



Fourier transform near-infrared spectroscopy (FT-NIRS) application to estimate Brazilian soybean [*Glycine max* (L.) Merrill] composition

Daniela Souza Ferreira ^{a,*}, Juliana Azevedo Lima Pallone ^a, Ronei Jesus Poppi ^b

^a Department of Food Science, Faculty of Food Engineering, University of Campinas (UNICAMP), P.O. Box 6121, 13083-862, Campinas, SP, Brazil

^b Institute of Chemistry, University of Campinas (UNICAMP), P.O. Box 6154, 13084-971, Campinas, SP, Brazil

ARTICLE INFO

Article history:

Received 13 July 2012

Accepted 14 September 2012

Keywords:

Multivariate data analysis

Composition analysis

Near-infrared reflectance spectroscopy

Soybean

ABSTRACT

This study examined the ability of near-infrared reflectance spectroscopy method (FT-NIRS) and multivariate calibration to estimate the concentration of moisture, protein, lipid, ash and carbohydrate of Brazilian soybeans. The spectra obtained in the range of 4000 to 10,000 cm^{-1} were preprocessed by several combinations of mathematical treatments: MSC (multiplicative scatter correction), SNV (standard normal variate) or first and second derivative and all data were mean centered before the calibration, for which was used the PLS method (partial least squares). The best calibration models found in this study were the ones used to determine protein and moisture contents ($R^2 = 0.81$, RMSEP = 1.61% and $R^2 = 0.80$, RMSEC = 1.55%, respectively). However, the technique shows high predictability for all parameters, including lipids, ashes and carbohydrates, with RMSECV of 0.40 to 2.30% and RMSEP of 0.38 to 3.71%. This result shows the viability of using NIR in controlling the quality parameters of soybeans.

© 2012 Elsevier Ltd. Open access under the [Elsevier OA license](#)

1. Introduction

The Fourier transform near-infrared spectroscopy (FT-NIRS) is a non-destructive fast technique capable of analyzing organic substances in a reliable way. Such analytical method has presented a great impact in the food industry and agriculture. According to Williams' (2001) work, a huge number of analytical methodologies based on near-infrared spectroscopy have already demonstrated the ability of such technique in dealing safely and quickly with so diverse problems as determining proteins, moistures, carbohydrates and fat content in different types of food.

The centesimal food content analysis aims to guarantee the product quality, to follow official supervision and regulation committee requirements, as well as, to contribute to the consumer health. This chemical analysis, which includes the determination of moisture, protein, ash, fat and carbohydrate contents, traditionally has been performed by laborious methods which are time demanding and generate much chemical waste. However, the methods that use the FT-NIRS technique provide precise and fast results, with minimal or null sample preparation and waste generation, being able to be an alternative to such determinations (Kandala, Sundaram, & Puppala, 2012; Salgó & Gergely, 2012; Zhou, Yang, Huang, & Han, 2012).

Nevertheless, due to spectral features in the NIR region and in order to simultaneously evaluate diverse parameters, multivariate calibration methods must be used to obtain quantitative information. Multivariate calibration methods are part of chemometrics, which

needs a huge number of results from analytical trials for the development of a mathematical model. Furthermore, once a model is built, it has to be validated by a group of data that were not part of the calibration (Agelet & Hurburgh, 2010; Jimaré Benito, Bosch Ojeda, & Sanchez Rojas, 2008). Toward this Sun, Han, Yan, Yang, and Sato (2008) estimated the fatty acid concentrations in one hundred and eight Chinese soybean varieties.

The FT-NIRS technique and the multivariate calibration have already been used to determine some major contents in food, like Haughey et al. (in press) who described the evaluation of melamine content in soybean, Kuligowski, Carrión, Quintás, Garrigues, and de la Guardia (2012) who evaluated the NIR method for the determination of polymerized triacylglyceride in vegetable oils and Szegedi, Lénárt, Dernovics, Turza, and Fodor (2011) who determined protein in different vegetables. Zhu et al. (2011) also measured soybean seeds, nevertheless by near-infrared hyperspectral imaging system, for the determination of protein and fat contents of samples.

The wide range of results of the developed models demonstrates that with only one spectroscopic technique, the FT-NIRS, it is possible to simultaneously analyze many parameters: moisture, protein, ash, lipid and carbohydrate. Munck et al. (2010) analyzed starch, protein and lipid contents of barley varieties by NIR only, while an application of NMR, FT-Raman, FT-IR and NIR was necessary for Holve, Larsen, Hansen, and Engelsens (2011) in order to characterize carbohydrates, proteins, dietetic fibers and unsaturated fatty acids in leguminous plants from Africa.

Brazil is the second country in the world with the largest production of soybean (CONAB, 2012), being it a very important food in Brazil and an inexpensive source of oil and protein. According to the

* Corresponding author. Tel.: +55 19 33053882; fax: +55 19 35212153.

E-mail address: eng.dsf@gmail.com (D.S. Ferreira).

Table 1
Descriptive statistics of the soybean chemical parameters used in the calibration.

Parameter	Calibration					Validation					SE ±
	Min. (%)	Max. (%)	Mean (%)	SD ±	No.	Min. (%)	Max. (%)	Mean (%)	SD ±	No.	
Moisture	8.16	18.10	13.70	2.14	70	10.40	16.566	13.61	1.60	30	0.086
Ashes	4.32	6.14	5.11	0.36	70	4.44	5.69	5.04	0.32	30	0.069
Lipids	12.55	26.96	21.91	2.20	70	12.93	24.27	21.61	1.94	30	0.162
Proteins	31.52	43.48	38.91	2.34	70	34.12	43.46	39.32	2.29	30	0.274
Carbohydrates	13.34	26.96	20.51	3.06	69	14.04	27.50	20.02	3.12	29	–

Min. – minimum, max. – maximum, SD – standard deviation, SE – standard error of analysis, no. – number of samples.

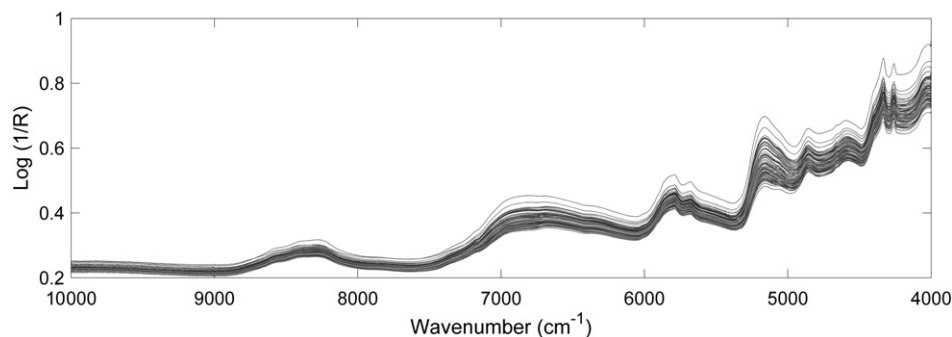


Fig. 1. NIR spectra of the soybean samples obtained in the range of 10,000 to 4000 cm^{-1} .

data of the Brazilian Agricultural Research Corporation (EMBRAPA, 2012), many varieties are produced and the evaluation of the quality of the beans depends on the centesimal content. Hence the aim of this work is to offer a fast and efficient method to predict the chemical characteristics of soybean with the application of FT-NIRS and multivariate calibration to several Brazilian varieties.

2. Materials and methods

2.1. Samples

A total of 100 soybean [*Glycine max* (L.) Merrill] varieties from the south and central parts of Brazil were donated by EMBRAPA Soybean. The landmark analyses were taken between 2009 and 2010.

2.2. Sample preparation

The 250 g received for each bean was crushed in a grinding mill (A11 model, IKA, Germany) and the soil gradation was standardized to 0.6 mm in a 28 mesh strainer. While a small portion (about 1 g) of such sample was immediately analyzed by near-infrared spectroscopy

by diffuse reflectance, another portion was used in chemical analyses by standard methods.

2.3. Landmark analysis

The total contents of moisture, protein and ash were determined by 925.09, 979.09, and 923.03 reference methods, respectively, as described by AOAC (2006). The total lipid content was determined by Bligh and Dyer (1959) and the total carbohydrate content was calculated by the formula (% carbohydrate = 100 – % moisture – % lipid – % protein – % ash). All the analyses were carried out in triplicate and right after the grinding in order to avoid any sort of alteration with the environment.

2.4. Near-infrared analysis

The grinded beans were directly analyzed by diffuse reflectance, using a NIR infrared reflectance accessory – NIRA (Perkin Elmer, United States), and a Spectrum 100N equipment (Perkin Elmer, United States), with a source of quartz halogen bulb. The software used for NIRA is Spectrum – V. 6.3.1.0132 (Perkin Elmer), the region measured is between 10,000 and 4000 cm^{-1} with 4 cm^{-1} of resolution and 64 scans and the analysis was run in duplicate.

Table 2
Wavenumber region and treatments used in calibration and validation.

Parameter	Wavenumber region	Derivative treatment ^a	Scatter correction	No of outliers ^b
Moisture	10,000–3500	X, 11, 2, 1		
Ashes	10,000–4900	X, 51, 3, 2		
Lipids	10,000–4000		SNV	
Proteins	7500–4900	X, 51, 3, 2		
Carbohydrates	7500–4500	X, 15, 2, 1		2

^a Savitzky–Golay employs the terms (y, width, order, deriv): y, the matrix of row vectors to be smoothed; width, number of points in the filter; order, order of the polynomial; and deriv, derivative.

^b Number of samples eliminated as outliers during calibration.

Table 3
Chemometrics results of PLS models by FT-NIRS of the evaluated parameters of soybeans.

Parameters	LV	R ²	RMSEC (%)	RMSECV (%)	RMSEP (%)	Bias	RPD
Moisture	5	0.80	0.28	2.30	1.55	–0.049	1.38
Ash	5	0.63	0.07	0.40	0.38	–0.080	0.95
Lipids	6	0.71	1.13	1.51	1.20	–0.023	1.83
Proteins	6	0.81	0.58	1.14	1.61	–0.020	1.45
Carbohydrates	8	0.50	1.11	1.62	3.71	–0.196	0.83

LV – latent variables, R² – correlation coefficient, RMSEC, RMSEP, RMSECV – root mean square error of the calibration, prediction and cross-validation, respectively, bias – systematic calibration error (the mean difference between the concentration predicted by NIR and the reference data), RPD (residual predictive deviation = standard deviation of the reference data/root mean square error of the prediction).

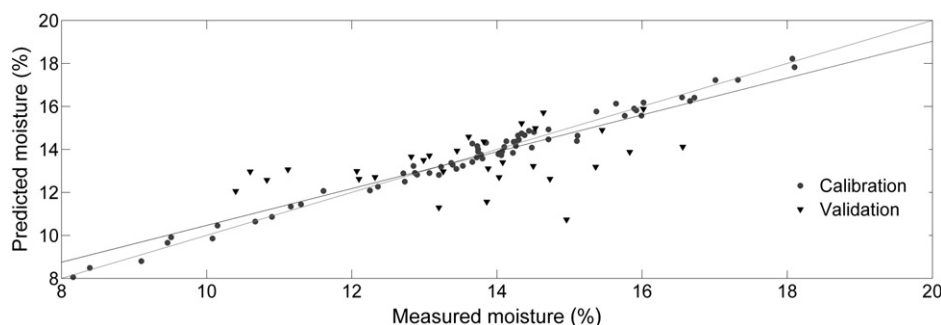


Fig. 2. Plot of the predicted values by NIR against the values measured by standard methods for moisture content.

2.5. Chemometrics analysis

2.5.1. Data modeling

The multivariate calibration model was carried out by partial least squares (PLS). The NIR spectra were treated by the PLS Toolbox v. 5.8 (Eigenvector Research Inc., United States) program developed for Matlab 7.8 (MathWorks, United States).

2.5.2. Sample choice for calibration and prediction

From the total of 100 soybean samples, 70 were used for the construction of the calibration model and 30 were used for the external validation (prediction set). The sample choice for the calibration and prediction was carried out by the Kennard–Stone algorithm (Kennard & Stone, 1969). After defining the number of samples to calibrate, this algorithm initiates the selection of two spectra: the closest one to the central area and the other with the longer Euclidian distance in relation to the first. Afterward, the other chosen sample is the one with the longest distance in relation to this last one and thus successively until the chosen number of samples to calibrate is reached. The rest of the samples are the prediction set. Finally, the quality of the model was measured by determining the concentration of each one of the 30 samples.

2.5.3. Pretreatment of spectra

The spectra in the infrared region usually present problems of baseline due to light scattering from the non-homogeneity of the matrix particle distribution (Bjorsvik & Martens, 2007). In order to minimize the effect of the change in the baseline, the spectra were preprocessed by the combination of many mathematical treatments: MSC (multiplicative scatter correction) (Geladi, MacDougall, & Martens, 1985), SNV (standard normal variate transformation) (Barnes, Dhanoa, & Lister, 1989) or first and second derivative using the Savitzky–Golay method (Brown, Vega-Montoto, & Wentzell, 2000), and all spectra data were mean centered before calibration. The number of PLS factors was determined by cross-validation (leave-one-out) procedure (Geladi & Kowalski, 1986).

2.5.4. Model evaluation

The performance of the final PLS model was evaluated according to: root mean square error calibration (RMSEC), root mean square error of cross-validation (RMSECV), root mean square error of prediction (RMSEP), coefficient of correlation (R^2) of the plot between the predicted values by FT-NIRS and the reference values, and bias (Eq. (1)), which consists of the difference between the mean value predicted by FT-NIRS and the mean value of the reference model and residual predictive deviation (RPD) (Williams, 2001; Williams & Sobering, 1995; Windham, Mertens, & Barton, 1989). High results of R^2 and low RMSE indicate the best model to predict the chemical content.

$$\text{bias} \equiv \frac{1}{n} \sum_{i=1}^n (x_i - y_i) \quad (1)$$

where n is the number of samples of the calibration, x_i is the result obtained by FT-NIRS and y_i is the result obtained by the reference method of the sample i .

In order to evaluate the RMSECV, a spectrum of a sample is removed from the calibration set and the PLS model is constructed with the remaining spectra. The removed sample was predicted with such model and the process was repeated with the removal of each sample. The RMSECV was then calculated as Eq. (2).

While the number of the PLS factors included in the model was chosen according to the lowest value of RMSECV, the RMSEP, which expresses the degree of accordance between the estimated values by a model previously constructed and the value considered real or of reference, was also calculated as in Eq. (2).

RMSEC, which provides information about the adjustment of the model to the calibration data, is calculated according to Eq. (2).

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \quad (2)$$

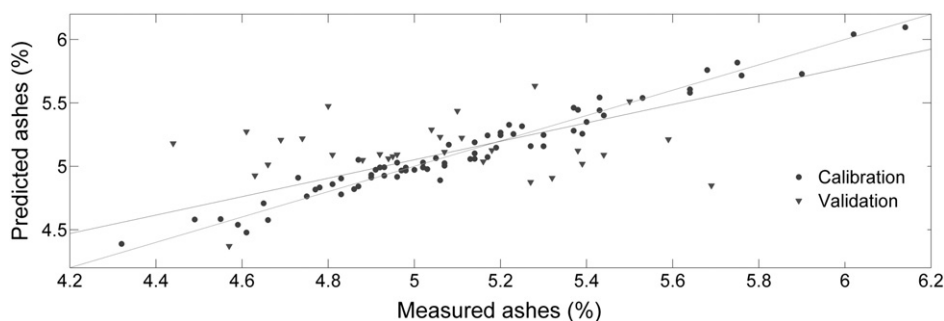


Fig. 3. Plot of the predicted values by NIR against the values measured by standard methods for ash content.

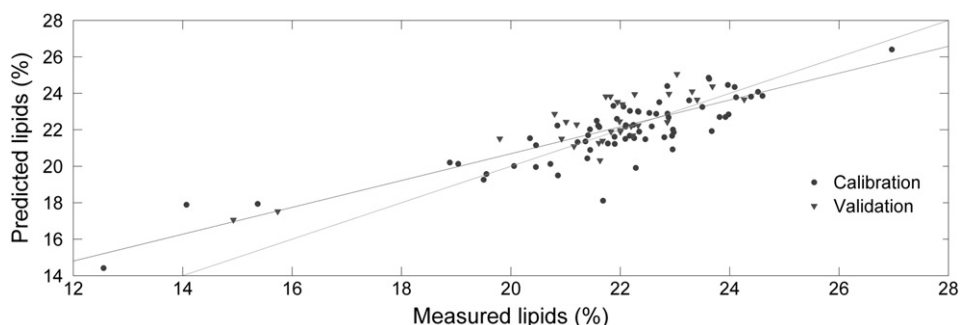


Fig. 4. Plot of the predicted values by NIR against the values measured by standard methods for lipid content.

where n is the number of samples (of the calibration set or prediction set), y_i is the result of the reference to the sample i and x_i is the estimated result by the NIR model.

Furthermore, the RPD represents the division between the reference data standard deviation and the mean error of the prediction (Williams & Sobering, 1995).

$$RPD = \frac{DP}{RMSEP} \quad (3)$$

The coefficient of correlation between the predicted value by NIR and the reference value was calculated for the prediction sets.

The outliers were identified and removed when necessary by $Q_{\text{residual}} \times T^2$ graphics and by leverage. Hotelling's T^2 means the square of the distance of the sample from the model mean that can be explained by normal variations within the group. The Q residuals mean the distance of the sample from the model that can be explained by random error (Naes, Isaksson, Fearn, & Davies, 2002).

3. Results and discussion

3.1. Composition analysis in soybean

The results obtained in the reference analysis of the moisture, protein, ash and carbohydrate contents of the samples that are used to predict and validate the model are presented in Table 1. By analyzing the values of the mean and standard deviation, the huge variability among the samples, especially considering lipid, protein and carbohydrate contents, is observed.

The variability of the results, represented as standard deviation, is placed in the prediction errors and the cross-validated results.

3.2. Development of calibration models and validation

In Fig. 1 all the spectra are presented, without any pretreatment, in the region of the near-infrared, obtained from the 100 soybean samples, in the range of 10,000 to 4000 cm^{-1} .

Among the pretreatments applied before calibration, a procedure that presented a good performance in the elimination of the multiplicative light scattering effect was the first derivative. Furthermore, such data were mean centered in order to reduce the matrix effects from the samples on the model.

Table 2 shows the wavenumber regions and final pretreatments used for each model. Examination of the loading vector indicated which bands of the spectrum were associated with the relevant chemical information (Haaland & Thomas, 1988).

The results used to evaluate the efficiency of predicting moisture, lipid, protein, ash and carbohydrate contents of the Brazilian soybean samples are presented in Table 3. The best models to predict the soybean chemical composition were chosen based on the lowest results of RMSECV and the highest results of R^2 (Windham et al., 1989).

All the models presented good correlation between the reference values and the NIR predicted ones. All the regression errors were low, demonstrating the capacity to predict the soybean chemical characteristics by FT-NIRS. Considering the expectations, the errors of prediction (RMSEP) were higher than the errors of the calibration group (RMSEC).

Low calibration and prediction errors of the moisture and ash contents (RMSEC 0.28 and 0.07%, RMSEP 1.55 and 0.28%, respectively) were obtained due to the few variables to describe the models (5 LV for each one). However, the correlation coefficient was better for the moisture content when compared to the ash content (R^2 0.80 and 0.63) due to the difficulty in predicting the compound present in the lower content and inorganic substance.

The wide dispersion of the results for the carbohydrate content in the reference analysis (Table 1, DP = 3.07) had as a consequence a lower correlation (Table 2, $R^2 = 0.50$) between the reference value and the predicted one by the model, indicating a lower accuracy to the prediction of such result, even higher than the latent variable numbers (Table 2, 8 LV). This behavior can be explained by the difficulty to obtain a model capable of appropriately describing the value range to the carbohydrate content in the matrix.

The calibration model obtained in this study presented high R^2 (0.81) and low prediction error (RMSEP = 1.61%) for the protein

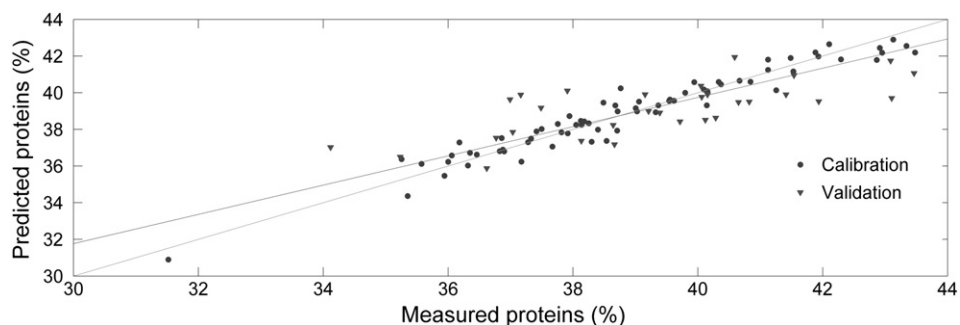


Fig. 5. Plot of the predicted values by NIR against the values measured by standard methods for protein content.

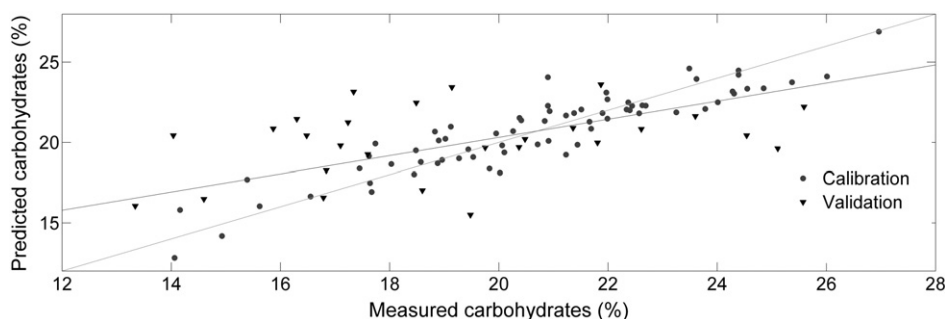


Fig. 6. Plot of the predicted values by NIR against the values measured by standard methods for carbohydrate content.

content with 6 latent variables, which is in conformity with other references that evaluated proteins of vegetables and soybeans (Kovalenko, Rippke, & Hurburgh, 2006; Szegedi et al., 2011).

The prediction errors (RMSEP) and bias obtained in this study were low and presented the same value of the results presented in the study of bread samples by Sorensen (2009), who also obtained a higher RMSEP for carbohydrate content in relation to other errors.

According to Williams (2001) a value of RPD > 2.4 is desirable to an appropriate model, while Williams and Sobering (1995) indicate that the value of 3 or more is the recommended value. This result depends on the errors of the prediction (RMSEP) which must be lower than the standard deviation of the data group. Therefore, the result of the RPD used to determine lipid content (RPD 1.83) was better, once the other results were low (RPD from 0.83 to 1.45), especially for the carbohydrate content which presented higher RMSEP than the standard deviation (RMSEP = 3.71%, DP = 3.08). However, evaluating only RPD does not determine the predictable capacity of the models, since the other RMSEPs were low when compared to the DP, indicating low prediction errors. Thus, a wide range of results was chosen in the group in order to evaluate the predictable ability of FT-NIRS.

The mean difference between the values predicted by the model and the reference values (bias) were not representative, the other parameters being higher for carbohydrates only (bias = -0.196). Despite such good result, the proximity between the reference values and the ones predicted by the model was only noted due to the analysis in the group with the other results of the model.

The ability of the PLS models to predict soybean chemical contents was also evaluated by the plot that relates the value measured by the reference method and the one predicted by FT-NIRS, presented in Figs. 2 to 6.

In all the figures the curve of the predicted values versus the real ones in the 30 samples of the predictable group presents an appropriate adjustment to the model. It is noted that the inclination of the line is close to 1 and the intercept is next to zero, two parameters that prove the absence of systematic errors.

It is observed that the parameters presented an appropriate relation between the evaluated and the predicted values, the models being closer to the real ones, confirming once the best results for moisture and protein contents (Figs. 2 and 5 – R^2 of 0.80 and 0.81, respectively).

In the next figure, the absolute errors from the predictions obtained by the models of moisture, ash, lipid, protein and carbohydrate contents of the 30 samples of the prediction set (Fig. 7) are presented.

Considering that the value of the bias is not enough to emphasize the difference between the mean values obtained by the conventional methods and the predicted values obtained by FT-NIRS technique, the absolute error demonstrated uniformity for all the parameters with the exception of the carbohydrate values, which presented higher differences, and the lipid content, which despite the lower unities, the values estimated by the model were higher than the content obtained by the reference method (Fig. 7). Nevertheless, low values to bias and RMSEP suggest that the lipid content can be predicted by NIR, but not so accurately as the other parameters.

4. Conclusion

This study demonstrated that it is possible to predict the parameters of the quality of moisture, ashes, proteins, lipids and carbohydrates in soybeans by associating near-infrared spectroscopy with multivariate calibration.

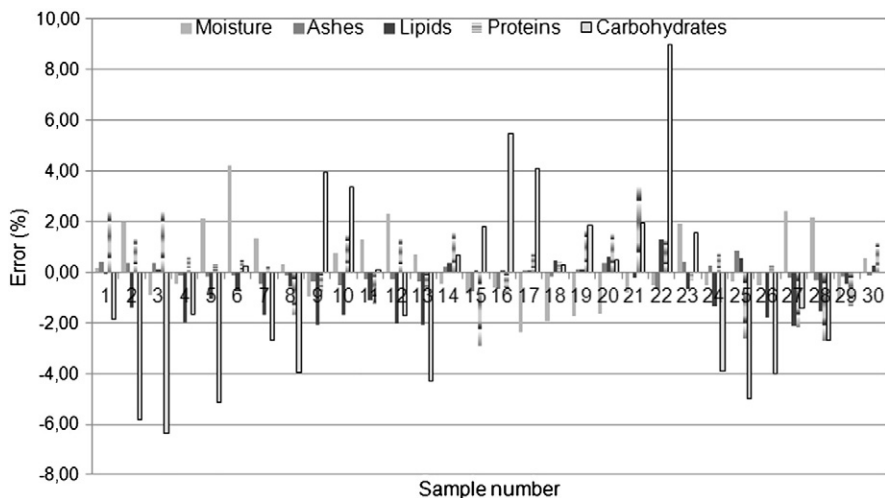


Fig. 7. Absolute errors: evaluated parameters measured minus the predicted values of the soybean validation group.

It can be concluded that despite the existence of a wide range of studies on the FT-NIRS technique, the methodology proposed in this work to develop a NIR model for a grain, controlling the quality of the soybean, has a high potential for implanting and developing calibration models in industries. The offer of the technique to inspection organs can be a solution for a faster and more efficient control of the quality of a wide range of food in Brazil, enabling the replacement of laborious methods.

Acknowledgments

The authors thank the EMBRAPA Soybean for providing the soybean samples and the FAPESP (2010/50418-5) and CNPq for the financial support.

References

- Ageel, L. E., & Hurburgh, C. R. (2010). A tutorial on near infrared spectroscopy and its calibration. *Critical Reviews in Analytical Chemistry*, 40(4), 246–260.
- AOAC (2006). Official methods of analysis of AOAC International. In W. Horwitz, & G. W. Latimer (Eds.), (18th ed.). Gaithersburg, Maryland: AOAC International (Vol. 18th Ed).
- Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy*, 43(5), 772–777.
- Bjorsvik, H. -R., & Martens, H. (2007). Calibration of NIR instruments by PLS regression. In E. W. Ciurczak, & D. A. Burns (Eds.), *Handbook of near-infrared analysis* (pp. 189–205). (3rd ed.). New York, USA: CRC Press Taylor & Francis Group.
- Bligh, E. G., & Dyer, W. J. (1959). A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology*, 37(8), 911–917.
- Brown, C. D., Vega-Montoto, L., & Wentzell, P. D. (2000). Derivative preprocessing and optimal corrections for baseline drift in multivariate calibration. *Applied Spectroscopy*, 54(7), 1055–1068.
- CONAB – Companhia Nacional de Abastecimento (2012). Retrieved May 8, 2012, from <http://www.conab.gov.br>
- EMBRAPA (2012). Brazilian Agricultural Research Corporation. Retrieved July 9, 2012, from <http://www.cnpso.embrapa.br/>
- Geladi, P., & Kowalski, B. R. (1986). Partial least-squares regression: A tutorial. *Analytica Chimica Acta*, 185, 1–17.
- Geladi, P., MacDougall, D., & Martens, H. (1985). Linearization and scatter-correction for near-infrared reflectance spectra of meat. *Applied Spectroscopy*, 39(3), 491–500.
- Haaland, D. M., & Thomas, E. V. (1988). Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry*, 60(11), 1193–1202.
- Haughey, S. a., Graham, S. F., Cancouët, E., & Elliott, C. T. (in press). The application of Near-Infrared Reflectance Spectroscopy (NIRS) to detect melamine adulteration of soya bean meal. *Food Chemistry*, 1–5. <http://dx.doi.org/10.1016/j.foodchem.2012.01.068>.
- Holse, M., Larsen, F. H., Hansen, Å., & Engelsen, S. B. (2011). Characterization of marama bean (*Tylosema esculentum*) by comparative spectroscopy: NMR, FT-Raman, FT-IR and NIR. *Food Research International*, 44(1), 373–384.
- Jimaré Benito, M. T., Bosch Ojeda, C., & Sanchez Rojas, F. (2008). Process analytical chemistry: Applications of near infrared spectrometry in environmental and food analysis: An overview. *Applied Spectroscopy Reviews*, 43(5), 452–484.
- Kandala, C. V., Sundaram, J., & Puppala, N. (2012). Analysis of moisture content, total oil and fatty acid composition by NIR reflectance spectroscopy: A review. *Lecture Notes in Electrical Engineering*, 146, 59–80.
- Kennard, R. W., & Stone, L. A. (1969). Computer aided design of experiments. *Technometrics*, 11(1), 137–148.
- Kovalenko, I. V., Rippke, G. R., & Hurburgh, C. R. (2006). Determination of amino acid composition of soybeans (*Glycine max*) by near-infrared spectroscopy. *Journal of Agricultural and Food Chemistry*, 54(10), 3485–3491.
- Kuligowski, J., Carrión, D., Quintás, G., Garrigues, S., & de la Guardia, M. (2012). Direct determination of polymerised triacylglycerides in deep-frying vegetable oil by near infrared spectroscopy using partial least squares regression. *Food Chemistry*, 131(1), 353–359.
- Munck, L., Jespersen, B. M., Rinnan, Å., Seefeldt, H. F., Engelsen, M. M., Nørgaard, L., et al. (2010). A physicochemical theory on the applicability of soft mathematical models—experimentally interpreted. *Journal of Chemometrics*, 24(7–8), 481–495.
- Naes, T., Isaksson, T., Fearn, T., & Davies, T. (2002). Outlier detection. A user-friendly guide to multivariate calibration and classification. Chichester, UK: NIR Publications.
- Salgó, A., & Gergely, S. (2012). Analysis of wheat grain development using NIR spectroscopy. *Journal of Cereal Science*, 56(1), 31–38.
- Sorensen, L. K. (2009). Application of reflectance near infrared spectroscopy for bread analyses. *Food Chemistry*, 113(4), 1318–1322.
- Sun, J. M., Han, F. X., Yan, S. R., Yang, H., & Sato, T. (2008). Rapid determination of fatty acids in soybeans [*Glycine max* (L.) Merr.] by FT-near-infrared reflectance spectroscopy. *Guang Pu Xue Yu Guang Pu Fen Xi/Spectroscopy and Spectral Analysis*, 28(6), 1290–1295.
- Szigedi, T., Lénárt, J., Dernovics, M., Turza, S., & Fodor, M. (2011). Protein content determination in *Brassica oleracea* species using FT-NIR technique and PLS regression. *International Journal of Food Science and Technology*, 47(2), 436–440.
- Williams, P. (2001). Near-infrared technology: In the agricultural and food industries. In P. Williams, & K. Norris (Eds.), (2nd ed.). *American Association of Cereal Chemists*. (pp. 296) St. Paul, Minn., USA: American Association of Cereal Chemists.
- Williams, P. C., & Sobering, D. (1995). How do we do it: A brief summary of the methods we use in developing near infrared calibrations. In A. M. C. Daves, & P. C. Williams (Eds.), *Near infrared spectroscopy: The future waves* (pp. 185–188). Chichester, UK: NIR Publications.
- Windham, W. R., Mertens, D. R., & Barton, F. E., II (1989). Protocol for NIRS calibration: Sample selection and equation development and validation. In C. G. Marten, J. S. Shenk, & F. E. Barton (Eds.), *Near infrared reflectance spectroscopy (NIRS): Analysis of forage quality* (pp. 96–103). (643rd ed.). Washington, DC: USDA-ARS: Agricultural Handbook.
- Zhou, X., Yang, Z., Huang, G., & Han, L. (2012). Non-invasive detection of protein content in corn distillers dried grains with solubles: Method for selecting spectral variables to construct high-performance calibration model using near infrared reflectance spectroscopy. *Journal of Near Infrared Spectroscopy*, 20(3), 407.
- Zhu, D., Wang, K., Zhang, D., Huang, W., Yang, G., Ma, Z., et al. (2011). Quality assessment of crop seeds by near-infrared hyperspectral imaging. *Sensor Letters*, 9(3), 1144–1150.