SEMINA
Ciências Agrárias

DOI: 10.5433/1679-0359.2023v44n2p859

Development and validation of regression models from NIR spectra to predict the composition of sugarcane, soybean meal, and cornmeal

Construção e validação de modelos de regressão a partir de espectros NIR para predição da composição da cana-de-açúcar, farelo de soja e fubá de milho

Nathália Veloso Trópia[1]*; Flávia Adriane de Sales Silva[2]; Dhones Rodrigues Andrade[1]; Éllem Maria de Almeida Matos[1]; Fernando Alerrandro Andrade Cidrini[3]; Karen Melo Borges[4]; Yuri Cesconetto Ebani[3]; Jussara Valente Roque[5]; Diego Zanetti[6]; Sebastião de Campos Valadares Filho[7]

Highlights

Portable NIR-generated spectra estimate the composition of ruminant feeds.
The composition of sugarcane, soybean meal, and cornmeal is estimated by NIR.
Nutrient concentration affects the quality of composition prediction by NIR.

Abstract

This study aimed to develop and assess regression models for predicting the chemical composition of sugarcane, soybean meal, and cornmeal using portable near-infrared (NIR) spectroscopy combined with chemometric techniques. A total of 95 sugarcane samples, 92 soybean meal samples, and 120 cornmeal samples were used. The samples were ground, and NIR spectra were obtained for each sample. Reference values were determined through conventional chemical analysis. Partial least squares regression and leave-one-out cross-validation were employed to construct the models. Models with the lowest root mean

[1] PhD Students in Post-Graduate Program in Animal Science, Department of Animal Science, Universidade Federal de Viçosa, UFV, Viçosa, MG, Brazil. E-mail: nathaliatropia@gmail.com; dhonesandrade2@gmail.com; ellem.matos@ufv.br
[2] Researcher, Department of Animal Science, UFV, Viçosa, MG, Brazil. E-mail: flaviasales_pf@hotmail.com
[3] Undergraduate Students in Animal Science, Department of Animal Science, Viçosa, UFV, MG, Brazil. E-mail: fernandocidrini@gmail.com; yuriebani@gmail.com
[4] M.e Student in Post-Graduate Program in Animal Science, Department of Animal Science, UFV, Viçosa, MG, Brazil. E-mail: karenmelozoo@gmail.com
[5] Researcher, Institute of Chemistry, Universidade Federal de Goiás, UFG, Goiânia, GO, Brazil. E-mail: jussara_roque@ufg.br
[6] PhD Prof., Department of Animal Science, Instituto Federal do Sul de Minas Gerais, IFSMG, Machado, MG, Brazil. E-mail: diego.zanetti@ifsuldeminas.edu.br
[7] PhD Prof., Department of Animal Science, UFV, Viçosa, MG, Brazil. E-mail: scvfilho@ufv.br
* Author for correspondence

Received: Nov. 22, 2022 Approved: June 13, 2023

squared error in cross-validation were further validated externally. The goodness-of-fit of the models was evaluated by comparing the predicted values with those obtained through conventional laboratory methods. The constructed models properly estimated all constituents evaluated for sugarcane, soybean meal, and cornmeal ($P \geq 0.056$). The models developed for predicting the contents of samples oven-dried at 55 °C (ADS) and 105 °C (ODS), total dry matter (DM), organic matter (OM), neutral detergent fiber (NDF), NDF corrected for ash and protein (NDFap), neutral detergent insoluble protein (NDIP), acid detergent fiber (ADF), crude protein (CP), non-fiber carbohydrates (NFC), and total digestible nutrients (TDN) in sugarcane; ODS, OM, NDF, ADF, indigestible NDF (iNDF), CP, TDN, and starch in soybean meal; and ODS and CP in cornmeal exhibited high accuracy and precision ($R^2 \geq 0.50$ and $CCC \geq 0.60$). However, the models developed for predicting the levels of neutral detergent insoluble ash (NDIA) in sugarcane; ether extract (EE) and NDIA in soybean meal; and NDF, iNDF, NDIA, NFC, and EE in cornmeal demonstrated accuracy but lacked precision ($R^2 \geq -0.04$ and $CCC \geq 0.03$). In conclusion, the portable NIR regression models provided accurate estimates and are therefore recommended for predicting the chemical composition of sugarcane, soybean meal, and cornmeal.
Key words: Chemometrics. Partial least squares regression. Spectroscopy.

Resumo

Objetivou-se desenvolver e avaliar modelos de regressão para a predição da composição química da cana-de-açúcar, farelo de soja e fubá de milho por NIR portátil aliado a técnicas quimiométricas. Foram utilizadas 95 amostras de cana-de-açúcar, 92 amostras de farelo de soja e 120 amostras de fubá de milho. Após a moagem das amostras, foi realizada aquisição dos espectros de cada amostra. Os valores referência foram obtidos através de análises químicas convencionais. Para construção dos modelos, foi utilizada a regressão por quadrados mínimos parciais e a validação cruzada *leave one out*. Os modelos com menor raiz quadrada do erro quadrático médio da validação cruzada foram submetidos a validação externa. Para avaliar a qualidade de ajuste dos modelos, os valores preditos foram comparados com os valores obtidos pelos métodos laboratoriais convencionais. Os modelos construídos estimaram corretamente todos os constituintes avaliados para a cana-de-açúcar, farelo de soja e fubá de milho ($P \geq 0,056$). Os modelos construídos para predição dos teores de amostra seca em estufa a 55°C (ASA) e a 105°C (ASE), matéria seca total (MS), matéria orgânica (MO), fibra insolúvel em detergente neutro (FDN), FDN corrigida para cinzas e proteína (FDNcp), proteína insolúvel em detergente neutro (PIDN), fibra insolúvel em detergente ácido (FDA), proteína bruta (PB), carboidratos não fibrosos (CNF) e nutrientes digestíveis totais (NDT) da cana-de-açúcar; ASE, MO, FDN, FDA, FDN indigestível (FDNi), PB, NDT e amido de farelo de soja; e ASE, PB do fubá de milho apresentaram elevada acurácia e precisão ($R^2 \geq 0,50$ e $CCC \geq 0,60$). Contudo os modelos construídos para predição dos teores de cinzas insolúveis em detergente neutro (CIDN) da cana-de-açúcar; extrato etéreo (EE) e CIDN do farelo de soja; e FDN, FDNi, CIDN, CNF e EE do fubá de milho foram acurados, porém pouco precisos ($R^2 \geq -0,04$ e $CCC \geq 0,03$). Conclui-se que os modelos de regressão por NIR portátil estimaram acuradamente e, portanto, são recomendados para estimar a composição química da cana-de-açúcar, farelo de soja e fubá de milho.
Palavras-chave: Espectroscopia. Quimiometria. Regressão por mínimos quadrados parciais.

Introduction

Sugarcane is a primary roughage source utilized in feedlots (Pinto & Millen, 2019; Silvestre & Millen, 2021) and Brazilian milk production systems (D. P. Silva et al., 2019b). This roughage source is valuable due to its low cost per ton of dry matter (DM), large availability, and ability to maintain its nutritional value unchanged during the dry season (Souza et al., 2015). Concentrate feed ingredients, such as soybean meal and cornmeal, are also essential for ruminant nutrition, particularly in intensive production systems (Pinto & Millen, 2019). These ingredients are produced on a large scale in Brazil (Companhia Nacional de Abastecimento [CONAB], 2021).

To formulate balanced diets, it is crucial to have accurate knowledge of the chemical composition of feed ingredients. Traditionally, this information is obtained through conventional chemical analyses, which provide estimates of the actual chemical composition. However, conventional analyses have limitations, including high costs, labor intensiveness, and time constraints that may render them infeasible. Additionally, these analyses are often destructive and environmentally harmful due to the use of various chemical reagents.

In this context, near-infrared (NIR) spectroscopy has emerged as an alternative to conventional chemical analysis methods. It has been successfully used to develop prediction models for the composition of feed ingredients for livestock (Thomson et al., 2018). Imported portable NIR devices are now available in the Brazilian market, enabling the application of this technology in the field. Nevertheless, it is important to consider that the chemical composition of feed ingredients may vary across regions due to factors such as soil type, variety/cultivar, precipitation, fertilizer use, and radiation. Therefore, prediction models generated in other countries may not be applicable to tropical conditions.

Thus, we hypothesize that portable NIR prediction models can replace conventional analysis methods for predicting the chemical composition of sugarcane, soybean meal, and cornmeal constituents. Therefore, the objective of this study was to develop and evaluate regression models utilizing portable NIR combined with chemometric techniques to predict the contents of oven-dried matter at 55 °C (ADS), oven-dried matter at 105 °C (ODS), total dry matter (DM), organic dry matter (OM), crude protein (CP), neutral detergent insoluble fiber (NDF), neutral detergent insoluble protein (NDIP), neutral detergent insoluble ash (NDIA), NDF corrected for ash and protein (NDFap), acid detergent fiber (ADF), acid detergent insoluble protein (ADIP), ether extract (EE), indigestible NDF (iNDF), non-fiber carbohydrates (NFC), lignin, starch, and total digestible nutrients (TDN) in sugarcane, soybean meal, and cornmeal.

Material and Methods

Sample collection and preparation

To compose the database, 95 sugarcane samples, 120 cornmeal samples, and 92 soybean meal samples were collected from rural properties, animal feed companies, and research institutions in different locations.

The 95 sugarcane samples consisted of different cultivars from different municipalities in the states of Minas Gerais (Porto Firme, Timóteo, Barra Longa, Viçosa, Coimbra, Divinnésia, Paula Candido, Ouro Preto, Felixlândia, Mariana, and Oratório). Of these, 59% included cultivar information, namely, RB097021 (2 samples), RB107221 (2 samples), RB107414 (2 samples), RB867515 (2 samples), RB966928 (2 samples), RB107264 (2 samples), RB087218 (2 samples), RB057310 (2 samples), RM107418 (2 samples), RB107277 (2 samples), RB107235 (2 samples), RB107070 (2 samples), RB097012 (2 samples), RB107306 (2 samples), RB107020 (2 samples), RB107382 (2 samples), RB107210 (2 samples), RB107247 (2 samples), RB107060 (2 samples), RB107224 (2 samples), RB037059 (2 samples), CTC4 (2 samples), CTC9001 (2 samples), RB107076 (2 samples), RB987935 (2 samples), RB991532 (2 samples), and RB037076 (2 samples).

The 120 cornmeal samples originated from different Brazilian states and municipalities, namely, Ceará (0.68% of the total samples; municipality: Sertão Central), Distrito Federal (0.68% of the total samples; municipality: Brasília), Goiás (13.70% of the total samples; municipalities: Mossâmedes, Caiapônia, Santa Fé de Goiás, and Goianira), Maranhão (0.68% of the total samples; municipality: Balsas), Minas Gerais (77.70% of the total samples; municipalities: Acaiaca, Aguanil, Alfenas, Alvinópolis, Barra Longa, Campo Belo, Coimbra, Contagem, Ervália, Fazenda Pimenta de Cima, Felixlândia, Formiga, Formoso, Granja Lago, Itatinga, Lagoa Dourada, Manga, Nazareno, Oratorios, Paraíba do Sul, Passos, Patos de Minas, Paula Cândido, Piranga, Porto Firme, Pratápolis, Rio Pomba, Santa Fé de Minas, Timóteo, Tupaciguara, Uberlândia, Unaí, and Viçosa), Pará (0.68% of the total of samples; municipality: Canoa do Pará), Piauí (0.68% of the total samples; municipality: Baixa grade), Paraná (5.48% of the total samples; municipality: Itatuba), and Rio de Janeiro (0.68% of total samples; municipality: Três Rios).

The 92 soybean meal samples originated from different Brazilian states and municipalities, namely, Minas Gerais (85.58% of the total samples; municipalities: Viçosa, Porto Firme, Piranga, Felixlândia, Coimbra, Alvinópolis, Formoso, Uberlândia, Primavera do Leste, Rio Pomba, and Oratório), Goiás (12.50% of the total samples; municipalities: Rio Verde, Palmeiras Goiás, and Anápolis), Piauí (0.96% of the total samples; municipality: Uruçuí), and Maranhão (0.96% of the total samples; municipality: Porto Franco). The samples were collected from different locations to ensure enough variation in the chemical composition for the development and evaluation of the models.

Once collected, the samples were frozen and immediately sent to the Ruminant Nutrition Laboratory (LabNUR) of the Federal University of Viçosa (UFV), where they were kept in a cold chamber (-10 ºC) for further laboratory analysis.

Sample composition

To obtain reference data, each sugarcane sample weighing approximately 500 g was dried at 55 ºC for 72 h in a forced-air oven. Subsequently, all feedstuffs were ground using a knife mill to particle sizes of 1 and 2 mm for further laboratory analysis. The following parameters were determined according to the specified methods: ADS, ODS, and DM (methods INCT G-001/2,

G-003/1); crude protein (CP) (method INCT N-001/2); mineral matter (MM) (method INCT M-001/2); ether extract (EE) (method INCT G-004/1); neutral detergent fiber (NDF) (method INCT F-001/2); acid detergent fiber (ADF) (method INCT F-003/2), along with the respective corrections for ash (neutral detergent insoluble ash, NDIA) and protein (neutral detergent insoluble protein, NDIP) and acid detergent insoluble protein, ADIP; methods INCT M-002/2, INCT M-003/2, INCT N-004/2, and INCT N-005/2, respectively); indigestible NDF (iNDF); and lignin (methods INCT F-008/2 and INCT F-005/02), as described by Detmann et al. (2021).

The organic matter (OM) content was calculated by difference using the following equation: OM = 100 - MM. The starch content was determined following the procedure outlined by B. C. Silva et al. (2019a). Non-fiber carbohydrates (NFC) were quantified according to Detmann et al. (2021) using the following formula: NFC = 100 - (%CP + %NDFap + %EE + %MM). The total digestible nutrient (TDN) content was calculated using equations proposed by Valadares et al. (2016): TDN = CPtd + NFCtd + NDFd + 2.25 × EEtd - FMTDN, in which CPtd, NFCtd, and EEtd represent the truly digestible fractions of CP, NFC, and EE, respectively; dNDF is the digestible fraction of NDF; FMTDN is the total fecal metabolic fraction using the value of 7.13, recommended for beef cattle; and 2.25 is the Atwater constant for the relationship between lipids and carbohydrates. The truly digestible fraction of CP (CPtd) was calculated using the formula below:

$$\mathrm{CP}_{td} = 0.95 \times (\mathrm{CP} - \mathrm{NDIP}) + \frac{kd}{kd+kp} \times \{NDIP \times \left[1 - e^{-(0.8188+1.1676 \times ADIP)}\right]\},$$

in which kd is the potentially digestible NDF degradation rate (pdNDF; h-1) and kp is the pdNDF ruminal passage rate (h-1). The kd and kp values estimated by CQBAL 4.0 (Valadares et al., 2018) were used. The truly digestible fraction of NFC (NFCtd) was calculated as NFCtd = 0.95 × NFC. The digestible fraction of NDF was calculated as displayed next:

$$\mathrm{NDF}_d = \left[\frac{kd}{kd+kp} \times \left(NDF_{ap} - iNDF\right)\right] \times IDF,$$

in which IDF is the intestinal digestibility correction factor (IDF = 1.12). The truly digestible fraction of EE was calculated as EEtd = 0.86 × EE.

Portable-NIR analyses

The samples processed to a particle size of 1 mm were thoroughly mixed, and each sample was divided into three sub-samples. These sub-samples were placed in Petri dishes in preparation for spectral reading. The spectra of the sub-samples were acquired using a portable near-infrared (NIR) spectrometer (ITPhotonics S.r.l., model poliSPECNIR 900-1700, Breganze, Italy) and recorded with the assistance of poliDATA software (ITPhotonics S.r.l., Breganze, Italy). Spectral readings were conducted in a controlled environment with a room temperature maintained at 21 °C. Three spectra were obtained for each sample, and the absorbance values were recorded in the range of 884.9 to 1702.9 nm, with measurements taken at intervals of 3.2 nm. For further analysis, the average of the three spectra for each sample was calculated. Figure 1 illustrates the spectra utilized in constructing the regression models.

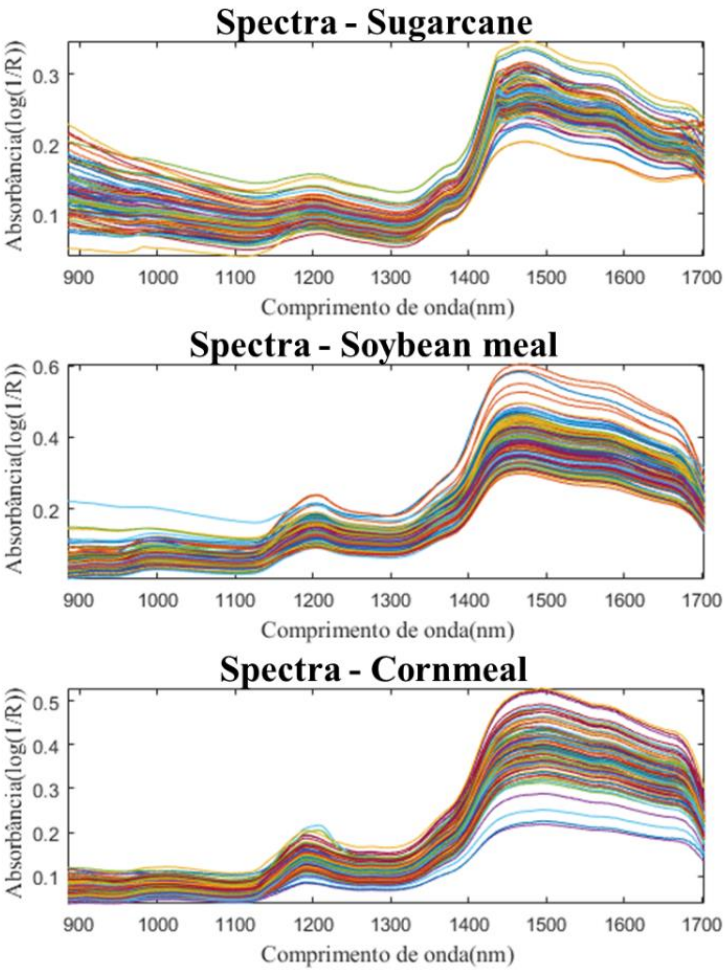

Figure 1. Spectral data used to build models of the chemical composition of sugarcane, cornmeal and soybean meal obtained by portable NIR.

Statistical analyses

For each feedstuff type examined, a matrix referred to as the X matrix was constructed from the mean spectra collected. The X matrix encompassed the independent variables of the dataset, where each row represented a sample and each column denoted the absorbance value at a specific wavelength. Simultaneously, a set of vectors was created, containing information regarding the chemical properties under investigation (dependent variables). These properties included ADS, ODS, DM, OM, CP, NDIP, NDIA, ADIP, EE, NDF, NDFap, iNDF, ADF, lignin, NFC, and TDN for sugarcane; and ODS, OM, CP, NDIP, ADIA, ADIP, EE, NDF, NDFap, iNDF, ADF, starch, NFC, and TDN for cornmeal and soybean meal.

Initially, outlier removal was performed using partial least squares (PLS) applied to the spectra after mean-centered pre-treatment. The sets of corresponding X and Y matrices were removed when they were detected as outliers based on graphical

analysis of Hotelling's T² vs Reduced Residual Q and Leverage vs. Student Residual Y tests, following the methods described by Peternelli et al. (2020) and Montgomery (2009). Subsequently, the dataset was divided into two subsets: a calibration set comprising 75% of the samples and an external evaluation set comprising the remaining 25%. The division of samples was accomplished using the Kennard-Stone algorithm (Kennard & Stone, 1969), which selects samples based on their distances. The average spectra files of the samples used for calibration were imported into PLS-toolbox 8.2.1 software, operating within the Matlab 2019b environment (Math Works, Natick, USA), for subsequent mathematical treatment and model development. The multivariate calibration approach employing PLS regression was employed to develop prediction models for each chemical constituent.

The number of latent variables was determined through leave-one-out cross-validation, selecting the number that yielded the lowest root mean squared error of cross-validation (RMSECV) value. The selection process also involved graphical evaluation of the number of latent variables in relation to RMSECV (Ferreira, 2015). Various pre-processing techniques, including multiplicative scatter correction, normalization, smoothing, first and second derivative, baseline correction, mean centering, standardized signal normalization, autoscaling, and detrend, were tested individually and in combination (Ferreira, 2015).

The models' performance was assessed using the RMSECV and the cross-validation correlation coefficient (RCV) parameters, calculated as follows:

$$RCV = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})(\hat{y}_i - \bar{y})}{\sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2 \overline{(y_i - \bar{y})^2}}$$

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

in which y_i represents the reference values in cross-validation; y is the mean of the reference values; yi represents the predicted values in cross-validation; y is the mean of the predicted values; and n is the number of samples in cross-validation.

The models with the lowest RMSECV values underwent external evaluation, comparing the chemical compositions as estimated using portable NIR and the values obtained through conventional laboratory methods. The Model Evaluation System (Tedeschi, 2006) was employed for this comparison. The predicted values were evaluated against the observed values using a linear regression model: y = β0 + β1 × X, in which X represents the predicted values, y represents the observed values, and β0 and β1 are the intercept and slope, respectively. The regression was evaluated based on the hypotheses of H0: β0 = 0 and H0: β1 = 1, with Ha being the alternative hypothesis. Models were considered good estimators when the regression's intercept and slope between predicted and observed values were equal to zero and one, respectively. The goodness-of-fit of the calibration models was further evaluated using the coefficient of determination (R^2), the concordance correlation coefficient or reproducibility index (CCC), and the mean squared error of prediction (Tedeschi, 2006) and its components: bias (SB), magnitude of random fluctuation (MaF), and random fluctuation of the model (MoF); Kobayashi and Salam (2000). Models were classified as having high

precision (R^2 and/or CCC ≥ 0.6), intermediate precision (0.4 ≤ R^2 and/or CCC < 0.6), or low precision (R^2 and/or CCC < 0.4) or precision and accuracy (CCC), respectively.

Results and Discussion

The chemical compositions of the feedstuffs examined in this study were found to be in line with the findings reported by Valadares et al. (2018) for samples of sugarcane, soybean meal, and cornmeal collected throughout Brazil.

Calibration

Table 1 shows the results obtained from conventional chemical analyses of sugarcane, cornmeal, and soybean meal in the calibration and external evaluation sets of the prediction models for all constituents studied, along with the sample sizes in each set.

Among the commonly employed pre-treatments for model development, the second derivative method was utilized in seven models for predicting the chemical composition of sugarcane and cornmeal. Smoothing was applied in seven models for predicting the chemical composition of sugarcane, while autoscaling alone or in combination with multiplicative scatter correction was used in 11 and eight models, respectively, for predicting the chemical composition of soybean meal and cornmeal (Table 2). These mathematical treatments consistently yielded lower RMSECV values and higher RCV values compared with the untreated models.

The second derivative treatment was frequently employed in models generated to predict the chemical composition of sugarcane (DM, OM, ADIP, ADF, CP, lignin, and TDN) and cornmeal (ODS, NDFap, ADIP, NDIA, NFC, EE, and TDN) (sugarcane = 44% and cornmeal = 46%). Additionally, the smoothing treatment was commonly used in models predicting the chemical composition of sugarcane (ADS, OM, NDF, NDFap, CP, NFC, and TDN) (44%). Multiplicative scatter correction transformation was applied to 30% of the models for predicting the chemical composition of soybean meal (ODS, iNDF, NDIA, CP, and starch) and 20% of the models for cornmeal (ODS, OM, and CP).

According to Ferreira (2015), the use of one or more combined spectral treatments during modeling is a regular practice in the development of NIR prediction models, since some transformations significantly reduce the observed errors. Second derivative pre-treatments, smoothing and multiplicative scattering correction are necessary when disturbances caused by noise are significant in the collected spectrum (stochastic contributions; Ciurczak et al., 2021). The presence of noise in the collected spectra may have been necessary due to reasons inherent to the NIR instrument, sampling, or the effects of physical phenomena (Ferreira, 2015). Disturbances can be corrected by correcting the baseline slope of the spectra, second derivative (Ferreira, 2015), reducing baseline and multiplicative spectral variations and, consequently, preserving the spectral band shapes, multiplicative scatter correction (Ciurczak et al., 2021) or eliminating high-frequency noises, smoothing (Ozaki et al., 2021).

Table 1
Descriptive statistics of data used to develop and validate models to predict the chemical composition of sugarcane, soybean meal, and cornmeal

Items[1]		Sugarcane					Soybean meal					Cornmeal				
	Set[2]	N[3]	Average	Minimum	Maximum	SD[4]	N	Average	Minimum	Maximum	SD	N	Average	Minimum	Maximum	SD
ADS	Cal	66	28.06	15.74	36.36	4.82	-	-	-	-	-	-	-	-	-	-
	Val	22	27.91	16.33	35.2	4.78	-	-	-	-	-	-	-	-	-	-
ODS	Cal	66	92.85	87.98	96.29	2.17	69	88.90	86.56	92.25	0.89	86	87.56	82.10	89.37	1.12
	Val	22	92.78	88.74	96.22	1.96	23	88.81	87.05	90.71	1.02	29	87.51	84.47	90.33	1.25
DM	Cal	65	25.52	14.66	32.83	4.32	-	-	-	-	-	-	-	-	-	-
	Val	21	26.42	18.85	31.59	3.68	-	-	-	-	-	-	-	-	-	-
OM	Cal	71	96.59	92.05	98.69	7.78	69	92.81	92.07	93.36	0.32	90	1.25	0.83	2.06	0.20
	Val	24	97.09	93.86	98.86	7.62	23	92.83	92.18	93.39	0.35	30	1.31	0.82	2.09	0.27
NDF	Cal	61	49.91	35.85	66.16	6.70	59	15.36	11.06	15.36	1.89	69	98.77	97.94	99.25	1.39
	Val	21	48.64	33.78	63.05	5.71	20	15.70	12.73	15.70	2.31	23	98.71	98.05	99.31	1.34
NDFap	Cal	64	48.35	34.69	61.36	3.04	63	12.96	9.91	12.96	1.80	77	13.97	10.17	15.98	1.35
	Val	21	48.79	36.64	58.47	3.54	21	13.02	9.06	13.02	1.94	26	14.45	11.49	16.27	1.39
iNDF	Cal	63	23.39	18.31	32.16	0.09	54	1.40	0.90	1.40	0.45	86	12.42	9.30	14.97	0.30
	Val	21	22.87	17.47	29.8	0.10	16	1.46	0.91	1.46	0.44	29	12.69	9.74	14.72	0.23
ADIP	Cal	50	0.33	0.23	0.66	0.42	60	0.27	0.16	0.27	0.11	85	1.56	1.04	2.47	0.11
	Val	17	0.34	0.16	0.50	0.38	20	0.29	0.16	0.29	0.10	28	1.63	1.16	2.14	0.11
NDIP	Cal	39	0.58	0.13	1.16	0.22	65	2.71	0.91	2.71	1.45	73	0.32	0.11	0.62	0.24
	Val	13	0.54	0.10	1.28	0.19	17	2.24	1.02	2.24	0.88	24	0.34	0.17	0.56	0.25
NDIA	Cal	58	0.62	0.12	1.07	4.74	64	0.54	0.25	0.54	0.17	84	1.29	0.80	1.81	0.16
	Val	20	0.62	0.43	1.10	4.41	21	0.51	0.28	0.51	0.12	28	1.27	0.78	1.74	0.24
ADF	Cal	52	29.49	20.91	40.67	0.74	67	6.65	4.06	6.65	1.11	79	0.59	0.11	1.20	0.58
	Val	18	30.36	23.90	42.87	0.62	22	6.98	4.27	6.98	1.22	26	0.58	0.15	1.13	0.72
CP	Cal	66	2.90	1.55	5.31	8.88	65	51.66	49.09	51.66	1.29	74	0.57	0.26	1.13	0.43
	Val	22	2.69	1.40	3.61	8.75	16	51.62	49.86	51.62	0.95	25	0.66	0.33	1.36	0.45
NFC	Cal	67	44.47	23.15	61.13	0.47	67	26.08	20.38	26.08	2.51	84	3.20	2.12	4.58	2.19
	Val	22	43.71	26.97	61.75	0.51	22	26.59	20.60	26.59	2.36	28	3.38	2.07	4.73	2.46

continue...

continuation...

EE	Cal	48	1.76	1.06	2.99	0.70	66	1.95	1.02	1.95	0.62	54	8.76	8.07	9.99	0.91
	Val	16	1.59	1.01	2.70	0.72	16	1.75	1.38	1.75	0.34	18	8.82	7.96	10.00	0.79
Lignin	Cal	52	4.44	3.14	5.95	0.70	58	5.29	3.00	5.29	1.36	52	73.06	67.27	76.60	2.84
	Val	17	4.17	3.05	5.25	0.72	16	5.88	3.62	5.88	1.18	18	73.21	68.14	76.78	3.28
Starch	Cal	-	-	-	-	-	58	79.39	77.01	79.39	1.36	52	4.26	2.54	5.90	2.84
	Val	-	-	-	-	-	16	78.98	77.42	78.98	1.18	18	4.06	2.71	5.56	3.28
TDN	Cal	53	62.16	55.38	68.1	3.64	63	15.36	11.06	15.36	1.35	77	71.95	65.72	76.47	1.12
	Val	18	62.15	55.02	66.0	3.14	21	15.70	12.73	15.70	1.17	26	72.47	65.38	78.61	1.43

[1]ADS= oven-dried matter at 55 °C, ODS = oven-dried matter at 105 °C, DM = total dry matter, OM = organic dry matter, NDF = neutral detergent insoluble fiber, NDFap = NDF corrected for ash and protein, iNDF = indigestible NDF, ADIP = acid detergent insoluble protein, NDIP = neutral detergent insoluble protein, NDIA = neutral detergent insoluble ash, ADF = acid detergent fiber, CP = crude protein, NFC = non-fiber carbohydrates, EE = ether extract, NDT = total digestible nutrients; [2]Cal = calibration dataset and Val = validation dataset; [3]number of evaluated samples e [4]Standard deviation.

Table 2
Information on data sets used to develop and the respective performance parameters of PLS models to predict the chemical composition of sugarcane, soybean meal, and cornmeal

Parameters[1]	Sugarcane					Soybean meal					Cornmeal				
	Trat.[2]	MS[3]	nVL[4]	RMSECV[5]	RVC[6]	Trat.	MS	nVL	RMSECV	RVC	Trat.	MS	nVL[4]	RMSECV	RVC
ADS	Det + Su	88x256	10	2.38	0.87	-	-	-	-	-	-	-	-	-	-
ODS	Det + Norm	88x256	9	0.99	0.89	Au + DMC	92x256	10	0.61	0.54	2nd d + DMC	115x256	9	0.44	0.92
DM	DMC + 2nd d	86x256	8	2.14	0.87		-	-	-	-	-	-	-	-	-
OM	Su + 2nd d	95x256	10	0.63	0.93	Au	92x256	11	0.21	0.57	DMC + Det	120x256	5	0.19	0.35
NDF	SNV + Su	82x256	8	2.32	0.96	Au + Det	79x256	6	1.79	0.67	Norm+ LB	92x256	6	1.28	0.44
NDFap	Su + DMC	85x256	10	2.15	0.95	Norm+ Au	84x256	3	1.49	0.56	2nd d + Su	103x256	9	1.35	0.34
iNDF	CM + 1st d	84x256	7	2.28	0.67	Au + DMC	70x256	5	0.31	0.53	Au + SNV	115x256	3	0.27	0.43
ADIP	2nd d + Det	67x256	3	0.06	0.69	Su + 1st d	89x256	8	0.58	0.85	Su + 2nd d	113x256	6	0.09	0.60
NDIP	Norm + 1st d	52x256	7	0.35	0.62	Au + SNV	82x256	11	1.27	0.27	Det + LB	97x256	4	0.21	0.49
NDIA	Au + LB	78x256	3	0.20	0.37	2nd d + DMC	85x256	3	0.17	0.26	2nd d + LB	112x256	2	0.16	0.23
ADF	CM + 2nd d	70x256	8	1.48	0.95	SNV + Su	91x256	10	0.60	0.70	Au + Norm	105x256	7	0.45	0.64
CP	Su + 2nd d	88x256	10	0.39	0.85	DMC + Su	81x256	1	1.23	0.09	DMC + Su	99x256	9	0.38	0.51
NFC	Su + SNV	89x256	8	2.77	0.95	CM + LB	89x256	7	2.00	0.39	2nd d + Au	112x256	2	1.93	0.49
EE	CM + Norm	64x256	3	0.46	0.24	Au + 2nd d	82x256	3	0.56	0.21	2nd d + Au	72x256	2	0.85	0.49

continue...

SEMINA Ciências Agrárias

continuation...

Lignin	2nd d + Au	69x256	3	0.58	0.56	-	-	-	-	-	-	-	-	-	-
Starch	-	-	-	-	-	Au + DMC	74x256	9	1.08	0.40	Det + 1st d	70x256	4	2.57	0.44
TDN	Su + 2nd d	71x256	7	1.06	0.96	Su + 1st d	84x256	7	0.72	0.85	2nd d + Det	103x256	4	1.02	0.45

[1]ADS= oven-dried matter at 55 °C, ODS = oven-dried matter at 105 °C, DM = total dry matter, OM = organic dry matter, NDF = neutral detergent insoluble fiber, NDFap = NDF corrected for ash and protein, iNDF = indigestible NDF, ADIP = acid detergent insoluble protein, NDIP = neutral detergent insoluble protein, NDIA = neutral detergent insoluble ash, ADF = acid detergent fiber, CP = crude protein, NFC = non-fiber carbohydrates, EE = ether extract, NDT = total digestible nutrients,
[2]Det = detrend, Su = smoothing, Norm = normalization, DMC = multiplicative scattering correction, 2nd d = second derivative, SNV = standardized signal normalization, CM = center on the mean, 1st d = first derivative, Au = autoscaling, LB = baseline correction, [3]Matrix size, [4]number of latent variables; [5]root mean square error of the cross-validation; 6cross-validation correlation coefficient.

Unlike benchtop NIR devices, in which samples are deposited in an isolated and sealed compartment, the spectra in this study were collected using a portable NIR device. Portable equipment is more susceptible to physical phenomena such as temperature, pressure, and humidity changes, as it lacks adequate protection against environmental effects. Consequently, the persistent selection of these treatments may be attributed to issues related to reduced spectral variation and greater environmental interference in the acquired spectra. Autoscaling treatments and/or multiplicative scatter correction were applied in 67% of the models for predicting the chemical composition of soybean meal (ODS, OM, NDF, NDFap, iNDF, NDIP, ADIN, CP, EE, and starch) and 50% of the models for cornmeal (ODS, OM, iNDF, ADF, CP, NFC, and EE).

In contrast, 53% of the models for predicting the chemical composition of soybean meal (ODS, OM, NDF, NDFap, iNDF, NDIP, EE, and starch) and 27% of the models for cornmeal (iNDF, ADF, NFC, and EE) solely or in combination used autoscaling treatment. Autoscaling involves centering the spectral matrix on the mean and scaling it by variance (Ferreira, 2015). The need for spectral variation is emphasized, as all variables exhibited similar variances before modeling (Ciurczak et al., 2021). Hence, the frequent use of autoscaling treatment in prediction models for chemical composition may be attributed to the small variation in constituents between concentrate feed samples, where the standard deviation was low (ODS = 4.81, OM = 4.16, NDF = 1.56, NDFap = 7.74, iNDF = 6.48, NDIP = 0.09, EE = 8.85, and starch = 0.47; cornmeal: iNDF = 0.28, ADF = 0, 62, NFC = 2.26, and EE = 0.88), making it challenging to distinguish between samples.

External evaluation

Tables 3, 4, 5, 6, 7, and 8 describe the results of the external evaluation of sugarcane, soybean meal, and cornmeal for each constituent. The regression analysis between the observed and predicted values indicated that the generated models exhibited good predictive capacity for all constituents of the evaluated feedstuffs, as evidenced by the lack of rejection ($P \geq 0.056$) of the hypotheses of intercept equal to zero and slope equal to one.

The models developed for predicting the contents of ADS, ASE, DM, OM, NDF, NDFap, NDIP, ADF, CP, NFC, and NDT in sugarcane; ODS, OM, NDF, ADF, iNDF, CP, TDN, and starch in soybean meal; ODS and CP in cornmeal demonstrated high accuracy and precision ($R^2 \geq 0.50$ and $CCC \geq 0.60$). Conversely, the models generated for predicting the levels of iNDF, ADIP, EE, and lignin in sugarcane; NDF, NDIP, NDFap, ADIP, and NFC in soybean meal; and OM, NDFap, ADIP, NDIP, ADF, starch, and TDN in cornmeal showed intermediate precision ($R2 \geq 0.41$ and $CCC \geq 0.29$).

However, the models developed for determining the NDIA contents in sugarcane ($R^2 = 0.06$), NDIA and EE in soybean meal ($R^2 = 0.01$), and NDIA in cornmeal ($R^2 = -0.04$) exhibited low precision. This may be attributed to the low concentration of these constituents in the analyzed samples, posing challenges for accurate prediction using the NIR technique

(Porep et al., 2015). Additionally, the limited variation range of certain constituents in sugarcane (ADIP = 0.16 to 0.66; NDIA = 0.10 to 1.10; lignin = 3.05 to 5.95), soybean meal (NDIA = 0.25 to 1.20; EE = 1.02 to 3.33), and cornmeal (NDF = 10.17 to 16.27; iNDF = 1.04 to 2 .47; ADIP = 0.11 to 0.62; NDIA = 0.26 to 1.36; and EE = 2.54 to 5.90) may have resulted in a diminished predictive capacity at the extremes of the studied values (Sarraguça & Lopes, 2009). Furthermore, the lower prediction quality of the model generated for estimating NFC levels in cornmeal samples (R^2 = 0.19 and CCC = 0.38) may be attributed to the accumulation of systematic errors inherent to this constituent, as it is estimated from at least six different analyses, NFC = 100 - (%CP + %NDFap + %EE + % MM (Detmann et al., 2021).

The models developed for the prediction of all constituents in cornmeal demonstrated low values of MSEP (MSEP < 14.9% of the observed means), whereas the models generated for the prediction of the chemical composition of sugarcane showed MSEP values ranging from low (MSEP < 13.2% of the observed means for ODS, OM, NDF, NDFap, ADIP, NDIA, ADF, CP, EE, LIG, and TDN) to moderate (MSEP < 39.6% of the observed means for DM, iNDF, and NFC). Additionally, the MSEP for the estimates of all constituents in sugarcane and cornmeal was predominantly associated with random errors (MoF ≥ 71% of the MSEP), indicating that these errors are mainly unrelated to the structure of the models.

In the case of soybean meal, the models developed for predicting the levels of ODS, OM, NDIA, iNDF, ADF, ADIP, EE, CP, and TDN exhibited low values of MSEP (MSEP < 9% of the observed means), while the models for predicting NDF, NDFap, NDIP, NFC, and starch contents showed proportionally higher MSEP values (MSEP < 30% of the observed means). With the exception of NDIP, these higher MSEP values were primarily attributed to random errors (MoF ≥ 78.7% of MSEP). The prediction of the NDIP content in soybean meal displayed a greater contribution of bias errors (SB = 55.56% of MSEP), indicating problems with the adjustment of this model. The observed minimum NDIP values in soybean meal (NDIP = 2.1) deviated significantly from the predicted values (NDIP = 1.0), particularly in samples with lower content. Moreover, the minimum values were overestimated by approximately 210% when comparing them to the maximum observed values (1.0-3.9). Consequently, further studies with a larger number of samples are necessary to enhance the accuracy and robustness of the model for predicting the NDIP content in soybean meal samples.

Table 3
Mean and descriptive statistics for the relationship between the observed and predicted values of the chemical components of sugarcane

Parameters	Obs.[9]	Pred.[10]	Obs.	Pred.[11]	Obs.	Pred.[12]	Obs.	Pred.[13]	Obs.	Pred.[14]	Obs.	Pred.[15]	Obs.	Pred.[16]	Obs.	Pred.[17]
Average (%)	27.9	28.9	92.8	92.9	26.4	26.2	97.1	97.0	48.6	48.9	48.8	48.1	22.9	23.7	0.30	0.30
SD[18]	4.70	4.58	1.95	2.06	3.60	2.75	1.2	1.20	7.44	6.73	5.62	6.11	3.53	2.66	0.09	0.06
Maximum	35.2	35.0	96.2	96.8	31.6	30.3	98.9	99.1	63.1	60.1	58.5	59.8	29.8	27.6	0.50	0.40
Minimum	16.3	19.1	88.7	90.1	18.9	20.4	93.9	93.9	33.8	37.0	36.6	38.4	17.5	18.5	0.20	0.20
R2[19]	-	0.68	-	0.68	-	0.48	-	0.71	-	0.84	-	0.83	-	0.27	-	0.21
CCC[20]	-	0.82	-	0.83	-	0.68	-	0.85	-	0.92	-	0.90	-	0.51	-	0.47
Regression																
Intercept																
Estimated	-	3.19	-	19.61	-	2.15	-	18.12	-	-1.35	-	8.39	-	5.48	-	0.08
SE[21]	-	3.69	-	10.89	-	5.58	-	10.48	-	4.81	-	4.12	-	6.04	-	0.12
P-value[22]	-	0.398	-	0.087	-	0.704	-	0.098	-	0.782	-	0.056	-	0.376	-	0.518
Slope																
Estimated	-	0.86	-	0.79	-	0.93	-	0.81	-	1.02	-	0.84	-	0.73	-	0.79
SE	-	0.13	-	0.117	-	0.21	-	0.11	-	0.10	-	0.09	-	0.25	-	0.34
P-value[23]	-	0.27	-	0.085	-	0.738	-	0.099	-	0.825	-	0.077	-	0.306	-	0.538
QMEP[24]	-	7.738	-	1.317	-	6.251	-	0.409	-	7.870	-	6.341	-	9.389	-	0.006
QV[25]	-	0.922	-	0.019	-	0.072	-	0.005	-	0.080	-	0.541	-	0.694	-	≤0.001
MaF[26]	-	0.414	-	0.183	-	0.037	-	0.048	-	0.021	-	0.903	-	0.478	-	≤0.001
MoF[27]	-	6.402	-	1.114	-	6.141	-	0.356	-	7.767	-	4.898	-	8.217	-	0.006

[1]oven-dried matter at 55 °C, [2]oven-dried matter at 105 °C, [3]total dry matter, [4]organic dry matter, [5]neutral detergent insoluble fiber, [6]NDF corrected for ash and protein, [7]indigestible NDF, [8]acid detergent insoluble protein, [9]observed values, [10]detrend and smoothing, [11]detrend and normalization, [12]multiplicative scattering correction and second derivative, [13]smoothing and second derivative, [14]standardized signal normalization and smoothing, [15]smoothing and multiplicative scattering correction, [16]center on the mean and first derivative, [17]second derivative and detrend, [18]standard deviation, [19]coefficient of determination, [20]coefficient of correlation and agreement or reproducibility index, [21]standard error, [22]H0: β0 = 0, [23]H0: β1 = 1, [24]mean square of prediction error, [25]bias, [26]magnitude of random fluctuation and [27]random fluctuation of the model.

Table 4
Mean and descriptive statistics for the relationship between the observed and predicted values of the chemical components of sugarcane

Item	NDIP[1]		NDIA[2]		ADF[3]		CP[4]		NFC[5]		EE[6]		LIG[7]		NDT[8]	
	Obs.[9]	Pred.[10]	Obs.	Pred.[11]	Obs.	Pred.[12]	Obs.	Pred.[13]	Obs.	Pred.[14]	Obs.	Pred.[15]	Obs.	Pred.[16]	Obs.	Pred.[13]
Average (%)	1.2	1.2	0.6	0.7	30.4	30.6	2.7	2.6	43.7	44.4	1.6	1.8	4.2	4.5	62.2	61.8
SD[17]	0.36	0.40	0.18	0.13	4.58	4.72	0.60	0.62	8.58	8.15	0.50	0.19	0.75	0.59	3.15	3.28
Maximum	2.2	1.9	1.1	1.0	42.9	44.2	3.6	3.6	61.8	58.4	2.7	2.0	5.3	5.8	66.0	66.2
Minimum	0.8	0.5	0.4	0.5	23.9	24.8	1.4	1.5	27.0	30.1	1.0	1.5	3.1	3.5	55.0	53.7
R2[18]	-	0.34	-	0.06	-	0.94	-	0.68	-	0.90	-	0.43	-	0.22	-	0.88
CCC[19]	-	0.63	-	0.29	-	0.97	-	0.81	-	0.94	-	0.42	-	0.45	-	0.94
Regression																
Intercept																
Estimated	-	0.50	-	0.30	-	1.50	-	0.63	-	-0.67	-	-1.52	-	1.19	-	6.14
SE[20]	-	0.28	-	0.22	-	1.80	-	0.32	-	3.34	-	0.89	-	1.27	-	4.96
P-value[21]	-	0.096	-	0.180	-	0.416	-	0.060	-	0.843	-	0.108	-	0.365	-	0.233
Slope																
Estimated	-	0.58	-	0.46	-	0.94	-	0.80	-	1.0	-	1.77	-	0.66	-	0.91
SE	-	0.22	-	0.31	-	0.06	-	0.12	-	0.07	-	0.50	-	0.28	-	0.08
P-value[22]	-	0.077	-	0.099	-	0.331	-	0.115	-	0.993	-	0.146	-	0.242	-	0.261
QMEP[23]	-	0.099	-	0.038	-	1.279	-	0.137	-	7.418	-	0.174	-	0.555	-	1.267
QV[24]	-	≤0.01	-	0.005	-	0.079	-	0.015	-	0.484	-	0.028	-	0.123	-	0.136
MaF[25]	-	0.025	-	0.005	-	0.071	-	0.015	-	≤0.01	-	0.021	-	0.038	-	0.089
MoF[26]	-	0.074	-	0.028	-	1.130	-	0.107	-	6.934	-	0.125	-	0.394	-	1.043

[1]neutral detergent insoluble protein, [2]neutral detergent insoluble ash, [3]acid detergent fiber, [4]crude protein, [5]non-fiber carbohydrates, [6]ether extract, [7]lignin, [8]total digestible nutrients, [9]observed values, [10]normalization and first derivative, [11]autoscaling and baseline correction, [12]center on the mean and second derivative, [13]smoothing and second derivative, [14]smoothing and standardized signal normalization, [15]center on the mean and normalization, [16]second derivative and autoscaling, [17]standard deviation, [18]coefficient of determination, [19]coefficient of correlation and agreement or reproducibility index, [20]standard error, [21]H0: β0 = 0, [22]H0: β1 = 1, [23]mean square of prediction error, [24]bias, [25]magnitude of random fluctuation and [26]random fluctuation of the model.

Table 5
Mean and descriptive statistics for the relationship between the observed and predicted values of the chemical components soybean meal

Item	ODS[1]		OM[2]		NDF[3]		NDFap[4]		NDIA[5]		NDIP[6]		iNDF[7]	
	Obs.[8]	Pred.[9]	Obs.	Pred.[10]	Obs.	Pred.[11]	Obs.	Pred.[12]	Obs.	Pred.[13]	Obs.	Pred.[14]	Obs.	Pred.[9]
Average (%)	88.8	88.6	92.8	92.9	15.7	15.8	13.0	13.2	0.5	0.5	2.2	3.0	1.5	1.5
SD[15]	1.02	1.02	0.35	0.29	2.31	1.73	1.94	1.13	0.12	0.08	0.88	0.57	0.44	0.38
Maximum	90.7	90.7	93.4	93.4	19.9	19.3	15.6	15.0	0.7	0.7	3.9	4.2	2.6	2.2
Minimum	87.1	87.3	92.2	92.4	12.7	13.2	9.1	11.1	0.3	0.4	1.0	2.1	0.9	1.0
R2[16]	-	0.56	-	0.38	-	0.46	-	0.31	-	0.06	-	0.41	-	0.34
CCC[17]	-	0.74	-	0.60	-	0.67	-	0.51	-	0.30	-	0.42	-	0.61
Regression														
Intercept														
Estimated	-	21.78	-	22.26	-	0.88	-	-0.33	-	0.26	-	-0.08	-	0.41
SE[18]	-	12.57	-	18.48	-	3.58	-	4.23	-	0.17	-	0.89	-	0.37
P-value[19]	-	0.098	-	0.242	-	0.809	-	0.939	-	0.140	-	0.383	-	0.277
Slope														
Estimated	-	0.76	-	0.76	-	0.94	-	1.01	-	0.47	-	1.03	-	0.72
SE	-	0.14	-	0.20	-	0.23	-	0.32	-	0.32	-	0.30	-	0.24
P-value[20]	-	0.100	-	0.240	-	0.780	-	0.969	-	0.118	-	0.918	-	0.272
QMEP[21]	-	0.520	-	0.084	-	2.613	-	2.383	-	0.014	-	0.913	-	0.125
QV[22]	-	0.042	-	0.010	-	0.017	-	0.026	-	≤0.01	-	0.507	-	≤0.01
MaF[23]	-	0.059	-	0.005	-	0.012	-	≤0.01	-	0.002	-	≤0.01	-	0.011
MoF[24]	-	0.419	-	0.069	-	2.584	-	2.357	-	0.012	-	0.405	-	0.115

[1]oven-dried matter at 105 °C, [2]organic dry matter, [3]neutral detergent insoluble fiber, [4]NDF corrected for ash and protein, [5]neutral detergent insoluble ash, [6]neutral detergent insoluble protein, [7]indigestible NDF, [8]observed values, [9]autoscaling and multiplicative scattering correction, [10]autoscaling, [11]autoscaling and detrend, [12]normalization and autoscaling, [13]second derivative and multiplicative scattering correction, [14]autoscaling and standardized signal normalization, [15]standard deviation, [16]coefficient of determination, [17]coefficient of correlation and agreement or reproducibility index, [18]standard error, [19]H0: β0 = 0, [20]H0: β1 = 1, [21]mean square of prediction error, [22]bias, [23]magnitude of random fluctuation and [24]random fluctuation of the model.

Table 6
ean and descriptive statistics for the relationship between the observed and predicted values of the chemical components soybean meal (continuation)

Item	ADF[1]		ADIP[2]		NFC[3]		EE[4]		PB[5]		NDT[6]		AMIDO	
	Obs.[7]	Pred.[8]	Obs.	Pred.[9]	Obs.	Pred.[10]	Obs.	Pred.[11]	Obs.	Pred.[12]	Obs.	Pred.[9]	Obs.	Pred.[13]
Average (%)	7.0	7.0	0.3	0.3	26.6	26.2	1.2	1.8	51.6	51.6	79.0	78.7	5.9	5.5
SD[14]	1.22	1.21	0.10	0.05	2.36	1.89	0.94	0.24	0.95	0.63	1.17	1.10	1.18	1.06
Maximum	9.5	9.7	0.6	0.4	30.1	29.8	2.3	2.1	53.0	52.6	81.7	81.5	7.5	7.4
Minimum	4.3	4.7	0.2	0.2	20.6	22.6	1.4	1.4	49.9	50.5	77.4	77.1	3.6	3.2
R2[15]	-	0.80	-	0.35	-	0.33	-	0.01	-	0.50	-	0.64	-	0.48
CCC[16]	-	0.89	-	0.43	-	0.58	-	0.26	-	0.67	-	0.78	-	0.67
Regression														
Intercept														
Estimated	-	0.65	-	-0.10	-	6.80	-	1.04	-	-4.40	-	11.44	-	1.35
SE[17]	-	0.72	-	0.12	-	5.84	-	0.66	-	14.12	-	11.13	-	1.20
P-value[18]	-	0.38	-	0.4117	-	0.257	-	0.139	-	0.759	-	0.317	-	0.2789
Slope														
Estimated	-	0.90	-	1.45	-	0.76	-	0.39	-	1.09	-	0.86	-	0.82
SE	-	0.10	-	0.43	-	0.22	-	0.36	-	0.27	-	0.14	-	0.21
P-value[19]	-	0.350	-	0.310	-	0.288	-	0.114	-	0.760	-	0.330	-	0.427
QMEP[20]	-	0.304	-	0.008	-	3.775	-	0.121	-	0.398	-	0.556	-	0.857
QV[21]	-	0.001	-	0.001	-	0.191	-	0.004	-	≤0.001	-	0.091	-	0.148
MaF[22]	-	0.013	-	≤0.001	-	0.202	-	0.020	-	≤0.001	-	0.023	-	0.033
MoF[23]	-	0.289	-	0.007	-	3.382	-	0.097	-	0.395	-	0.441	-	0.676

[1]acid detergent fiber, [2]acid detergent insoluble protein, [3]non-fiber carbohydrates, [4]ether extract, [5]crude protein, [6]total digestible nutrients, [7]observed values, [8]standardized signal normalization and smoothing, [9]smoothing and first derivative, [10]center on the mean and baseline correction, [11]autoscaling and second derivative, [12]multiplicative scattering correction and smoothing, [13]autoscaling and multiplicative scattering correction, [14]standard deviation, [15]coefficient of determination, [16]coefficient of correlation and agreement or reproducibility index, [17]standard error, [18]H0: β0 = 0, [19]H0: β1 = 1, [20]mean square of prediction error, [21]bias, [22]magnitude of random fluctuation and [23]random fluctuation of the model.

Table 7
Mean and descriptive statistics for the relationship between the observed and predicted values of the chemical components of cornmeal

Item	ADF[1]		OM[2]		NDF[3]		NDFap[4]		iNDF[5]		ADIP[6]		NDIP[7]	
	Obs.[8]	Pred.[9]	Obs.	Pred.[10]	Obs.	Pred.[11]	Obs.	Pred.[12]	Obs.	Pred.[13]	Obs.	Pred.[14]	Obs.	Pred.[15]
Average (%)	87.5	87.4	98.7	98.7	14.5	14.0	12.7	12.9	1.6	1.6	0.3	0.3	1.3	1.3
SD[16]	1.23	1.11	0.27	0.13	1.35	0.94	1.41	1.10	0.24	0.09	0.11	0.08	0.25	0.14
Maximum	90.3	90.0	99.3	98.9	16.3	15.7	14.7	15.2	2.1	1.8	0.6	0.5	1.7	1.5
Minimum	84.5	84.5	98.1	98.3	11.5	11.6	9.7	10.9	1.2	1.3	0.2	0.2	0.8	1.0
R2[17]	-	0.88	-	0.47	-	0.11	-	0.36	-	0.05	-	0.29	-	0.33
CCC[18]	-	0.93	-	0.55	-	0.34	-	0.59	-	0.18	-	0.54	-	0.51
Regression														
Intercept														
Estimated	-	-3.84	-	-41.42	-	6.60	-	2.41	-	0.38	-	0.09	-	-0.12
SE[19]	-	6.43	-	27.19	-	403.00	-	2.66	-	0.77	-	0.07	-	0.39
P-value[20]	-	0.555	-	0.139	-	0.116	-	0.374	-	0.625	-	0.256	-	0.770
Slope														
Estimated	-	1.04	-	1.42	-	0.56	-	0.79	-	0.80	-	0.76	-	1.07
SE	-	0.07	-	0.28	-	0.29	-	0.20	-	0.50	-	0.22	-	0.3
P-value[21]	-	0.549	-	0.139	-	0.140	-	0.326	-	0.694	-	0.274	-	0.811
QMEP[22]	-	0.180	-	0.039	-	1.826	-	1.277	-	0.058	-	0.008	-	0.039
QV[23]	-	0.004	-	≤0.001	-	0.186	-	0.060	-	0.006	-	≤0.001	-	≤0.001
MaF[24]	-	0.002	-	0.003	-	0.165	-	0.049	-	≤0.001	-	≤0.001	-	≤0.001
MoF[25]	-	0.173	-	0.036	-	1.414	-	1.168	-	0.052	-	0.008	-	0.039

[1]oven-dried matter at 105 °C, [2]organic dry matter, [3]neutral detergent insoluble fiber, [4]NDF corrected for ash and protein, [5]indigestible NDF, [6]acid detergent insoluble protein, [7]neutral detergent insoluble protein, [8]observed values, [9]second derivative and multiplicative scattering correction, [10]multiplicative scattering correction and detrend, [11]normalization and multiplicative scattering correction, [12]second derivative and standardized signal normalization, [13]second derivative and standardized signal normalization, [14]smoothing and second derivative, [15]detrend and baseline correction, [16]standard deviation, [17]coefficient of determination, [18]coefficient of correlation and agreement or reproducibility index, [19]standard error, [20]H0: β0 = 0, [21]H0: β1 = 1, [22]mean square of prediction error, [23]bias, [24]magnitude of random fluctuation and [25]random fluctuation of the model.

Table 8
Mean and descriptive statistics for the relationship between the observed and predicted values of the chemical components of cornmeal (continuation)

Item	NDIA[1]		ADF[2]		CP[3]		NFC[4]		EE[5]		Starch		NDT[6]	
	Obs.[7]	Pred.[8]	Obs.	Pred.[9]	Obs.	Pred.[10]	Obs.	Pred.[11]	Obs.	Pred.[11]	Obs.	Pred.[12]	Obs.	Pred.[13]
Average (%)	0.7	0.6	3.4	3.2	8.8	8.8	73.2	73.1	4.1	4.1	72.5	72.1	89.8	90.0
SD[14]	0.23	0.08	0.71	0.30	0.45	0.38	2.45	1.24	0.83	0.60	3.20	1.78	1.43	0.73
Maximum	1.4	0.8	4.7	3.7	10.0	10.1	76.8	75.6	5.6	4.9	78.6	75.2	91.8	90.9
Minimum	0.3	0.5	2.1	2.5	8.0	8.3	68.1	71.2	2.71	2.9	65.4	68.7	86.9	88.3
R2[15]	-	-0.04	-	0.42	-	0.37	-	0.19	-	0.11	-	0.21	-	0.46
CCC[16]	-	0.03	-	0.45	-	0.62	-	0.38	-	0.38	-	0.43	-	0.55
Regression														
Intercept														
Estimated	-	0.58	-	-1.71	-	2.40	-	5.88	-	1.74	-	7.18	-	-32.70
SE[17]	-	0.34	-	1.17	-	1.67	-	24.96	-	1.33	-	27.86	-	26.20
P-value[18]	-	0.105	-	0.157	-	0.165	-	0.816	-	0.209	-	0.800	-	0.224
Slope														
Estimated	-	0.14	-	1.59	-	0.73	-	0.92	-	0.56	-	0.91	-	1.36
SE	-	0.56	-	0.36	-	0.19	-	0.34	-	0.32	-	0.39	-	0.29
P-value[19]	-	0.132	-	0.119	-	0.164	-	0.820	-	0.187	-	0.811	-	0.227
QMEP[20]	-	0.060	-	0.329	-	0.127	-	4.530	-	0.620	-	7.375	-	1.129
QV[21]	-	0.002	-	0.030	-	≤0.01	-	0.017	-	0.006	-	0.160	-	0.030
MaF[22]	-	0.005	-	0.029	-	0.015	-	0.009	-	0.065	-	0.027	-	0.066
MoF[23]	-	0.053	-	0.270	-	0.116	-	4.504	-	0.549	-	7.189	-	1.033

[1]neutral detergent insoluble ash, [2]acid detergent fiber, [3]crude protein, [4]non-fiber carbohydrates, [5]ether extract, [6]total digestible nutrients, [7]observed values, [8]second derivative and baseline correction, [9]autoscaling and normalization, [10]multiplicative scattering correction and smoothing, 11second derivative and autoscaling, [12]detrend and first derivative, [13]second derivative and detrend, [14]standard deviation, [15]coefficient of determination, [16]coefficient of correlation and agreement or reproducibility index, [17]standard error, [18]H0: β0 = 0, [19]H0: β1 = 1, [20]mean square of prediction error, [21]bias, [22]magnitude of random fluctuation and [23]random fluctuation of the model.

Conclusion

The portable-NIR regression models accurately estimate and therefore are recommended for estimating the chemical composition of sugarcane, soybean meal, and cornmeal. The portable NIR thus offers a viable alternative to conventional laboratory methods for determining the composition of these feedstuffs, having advantages such as cost reduction, decreased labor requirements, faster results, and reduced generation of potential pollutants. However, to enhance the accuracy and robustness of the prediction equations for the chemical composition of sugarcane, soybean meal, and cornmeal, further studies with larger sample sizes and increased variation in the origin of the samples are warranted.

Acknowledgments

Thanks are due to the following research funding bodies and funding institutions: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

References

Ciurczak, E. W., Igne, B., Workman, J., Jr., & Burns, D. A. (2021). *Handbook of near-infrared analysis.* CRC Press.

Companhia Nacional de Abastecimento (2021). *Boletim da safra de grãos, 3º Levantamento - Safra 2020/21.* https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos

Detmann, E., Silva, L. F. C., Rocha, G. C., Palma, M. N. N., & Rodrigues, J. P. P. (2021). *Métodos para análise de alimentos.* Suprema.

Ferreira, M. M. C. (2015). *Quimiometria: conceitos, métodos e aplicações.* Editora da Unicamp.

Kennard, R. W., & Stone, L. A. (1969). Computer aided design of experiments. *Technometrics, 11*(1), 137-148. doi: 10.1080/00401706.1969.10490666

Kobayashi, K., & Salam, M. U. (2000). Comparing simulated and measured values using mean squared deviation and its components. *Agronomy Journal, 92*(2), 345-352. doi: 10.2134/agronj2000.922345x

Montgomery, D. C. (2009). *Introdução ao controle estatístico da qualidade.* Grupo Gen-LTC.

Ozaki, Y., Huck, C., Tsuchikawa, S., & Engelsen, S. B. (2021). *Near-infrared spectroscopy: theory, spectral analysis, instrumentation, and applications.* Springer.

Peternelli, L. A., Gonçalves, M. T. V., Fernandes, J. G., Brasileiro, B. P., & Teófilo, R. F. (2020). Selection of sugarcane clones via multivariate models using near-infrared (NIR) spectroscopy data. *Australian Journal of Crop Science, 14*(6), 889-895. doi: 10.21475/ajcs.20.14.06.p2099

Pinto, A. C. J., & Millen, D. D. (2019). Nutritional recommendations and management practices adopted by feedlot cattle nutritionists: the 2016 Brazilian survey. *Canadian Journal of Animal Science, 99*(2), 392-407. doi: 10.1139/cjas-2018-0031

Porep, J. U., Kammerer, D. R., & Carle, R. (2015). On-line application of near infrared (NIR) spectroscopy in food production. *Trends in Food Science & Technology, 46*(2), 211-230. doi: 10.1016/j.tifs.2015.10.002

Sarraguça, M. C., & Lopes, J. A. (2009). Quality control of pharmaceuticals with NIR: from lab to process line. *Vibrational Spectroscopy, 49*(2), 204-210. doi: 10.1016/j.vibspec.2008.07.013

Silva, B. C., Godoi, L. A., Valadares, S. C., Fº., Zanetti, D., Benedeti, P. D. B., & Detmann, E. (2019a). A suitable enzymatic method for starch quantification in different organic matrices. *MethodsX, 6*(2019), 2322-2328. doi: 10.1016/j.mex.2019.09.040

Silva, D. P., Pedroso, A. M., Pereira, M. C. S., Bertoldi, G. P., Watanabe, D. H. M., Melo, A. C. B., & Millen, D. D. (2019b). Survey of management practices used by brazilian dairy farmers and recommendations provided by 43 dairy cattle nutritionists. *Canadian Journal of Animal Science, 99*(4), 890-904. doi: 10.1139/cjas-2018-0004

Silvestre, A. M., & Millen, D. D. (2021). The 2019 brazilian survey on nutritional practices provided by feedlot cattle consulting nutritionists. *Revista Brasileira de Zootecnia, 50*(9), 1-25. doi: 10.37496/RBZ50202 00189

Souza, R. C., Reis, R. B., Lopez, F. C. F., Mourthe, M. H. F., Lana, A. M. Q., Barbosa, F. A., & Sousa, B. M. (2015). Efeito da adição de teores crescentes de ureia na cana-de-açúcar em dietas de vacas em lactação sobre a produção e composição do leite e viabilidade econômica. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia, 67*(2-3), 564-572. doi: 10.1590/1678-7799

Tedeschi, L. O. (2006). Assessment of the adequacy of mathematical models. *Agricultural Systems, 89*(1-2), 225-247. doi: 10.1016/j.agsy.2005.11.004

Thomson, A. L., Humphries, D. J., Rymer, C., Archer, J. E., Grant, N. W., & Reynolds, C. K. (2018). Assessing the accuracy of current near infra-red reflectance spectroscopy analysis for fresh grass-clover mixture silages and development of new equations for this purpose. *Animal Feed Science and Technology, 239*(2018) 94-106. doi: 10.1016/j.anifeedsci.2018.03.009

Valadares, S. C., F., Lopes, S. A., Silva, B. C., & Chizzotti, M. L. L. Z. B. (2018). *Tabelas brasileiras de composição de alimentos para ruminantes (CQBAL 4.0).*

Valadares, S. C. F., L., Silva F. C., Gionbelli, M. P., Rotta, P. P., Marcondes, M. I., Chizzotti, M. L., & Prados, L. F. (2016). *Exigências Nutricionais de Zebuínos Puros e Cruzados - BR-CORTE.* Editora Federal de Viçosa.