



Detection and Quantification of Adulterants in Roasted and Ground Coffee by NIR Hyperspectral Imaging and Multivariate Curve Resolution

Débora A. P. Forchetti¹ · Ronei J. Poppi¹

Received: 13 December 2018 / Accepted: 9 April 2019 / Published online: 25 April 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

In this work, a methodology was proposed for detection and quantification of the four main *roasted and ground* coffee adulterants: coffee husks, roasted and powdered corn kernels, wood sticks, and soil, based on the combination of near-infrared hyperspectral imaging and multivariate curve resolution (MCR). By using this procedure, it was possible to detect the adulterants and their confirmation was accomplished comparing the reference spectra of each adulterant with the spectra recovered by the MCR model. The methodology was suitable for the quantification of adulterants in the mixtures which ranged from 1 to 40% (*w/w*), with errors lower than 4%. Also, a chart control card based on the scores of the MCR was developed, in order to detect the adulterated samples in a qualitative way, without the need of concentration values.

Keywords *Coffee* · Adulterant · Near-infrared spectroscopy · Hyperspectral imaging · Multivariate curve resolution

Introduction

Coffee is a prominent world economic product due to the high consumption. According to the ICO (International Coffee Organization), the crop of 2014/2015 reached about 142,000,000 bags of 60 kg. During this period, Brazil remained as the first on the list of producers and exporters of coffee with 43,200,000 bags produced, with approximately 84% intended for export (ICO 2014). Despite the great movement of capital raised in the marketing of this product, the coffee has been portrayed in literature as the target of frequent tampering with additions of lower-cost products such as the shells of coffee husks, roasted corn, wood twigs or sticks, and soil, among others (Pauli et al. 2014; Reis et al. 2013a, b; Sano et al. 2003).

The ease of adulteration lies in the fact that the products added when roasted and ground, in special conditions, are confused with the original dust. Furthermore, in Brazil, there is a standard technique for detecting adulterants in coffee that is not subjective. According to the norms of ABIC (Brazilian Coffee Industry Association), created to ensure the quality of

the coffee, they are allowed up to 1% of foreign material in the ground roasted coffee and this investigation has been accomplished through techniques such as optical microscopy, chromatography, and infrared spectroscopy (ABIC 2013).

The near-infrared hyperspectral imaging is a promising technique which has been widely used for drug analysis and forensic chemistry in general, with advantage especially in the studies of heterogeneous samples due to the ability to provide chemical information relevant to each pixel in the mapped area, even in lower concentrations, without sample preparation (Amigo et al. 2008; Fu et al. 2014; Lim et al. 2016; Silva et al. 2014; Piqueras et al. 2012). This technique consists in obtaining a three-dimensional data array, also known as hypercube, containing two spatial information (from the selected area to be mapped) and one spectral (wavenumber). The spectrum in each pixel of the image contains the chemical information about the species present and this signal is not proportional to the whole concentration of the analyte on the mapped surface, but its abundance in the particular sampling point (Sun 2010; Gowen et al. 2007; Kim et al. 2011).

The massive quantity of data generated using hyperspectral imaging that depends of the number of scans, the size of the pixels, and the selected area, often prevents the use of traditional methods of processing and the use of chemometric tools becomes imperative.

The application of chemometric tools involves distinct situations, depending of the goals to be reached, as classification,

✉ Ronei J. Poppi
rjpoppi@unicamp.br

¹ Institute of Chemistry, University of Campinas (UNICAMP), 6154, Campinas, São Paulo, CP 13084-971, Brazil

Multivariate regression, or multivariate curve resolution. Multivariate curve resolution methods present the advantage of spectra recovered of the pure species participating in the composition of the sample without prior knowledge of the same (Parasta and Tauler 2014; de Juan and Tauler 2006; de Juan et al. 2004). These tools provide the development of a mathematical model able to correlate the spectra of the reference species participating in the mixture and their concentration, making possible the detection and quantification in an efficiently way (Rutan et al. 2009).

The principle of the MCR-ALS (multivariate curve resolution with alternating least squares) (Jaumont et al. 2005) is to resolve, subject to certain constraints, the concentration matrix **C** and the pure spectra matrix **S**. The array of initial data **D** is decomposed according to Eq. 1, where **E** is the error matrix relating to residual variation of the data.

$$\mathbf{D} = \mathbf{CS}^T + \mathbf{E} \quad (1)$$

The main condition for a suitable resolution is to begin the process with the correct number of chemical components of the mixture that, in case of first-order data, can be estimated by the percentage of variance explained by the singular value decomposition (SVD) or principal component analysis (PCA) (de Juan and Tauler 2006). In order to obtain more appropriate models, some constraints are selected regarding non-negativity of concentration and/or spectra, unimodality, mass closure, etc. (de Juan et al. 2009; de Juan et al. 2014; Tauler et al. 2009).

The correlation constraint is an advanced tool available in MCR-ALS and it implies a concentration correlation information for each constituent of the samples, established by the model compared to the reference values (de Juan et al. 2008). In this case, an augmented matrix (**D_{AUG}**) is built, including a set of calibration samples along with the initial array (Alexandrino et al. 2015). By using this constraint, the samples are forced to correlate with the concentration of values previously known, establishing a linear model. In hyperspectral imaging data, the concentration values obtained by the model for each pixel are used for creation of maps of distribution of concentration of the chemical species (Antunes et al. 2002; Jaumont et al. 2013).

In this work, we performed the detection and quantification of main coffee adulterants, such as coffee husks, roasted and powdered corn kernels, wood sticks, and soil, based on the combination of near-infrared hyperspectral imaging and multivariate curve resolution with alternating least squares (MCR-ALS). The detection was accomplished comparing the reference spectra of each adulterant with the spectra recovered by the MCR model using only the adulterated samples, and the quantification was based on the use of the MCR correlation constraint.

Another proposal for this work was to ensure the quality of the *roasted and ground* coffee through the development of a

control chart associated with the results of the MCR-ALS applied to hyperspectral imaging data. In this procedure, it is possible to check the purity of the coffee quickly and efficiently, as well as identify samples adulterated with coffee husks, roasted and powdered corn kernels, soil, and wood sticks, simultaneously, in concentrations above 1%. Although the term quality can present numerous definitions, according to experts, there is a consensus that it refers to the needs and expectations of consumers in relation to the product and are also associated with indirect factors such as price and presentation of the packaging, among others (Does et al. 1997; Caplen 1988). To better monitor the attributes related to the quality of a product, it is possible to establish a statistical control, assessing the possible variations in relation to the known standards (Dale 1994; Miller 1998). In this case, a chart control was developed, based on a center line and two control lines (Linna and Woodall 2001). Initially, it was defined a set of purity samples of certified coffees to be used for control. From MCR-ALS scores (**C** matrix) for each control sample, it was calculated the average of the values obtained related to coffee constituent. The acceptance of samples within the quality profile of purity was established by the lines of upper limits and lower limit corresponding to two times the standard deviation of the average of these score values.

Materials and Methods

Samples

In this work, Robusta coffee beans without adulterants certified by producers were used as reference samples. Corn kernels were purchased at local stores, while soil, coffee husks, and wood sticks were obtained in the region's crops. To make the adulterants with the same coffee appearance, they were placed separately in aluminum trays and heated in oven at 250 °C for 30 min, in a similar procedure already described in literature (Schmidt et al. 2008). Finally, the adulterants were ground up and added to the ground roasted coffee in concentrations that ranged from 1.0 to 40.0% (w/w). The mixtures were stirred in vortex and pressed at 6 tons for making tablets, ensuring a smooth surface.

Equipment

The spectra were obtained using a Spotlight 400 N FT-NIR Imaging System from Perkin-Elmer. The tablets were placed on the microscopy support, and with the help of a joystick, the focus was adjusted for each sample. The conditions for the analysis were 32 scans in the range of 4000 to 7800 cm⁻¹, pixels with 25 μm², and a monitored area of 4 mm², resulting in 160 × 160 pixels or 25,600 spectra per sample.

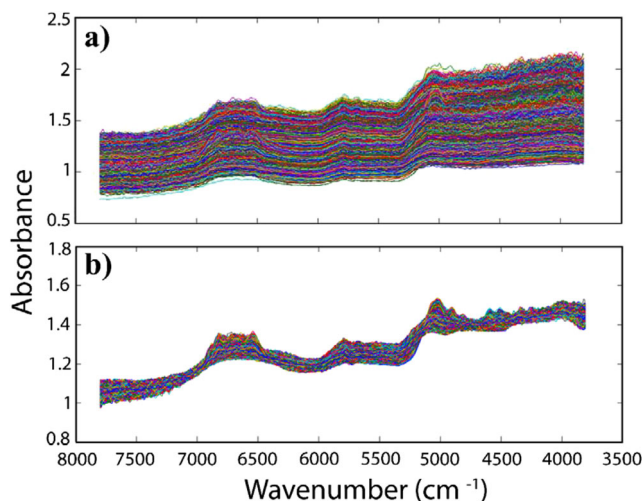


Fig. 1 Spectral profile of (a) a sample of the coffee adulterated and (b) after the pre-processing by MSC

The data were analyzed using the MCR Toolbox 2.0 for Matlab, freely available at <https://mcrals.wordpress.com/download/mcr-als-2-0-toolbox>.

Results and Discussion

Data Pre-processing

The first step in the development of a multivariate model is the appropriate choice of the pre-processing steps. In this work, the near-infrared spectra obtained in reflectance mode (R) were converted to pseudo-absorbance by a logarithmic function ($\log 1/R$), and subsequently was applied the

multiplicative scatter correction (MSC) to reduce the scattering effect, as shown in Fig. 1.

Detection and Quantification of the Adulterants by MCR-ALS

In a first step, a model was generated for each set of samples containing the coffee plus one adulterant, and the simplisma was used to obtain the coffee and adulterant spectra. In this situation, from the recovered spectrum by MCR, the adulterant can be identified, and the quantification can be performed using the calibration model generated by MCR-ALS with correlation constraint for this adulterant. This procedure is important to ensure the reliability that the generated concentration maps correspond to the species of interest (coffee and adulterant). The spectra recovered by the MCR-ALS algorithm using the simplisma for each adulterant, as well as the reference spectra of this adulterant, are presented in the Fig. 2. For each recovered spectrum by the MCR-ALS, the coefficient of correlation was determined in terms of similarity to the spectrum of the reference species. For the adulterants studied, these values ranged from 0.90 to 0.99. These values indicate that it was possible to identify all adulterants with great reliability, showing the potentiality of the proposed methodology.

The number of factors used in the MCR calculations was determined by SVD analysis, and it was equal to two for all samples.

In order to perform the quantification using the correlation constraint in the MCR-ALS algorithm, 5 samples were used for calibration at concentrations of 5.0%, 10.0%, 20.0%, 30.0%, and 40.0 % (w/w), for each adulterant (coffee husks, roasted and powdered corn kernels, wood sticks,

Fig. 2 Recovered spectra by Simplisma. (a) Coffee husks. (b) Roasted and powdered corn kernels. (c) Wood sticks. (d) Soil

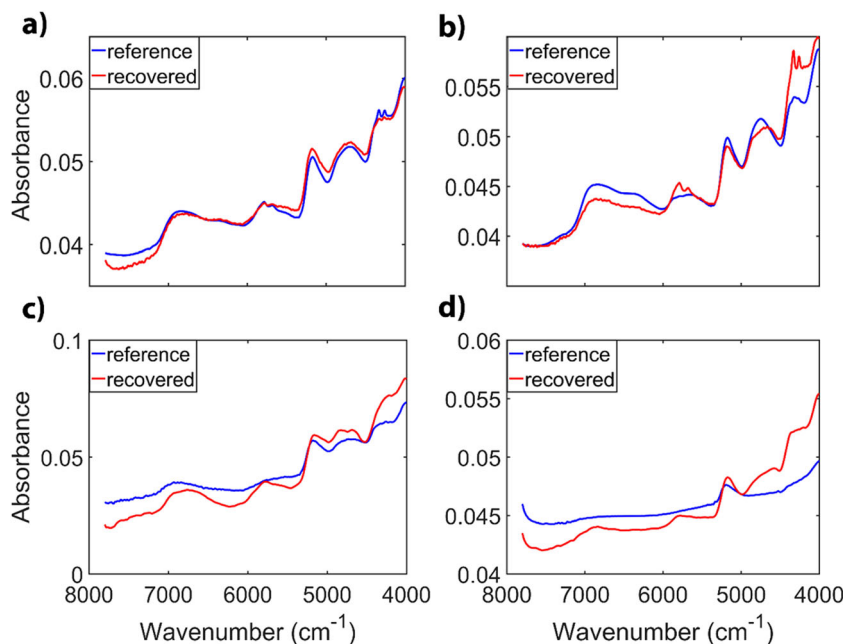
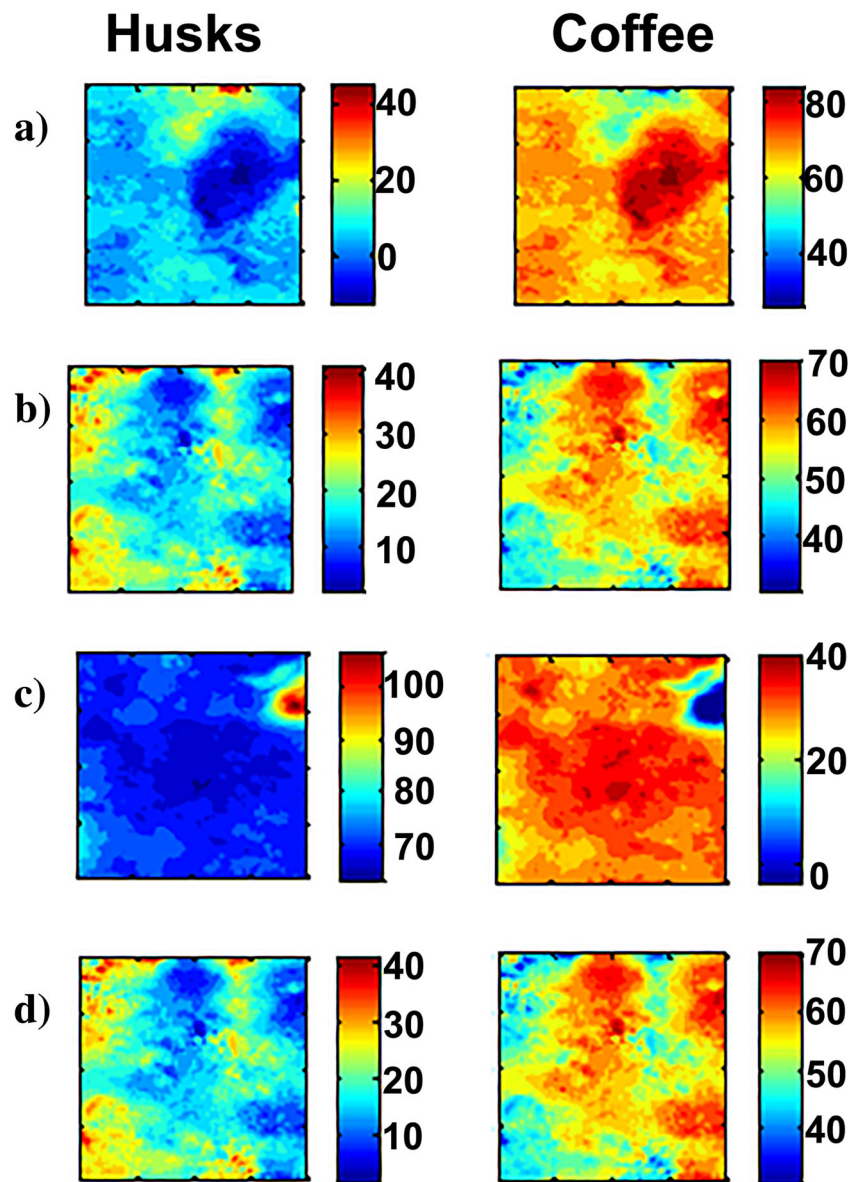


Fig. 3 Distribution concentration maps for coffee husk adulterant (area of 4 mm², resulting in 160 × 160 pixels). (a) 10% adulteration. (b) 15% adulteration. (c) 20% adulteration. (d) 30% adulteration



and soil) added to the coffee. In this situation, the matrix to be decomposed by the MCR-ALS was formed by the mean spectra of the calibration samples (used in the calibration step), together with the spectra of each pixel of the image. The concentration values of the calibration samples are fixed by the algorithm that sets these values in the concentration matrix.

Subsequently, from the MCR-ALS results for each adulterant, it was built the distribution concentration maps of the constituents of each sample and the corresponding histograms of adulterant concentration. In Fig. 3, it is presented, as an example, the distribution concentration maps for coffee adulterated with husks. In this figure, it is possible to verify in the concentration maps a complementarity of the images as to the regions of presence and/or absence of adulterants. The

heterogeneous distribution of the adulterant visualized in the chemical maps is an aspect that emphasizes the importance of the use of hyperspectral imaging technique to guarantee a better result regarding the quantification of the adulterant, since depending on the region selected for the analysis and/or the resolution used, one would not be guaranteed the presence of the adulterant.

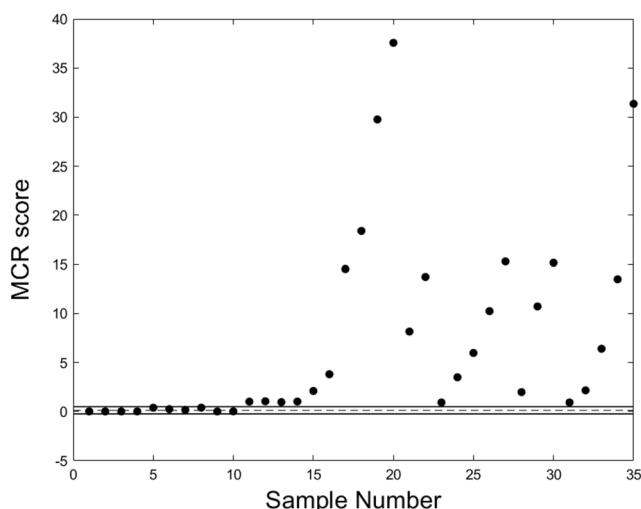
By calculating the concentration averages in all pixels of the image, it was possible to determine the concentration of the adulterant in the sample. Table 1 shows the expected and predicted values (with their absolute errors) for the studied adulterants. It is possible to verify that the absolute errors for quantification of the adulterants in coffee in the models developed were less than 4%, which indicated the feasibility in the application of the proposed methodology.

Table 1 Predicted values and errors for the adulterant quantification

Adulterant	Expected (%)	Predicted (%)	Absolute error (%)
Coffee husks	10.0	11.3	1.3
	15.0	15.3	0.3
	20.0	23.7	3.7
	25.0	26.8	1.8
	30.0	30.2	0.2
Roasted and powdered corn kernels	10.0	10.1	0.1
	15.0	14.8	0.2
	20.0	18.4	1.6
	30.0	29.8	0.2
Wood sticks	5.0	6.4	1.4
	15.0	15.8	0.8
	20.0	21.0	1.0
	25.0	23.6	1.4
	30.0	31.5	1.5
Soil	5.0	5.6	0.6
	10.0	10.7	0.7
	15.0	14.8	0.2
	20.0	22.6	2.6
	25.0	25.0	0.0

Control Charts

In order to guarantee the efficiency of the quality characterization of coffee samples using the NIR hyperspectral imaging technique and the curve resolution method, a control chart was developed, with a total of 35 samples containing different concentrations of all types of adulterants previously mentioned. The concentration score values estimated by the MCR-ALS were used for the preparation of the chart, without the use of any restriction and quantitative value. In this case, a single matrix was formed, with all

**Fig. 4** Control chart for coffee adulteration

images obtained, including 10 images without any adulterant to serve as the basis of the control charts.

The limits used in the proposed control chart were based on the standard procedure for control chart development. In this procedure, a target value should be defined that in this work was the average of the scores obtained from the images without adulteration. Also, two limits should be defined: warning (average plus or minus two standard deviations—95% of the samples inside this range) and action (average plus or minus three standard deviations—99.7% of the samples inside this range). In this work, we considered as out of control the samples with score values above the warning limit. Figure 4 shows the control chart developed for detection of the adulteration in coffee. In this figure, the first 10 samples correspond to the coffee without adulterants, whose average score was 0.12. The other samples contain at least 1% of some adulterant and therefore are outside the control lines with 95% confidence. This result indicates that a simple MCR-ALS model can be developed, using only the hyperspectral images, to be used in cases where it is desired to know if the sample is adulterated or not.

Conclusions

The use of NIR hyperspectral imaging combined with MCR-ALS algorithm for detection and quantification of various types of adulterants (coffee husks, roasted and powdered corn kernels, wood sticks, and soil) in coffee samples proved to be an effective technique, with errors smaller than 4% (*w/w*). An additional advantage is the identification of the adulterant present in the sample comparing the MCR recovered spectra with the reference spectra of the adulterants.

Control charts were used to analyze the quality of the samples relating the scores obtained by the MCR-ALS method to adulterated samples of coffee, using only the hyperspectral images. It was possible to identify coffee adulterated in levels above 1% in a practical and effective way.

Funding Information This study was funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (proc. 303994/2017-7) and by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES) (Finance Code 001).

Compliance with Ethical Standards

Conflict of Interest Debora de Andrade Penteadó Forchetti declares that he has no conflict of interest. Ronei Jesus Poppi declares that he has no conflict of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed Consent Not applicable

References

- ABIC (2013) Associação Brasileira das Indústrias de Café. Norma de Qualidade Recomendável e Boas Práticas de Fabricação de Cafês Torrados em Grãos e Cafês Torrados e Moídos. http://www.abic.com.br/publique/media/NMQ_LEGISLACAO_NORMAQUALIDADE.pdf, accessed in december 2018
- Alexandrino GL, Khorasani MR, Amigo JM, Rantenen J, Poppi RJ (2015) Monitoring of multiple solid-state transformations at tablet surfaces using multi-series near-infrared hyperspectral imaging and multivariate curve resolution. *Eur J Pharm Biopharm* 93:224–230
- Amigo JM, Cruz J, Bautista M, MasPOCH S, Coelho J, Blanco M (2008) Study of pharmaceutical samples by NIR chemical-image and multivariate analysis. *Trend Anal Chem* 27:696–713
- Antunes MC, Simão JEJ, Duarte AC, Tauler R (2002) Multivariate curve resolution of overlapping voltammetric peaks: quantitative analysis of binary and quaternary metal mixtures. *Analyst* 127:809–817
- Caplen R. H. (1988) A practical approach to quality control. Rev Ed edition, Random House Business Books, London
- Dale B (1994) Managing quality, 2nd edn. Prentice-Hall, London
- de Juan A, Tauler R (2006) Multivariate curve resolution (MCR) from 2000: progress in concepts and applications. *Crit Rev Anal Chem* 36:163–176
- de Juan A, Navea S, Diewok J, Tauler R (2004) Local rank exploratory analysis of evolving rank deficient systems. *Chemom Intell Lab Syst* 70:11–21
- de Juan A, Maeder M, Hancewicz T, Tauler R (2008) Use of local rank-based special information for resolution spectroscopic images. *J Chemom* 22:291–298
- de Juan A, Maeder M, Hancewicz T, Duponchel L, Tauler R (2009) Chemometric tools for image analysis. In: Salzer R, Siesler HW (eds) *infrared and raman spectroscopic imaging*. Wiley-VCH, Weinheim, pp 65–106
- de Juan A, Jaumot J, Tauler R (2014) Multivariate curve resolution (MCR) solving the mixture analysis problem. *Anal Methods* 6: 4964–4976
- Does RJMM, Schippers WAJ, Trip A (1997) A framework for implementation of statistical process control. *Int J Qual Sci* 2:3
- Fu X, Kim MS, Chao K, Quin J, Lim J, Lee H (2014) Detection of melamine in powders based on NIR hyperspectral imaging and spectral similarity analyses. *J Food Eng* 124:97–104
- Gowen AA, O'Donnell CP, Cullen PJ, Downey G, Frias JM (2007) Hyperspectral imaging—an emerging process analytical tool for food quality and safety control. *Trends Food Sci Technol* 18:590–598
- ICO (2014) International Coffee Organization (<http://www.ico.org>), accessed in december 2018.
- Jaumot J, Gargallo R, de Juan A, Tauler R (2005) An user friendly interface for MCR-ALS: a new tool for multivariate curve resolution in MATLAB. *Chemom Intell Lab Syst* 76:101–110
- Jaumot J, Igne B, Anderson CA, Drennen JK, de Juan A (2013) Blending process modeling and control by multivariate curve resolution. *Talanta* 117:492–504
- Kim MS, Chao K, Chan DE, Jun W, Lefcourt AM, Delwiche SR, Kang SW, Lee KJ (2011) Line-scan hyperspectral imaging platform for agro-food safety and quality evaluation: system enhancement and characterization. *Trans ASABE* 54:703–711
- Lim J, Kim G, Mo C, Kim MS, Chao K, Quin J (2016) Detection of melamine in milk powders using near-infrared hyperspectral imaging combined with regression coefficient of partial least square regression model. *Talanta* 151:183–191
- Linna KW, Woodall WH (2001) Effect of measurement error on Shewhart Control Chart. *J Qual Technol* 33:213–222
- Miller H (1998) Variation, innovation and dynamic quality. *TQM Mag* 10:447–451
- Parasta H, Tauler R (2014) Multivariate curve resolution of hyphenated and multidimensional chromatographic measurements: a new insight to address current chromatographic challenges. *Anal Chem* 86:286–297
- Pauli ED, Barbieri F, Garcia PS, Madeira TB, Acquaro Junior VRA, Scarminio IS, Scarminio CAP, Nixdorf SL (2014) Detection of ground roasted coffee adulteration with roasted soybean and wheat. *Food Res Int* 61:112–119
- Piqueras S, Burger J, Tauler R, de Juan A (2012) Relevant aspects of quantification and sample heterogeneity in hyperspectral image resolution. *Chemom Intell Lab Syst* 117:169–182
- Reis N, Franca AS, Oliveira LS (2013a) Discrimination between roasted coffee, roasted corn and coffee husks by diffuse reflectance infrared fourier transform spectroscopy. *LWT Food Sci Technol* 50:715–722
- Reis N, Franca AS, Oliveira LS (2013b) Performance of diffuse reflectance infrared Fourier transform spectroscopy and chemometrics for detection of multiple adulterants in roasted and ground coffee. *LWT Food Sci Technol* 53:395–401
- Rutan SC, de Juan A, Tauler R (2009) Introduction to multivariate curve resolution. In: Brown S, Tauler R, Walczak B (eds) *Comprehensive Chemometrics*, vol 2. Amsterdam, Elsevier, pp 249–259
- Sano EE, Assad ED, Cunha SAR, Correa TBS, Rodrigues HR (2003) Quantifying adulteration in roast coffee powders by digital image processing. *J Food Qual* 26:123–134
- Schmidt CAP, Miglioranza E, Prudêncio SH (2008) Interação da torra e moagem do café na preferência do consumidor do oeste paranaense. *Cienc Rural* 38:1111–1117
- Silva CS, Pimentel MF, Honorato RS, Pasquini C, Prats-Montalbán JM, Ferrer A (2014) Near infrared hyperspectral imaging for forensic analysis of document forgery. *Analyst* 139:5176–5184
- Sun D (2010) *Hyperspectral imaging for food quality analysis and control*. Academic Press / Elsevier, San Diego
- Tauler R, Maeder M, de Juan A (2009) Multiset data analysis: extended multivariate curve resolution. In: Tauler R, Walczak B (eds) Brown, S. *Comprehensive Chemometrics*, Amsterdam

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.