B., 2013. Use of Reflectance Sensors to Monitor Plant Nitrogen Status in Horticultural Plants. <http://dasnr22.dasnr.okstate.edu/docushare/dsweb/Get/Version-14828/HLA-6719web.pdf>.
- Botero, J.M., Parra z, L.N., Cabrera, K.R., 2009. Determinación del nivel de nutrición foliar en banano por espectrometría de reflectancia. *Rev. Fac. Nac. Agron. Medellín* 62 (2), 5089–5098. <http://revistas.unal.edu.co/index.php/refame/article/view/24919>.
- Bulcock, H.H., Jewitt, G.P.W., 2010. Spatial Mapping of Leaf Area index Using Hyperspectral Remote Sensing for Hydrological Applications with a Particular Focus on Canopy Interception, pp. 383–392.
- CENICAÑA, 2018. Características agronómicas y de productividad de la variedad Cenicana Colombia (CC) 01-1940.
- Cordon, G.B., 2009. Métodos ópticos no destructivos para monitoreo de salud vegetal. Universidad de Buenos Aires, Argentina. http://digital.bl.fcen.uba.ar/gsd/282/cgi-bin/library.cgi?a=d&c=tesis&d=Tesis_4575_Cordon.
- Darvishzadeh, R., Skidmore, A., Schlerf, M., Atzberger, C., Corsi, F., Cho, M., 2008. LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. *ISPRS J. Photogrammetry Remote Sens.* 63 (4), 409–426.
- Franco, H.C.J., Otto, R., Faroni, C.E., Vitti, A.C., Almeida de Oliveira, E.C., Trivelin, P.C.O., 2011. Nitrogen in sugarcane derived from fertilizer under Brazilian field conditions. *Field Crop. Res.* 121 (1), 29–41.
- Hansen, P.M., Schjoerring, J.K., 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens. Environ.* 86 (4), 542–553.
- Henrich, V., Krauss, G., Götze, C., Sandow, C., 2012. IDB. In: *Entwicklung einer Datenbank für Fernerkundungsindizes*, 4. AK Fernerkundung, Bochum, 5. 10. 2012. www.indexdatabase.de.
- Huber, S., Kneubühler, M., Psomas, A., Itten, K., Zimmermann, N.E., 2008. Estimating foliar biochemistry from hyperspectral data in mixed forest canopy. *For. Ecol. Manag.* 256 (3), 491–501.
- Katsoulas, N., Elvanidi, A., Ferentinos, K.P., Kacira, M., Bartzanas, T., Kittas, C., 2016. Crop reflectance monitoring as a tool for water stress detection in greenhouses : a review. *Biosyst. Eng.* 151, 374–398.
- le Maire, G., François, C., Soudani, K., Berveiller, D., Pontailleur, J.-Y., Bréda, N., Genet, H., Davi, H., Dufréne, E., 2008. Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sens. Environ.* 112 (10), 3846–3864.
- Lebourgeois, V., Bégue, A., Labbé, S., Houles, M., Martiné, J.F., 2012. A light-weight multi-spectral aerial imaging system for nitrogen crop monitoring. *Precis. Agric.* 13 (5), 525–541.

- Lepine, L.C., Ollinger, S.V., Ouimette, A.P., Martin, M.E., 2016. Remote Sensing of Environment Examining spectral reflectance features related to foliar nitrogen in forests : implications for broad-scale nitrogen mapping. *Remote Sens. Environ.* 173, 174–186.
- Li, D., Wang, C., Liu, W., Peng, Z., Huang, S., Huang, J., Chen, S., 2016a. Estimation of litchi (*Litchi chinensis* Sonn.) leaf nitrogen content at different growth stages using canopy reflectance spectra. *Eur. J. Agron.* 80, 182–194.
- Li, L., Lu, J., Wang, S., Ma, Y., Wei, Q., Li, X., Cong, R., Ren, T., 2016b. Methods for estimating leaf nitrogen concentration of winter oilseed rape (*Brassica napus* L.) using in situ leaf spectroscopy. *Ind. Crop. Prod.* 91, 194–204.
- Lofton, J., Tubana, B.S., Kanke, Y., Teboh, J., Viator, H., Dalen, M., 2012. Estimating sugarcane yield potential using an in-season determination of normalized difference vegetative index. *Sensors (Switzerland)* 12 (6), 7529–7547.
- Miphokasap, P., Honda, K., Vaiphasa, C., Souris, M., Nagai, M., 2012. Estimating canopy nitrogen concentration in sugarcane using field imaging spectroscopy. *Rem. Sens.* 4 (6), 1651–1670.
- Mutanga, Onesimo, Skidmore, A.K., 2007. Red edge shift and biochemical content in grass canopies. *ISPRS J. Photogrammetry Remote Sens.* 62 (1), 34–42.
- Portz, G., Amaral, L.R., Molin, J.P., Adamchuk, V.I., 2013. Field comparison of ultrasonic and canopy reflectance sensors used to estimate biomass and N-uptake in sugarcane. In: *Precision Agriculture 2013 - Papers Presented at the 9th European Conference on Precision Agriculture, ECPA 2013, January*, 111–117.
- Portz, G., Molin, J.P., Jasper, J., 2012. Active crop sensor to detect variability of nitrogen supply and biomass on sugarcane fields. *Precis. Agric.* 13 (1), 33–44.
- Robles, W., Madsen, J.D., Wersal, R.M., 2010. Potential for remote sensing to detect and predict herbicide injury on waterhyacinth (*Eichhornia crassipes*). *Invasive Plant Sci. Manag.* 3 (4), 440–450.
- Robson, A., Rahman, M.M., Falzon, G., Verma, N.K., Johansen, K., Robinson, N., Lakshmanan, P., Salter, B., Skocaj, D., 2016. Evaluating remote sensing technologies for improved yield forecasting and for the measurement of foliar nitrogen concentration in sugarcane. *Int. Sugar J.* 118 (1416), 936–942.
- Schultz, N., de Morais, R.F., da Silva, J.A., Baptista, R.B., Oliveira, R.P., Leite, J.M., Pereira, W., de Barros Carneiro Júnior, J., Alves, B.J.R., Baldani, J.I., Boddey, R.M., Urquiaga, S., Reis, V.M., 2012. Agronomic evaluation of varieties of sugar cane inoculated with diazotrophic bacteria and fertilized with nitrogen. *Pesqui. Agropecu. Bras.* 47 (2), 261–268. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84862568806&partnerID=40&md5=a78ca83a4b8b1872c4fcc72839c8c2f3>.
- Steidle Neto, A.J., Toledo, J.V., Zolnier, S., Lopes, D.D.C., Pires, C.V., Silva, T.G.F.D., 2017. Prediction of mineral contents in sugarcane cultivated under saline conditions based on stalk scanning by Vis/NIR spectral reflectance. *Biosyst. Eng.* 156, 17–26.
- Stevens, A., Ramirez-López, L., 2014. Miscellaneous functions for processing and sample selection of vis-NIR diffuse reflectance data. In: *CRAN 'prospectr' R*, p. 32. <https://github.com/antoinestevens/prospectr>.
- Suarez, L.A., Apan, A., Werth, J., 2016. Hyperspectral sensing to detect the impact of herbicide drift on cotton growth and yield. *ISPRS J. Photogrammetry Remote Sens.* 120, 65–76.
- Tewari, J.C., Dixit, V., Cho, B.-K., Malik, K.A., 2008. Determination of origin and sugars of citrus fruits using genetic algorithm, correspondence analysis and partial least square combined with fiber optic NIR spectroscopy. *Spectrochim. Acta Mol. Biomol. Spectrosc.* 71 (3), 1119–1127.
- Wang, S.Q., Li, W.D., Li, J., Liu, X.-S., 2013. Prediction of soil texture using FT-NIR spectroscopy and PXRF spectrometry with data fusion. *Soil Sci.* 178 (11), 626–638.
- Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. *Chemometr. Intell. Lab. Syst.* 58 (2), 109–130.
- Xiuhua, L., Xiao, C., Yonghua, Z., Mengling, N., Xiaoyang, L., Jiaoyan, A., 2015. Spectral characteristics analysis and chlorophyll content detection of sugarcane leaves under different fertilizer treatments. *Trans. Chin. Soc. Agric. Eng.* 31 (2), 118–123.
- Zhao, D., Reddy, K.R., Kakani, V.G., Read, J.J., Koti, S., 2007. Canopy reflectance in cotton for growth assessment and lint yield prediction. *Eur. J. Agron.* 26, 335–344.
- Zhu, L., Chen, Z., Wang, J., Ding, J., Yu, Y., Li, J., Xiao, N., Jiang, L., Zheng, Y., Rimmington, G.M., 2014. Monitoring plant response to phenanthrene using the red edge of canopy hyperspectral reflectance. *Mar. Pollut. Bull.* 86 (1), 332–341.