ISSN 0101-2061 (Print)
ISSN 1678-457X (Online)
Food Science and Technology

DOI: https://doi.org/10.1590/fst.02119

Prediction of quality parameters of food residues using NIR spectroscopy and PLS models based on proximate analysis

Magale Karine Diel RAMBO[1*], Márcia Miguel Castro FERREIRA[2], Polyana Morais de MELO[3], Claúdio Carneiro SANTANA JUNIOR[4], Daniel Assumpção BERTUOL[5], Michele Cristiane Diel RAMBO[6]

Abstract

The real-time prediction in biorefinery industries has become essential. Models using partial least square regression (PLS) were developed to predict moisture, ash, volatile matter, fixed carbon and organic matter of coconut and coffee residues. In this study, 49 samples were collected and near infrared spectroscopy were applied to predict moisture, ash, volatile matter, fixed carbon and organic matter. For external validation 25% of the set samples were used. Moisture and volatile matter were predicted with coefficients of determination (R^2cal) above 0.90, and standard errors (RSD) of the estimate of 14.4% and 2.26%, respectively. Models of ash and organic matter show R^2cal > 0.77 and RSD values < 20.4%. For the external validation, the low deviations show the approximation between reference and predicted values and good prediction with R^2pred > 0.70. All calibration models were acceptable for sample screening. This study demonstrates that PLS can be used to predict biomass composition of different species, with very low costs and time.

Keywords: biomass; figures of merit; pre-treatments; external validation.

Practical Application: The Production of Sustainable Bioproducts/Fuels from the food residues of Brazil helps to reduce the need for fossil diesel imports. Therefore rapid analytical methods such as spectroscopy and chemometrics can be used to analysis feedstocks in food industries and thus identifying biomass with potential.

1 Introduction

The proximate composition of biomass is necessary to analyze the overall process of any thermochemical conversion. The proximate composition is an important property for food industries utilization and for heat and power generation (Hosseinpour et al., 2017). Focusing on the proximate analyses, it is possible to determine a large number of fuel and bioproducts parameters. In this way, the proposed models would assist in the optimal use of the feedstock based on this biomass properties (Feng et al., 2015).

Near infrared spectroscopy (NIRS) is one technique that measure the vibration of functional groups, relating the chemical composition and spectral data that allows for the development of multivariate calibration models (Rambo et al., 2016; Xu et al., 2013). NIRS has been shown to be fast and precise in the prediction of proximate analysis from biomasses using multivariate models (Qi et al., 2016; Rambo et al., 2015a). Such models have been frequently used in combination with different pretreatment methods for treating the complex spectroscopic data.

Recently relationships were developed using proximate analyses and chemometrics (NIRS coupled to multivariate methods). Hosseinpour et al. (2017, 2018) developed a prediction method to estimate higher heating value (HHV) using iterative neural network-adapted partial least squares (INNPLS) and iterative network-based fuzzy partial least squares coupled with principle component analysis (PCA-INFPLS), respectively. Good correlations were found (R_2> 0.86). Uzun et al. (2017) developed a prediction method using artificial neural network models, taking the data from proximate and ultimate analyses. The model presented a considerably higher correlation coefficient (0.96) and low root mean square error (0.375). Qi et al. (2016) using proximate analysis of sawdust biomass with NIR spectroscopy and locally weighted partial least squares, obtained good prediction results (> 0.80).

This study aims to create correlation models for predicting the contents of moisture, ash, fixed carbon (FC), organic matter (OM) and volatile matter (VM) in coconut and coffee biomasses, so facilitating the application of the screening analysis in biorefinery industries.

2 Materials and methods

2.1 Coconut and coffee samples

Thirty two samples of coconut were collected in Brazil; from them, 21 came from the Southeast regions where the others 12 ones came from the North region. Also, 16 coffee husks

Received 04 Feb., 2019
Accepted 23 Jul., 2019
[1] *Programa de Pós-graduação em Ciências Ambientais e Ciências dos Alimentos, Universidade Federal do Tocantins – UFT, Palmas, TO, Brasil*
[2] *Instituto de Química, Universidade Estadual de Campinas – UNICAMP, Campinas, SP, Brasil*
[3] *Programa de Pós-graduação em Ciências dos Alimentos, Universidade Federal do Tocantins – UFT, Palmas, TO, Brasil*
[4] *Curso de Engenharia Ambiental, Universidade Federal do Tocantins – UFT, Palmas, TO, Brasil*
[5] *Laboratório de Processos Ambientais – LAPAM, Departamento de Engenharia Química, Universidade Federal de Santa Maria — UFSM, Santa Maria, RS, Brasil*
*Corresponding author: magalerambo@uft.edu.br

samples (previously collected) were used. From each biomass a sufficient amount was collected, dried at 105 °C, ground using a Romer micromill (RomerLabs, São Paulo, Brazil), sieved for 20 min and the fractions of 48 mesh were used.

2.2 Proximate analysis

The proximate analysis is performed according to the American Society for Testing and Materials (ASTM).

The ASTM D 3173-87 (American Society for Testing and Materials, 2003) was using for the moisture content determination, after the sample was heated in an oven at 105° ± 5 °C (SP 100, SP Labor) during 12 hours, or until constant mass. The volatile matter (VM) was determined by the ASTM D 3175-07 (American Society for Testing and Materials, 2007) using 1g of sample, previously dried in muffle (1200DRP7, SP Labor) at 800 ± 10 °C for 8 minutes. The sample was removed and placed in a desiccator to be cooled for 60 minutes, finally being weighed and its VM content was calculated.

The ASTM D 3174-04 (American Society for Testing and Materials, 2004) was using for the ash content determination, involved the removal of organic constituents through high temperatures in furnace (1200DRP7, SP Labor) by 4 hours at 600 ± 10 °C.

The values of fixed carbon (FC) and organic matter (OM) were indirectly obtained by using the following equations:

$$FC\ (\%) = 100 - VM - Ash \quad (1)$$

$$OM\ (\%): 100 - Ash \quad (2)$$

2.3 Near infrared spectroscopy

The near-infrared spectra (triplicate) were obtained (FOSS, Hillerød, Denmark instrument) between 1100–2500 nm in a diffuse reflectance detector, with 1 nm increments and by averaging 32 successive scans.

2.4 Multivariate data analysis

The UNSCRAMBLER 10.3 software (Camo Software, Oslo, Norway) was using for the multivariate data analyses.

PLS-1 (Martens & Naes, 1996) was used for constituent quantification of moisture, ash, FC, VM and OM. The data were pre-treated by mean-centering and to determine the number of latent variables (LV) in the models, the leave-one-out cross-validation method was used.

Several transformations were applied to choose the pre-treatment with the best results. Initially the raw data was tested, followed by the second (D2) derivative, and by combining the D2 with Detrend (DT) and D2 with standard normal variate (SNV) (Wise et al., 2006). The size of the optimal window to be used in the Savitzky-Golay algorithm (Enke & Nieman, 1976; Savitzky & Golay, 1964) was also evaluated.

Different statistical parameters were evaluated based on the coefficients of determination (R^2_{cal} and R^2_{val}); the root mean square error of calibration (RMSEC); root mean square error of cross validation (RMSECV); root mean square error of prediction (RMSEP); bias; the relative standard deviation (RSD); the relative error of calibration and prediction sets (RE%); the range error ratio (RER), LV numbers and the outliers excluded (using the Student residues versus Leverage). The equations for the calculation of all the figures of merit used in this work are according to Rambo et al. (2013).

To assure a good prediction for new samples, the data set was split into calibration (75% of the samples) and in external validation set (25% of samples), randomly.

According the International Organization for Standardization (1994), the residual distribution and the linearity were analyzed by the graphical evaluation.

According to the ASTM E1655-00 (American Society for Testing and Materials, 2005), the bias and the R^2 values were investigated by a t-test and F-test, respectively.

The regression coefficient plots were interpreted for the parameters ensuring a real correlation and not due to chance.

3 Results and discussion

3.1 Reference analysis

The Figure 1 shows the NIR raw spectra of the coconut and coffee samples. At 1450/1470 nm, 1920 nm, and 2090 nm are located the main absorption bands, associated respectively to H-bonds of the OH groups; O–H stretch from water and the O–H combination from polysaccharides.

Less intense bands appearing at 1170/1270 nm and 2274 nm, associated respectively to lignin and polysaccharides components, in the C–H stretch 2nd overtone and O–H stretch of combination bands (Shenk et al., 2008).

The results for the descriptive analysis are listed in Table 1. The highest variability was observed for FC and ash, both for coconut samples, with a high coefficient of variation (37.1 and 35.0, respectively).

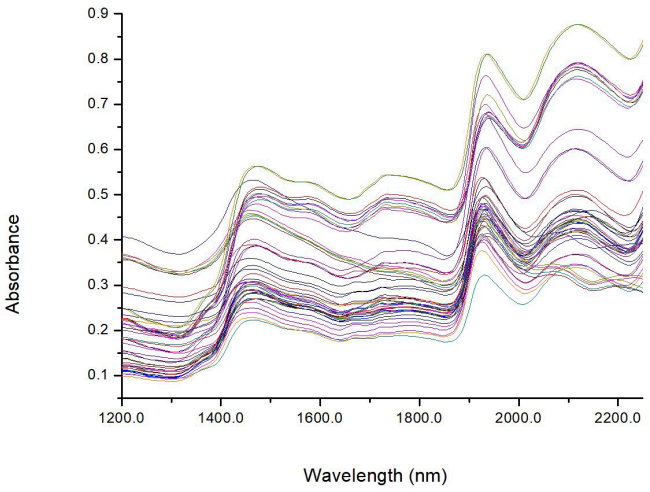

Figure 1. Raw coconut and coffee NIR spectra.

Table 1. Descriptive statistics for proximate analysis (%) of coconut and coffee samples.

Statistic parameters/ constituents	Mean value	Minimum value	Maximum value	Range	Standard deviation	Variance
Moisture	9.56	4.22	13.7	9.49	2.29	5.27
Ash	1.41	0.36	3.24	2.88	0.72	0.52
FC	9.76	2.36	15.55	13.19	3.16	10.0
OM	98.6	96.8	99.64	2.88	0.71	0.51
VM	88.9	82.1	97.17	14.99	3.22	10.4

FC: fixed carbon; OM: organic matter; VM: volatile matter.

Parameters with low variations (VM and OM) were easier to be modeled, because they present a small range for fitting the models. Although the external set to contain only 7 samples, a good variation for all parameters was observed, it is very important for the future prediction for the calibration models.

The mean values found for coconut and coffee samples are consistent with those obtained in literature for these constituents (Balasundram et al., 2017; Oliveira et al., 2018a). The composition of the raw coconut and coffee samples usually varies because it depends on factors such as cultivation conditions, crop variety, and the processing method, which explains the differences in the composition reported by other authors (Araújo et al., 2017; Oliveira et al., 2018b).

3.2 Pre-treatments

The best pre-treatment on the spectra was selected based on the regression models that provided the lowest RE (%) and the highest RER values (Figure 2).

The second one derivative from the Savitzky-Golay algorithm (second-order polynomial) with a window of 7 points (SG7) combined with detrend or SNV for the moisture parameter was the best choice. For the ash and volatile matter models, the D2 (SG7) provided the best results. For the organic matter constituent, the raw spectra presented the highest RER value. Caliari et al. (2017) when comparing different spectral treatment, found that the best choice to be used was taking the first derivative with SG5 in sugarcane biomass when building regression models to estimate the cellulose crystallinity.

Xie et al. (2018) found the multiplicative scatter correction (MSC) + D2 as the best pre-treatment when modeling ash, FC and VM of biochar from very different feedstocks.

3.3 Partial least square regression

To evaluate the prediction performance of the PLS models, the predicted physico-chemical constituents of coconut and coffee biomasses were plotted against reference values for all dataset (Figure 3). The agreement between these values shows that the models fit well and are not over fitted (built with a low number of factors). One forced fit use more PLS components in the model what need, where R^2 have a high value, but the model fail in predicting new samples (Strandberg et al., 2017).

The samples are linearly distributed around a diagonal line in the reference *versus* predicted values by PLS models plot (Figure 3). The PLS residuals are distributed randomly in

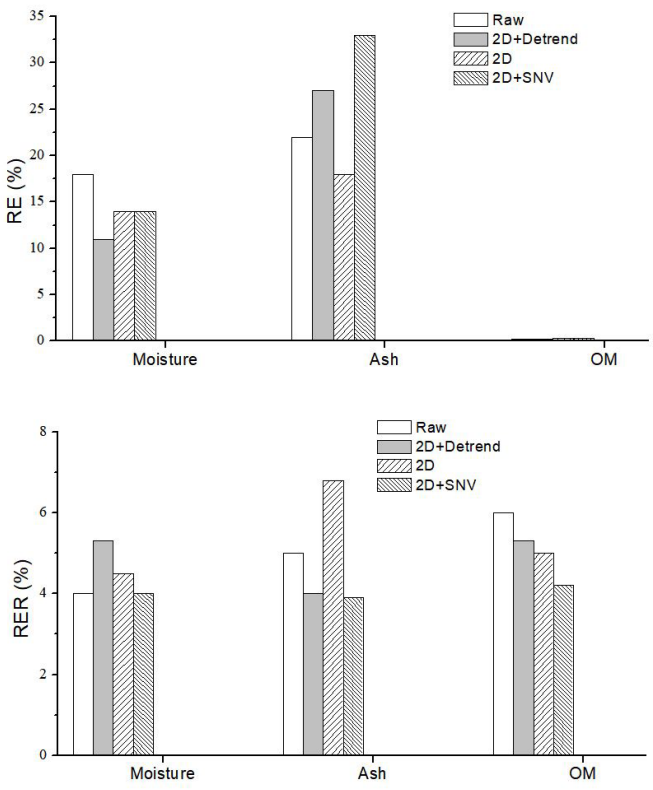

Figure 2. Values of relative error (RE) and range error ratio (RER) for the PLS models built applying different pretreatments.

the plots as shown the Figure 3 (B, D) indicating absence of systematic trends in the building models, that they presented a normal distribution with satisfactory linearity.

From the results summarized in Table 2, it is possible to verify that the models show good correlation coefficient of calibration (R^2_{cal}) cross validation (R^2_{cv}) and prediction (R^2_{pred}) for ash (R^2_{cal} 0.83, R^2_{cv} 0.72 and R^2_{pred} 0.79), and reasonable for moisture and OM. Low (< 2.0%) and quite close RMSEC and RMSECV values were observed, except for OM, where a significant difference was observed. Also, the OM model was not able to predict ($R^2_{pred} < 0.70$). FC and VM were the parameters that presented poor results ($R^2_{cal,val} < 0.70$), independently of the pre-treatment applied. These poor values can be attributed to an indirect measure calculated for these parameters.

All the models were built with a maximum of 7 LV and no more than 8.3% of outliers were remove. It is possible to observe that Gómez et al. (2018) have developed predictive models

for ash, VM and FC using PLS models and fourier transform infrared spectroscopy (FTIR) for coal samples. For the validation parameters were necessary to use 6 and 8 LV, respectively, for ash and VM, and RSD of 20.89% and 1.95% were found. In the present work, similar values of RSD and LV were obtained.

Still, the models obtained in this study when compared with literature data, (Xue et al., 2015) required lower numbers of LV, similar RSD and better $R^2_{cal,cv}$ (except for ash model, however they have used 11 LV). Strandberg et al. (2017), using multivariate analysis for predict the fuel properties of biomass, found low RMSEP (0.54) and high $R^2 > 0.9$ for VM. The RSD was high only for ash model (26.3%). Satisfactory results of RSD were found for the other models (≤14%), as well as good RER values, above > 5.0 indicating the calibration is acceptable for sample screening. It is considered that the RER value is a

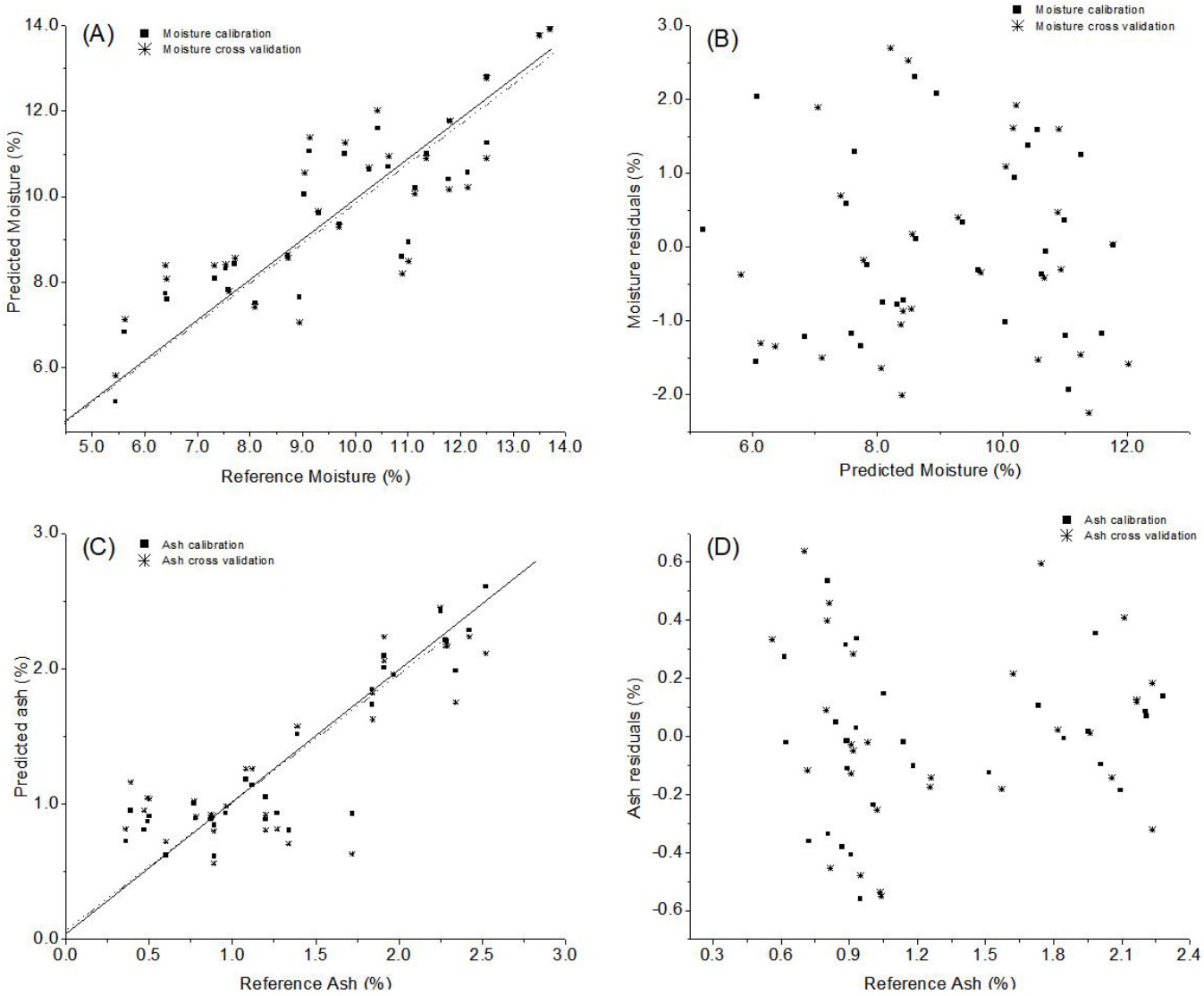

Figure 3. Plot of reference vs. predicted values of calibration and cross validation for (A) moisture and (C) ash of coconut+coffee models. (B) Plot of predicted vs. residual values of cross validation for (B) moisture and (D) ash from coconut+coffee models.

Table 2. Parameters and figures of merit for validation of the best PLS models.

Constituents	Pretreatment	LVs	Outliers	R^2_{cal}	R^2_{cv}	R^2_{pred}	RMSEC	RMSECV	RMSEP	RSD	RER
Moisture	2D(7)+Detrend	4	4	0.74	0.65	0.72	1.085	1.367	1.308	13.3	5.2
Ash	2D(7)	5	3	0.83	0.72	0.79	0.286	0.385	0.367	26.3	5.4
FC	-	-	-	-	-	-	-	-	-	-	-
OM	Raw	7	3	0.86	0.71	-	0.236	0.356	-	0.40	5.7
VM	-	-	-	-	-	-	-	-	-	-	-

FC: fixed carbon; OM: organic matter; VM: volatile matter; LVs: latent variables; R^2_{cal}: coefficient of calibration; R^2_{cv}: coefficient of cross validation; R^2_{pred}: coefficient of prediction; root mean square error of calibration (RMSEC); root mean square error of cross validation (RMSECV); root mean square error of prediction (RMSEP); RSD: standard error; RER: range errror ratio.

better parameter to test the quality of the fit, since no outliers are more present in the models and so the concentration range of the constituent is well represented (Hayes, 2011).

The number of LVs used to explain the variability for the models were high for OM, due to the existence of different types of molecular interactions.

The cross-validation and external validation results showed that the models developed here are promising for prediction future proximate analysis of coconut and coffee biomass. The results showed that one independent validation with new biomass species/samples was able to predict chemical composition and to make the NIR spectroscopy more robust and practical for industrial applications. Liu et al. (2010), used one independent validation including a different species than was used for calibration, and good results were obtained, with RSD < 14%.

Fagan et al. (2011) predicted the moisture, ash and carbon content of two crops (164 samples) using PLS and NIR obtaining a RMSECV above R2 > 0.88, except for ash with a R2 210 of 0.58, 211 demonstrating the application for screening calibration, while Rambo et al. (2016), using 26 samples were able to determine the ash and moisture content with R2 212 >0.80, similar to Gómez et al. 213 (2018) with only 28 samples they got models with a true correlation between the reference and 214 predicted values wit error less than 1.2%.

The results showed that the NIR in combination with PLS is able to quantify the composition of samples using multi-product calibration models.

In order to establish whether there is a true difference between the standard errors values (RMSEC and RMSEP) the F_{test} was applied. All F_{calc} were lower than $F_{tab,}$ (at 95% confidence level), it can be concluded that there is no statistically significant difference between the values, except for the OM model that was inaccurate for external validation. All the t value is less than the critical t value, suggesting that the results provided by the models show the same values as the standard method.

Finally, the low bias values are indicative of the absence of systematic errors, with little significance.

3.4 Regression coefficients interpretation

The interpretation of the regression coefficients is necessary to avoid possible accidental correlations (Rambo et al., 2015b). So, it is important to assign the observed signals to the constituents in question. Figure 4 (A) shows the regression vectors for moisture, where the significant bands at 1350 nm and 1920 nm are assigned to the water. The weak band at 2270 nm (O-H stretch/C-O stretch combination) confirms the contribution of polysaccharides. The negative regions correlated with moisture were mainly due to the presence of carboxylic acids (1885 nm), which together with the alkyl stretching of the first supernatant, can be attributed to fatty acids. Proteins (1640 nm) and cellulose (2000 nm) also present a negative correlation with the moisture content (Yonenobu et al., 2009).

For the ash vector the interpretation of the regression coefficients (Figure 4B) is not straightforward, since the correlation is probably indirect, and the NIR spectroscopy is limited to organic substances. But a more general analysis allows correlating inorganic complex compounds with the ash content, though with spectral bands broader and fewer in number (Rambo et al., 2013a). A correlation with the crystalline (1480 nm) and semi-crystalline cellulose (1488 nm), assigned to 1st overtone of O-H stretch, (Gómez et al., 2018; Zidan et al., 2012) was observed. A negative correlation with water was observed at 1780 nm attributed to H-O-H symmetric bending

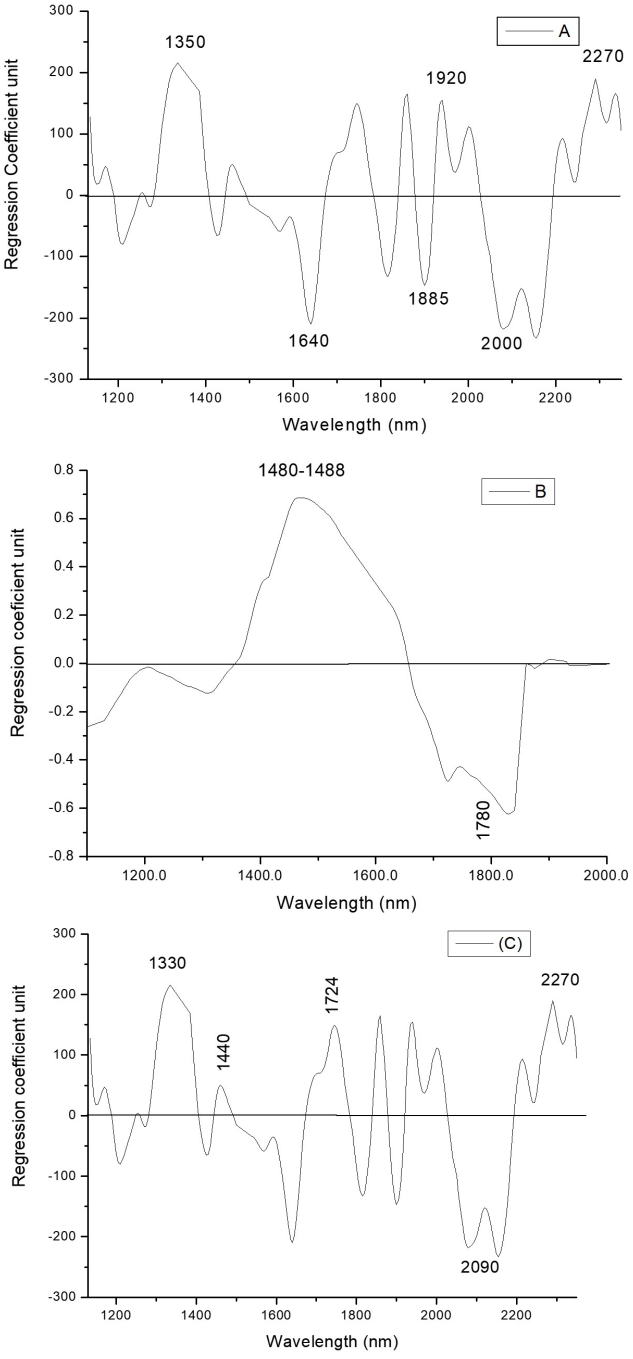

Figure 4. Regression coefficients of the (A) moisture, (B) ash and (C) OM for the PLS models.

Table 3. Predicted and reference values and corresponding deviation for the external coffee set samples in the PLS models of coconut+coffee.

y	Samples	Measu. value	Predicted value	Relative error	y	Samples	Measu. value	Predicted value	Relative error
Moisture	1	12.50	11.83	5.360	Ash	1	1.340	1.116	16.71
	2	8.340	6.948	16.70		2	0.540	0.641	-18.70
	3	11.43	9.466	17.20		3	1.000	0.735	**26.50**
	4	9.670	9.346	3.350		4	1.040	0.934	10.20
	5	7.930	8.728	-10.06		5	0.580	0.740	**-27.60**
	6	6.640	7.590	-14.30		6	1.390	1.115	19.80
	7	11.20	10.87	2.946		7	2.530	2.465	2.60
	8	11.14	12.53	-12.50		8	2.280	2.347	-2.93
	9	5.660	5.392	4.735		9	1.850	2.362	**-27.70**
	10	12.06	9.689	**19.70**		10		*inlier*	
	11	12.40	11.61	6.370		11	1.240	0.933	**24.75**
	12		*inlier*			12		*inlier*	

Measu.value: measurement value.

(Shenk et al., 2008). For the ash regression vector, it is possible to observe that the mineral regions (N-H bend/N-H stretch between 1490-1550 nm) with low frequencies are more influential than organic matter, more prominent in the other regression vectors (Gómez et al., 2018).

The organic matter (Figure 4C) is assigned mainly as carbon bonds; C-H stretch combination of amorphous cellulose (1330 nm), C-H stretch combination (1440 nm), C-H stretch of 1st overtone of furanose/pyranose due to hemicelluloses (1724 nm), and C-O stretch combination of polysaccharides (2270 nm). The negative correlation at 2090 nm is attributed to O-H combination (Shenk et al., 2008; Zidan et al., 2012).

3.5 External validation with coffee samples

The external validation with different species samples is one alternative sampling that avoids some problems, such as the collection of diverse sample, high costs and time-consuming favoring the fast biomass compositional analysis (Rambo et al., 2016), and it is used to evaluate the predictability of the models. They included different biomass species and the relative errors are listed in Table 3.

The ash model had the highest relative error (~ 20% of the mean), whereas the moisture models presented the least deviations (< 20%). With these results, it is possible to conclude that the models are useful and reliable for screening calibration in biorefineries industries, using multiple biomass species, with great variability in the chemical composition.

4 Conclusions

PLS and NIR spectroscopy were useful for predicting fuel properties from proximate analysis of coconut and coffee samples, reducing so the real-time prediction in biorefinery industries. After removal of outlier samples and acquisition of spectra, the best pre-treatment was the D2 (SG7) and D2 (SG7) + Detrend. The approach proposed presented an error of prediction lower than 27.70%, error of calibration < 26.3% and $R^2_{cal,cv,pred}$ higher than 0.70, showing the useful for screening routine 272 analyses (RER > 5), providing fast and inexpensive results for ash, moisture and OM parameters. The regression coefficients have shown a real correlation with the modeled parameter. Thus, the proposed method is useful for screening routine analyses (RER > 5), providing fast and inexpensive results for ash, moisture and OM parameters. For the VM and FC models, the calibration accuracy were inadequate (R^2<0.70) and could not be applied for screening analysis in coconut/coffee biomasses.

Acknowledgements

The authors thanks the Federal University of Tocantins (UFT) for its financial support by the Productivity Research Program (nº 17/2018) and (nº 64/2018), also thanks to Federal Institute of Tocantins – IFTO (nº 51/2018) by the Research Support Program – PAP/INOVA.

References

American Society for Testing and Materials – ASTM. (2003). *Standard method for determination of moisture content in biomass* (ASTM D 3173-87). West Conshohocken: ASTM.

American Society for Testing and Materials – ASTM. (2004). *Standard method for ash in the analysis sample of coal and coke* (ASTM D 3174-04). West Conshohocken: ASTM.

American Society for Testing and Materials – ASTM. (2005). *Standard practices for infrared multivariate quantitative analysis* (ASTM E 1655-00). West Conshohocken: ASTM.

American Society for Testing and Materials – ASTM. (2007). *Standard method for volatile matter in the analysis sample of coal* (ASTM D 3175-07). West Conshohocken: ASTM.

Araújo, C. K. C., Campos, A. O., Padilha, C. E. A., Sousa, F. C. Jr., Nascimento, R. J. A., Macedo, G. R., & Santos, E. S. (2017). Enhancing enzymatic hydrolysis of coconut husk through Pseudomonas aeruginosa AP 029/GLVIIA rhamnolipid preparation. *Bioresource Technology*, 237, 20-26. http://dx.doi.org/10.1016/j.biortech.2017.03.178. PMid:28411051.

Balasundram, V., Ibrahim, N., Kasmani, R. M., Hamid, M. K. A., Isha, R., Hasbullah, H., & Ali, R. R. (2017). Thermogravimetric catalytic pyrolysis and kinetic studies of coconut copra and rice husk for possible maximum production of pyrolysis oil. *Journal*

of Cleaner Production, 167, 218-228. http://dx.doi.org/10.1016/j.jclepro.2017.08.173.

Caliari, Í. P., Barbosa, M. H. P., Ferreira, S. O., & Teófilo, R. F. (2017). Estimation of cellulose crystallinity of sugarcane biomass using near infrared spectroscopy and multivariate analysis methods. *Carbohydrate Polymers*, 158, 20-28. http://dx.doi.org/10.1016/j.carbpol.2016.12.005. PMid:28024538.

Enke, C. G., & Nieman, T. A. (1976). Signal to noise ratio enhancement by least-squares polynomial smoothing. *Analytical Chemistry*, 48(8), 705-712. http://dx.doi.org/10.1021/ac50002a769.

Fagan, C. C., Everard, C. D., & McDonnell, K. (2011). Prediction of moisture, calorific value, ash and 299 carbon content of two dedicated bioenergy crops using near-infrared spectroscopy. *Bioresource Technology*, 102(8), 5200-5206. http://dx.doi.org/10.1016/j.biortech.2011.01.087. PMid:21349705.

Feng, Q., Zhang, J., Zhang, X., & Wen, S. (2015). Proximate analysis based prediction of gross calorific value of coals: A comparison of support vector machine, alternating conditional expectation and artificial neural network. *Fuel Processing Technology*, 129, 120-129. http://dx.doi.org/10.1016/j.fuproc.2014.09.001.

Gómez, Y. R., Hernández, R. C., Guerrero, J. E., & Mejía-Ospino, E. (2018). FTIR-PAS coupled to partial least squares for prediction of ash content, volatile matter, fixed carbon and calorific value of coal. *Fuel*, 226, 536-544. http://dx.doi.org/10.1016/j.fuel.2018.04.040.

Hayes, D. (2011). *Analysis of lignocellulosic feedstocks for biorefineries with a focus on the development of near infrared spectroscopy as a primary analytical tool*. Limerick.

Hosseinpour, S., Aghbashlo, M., & Tabatabaei, M. (2018). Biomass higher heating value (HHV) modeling on the basis of proximate analysis using iterative network-based fuzzy partial least squares coupled with principle component analysis (PCA-INFPLS). *Fuel*, 222, 1-10. http://dx.doi.org/10.1016/j.fuel.2018.02.126.

Hosseinpour, S., Aghbashlo, M., Tabatabaei, M., & Mehrpooya, M. (2017). Estimation of biomass higher heating value (HHV) based on the proximate analysis by using iterative neural network-adapted partial least squares (INNPLS). *Energy*, 138, 473-479. http://dx.doi.org/10.1016/j.energy.2017.07.075.

International Organization for Standardization – ISO. (1994). *International standard ISO 5725-1*. Geneva: ISO.

Liu, L., Ye, X. P., Womac, A. R., & Sokhansanj, S. (2010). Variability of biomass chemical composition and rapid analysis using FT-NIR techniques. *Carbohydrate Polymers*, 81(4), 820-829. http://dx.doi.org/10.1016/j.carbpol.2010.03.058.

Martens, H., & Naes, T. (1996). *Multivariate calibration* (1st ed.). New York: John Wiley and Sons.

Oliveira, J. L., Silva, J. N., Martins, M. A., Pereira, E. G., & Oliveira, M. C. T. B. (2018a). Gasification of waste from coffee and eucalyptus production as an alternative source of bioenergy in Brazil. *Sustainable Energy Technologies and Assessments*, 27, 159-166. http://dx.doi.org/10.1016/j.seta.2018.04.005.

Oliveira, F. C., Srinivas, K., Helms, G. L., Isern, N. G., Cort, J. R., Gonçalves, A. R., & Ahring, B. K. (2018b). Characterization of coffee (Coffea arabica) husk lignin and degradation products obtained after oxygen and alkali addition. *Bioresource Technology*, 257, 172-180. http://dx.doi.org/10.1016/j.biortech.2018.01.041. PMid:29500951.

Qi, S., Yan, Y., Jinhui, C., & Yingying, Z. (2016). Proximate analysis of sawdust using Near Infrared Spectroscopy and locally weighted partial least squares. *Energy Procedia*, 88, 600-607. http://dx.doi.org/10.1016/j.egypro.2016.06.085.

Rambo, M. K. D., Alves, A. R., Garcia, W. T., & Ferreira, M. M. C. (2015a). Multivariate analysis of coconut residues by near infrared spectroscopy. *Talanta*, 138, 263-272. http://dx.doi.org/10.1016/j.talanta.2015.03.014. PMid:25863400.

Rambo, M. K. D., Macêdo, L. P. R., Silva, M. M., & Rambo, M. C. D. (2015b). Prediction by multivariate calibration of quality parameters of coffee residues. *Ciência e Natura*, 37, 374-380.

Rambo, M. K. D., Ferreira, M. M. C., & Amorim, E. P. (2016). Multi-product calibration models using NIR spectroscopy. *Chemometrics and Intelligent Laboratory Systems*, 151, 108-114. http://dx.doi.org/10.1016/j.chemolab.2015.12.013.

Rambo, M. K., Amorim, E. P., & Ferreira, M. M. C. (2013). Potential of visible-near infrared spectroscopy combined with chemometrics for analysis of some constituents of coffee and banana residues. *Analytica Chimica Acta*, 775, 41-49. http://dx.doi.org/10.1016/j.aca.2013.03.015. PMid:23601973.

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8), 1627-1639. http://dx.doi.org/10.1021/ac60214a047.

Shenk, J. S., Workman, J. J., & Westerhaus, M. O. (2008). Application of NIR spectroscopy to agricultural products. In D. A. Burns & E. W. Ciurczak (Eds.), *Handbook of near-infrared analysis* (3rd ed.). Florida: CRC Press.

Strandberg, A., Holmgren, P., & Brostrom, M. (2017). Predicting fuel properties of biomass using thermogravimetry and multivariate data analysis. *Fuel Processing Technology*, 156, 107-112. http://dx.doi.org/10.1016/j.fuproc.2016.10.021.

Uzun, H., Yildiz, Z., Goldfarb, J.L., & Ceylan, S. (2017). Improved prediction of higher heating value of biomass using an artificial neural network model based on proximate analysis. *Bioresource Technology*, 234, 122-130. http://dx.doi.org/10.1016/j.biortech.2017.03.015. PMid:28319760.

Wise, B. M, Gallagher, N. B., Shaver, J. M., Windig, W., & Koch, R. S. (2006). *PLS toolbox version 4.0 for use with MATLABTM*. Wenatchee: Eigenvector Research.

Xie, Y., Li, F.-Y., Fan, X.J., Hu, S.-J., Xiao, X., & Wang, J.-F. (2018). Components analysis of biochar based on near infrared spectroscopy technology. *Chinese Journal of Analytical Chemistry*, 46(4), 609-615. http://dx.doi.org/10.1016/S1872-2040(17)61081-8.

Xu, F., Yu, J., Tesso, T., Dowell, F., & Wang, D. (2013). Qualitative and quantitative analysis of lignocellulosic biomass using infrared techniques: a mini-review. *Applied Energy*, 104, 801-809. http://dx.doi.org/10.1016/j.apenergy.2012.12.019.

Xue, J., Yang, Z., Han, L., Liu, Y., Liu, Y., & Zhou, C. (2015). On-line measurement of proximates and lignocellulose components of corn stover using NIRS. *Applied Energy*, 137, 18-25. http://dx.doi.org/10.1016/j.apenergy.2014.09.089.

Yonenobu, H., Tsuchikawa, S., & Sato, K. (2009). Near-infrared spectroscopic analysis of aging degradation in antique washi paper using a deuterium exchange method. *Vibrational Spectroscopy*, 51(1), 100-104. http://dx.doi.org/10.1016/j.vibspec.2008.11.001.

Zidan, A. S., Rahman, Z., Sayeed, V., Raw, A., Yu, L., & Khan, M. A. (2012). Crystallinity evaluation of tacrolimus solid dispersions by chemometric analysis. *International Journal of Pharmaceutics*, 423(2), 341-350. http://dx.doi.org/10.1016/j.ijpharm.2011.11.003. PMid:22100517.