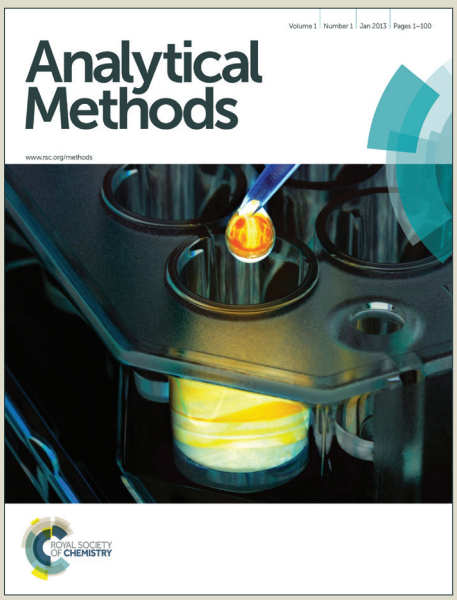
Analytical Methods

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

REVISED MANUSCRIPT

Estimating of Brazilian charcoal properties using attenuated total reflectance-Fourier transform infrared (ATR-FTIR) spectrometry coupled with multivariate analysis

Rafaela M. R. Bezerra[1], Ana C. O. Neves[2], Alexandre S. Pimenta[1], Kássio M. G. Lima[2*]

[1]*Universidade Federal do Rio Grande do Norte, Unidade Acadêmica Especializada em Ciências Agrárias, Grupo de Pesquisa Florestas, Bioenergia e Meio Ambiente, 59072-970 – Natal, RN, Brazil*
[2]*Biological Chemistry and Chemometrics, Institute of Chemistry, Federal University of Rio Grande of Norte, Natal 59072-970, RN-Brazil*

The aim of the present work was to estimate fixed-carbon, volatile matter content and ash in Brazilian commercial charcoal by using attenuated total reflectance-Fourier transform infrared (ATR-FTIR) together with multivariate calibration methods. Several multivariate calibration techniques, including partial least squares (PLS), interval partial least squares (iPLS), genetic algorithm (GA), were compared and validated by establishing significance testing. Charcoal samples (n = 72) were divided into calibration (n = 52) and validation sets (n = 20) by applying the classic Kennard-Stone (KS) selection algorithm to the ATR-FTIR spectra. For fixed-carbon content, the result obtained using PLS-GA for the root mean square error of cross validation (RMSECV) and prediction (RMSEP) were 3.77% and 4.29%, respectively. For volatile matter, RMSECV and RMSEP of 4.36% and 4.65% was achieved by PLS model using seven latent variables (LV). Finally, for ash, RMSECV and RMSEP of 0.58% and 0.38% was achieved by PLS model using eight latent variables (LV). A t-test and Quantile-quantile (Q–Q) plot were performed to compare the results of the models with each other and with a reference method. These results suggest that ATR–FTIR spectroscopy and

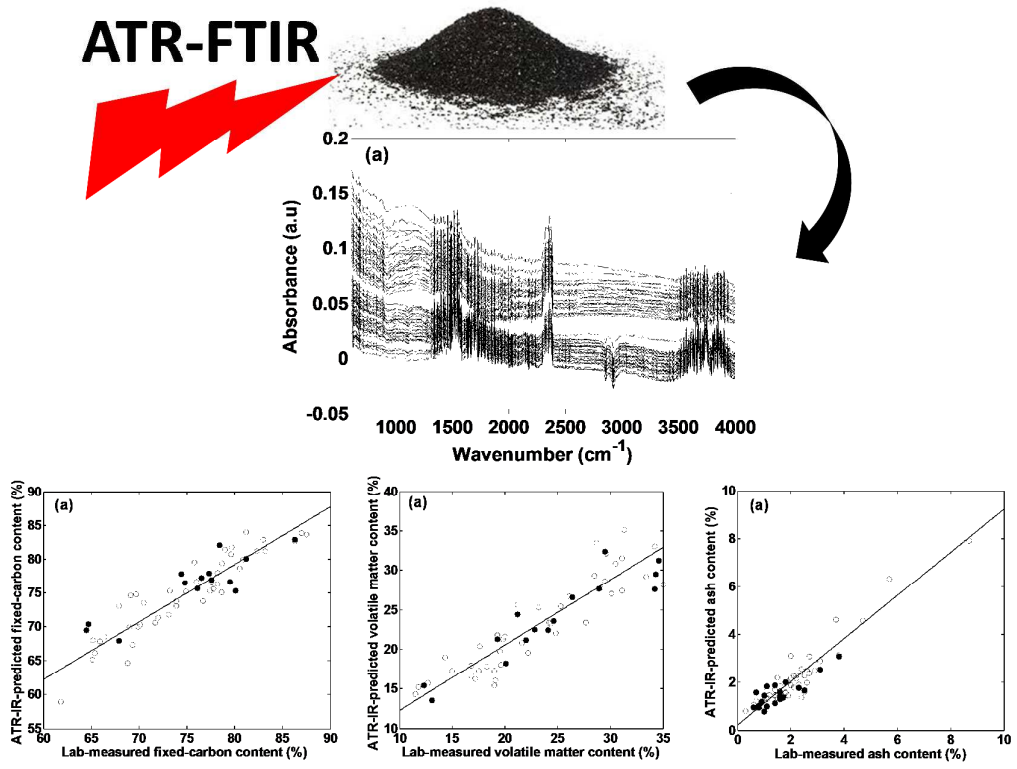
multivariate calibration can be effectively used to determine fixed-carbon, volatile matter content and ash content in Brazilian charcoal.

*Key-Words:*charcoal; Attenuated total reflectance Fourier transform infrared spectroscopy; Near Infrared Spectroscopy; Partial Least Squares; interval-Partial Least Squares; Genetic Algorithm;

* Corresponding author. Tel.: +55 84 3342 2323

E-mail address: kassiolima@gmail.com (K.M.G Lima)

Graphical Abstract



Introduction

Charcoal is the residue of solid non-agglomerating organic matter, of vegetable or animal origin, that results from carbonization by heat in the absence of air at a temperature above 300 degrees celsius.[1,2]The characteristics of wood charcoal are effectively associated to the chemical structures formed during the heating process.[3,4] The structure of charcoal is believed to be closely related to that of activated carbon (AC), which is primarily composed of short stacks of graphene sheets rimmed with O-containing groups (-OH, -CO_2H, -O-, =O, -CHO, etc.) to form a microporous network.[5] Chemical (fixed-carbon, volatile matter, ash, sulfur and phosphorus) and physical (hardness, specific weight, yield and moisture) properties are greatly influenced by three factors – raw material type,[6] process characteristics, and after-treatment.[7]

Most of the techniques used for chemical and physical properties in charcoal are time-consuming and expensive, while rapid methods that require little or no sample preparation are needed for large scale surveys. Alternative methods such as X-ray photoelectron spectroscopy,[8]emission scanning electron microscopy,[9]Near Infrared spectroscopy,[10] Nuclear Magnetic Resonance spectroscopy,[11] Raman spectroscopy[12] and Infrared spectroscopy[13–15]can suitably replace usual physico-chemical analysis. Specifically, one recent development in FTIR techniques applied to coal is the incorporation of an attenuated total reflectance (ATR) crystal (~100 μm in diameter), in which the standard polished block samples can be used without further sample preparation.[16]

Furthermore, ATR-FTIR spectroscopy can distinguish components (macerals) of charcoal which have diverse chemical compositions and physical properties, quantifying the abundance of chemical functional groups.[13] Limited studies have applied ATR-FTIR

spectra for qualitatively evaluating charcoal. For instance, Guo and Bustin[17] have used ATR-FTIR to establish the relationships between both temperature and duration of heating of charcoal formation, reflectance values and spectral characteristics of charcoals such as coalification maturation. In addition, reflectance and FTIR spectra indicate that fungal-decayed wood is particularly susceptible to formation of charcoal and thus inertinite. Labbé and colleagues[13] employed ATR-FTIR to investigate the chemical structure of charcoal made from different maple species: sugar maple (*Acer saccharum*), red maple (*Acer rubrum*), and silver maple (*Acer saccharinum*).In the second part of the study, the authors investigated the effect of thermal treatments on the chemical structure of white oak to have a better understanding of maturation process in toasted charred white oak barrels.

However, the use of appropriate tools for multivariate calibration is mainly responsible for the advancement of the ATR-FTIR technique to give a complete and fast characterization of charcoal or coal including partial least squares (PLS)[18] and methods based on the selection of variable intervals or spectral bands, such as iPLS (interval partial least squares),[19] GA (genetic algorithm)[20] and successive projection algorithm (SPA).[21]These last methods eliminate variables that do not directly correlate with the property of interest. They also eliminate potential interferences and variables that generate a lower signal/noise ratio, which is indicative of low sensitivity.

Herein, we have attempted to make a comparison of the full-spectrum PLS (full-PLS), interval PLS (iPLS), successive projection algorithm (SPA), and genetic algorithm-PLS (GA-PLS), using ATR-FTIR for estimating the properties of charcoal made from wood such as volatile matter, fixed carbon content and ash. The topic of wavelength selection is of particular importance in ATR-FTIR spectra since it generally shows relatively high sensitivity to small perturbations in the experimental conditions,

as well as the physical and chemical properties of samples. To our knowledge, there is no report presenting ATR-FTIR-based calibrations for estimating the properties of charcoal made from wood such as volatile matter, fixed-carbon content and ash. Wavelength selection with interval base algorithms such as iPLS, SPA and GA are also not mentioned for charcoal analysis.

Materials and methods

Samples of charcoal

In this study, seventy-two commercial charcoal samples acquired from some locations in and around the Natal-Brazil region are the whole sample set. All samples were ground to a particle size of 40 mesh with a Wiley mill (Thomas Scientific, Philadelphia, PA).

Chemical properties

Fixed-carbon, volatile matter and ash content were determined using the proximate chemical analysis of wood charcoal according to the procedure D-1762-84 of ASTM[22] and ABNT NBR 8112/83[23],while fixed carbon content was calculated following the equation of Anon.[24] To determine volatile matter, the furnace was then preheated to 950°C with the vent port capped. Samples were introduced into the furnace as quickly as possible, rather than preheating crucibles by placing them on the outer ledge of the furnace for 2 min, then on the edge of the furnace for 3 min, as described in the ASTM method. It should be noted that the furnace temperature would not rebound to 950°C until approximately 8 min after the samples had been introduced. Samples were removed from the furnace after 10 min and placed on a refractory brick to cool until they could be safely transferred into dessicators, at a point above 200°C. Covered

crucibles were weighed after cooling to ambient temperature and volatile matter content was calculated as follows:

$$Volatile\ matter\ \% = \frac{weight_{105°C\ dried} - weight_{950°C\ devolatilized}}{weight_{105°C\ dried}} \times 100 \quad (1)$$

For determination of ash contents, covers were removed from the crucibles and the furnace vent port was connected to the fume hood exhaust. Following this, samples were placed in the furnace and the temperature was increased from 105°C to 750°C at 5°C/min, and then held at 750°C for 6 hours. The furnace was allowed to cool to 105°C before samples were transferred to dessicators. Ash content was determined by weight loss according to the following:

$$Ash\ \% = \frac{weight_{residue\ after\ 750°C}}{weight_{105°C\ dried}} \times 100 \quad (2)$$

Volatile and ash contents were used to calculate the fixed carbon content according to the following:

$$Fixed\ carbon\ \% = \frac{weight_{105°C\ dried} - weight_{950°C\ devolatilized} - weight_{residue\ after\ 750°C}}{weight_{105°C\ dried}} \times 100 \quad (3)$$

It should be noted that the so-called fixed carbon content is given as the mass residue, and is not strictly a C content. All analyses were done in duplicate. Further details of the samples, including chemical analysis of the individual charcoal and reference method used in each parameter, are summarized in Table 1.

ATR-FTIR spectra measurement

Spectral measurements were performed using a Bruker ALPHA FTIR spectrometer equipped with an ATR accessory. Spectra (8 cm^{-1} spectral resolution

giving 4 cm^{-1} data spacing equivalent to 258 wavenumbers, co added for 32 scans) were converted into absorbance by Bruker OPUS software. For the infrared measurements, the powder for each sample was placed on the diamond crystal of an ATR accessory. The average value from two different measurements of each sample was properly stored, and the mean spectrum was then calculated for each sample, giving a total of 72 ATR-FTIR spectra. After each measurement, the ATR plate was washed with ethanol (70% v/v) and dried using tissue paper. Cleanliness of the ATR plate was verified by collecting an absorbance spectrum of the crystal using the most recently collected background as a reference. Spectral measurements were done in an acclimatized room under controlled temperature of 22^{o}C, and 60% relative air humidity.

Chemometrics procedure and software

All the data set were exported to MATLABversion7.12 (The Math-Works, Natick, USA). Data analysis was performed using the PLS-toolbox (Eigenvector Research, Inc., Wenatchee, WA, USA, version 7.8). Cross validation was employed to optimize the number of PLS factors and to guide the selection process in PLS models. Before computing variable selections and calibrations, different preprocessing methods were used, including the multiplicative scattering correction (MSC), first and second derivative and smoothing Savitzky-Golay methods by varying the number of window points (3,5,7,9 and 11 points) using a first-order polynomial. Mean centering was applied to all spectra before performing variable subset selection and calibration.

All the samples were divided into calibration and prediction sets using the SPXY algorithm.[25] Then, $n_{calibration}$ = 52 and $n_{prediction}$ = 20 samples were used. To verify the capability of the calibration models based on the selected region by different methods (full-PLS, iPLS, GA-PLS and SPA), each model mentioned above was used to predict the calibration data set and the prediction data set. The RMSEC (root mean square error

of calibration), RMSEP (root mean square error of prediction) and correlation coefficients of each model for calibration data set (r_c) and prediction data set (r_p) were taken into account. For an ideal model, correlation coefficients (r_c and r_p) should be close to 1 while RMSECV/RMSEP is close to 0. Furthermore, root mean square error of both calibration and prediction samples was proposed for assessing the overall performance of the model. Smaller RMSECV/RMSEP value indicates better model quality. Additionally, we used t-pared statistic for a significant ($P < 0.05$) difference or trend in the concentration of each parameter with reference method. If the *t* calculated is higher than the critical *t*-value at the 95% confidence level, there is evidence that the bias included in the multivariate model is significant. The Quantile-quantile (Q–Q) plot compares the ordered distribution of a test sample with the quantiles of a standard Normal distribution indicated by the straight line. If the sample is Normally distributed, the points will lie along this line.[26]

Results and discussion

ATR-FTIR spectroscopic charcoal properties

The original spectra (calculated from the average between the two readings) giving a total of 72 ATR-FTIR spectra, are shown in Fig. 1a. As can be seen in the ATR-FTIR spectra of charcoal containing information regarding its chemical composition and molecular structure, there were clear variations in the IR spectra of the charcoal samples. Although the direct interpretation of IR spectrum is complicated, it is possible to assign some bands (inorganic structures) such as O–H stretching modes (3700−3600 cm^{-1}), aliphatic C–H stretching modes (2900−2800 cm^{-1}), Si–O stretching modes (1100−900 cm^{-1}) stretching modes of aromatic rings and carbonyl groups (1600

and 1400 cm^{-1}), C−O−C stretching (1030 cm^{-1}), aromatic C–H (900−700 cm^{-1}), Si–O–Si and Si–O–Al bending modes (700−400 cm^{-1}).

In addition, Fig. 1a shows baseline shifts and bias present in the spectra; undesirable features which need to be removed using some pre-treatments, such as smoothing (first-order), multiplicative scattering correction (MSC) and first- and second-order derivatives (Savitzky-Golay). Fig. 1b shows 72 ATR-FTIR obtained during the pretreatment stage utilized Savitzky-Golay smoothing (with a window of 5 points), MSC and the first derivative of the Savitzky-Golay polynomial (with a window of 5 points). Mean centering was also applied to all spectra before performing variable subset selection and calibration.

Fixed-carbon content

Fixed carbon content in the charcoal specimen was determined following Equation 3,[24], as the difference between 100 and the sum of moisture content, volatile matter and ash content. In other words, the carbon content can be estimated as a difference; all the other constituents are deducted from 100 as percentages and the remainder is assumed to be the fixed carbon. The results obtained for the calibration and predicted models in the ATR-FTIR region for the fixed-carbon content of commercial charcoal are displayed in Table 2. In addition to the full PLS models, the results of the *i*PLS, PLS-SPA and PLS-GA models are shown. Only the best results from the tested pre-processing techniques are presented. The number of Latent Variables (LV) calculated for each model corresponded to the first minimal residual variance. As can be shown in Table 2, the performance of the full PLS model is slightly better than that of the *i*PLS, GA and SPA models for fixed-carbon content. The correlation coefficient for the prediction set ranged from 0.60 to 0.80. In addition, it was also observed that models

with wavelength selection in the ATR-FTIR spectral region(iPLS, GA and SPA) achieved RMSEP values between 3.01 and 4.56 (%). The number of LV used for the PLS, iPLS, SPA and GA models using ATR-FTIR spectra for fixed-carbon content varied between 5 and 8. The calibration set was optimized by the exclusion of the samples that presented leverage, non-modeled residuals in the parameter (fixed-carbon content). Five outliers were excluded from the calibration set, and the best PLS model for fixed-carbon content achieved RMSECV and RMSEP of 3.06 and 3.05, respectively. In addition, the correlation coefficient for the calibration and validation set for this model were 0.77 and 0.78, respectively, using 5 latent variables.

These results are corroborated by the graph of predicted versus reference values obtained by full PLS using 1666 spectral variables and a correlation coefficient of 0.78 for the prediction set using 5 LV, as shown in Fig. 2a. Moreover, to obtain a better inside of improvement in predictive ability for this model, *t*-test suggested by ASTM E1655-00[27] and normal ($P < 0.05$, Quantile-quantile (Q-Q) plot)were calculated. The results showed that the bias included in the model was not significant, since the *t* value obtained 0.67 for fixed-carbon content was lower that the critical value of 2.14 with 95% of confidence. The Q–Q plot is an excellent graphical test of the Normality of a sample and is commonly used for that purpose. The full PLS model for fixed-carbon content was subjected to the Q–Q plot univariate normality test, and indicated a univariate normal data distribution as show in Fig.2b. It therefore was concluded that the dataset could be multivariate normally distributed.

Volatile matter content

Volatile matter was extracted by pre-heating the specimen in a tube furnace for 2 min at 300°C, then heating for 3 min at 500°C and for 6 min at 950°C. Volatile matter

content was determined as a proportion of the oven-dry weight of the charcoal specimen. Table 3 displays the results for the analysis of the volatile matter of the charcoals. As can be seen for PLS models, better values were obtained for the RMSEP with smoothed data and MSC treatment compared to the models obtained with original raw. For this parameter, the variable selection using the iPLS, the GA and SPA algorithms produced inferior results to those of full PLS. The best model found for this parameter was achieved using full PLS after exclusion of the outliers. When 1666 spectral variables were used to build the full PLS (6) model, we found a correlation coefficient of 0.85 for the prediction set. The plot of laboratory-determined volatile matter versus ATR-FTIR-predicted volatile matter is given in Figure 3a, using 6 VL. We tested the presence of relevant bias with the prediction results for the prediction samples using the full PLS of the *t*-test suggested by ASTM E1655-00. The results showed that the bias included in the model was not significant ($t_{calculated}$= 0.94, $t_{critical}$ = 2.14, 95% confidence level).Volatile matter content was also subjected to the Q–Q plot univariate normality test, and indicated a univariate normal data distribution as shown in Fig.3b.

Ash content

Ash content was calculated as a proportion of the oven-dry weight of the residue to the oven-dry weight of charcoal specimen. Table 4 presents the model statistics of the ATR-FTIR models for ash content. As can be seen in Table 4, the performance of the full PLS model was better than the wavelength selection models (*i*PLS, GA and SPA) for ash content. The correlation coefficient for the prediction set ranged from 0.01 to 0.88. The number of LV used for the PLS, iPLS, SPA and GA models using ATR-FTIR spectra for ash content varied between 3 and 9. The best model found for ash content was achieved using full PLS after exclusion of the outliers and smoothing (5 points

window) and MSC as preprocessing methods. When 1662 spectral variables were used to build the full PLS (8) model, a correlation coefficient of 0.75 for the prediction set was achieved. These results are corroborated by the graph of predicted versus reference values obtained by full PLS, as shown in Fig. 4a.

In addition, to obtain a better inside of the improvement in predictive ability for this full PLS model, the model was not significantly different using prediction samples for ash content when compared with the reference values according to a paired *t*-test ($t_{calculated}$=0.39, $t_{critical}$ = 2.09, 95% confidence level). Lastly, the full PLS model for ash content was subjected to the Q–Q plot univariate normality test, and indicated a univariate normal data distribution as shown in Fig.4b. It therefore was concluded that the dataset could be multivariate normally distributed.

Conclusions

In this study, we demonstrated ATR-FTIR based in full PLS and wavelength variable models (*i*SPA, GA and SPA) for estimating fixed-carbon, volatile matter content and content of commercial charcoal. It can be concluded that ATR-FTIR is a very promising technique for the non-destructive quantification of important parameters in charcoals. An advantage of the ATR method applied to charcoal is that standard polished block samples can be used without further sample preparation. For instance, the full PLS models developed for each parameter can be useful for monitoring charcoal quality in steel industries. These models were validated by cross-validations and independent statistic tests. The findings presented in this paper provide a detailed analytical view of real ATR-FTIR data and they could be applied to other spectral signals as well.

Acknowledgements

The authors would like to acknowledge the financial support from the PPGQ/UFRN/CAPES of a fellowship for Ana Carolina de Oliveira Neves. We are grateful to Fabio Godoy (Bruker Optics Ltd.) for the excellent technical assistance in this study of the Bruker ALPHA FT-IR spectrometer. K.M.G. Lima acknowledges the CNPq/CAPES project (Grant 070/2012 and 442087/2014-4) and FAPERN (Grant 005/2012) for financial support.

References

1 W. Emrich, in *Solar Energy R&D in the European Community*, Springer-Science+Business Media, 1985, p. 278.

2 M. J. Antal, *Ind. Eng. Chem. Res.*, 2003, **42**, 1619–1640.

3 P. L. Ascough, M. I. Bird, P. Wormald, C. E. Snape and D. Apperley, *Geochim. Cosmochim. Acta*, 2008, **72**, 6090–6102.

4 M. Somerville and S. Jahanshahi, *Renew. Energy*, 2015, **80**, 471–478.

5 J. Bourke, M. Manley-Harris, C. Fushimi, K. Dowaki, T. Nunoura and M. J. Antal, *Ind. Eng. Chem. Res.*, 2007, **46**, 5954–5967.

6 A. L. Missio, B. D. Mattos, D. A. Gatto and E. A. de Lima, *J. Wood Chem. Technol.*, 2013, **34**, 191–201.

7 L. Wang, Y. Xin, Z. Zhou, X. Xu and H. Sun, *J. Hazard. Mater.*, 2013, **244-245**, 268–75.

8 K. Nishimiya, T. Hata, Y. Imamura and S. Ishihara, *J. Wood Sci.*, 1998, **44**, 56–61.

9 E. Hobley, G. R. Willgoose, S. Frisia and G. Jacobsen, *Eur. J. Soil Sci.*, 2014, **65**, 751–762.

10 C. Andrade, P. Trugilho, P. Gherardi Hein, J. Lima and A. Napoli, *J. Near Infrared Spectrosc.*, 2012, **20**, 657.

11 A. V. McBeath, R. J. Smernik, M. P. W. Schneider, M. W. I. Schmidt and E. L. Plant, *Org. Geochem.*, 2011, **42**, 1194–1202.

12 O. Francioso, S. Sanchez-Cortes, S. Bonora, M. L. Roldán and G. Certini, *J. Mol. Struct.*, 2011, **994**, 155–162.

13 N. Labbé, D. Harper, T. Rials and T. Elder, *J. Agric. Food Chem.*, 2006, **54**, 3492–7.

14 B. Pizzo, E. Pecoraro, A. Alves, N. Macchioni and J. C. Rodrigues, *Talanta*, 2015, **131**, 14–20.

15 H. Li, Y. Yang, S. Yang, A. Chen and D. Yang, *J. Spectrosc.*, 2014, **2014**, 1–7.

16 J. Thomasson, C. Coin, H. Kahraman and P. M. Fredericks, *Fuel*, 2000, **79**, 685–691.

17 Y. Guo and R. . Bustin, *Int. J. Coal Geol.*, 1998, **37**, 29–53.

18 S. Wold and M. Sjostrom, *Chemom. Intell. Lab. Syst.*, 2001, **58**, 109–130.

19 L. Nørgaard, A. Saudland, J. Wagner, J. P. Nielsen, L. Munck and S. B. Engelsen, *Appl. Spectrosc.*, 2000, **54**, 413–419.

20 M. Ferrand, B. Huquet, S. Barbey, F. Barillet, F. Faucon, H. Larroque, O. Leray, J. M. Trommenschlager and M. Brochard, *Chemom. Intell. Lab. Syst.*, 2011, **106**, 183–189.

21 S. F. C. Soares, A. A. Gomes, A. R. G. F. Filho, M. C. U. Araujo and R. K. H. Galvão, *Trends Anal. Chem.*, 2013, **42**, 84–98.

22 *American Society for Testing and Materials—ASTM, Standard methods for chemical analysis of wood charcoal, D1762-84.*, American Society for Testing and Materials, Philadelphia, PA, USA, 1989.

23 *Associação brasileira de normas técnicas - NBR 8112/83 Carvão vegetal – análise imediata.*, 1983.

24 Anon, *Simple technologies for charcoal making, Forestry paper No. 41*, Food and Agriculture Organization of the United Nations (FAO), Rome, Italy, 1987.

25 R. K. H. Galvão, M. C. U. Araujo, G. E. José, M. J. C. Pontes, E. C. Silva and T. C. B. Saldanha, *Talanta*, 2005, **67**, 736–740.

26 A. R. Henderson, *Clin. Chim. Acta.*, 2006, **366**, 112–29.

27 ASTM International, Ed., *Annual Book of ASTM Standards, Standard Practices for Infrared Multivariate Quantitative Analysis - E1655-00*, West Conshohocken, Pennsylvania, USA., 2000.

Captions for Figure

Figure 1: (a) Original ATR-FTIR average spectra of 72 samples of commercial charcoals. (b) First derivative spectra of the original 72 samples of charcoal after pretreatment (Savitzky-Golay smoothing, MSC and a Savitzky-Golay derivatives).

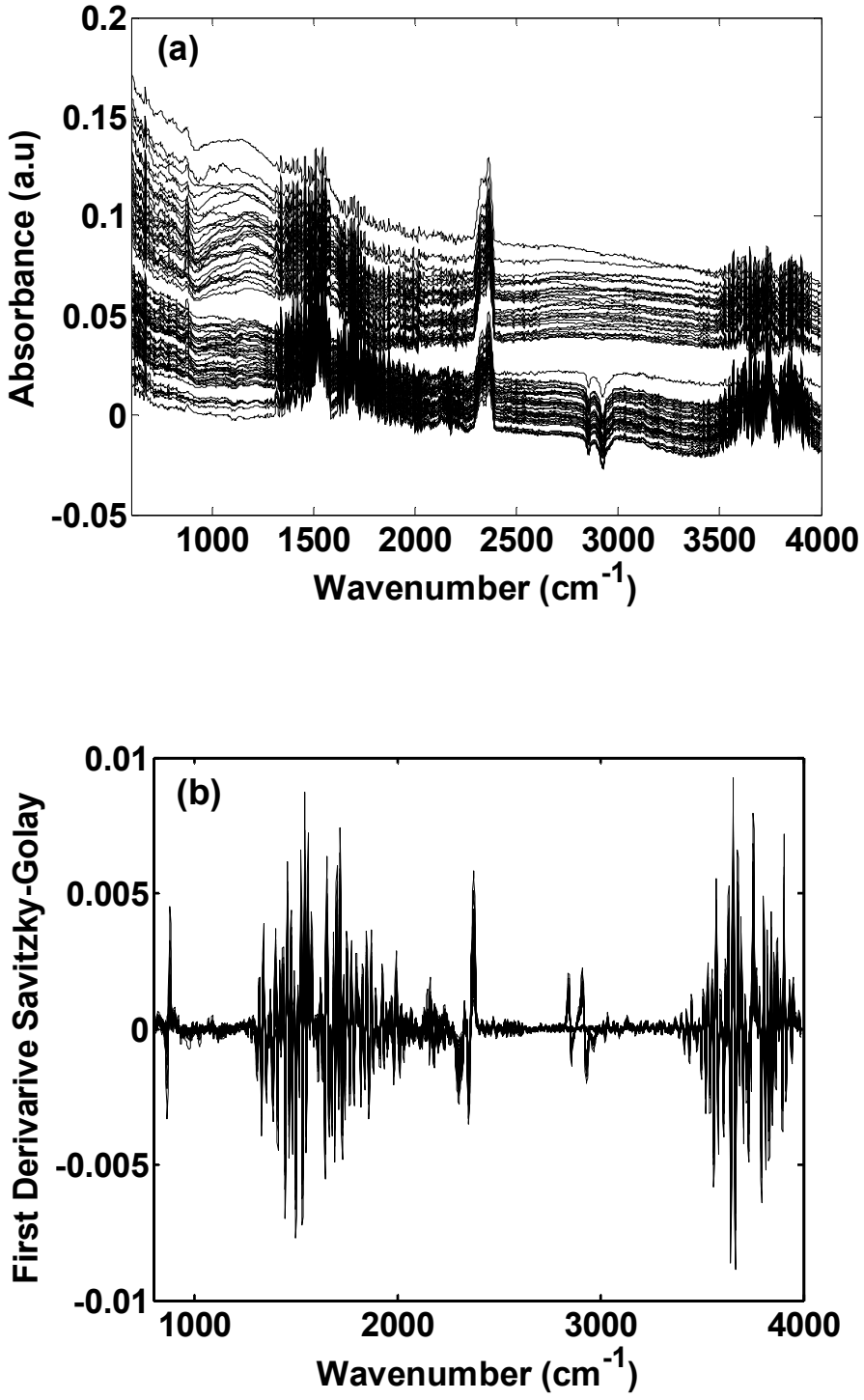
Figure 2 (a) Predicted concentration *vs.* reference measured concentration of calibration and validation samples for fixed-carbon content in commercial charcoals using full PLS model after outlier test, (○) calibration set and (●) prediction set. (b) Quantile-quantile (Q-Q) plot normal distribution for fixed-carbon content.

Figure 3 (a) Predicted concentration *vs.* reference measured concentration of calibration and validation samples for volatile matter content in commercial charcoals using full PLS model after outlier test, (○) calibration set and (●) prediction set. (b) Quantile-quantile (Q-Q) plot normal distribution for volatile matter content.

Figure 4 (a) Predicted concentration *vs.* reference measured concentration of calibration and validation samples for ash content in commercial charcoals using full PLS model after outlier test, (○) calibration set and (●) prediction set. (b) Quantile-quantile (Q-Q) plot normal distribution for ash content.

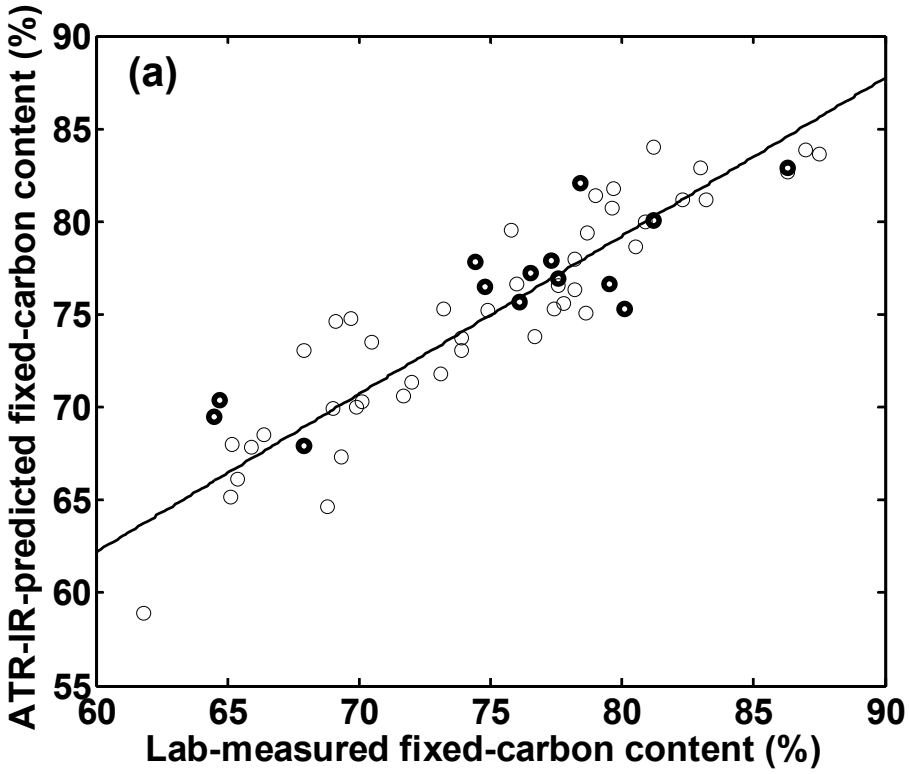
16

Figure 1

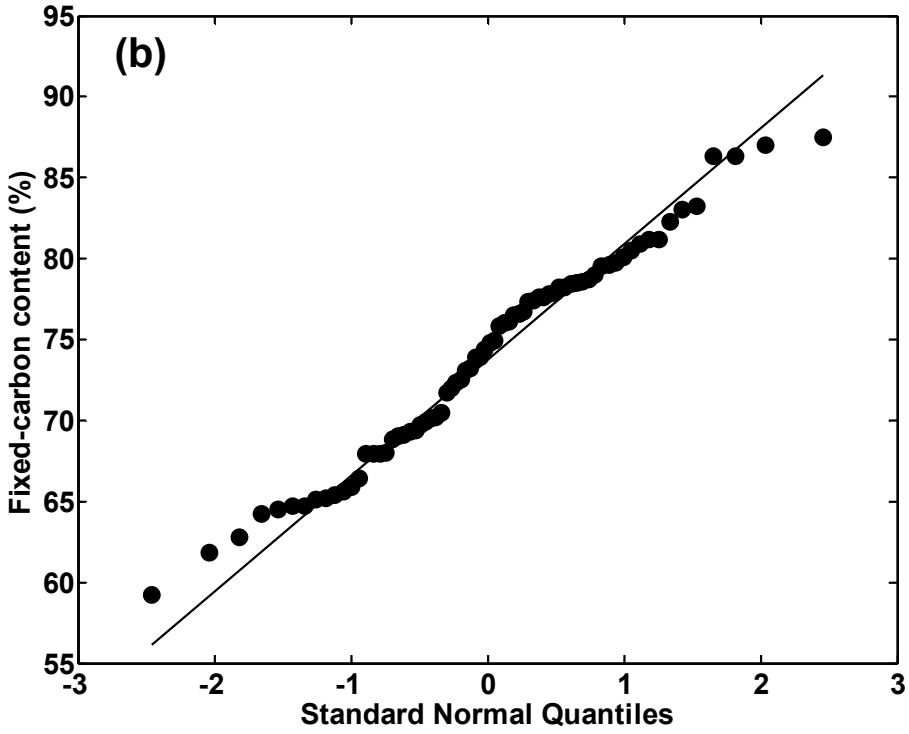


17

386 **Figure 2**

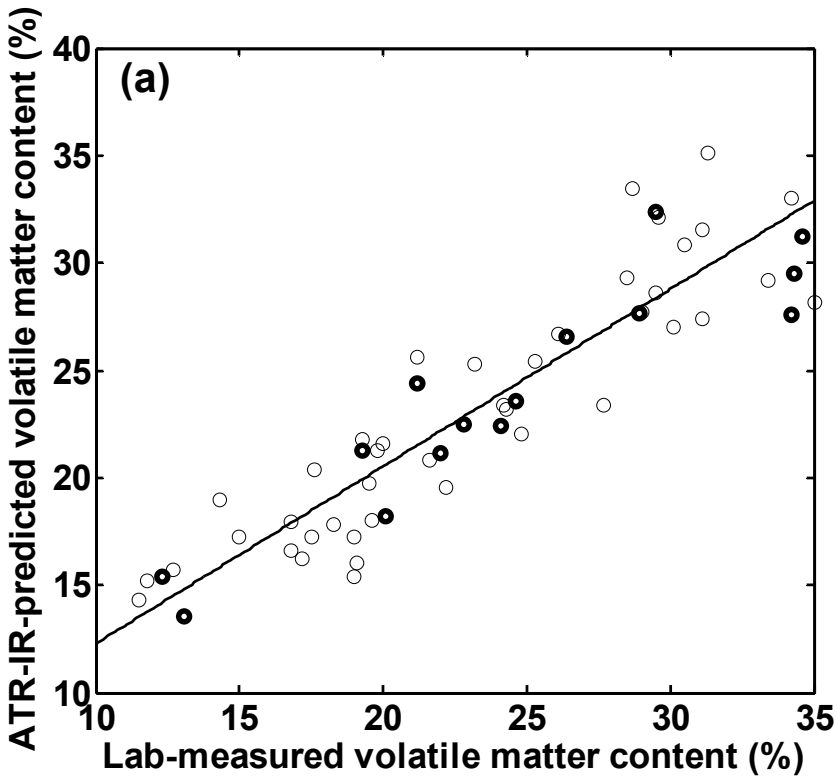


387



388

Figure 3



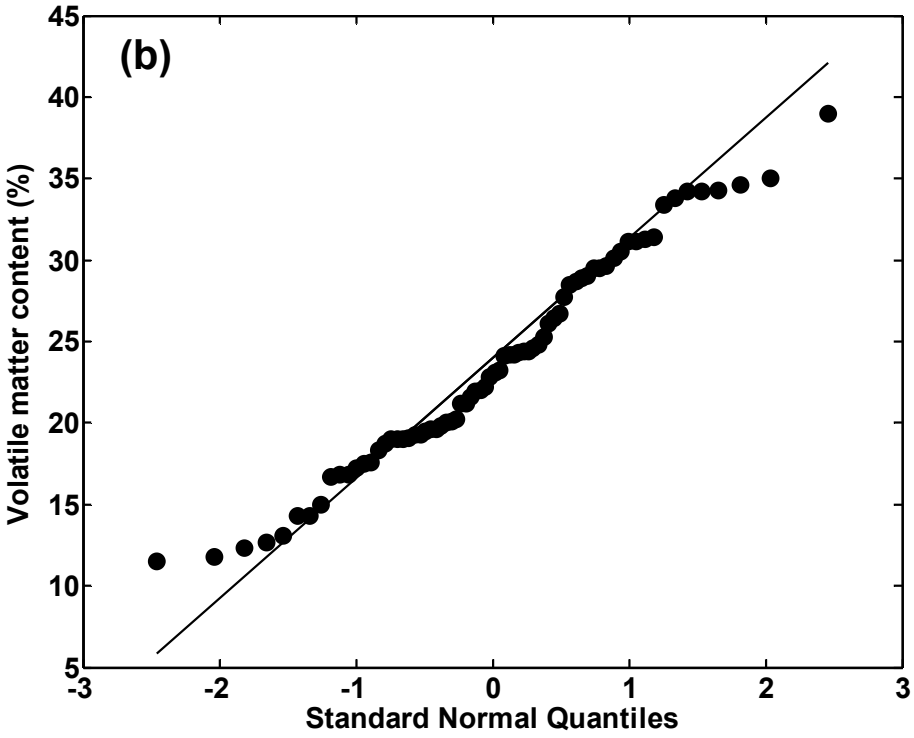
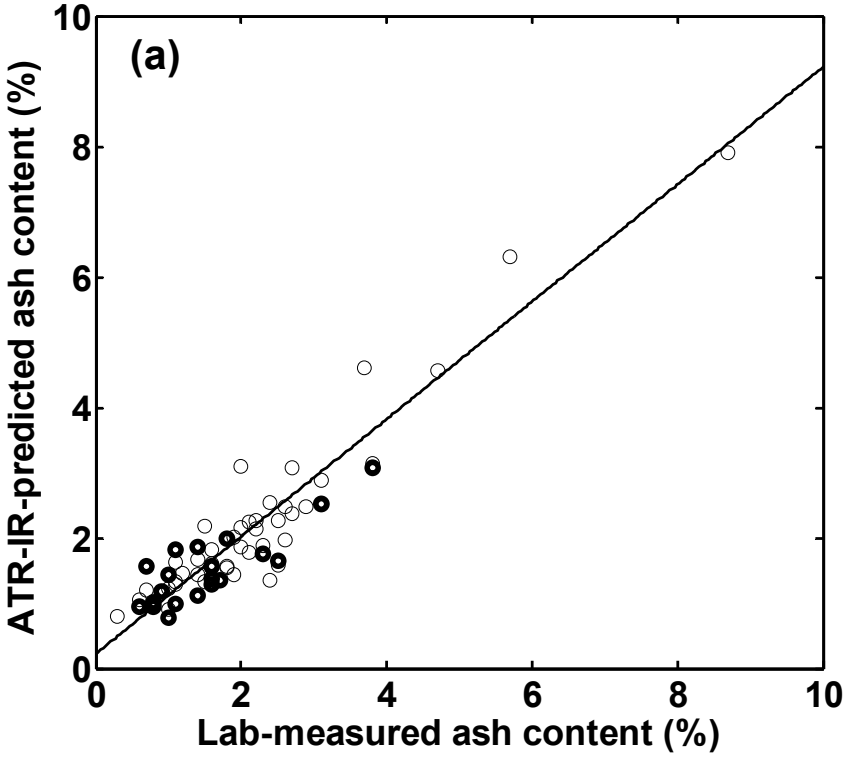


Figure 4



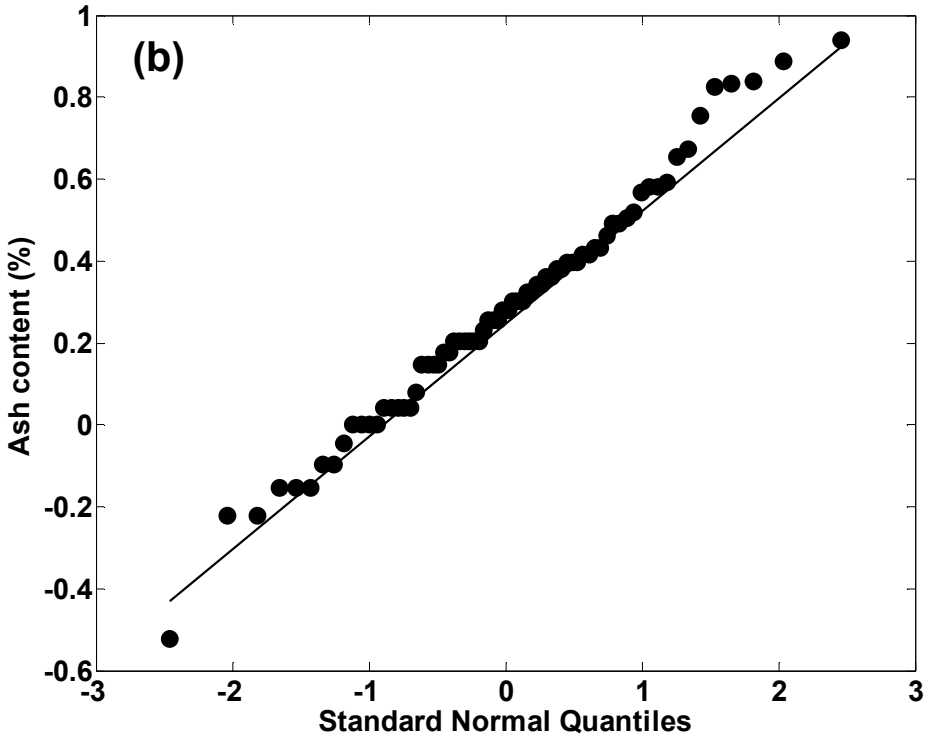


Table 1: Statistical results of full set sample Brazilian charcoal analysis (72 samples) and the reference method applied in each case.

Property	Minimum	Maximum	Mean	S.D.	Reference method
Fixed carbon (%)	59.2000	87.5000	73.9042	6.6124	ASTM D 1762-84
Volatile matter (%)	11.5000	39	23.5028	6.5554	ASTM D 1762-84
Ash (%)	0.3000	8.7000	2.3681	1.7368	ASTM D 1762-84

Table 2: Results for calibration and the external validation set for fixed-carbon content (%): root mean square error of cross validation (RMSECV) and prediction (RMSEP), correlation coefficient for calibration set (r_c) and prediction set (r_p) and the number of used spectral variables (Size). The number of variable latent in PLS, iPLS, PLS-SPA and PLS-GA models are shown in brackets.

Models	Calibration		Prediction		
	r_c	RMSECV (%)	r_p	RMSEP (%)	Size
PLS (6)	0.65	3.92	0.63	4.37	1666
PLS (6)[a]	0.66	3.91	0.63	4.37	1656
PLS (8)[b]	0.57	4.18	0.77	3.99	1658
PLS (8)[c]	0.57	4.59	0.66	3.09	1658
PLS (5)[d]	0.54	4.37	0.71	3.67	1666
*i*PLS (5)	0.44	5.06	0.80	3.01	166
*i*PLS (5)	0.61	4.18	0.65	4.08	833
PLS-SPA (7)	0.72	3.50	0.60	4.56	20
PLS-GA (7)	0.68	3.77	0.66	4.29	407
PLS (7)[e]	0.62	4.04	0.79	3.35	1666
PLS (5)[f]	0.77	3.06	0.78	3.05	1666

[a]smoothing (11 points); [b]first derivative (9 points);[c]smoothing (5 points), first derivative (5 points) and MSC; [d]MSC; [e]One application of outlier detection; [f]second application of outlier detection.

Table 3: Results for calibration and the external validation set for volatile matter content (%): root mean square error of cross validation (RMSECV) and prediction (RMSEP), correlation coefficient for calibration set (r_c) and prediction set (r_p) and the number of used spectral variables (Size). The number of variable latent in PLS, iPLS, PLS-SPA and PLS-GA models are shown in brackets.

Models	Calibration		Prediction		
	r_c	RMSECV (%)	r_p	RMSEP (%)	Size
PLS (7)	0.57	4.36	0.55	4.65	1666
PLS (6)[a]	0.56	4.46	0.56	4.50	1656
PLS (9)[b]	0.61	3.98	0.66	4.33	1656
PLS (8)[c]	0.44	5.09	0.72	3.10	1658
PLS (5)[d]	0.47	4.67	0.81	3.18	1666
*i*PLS (5)	0.56	4.42	0.53	4.59	166
*i*PLS (6)	0.53	4.57	0.58	4.22	833
SPA (7)	0.59	4.28	0.53	4.75	20
GA (7)	0.60	4.23	0.54	4.75	403
PLS (5)[e]	0.65	3.87	0.74	3.43	1666
PLS (6)[f]	0.70	3.51	0.85	2.83	1666

[a]smoothing (11 points); [b]first derivative (9 points); [c]smoothing (5 points), first derivative (5 points) and MSC; [d]MSC; [e]One application of outlier detection; [f]second application of outlier detection.

Table 4: Results for calibration and the external validation set for ash content (%): root mean square error of cross validation (RMSECV) and prediction (RMSEP), correlation coefficient for calibration set (r_c) and prediction set (r_p) and the number of used spectral variables (Size). The number of variable latent in PLS, iPLS, PLS-SPA and PLS-GA models are shown in brackets.

Models	Calibration		Prediction		
	r_c	RMSECV (%)	r_p	RMSEP (%)	Size
PLS (7)	0.72	1.00	0.34	0.65	1666
PLS (8)[a]	0.73	0.97	0.20	0.76	1662
PLS (3)[b]	0.62	1.11	0.79	0.81	1664
PLS (3)[c]	0.64	1.10	0.88	0.71	1658
PLS (6)[d]	0.75	0.96	0.65	0.46	1662
PLS (5)[e]	0.73	1.00	0.38	0.63	1666
***i*PLS (3)**	0.65	1.13	0.65	0.57	166
***i*PLS (5)**	0.53	1.33	0.01	1.42	831
SPA (5)	0.68	1.09	0.34	0.65	20
GA (9)	0.89	0.63	0.67	0.48	413
PLS (8)[f]	0.83	0.77	0.74	0.41	1662
PLS (8)[g]	0.89	0.58	0.75	0.38	1662

[a]smoothing (5 points); [b]first derivative (3 points); [c]smoothing (5 points), first derivative (5 points) and MSC; [d]smoothing (5 points) and MSC; [e]MSC; [f]One application of outlier detection; [g]second application of outlier detection.