Talanta 115 (2013) 563–568

Contents lists available at SciVerse ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Quantitative evaluation of multiple adulterants in roasted coffee by Diffuse Reflectance Infrared Fourier Transform Spectroscopy (DRIFTS) and chemometrics

Nádia Reis [a], Adriana S. Franca [b,*], Leandro S. Oliveira [b]

[a] PPGCA, Universidade Federal de Minas Gerais, Avenida Antônio Carlos 6627, 31270-901 Belo Horizonte, MG, Brazil
[b] DEMEC, Universidade Federal de Minas Gerais, Avenida Antônio Carlos 6627, 31270-901 Belo Horizonte, MG, Brazil

ARTICLE INFO

Article history:
Received 6 March 2013
Received in revised form
3 June 2013
Accepted 5 June 2013
Available online 14 June 2013

Keywords:
Coffee adulteration
Fourier Transform Infrared Spectroscopy
FTIR
Spent coffee grounds
Corn
Barley
Coffee husks
PLS

ABSTRACT

The current study presents an application of Diffuse Reflectance Infrared Fourier Transform Spectroscopy for detection and quantification of fraudulent addition of commonly employed adulterants (spent coffee grounds, coffee husks, roasted corn and roasted barley) to roasted and ground coffee. Roasted coffee samples were intentionally blended with the adulterants (pure and mixed), with total adulteration levels ranging from 1% to 66% w/w. Partial Least Squares Regression (PLS) was used to relate the processed spectra to the mass fraction of adulterants and the model obtained provided reliable predictions of adulterations at levels as low as 1% w/w. A robust methodology was implemented that included the detection of outliers. High correlation coefficients (0.99 for calibration; 0.98 for validation) coupled with low degrees of error (1.23% for calibration; 2.67% for validation) confirmed that DRIFTS can be a valuable analytical tool for detection and quantification of adulteration in ground, roasted coffee.

1. Introduction

Coffee is one of the most widely traded food products and the world's second largest industrial commodity [1]. Such highly-priced commodities are usually a target for adulteration, and ground, roasted coffee, whose appearance can easily be reproduced by roasting and grinding a variety of materials, is rather vulnerable to this type of adulteration [2]. The major adulterants of coffee include by-products of coffee processing such as coffee husks, parchment, spent coffee grounds, cheaper grains (barley, corn, soybean, maize and others), and lower quality coffees [1–5]. Some recent studies have targeted the detection of coffee husks and roasted grains in ground, roasted coffee and instant or soluble coffees [2,6–9]. Although effective, the methods employed (gas chromatography–mass spectrometry, high performance liquid chromatography, high performance anion-exchange chromatography with pulsed amperometric detection, and solid phase microextraction) were time demanding, expensive, laborious, and, in most cases, not appropriate for routine analysis.

Over the last decades, the need for new and rapid analytical methods in the field of food adulteration has prompted extensive research on spectroscopic methods, such as near infrared spectroscopy (NIRS), Raman spectroscopy (RS) and Fourier Transform Infrared (FTIR) spectroscopy [10–12]. Recent applications of such methods to coffee quality analysis include discrimination between Arabica and Robusta species [13,14], discrimination between high and low quality coffees [4,15–17] and discrimination between pure and adulterated coffee samples [1,5]. Ebrahimi-Najafabadi et al. [1] employed NIRS for the identification and quantification of the fraudulent addition of barley to ground, roasted coffee samples. The authors employed different species of coffee (pure Arabica, Robusta and mixtures), with different degrees of roast, and four types of barley at adulteration levels ranging from 2% to 20% w/w of barley. Genetic algorithms were used to determine the spectral regions that would be most useful for identifying the adulteration of coffee with barley. The models presented excellent predictive abilities, with quite low root mean square errors for both calibration (1.4%) and validation (0.8%) sets. The feasibility of applying

Abbreviations: DR, diffuse reflectance; DRIFTS, Diffuse Reflectance Infrared Fourier Transform Spectroscopy; DLATGS, Deuterated Triglycine Sulfate Doped with L-Alanine; LDA, Linear Discriminant Analysis; MSC, multiple scatter correction; NIRS, near infrared spectroscopy; FTIR, Fourier Transform Infrared Spectroscopy; PLS, Partial Least Squares Regression; RMSEC, root mean square error for calibration; RMSECV, root mean square error for cross validation; RMSEP, root mean square error for validation; RS, Raman spectroscopy; SNV, standard normal variates.

* Corresponding author. Tel.: +55 31 34093512; fax: +55 31 34433783.
E-mail addresses: drisfranca@gmail.com, adriana@demec.ufmg.br.(A.S. Franca)

0039-9140/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2013.06.004

Diffuse Reflectance Infrared Fourier Transform Spectroscopy (DRIFTS) for detection of adulteration of coffee was established in a recent study [5]. Two different types of adulterants (roasted coffee husks and roasted corn) were mixed with roasted Arabica coffee under different roasting conditions (light, medium and dark roasts and roasting temperatures ranging from 200 to 260 °C). Linear Discriminant Analysis (LDA) was employed to construct classification models that were able to discriminate between pure coffee and mixtures of coffee, corn and coffee husks. Such models were able to provide complete discrimination (100% recognition and prediction) between pure coffee and adulterated coffee samples at adulteration levels of 10% and above.

It is clear from these studies that spectroscopic techniques offer promise for the detection of adulteration in ground, roasted coffee. However, in the aforementioned studies, only one or two types of adulterants were evaluated, and, the models are only applicable for these specific adulterants. A larger variety of adulterants should be employed when attempting to quantify the adulteration levels to obtain more representative and, therefore, more widely applicable models. In the present study, we sought to confirm the potential of DRIFTS for the detection of multiple adulterants in ground, roasted coffee. The adulterants were coffee by-products (roasted coffee husks and spent coffee grounds) and roasted grains (corn and barley). A Partial Least Squares Regression (PLS) was employed to construct models for the prediction of the levels of adulteration in coffee samples.

2. Experimental

2.1. Samples

Green Arabica coffee (*Coffea arabica*), barley and corn samples were acquired from local markets. Coffee husks (residue obtained after dehulling dried coffee beans) were provided by the Minas Gerais State Coffee Industry Union (Sindicato da Indústria de Café do Estado de Minas Gerais, Brazil). Spent coffee grounds were provided by a local soluble coffee manufacturer (Café Brasília, Minas Gerais, Brazil) and kept frozen (−12 °C) until needed.

Spent coffee grounds (three lots of 2 kg each) were defrosted (18 h at 25 °C) and washed with distilled water to remove impurities. Three 200 g samples were randomly selected from each lot and submitted to drying in a convection oven (Model 4201D Nova Ética, SP, Brazil) at 100 °C for 5 h to reduce the moisture content to that of ground roasted coffee (~5 g/100 g). Coffee beans (50 g), coffee husks (30 g), barley (50 g) and corn samples (30 g) were submitted to roasting in a convection oven (Model 4201D Nova Ética, São Paulo, Brazil) at 200, 220, 240, 250 and 260 °C. The samples were ground ($D < 0.85$ mm) after roasting and submitted to color evaluation. Color measurements were performed using a tristimulus colorimeter (HunterLab Colorflex 45/0 Spectrophotometer, Hunter Laboratories, VA, USA) with standard D_{65} illumination and normal colorimetric observer angle of 10°. Measurements were based on the CIE $L^*a^*b^*$ three-dimensional cartesian (xyz) color space represented by Luminosity (L^*), ranging from 0 (black) to 100 (white) – z axis; parameter a^*, representing the green–red color component – x axis; and parameter b^*, representing the blue–yellow component – y axis. Previous studies [2,5,18] have shown that the degree of roast will be dependent on the type of sample and on the roasting temperature. Preliminary tests showed that it would take higher temperatures (over 240 °C or 250 °C) to roast corn and barley, whereas coffee husks required milder processing conditions (temperatures equal to or below 240 °C). Therefore, roasting conditions were established for each type of sample according to the results of luminosity (L^*) measurements. Previous studies have shown that L^* can be employed as a reference of roasting degree, given that darker roasts will result in coffees with smaller values of luminosity [5,18]. Degrees of roasting were then defined by comparison with commercially available coffee samples ($19.0 < L^* < 25.0$) as light ($23.5 < L^* < 25.0$), medium ($21.0 < L^* < 23.5$) and dark ($19.0 < L^* < 21.0$) roasts.

2.2. FTIR analysis

A Shimadzu IRAffinity-1 FTIR Spectrophotometer (Shimadzu, Japan) with a Deuterated Triglycine Sulfate Doped with L-Alanine (DLATGS) detector was used for the measurements that were all performed in a dry, controlled atmosphere at room temperature (20 ± 0.5 °C). Diffuse reflectance (DR) measurements were performed in diffuse reflection mode with a Shimadzu sampling accessory (DRS8000A). Each sample was mixed with KBr, and 23 mg of this mixture was placed inside the sample port. Pure KBr was employed as the reference material (background spectrum). All spectra were recorded within a range of 4000–400 cm^{-1} with 4 cm^{-1} resolution and 20 scans, and submitted to subtraction of background (pure KBr spectra). They were also truncated to 2500 data points in the range from 3200 to 700 cm^{-1} to eliminate noise present in the upper and lower ends of the spectra. Preliminary tests were performed to evaluate the effect of particle size ($0.39\ mm < D < 0.5$ mm; $0.25\ mm < D < 0.39$ mm; $0.15\ mm < D < 0.25$ mm; and $D < 0.15$ mm) and sample/KBr mass ratio (1%, 5%, 10%, 20% and 50%) on the quality of the spectra. The conditions that provided the best quality spectra (higher intensity and lower noise interference) were $D < 0.15$ mm and 10% sample/KBr mass ratio.

2.3. Data analysis

PLS was employed for quantification of adulterants (pure or mixed) in roasted coffee samples using the DR spectra as chemical descriptors, with adulteration levels ranging from 1% to 66% in mass (see Table 1). To reduce the effect of noise, remove redundant information and enhance sample-to-sample differences, the following data pre-processing (pretreatment) techniques were evaluated: (1) no additional processing (raw data), (2) mean centering, (3) absorbance normalization, (4) absorbance normalization followed by mean centering, (5) area normalization, (6) area normalization followed by mean centering, (7) first

Table 1
Mass composition of adulterated coffee samples.

Sample	Adulteration level	Mass fraction (%)				
		Coffee	Spent coffee grounds	Coffee husks	Barley	Corn
1	66	33.3		33.3		33.3
2	50	50		50		
3	50	50				50
4	40	60	10	10	10	10
5	40	60		20		20
6	40	60	20		20	
7	20	80	5	5	5	5
8	20	80		10		10
9	20	80	10		10	
10	10	90		5		5
11	10	90	5		5	
12	10	90	3.33	3.33		3.33
13	10	90	10			
14	10	90		10		
15	10	90			10	
16	10	90				10
17	1	99	1			
18	1	99		1		
19	1	99			1	
20	1	99				1

derivative followed by smoothing and mean centering, (8) multiple scatter correction (MSC) and (9) standard normal variates (SNV).

Mean centering corresponded to subtraction of the average absorbance value of a given spectrum from each data point. Absorbance normalization consisted of dividing (i) the difference between the absorbance value at each data point and the minimum absorbance value by (ii) the difference between the maximum and minimum absorbance values. Spectra derivatives are commonly used for baseline correction, because they provide visualization of small peaks that are difficult to detect in the original spectra. However, application also leads to a decrease in the signal-to-noise ratio so a smoothing filter (Savitzky–Golay) was employed to provide noise reduction. Multiplicative signal correction (MSC), originally developed to compensate for the effects of light scattering in reflectance spectroscopy, has become a widely employed technique for removing general spectra drift features such as day-to-day intensity variations. SNV is applied to every spectrum individually. Once the mean and standard deviation of all the data points of the spectra are calculated, every data point is subtracted from the mean and divided by the standard deviation [19].

The optimum number of latent variables (LV) for each model was estimated by a cross-validation method (continuous blocks with nine data splits) based on the smallest root mean square error for cross validation (RMSECV). Model performance was measured by evaluation of the root mean square errors for both the calibration (RMSEC) and evaluation (RMSEP) sets, calculated as follows:

$$\mathrm{RMSEC} = \sqrt{\frac{\sum_{t=1}^{I_C}(y_i - \hat{y}_i)^2}{I_C}} \quad (1)$$

$$\mathrm{RMSEP} = \sqrt{\frac{\sum_{t=1}^{I_P}(y_i - \hat{y}_i)^2}{I_P}} \quad (2)$$

where y_i and $\hat{y}_i$ correspond to the actual and predicted adulteration levels of sample i, and I_C and I_P are the total number of samples in the calibration and prediction (validation) sets, respectively.

Model optimization was performed by detection and elimination of outliers. Outliers correspond to samples that are very different from the rest of the data set, and their detection is crucial when developing multivariate models. In this study, outlier detection was based on the method discussed by Botelho et al. [20], which is appropriate for the detection of samples with extreme leverages, e.g., large residuals in the X block (data) or large residuals in the Y block (model response). If a sample presents a leverage (measure of the influence of each sample on the PLS model) larger than a limit value, it is considered an outlier. This limit can be defined as three times the ratio of the number of latent variables to the number of samples [21]. Outliers were detected by evaluating the plots of leverage vs. Student t residues (for calibration data) and of Hotelling's T-squared distribution vs. Q-residues (for calibration and validation data). MATLAB (The Math Works, Natick, Massachusetts) and PLS Toolbox (Eigenvector Technologies, Manson, Washington) statistical packages were employed for the chemometric calculations.

3. Results and discussion

PLS models were constructed by using the data obtained from 100 samples of adulterated coffee and 45 samples of pure coffee. The calibration and validation sets consisted of a total of 97 and 48 samples, respectively. The results obtained for the PLS models based on the full-spectrum approach and employing the different pre-processing techniques cited in Section 2 are presented in Table 2. The model that was based on raw data was not considered to be satisfactory because it presented a larger calibration error and a smaller correlation factor than the models that were based on treated data. The pretreatments that provided a significant improvement in model performance, i.e., a decrease in RMSEC and RMSEP values coupled with an increase in correlation values, were the following: normalization, first derivative, MSC and SNV. Mean centering alone was not effective, although it improved model performance when coupled with absorbance normalization. The model obtained with the data submitted to first derivative followed by smoothing and mean centering presented the best performance with respect to the calibration set. Therefore, it was chosen as the most appropriate of the pre-treatment schemes tested. Although other techniques such as MSC and SNV presented slightly better

Table 2
Performance results of full spectrum PLS models based on different data pre-processing techniques.

Data pre-treatment	LV	RMSEC (%)	Rc	RMSEP (%)	Rv
None (raw data)	7	5.84	0.90	5.72	0.91
Mean centering (MC)	7	5.71	0.90	5.83	0.90
First derivatives+smoothing+MC	**10**	**2.01**	**0.99**	**3.70**	**0.96**
Absorbance normalization	12	2.68	0.98	3.55	0.96
Absorbance normalization+MC	12	2.28	0.98	3.18	0.97
Area normalization	12	2.71	0.98	3.59	0.96
Area normalization+MC	11	2.71	0.98	3.59	0.96
Multiple scatter correction (MSC)	12	2.40	0.98	3.10	0.97
Standard normal variates (SNV)	12	2.36	0.98	3.03	0.97

LV: latent variables; Rc: calibration correlation coefficient; Rv: validation correlation coefficient.

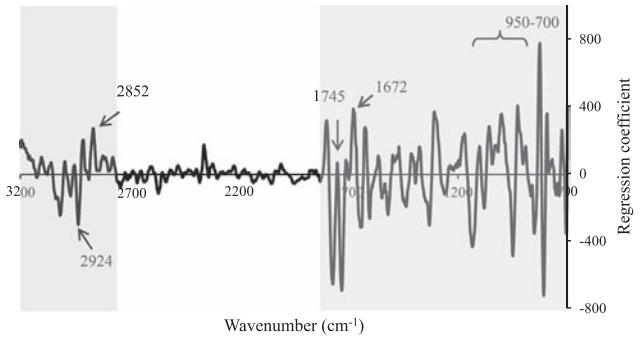

Fig. 1. Regresssion coefficients of the full-spectrum (3200–700 cm^{-1}) PLS model based on the first derivative of spectra, followed by smoothing and mean centering.

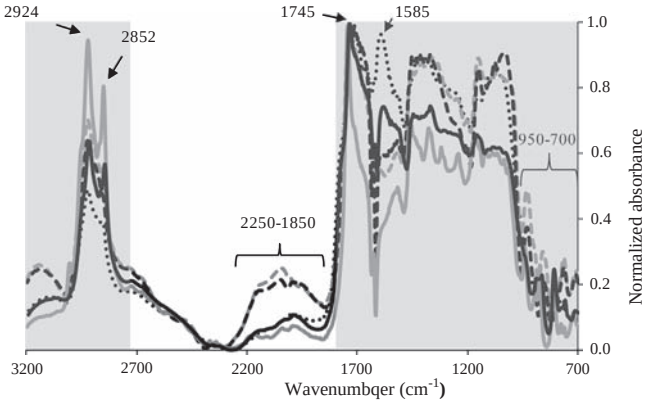

Fig. 2. Average normalized diffuse reflectance spectra obtained for (—) roasted coffee, (· · · · · ·) roasted coffee husks, (— — —) roasted corn, (- - -) roasted barley and (————) spent coffee grounds.

performance in terms of validation, the number of latent variables (LV) was higher. LV selection for each model was based on the smallest value of RMSECV.

Table 3
Performance results of PLS models based on different data ranges.

Wave number range (cm^{-1})	LV	RMSEC (%)	Rc	RMSEP (%)	Rv
3200–700 (full spectrum)	10	2.01	0.99	3.70	0.96
3200–2730 and 1800–700	**10**	**1.96**	**0.99**	**3.74**	**0.96**
1800–700	10	2.15	0.99	3.86	0.96
1800–1200	6	5.16	0.92	5.60	0.91
1200–700	10	2.43	0.98	6.18	0.90

LV: latent variables; Rc: calibration correlation coefficient; Rv: validation correlation coefficient.

Table 4
Results for the optimization of the PLS models by detection of outliers.

Model	1st	2nd	**3rd**
Number of calibration samples	97	90	**88**
Number of validation samples	48	45	**44**
LV	10	8	**10**
RMSEC (%)	1.96	1.76	**1.61**
RMSEP (%)	3.74	2.77	**2.34**
Rc	0.99	0.99	**0.99**
Rv	0.96	0.98	**0.98**

LV: latent variables; Rc: calibration correlation coefficient; Rv: validation correlation coefficient.

We evaluated the plot of model regression coefficients (see Fig. 1) to verify the feasibility of improving the predictive ability of the model. The spectral regions that affect model performance are those that present higher absolute values for regression coefficients, e.g., 3200–2730 (moderate intensity) and 1800–700 cm^{-1} (high intensity). Such regions are also highlighted in Fig. 2, where average normalized spectra of pure coffee and adulterants are shown.

The two sharp absorption bands seen at 2924–2925 and 2852 cm^{-1} are indicative of compounds containing long linear aliphatic chains and, given the presence of absorption bands above 3000 cm^{-1}, some of these might be unsaturated. Thus, these bands can be partly assigned to unsaturated and saturated lipids [5] present in coffee, corn and barley oils. Similar bands have also been identified in spectra of roasted [4,5,22] and raw coffee samples [15,16], as well as in spectra of caffeinated beverages such as coffee, tea and soft drinks [23], with the second band (~2852 cm^{-1}) attributed to stretching of C–H bonds of the methyl ($-CH_3$) groups in the caffeine molecule. It is interesting to notice that the second band is less evident in the spectra of barley and corn than in the other spectra because these grains do not contain any caffeine. The amounts of lipids reported to be present in barley (1.9–2.87%) and coffee husks (1.5–3%) are lower than those in coffee beans (12–16%) and corn kernels (3–5%) [5,24]. Thus, such bands are probably affected by both caffeine and lipid levels in the case of coffee, but only by caffeine in the case of coffee husks. These bands are attributed to lipids in roasted corn, roasted barley and spent coffee grounds. Most of the caffeine present in roasted coffee will be extracted during the production of soluble coffee, whereas the lipid fraction is only partially extracted. Thus, the

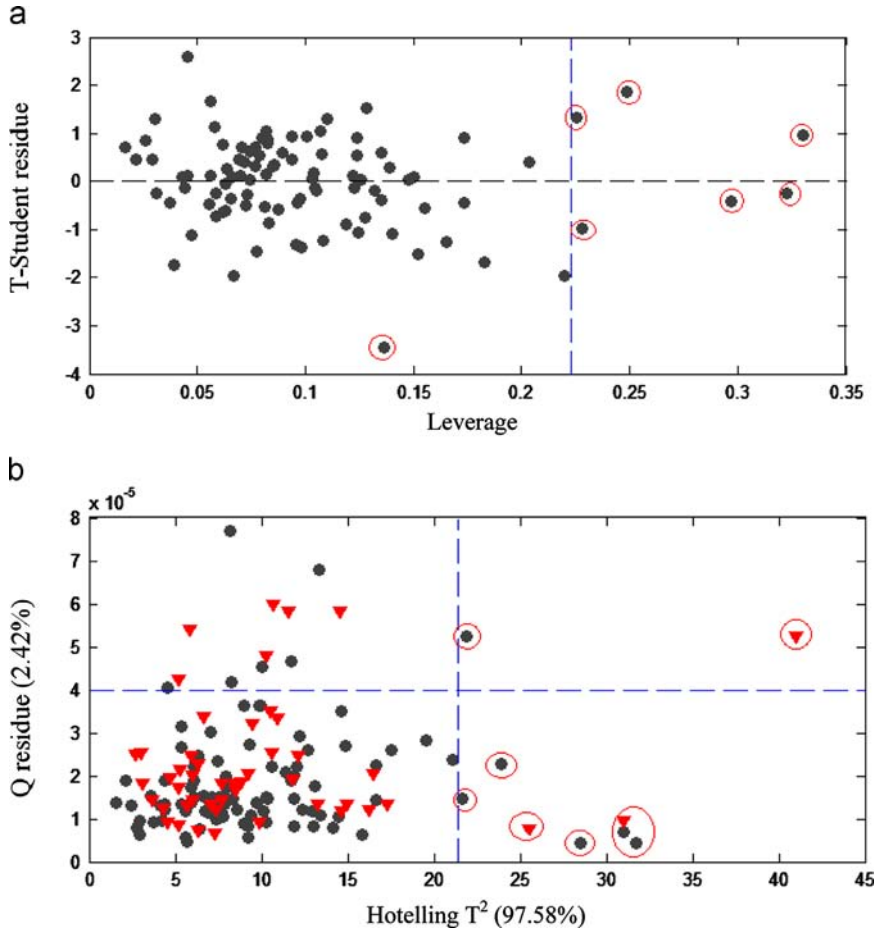

Fig. 3. Outlier removal strategy based on (a) leverage vs. Student-*t*-student residues plot and (b) of Hotelling's *T*-squared distribution vs. *Q*-residues plot (● calibration samples; ▼ validation samples).

resulting spent coffee grounds will be essentially devoid of caffeine but still contain lipids.

Absorbance peaks in the 2730–1800 cm^{-1} range did not influence PLS results. Nonetheless, bands in the 2250–1850 cm^{-1} range are expected to appear in all the matrices studied because they all present complex chemical compositions and these bands are usually associated with harmonic bands and combination deformation bands of aromatic compounds. The fact that these bands are significantly more intense in the spectra of roasted corn and barley than they are in the spectra of roasted coffee and husks and spent coffee is an indication that they might be associated with starch. They are attributed to the combination absorption bands of bound phenolic compounds [25,26] such as ferulic and coumaric acids and their derivatives or to absorption in the C–O stretching region resulting from the interactions of starch and the residual gluten in the presence of water.

The sharp bands at 1745 cm^{-1} have previously been identified in FTIR spectra of roasted coffee [5] in association with the carbonyl (C═O) vibration in esters (triglycerides) and aldehydes and are also related to the lipid concentration. Many bands can be viewed in all the spectra in the range from 1700 to 700 cm^{-1}. This result was expected because many substances that naturally occur in coffee are reported to present absorbance bands in this range, which is the double bond region [5]. One example is trigonelline, a pyridine compound that is present in both crude and roasted coffee and has been reported to present several bands in the range from 1650 to 1400 cm^{-1} [27,28]. Chlorogenic acids present a strong absorption in the 1450–1000 cm^{-1} range [29]. Carbohydrates also exhibit several absorption bands in the 1500–700 cm^{-1} region, so it is expected that this class of compounds will contribute to many of the observed bands. The skeletal mode vibrations of the glycoside bonds in starch are usually observed in the 950–700 cm^{-1} range [30]. The sharp bands in the 950–700 cm^{-1} region coincide in the spectra of corn and barley and are slightly shifted in relation to the bands in the spectra of coffee, spent coffee and coffee husks. The differences in this region could be associated with the differences in the types of polysaccharide found in coffee and coffee by-products from those found in the roasted grains, i.e., β-glycoside bonds in coffee and its by-products (arabinogalactans, galactomannans and cellulose) and α-glycoside bonds in corn and barley (starch).

Therefore, PLS models were rebuilt on the basis of the following regions of the spectrum: (i) 3200–2730 and 1800–700 cm^{-1}, (ii) 1800–700 cm^{-1}, (iii) 1800–1200 cm^{-1} and (iv) 1200–700 cm^{-1}, with the results displayed in Table 3 in comparison with the full spectrum model (3200–700 cm^{-1}). There was a slight improvement in model performance, e.g., reduction in RMSEC, when the model was restricted to the 3200–2730 and 1800–700 cm^{-1} regions. This model was chosen for optimization by detection and elimination of outliers [20].

The outliers were detected at the 99% confidence level, and the results are summarized in Table 4. Outliers can be removed up to a limit of 22% of the total number of samples, according to the Brazilian and international guidelines [20]. Optimization of the validation set was performed after terminating the optimization of the calibration set. The rounds of outlier detection were limited to at most three, as recommended in the literature [20,31]. The first round of outlier removal is illustrated in Fig. 3a and b. Seven outliers were removed in Fig. 3a and nine (six calibration and three validation) were removed in Fig. 3b. The calibration data targeted for removal in the leverage vs. the plot of Student-*t* residues (Fig. 3a) included those selected in the Hotelling's *T*-squared distribution vs. *Q*-residues plot (Fig. 3b), so a total of 10 outliers (seven calibration, three validation) were removed in the first round. As it can be seen in Table 4, the best model was obtained after two rounds of outlier removal, with three more samples being removed in the second round. A third round of outlier detection did not improve model performance. Therefore, nine outliers were eliminated from the calibration set (corresponding to 9.3% of the samples) and four from the validation set (8.3% of the validation samples). Thus, the optimized PLS model for quantification of adulteration was constructed with 88 calibration and 44 validation samples, with 10 LVs, and accounted for 98.1% and 99.1% of the variance in *X* (spectral data) and *Y* (adulteration concentration), respectively. The parity plot of the optimized model is shown in Fig. 4a. As it can be seen by examination of the plot, the model is capable of predicting adulteration levels with accuracy. Residuals are randomly distributed about the mean value, which is satisfactorily close to zero, as is shown in Fig. 4b.

We further evaluated the performance of the model by adding new samples at different levels of adulteration (0.5%, 2%, 6% and 8%) to confirm the applicability of DRIFTS for the detection and quantification of adulterants in roasted and ground coffee. Also, roasted coffee samples collected from supermarkets were added to the validation set. These samples contained both coffee husks and twigs (an adulterant that was not present in the calibration set) at adulteration levels ranging from 2% to 3%. Because these samples were not prepared in the laboratory and, thus, could be regarded as "unknown" to the model, adulteration levels were established by another laboratory (FUNED – Fundação Ezequiel Dias) according to the official method employed by the Brazilian food regulation agency (ANVISA), based on microscopic analysis [32]. The results are summarized in Table 5.

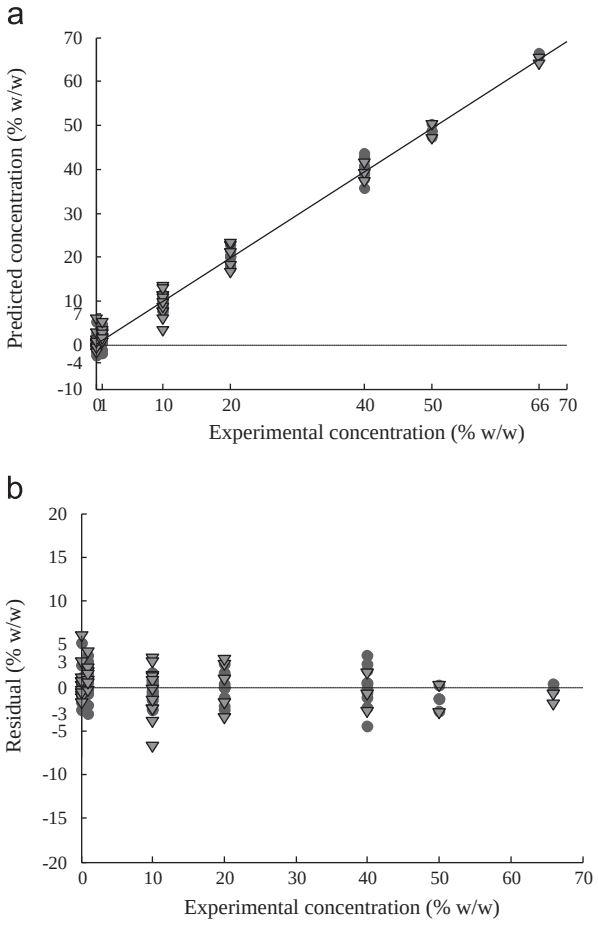

Fig. 4. (a) Experimental vs. predicted values of adulteration (% w/w) of coffee samples based on the optimized PLS model (3200-2730 and 1800-700 cm^{-1}) after outlier removal. (b) Residual vs. adulteration levels (% w/w) of coffee samples based on the optimized PLS model (3200-2730 and 1800-700 cm^{-1}) after outlier removal (● calibration samples; ▽ validation samples).

Table 5
Performance results of PLS models obtained after inclusion of new adulterated samples.

Model	LV	RMSEC (%)	Rc	RMSEP (%)	Rv
1 (reference model)	10	1.96	0.99	3.74	0.96
2 (new adulteration levels)	10	2.47	0.98	3.77	0.95
3 (samples from supermarket)	10	2.47	0.98	11.0	0.62
4 (without samples at adulteration levels over 20%)	8	2.08	0.87	3.05	0.71
5 (after removal of outliers)	**12**	**1.23**	**0.95**	**2.50**	**0.82**

LV: latent variables; Rc: calibration correlation coefficient; Rv: validation correlation coefficient. All models based on the data in the following wave number range: 3200–2730 and 1800–700 cm^{-1}.

All the models presented on Table 5 were based on data obtained in the 3200–2730 and 1800–700 cm^{-1} regions. For the sake of comparison, the original model obtained in that range, as presented in Table 3 (Model 1), was employed as a reference. Model 2 was obtained after inclusion of new samples roasted in our laboratory employing the same adulterants (barley, coffee husks, corn, and spent coffee grounds) at adulteration levels different from those presented in Table 1. There was an increase in RMSEC and RMSEP values coupled with a slight decrease in correlation coefficients. This result was expected because of the increase in the number and variability of samples. Model 3 was obtained after adding the adulterated samples collected from supermarkets to the validation set, while keeping the calibration set previously employed. The main goal was to verify whether the method would still be valid for samples that contained other adulterants. RMSEC and Rc remained unchanged, given that no new samples were added to the calibration set. There was a significant increase in RMSEP accompanied by a decrease in Rv, indicating that the proposed model was not sufficiently robust to deal with other adulterants. However, given that the range of adulteration levels considered for building the model (0.5–66%) was quite broad in comparison with expected adulteration levels in commercial samples, we decided to restrict the upper limit of adulteration to 20% in the calibration set, while still keeping the supermarket samples in the validation set (Model 4). There was a significant decrease in both RMSEC and RMSEP values, which are now similar to the values obtained for the reference model. Model performance was further improved (Model 5) by removal of outliers (two samples from calibration and one sample from validation), with RMSEC and RMSEP values now quite similar to those obtained for the optimized model presented in Table 4. Outlier removal also provided an increase in correlation factors. Such results confirm that DRIFTS is a robust technique that can be employed for detection and quantification of multiple adulterants in roasted and ground coffee.

4. Conclusions

Partial Least Squares Regression-first derivative calibration and validation models were successfully developed and applied for the detection and quantification of multiple adulterants in ground, roasted coffee, with adulteration levels ranging from 1% to 66% w/w. The model was developed using a combination of the spectral regions of 3200–2730 cm^{-1} and 1800–700 cm^{-1}. The determination coefficients were 0.99 and 0.98 for the calibration and validation sets, respectively, and the errors observed during calibration and validation were quite low, 1.23% and 2.67%, respectively. The main advantages of using DRIFTS for the detection and quantification of adulterants in ground, roasted coffee are the simplicity and the fact that hazardous solvents and reagents need not be used; highly trained personnel are not needed to perform routine analysis in quality-monitoring labs.

Acknowledgments

The authors acknowledge financial support from the following Brazilian government agencies: CAPES, CNPq and FAPEMIG. The authors thank Dr. Shawn Steiman and Prof. David Lee Nelson for proofreading the paper and FUNED (Fundação Ezequiel Dias) for providing commercially available adulterated coffee samples.

References

[1] H. Ebrahimi-Najafabadi, R. Leardi, P. Oliveri, M.C. Casolino, M. Jalali-Heravi, S. Lanteri, Talanta 99 (2012) 175–179.
[2] R.C.S. Oliveira, L.S. Oliveira, A.S. Franca, R. Augusti, J. Food Compos. Anal. 22 (2009) 257–261.
[3] R.C. Alves, S. Casal, M.R. Alves, M.B. Oliveira, Food Chem. 114 (2009) 295–299.
[4] A.P. Craig, A.S. Franca, L.S. Oliveira, Lebensm.-Wiss. Technol. 47 (2012) 505–511.
[5] N. Reis, A.S. Franca, L.S. Oliveira, Lebensm.-Wiss. Technol. 50 (2013) 715–722.
[6] G.N. Jham, J.K. Winkler, M.A. Berhow, S.F. Vaughn, J. Agric. Food Chem. 55 (2007) 5995–5999.
[7] T. Nogueira, C.L. Lago, J. Sep. Sci 32 (2009) 3507–3511.
[8] L.M.Z. Garcia, E.D. Pauli, V. Cristiano, C.A.P. Camara, I.S. Scarminio, S.L. Nixdorf, J. Chromatogr. Sci 47 (2009) 825–832.
[9] E.D. Pauli, V. Cristiano, S.L. Nixdorf, Quim. Nova 34 (2011) 689–694.
[10] L.S. Oliveira, A.S. Franca, Food Quality: Control, Analysis and Consumer Concerns, in: D.M. Medina, A.M. Laine (Eds.), Nova Publishers, New York, 2011, pp. 131–179.
[11] R. Kizil, J. Irudayaraj, Nondestructive Testing of Food Quality, in: J. Irudayaraj, C. Reh (Eds.), Blackwell Publishing Ltd, Oxford, 2008, pp. 143–163.
[12] L. Rodriguez-Saona, M.E. Allendorf, Annu. Rev. Food Sci. Technol. 2 (2011) 467–483.
[13] I. Esteban-Díez, J.M. González-Sáiz, C. Sáenz-González, C. Pizarro, Talanta 71 (2007) 221–229.
[14] R.M. El-Abassy, P. Donfack, A. Materny, Food Chem. 126 (2011) 1443–1448.
[15] A.P. Craig, A.S. Franca, L.S. Oliveira, J. Food Sci. 76 (2011) C1162–C1168.
[16] A.P. Craig, A.S. Franca, L.S. Oliveira, Food Chem. 132 (2012) 1368–1374.
[17] J.R. Santos, M.C. Sarraguça, A.O.S.S. Rangel, J.A. Lopes, Food Chem. 135 (2012) 1828–1835.
[18] A.S. Franca, L.S. Oliveira, R.C.S. Oliveira, P.C. Mancha Agresti, R. Augusti, J. Food Eng. 92 (2009) 345–352.
[19] K.E. Kramer, R.E. Morris, S.L. Rose-Pehrsson, Chemometr. Intell. Lab. 92 (2008) 33–43.
[20] B.G. Botelho, B.A.T. Mendes, M. Sena, Food Anal. Method 6 (2013) 881–891.
[21] M.C. Ferreira, A.M. Antunes, M.S. Melgo, P.L. Volpe, Quím. Nova 22 (1999) 724–731.
[22] N. Wang, L.T. Lim, J. Agric. Food Chem. 60 (2012) 5446–5453.
[23] M.M. Paradkar, J. Irudayaraj, Food Chem. 78 (2002) 261–266.
[24] R.O. Osman, F.M. Abd El Gelil, H.M. El-Noamany, M.G. Dawood, Grasas Aceites 51 (2000) 157–162.
[25] L.X. Lopez-Martinez, R.M. Oliart-Ros, G. Valerio-Alfaro, C.H. Lee, K.L. Parkin, H.S. Garcia, Lebensm.-Wiss. Technol. 42 (2009) 1187–1192.
[26] M. Omwamba, Q. Hu, J. Food Sci. 75 (2010) C66–C73.
[27] M. Szafran, J. Koput, Z. Dega-Szafran, M. Pankowski, J. Mol. Struct. 614 (2000) 97–108.
[28] M.R. Silverstein, F.X. Webster, D. Kiemle, Spectrometric Identification of Organic Compounds, Wiley, Hoboken, 2005.
[29] R. Briandet, E.K. Kemsley, R.H. Wilson, J. Sci. Food Agric. 71 (1996) 359–366.
[30] R. Kizil, J. Irudayaraj, K. Seetharaman, J. Agric. Food Chem. 50 (2002) 3912–3918.
[31] P. Valderrama, J.W.B. Braga, R.J. Poppi, J. Agric. Food Chem. 55 (2007) 8331–8338.
[32] F.C. Lopez, Rev. Inst. Adolfo Lutz 43 (1983) 3–8.