CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Elsevier - Publisher Connector

LWT - Food Science and Technology 47 (2012) 505–511

Contents lists available at SciVerse ScienceDirect

LWT - Food Science and Technology

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

Discrimination between defective and non-defective roasted coffees by diffuse reflectance infrared Fourier transform spectroscopy

Ana Paula Craig [a], Adriana S. Franca [a,b,*], Leandro S. Oliveira [a,b]

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

ARTICLE INFO

Article history:
Received 24 November 2011
Received in revised form
10 February 2012
Accepted 17 February 2012

Keywords:
Coffee
Diffuse reflectance
Fourier transform infrared spectroscopy
DRIFTS
FTIR

ABSTRACT

The objective of this work was to evaluate the feasibility of employing Diffuse Reflectance Infrared Fourier Transform Spectroscopy (DRIFTS) for discrimination between defective and non-defective coffees after roasting and grinding. Defective (black, immature and sour) and non-defective Arabica coffee beans were submitted to light, medium and dark roasts at 220, 235 and 250 °C. Principal Components Analysis of the DRIFTS spectra (normalized or not) and of the first derivatives of the spectra provided separation of the samples into four groups: non-defective, black, dark sour and light sour, with immature beans scattered among the sour samples. Classification models were developed based on Linear Discriminant Analysis and recognition and prediction abilities of these models ranged from 95 to 100%. Such results indicate that DRIFTS presents potential for the development of a fast and reliable analytical methodology for discrimination between defective and non-defective coffee after roasting and grinding.

© 2012 Elsevier Ltd. Open access under the Elsevier OA license.

1. Introduction

The presence of defective coffee beans depreciates the quality of the coffee beverage consumed worldwide. These beans represent about 20% of the total coffee produced in Brazil and similar amounts can be expected in other producing areas around the world (Mendonça, Franca, Oliveira, & Nunes, 2008; Ramalakshmi, Kubra, & Rao, 2007). Although separated from the non-defective beans prior to commercialization in external markets, the majority of the defective beans are dumped in the Brazilian internal market and, overall, a low-grade roasted coffee is consumed in the country (Craig, Franca, & Oliveira, 2011). The negative effect that such beans have on coffee quality can be associated to specific problems that occur during harvesting and post-harvest processing operations. Black beans result from dead beans within the coffee cherries or from beans that fall naturally on the ground by action of rain or over-ripening (Mazzafera, 1999). The presence of sour beans can be associated with 'overfermentation' during wet processing and with improper drying or picking of overripe cherries, whereas immature beans come from immature fruits (Clarke & Macrae, 1987; Mendonça et al., 2008). The chemical changes due to the extraneous factors acting upon the beans (e.g., microbial fermentation) and due to the maturity stage of the beans (e.g., immature vs. mature) exert a perceptive effect in the sensory quality of the coffee beverage when determined by a trained sensory panel, but can be subtle enough not to be detected by analytical instruments depending on the technique being employed for that purpose. Considering that the defective coffee is separated from the non-defective prior to commercialization, and is also cheaper than non-defective coffee, the amount of defective beans to be used for roasting is dependent exclusively on the types of blends defined by the roasters themselves. Thus, the ultimate quality of a brand of coffee will be dictated by the amount of defective beans used for roasting, with higher qualities being expected for blends with small amounts of these beans and lower qualities for blends with greater amounts. The presence of black beans in a roasted batch usually imparts a heavy flavor to the beverage; sour beans contribute to sour and oniony tastes, while immature beans impart astringency (Clarke & Macrae, 1987).

Research interest on defective and low quality coffee beans has intensified over the past years, given the increasing awareness regarding the negative aspects they impart to the quality of the roasted and ground coffee used for beverage preparation and consumption (Craig et al., 2011; Craig, Franca, & Oliveira, 2012;

Abbreviations: ATR-FTIR, Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy; DR, diffuse reflectance; DRIFTS, Diffuse Reflectance Infrared Fourier Transform Spectroscopy; DLATGS, Deuterated Triglycine Sulfate Doped with L-Alanine; ESI-MS, electrospray-ionization mass spectrometry; LDA, linear discriminant analysis; FTIR, Fourier Transform Infrared Spectroscopy; PCA, principal components analysis; PR, pattern recognition.

* Corresponding author. DEMEC/Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, 31270-901 Belo Horizonte, MG, Brazil. Tel.: +55 31 34093512; fax: +55 31 34433783.
E-mail addresses: adriana@demec.ufmg.br, drisfranca@gmail.com (A.S. Franca).

0023-6438 © 2012 Elsevier Ltd. Open access under the Elsevier OA license.
doi:10.1016/j.lwt.2012.02.016

Farah, Monteiro, Calado, Franca, & Trugo, 2006; Franca, Mendonça, & Oliveira, 2005; Franca, Oliveira, Mendonça, & Silva, 2005; Mancha Agresti, Franca, Oliveira, & Augusti, 2008; Mendonça et al., 2008; Mendonça, Franca, & Oliveira, 2009; Mendonça, Franca, Oliveira, & Afonso, 2009; Oliveira, Franca, Mendonça, & Barros-Junior, 2006; Ramalakshmi et al., 2007; Vasconcelos, Franca, Glória, & Mendonça, 2007). Such studies have shown that there are physical and chemical differences between defective and non-defective coffee beans prior to roasting, but only a few have attained some success regarding discrimination of defective and non-defective coffees after roasting. Mancha Agresti et al. (2008) showed that roasted defective and non-defective coffees could be separated into two distinct groups based on their volatile profiles: immature/black beans and non-defective/sour coffees. Mendonça, Franca, and Oliveira (2009) showed that, for Arabica coffees, defective and non-defective roasted coffees could be separated by sieving. However, the majority of the commercially available roasted coffee is ground. Mendonça et al. (2008) and Mendonça, Franca, Oliveira et al. (2009) attempted to employ electrospray-ionization mass spectrometry (ESI-MS) for discrimination of defective and non-defective coffees before and after roasting. ESI-MS profiles in the positive mode (ESI(+)-MS) provided separation between defective and non-defective green coffees prior to roasting, but could not provide separation of roasted coffees.

Recent studies have shown that methods based on Fourier Transform Infrared spectroscopy (FTIR) in combination with chemometric techniques have been successfully applied for food quality evaluation (Rodriguez-Saona & Allendorf, 2011). FTIR-based methods are fast, reliable and simple to perform. They can be based on transmittance or reflectance readings, and although both techniques are appropriate for analyzing either solid or liquid samples, reflectance-based methods require none or very little sample pretreatment, being thus more commonly employed as routine methodologies for food analysis (Bauer et al., 2008; Rodriguez-Saona & Allendorf, 2011). Reflectance methods that are appropriate for non specular solid samples are divided into Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy (ATR-FTIR) and Diffuse Reflectance Fourier Transform Infrared Spectroscopy (DRIFTS). While ATR collects information mainly from the solid surface, DRIFTS provides information from the entire solid matrix, given that it is a combination of internal and external reflection. Both techniques have been employed for coffee quality analysis, with most of the ATR-based studies focusing on analysis of liquid samples, i.e., the coffee beverage (Briandet, Kemsley, & Wilson, 1996; Lyman, Benck, Dell, Merle, & Murray-Wijelath, 2003; Wang, Jun, Bittenbender, Gautz, & Li, 2009). DRIFTS has been also successfully applied for analysis of coffee, specifically targeting discrimination between Arabica and Robusta varieties (Kemsley, Ruault, & Wilson, 1995), detection of glucose, starch or chicory as adulterants of freeze-dried instant coffees (Briandet et al., 1996) and separation between decaffeinated and regular roasted coffees (Ribeiro, Salva, & Ferreira, 2010). We have shown, in recent studies, that DRIFTS provides satisfactory discrimination of non-defective/defective and immature/mature coffees prior to roasting (Craig et al., 2011, 2012). In view of the aforementioned, the objective of this work was to evaluate the potential of this technique in the discrimination of defective and non-defective coffee beans after roasting and grinding.

2. Materials and methods

Arabica green coffee samples were acquired from a Coffee Roasting Company located in Minas Gerais (MG) State, Brazil (Café Fino Grão, Contagem, MG). The samples consisted of three 60 kg bags of coffee beans (harvested by the strip-picking method) that were rejected by color sorting machines. Four samples of 2 kg of whole beans were randomly taken from each bag, mixed and their beans were manually sorted (by a professional trained and certified for green coffee classification) into five lots: non-defective, immature, black and sour (separated into light and dark colored). Coffee samples (25 g) were taken from each lot and submitted to roasting in a convection oven (Model 4201D Nova Ética, São Paulo, Brazil), at 220, 235 and 250 °C. After roasting, the samples were ground ($D < 0.5$ mm) and submitted to color evaluation. Color measurements were performed using a tristimulus colorimeter (HunterLab Colorflex 45/0 Spectrophotometer, Hunter Laboratories, VA, USA) with standard illumination D_{65} and colorimetric normal 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. Roasting conditions were established for each specific lot, given that defective coffee beans have been reported to roast to a lesser degree than non-defective coffee beans when submitted to the same processing conditions (Mancha Agresti et al., 2008). Roasting degrees were then defined according to luminosity (L^*) measurements similar to commercially available coffee samples ($19.0 < L^* < 25.0$), corresponding to light ($23.5 < L^* < 25.0$), medium ($21.0 < L^* < 23.5$) and dark ($19.0 < L^* < 21.0$) roasts. The corresponding roasting times ranged from 7 to 10 min (250 °C), 9–16 min (235 °C) and 12–33 min (220 °C), with the smaller and larger times for a given temperature corresponding to the light and dark roasts, respectively.

A Shimadzu IRAffinity-1 FTIR Spectrophotometer (Shimadzu, Japan) with a DLATGS (Deuterated Triglycine Sulfate Doped with L-Alanine) detector was used in 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). The ground coffee sample was mixed with KBr (100 mg) and then 23 mg of this mixture was placed inside the sample port. Pure KBr was employed as reference material (background spectrum). All spectra were recorded within a range of 4000–400 cm^{-1} with a 4 cm^{-1} resolution and 20 scans, and submitted to background subtraction. The spectra were also truncated to 2500 data points in the range of 3100–600 cm^{-1}, in order to eliminate noise readings present in the upper and lower ends of the spectra. Preliminary tests were performed in order 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 coffee/KBr mass ratio (2, 5, 10, 20, 30, 40 and 50%) on the quality of the obtained spectra. The conditions that provided the best quality spectra (higher intensity and lower noise interference) were $D <$ 0.15 mm and 10% coffee/KBr mass ratio. In order to improve performance of prediction models, the following data pretreatment techniques were evaluated: (0) no additional processing (raw data), (1) mean centering, (2) normalization, (3) baseline correction, (4) first derivatives and (5) second derivatives. Mathematical treatments such as mean centering and normalization are commonly applied to data in order to remove redundant information and enhance sample-to-sample differences (Wang et al., 2009). Mean centering corresponds to subtraction of the average absorbance value of a given spectrum from each data point. Normalization is calculated by dividing the difference between the response at each data point and the minimum absorbance value by the difference between the maximum and minimum absorbance values. Baseline correction and derivative transformations are usually performed in order to compensate for baseline offset between samples and also to reduce instrument variations (Esteban-Díez, González-Sáiz, Sáenz-González, & Pizarro, 2007).

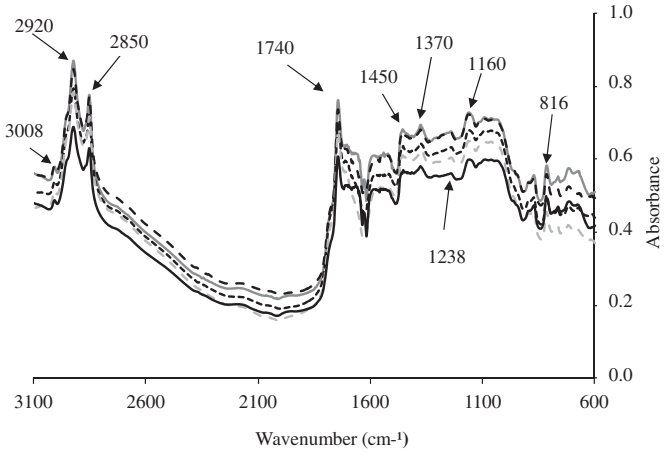

Fig. 1. Average DR spectra obtained for defective and non-defective roasted coffee beans (non-defective; immature; sour (light); sour (dark); black; each spectrum represents an average of 25 samples).

The statistical software XLSTAT Sensory 2010 (Addinsoft, New York) was employed for all the chemometric calculations.

3. Results and discussion

Average spectra obtained for defective and non-defective roasted coffee samples are shown in Fig. 1. A comparative evaluation of these spectra indicates that they are quite similar, although variations in band intensity are perceived, with absorbance values being higher for non-defective and light sour beans and lower for black beans. The two sharp bands at 2920 and 2850 cm^{-1} have been previously identified in Arabica and Robusta roasted coffee samples (Kemsley et al., 1995) and also on Arabica green coffee samples (Craig et al., 2011, 2012), in association to asymmetric and symmetric stretching of C–H bonds. Studies of FTIR analysis of caffeine on soft drinks have reported two sharp bands at 2882 and 2829 cm^{-1}, with the latter being due to the asymmetric stretching of C–H bonds of methyl ($-CH_3$) group in the caffeine molecule (Paradkar & Irudayaraj, 2002). Other FTIR studies on corn and corn flour have also reported two bands at 2927–2925 and 2855 cm^{-1}, being respectively attributed to asymmetric and symmetric C–H stretching in lipids (Cremer & Kaletunç, 2003; Gordon, Schudy, Wheeler, Wicklow, & Greene, 1997). Thus, the sharp bands at 2920 and 2850 cm^{-1} observed in the spectra presented for coffee in Fig. 1 can be attributed to combination bands to which both caffeine and lipids contribute. The sharp band at 1740 cm^{-1} was also reported on previous FTIR studies on roasted coffee, in association to

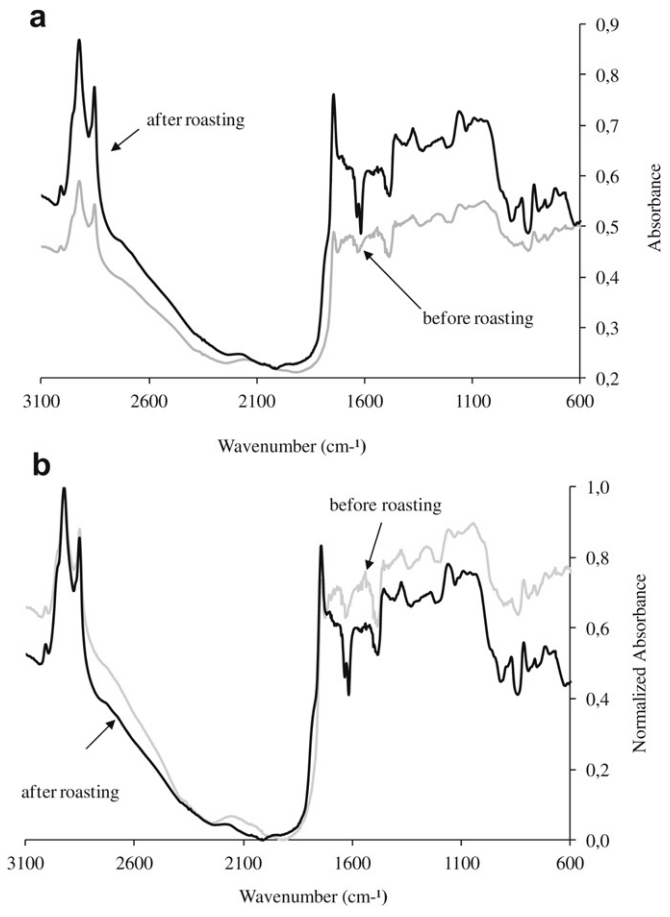

Fig. 2. Comparison of (a) raw and (b) normalized spectra obtained for coffee beans before (gray curves) and after roasting (black curves).

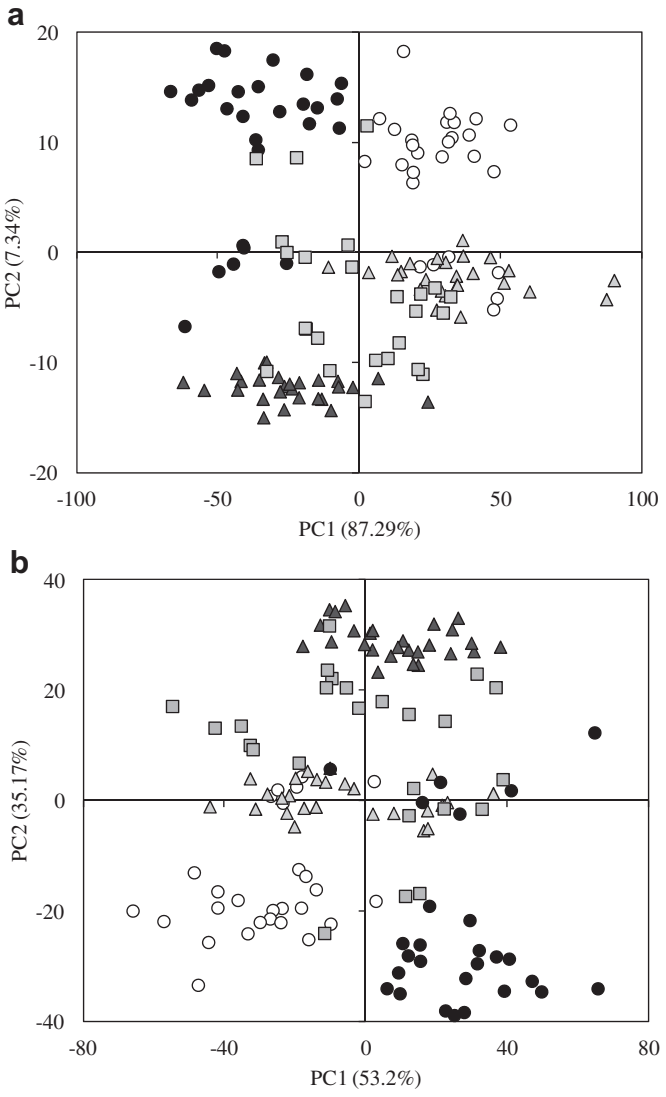

Fig. 3. PCA scores scatter plot (PC1 vs. PC2) of diffuse reflectance spectra (3100–600 cm^{-1}) (a) raw spectra (b) after mean centering; (c) after normalization; and (d) after first derivatives. ○ non-defective; □ immature; △ sour (light); ▲ sour (dark); ● black.

carbonyl (C═O) vibration of the ester group in triglycerides (Kemsley et al., 1995) or to aliphatic esters (Lyman et al., 2003), indicating that this band could be associated to lipids. The combination of absorptions at 1740 cm^{-1} (C═O stretch) and at 2830-2695 cm^{-1} (H–C═O stretch) with a weak shoulder-type peak at 2725–2740 cm^{-1} could be interpreted as a presence of aldehydes (Miller, Mayo, & Hannah, 2003), which are volatile compounds found aplenty in roasted coffee, as a result of the thermal degradation of unsaturated fatty acids, such as linoleic acid, which is quite abundant in the coffee lipid fraction (Oliveira et al., 2006). The wavenumber 1659 cm^{-1} has been identified by Garrigues, Bouhsain, Garrigues, and De La Guardia (2000) as due to the presence of carbonyl groups in caffeine in their FTIR analysis of trichloromethane extracts of roasted coffee, and was further used as the determinant band in their quantitative analytical procedure for caffeine in roasted coffee samples. However, in our study, this band appears rather modestly in the spectra for roasted and ground coffee. Thus, it can be assumed that several other compounds in roasted coffee also absorb in that range of wavenumbers and that, apparently, trichloromethane does not extract them, since in the work by Garrigues et al. (2000) the 1659 cm^{-1} was quite sharp in the trichloromethane extract.

A comparison of average DR spectra obtained for green and roasted coffees is shown in Fig. 2a. The spectra are qualitatively similar, even though roasted coffees presented higher absorbance in comparison to green coffees. It is interesting to observe that, once the spectra were normalized (see Fig. 2b), all the previously cited bands (2920, 2850 and 1740 cm^{-1}) presented similar levels of absorbance in green and roasted coffees. This could be associated to the fact that both caffeine and lipids levels are not expected to vary significantly during roasting (Franca, Mendonça et al., 2005; Franca, Oliveira et al., 2005; Vasconcelos et al., 2007). Evaluation of Fig. 2b

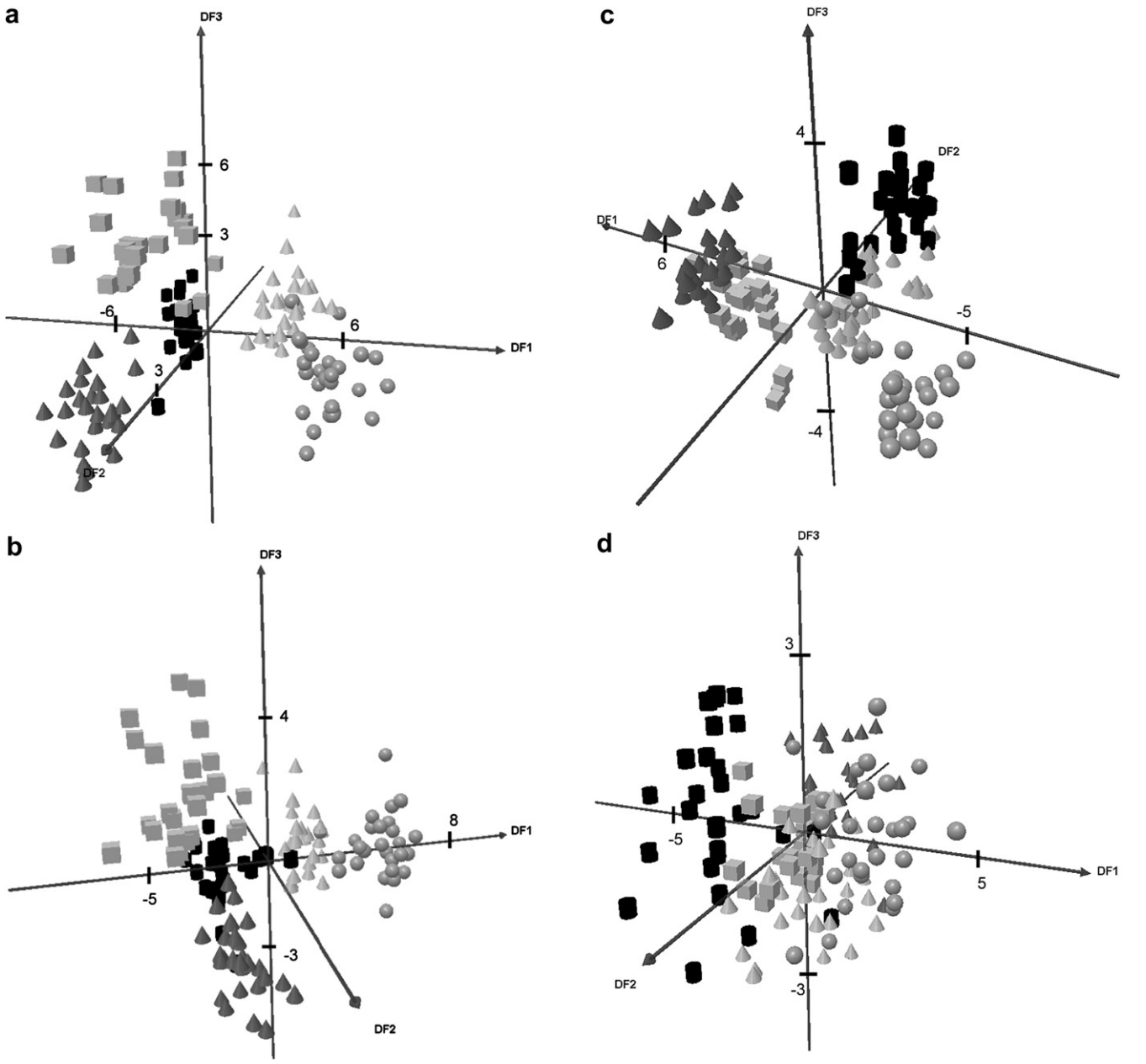

Fig. 4. Scores on the discriminant functions provided by the 8 variables LDA models of diffuse reflectance spectra (3100–600 cm^{-1}) (a) raw spectra (b) after mean centering; (c) after normalization; and (d) after first derivatives. ● non-defective; ■ immature; ▲ sour (light); ▲ sour (dark); ■ black.

also shows no significant differences between green and roasted coffees regarding absorbance values of the small band at 3008 cm^{-1}. This band can be attributed to the symmetric stretching vibration of C–H cis-olefinic groups (=C–H in *cis* RHC = CHR) and can be also associated to the presence of lipids (Yang & Irudayaraj, 2001). The fact that it was not significantly changed from the spectrum for the green beans to that of roasted ones indicates that the double bonds of unsaturated fatty acids did not undergo isomerization from *cis* to *trans* during roasting.

Several bands can be viewed in the range of 1700–600 cm^{-1}. The wavenumber range of 1400–900 cm^{-1} is characterized by vibrations of several types of bonds, including C–H, C–O, C–N and P–O (Sablinskas, Steiner, & Hof, 2003; Wang et al., 2009). Other studies on FTIR analysis of roasted coffees (Briandet et al., 1996; Kemsley et al., 1995) have reported that carbohydrates exhibit several absorption bands in this region, so it is expected that this class of compounds will contribute to several of the observed bands. According to Kemsley et al. (1995), Briandet et al. (1996), and Lyman et al. (2003), chlorogenic acids also present absorption in the region of 1450–1000 cm^{-1}. Chlorogenic acids represent a family of esters formed between quinic acid and one to four residues of certain *trans*-cinnamic acids, most commonly caffeic, *p*-coumaric and ferulic (Clifford, Kirkpatrick, Kuhnert, Roozendaal, & Salgado, 2008). Axial C–O deformation of the quinic acid occurs in the range of 1085–1050 cm^{-1}, and O–H angular deformation occurs between 1420 and 1330 cm^{-1}. The C–O–C ester bond also absorbs in the 1300–1000 cm^{-1} range (Silverstein, Webster, & Kiemle, 2005) and therefore the bands located in the range of 1450–1050 cm^{-1} could be partially due to chlorogenic acids. Hashimoto et al. (2009) studied the influences of coffee varieties, geographical origin and of roasting degree on the mid-infrared spectral characteristics of brewed coffee, and also developed a fast and reliable procedure to determine the caffeine and chlorogenic acid contents in brewed coffee using the ATR-FTIR method. In their method, developed based on the spiking of the coffee brew with different amounts of caffeine, they identified the band at 1242 cm^{-1} as the most relevant absorption band for characterization of the caffeine content in the brew. In the roasted and ground coffee IR spectra herein obtained for defective and non-defective coffee beans this peak appears shifted to a slightly lower band (1238 cm^{-1}), but it is present in all spectra. Another substance that can be associated to peaks in the 1600–1300 cm^{-1} range is trigonelline, a pyridine derivative that has been reported to present four bands in this range, due to axial deformation of C=C and C=N bonds (Silverstein et al., 2005). A comparison of the average spectra of green and roasted coffees presented in Fig. 2b shows a decrease in the relative absorbance of several bands in the 1700–600 cm^{-1} region after roasting. Several literature reports confirm that the levels of carbohydrates, trigonelline and chlorogenic acids diminish upon roasting (Farah et al., 2006; Franca, Oliveira et al., 2005), so such variations in chemical composition are expected to affect the spectra in the 1700–600 cm^{-1} range.

Using the DR spectra as chemical descriptors, pattern recognition (PR) methods (principal components analysis – PCA and linear discriminant analysis – LDA) were applied in order to establish whether or not specific types of beans could be discriminated within roasted coffee samples. Data matrices were constructed so that each row corresponded to a sample and each column represented the spectra datum at a given wavenumber, after processing as described in the previous section. The spectra pretreatment steps that provided a satisfactory level of discrimination between defective and non-defective coffees were the following: (0) no additional treatment of raw data, (1) mean centering, (2) normalization and (4) first derivatives. Pretreatments (3) and (5), baseline correction and second derivatives, did not provide satisfactory separation between defective and non-defective coffees. Furthermore, baseline correction (3) provided undesirable separation by roasting temperature.

Table 1
Calculated values of the first three discriminant functions at each sample group centroid.

Model	Non-defective	Immature	Dark sour	Light sour	Black
Raw spectra					
DF1	5.683	−2.422	−3.816	3.483	−3.013
DF2	0.432	0.314	2.544	−0.304	−3.122
DF3	−1.034	2.851	−1.479	0.996	−1.220
Mean-centered spectra					
DF1	4.695	−3.409	−1.918	1.614	−1.115
DF2	0.577	−0.040	2.880	−0.489	−2.975
DF2	0.454	2.121	−1.493	0.350	−1.437
Normalized spectra					
DF1	−3.621	3.274	2.847	−1.266	−0.945
DF2	−1.691	0.506	−1.588	0.828	2.513
DF2	−0.507	3.274	1.549	−0.531	1.283
First derivatives					
DF1	2.402	−0.711	0.094	0.376	−2.078
DF2	0.885	−0.696	−2.073	0.625	1.280
DF2	0.423	−0.410	0.388	−0.992	0.496

DF1, DF2 and DF3 represent the first, second and third discriminant functions, respectively.

The scatter plots obtained by PCA analysis are displayed in Fig. 3. A clear separation between categories can be observed, with four distinct major groups: non-defective (○), black (●), dark (▲) and light sour (△), with some outlier points. The few outlier samples from each group that were present in other classes (for example, a few non-defective and black beans in the light sour group) correspond to samples subjected to extreme roasting conditions (light roast/lower temperature and dark roast/higher temperature). Regardless of the employed spectra processing technique, immature beans (□) are somewhat scattered between light and dark sour defects. Clustering of immature and sour defects was also observed in the analysis of green coffees by ESI (+)-MS profiles (Mendonça et al., 2008) or DRIFTS (Craig et al., 2011), whereas Mancha Agresti et al. (2008) reported grouping of immature and black roasted coffee beans according to their volatile profiles.

A clear separation between non-defective and defective coffee beans can be observed in all the plots displayed in Fig. 3. Evaluation of the loadings plots obtained after PCA analysis of raw and processed spectra (not shown) indicated that the spectral ranges that presented the highest influence on PC1 and PC2 values in association with the non-defective coffees (PC1 and PC2 positive for spectra without further treatment, PC1 and PC2 negative for spectra submitted to mean centering, and PC1 negative and PC2 positive for normalized

Table 2
Correct classification rates (%) for the LDA models.

Model	Non-defective	Immature	Dark sour	Light sour	Black	Total
Raw spectra						
Recognition	87.0	90.9	100.0	100.0	100.0	95.5
Prediction	100.0	100.0	100.0	100.0	100.0	100.0
Mean-centered spectra						
Recognition	83.3	87.0	100.0	100.0	78.3	89.6
Prediction	100.0	100.0	100.0	100.0	100.0	100.0
Normalized spectra						
Recognition	84.0	90.0	100.0	100.0	72.7	89.3
Prediction	100.0	80.0	100.0	100.0	100.0	95.0
First derivatives						
Recognition	82.6	75.0	77.3	70.0	87.0	78.6
Prediction	75.0	100.0	100.0	66.7	75.0	80.0

Classification rates were evaluated as the percent ratio between the number of samples correctly classified in a specific group and the total number of samples of that group.

Table 3
Model equations and correct classification rates (%) based on generic discrimination between defective and non-defective coffees.

Model			
Raw spectra: $DF = -5.0 + 93.7A_{2924} - 110.8A_{2852} + 53A_{1743} - 23.9A_{1541} - 86.4A_{1377} - 128.6A_{1076} + 25.4A_{910} + 9.4A_{816}$			
	Non-defective	Defective	Total
Recognition	84.0	100.0	96.4
Prediction	100.0	100.0	100.0
Mean-centered spectra: $DF = -6.7 + 63.3B_{2924} - 69.9B_{2852} + 66.9B_{1743} - 21.9B_{1541} - 60.7B_{1377} - 111.0B_{1076} + 49.3B_{910} + 19.7B_{816}$			
	Non-defective	Defective	Total
Recognition	87.5	100.0	97.3
Prediction	100.0	100.0	100.0
Normalized spectra: $DF = -251.2 + 175.4C_{2924} + 93.6C_{2852} - 36.0C_{1743} + 18.9C_{1541} - 58.8C_{1377} + 86.6C_{1076} - 29.4C_{910} + 3.0C_{816}$			
	Non-defective	Defective	Total
Recognition	84.0	100.0	96.4
Prediction	100.0	100.0	100.0
First derivatives: $DF = -6.2 - 109.0D_{2924} - 815.8D_{2852} - 433.5D_{1743} - 615.2D_{1541} - 715.4D_{1377} + 2560.3D_{1076} + 859.2D_{910} - 486.3D_{816}$			
	Non-defective	Defective	Total
Recognition	88.0	96.6	94.6
Prediction	94.0	100.0	95.0

DF represents the discriminant function. A_n corresponds to the absorbance value at wavenumber n; B_n corresponds to the absorbance value at wavenumber n, after mean centering; C_n corresponds to the absorbance value at wavenumber n, after normalization; and D_n corresponds to the absorbance first derivative at wavenumber n. Classification rates were evaluated as the percent ratio between the number of samples correctly classified in a specific group and the total number of samples of that group.

spectra) were the following: 1700–1500 and 970–600 cm^{-1}, in general representing the regions in which non-defective coffees presented higher absorbance intensity in comparison to all defective categories (see Fig. 1). Loadings obtained for first derivatives could not be associated to specific regions in the spectra.

Results from the principal components analysis indicate that the obtained spectra could provide enough information to develop classification models for non-defective and each specific class of defective roasted coffees. Thus, linear discriminant analysis (LDA) was performed with the purpose of obtaining classification models for assigning categories to samples. Model validation was performed using ~25% of the samples as the evaluation set. Recognition ability was calculated as the percentage of members of the calibration set that were correctly classified, and prediction ability was calculated as the percentage of members of the validation set that were correctly classified. LDA models were constructed employing different numbers of variables (wavenumbers), starting with the entire spectrum and decreasing the number of variables. It was observed that model recognition ability varied significantly with the number of variables, with the best correlations being provided by eight-variable models. In general the models were satisfactory (average recognition and prediction abilities above 75%) as long as the selected wavenumbers presented high loading values. Therefore, the following wavenumbers, that have been previously reported in other FTIR studies on coffee, were selected for the final models: 2924, 2852, 1743, 1541, 1377, 1076, 910 and 816 cm^{-1}, with possible association to caffeine, carboxylic acids, lipids, chlorogenic acids, trigonelline and carbohydrates. The score plots for the first three discriminant functions are shown in Fig. 4. The first three discriminant functions accounted for 96.2, 95.2, 95.3 and 97.6% of of the total sample variance, for the models based on raw spectra, media-centered spectra, normalized spectra and first derivatives, respectively. A clear separation of all groups (non-defective, black, immature, dark sour and light sour) can be observed for the models based on DR spectra (see Figs 4a–c), whereas some level of group overlapping was observed for the model based on spectra derivatives (Fig. 4d). The calculated values of each discriminant function at the group centroids are displayed in Table 1. It is interesting to point out that, for all the developed models, the first three discriminant functions are enough to provide sample classification. For example, considering the model based on the raw spectra, it can be observed that non-defective coffees present positive values for DF1 and DF2 and negative values for DF3, whereas black beans present negative values for DF1, DF2 and DF3. The corresponding values obtained for correct classification rates for each specific model and group are shown in Table 2. Recognition and prediction abilities were quite similar for all the developed models.

The data were further evaluated in order to develop a more generic classification model, i.e., only one discrimination function that would provide discrimination between non-defective and defective beans, without separating the defects into specific groups. The classification functions and respective correct classification rates are shown in Table 3. Respective average values of recognition and prediction abilities were 96.4 and 100%, for the model based on raw spectra, 97.3 and 100%, for the model based on media-corrected spectra, 96.4 and 100%, for the model based on normalized spectra, and 94.6 and 95%, for the model based on first derivatives. Such results confirm that DRIFTS provides satisfactory discrimination between defective and non-defective roasted coffees, demonstrating its potential for detection of defective beans in mixtures with non-defective ones after roasting. Regarding the application of such methodology for routine analyses of roasted coffee quality, further studies are still necessary, involving a trained panel of coffee tasters, to establish the minimum amount, if any, in which defective beans can be introduced to a non-defective coffee batch and changes in the beverage quality would still not be perceived in relation to one without defective beans. With the minimum amounts effectively established, mixtures of defective and non-defective roasted beans can be suitably prepared and duly tested for the discrimination capability of the developed models.

4. Conclusion

The feasibility of employing DRIFTS as a methodology for discrimination between defective and non-defective roasted coffees was evaluated. The obtained spectra were similar, with small differences in absorbance intensity between non-defective and defective coffees. PCA results based on DR spectra and first derivatives indicated separation of the samples into four major groups: non-defective, black, dark sour and light sour, with immature beans scattered among the sour samples. LDA classification models, based on absorbance readings and derivatives at eight wavenumbers (2924, 2852, 1743, 1541, 1377, 1076, 910 and 816 cm^{-1}), provided separation of the samples into five groups: non-defective, black, dark sour, light sour and immature beans. Average recognition and prediction abilities ranged from 79 to 96% and from 80 to 100%, respectively. Discrimination functions for

generic classes of defective and non-defective coffee samples were also developed. For these generic models, recognition and prediction abilities ranged from 95 to 97% and from 95 to 100%, respectively. The results obtained in the present study confirm that DRIFTS provides satisfactory levels of discrimination between defective and non-defective coffee beans after roasting.

Acknowledgments

The authors acknowledge financial support from the following Brazilian Government Agencies: CNPq and FAPEMIG.

References

Bauer, R., Nieuwoudt, H., Bauer, F. F., Kossmann, J., Koch, K. R., & Esbensen, K. H. (2008). FTIR spectroscopy for grape and wine analysis. *Analytical Chemistry, 80*, 1371–1379.

Briandet, R., Kemsley, E. K., & Wilson, R. H. (1996). Approaches to adulteration detection in instant coffees using infrared spectroscopy and chemometrics. *Journal of the Science of Food and Agriculture, 71*, 359–366.

Clarke, R. J., & Macrae, R. (1987). *Coffee*In *Technology, Vol. 2*. London: Elsevier Applied Science.

Clifford, M. N., Kirkpatrick, J., Kuhnert, N., Roozendaal, H., & Salgado, P. R. (2008). LC-MSn analysis of the cis isomers of chlorogenic acids. *Food Chemistry, 106*, 379–385.

Craig, A. P., Franca, A. S., & Oliveira, L. S. (2011). Discrimination between immature and mature green coffees by attenuated total reflectance and diffuse reflectance Fourier transform infrared spectroscopy. *Journal of Food Science, 76*, C1162–C1168.

Craig, A. P., Franca, A. S., & Oliveira, L. S. (2012). Evaluation of the potential of FTIR and chemometrics for separation between defective and non-defective coffees. *Food Chemistry, 132*, 1368–1374.

Cremer, D. R., & Kaletunç, G. (2003). Fourier transform infrared microspectroscopic study of the chemical microstructure of corn and oat flour-based extrudates. *Carbohydrate Polymers, 52*, 53–65.

Esteban-Díez, I., González-Sáiz, J. M., Sáenz-González, C., & Pizarro, C. (2007). Coffee varietal differentiation based on near infrared spectroscopy. *Talanta, 71*, 221–229.

Farah, A., Monteiro, M. C., Calado, V., Franca, A. S., & Trugo, L. C. (2006). Correlation between cup quality and chemical attributes of Brazilian coffee. *Food Chemistry, 98*, 373–380.

Franca, A. S., Mendonça, J. C. F., & Oliveira, S. D. (2005). Composition of green and roasted coffees of different cup qualities. *LWT – Food Science and Technology, 38*, 709–715.

Franca, A. S., Oliveira, L. S., Mendonça, J. C. F., & Silva, X. A. (2005). Physical and chemical attributes of defective crude and roasted coffee beans. *Food Chemistry, 90*, 89–94.

Garrigues, J. M., Bouhsain, Z., Garrigues, S., & De La Guardia, M. (2000). Fourier transform infrared determination of caffeine in roasted coffee samples. *Fresenius Journal of Analytical Chemistry, 366*, 319–322.

Gordon, S. H., Schudy, R. B., Wheeler, B. C., Wicklow, D. T., & Greene, R. V. (1997). Identification of Fourier transform infrared photoacoustic spectral features for detection of *Aspergillus flavus* infection in corn. *International Journal of Food Microbiology, 35*, 179–186.

Hashimoto, A., Mori, H., Kanou, M., Kameoka, T., Shimazu, H., & Kobayashi, K. (2009). Mid-infrared spectroscopic characteristics of brewed coffee and simultaneous content determination of main components using an attenuated total reflection method. *Journal of the Illuminating Engineering Institute of Japan (Shomei Gakkai Shi), 93*, 501–509, (in Japanese).

Kemsley, E. K., Ruault, S., & Wilson, R. H. (1995). Discrimination between *Coffea arabica* and *Coffea canephora* variant robusta beans using infrared spectroscopy. *Food Chemistry, 54*, 321–326.

Lyman, D. J., Benck, R., Dell, S., Merle, S., & Murray-Wijelath, J. (2003). FTIR-ATR analysis of brewed coffee: effect of roasting conditions. *Journal of Agricultural and Food Chemistry, 51*, 3268–3272.

Mancha Agresti, P. C. M., Franca, A. S., Oliveira, L. S., & Augusti, R. (2008). Discrimination between defective and non-defective Brazilian coffee beans by their volatile profile. *Food Chemistry, 106*, 787–796.

Mazzafera, P. (1999). Chemical composition of defective coffee beans. *Food Chemistry, 64*, 547–554.

Mendonça, J. C. F., Franca, A. S., & Oliveira, L. S. (2009). Physical characterization of non-defective and defective Arabica and Robusta coffees before and after roasting. *Journal of Food Engineering, 92*, 474–479.

Mendonça, J. C. F., Franca, A. S., Oliveira, L. S., & Afonso, R. J. C. F. (2009). Application of electrospray ionization-mass spectrometry to the discrimination of green and roasted coffees by species and quality. In K. D. Lang (Ed.), *Handbook on mass spectrometry: Instrumentation, data and analysis, and applications* (pp. 269–290). New York: Nova Publishers.

Mendonça, J. C. F., Franca, A. S., Oliveira, L. S., & Nunes, M. (2008). Chemical characterization of non-defective and defective green Arabica and Robusta coffees by Electrospray Ionization-Mass Spectrometry (ESI-MS). *Food Chemistry, 111*, 490–497.

Miller, F. A., Mayo, D. W., & Hannah, R. W. (2003). *Course notes on the interpretation of infrared and Raman spectra*. Hoboken, NJ, USA: Wiley.

Oliveira, L. S., Franca, A. S., Mendonça, J. C. F., & Barros-Junior, M. C. (2006). Proximate composition and fatty acids profile of green and roasted defective coffee beans. *LWT – Food Science and Technology, 39*, 235–239.

Paradkar, M. M., & Irudayaraj, J. (2002). Rapid determination of caffeine content in soft drinks using FTIR-ATR spectroscopy. *Food Chemistry, 78*, 261–266.

Ramalakshmi, K., Kubra, I. R., & Rao, L. J. M. (2007). Physicochemical characteristics of green coffee: comparison of graded and defective beans. *Journal of Food Science, 72*, S333–S337.

Ribeiro, J. S., Salva, T. J., & Ferreira, M. M. (2010). Chemometric studies for quality control of processed Brazilian coffees using drifts. *Journal of Food Quality, 33*, 212–227.

Rodriguez-Saona, L., & Allendorf, M. E. (2011). Use of FTIR for rapid authentication and detection of adulteration of food. *Annual Review of Food Science and Technology, 2*, 467–483.

Sablinskas, V., Steiner, G., & Hof, M. (2003). Aplications. In G. Gauglitz, & T. Vo-Dinh (Eds.), *Handbook of spectroscopy* (pp. 89–168). Weinheim, Germany: Wiley-VCH.

Silverstein, M. R., Webster, F. X., & Kiemle, D. (2005). *Spectrometric identification of organic compounds*. Hoboken, NJ, USA: Wiley.

Vasconcelos, A. L. S., Franca, A. S., Glória, M. B. A., & Mendonça, J. C. F. (2007). A comparative study of chemical attributes and levels of amines in defective green and roasted coffee beans. *Food Chemistry, 101*, 26–32.

Wang, J., Jun, S., Bittenbender, H. C., Gautz, L., & Li, Q. X. (2009). Fourier transform infrared spectroscopy for Kona coffee authentication. *Journal of Food Science, 74*, C385–C391.

Yang, H., & Irudayaraj, J. (2001). Comparison of near-infrared, Fourier transform-infrared, and Fourier transform-Raman methods for determining olive pomace oil adulteration in extra virgin olive oil spectroscopy. *Journal of the American Oil Chemists' Society, 78*, 889–895.