Talanta 134 (2015) 379–386

Contents lists available at ScienceDirect

Talanta

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

# Fourier transform infrared spectroscopy and near infrared spectroscopy for the quantification of defects in roasted coffees

Ana Paula Craig [a,b], Adriana S. Franca [c,*], Leandro S. Oliveira [a,c], Joseph Irudayaraj [b], Klein Ileleji [b]

[a] PPGCA, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, 31270-901 Belo Horizonte, MG, Brazil
[b] Department of Agricultural and Biological Engineering, Purdue University, 225 S. University Street, West Lafayette, IN 47907, USA
[c] DEMEC, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, 31270-901 Belo Horizonte, MG, Brazil

ARTICLE INFO

*Article history:*
Received 16 September 2014
Received in revised form
17 November 2014
Accepted 18 November 2014
Available online 27 November 2014

*Keywords:*
Defective coffee
PLSR
FTIR
NIR

ABSTRACT

The coffee strip-picking harvesting method, preferred in Brazil, results in high percentages of immature and overripe beans, as the fruits in a single tree branch do not reach ripeness at the same time. This practice, together with inappropriate processing and storage conditions, contribute to the production of high amounts of defective coffee beans in Brazil, which upon roasting will impart negative sensory aspects to the beverage. Therefore, the development of analytical methodologies that will enable the discrimination and quantification of defective and non-defective coffees after roasting is rather desirable. Given that infrared spectroscopy has been successfully applied to coffee analysis, the objective of this work was to evaluate and to compare the performances of Fourier transform infrared (FTIR) and near infrared (NIR) spectroscopies for the quantification of defective beans in roasted coffees. Defective and non-defective Arabica coffee beans were manually selected, roasted, ground and sieved. Mixtures of defective and non-defective roasted and ground coffees were produced and analyzed, with % defects ranging from 0% to 30%. FTIR and NIR spectra were recorded, respectively, within a range of 3100–800 $cm^{-1}$ and 1200–2400 nm and submitted to mathematical processing. Quantitative models were developed by partial least squares regression (PLSR). Excellent predictive results were obtained indicating that defective coffees could be satisfactorily quantified. The correlation coefficients and the root mean squared errors of validation for the FTIR and NIR models developed to quantify the amount of defective roasted coffees mixed with non-defective ones were, respectively, as high as 0.891 and as low as 0.032, and as high as 0.953 and as low as 0.026. A comparison between the two techniques indicated that NIR provided more robust models.



## 1. Introduction

For every ten cups of coffee consumed in the world, approximately three come from beans produced in Brazil. In 2013, Brazil produced 45,152 million bags of coffee, which was almost twice the amount produced by the second largest producer, Vietnam (27,500 million bags) [1]. In order to achieve such a high number of coffee bags produced every year, the strip-picking harvesting practices are preferred in Brazil, and such practices usually result in coffees with high amounts of immature or unripe and overripe beans, as the coffee fruits in a single branch of the coffee plant do not reach ripeness at the same time. Furthermore, the harvest of fallen and fermented fruits in contact with the ground may also result in low quality beans. These practices, together with inappropriate processing and storage conditions contribute to the production of defective beans that comprise about 20% of the total coffee produced in Brazil [2–5]. Considered improper for exportation, defective beans are separated from non-defective ones by optical sorting machines prior to commercialization [3,4]. However, as these beans represent an investment in growing, harvesting and handling in the coffee production chain, coffee producers have adopted the practice of incorporating the separated beans into the Brazilian internal market in mixtures with non-defective ones, giving rise to a low-grade roasted and ground coffee [3].

*Abbreviations:* ATR, attenuated total reflectance; D, particle diameter; DLATGS, deuterated triglycine sulphate doped with L-alanine; FTIR, Fourier transform infrared spectroscopy; LV, latent variable; MSC, multiplicative scatter correction; NIR, near infrared; PLSR, partial least squares regression; RH, relative humidity; RMSEC, root mean square error of calibration; RMSECV, root mean square error of cross validation; RMSEP, root mean square error of prediction.

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

http://dx.doi.org/10.1016/j.talanta.2014.11.038
0039-9140/

Among the beans with irregular visual appearance, the most detrimental ones to the brew's flavor are immature, sour and black. Immature beans contribute to astringency, bitterness and metallic tastes. Sour beans are associated with 'overfermentation' caused by unfavorable conditions of temperature and humidity during processing, storage or transportation, being fermented by bacteria or xenophilic moulds. These beans impart sour, oniony and fermented taste and smell to the beverage. Black beans derive from beans that died within the cherry while still on the tree, from over-ripe cherries fallen on the ground or from beans that were attacked by fungi and other pests. This defect is generally regarded as giving a 'heavy' and 'ashy' flavor to the beverage and is considered the worst intrinsic defect [4,6].

Coelho et al. [7], studying the sensory impact of the inclusion of defective beans into coffees, observed that the transition of the cup quality from 'strictly soft' to 'hard' occurred after the addition of 19.5%, 16.4% and 14.3% of immature, sour or black beans, respectively. Puerta-Quintero [8] reported that 2.5% of immature beans mixed with non-defective ones was sufficient to promote rejection of 30% of samples by cuppers due to unpleasant tastes, while Bee and coworkers [4] reported that, in espresso coffee, the astringency and metallic tastes of immature beans can be perceived at quantities as low as 1%. In an attempt to identify the compounds that explain these effects, Toci and Farah [9] quantitatively determined 159 volatile compounds in defective and non-defective crude and roasted coffees. Defective beans presented a broader spectrum of volatile compounds than those presented by non-defective ones. Also, the volatile compounds identified in both defective and non-defective beans were present in higher concentrations in the latter, especially pyrazines, pyrroles and phenols.

In order to mask these detrimental flavors and/or aromas caused by the presence of defective beans, low-grade coffees are generally over roasted to a dark roasting degree. This procedure makes a suitable assessment of the quality of commercial coffees impractical, as the sensory analysis or cup-test still remains the ultimate tool to assess coffee quality. Aside from the undesirable sensory impact, studies suggest that the presence of defective beans may pose risks for human health due to the high incidence of ochratoxin A [10,11], which is only partially reduced during the roasting and extraction processes [12]. In the recent study by Taniwaki and coworkers [11], the presence of ochratoxigenic fungi and ochratoxin A in non-defective and in defective raw coffee beans was evaluated. The results indicated that all defective beans were infected with *Aspergillus carbonarius*, *A.* section *Nigri*, *A. westerdijkiae* or *A.* section *circumdati*, and highest ochratoxin A concentrations were observed in sour and black beans.

Good examples of rapid, reliable and promising fingerprint techniques that could be used to assess the overall coffee quality attributes are mid and near infrared spectroscopy. The mid infrared region (4000–400 $cm^{-1}$), with the corresponding spectroscopic method referred as Fourier transform infrared (FTIR), detects fundamental vibrations bands whereas the near infrared (NIR) spectrum (800–2500 nm) arises from the molecular absorptions of overtones and combinations of fundamental vibration bands in the mid infrared region [13,14]. A literature review clearly reveals that, among other applications [15], both FTIR and NIR can be effective for characterization and quantification of chemical attributes such as ash, lipids and caffeine content [16,17], for discrimination and quantification of arabica and robusta blends [18], for detection of adulterants [19–21] and for prediction of sensory properties and roasting degree [22,23]. In particular, previous works have shown that infrared spectroscopy is capable of discriminating non-defective from black, sour and immature defective beans in crude and roasted coffees [24–27].

Santos et al. [27] successfully developed a methodology based on NIR to quantify crude defective beans among non-defective ones which could enable the fast assessment of coffee grade. More recently, we have presented a comparative evaluation of the performances of FTIR and NIR for the qualitative discrimination of roasted defective and non-defective coffees, employing a novel statistical approach, Elastic Net [28]. The Elastic net models exhibited high percentages of correct classification. Furthermore, they provided insights on the characterization of the samples and on the visualization of discrete spectral bands associated with the correct classification of defective and non-defective coffees. The correct classification of non-defective coffees was associated to absorbance regions that are characteristic of carbohydrates (1138–1165 $cm^{-1}$, 1760–1871 nm) and lipids (1722–1759 $cm^{-1}$, 2810–2848 $cm^{-1}$, 2908–2920 $cm^{-1}$, 1680–1755 nm, 2132–2166 nm). Although the understanding of the chemical differences between high and low quality beans is scientifically relevant, in practice, commercial roasted coffees comprise a mixture of defective and non-defective beans. Therefore, the development of a methodology aiming to detect and quantify defective beans mixed with non-defective ones must be considered as a reliable analytical tool to regulate coffee quality.

In view of the aforementioned, the objective of this work was to further investigate the potential of FTIR and NIR spectroscopies to evaluate the quality of coffees based on the presence of defective beans. The major goals were to develop quantitative models based on partial least squares regression (PLSR) to predict the percentage of defective coffees in admixtures with non-defective ones and to compare the performance of FTIR and NIR techniques for this purpose.

## 2. Material and methods

### 2.1. Preparation of coffee samples and standard mixtures

Arabica green coffee samples were acquired from a roasting company located in Minas Gerais State, Brazil. Samples consisted of coffee beans harvested by strip-picking that were rejected by color sorting machines. The beans were manually sorted (by a professional trained and certified for green coffee classification) into five lots or sample classes: non-defective, immature, black, light sour and dark sour beans.

Samples of 25 g were taken from each lot and roasted in a convection oven (Model 4201D Nova Ética, SP, Brazil) at 235 °C. In our previous study [28], we showed that defective and non-defective coffees can be successfully discriminated based on their infrared spectra regardless of the roasting condition of the beans, which means that the variance due to beans quality is larger than the variance due to roasting degree. Therefore, in this study, all samples were roasted to a medium roasting degree similar to commercially available coffee samples. In order to achieve this roasting degree, each sample was roasted to a specific roasting time. Roasting times ranged from 10 to 15 min. Samples were then ground in a coffee grinder (Arbel, Brasil) and color evaluation was 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°. Roasting degree was evaluated on ground samples by luminosity ($L^*$) measurements. Based on previous analysis of commercial coffees, a medium roasting degree was defined as $21 < L^* < 23.5$ [28]. Sequentially, samples were sieved. Fractions with 0.25 > particles diameter > 0.15 mm and 0.84 > particles diameter > 0.39 mm were employed for the FTIR and NIR experiments, respectively. The appropriate particle size ranges were chosen based on preliminary tests performed in previous studies [26,28], aiming at the conditions that provided the best quality spectra (higher intensity and lower noise interference).

Coffee mixtures (blends) were prepared in five replicates by mixing each type of defective coffee with non-defective ones, with the amount of defects ranging from 3 to 30% in steps of 3% (10 blends for each of the four defects). In addition, blends containing a mixture of the four defects (25% of each defect) with non-defective coffee were produced. Pure samples of non-defective coffee, representing 0% of defects, were also taken and analyzed. Therefore, 55 samples were obtained comprising the following blends: (a) light sour in admixture with non-defective coffee, (b) dark sour in admixture with non-defective coffee, (c) black in admixture with non-defective coffee, (d) immature in admixture with non-defective coffee and (e) defects (25% of each defect) in admixture with non-defective coffee. The blends were placed in Falcon tubes and shaken for 1 min in a tube shaker (Fisatom, Brazil). All samples were stored at room temperature (20 °C).

### 2.2. FTIR and NIR measurements and spectral collection

A Shimadzu IRAffinity-1 FTIR Spectrophotometer (Shimadzu, Japan) with a DLATGS detector was used in the FTIR measurements that were performed in dry atmosphere (20 ± 0.5 °C, RH=25%). A horizontal Attenuated Total Reflectance (ATR) sampling accessory (ATR-8200HA) equipped with ZnSe cell was employed. Approximately 2 g of each sample were placed in the sampling accessory and pressed in order to obtain the best contact with the crystal surface. The empty accessory was used to generate the background spectrum. The approximate total time required for analysis (including adequately pressing the sample onto the ATR crystal and obtaining both the background and sample spectra) was 5 min. All spectra were recorded within a range of 3100–800 $cm^{-1}$, with a 4 $cm^{-1}$ resolution. Each spectrum was calculated as the average of 20 scans and submitted to background subtraction.

A SpectraStar 2400 Drawer NIR spectrophotometer (Unity Scientific) with an InGaAs detector was used in the measurements. Approximately 3 g of samples were placed inside a glass cup, filling the entire empty space, and covered. Atmosphere air was used to obtain the background spectra. The approximate total time required for analysis was 2 min. All spectra were recorded within a range of 1200–2400 nm with 1 nm resolution. Each spectrum was calculated as the average of 30 scans and submitted to background subtraction. In both FTIR and NIR experiments, five replicates of each sample were analyzed, resulting in a total of 275 spectra for each technique.

### 2.3. Data analysis

Data processing (pretreatment) techniques were applied to the raw data to compensate for any changes in experimental conditions and enhance the results. The processing methods that provided the best performances in terms of model prediction were the following: baseline correction followed by area normalization (FTIR); 1st derivative (FTIR); baseline correction (NIR); and multiplicative scatter correction (MSC) (NIR). All datasets were mean-centered prior to statistical analyses.

PLSR was the technique of choice for the quantification of mixtures of defective and non-defective coffees. The optimum number of latent variables (LV) employed in each model was chosen by leave-one-out cross-validation based on the minimum value of root mean square error for cross validation (RMSECV). The combination of *Q*-residues and the Hotelling's *T*-squared distribution (*T2*) was used to detect abnormal observations in the calibration set. Given the significance level for the *Q* and *T2* statistics, in this case, 99%, observations with *Q* and/or *T2* values above the threshold were classified as outliers. After the elimination of the outlier observations from the model, the procedure was continually repeated until no outliers were identified. The random behavior of the residuals of the fits was verified by visual inspection. The evaluation of the accuracy of the models was based on the following parameters: the correlation coefficient (*R*) that should be as close to 1 as possible; and the root mean square errors for both the calibration (RMSEC) and validation (RMSEP) sets, that should be as small as possible. The latter parameters were calculated as follows:

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

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

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, and $v$ is the number of degrees of freedom, or the number of latent variables used in the model plus 1 for mean centered data. The softwares Matlab (The MathWorks, Co., Natick, MA) and the computational package PLS_Toolbox (Eigenvector Research, Inc.) were employed for the statistical calculations.

## 3. Results and discussion

### 3.1. Overall characteristics of the FTIR and NIR spectra of defective and non-defective coffees

Fig. 1 shows the average original spectra (Fig. 1a), the average spectra after baseline correction and normalization (Fig. 1b), and the average 1st derivative spectra (Fig. 1c) of defective and non-defective pure coffees obtained by FTIR. Major peaks were observed at 2920 $cm^{-1}$, 2859 $cm^{-1}$, 1747 $cm^{-1}$ and at 1400–900 $cm^{-1}$. These bands have been previously identified in arabica and robusta coffees, prior to and after roasting [24,26,29–30]. In our previous study [28], the aforementioned regions of the FTIR spectra were selected as important variables for the discrimination of defective and non-defective coffee. The region 2940–2820 $cm^{-1}$ is associated with symmetric (sym) and asymmetric (asym) stretching of CH bonds in $CH_2$ and $CH_3$ groups [31]. The sym and asym stretching of $CH_2$ is highly related to the presence of lipids [22,32], while the vibration of $CH_3$ presents great relevance in the identification of caffeine [33].

The sharp band at 1747 $cm^{-1}$ is assigned to C=O stretch of aliphatic ester groups, thus it is mostly related to the presence of lipids. The regions around 1747 $cm^{-1}$, also related to C=O stretch, are assigned to different functional groups including aliphatic and aromatic acids, aldehydes, ketones and lactones. Such compounds confer different aromas to the coffee, making this an important region of the spectrum from a sensory point of view [29,34–36]. The third region, from 1400 to 900 $cm^{-1}$, is commonly called the *fingerprint* region because of the large amount of characteristic single bands attributed to specific functional groups. Among these groups, C–H, C–O, C–N and P–O bonds are included [31]. In particular, carbohydrates exhibit large features in this region [30]. The most relevant visual difference between the defective and non-defective coffees original spectra (Fig. 1a) was associated with a shift in the baseline, which was eliminated with the application of processing techniques.

The average original and processed spectra of defective and non-defective pure coffees obtained by NIR are shown in Fig. 1d–f. The shape of the spectra was particularly dominated by broad water absorbance bands at 1440–1480 nm (1st overtone of O–H stretching) and 1930–1950 nm (combination band of O–H

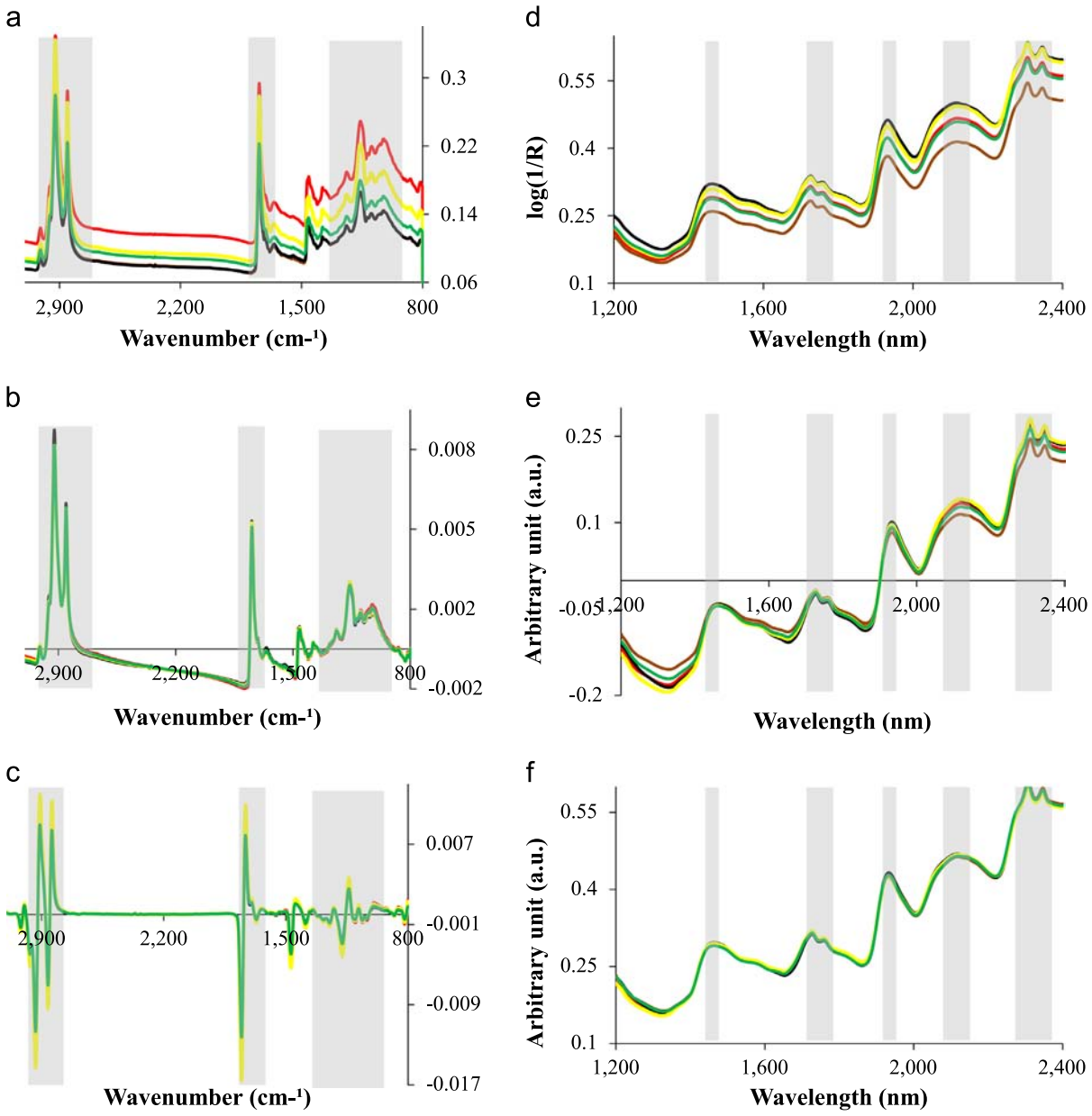


**Fig. 1.** Mean average spectra of pure defective and non-defective coffees obtained by FTIR, (a) original and preprocessed with (b) baseline correction and normalization, and (c) 1st derivative. Mean average spectra of pure defective and non-defective coffees obtained by NIR, (d) original and preprocessed with (e) baseline correction and (f) MSC. non-defective; –immature; sour (light); sour (dark); –black.

stretching and O–H deformation). Other regions that could be visually identified have been reported in the literature as characteristic absorbance regions of specific compounds such as lipids, which absorb at two well defined regions at 1715–1760 nm and 2300–2350 nm [23,37]. Although it is not possible to visually discriminate the NIR spectra of defective and non-defective coffees, in our previous study [28], the region of 1715 to 1760 nm exhibited positive regression coefficients associated with the correct classification of non-defective and light sour coffee. In the employed statistical approach (discrimination by Elastic Net), a non-zero regression coefficient indicates that correct classification of a given sample class is associated with the corresponding spectral region. Positive coefficients indicate higher absorbance intensity at that range of the spectrum, possibly associated with higher concentration of a specific compound, and negative coefficients indicate the opposite. Also in that study, the correct classification of black and immature coffees was associated with negative coefficients in this region [28]. The region around 2100 nm is related to carbohydrates, caffeine, chlorogenic acids and/or proteins [23,37]. In our previous work, non-defective and black coffee exhibited positive, while dark sour exhibited negative coefficients around this region. Nevertheless, the significant overlap of combination bands at this region hindered a precise interpretation of the samples discrimination.

### 3.2. PLSR models for quantitative analysis of defective and non-defective coffees

Due to the complexity of the FTIR and NIR spectra, the development of the PLSR models was based on a full-spectrum approach, including a number of structural information of compounds. The quantitative datasets were split into calibration (73% of the spectra) and validation (27% of the spectra) sets. The procedure for outlier detection and removal in the calibration

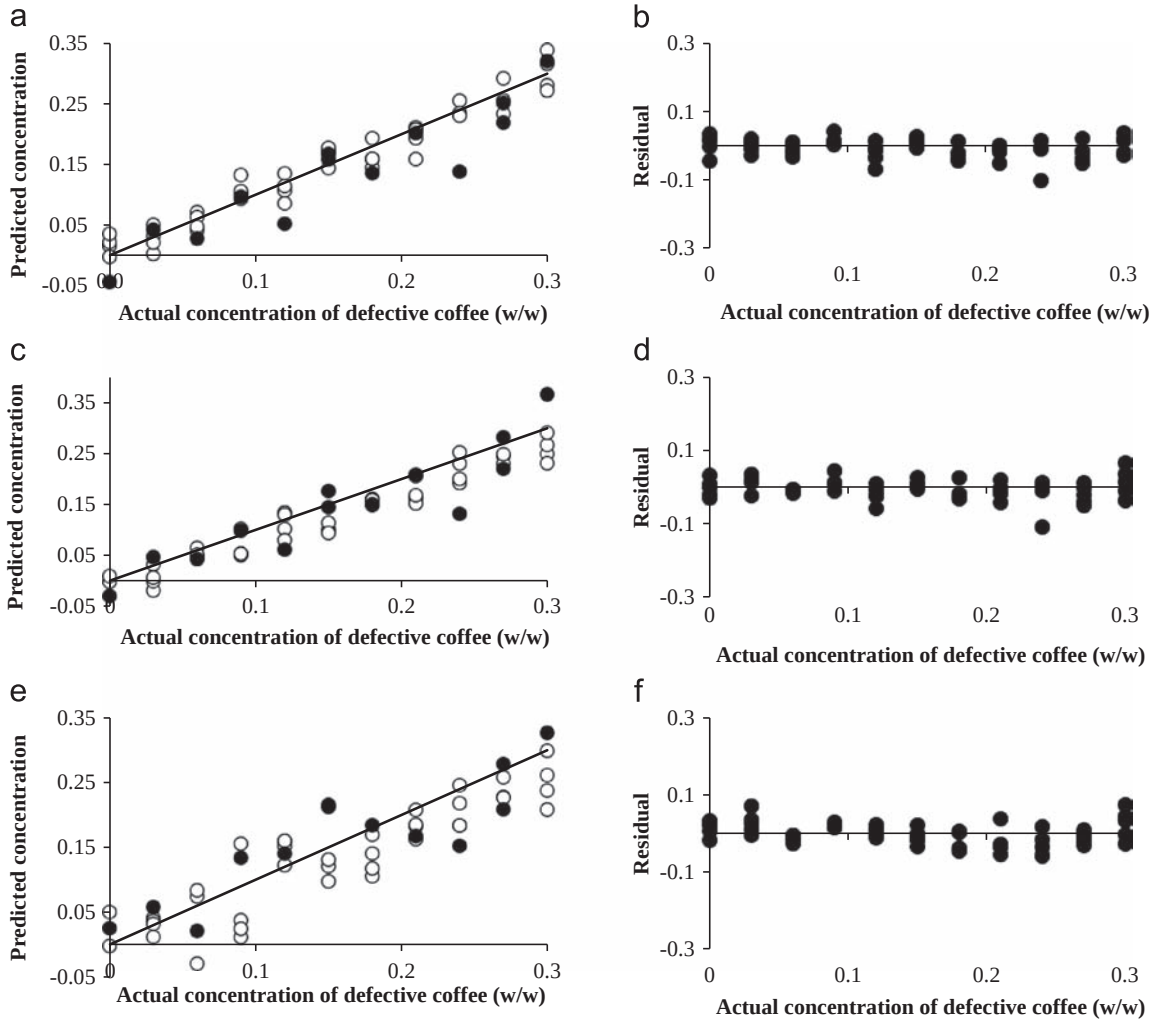


**Fig. 2.** Actual vs. predicted concentration of defective coffee (w/w), and residual vs. actual concentration of defective coffee for the FTIR models constructed with (a and b) original spectra and spectra preprocessed with (c and d) baseline correction and normalization and (e and f) 1st derivative. ○ calibration samples; • validation samples.

set, at 99% confidence level, was performed and repeated until no outliers were identified. In this scenario, the presence of outliers can be related to operational errors, instrumental noise or abnormal observations originated from errors or differences during the sample weighing and production of the mixtures. According to the Protocol for Design, Conduct and Interpretation of Method-Performance Studies [38], outliers can be removed up to a limit of 22% of the total number of samples. The number of outliers detected and removed from the FTIR and NIR models varied, respectively, from 0 to 4% and from 0 to 2% of the total number of samples.

The scatter plots of actual and predicted values for percentage of the mixture of defects in admixtures with non-defective coffees are shown in Fig. 2 (FTIR) and 3 (NIR). Visual inspection of the figures suggests that the models could predict the percentage of defects satisfactorily. The residuals, which were plotted vs. sample percentage values, were distributed randomly and satisfactorily close to zero, indicating no apparent systematic trend (Figs. 2b, d, f and 3b, d, f).

Tables 1 and 2 show the performance results of the optimized PLSR models. The number of LVs used in the FTIR models (Table 1) ranged from 3 to 10 and accounted for at least 94.6% of the variance in *X* (spectral data) and 84.3% of the variance in *Y* (percentage of defects). In the NIR models (Table 2), the number of LVs ranged from 4 to 7 and accounted for at least 89.9% and 92.1% of the variance in *X* and *Y*, respectively.

A comparison between the FTIR models constructed with original (only mean centered) spectra and spectra submitted to baseline correction followed by area normalization and 1st derivative shows that the application of mathematical processing provided potential benefits to the regression models, as expected. Overall, the errors (RMSEC and RMSEP) were reduced and the correlation coefficient (*R*) of the models increased. The same observation holds true for all NIR models. Beyond that, the number of LVs was a key feature. To achieve satisfactory predictive results, a higher number of LVs was used in the development of most of the models constructed with original spectra. In reference to ASTM [39], the determination of the number of LVs to be used is a critical step in the model development. In general, if too few variables are used, a less accurate model will result. If too many variables are used, the estimates from the model will be unstable, which means that small changes in the spectrum, on the order of the spectral noise, may produce statistically significant changes in the estimates. Thus, models with fewer factors are less likely to exhibit over fitting and tend to have better generalization capabilities.

Overall, all classes of defects could be satisfactorily quantified in admixtures with non-defective coffee. The quantification results for light sour coffee, solely by the preprocessed FTIR method were, however, found not to be as satisfactory, with greater number of LVs employed and relatively high errors and low correlation coefficient in the validation set. It was observed in our previous studies that normal and light sour coffees could not be effectively

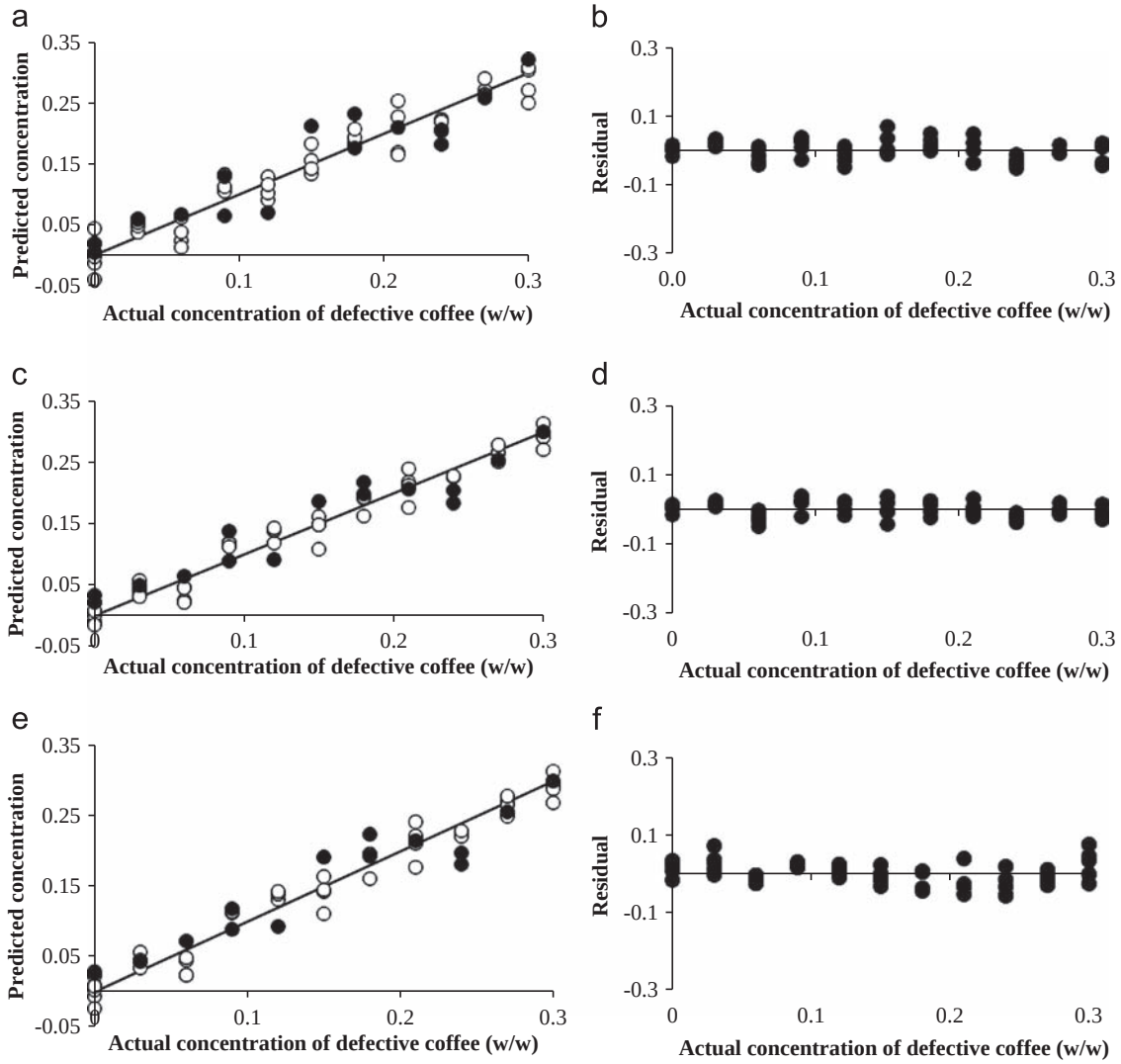


**Fig. 3.** Actual vs. predicted concentration of defective coffee (w/w), and residual vs. actual concentration of defective coffee for the NIR models constructed with (a and b) original spectra and spectra preprocessed with (c and d) baseline correction and (e and f) MSC. ○ calibration samples; • validation samples. The straight lines (a, c and e) indicate equal values for predicted and actual concentrations.

discriminated by FTIR, and these classes were clustered together on PCA analysis [24,26]. The official New York Coffee and Sugar Exchange type-classification system, in which coffee are classified as defective and non-defective, considers several types of defects with sour being one of them. Our previous work [24] has shown that, when adopting this type-classification system for studies on discrimination of defects, the defect 'sour' tended to cause a natural separation within its class, with a portion of them grouping together with the 'graded' coffees and the other portion grouping with the 'black' coffees in Principal Components Analysis. The sour-type defect is characterized by a wide spectrum of color shades between that of graded and that of black beans, representing distinct levels of fermentation. Thus, this type of defect was further divided into two distinct types: light sour representing the lightly fermented beans; and dark sour, representing the beans that were more intensely fermented during processing. Therefore, it seems that the level of fermentation that the light sour beans have undergone did not promote significant changes to the beans in order for them to be dully differentiated from the graded ones.

A comparative evaluation between the correlation coefficients and errors of prediction indicates that the preprocessed NIR spectra models were superior to the preprocessed FTIR spectra models. The correlation coefficients and mean squared errors of prediction of the preprocessed FTIR spectra models ranged, respectively, from 0.747 to 0.9 and from 0.032 to 0.048. The same parameters for the preprocessed NIR models ranged, respectively, from 0.799 to 0.941 and from 0.027 to 0.043. Most importantly, the number of LVs varied from 5 to 10 for the processed FTIR spectra models and from 4 to 6 for the processed NIR spectra models, indicating that NIR provided more robust models. It must be pointed out that, although NIR provided more robust quantitative models, in our previous work [28] we have demonstrated that FTIR can provide more chemical information and selectivity on the discriminating group frequencies of defective and non-defective coffees. Indeed, it is well known that precise band assignments are difficult in the near infrared region due to the fact that a single band may be attributable to several possible combinations of fundamental and overtone vibrations overlapped.

From our knowledge, this is the first time an analytical methodology able to quantify defective beans in roasted coffees is reported. Santos and coworkers [27] applied NIR spectroscopy and PLSR to quantify the mass fraction of defective beans in crude and whole Arabica coffee beans. According to the authors a major difficulty when analyzing the whole crude beans is that defects may be present only in some parts of the beans (e.g., partially

**Table 1**
Performance results of optimized PLSR models based on FTIR spectra.

| Models | LV | Variance accounted | | Calibration | | Validation | |
|---|---|---|---|---|---|---|---|
| | | *X* | *Y* | RMSEC | $R_c$ | RMSEP | $R_v$ |
| Mixture of defects | | | | | | | |
| Original | 8 | 99.96 | 84.26 | 0.021 | 0.949 | 0.042 | 0.863 |
| Baseline+normalization | 7 | 97.46 | 95.20 | 0.021 | 0.932 | 0.044 | 0.832 |
| 1st derivative | 5 | 98.19 | 92.58 | 0.026 | 0.926 | 0.032 | 0.891 |
| Light sour | | | | | | | |
| Original | 10 | 99.86 | 73.68 | 0.014 | 0.977 | 0.043 | 0.786 |
| Baseline+normalization | 8 | 98.31 | 97.66 | 0.014 | 0.977 | 0.045 | 0.784 |
| 1st derivative | 9 | 99.36 | 97.79 | 0.048 | 0.978 | 0.048 | 0.747 |
| Dark sour | | | | | | | |
| Original | 8 | 99.97 | 94.87 | 0.021 | 0.949 | 0.033 | 0.881 |
| Baseline+normalization | 5 | 96.69 | 94.26 | 0.023 | 0.943 | 0.034 | 0.857 |
| 1st derivative | 5 | 96.59 | 96.47 | 0.018 | 0.965 | 0.039 | 0.847 |
| Black | | | | | | | |
| Original | 3 | 99.92 | 90.97 | 0.028 | 0.911 | 0.042 | 0.817 |
| Baseline+normalization | 5 | 96.98 | 92.99 | 0.025 | 0.930 | 0.042 | 0.839 |
| 1st derivative | 10 | 99.76 | 99.70 | 0.005 | 0.997 | 0.039 | 0.847 |
| Immature | | | | | | | |
| Original | 7 | 99.96 | 86.47 | 0.035 | 0.889 | 0.039 | 0.865 |
| Baseline+normalization | 6 | 94.61 | 87.13 | 0.034 | 0.884 | 0.043 | 0.871 |
| 1st derivative | 6 | 98.05 | 90.33 | 0.029 | 0.903 | 0.034 | 0.902 |

**Table 2**
Performance results of PLSR models based on NIR spectra.

| Models | LV | Variance accounted | | Calibration | | Validation | |
|---|---|---|---|---|---|---|---|
| | | *X* | *Y* | RMSEC | $R_c$ | RMSEP | $R_v$ |
| Mixture of defects | | | | | | | |
| Original | 5 | 99.67 | 92.13 | 0.024 | 0.940 | 0.034 | 0.871 |
| Baseline | 5 | 97.61 | 94.73 | 0.018 | 0.964 | 0.029 | 0.913 |
| MSC | 4 | 97.27 | 94.80 | 0.019 | 0.963 | 0.030 | 0.906 |
| Light sour | | | | | | | |
| Original | 7 | 99.63 | 97.72 | 0.015 | 0.977 | 0.038 | 0.837 |
| Baseline | 6 | 97.02 | 96.75 | 0.018 | 0.968 | 0.041 | 0.818 |
| MSC | 6 | 92.16 | 97.04 | 0.017 | 0.970 | 0.043 | 0.799 |
| Dark sour | | | | | | | |
| Original | 6 | 99.89 | 97.52 | 0.016 | 0.975 | 0.026 | 0.953 |
| Baseline | 5 | 99.21 | 97.01 | 0.017 | 0.970 | 0.027 | 0.943 |
| MSC | 4 | 91.33 | 96.35 | 0.019 | 0.963 | 0.029 | 0.941 |
| Black | | | | | | | |
| Original | 6 | 99.54 | 96.42 | 0.019 | 0.964 | 0.028 | 0.918 |
| Baseline | 5 | 96.48 | 95.76 | 0.021 | 0.958 | 0.029 | 0.915 |
| MSC | 4 | 87.77 | 95.21 | 0.022 | 0.979 | 0.030 | 0.905 |
| Immature | | | | | | | |
| Original | 5 | 99.14 | 96.17 | 0.019 | 0.962 | 0.055 | 0.723 |
| Baseline | 5 | 96.35 | 98.62 | 0.012 | 0.986 | 0.029 | 0.903 |
| MSC | 5 | 89.86 | 99.50 | 0.007 | 0.995 | 0.031 | 0.895 |

black beans), inducing the irreproducibility. In addition, using a spinning accessory for spectral acquisition does not assure that all the surface area of the beans will face the acquisition window. Still, the strategy proposed by Santos and coworkers [27] can be a valuable tool for fast assessment of crude coffee grade following general international guidelines of coffee classification [40,41]. The region 2000–2500 nm of the spectra, where a number of combination bands most likely associated with lipids, proteins, chlorogenic acids and caffeine take place [28], was reported by the authors as the region that contributed the most to the PLSR quantification models. An evaluation of the loadings plot of the NIR models constructed in our study also indicated that the aforementioned region greatly contributed to the quantification of defective beans in roasted coffee. Other regions of the spectra that significantly contributed to the models were 1200–1400 nm, characterized by 1st overtone vibrations of C–H and O–H, and 1840–2050 nm, where a number of combination and 2nd overtone bands take place [28].

## 4. Conclusion

In this study, we presented an evaluation of the potential of FTIR- and NIR-based methods for the quality assessment of roasted coffees regarding the presence of defective beans. The methods involved the development of PLSR models for the quantification of defective beans in concentrations ranging from 0% to 30%, in mixtures with non-defective coffee. The correlation coefficients and the root mean squared error of validation for the FTIR model developed to quantify a mixture of the four defects in blends with non-defective coffee were, respectively, as high as 0.891 and as low as 0.032. The same parameters were, respectively, as high as 0.953 and as low as 0.026 for the NIR model, indicating that both techniques provided accurate predictive results. A comparative evaluation between the two spectroscopic techniques, taking into account the aforementioned quality parameters and the number of latent variables employed in each model, indicated that NIR provided quantitative models that were slightly more robust than the ones based on FTIR for this application.

## Acknowledgments

The authors acknowledge financial support from the Brazilian Government Agencies CNPq (Grants # 505001/2013-6, 306139/2013-8 and 475746/2013-9) and FAPEMIG.

## References

[1] ICO, International Coffee Organization. Trade statistics, 2014.
[2] F.M. Borém, Handbook of Coffee Post-harvest Technology, Gin Press, Norcross, GA, 2014.
[3] A.S. Franca, L.S. Oliveira, Chemistry of defective coffee beans, in: E.N. Koeffer (Ed.), Progress in Food Chemistry, Nova Publishers, New York, NY, 2008, pp. 105–138.
[4] S. Bee, C. Brando, G. Brumen, N. Carvalhaes, I. Kolling-Speer, K. Speer, F. Suggi Liverani, A. Texeira, R. Thomaziello, R. Viani, The raw bean, in: A. Illy, R. Viani (Eds.), Espresso Coffee: The Science of Quality, second ed., Elsevier Academic Press, San Diego, 2005, pp. 87–178.
[5] A.T. Toci, A. Farah, Food Chem. 108 (2008) 1133–1141.
[6] R.J. Clarke, Grading, storage, pre-treatments and blending, in: R.J. Clarke, R. Macrae (Eds.), Coffee: Volume 2, Technology, Elsevier Applied Science, New York, NY, 1987, pp. 35–58.
[7] K.F. Coelho, R. Pereira, E. Vilela, Carvalho, V.D, Avaliação de algumas características sensoriais do café de bebida estritamente mole após a inclusão de grãos defeituosos, in: Congresso Brasileiro de Pesquisas Cafeeiras, 26, IBC, Rio de Janeiro, 2000, pp. 119–120.
[8] G. Puerta-Quintero, Cenicafé 51 (2000) 136–150.
[9] A.T. Toci, A. Farah, Food Chem. 153 (2014) 298–314.
[10] C.F. Silva, R.F. Schwan, Ë. Sousa Dias, A.E. Wheals, Int. J. Food Microbiol 60 (2000) 251–260.
[11] M. Taniwaki, A. Teixeira, A. Teixeira, M. Copetti, B. Iamanaka, Food Res. Int. 61 (2014) 161–166.
[12] H. Tsubouchi, H. Terada, K. Yamamoto, K. Hisada, Y. Sakabe, J. Agric. Food Chem. 36 (1988) 540–542.
[13] A.P. Craig, A.S. Franca, J. Irudayaraj, Vibrational spectroscopy for food quality and safety screening, in: A.K. Bhunia, M.S. Kim, C.R. Taitt (Eds.), High Throughput Screening for Food Safety Assessment: Biosensor Technologies, Hyperspectral Imaging and Practical Applications, Elsevier Applied Science, London, 2015, pp. 165–194.
[14] D.F. Barbin, A.L.d.S.M. Felicio, D.-W. Sun, S.L. Nixdorf, E.Y. Hirooka, Food Res. Int. 61 (2014) 23–32.
[15] P.D.C. Mancha Agresti, A.S. Franca, L.S. Oliveira, R. Augusti, Food Chem. 106 (2008) 787–796.
[16] C.W. Huck, W. Guggenbichler, G.K. Bonn, Anal. Chim. Acta 538 (2005) 195–203.
[17] C. Pizarro, I. Esteban-Díez, A.-J. Nistal, J.-M. González-Sáiz, Anal. Chim. Acta 509 (2004) 217–227.
[18] I. Esteban-Díez, J.M. González-Sáiz, C. Sáenz-González, C. Pizarro, Talanta 71 (2007) 221–229.
[19] H. Ebrahimi-Najafabadi, R. Leardi, P. Oliveri, M. Chiara Casolino, M. Jalali-Heravi, S. Lanteri, Talanta 99 (2012) 175–179.

[20] N. Reis, A.S. Franca, L.S. Oliveira, LWT Food Sci. Technol 50 (2013) 715–722.
[21] N. Reis, A.S. Franca, L.S. Oliveira, Talanta 115 (2013) 563–568.
[22] I. Esteban-Diez, J.M. Gonzalez-Saiz, C. Pizarro, Anal. Chim. Acta 525 (2004) 171–182.
[23] J.S. Ribeiro, M.M.C. Ferreira, T.J.G. Salva, Talanta 83 (2011) 1352–1358.
[24] A.P. Craig, A.S. Franca, L.S. Oliveira, Food Chem. 132 (2012) 1368–1374.
[25] A.P. Craig, A.S. Franca, L.S. Oliveira, J. Food Sci. 76 (2011) C1162–C1168.
[26] A.P. Craig, A.S. Franca, L.S. Oliveira, LWT Food Sci. Technol 47 (2012) 505–511.
[27] J.R. Santos, M.C. Sarraguça, A.O.S.S. Rangel, J.A. Lopes, Food Chem. 135 (2012) 1828–1835.
[28] A.P. Craig, A.S. Franca, L.S. Oliveira, J. Irudayaraj, K. Ileleji, Talanta 128 (2014) 393–400.
[29] E.K. Kemsley, S. Ruault, R.H. Wilson, Food Chem. 54 (1995) 321–326.
[30] G. Downey, R. Briandet, R.H. Wilson, E.K. Kemsley, J. Agric. Food Chem. 45 (1997) 4357–4361.
[31] R.M. Silverstein, F.X. Webster, D.J. Kiemle, Spectrometric Identification of Organic Compounds, seventh ed., John Wiley & Sons, Hoboken, NJ, 2005.
[32] N. Wang, L.-T. Lim, J. Agric. Food Chem. 60 (2012) 5446–5453.
[33] M.M. Paradkar, J. Irudayaraj, Food Chem. 78 (2002) 261–266.
[34] R. Karoui, G. Downey, C. Blecker, Chem. Rev. 110 (2010) 6144–6168.
[35] D.J. Lyman, R. Benck, S. Dell, S. Merle, J. Murray-Wijelath, J. Agric. Food Chem. 51 (2003) 3268–3272.
[36] J. Wang, S. Jun, H.C. Bittenbender, L. Gautz, Q.X. Li, J. Food Sci. 74 (2009) C385–C391.
[37] J.S. Shenk, M.O. Westerhaus, J. Workman, Application of NIR spectroscopy to agricultural products, in: E.W. Ciurczak, D.A. Burns (Eds.), Handbook of Near-Infrared Analysis, third ed.,CRC Press, Boca Raton, FL, 2007, pp. 347–386.
[38] W. Horwitz, Pure Appl. Chem. 67 (1995) 331–342.
[39] ASTM, Standard E1655-05, 1997, Standard practices for infrared multivariate quantitative analysis, ASTM International, West Conshohocken, PA, 2012.
[40] ISO, Standard 4149:1980, Green Coffee—Olfactory and Visual Examination and Determination of Foreign Matter and Defects, International Organization for Standardization, Geneva, 1980.
[41] ISO, Standard 10470:2004, Green Coffee—Defect Reference Chart, International Organization for Standardization, Geneva, 2004.