Ciência Rural, Santa Maria, v.52:7, e20200630, 2022 http://doi.org/10.1590/0103-8478cr20200630

ISSNe 1678-4596
RURAL ENGINEERING

Estimation of leaf nitrogen levels in sugarcaneusing hyperspectral models

Pedro Paulo da Silva Barros[1*] **Peterson Ricardo Fiorio**[2] **José Alexandre de Melo Demattê**[3] **Juliano Araújo Martins**[4] **Zaqueu Fernando Montezano**[5] **Fábio Luis Ferreira Dias**[6]

[1]Faculdade de Engenharia Civil, Universidade Federal de Uberlândia (UFU), Campus de Monte Carmelo, 38500-000, Monte Carmelo, MG, Brasil. E-mail: pedropaulo76@gmail.com. *Corresponding author.
[2]Departamento de Engenharia de Biossistemas, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo (USP), Piracicaba, SP, Brasil.
[3]Departamento de Ciência do Solo, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo (USP), Piracicaba, SP, Brasil.
[4]Instituto Federal de Mato Grosso (IFMT), Campus Sorriso, Sorriso, MT, Brasil.
[5]Innovak Global, Campinas, SP, Brasil.
[6]Instituto Agronômico de Campinas (IAC), Campinas, SP, Brasil.

ABSTRACT: *Sugarcane is a good source of renewable energy and helps reduce the emission of greenhouse gases. Nitrogen has a critical role in plant growth; therefore,estimating nitrogen levels is essential, and remote sensing can improve fertilizer management. This field study selects wavelengths from hyperspectral data on a sugarcane canopy to generate models for estimating leaf nitrogen concentrations. The study was carried out in the municipalities of Piracicaba, Jaú, and Santa Maria da Serra, state of São Paulo, in the 2013/2014 growing season. The experiments were carried out using a completely randomized block design with split plots (three sugarcane varieties per plot [variety SP 81-3250 was common to all plots] and four nitrogen concentrations [0, 50, 100, and 150 kgha^{-1}] per subplot) and four repetitions. The wavelengths that best correlated with leaf nitrogen were selected usingsparse partial least square regression. The wavelength regionswere combinedby stepwise multiple linear regression. Spectral bands in the visible (700–705 nm), red-edge (710–720 nm), near-infrared (725, 925, 955, and 980 nm), and short-wave infrared (1355, 1420, 1595, 1600, 1605, and 1610 nm) regions were identified. The R^2 and RMSE of the model were 0.50 and 1.67 g.kg^{-1}, respectively. The adjusted R^2 and RMSE of the models for Piracicaba, Jaú, and Santa Maria were 0.31 (unreliable) and 1.30 g.kg^{-1}, 0.53 and 1.96 g.kg^{-1}, and 0.54 and 1.46 g.kg^{-1}, respectively. Our results showed that canopy hyperspectral reflectance can estimate leaf nitrogen concentrations and manage nitrogen application in sugarcane.*
Key words: *remote sensing, Saccharumspp, nitrogen fertilization, reflectance, sPLS, regression model.*

Modelagem hiperespectral na quantificação de nitrogênio foliar em cana-de-açúcar

RESUMO: *A cana-de-açúcar se destaca como uma das fontes de energia renovável frente às estratégias para reduzir a emissão de gases causadores do efeito estufa. O nitrogênio é um dos mais significativos devido ao seu impacto no crescimento de folhas e colmos. Portanto, o monitoramento eficiente do nitrogênio aplicado é essencial e o sensoriamento remoto se apresenta como uma alternativa na melhoria do gerenciamento da adubação. O presente trabalho teve por objetivo selecionar comprimentos de onda a partir de dados hiperespectrais de dossel da cana-de-açúcar para geração de modelos na predição da concentração de Nitrogênio. O estudo foi realizado em experimentos de campo instalados nos municípios de Piracicaba, Jaú e Santa Maria da Serra, estado São Paulo, na safra 2013/2014. Cada experimento foi alocado em blocos ao acaso, com parcelas subdivididas e quatro repetições, com variedades de cana-de-açúcar na parcela (três variedades por local, sendo a SP 81-3250 comum à todos) e doses de nitrogênio (0, 50, 100 e 150 kg.ha^{-1}) na subparcela. Na seleção dos comprimentos de onda que melhor se correlacionam com o TFN foi utilizada a metodologia sPLS. Posteriormente, foi realizada a combinação linear dos comprimentos de onda selecionados pela metodologia sPLS, por meio de Regressão Linear Múltipla por Stepwise (SMLR). Foram identificadas bandas importantes nas regiões do visível (700 a 705 nm), red-edge(710 a 720 nm), infravermelho próximo (725, 925, 955 e 980 nm) e infravermelho de ondas curtas (1355, 1420, 1595, 1600, 1605 e 1610 nm). O modelo de predição de TFN teve valores de R^2 de 0,50 e o RMSE de 1,67 g.kg^{-1}. Os modelos gerados para Piracicaba, Jaú e Santa Maria obtiveram R^2 ajustados e RMSE, respectivamente, de 0,31 considerado não confiável (1,30 g.kg^{-1}), 0,53 (1,96 g.kg^{-1}) e 0,54 (1,46 g.kg^{-1}). Os sensores hiperespectrais de dossel podem ser utilizados para predição do TFN e monitoramento de aplicação de nitrogênio em cana-de-açúcar.*
Palavras-chave: *sensoriamento remoto, Saccharumspp,adubação nitrogenada, reflectância, sPLS, modelo de regressão.*

INTRODUCTION

According to the Food and Agriculture Organization Corporate Statistical Database (FAOSTAT, 2019), Brazil is the world's largest producer of sugarcane, with 40% of the global production, and has developed successful initiatives in the search for renewable energy sources. Finite natural resources and adverse effects on the environment and human health due to the use of non-

Received 07.07.20 Approved 08.11.21 Returned by the author 10.22.21
CR-2020-0630.R3
Editors: Leandro Souza da Silva Tales Tiecher

renewable energy sources have increased the global interest in renewable sources, including bioethanol, derived from energy crops.

Sugarcane, like any other crop, depends on nitrogen (N) for biomass production. Nitrogen is the most significant limiting factor, after water deficit, for biomass production. Nitrogen fertilizers can remarkably increase sugarcanetillering and yield (AMARAL et al., 2015). However, because of climate variations, farmers seeking to achieve high yields often use more nitrogen than necessary in production fields.

ALI et al. (2016) reported that estimating nitrogen levels in crops is fundamental in agronomic studies. Both destructive and non-destructive techniques are used to calculate nitrogen concentration. The destructive technique accurately estimates leaf nitrogen levels under laboratory conditions; however, the time from the sample collection to the results of the analysis exceeds the ideal time to carry out agricultural interventions during the growth cycle. Furthermore, laboratory analyses involve using chemicals, generating toxic waste and potentially leading to environmental contamination (IZNAGA et al., 2014; ZHAO et al., 2018).

Non-destructive methods are fast and cost-effective; nonetheless, their complexity varies because obtaining spectral data with passive sensors without considering important parameters (solar azimuth, solar elevation angle, and plant biophysical parameters) limits data analysis; thus,the methods requirethe accurate calibration of sensors (MARTINS et al., 2021; MOKHELE& AHMED, 2010; ZHAO et al., 2012; MAHAJAN et al., 2014; LISBOA et al., 2018; MORIYA et al., 2018).

To better understand nitrogen dynamics in sugarcane crops, wavelengths potentially associated with nitrogen concentrations have been evaluated using hyperspectral sensors (ABDEL-RAHMAN et al., 2010; MIPHOKASAP et al., 2012; ABDEL-RAHMAN et al., 2013; MIPHOKASAP & WANNASIRI, 2018). However, few studies have evaluated the use of these sensors in sugarcane in Brazil.

Hyperspectral sensors facilitate the analysis of specific regions of the electromagnetic spectrum to accurately model the attributes of interest, enabling detailed analysis of crop characteristics (THENKABAIL et al., 2010; MULLA, 2013). Thus, studies based on hyperspectral data can help develop sensors and crop analysis methods.

However, the analysis of many dependent variables andfew independent variables by hyperspectral sensors is complex because of the existence of multicollinearity between variables, leading to overestimations of the regression coefficients of the adjusted models (COIMBRA et al., 2005).

This issue can be resolved by multivariate techniques that identify regions of the electromagnetic spectrum associated with the attributes of interest, thus avoiding overestimations (ABDEL-RAHMAN et al., 2014; FIORIO et al., 2018; TAVARES et al., 2020, MARTINS et al., 2021).

This study selects wavelengths in canopy hyperspectral reflectance data to generate models for estimating leaf nitrogen concentration in sugarcane. We started from the premise that there is a strong relationship between reflectance and plant structures, and that alterations in these structures due to nutrient deficiency cause changes in canopy reflectance in specific spectral regions; these data could be used to measure the attributes of interest (GITELSON et al., 2005).

MATERIALS AND METHODS

Study area

This field study was conducted in two units of the São Paulo Agribusiness Technology Agency (Agência Paulista de Tecnologia dos Agronegócios - APTA) in the municipalities of Jaú and Piracicaba, São Paulo, Brazil, and at a Raízen unit in Santa Maria da Serra, São Paulo (Table 1).

According to Köppen's classification, the climate of the region is humid subtropical (Cwa), with an average annual rainfall of less than 1400 mm, with rainy summers and dry winters. The experiments were performed in March 2010 using completely randomized block and split-plot designs, with sugarcane varieties in plots and nitrogen concentrations in subplots. In the 2013/2014 growing season, the study areas were in the fourth growth cycle, i.e., the cropwas under nitrogen deficiency stress. Three sugarcane varieties were planted in each area, as follows: SP 81-3250, IAC SP 95-5000, and RB 85-5536 in Jaú; SP 81-3250, IAC 87-3396, and CTC 14 in Piracicaba; SP-81- 3250, RB 93579, and RB 86-7515 in Santa Maria da Serra. In the 0–40 cm soil layer, the soil type was classified as medium, clayey, and sandy in Jaú, Piracicaba, and Santa Maria da Serra.

The concentrations of nitrogen and other nutrients for crop growth were defined according to the official table of fertilizer recommendations for sugarcane in the state of São Paulo (RAIJ & CANTARELLA, 1997). The nitrogen doses applied for ratoon cane were 0, 50, 100, and 150 kg ha^{-1}, and

Table 1 - Experimental areas.

	Area 1	Area 2	Area 3
Location	Jaú	Piracicaba	Santa Maria da Serra (STM)
Description	Center-west branch (APTA)	Center-south branch (APTA)	Itaúna Farm - Raízen
Coordinates	22°15'08"S; 48°34'04"O	22°41'05"S; 47°38'54"O	22°33'26"S; 48°16'42"W
Soil classification	Red Latosol	Red Ultisol	Quartzarenic Neosol
Soil texture	Medium sandy (0-20 cm), Medium clayey (20-40 cm)	Clayey (0-20 cm) Very Clayey (20-40 cm)	Sandy (0-40 cm)
Varieties	SP 81-3250 IAC 95-5000 RB 85-5536	SP 81-3250 IAC 87-3396 CTC 14	SP 81-3250 RB 92579 RB 86-7515
Days after cutting (DAC)	148	169	146

the nitrogen fertilizer source was ammonium nitrate, which was applied under straw from the previous harvest. Additionally, P_2O_5 (40 kg ha^{-1}) and K_2O (150 kg ha^{-1}) were applied to each plot.

Meteorological data were obtained at station A741 (latitude, 22°28’16” South; longitude, 48°33’27” West; and altitude, 534 m) (built in 2008) located at the National Institute of Meteorology (*Instituto Nacional de Meteorologia* - INMET), in Barra Bonita, São Paulo,and at the meteorological station (latitude, 22°42’30” South; longitude, 47°38’00” West; altitude, 546m) (constructed in 1917) located at the Higher School of Agriculture “Luiz de Queiroz” (*Escola Superior de Agricultura “Luiz de Queiroz”*), in Piracicaba, São Paulo.

Collection of canopy spectral data

Each subplot consisted of five sugarcane rows, and spectral data were collected from the three central rows. Four plants were randomly selected in the first and third rows (two in each) and one in the middle row.

Canopy spectral data were acquired using a FieldSpec®3 spectroradiometer (Analytical Spectral Device, CO, USA). This sensor records signals in the visible-near infrared region (350–1000 nm) and two short-wave infrared regions (1001–1800 nm and 1801–2500 nm) at a sampling interval of 1.4 and 2.0 nm and spectral resolutions (full width at half maximum) of 3 nm and 10 nm, respectively.

Data were collected on sunny days between 10:00 am and 2:00 pm. The sensor was positioned 1 m above the average height of the plant stalks, with a field of view of 25°, allowing the measurement of an area of approximately 0.25 m^2. Calibration was performed in each plot using a standard scale provided by the manufacturer, and readings were done in each subplot in five previously demarcated areas. This process was performed in all plots.

Analysis of leaf nitrogen concentration

For all subplots where canopy spectral data were obtained, leaves were collected to determine the nitrogen concentration. In each subplot, 10 + 1 leaves (two leaves on each side of five plants) were collected 4 to 5 months after nitrogen fertilization. The denomination +1 refers to the first fully expanded leaf from the plant apex.

The leaves were sequentially washed in running water, distilled water with detergent, and distilled water. Subsequently, they were dried in an oven with forced ventilation at 65 °C and ground. Chemical analyses were performed using the semi-micro Kjeldahl method in extracts obtained by sulfuric digestion (MALAVOLTA et al.,1997).

Statistical analysis

Data processing

The normality of the frequency distribution of the nitrogen data was analyzed using the Shapiro-Wilk test (1965). The test is based on squared values and is the most common in the normality test.

Regions known to produce spectral noise (350–399, 1355–1420 nm, and 1800–2500 nm) due to interference from environmental moisture (ABDEL-RAHMAN et al., 2013) were excluded from the spectral curves and final analysis.

The characteristics of spectral curves in each subplot of each demarcated area were assessed by multivariate analysis, checking if the mean represented the subplot. Outliers were identified using a normality test (Shapiro-Wilk) and analysis of variance (ANOVA). This analysis was performed using R software.

A median filter for noise smoothing followed by multiplicative scatter correction (MSC) (ISAKSSON & NAES, 1988) was applied to the average spectral reflectance curves. MSC is a transformation method used to compensate for the additive and/or multiplicative effects on spectral data. This analysis was performed using Parles software version 3.1 (VISCARRA ROSSEL, 2008).

Dimension reduction and wavelength selection

A major problem with multivariate data is that the number of observations is greater than that of the predictive wavelengths, even in cases where these observations are highly correlated. The sparse partial least square (sPLS) methodology is a multivariate procedure based on partial least squares (PLSs). Its central principle is the measurement of the wavelength dispersion through PLS regression, allowing the efficient selection of wavelengths. This methodology was implemented using the "sPLS" analysis package in R software (CHUNG et al., 2012).

The results were calculated using the entire calibration dataset (48 samples per area) and generated coefficients that indicated the importance of each wavelength for the nitrogen level estimation; the 15 most important wavelengths were selected. This number proved to be adequate because, in the later phase, none of the final models calibrated by stepwise multiple linear regression (SMLR) required more than four wavelengths.

SMLR

Models that estimated the leaf nitrogen concentrations in sugarcane were generatedby SMRL using the "MASS" analysis package in R (DARVISHZADEH et al., 2008). The initial dataset contained 48 spectral curves from each study area; approximately 2/3 (n = 33) were used for calibrating the spectral models, which were later validated using the remaining data (n = 15). The calibration and validation datasets were randomly selected from the initial dataset.

Calibration by SMLR began with a model without predictor variables. At each step, a variable was added beginning with the most significant (highest p-value). The process ended when the inclusion of a new predictor variable did not improve the accuracy of the model. This approach guarantees that the model has the highest performance and the lowest possible number of variables (MIPHOKASAP et al., 2012).

Assessment of model accuracy

The accuracy of the generated models was assessed based on the coefficient of determination (R^2) (equation 1) and the root mean square error (RMSE) (equation 2).

$$R^2 = \frac{\sum(\widehat{y_i} - \bar{y_i})^2}{\sum(y_i - \bar{y})^2} \quad (1)$$

where $\widehat{y_i}$, $\bar{y_i}$, and y_i are the predicted, average measured, and measured values, respectively.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(Pi - Oi)^2}{n}} \quad (2)$$

where *Pi* and *Oi* are the predicted and actual nitrogen concentrations, respectively, and *n* is the number of observations used to generate or validate the model.

RMSE is commonly used to measure the accuracy of numerical predictions, with excellent results for the overall behavior of the models. This indicator has been used in several areas of science, including spectroradiometry (SENANAYAKE et al., 2019). One of the advantages of this error metric is that the results are given in the same unit as the study variables.

RESULTS AND DISCUSSION

Evaluation of leaf nitrogen levels and monthly rainfall distribution in the experimental areas

Results showed that the residuals followed a normal distribution, according to the Shapiro Wilk test (W=0.9856; P=0.00542). The regression model for each experimental area and a general regression model are shown in figure 1.

The nitrogen levels in sugarcane leaves range from 18 to 25 gkg^{-1} (RAIJ & CANTARELLA, 1997). The results showed that the crop was under nitrogen stress (Figure 1), andthe low rainfall during the growing season explainedthe low nitrogen levels.

Extreme environmental events are more common in the southeast of Brazil because of longer and more severe droughts; rainfall was below the

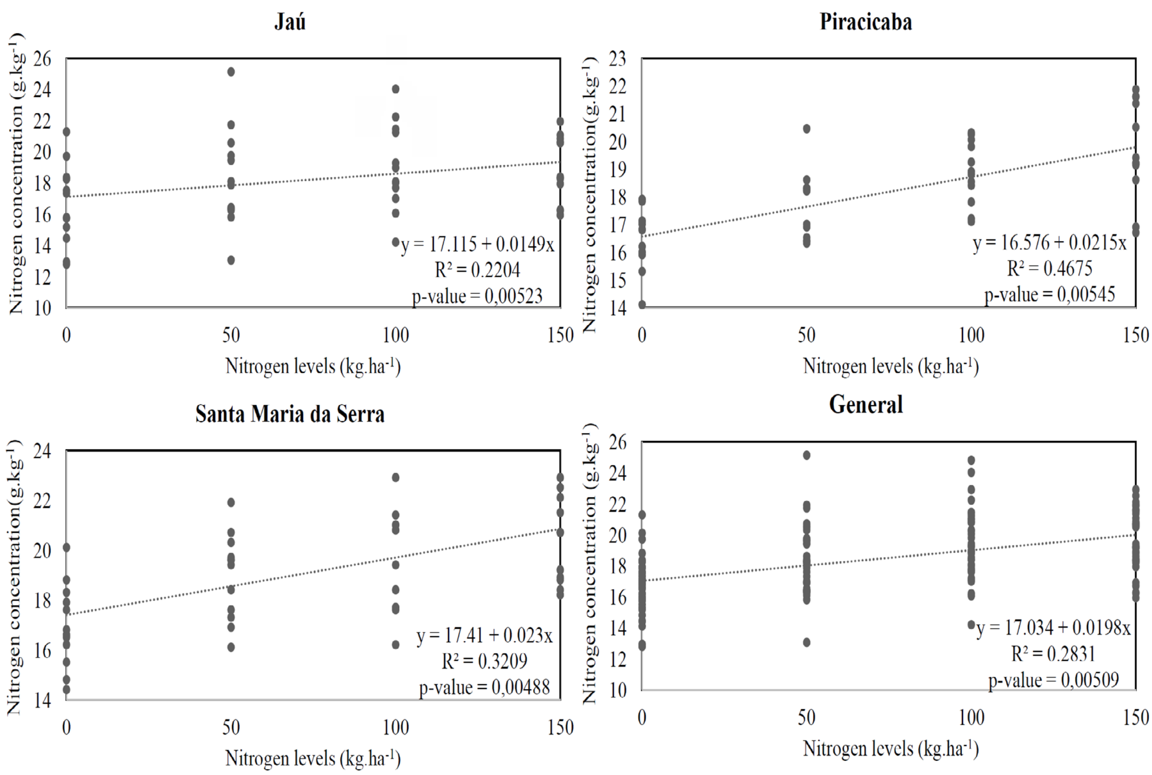

Figure 1 - Regression between nitrogen application and leaf nitrogen concentration (g.kg-1).

historical average for several months in 2014 (Figure 2) (MONTEIRO & SENTELHAS, 2017). For instance, in Piracicaba, rainfall from November 2013 to February 2014 was below historical averages and increased slightly in March 2014 (Figure 2B).

The water deficit directly affects the crop development and causes morphophysiological changes, including leaf curling, changes in the leaf angle, and a decrease in the leaf area, depending on the cultivar genotype, severity of change, and development stage of the plant (CHAVES et al., 2009). These changes can affect the spectral characteristics of the crop.

Selection of wavelengths using the sPLS method

Although, area-specific regression models and a general model were generated in the calibration phase, the complete dataset (48 samples per experimental areaand a total of 144 samples) for the 2013/2014 growing season was used for wavelength selection bysPLSregression. These coefficients indicated the contribution of each wavelength across the electromagnetic spectrum; no specific region can individually describe the variability in nitrogen levels in sugarcane crops (ABDEL-RAHMAN et al., 2014). This method is useful in conditionswith large datasets and many independent variables (DEMATTÊ et al., 2015). However, few studies have used the sPLS methodology to select variables in hyperspectral data, especially in crops.

PEERBHAY et al. (2014) compared PLS and sPLS to select variables for the discriminant analysis of pine varieties in South Africa and reported that the efficiency of the analysis increased from 71.88% using PLS to 80.21% using sPLS. ABDEL-RAHMAN et al. (2014) compared the two methods to select variables and generate models to predict the productivity of vegetables in South Africa.

Selecting wavelengths associated with changes in nitrogen levels is crucial because, although

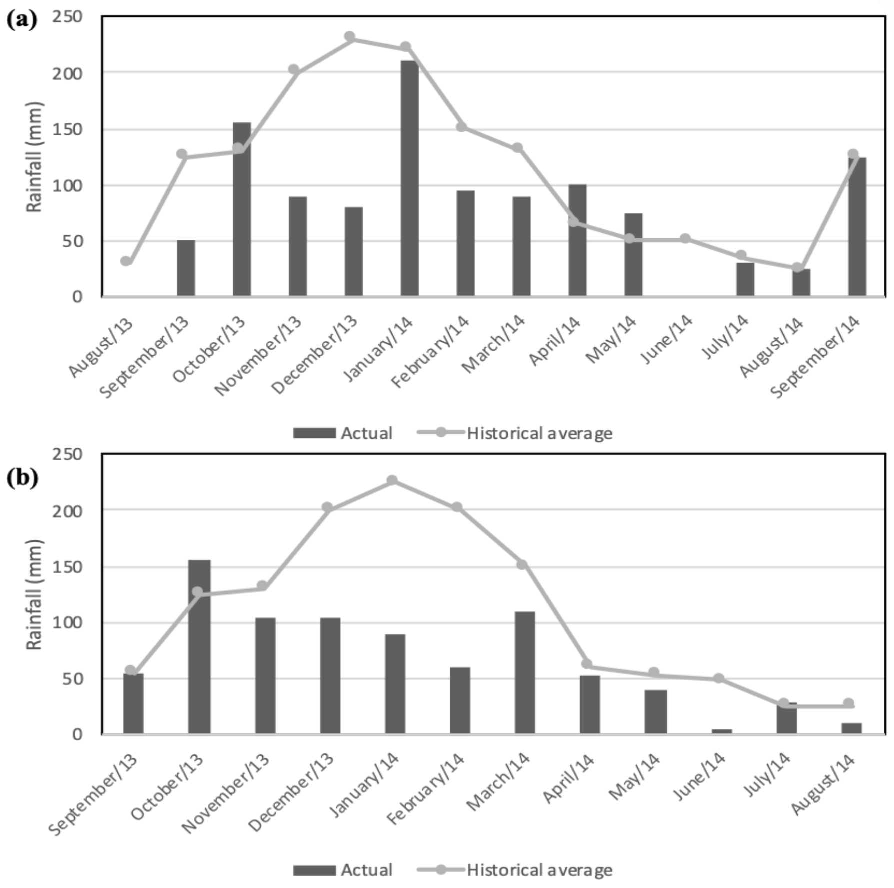

Figure 2 - Monthly accumulated rainfall in Jaú and Santa Maria da Serra. Source: INMET – A741 from Barra Bonita, state of São Paulo (A), Monthly accumulated rainfall in Piracicaba. Source: ESALQ/USP weather station (B).

a plant variety/environment is represented by specific wavelengths, sensors capable of detecting a few wavelengths can be developed for a specific crop using algorithms that can select different spectral regions; however, developing sensors for specific conditions is not feasible.

Wavelengths with zero values cannot estimate nitrogen levels. A total of 116 wavelengths presented positive or negative values, reducing the dimensionality of hyperspectral data by more than 91%. The 15 most significant wavelengths were selected to generate the models andbelonged to the following spectral regions: visible red (700–705 nm), red-edge (710, 715, and 720 nm), near-infrared (725, 925, 955, and 980 nm), and short-wave infrared (1355, 1420, 1595, 1600, 1605, and 1610 nm).

MIN & LEE (2005) evaluated five wavelength selection methods to determine the nitrogen concentration in citrus and found that the higher number of variables relative to the number of samples increased the collinearity, potentially creating instability in the regression model.

MIPHOKASAP & WANNASIRI (2018) evaluated three methods of generating a hyperspectral model using the Hyperion orbital sensor to determine the nitrogen levels in four sugarcane varieties and observed that the models with the best fit required two to four wavelengths. MARTINS et al. (2021) reported that five to six wavelengths were required to generate area-specific models and ten wavelengths to develop a general model.

Calibration of SMLR models

In the calibration phase, the performance of the models created using reflectance data was similar, with an adjusted R^2 ranging from 0.30 to

0.55 for site-specific models and 0.39 for the general model (Table 2). MARTINS et al. (2021) used a hyperspectral sensor to analyze the reflectance data on crop leaves at 90 days after cutting (DAC) from three study areas and four nitrogen doses. They observed that R^2 ranged from 0.61 to 0.71 and the RMSE ranged from 0.80 to 1.14.

In our sample, the wavelengths most sensitive to variations in nitrogen (Table 2) were reported in the red-edge region (710, 715, and 720) because all models selected at least one wavelength in this region, and only wavelengths from this spectral range were selected in the Santa Maria area. A possible explanation is that these wavelengths are highly responsive to variations in nitrogen levels, whereas other wavelengths are influenced by other factors not controlled in this study, as observed by MUTANGA & SKIDMORE (2007) and MARTINS et al. (2021).The red-edge region (680–780 nm) is associated with chlorophyll, nitrogen, and water content, and crop characteristics (JENSEN, 2009; HENNESSY et al., 2020).

In this study, in the near-infrared region, only the wavelength at 980 nm, associated with water absorption, was used to generate the models (STRACHAN et al., 2002). Nitrogen affects important physiological processes in plants, and nitrogen stress can affect the plant cellular structure and thus, the near-infrared reflectance (CECCATO et al., 2001).

Four wavelengths (1355, 1420, 1600, and 1605 nm) in the short-wave infrared region significantly contributed to the generation of the models. This region is responsive to variations in the leaf water content (JENSEN, 2009). Although, several studies haveshown the importance of this region because water content affects the leaf reflectance, the effect of other factors on reflectance should be further evaluated.

Validation of the SMRM models

The performance of the spectral models by area was satisfactory, especially in Jaú and Santa Maria (Figure 3). In Piracicaba, the adjusted R^2 and RMSE were 0.31 and 1.30 g kg^{-1}, respectively, and these values are considered unreliable (MALLEY et al., 2004). These values are lower than those of sugarcane crops in Thailand (R^2 of 0.73) (MIPHOKASAP et al., 2012). The accuracy of the estimates using the general model was slightly higher than that during the calibration phase (Table 2).

MUTANGA & SKIDMORE (2007) showed that the maximum change in the slope of the reflectance spectra in the red-edge region usually occurred at 720 nm and concluded that variations in the crop growth, plant stress, leaf area index, and chlorophyll and nitrogen concentrations could be detected at this wavelength. Similar results were obtained in this study, in which two calibration models (general and Santa Maria) used this wavelength, and the adjusted R^2 of the latter was higher (0.54) (Table 2).MIPHOKASAP et al. (2012) developed a model to explain variations in nitrogen levels in sugarcane and chose five wavelengths: 410, 426, 720, 754, and 1216 nm.

General models are more stable under heterogeneous conditions because they describe the intrinsic characteristics of crops and are less sensitive to local conditions. Area-specific models describe the crop characteristics in a particular cultivation environment; however, these data cannot be applied to other areas with different conditions.

The results obtained using our models may be related to the variable response of the crop to nitrogen fertilization and the fact that canopy spectral data are influenced by factors other than nitrogen; therefore, nitrogen concentration is estimated indirectly (MIPHOKASAP et al., 2012). INOUE et

Table 2 - Calibration of stepwise multiple linear regression models for predicting leaf nitrogen concentration in sugarcane based on reflectance data. In the equations, the letter B represents the wavelength used in the model.

Area	Equation	Adjusted R^2	RMSE (g kg^{-1})	AIC
Jaú	Y = 38.26 – 107.08 * B715 – 44.67 * B1420 + 1583.08 * B1600 – 1549.21 * B1605	0.39	1.64	235
Piracicaba	Y = 100.83 – 40.53 * B715 – 186.57 * B980 + 8.38 * B1355	0.30	1.32	269
Santa Maria	Y = 53.32 + 227.09 * B710 – 350.54 * B720	0.55	1.38	154
General	Y = 40.16 – 99.03 * B720 – 36.03 * B1420	0.39	1.72	168

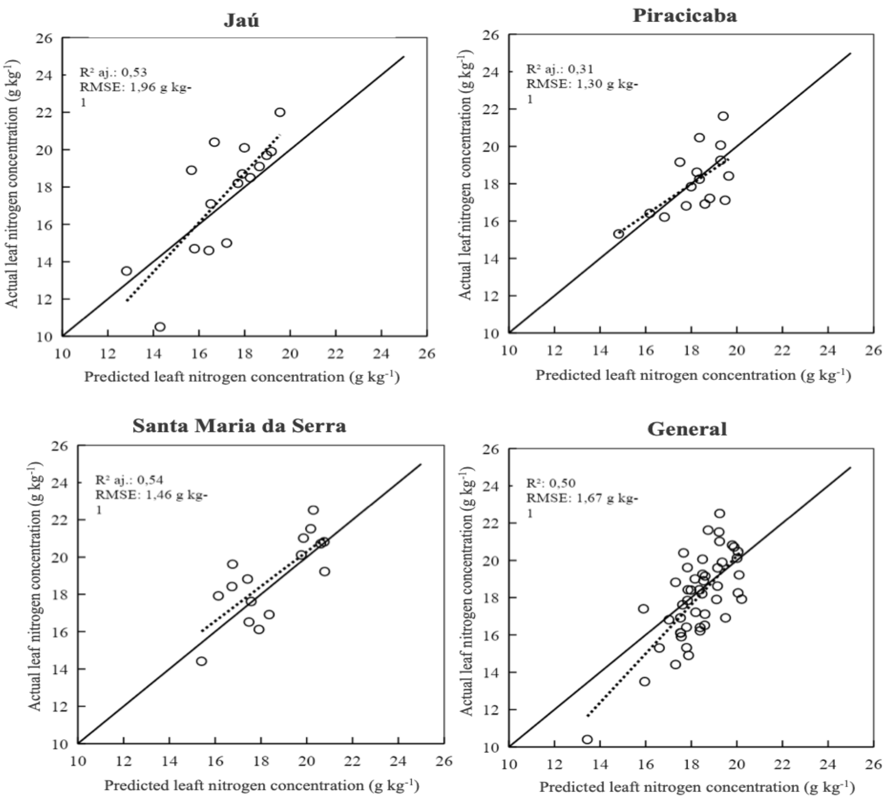

Figure 3 - Validation of stepwise multiple regression models used to estimate leaf nitrogen concentration in sugarcane.

al. (2016) reported that the accuracy and applicability of the models were highly dependent on the size and quality of the dataset because the number of samples directly affected the performance of the results during the modeling of plant biochemical parameters.

This limitation may be one of the reasons for the low adjusted R^2 in the study area. In this respect, ROSA et al. (2015) evaluated four regions of the state of São Paulo (two of them close to Jaú and Piracicaba) and reported that different plant varieties responded differently to the climatic factors, soil type, and management conditions, resulting in variability in the crop vigor and canopy volume in the same field.

GAVA et al. (2001) assessed the growth and nitrogen accumulation in sugarcane cultivated in straw-covered soil and observed that translocation occurred at 204–237 DAC, with the emission of new roots, allowing better use of soil volume.

Our models were impacted by climatic factors, especially the water deficit during the 2013/2014 growing season. This result indicated that hyperspectral data on nitrogen concentration are strongly susceptible to variations in environmental factors. The water deficit in the 2013/2014 growing season was severe, and at the time of data collection in Piracicaba, the accumulated monthly rainfall in February (Figure 2B) was only 31% of the historical average.

High nitrogen levels in the soil due to mineralization of organic matter from the previous season can affect the relationship between the leaf nitrogen and biomass and;consequently, the spectral readings (SANTANA et al., 2020). Critical nitrogen concentration in dry biomass is the minimum amount of nitrogen required for maximum crop growth. If nitrogen supply is not limited, nitrogen concentration generally decreases as the dry biomass increases during the growing season. This allometric relationship can be expressed using a negative power function designated dilution curve (TILLY & BARETH, 2019).

The adopted methodology efficiently selected wavelengths and generated models for estimating nitrogen concentrations from the hyperspectral canopy data. Results of the calibration and validation models showed the potential of using reflectance data to monitor variations in leaf nitrogen levels in sugarcane.

CONCLUSION

The sPLS methodology facilitated the selection of spectral regionsstrongly associated with the leaf nitrogen content in sugarcane crops. The most suitable wavelengths were reportedin the following spectral regions: visible (700–705 nm), red-edge (710–720 nm) (used by all study models), near-infrared (725, 925, 955, and 980 nm), and short-wave infrared (1355, 1420, 1595, 1600, 1605, and 1610 nm).In line with previous studies, the number of variables needed to predict the leaf nitrogen concentration in sugarcane was small, and the calibrated models (general and site-specific) used two to four wavelengths.

Our results demonstrated that hyperspectral data are strongly influenced by several factors, including the crop environment, cultivar genotype, and climatic factors.

Identifying specific spectral regions allows users to select and use hyperspectral data to monitor nitrogen levels in sugarcane.This should be carried out beforeusing sensors with predetermined wavelengths to complement laboratory data and improve crop management.

ACKNOWLEDGMENTS

This research was supported by grants from the Financiadora de Estudos e Projetos (FINEP), allowing the purchase of a radiometer through the PROSENSAP project, and from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (Project Nº 2013/22435-9). We are grateful to Dr. HeitorCantarella and his research team at the Instituto Agronômico de Campinas (IAC), who, through partnerships, enabled the development of this field study (Thematic Project BIOEN/FAPESP 2008/56147-1). Andwasfinanced in partbythe Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brasil - Financecode 001.

DECLARATION OF CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

AUTHORS' CONTRIBUTIONS

All authors contributed equally to the design and writing of the manuscript. Additionally, all authors critically revised the manuscript for important intellectual content.

REFERENCES

ABDEL-RAHMAN, E.M.; AHMED, F.B.; ISMAIL, R. Random forest regression and spectral band selection for estimating sugarcane leaf nitrogen concentration using EO-1 Hyperion hyperspectral data. **International Journal of Remote Sensing**, v. 34, n. 2, p. 712-728, 2013. Available from: <https://doi.org/10.1080/01431161.2012.713142>. Accessed: Mar. 18, 2019. doi: 10.1080/01431161.2012.713142.

ABDEL-RAHMAN, E.M. et al. Estimation of sugarcane leaf nitrogen concentration using in situ spectroscopy. **International Journal of Applied Earth Observation and Geoinformation**, v. 12, Supplement 1, p. S52-S57, 2010. Available from: <https://doi.org/10.1016/j.jag.2009.11.003>. Accessed: Sep. 16, 2019. doi: 10.1016/j.jag.2009.11.003.

ABDEL-RAHMAN, E.M. et al. Comparison of partial least squares (PLS) and sparse PLS regressions for predicting yield of Swiss chard grown under different irrigation water sources using hyperspectral data. **Computers and Electronics in Agriculture**, v. 106, p. 11-19, 2014. Available from: <https://doi.org/10.1016/j.compag.2014.05.001>. Accessed: Jul. 16, 2018. doi: 10.1016/j.compag.2014.05.001.

ALI, M.M. et al. Leaf Nitrogen Determination Using Non-Destructive Techniques – A Review. **Journal of Plant Nutrition**, p. 928-953, 2016. Available from: <https://doi.org/10.1080/01904167.2016.1143954>. Accessed: Apr. 08, 2020. doi: 10.1080/01904167.2016.1143954.

AMARAL, L.R. et al. Comparison of crop canopy reflectance sensors used to identify sugarcane biomass and nitrogen status. **Precision Agriculture**, v. 16, n. 1, p. 15-28, 2015. Available from: <http://afurlan.com.br/lap/cp/assets/layout/files/articles/art-64_i.pdf>. Accessed: Feb. 21, 2020. doi: 10.1007/s11119-014-9377-2.

CECCATO, P.; et al., Detecting vegetation leaf water content using reflectance in the optical domain. **Remote Sensing of Environment**, v. 77, p. 22–33, 2001. Available from: <https://doi.org/10.1016/S0034-4257(01)00191-2>. Accessed: Mar. 22, 2020. doi: 10.1016/S0034-4257(01)00191-2.

CHAVES, M.M. et al. Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. **Annals of Botany**, v. 103, n. 4, p. 551-560, 2009. Available from: <https://doi.org/10.1093/aob/mcn125>. Accessed: Nov. 02, 2019. doi: 10.1093/aob/mcn125.

CHUNG, D.; CHUN, H.; KELES, S. **An introduction to the "spls" Package**, Version 1.0. R Foundation for Statistical Computing: Vienna, Austria, 2012. Available from: <http://cran.r-project.org/web/packages/spls/vignettes/spls-example.pdf>. Accessed: Nov. 26, 2019.

COIMBRA, J. L. M. et al. Conseqüências da multicolinearidade sobre a análise de trilha em canola. **Ciência Rural**, v. 35, p. 347-352, 2005. Available from: <https://doi.org/10.1590/S0103-84782005000200015>. Accessed: Jun. 22, 2021. doi: 10.1590/S0103-84782005000200015.

DARVISHZADEH, R. et al. LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. **ISPRS journal of photogrammetry and remote sensing**, v. 63, n. 4, p. 409-426, 2008. Available from: <https://doi.org/10.1016/j.isprsjprs.2008.01.001>. Accessed: Oct. 27, 2020. doi: 10.1016/j.isprsjprs.2008.01.001.

DEMATTÊ, J.A.M. el al., .Hyperspectral remote sensing as an alternative to estimate soil attributes. **Revista Ciência Agronômica**, v. 46, n. 2, p. 223–232, 2015. Available from: <https://doi.org/10.5935/1806-6690.20150001>. Accessed: Feb. 25, 2021. doi: 10.5935/1806-6690.20150001.

FAOSTAT. 2019. Statistical Database of Food and Agriculture Organization of the United Nations (FAO). Available from: <http://faostat.fao.org/site/291/default.aspx>. Accessed: May. 29,2020.

FIORIO, P.R. et al. Spectral behavior of sugarcane's leaves under the water deficit conditions. **Irriga**, v. 23, n. 3, p. 609-621, 2018. Available from: <https://doi.org/10.15809/irriga.2018v23n3p609-621>. Accessed: Jun. 08, 2020. doi: 10.15809/irriga.2018v23n3p609-621.

GAVA, G.J.C. et al. Growth and accumulation of nitrogen by sugarcane cultivated in soil covered with cane trash. **Pesquisa Agropecuária Brasileira**, v. 36, n. 11, p. 1347-1354, 2001. Available from: <https://doi.org/10.1590/S0100-204X2001001100004>. Accessed: Oct. 08, 2018. doi: 10.1590/S0100-204X2001001100004.

GITELSON, A. A. et al. Remote estimation of canopy chlorophyll content in crops. **Geophysical Research Letters**, v. 32, n. 8, 2005. Available from: <https://doi.org/10.1029/2005GL022688>. Accessed: Jun. 12, 2021. doi: 10.1029/2005GL022688.

HENNESSY, A. et al Hyperspectral classification of plants: a review of waveband selection generalizability. **Remote Sensing**, v. 12, n. 113, p. 1-27, 2020. Available from: <http://dx.doi.org/10.3390/rs12010113>. Accessed: Jun. 02, 2020. doi: 10.3390/rs12010113.

INOUE, Y et al., . Simple and robust methods for remote sensing of canopy chlorophyll content: a comparative analysis of hyperspectral data for different types of vegetation. **Plant, cell & environment**, v. 39, n. 12, p. 2609-2623, 2016. Available from: <https://doi.org/10.1111/pce.12815>. Accessed: Apr. 01, 2021. doi: 10.1111/pce.12815.

ISAKSSON, T.; NAES, T. The Effect of Multiplicative Scatter Correction and Linearity Improvement in NIR Spectroscopy. **Applied Spectroscopy**, v. 42, p. 1273-1284, 1988. Available from: <https://doi.org/10.1366/0003702884429869>. Accessed: Apr. 01, 2021. doi: 10.1366/0003702884429869.

IZNAGA, A. C. et al. VIS/NIR spectroscopic measurement of selected soil fertility parameters of Cuban agricultural Cambisols. **Biosystems Engineering**, v. 125, p. 105–121, 2014. Available from: <https://doi.org/10.1016/j.biosystemseng.2014.06.018>. Accessed: Oct. 18, 2020. doi: 10.1016/j.biosystemseng.2014.06.018.

JENSEN, J. R**. Remote sensing of the environment: An earth resource perspective**. 2. New Jersey: Pearson Education India, 2009.

LISBOA, I. P. et al. Prediction of sugarcane yield Based on NDVI and Concentration of Leaf-Tissue Nutrients in Fields Managed with Straw Removal. **Agronomy**, v. 8, n. 9, p. 196, 2018. Available from: <http://dx.doi.org/10.3390/agronomy8090196>. Accessed: Mar. 19, 2020. doi: 10.3390/agronomy8090196.

MAHAJAN, G.R. et al. Using hyperspectral remote sensing techniques to monitor nitrogen, phosphorus, sulphur and potassium in wheat (Triticum aestivum L.). **Precision Agriculture**, v. 15, n. 5, p. 499-522, 2014. Available from: <http://dx.doi.org/10.1007/s11119-014-9348-7>. Accessed: Oct. 08, 2018. doi: 10.1007/s11119-014-9348-7.

MALAVOLTA, E. et al. **Avaliação do estado nutricional das plantas**: **princípios e aplicações**. ed. Piracicaba: Associação Brasileira para pesquisa da Potassa e do Fosfato, 1997. 319p. Accessed: Oct. 25, 2018.

MALLEY, D.F. et al. Application in analysis of soils. *In*: ROBERTS, C.A.; WORKMAN JR., J.W.; REEVES III, J.B.(Ed.). **Near-infrared spectroscopy in agriculture**. Madison: American Society of Agronomy, 2004. p. 729-784. Accessed: Oct. 08, 2018.

MARTINS, J. A. et al. Potential use of hyperspectral data to monitor sugarcane nitrogen status. **Acta Scientiarum**. **Agronomy**, v. 43, e47632, 2021. Available from: <https://doi.org/10.4025/actasciagron.v43i1.47632>. Accessed: Feb. 25, 2021. doi: 10.4025/actasciagron.v43i1.47632.

MIN, M.; LEE, W.S. Determination of significant wavelengths and prediction of nitrogen content for citrus. **Transaction of the ASAE**, v. 48, n. 2, p. 455-461, 2005. Available from: <https://doi.org/10.13031/2013.18308>. Accessed: Feb. 25, 2021. doi: 10.13031/2013.18308.

MIPHOKASAP, P. et al. Estimating Canopy Nitrogen Concentration in Sugarcane Using Field Imaging Spectroscopy. **Remote Sensing**, v. 4, n. 12, p. 1651-1670, 2012. Available from: <http://dx.doi.org/10.3390/rs4061651>. Accessed: Oct. 08, 2018. doi: 10.3390/rs4061651.

MIPHOKASAP, P.; WANNASIRI, W. Estimations of nitrogen concentration in sugarcane using hyperspectral imagery. **Sustainability**, v. 10, p. 1266, 2018. Available from: <http://dx.doi.org/10.3390/su10041266>. Accessed: May. 30, 2020. doi: 10.3390/su10041266.

MOKHELE, T.A.; AHMED, F.B. Estimation of leaf nitrogen and silicon using hyperspectral remote sensing. **Journal of Applied Remote Sensing**, v. 4, n. 1, p. 043560, 2010. Available from: <http://dx.doi.org/10.1117/1.3525241>. Accessed: Oct. 08, 2018. doi: 10.1117/1.3525241

MONTEIRO, L.; SENTELHAS, P. Sugarcane yield gap: can it be determined at national level with a simple agrometeorological model? **Crop and Pasture Science**, v. 68, n. 3, p. 272-284, 2017. Available from: <https://doi.org/10.1071/CP16334>. Accessed: Oct. 08, 2018. doi: 10.1071/CP16334.

MORIYA, E. A. S. et al. A study on the effects of viewing angle variation in sugarcane radiometric measures. **Boletim de Ciências Geodésicas**, v. 24, n. 1, p. 85-97, 2018. Available from: <https://doi.org/10.1590/s1982-21702018000100007>. Accessed: Oct. 24, 2020.

MULLA, D. J. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. **Biosystems Engineering**, v. 114, n. 4, p. 358–371, 2013. Available from: <http://dx.doi.org/10.1016/j.biosystemseng.2012.08.009>. Accessed: Oct. 12, 2020. doi: 10.1016/j.biosystemseng.2012.08.009.

MUTANGA, O.; SKIDMORE, A.K. Red edge shift and the quality of tropical grass canopies. **ISPRS Journal of Photogrammetry and Remote Sensing**. 2007, v. 62, n. 1, 34–42. Available from: <https://doi.org/10.1016/j.isprsjprs.2007.02.001>. Accessed: May. 28, 2020. doi: 10.1016/j.isprsjprs.2007.02.001.

PEERBHAY, K.Y.; MUTANGA, O.; ISMAIL, R. Does simultaneous variable selection and dimension reduction improve the classification of Pinus forest species? **Journal of Applied Remote Sensing**, v. 8, n. 1, p. 85194-85194, 2014. Available from: <https://doi.org/10.1117/1.JRS.8.085194>. Accessed: Feb. 28, 2021. doi: 10.1117/1.JRS.8.085194.

RAIJ, B.V.; CANTARELLA, H. Outras culturas industriais. *In*: RAIJ, B. V. *et al*. (Ed.). **Recomendações de adubação e calagem para o Estado de São Paulo**. 2. Campinas: IAC, v. 100, 1997. p. 8-13. Available from: <http://www.academia.edu/download/56601012/boletim_tecnico_100_pdf.pdf>. Accessed: Oct. 08, 2018.

ROSA, H.J.A. et al. Sugarcane response to nitrogen rates, measured by a canopy reflectance sensor. **Pesquisa Agropecuária Brasileira**, v. 50, n. 9, p. 840-848, 2015. Available from: <https://doi.org/10.1590/S0100-204X2015000900013>. Accessed: Oct. 08, 2018. doi: 10.1590/S0100-204X2015000900013.

SANTANA, A.C.A. et al. Critical nitrogen dilution curves and productivity assessments for plant cane. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 24, n. 4, p. 244-251, 2020. Available from: <http://dx.doi.org/10.1590/1807-1929/agriambi.v24n4p244-251>. Accessed: Jun. 10, 2020. doi: 10.1590/1807-1929/agriambi.v24n4p244-251.

SENANAYAKE, I.P.; et al., . An in-situ data based model to downscale radiometric satellite soil moisture products in the Upper Hunter Region of NSW, Australia. **Journal of Hydrology**, v. 572, p. 820-838, 2019. Available from: <https://www.scopus.com/record/display.uri?eid=2-s2.0-85063260068&origin=inward&txGid=4cb1bb9b70e377847cd5cf5b725c0a38>. Accessed: Oct. 20, 2020. doi: 10.1016/j.jhydrol.2019.03.014.

SHAPIRO, S.S.; WILK, M.B. An analysis of variance Test for Normality (Complete Samples). **Biometrika**, v. 52, n. 3/4, p. 591-611, 1965. Available from: <https://www.jstor.org/stable/pdf/2333709.pdf?casa_token=w_ecvlIzYakAAAAA:w5uy_rf1Yx85ZPhho7NnyUfCMwhIwfSfOhu7t_dHymEnyiXlZzCdvsjarNfxbpl9bu7JR04cjldetZNktS0kAwThRr_a9J4Q_fiCu1Jk0klTa4olA>. Accessed: Feb. 23, 2019.

STRACHAN, I.B.; PATTEY, E.; BOISVERT, J.B. Impact of nitrogen and environmental conditions on corn as detected by hyperspectral reflectance. **Remote Sensing of Environment**, v. 80, n. 2, p. 213-224, 2002. Available from: <https://doi.org/10.1016/S0034-4257(01)00299-1>. Accessed: Mar. 23, 2021. doi: 10.1016/S0034-4257(01)00299-1.

TAVARES, T.R. et al. Effects of storage on vis-NIR-SWIR reflectance spectra of Mombassa grass leaf samples. **Ciência Rural**, v. 50, n. 3, p. e20190587, 2020. Available from: <https://doi.org/10.1590/0103-8478cr20190587>. Accessed: Mar. 03, 2020. doi: 10.1590/0103-8478cr20190587.

THENKABAIL, P. S. et al. Hyperspectral remote sensing of vegetation and agricultural crops: knowledge gain and knowledge gap after 40 years of research. In: THENKABAIL, P. S.; LYON, J. G. (Ed.). **Hyperspectral remote sensing of vegetation**. 1. Boca Raton: CRC Press, 2010. p. 705. Available from: <http://oar.icrisat.org/9223/1/06_PERS-hyperspectral-special-issue-%28press%29.pdf>. Accessed: Oct. 18, 2020.

TILLY, N.; BARETH, G. Estimating nitrogen from structural crop traits at field scale—A novel approach versus spectral vegetation indices. **Remote Sensing**, v. 11, n. 17, p. 2066, 2019. Available from: <https://doi.org/10.3390/rs11172066>. Accessed: Mar. 03, 2020. doi: 10.3390/rs11172066.

VISCARRA ROSSEL, R.A. ParLeS: Software for chemometric analysis of spectroscopic data. **Chemometrics and Intelligent Laboratory Systems**, v. 90, n. 1, p. 72-83, 2008. Available from: <https://doi.org/10.1016/j.chemolab.2007.06.006>. Accessed: Oct. 08, 2018. doi: 10.1016/j.chemolab.2007.06.006.

ZHAO, B. et al. Exploring new spectral bands and vegetation indices for estimating nitrogen nutrition index of summer maize. **European Journal of Agronomy**, v. 93, n. 1, p. 113-125, 2018. Available from: <https://doi.org/10.1016/j.eja.2017.12.006>. Accessed: Oct. 08, 2018. doi: 10.1016/j.eja.2017.12.006.

ZHAO, C. et al. Relationships of leaf nitrogen concentration and canopy nitrogen density with spectral features parameters and narrow-band spectral indices calculated from field winter wheat (*Triticum aestivum L.*) spectra. **International Journal of Remote Sensing**, v. 33, n. 11, p. 3472-3491, 2012. Available from: <https://doi.org/10.1080/01431161.2011.604052>. Accessed: Oct. 08, 2018. doi: 10.1080/01431161.2011.604052.