CrossMark

ORIGINAL PAPER

Reliable discriminant analysis tool for controlling the roast degree of coffee samples through chemical markers approach

Paulo R. A. B. de Toledo[1] · Marcelo M. R. de Melo[2] · Helena R. Pezza[1] ·
Leonardo Pezza[1] · Aline T. Toci[3] · Carlos M. Silva[2]

Abstract Roasting is one of the most influencing stages of coffee processing. Accordingly, a discriminant analysis (DA) was carried out with the objective of identifying key compounds (chemical markers) that enable a differentiation of coffee samples according to their roasting degree. For this, chromatographic data of the volatile fraction of 21 coffee samples submitted to distinct roasting treatments (Light, Medium, Dark, and French Roasts) were employed. Using three discriminant functions that rely on only ten chemical markers, it was possible to explain 100 % of the variance of the data points. If two functions are used, the surprisingly high value of 99.4 % is achieved. The model was cross-validated, and the main function successfully passed a permutation test using two statistical indicators. It was found that half of the markers belong to the pyrazines family, known to grant sensorial notes related to roasted hazelnut and peanuts. In the whole, this essay demonstrates the usefulness of DA as a tool to control the quality of roasting treatment of coffee and can be further extended with advantage to the eight roasting degrees of the AGTRON Roasting Classification as soon as larger databases become available.

Introduction

The research on coffee extractives is a pertinent subject that covers the progressive processing stages (raw, roasted, brewed, and spent) of this raw material [2, 22, 23, 26] and involves challenges such as the quality control of samples [4, 12, 14, 24] and valorization of residues [3, 6, 10].

One of the most influencing stages of coffee processing comprises the roasting of raw beans. While decisive, roasting is a highly complex process in which the coffee is exposed to temperatures up to 300 °C plus variable conditions of time, heating rate, etc. During this process, a series of physical and chemical modifications take place, leading to the formation of an amalgam of volatile compounds that are known to determine in great extent the final flavor of coffee drinks.

The reactions most responsible for the alterations suffered by coffee during roasting are Maillard, Strecker, hydrolysis, and pyrolysis reactions. The Maillard reaction occurs between a reducing sugar and an amino acid or a protein [29]. It has low activation energy and is thermodynamically favored under roasting conditions. In addition to the temperature and to the presence of reducing sugars, other variables that influence the reaction include pH, moisture, and the presence of certain metals [29]. While it is known that many volatile products are prone to arise from the Maillard reaction, such as pyridines, pyrazines, furans, and pyrroles, there is still a lack of elucidation on the governing mechanisms. Nevertheless, some of the carbonyl derivative products obtained from the Maillard reaction can also react with free amino acids, resulting in their

✉ Carlos M. Silva
  carlos.manuel@ua.pt

[1] Institute of Chemistry, State University of São Paulo – UNESP, Araraquara, SP 14800-060, Brazil

[2] CICECO – Aveiro Institute of Materials, Department of Chemistry, University of Aveiro, Aveiro 3810-193, Portugal

[3] Latin American Institute of Science of Life and Nature, Federal University of Latin American Integration – UNILA, Foz do Iguaçú, PR 85867-970, Brazil

degradation to aldehydes, ammonia, and carbon dioxide. This is known as the Strecker degradation [5, 21].

Various roasting procedures for raw coffee are known, for instance "high temperature and short time" (HTST) and "low temperature and long time" (LTLT) [1, 8, 15]. The final degree of roast can be monitored by measuring reflectance, visual inspection of color, or even by the loss of mass. According to the colorimetric system of the Specialty Coffee Association of America, known as "AGTRON Roasting Classification," the degree of roast can be systematized in eight categories, as follows: Very Light, Light (American Roast), Moderately Light, Light Medium, Medium, Moderately Dark, Dark, and Very Dark (French Roast).

The influence of roasting on the volatiles composition and on the final quality of the coffee has been investigated and reported in various works [9, 15, 16, 18, 20, 23, 25]. Accordingly, several classes of compounds have been associated with different roasting degrees, being examples of it the 2,3-pentanedione, 2,5-dimethylpyrazine, 2-ethylpyrazine, 2,3-dimethylpyrazine, 3-ethyl-2,5-dimethylpyrazine, 2-methylbutanal, 2-ethylguaiacol, and 4-vinylguaiacol. In addition, it was found that a Light Roast (also known as American Roast due to the wide use in the USA) leads to volatile compounds with strong impacts on the aroma, such as furfural (sweet almond notes) and 2-methylbutanal (buttery notes) [9, 15, 18]. Furthermore, the most volatile compounds (e.g., volatile acids and furans) become less important as the coffee roast evolves from Light to Medium, probably due to losses as the roast temperature is increased or after exposure to high temperatures for longer periods. At the same time, increments in less volatile compounds, such as pyrazines and pyridine, are observed [15, 23]. Moreover, Toci et al. [25] investigated recently the effect of the roasting rate on three *C. arabica* mixtures. Seven chemical compounds exhibited a correlation with the studied roasting rates, namely 2,5-dimethylpyrazine, 2-ethylpyrazine, 2,3-dimethylpyrazine, 2,5-dimethyl-3-ethylpyrazine, guaiacol, 2-ethylguaiacol, and 4-vinylguaiacol. Using a semi-fluidized bed roaster, which allows a better control of temperature and time parameters, the roasting rate was showed to affect pyrazines, furans, pyrroles, and phenolics chemical classes, through an increase in their content, specially on molecules like 2,5-dimethylpyrazine, 2-ethylpyrazine, guaiacol, and 4-ethylguaiacol.

Considering that the cup tests currently used to classify coffee beverages may be inherently subjective, complementary methods of assessment are sought by researchers. While at laboratory scale a control of mass loss is usually possible, neither color-based nor weight loss procedures have proven to be effective in practice [9, 27].

This work focuses on the application of the multivariate statistical technique known as discriminant analysis to the volatiles composition of coffee samples submitted to different roasting degrees. The goal is to identify chemical markers capable of labeling their roasting category and combine them in a global model for subsequent food control applications.

The article is structured as follows: "Modeling" section is devoted to modeling; "Materials and methods" section comprises the database information, namely "Coffee samples" and "Chemical characterization" sections; results and discussion are presented in "Results and discussion" section, in the following sequence: "Preliminary normalization of results" and "Discrimination of samples according to roasting degrees"; and finally, the main conclusions are drawn in "Conclusions" section.

Modeling

Once this work aims to identify compositional parameters prone to be used as predictors of roasting degrees, discriminant analysis (DA) was chosen for screening appropriate chemical markers (coffee volatiles) from a database encompassing chromatographic characterization data.

Briefly, DA belongs to the class of multivariate statistics and is able to reduce data redundancy [13] through the fitting of first-degree polynomial functions that sort the distribution of each data point within a desired grouping (classes). Examples of research studies on coffee samples that have employed DA technique were published by Powers and Keith [19], Murota [17], Maeztu et al. [11], and de Toledo et al. [7].

From a mathematical point of view, DA procedure is such that key independent variables are chosen as predictors and combined linearly to ensure a maximized capacity of prediction of class memberships, as follows:

$$Y = \beta_0 + \beta_1 V_1 + \beta_2 V_2 + \beta_3 V_3 + \cdots + \beta_n V_n \tag{1}$$

where Y is the discriminant function, $\beta_0, \beta_1, \ldots, \beta_n$ are the linear discriminant coefficients, and $V_1, V_2, \ldots, V_n$ are the corresponding abundance ratios (i.e., normalized concentrations) of the volatiles. The DA statistical modeling was performed using the software SPSSv.23 and MATLAB v.7.8.0. For this, a data set comprising chromatograms of 21 coffee samples submitted to four distinct roasting degrees was employed, which implied extracting information from two publications (see Table 1).

Beforehand the fitting of the DA model, a preliminary study regarding the pretreatment of data was accomplished to enable the utilization of chromatographic data from different

Table 1 Summary of *C. Arabica* samples database compiled for the discriminant analysis

Sample	Origin	Roasting degree	Analytical technique	No. of compounds	References
1	Ethiopia	Light	GC–MS	53	Moon and Shibamoto [15]
2		Medium			
3		Dark			
4		French			
5	Nicaragua	Light			
6		Medium			
7		Dark			
8		French			
9	Indonesia	Light			
10		Medium			
11		Dark			
12		French			
13	Ethiopia	Light	HS-SPME/GC–MS	80	Akiyama et al. [1]
14		Medium			
15		Dark			
16	Tanzania	Light			
17		Medium			
18		Dark			
19	Guatemala	Light			
20		Medium			
21		Dark			

articles. Hence, the normalization of the individual peaks using several compounds appearing in the chromatograms was tested with the objective to disclose the molecule that could lead to the best discrimination results and functions. In this effort, the following molecules were evaluated: 2-hydroxy-3-methyl-2-cyclopenten-1-one, 2-methylpirrole, 2,5-dimethylpyrazine, 2-ethylpyrazine, 2- methylpyrazine, and pyridine.

The fitting of the DA model was achieved with and without cross-validation, with the latter taking place by rotating coffee samples between the training and validation subsets, to check the reliability and dependence of the DA model on the training dataset. Afterward, the main function of the DA model was tested against the chance of overfitting and also regarding its classification performance. For this, a permutation test was also accomplished [28], where the original data were permuted 10000 times, and classification errors were assessed in terms of both the number of misclassification (NoM) and the individual prediction error (Q^2) defined by:

$$Q^2 = 1 - \frac{\sum_i \left(y_i - \widehat{y_i}\right)^2}{\sum_i \left(y_i - \overline{y}\right)^2} \tag{2}$$

where y_i is the discriminant function score of the sample i, $\widehat{y_i}$ is the predicted value of class membership for sample i, and $\overline{y}$ is the mean value of all samples. In summary, the closer the value of Q^2 is from 1, the better the class prediction will be in relation to the class label expectations.

Materials and methods

Coffee samples

The six *C. arabica* samples that were picked to build the roasted coffee database are presented in Table 1. These belong to countries like Ethiopia, Nicaragua, and Indonesia [15], and Ethiopia, Tanzania, and Guatemala [1].

The samples from Moon and Shibamoto [15] were roasted with a Gene Cafe coffee bean roaster (Fresh Beans Inc., Park City, UT) at 230 °C for 12 min (Light), at 240 °C for 14 min (Medium), at 250 °C for 17 min (City, i.e., Dark), or at 250 °C for 21 min (French or Very Dark). After roasting, the coffee beans were ground with a Starbucks Barista coffee grinder (Seattle, WA). The remaining samples [1] were roasted to three roasting degrees: Light (L value 26), Medium (L value 23), and Dark (L value 18), using a Probat (Emmerich, Germany) G-12 roaster. The degree of roasting was represented as an L value, which was determined by measuring ground-roasted coffee (particle size < 500 μm) using a color meter ZE-2000 (Nippon Denshoku Industries Co. Ltd., Tokyo, Japan).

Table 2 Compound suitability results for usage as reference for database normalization

Reference compound under investigation	% of variance explained by discriminant functions			Correct classification (%)
	Y_1 (%)	Y_2 (%)	$Y_1 + Y_2$ (%)	
2-Hydroxy-3-methyl-2-cyclopenten-1-one	59.8	26.9	86.7	76.2
2-Methylpirrole	91.5	6.0	97.5	85.7
2,5-Dimethylpyrazine	73.4	24.0	97.4	100
2-Ethylpyrazine	88.2	10.1	98.3	100
2-Methylpyrazine	93.4	5.2	98.6	100
Pyridine	97.9	1.5	99.4	100

For this DA study, the aforementioned roasted coffee samples were grouped in four classes (shown in Table 1): Light, Medium, Dark, and French (Very Dark) roasting degrees.

Chemical characterization

The volatiles of the coffee samples studied in this essay were obtained by two distinct ways: headspace (HS) solid-phase microextraction (SPME) in the case of the work of Akiyama et al. [1], and Soxhlet extraction with dichloromethane according to Moon and Shibamoto [15]. Despite the different techniques, in both cases the extracts were analyzed by gas chromatography—mass spectrometry (GC–MS).

Moon and Shibamoto [15] used an Agilent model 6890 GC, equipped with a 60-m and 0.25-mm i.d. (film thickness of 0.5 μm) DB-Wax bonded-phase fused silica capillary column (Agilent, Folsom, CA), interfaced to an Agilent 5971A mass selective detector (MS) used for mass spectral identification of the GC components at MS ionization voltage of 70 eV. In turn, the GC used by Akiyama et al. [1] was a Hewlett-Packard (HP) 5973 (Agilent Technologies, Palo Alto, Calif., USA) with a fused silica capillary column DB-Wax (60 m × 0.25 mm, 0.25 μm film thickness, Agilent Technologies).

Results and discussion

Preliminary normalization of results

The first stage of our approach concerned treating the analytical data in order to guarantee a valid comparison basis between the results reported by different authors in different studies. While it is known that the abundance of volatiles is related to the roasting degree attained by the raw coffee samples, the strategy to ensure this comparison was to find the best compound to normalize the chromatograms.

The six candidates to reference compound (2-hydroxy-3-methyl-2-cyclopenten-1-one, 2-methylpirrole, 2,5-dimethylpyrazine, 2-ethylpyrazine, 2-methylpyrazine, and pyridine)

produced the results presented in Table 2, under the form of the following performance indicators: % of variance explained by the first two discriminant functions (Y_1, Y_2, and $Y_1 + Y_2$), and % of correct classification attained. The ideal compound would be one that allows the full explanation of the variance exhibited by data based on only one function, while ensuring also a 100 % correct classification.

In this context, 2-hydroxy-3-methyl-2-cyclopenten-1-one and 2-methylpirrole are two compounds that clearly did not fulfill the expectations: from Table 2, the former led to two functions that are able to explain only 86.7 % of data variance, and both fail the correct classification criterion of the roasting class labels (76.2 and 85.7 %, respectively). On the other hand, the bottom four compounds of Table 2 (i.e., 2,5-dimethylpyrazine, 2-ethylpyrazine, 2-methylpyrazine, and pyridine) enabled an explanation of more than 97 % of data variance with $Y_1 + Y_2$ and reached the maximum score in the correct classification (100 %). In fact, they only differ in the importance of the main discriminant function (Y_1) to allow such variance explanation. Accordingly, pyridine is the reference compound that scored best in both tests: 97.9 % in variance explanation through only Y_1 and 100 % of correct classifications. When jointly considered, no other compound exhibited such a good performance. For this reason, pyridine was chosen as the reference molecule for the discriminant analysis on roasting degrees.

Discrimination of samples according to roasting degrees

In order to explain 100 % of the variance of the coffee samples that were object of our statistical modeling, three discriminant functions were found necessary ($Y_1 + Y_2 + Y_3$). The resulting functions are provided in Table 3, where it can be noticed that two of them are able to cover up to 99.4 % of the variance through a bidimensional assignment of the data. For this reason, the third function was considered negligible for the goal of the differentiation of the coffee samples. In fact, if only Y_1 is used, the substantial 97.4 % of the variance of the data is explained with a correlation of experimental data of 0.995.

Table 3 Performance of the discriminant functions and respective data variance explanation

Function	Variance explanation (%)	Cumulative variance explanation (%)	Canonical correlation
Y_1	97.9	97.9	0.995
Y_2	1.5	99.4	0.787
Y_3	0.6	100.0	0.608

The reference compound for chromatograms normalization is pyridine

Table 4 Non-standardized coefficients (β_i) of the two canonical discriminant functions for the roasting degree of coffee samples

Compounds (Vi)	Y_1	Y_2
Constant	−10.305	−4.664
Pyrazine	1.313	108.167
2-Methylpyrazine	−16.195	2.492
2,5-Dimethylpyrazine	19.100	5.613
2-Ethylpyrazine	24.526	−24.077
2,3-Dimethylpyrazine	−53.733	−37.723
Dihydro-2-methyl-3(2H)-furanone	33.981	−1.656
5-Methyl-2(5H)-furanone	1.900	1.287
Furfural	−1.241	1.291
2,3-Dimethyl-2-cyclopenten-1-one	8.586	106.019
Acetic acid	1.586	1.708

The reference compound for chromatograms normalization is pyridine

The fitted discriminant functions rely on a total of ten chemical compounds (see Table 4) from the full list of 53 volatile organic compounds (VOCs) detected/identified via GC–MS by Moon and Shibamoto [15] and 80 VOCs detected/identified via HS-SPME/GC–MS by Akiyama et al. [1] (see Table 1). The respective coefficients are also provided for each function. It is worth noting that half of the model compounds belong to the pyrazines family, comprising a set of molecules known to grant sensorial notes related to roasted hazelnut and peanuts. Moreover, 2-methylpirazyne and 2,3-dimethylpirazyne are highly linked to intense roasting such as Dark and French classes, while dihydro-2-methyl-3(2H)-furanone is more linked to light(er) roasting degree.

The graphical representation of the roasting discriminant functions is furnished in Fig. 1. For each roasting group, the data points are graphed in their individual coordinates connected by a line to the respective group centroid (filled marks). To emphasize the areas belonging to the different groups, colored circles were drawn around each roasting class. It becomes clear from this figure that the roasting classes are ranked by Y_1 function in an inverse way to the
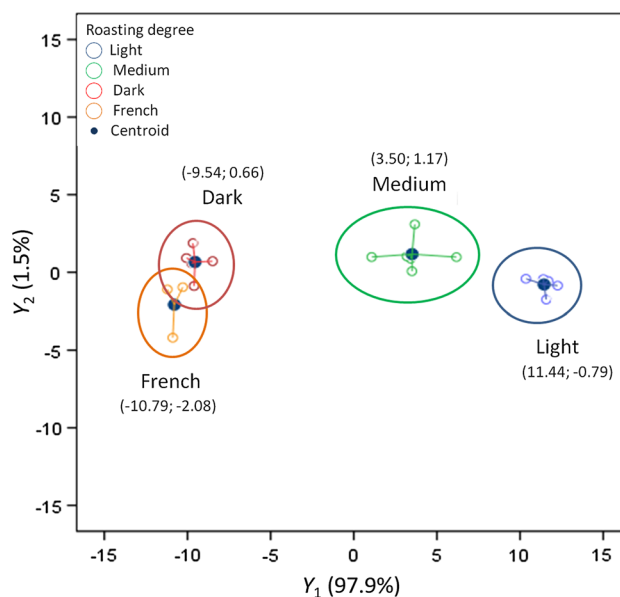
Fig. 1 Discriminant scores and centroid values of the DA modeling of four coffee roasting degrees

respective roasting intensity: as Y_1 increases, the French, Dark, Medium, and Light classes are sequentially reached. However, no clear hierarchy is found in Y_2.

As noticeable in Fig. 1, Dark and French classes exhibit partial overlapping, with a possible misleading in the distinction of these two methods. While distinct in terms of processing, the DA results evidence that Dark and French Roasts are very close to each other from the perspective of chemical markers. On the other hand, the results also point out that Medium class is closer to Light class than to Dark class, which highlights the existence of a gap between Dark and Medium Roasts.

It is also worthwhile to stress that the dispersion of data points inside each roasting group may be due to two specific factors beyond the experimental errors: (1) the different geographic provenance of the samples (e.g., places like Indonesia or Nicaragua) and (2) the distinct roasting methods adopted, which can impart fluctuations on the profiles of volatiles compounds. Nevertheless, such intragroup dispersions are of small magnitude in comparison with the variations owed to changes in the roast degrees (between groups), which demonstrates the prevailing influence of the roasting stage on the values of the chemical markers.

With reference to the quality of the proposed DA model, an assessment of the classification performance is provided in Table 5, where non-cross-validated results are contrasted with predictions when cross-validation (CV) is accomplished using the training data. Without cross–validation, the attained model is able to predict correctly all the class labels of coffee samples, which is an important initial achievement. Even under cross-validation, the DA model

was able to predict correctly the samples with French Roast. Also for the Dark class, only one coffee sample out of six was wrongly assigned. In the case of Light Roast, the successful counts dropped to 50 %, with the worst results being obtained for the Medium class, whose correct classifications fell to 16.7 %. These results suggest that future database upgrades should privilege the addition of coffee samples from Medium and Light roasting degrees, though the potential of the DA approach was demonstrated, mainly if one takes into account also the next permutation tests analysis.

The evaluation of the statistical significance of the DA model was accomplished through the realization of a permutation test to Y_1, the main function of the model. The results of this assessment are plotted in Fig. 2, under the

form of frequency histograms of the individual prediction error (Q^2, Fig. 2a) and the number of misclassifications (NoM, Fig. 2b). Upon calculation of the 95 % confidence bounds for the permuted data-based models, the statistical significance of the original Y_1 function was demonstrated. These threshold bounds are represented by vertical dashed lines on the right side of the histogram for Q^2 (Fig. 2a) and on the left side for NoM response (Fig. 2b). The function Y_1 led to scores that clearly surpass the threshold p values, i.e., higher Q^2 prediction and lower NoM as it is desirable for a good discrimination performance. As a result, the permutation test allows one to conclude that the decisive function of the DA model (Y_1) does not lead to good performances at the expenses of pure chance and/or overfitting problems.

Table 5 Overall classification performance (count and %) of the DA model, with and without cross-validation (CV)

Original class	Predicted class (DA model)							
	Light roast		Medium roast		Dark roast		French roast	
	Without CV	With CV	Without CV	With CV	Without CV	With CV	Without CV	With CV
Light	6 (100 %)	3 (50.0 %)	0 (0 %)	3 (50.0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
Medium	0 (0 %)	2 (33.3 %)	6 (100 %)	1 (16.7 %)	0 (0 %)	1 (16.7 %)	0 (0 %)	0 (0 %)
Dark	0 (0 %)	0 (0 %)	0 (0 %)	2 (33.3 %)	6 (100 %)	5 (83.3 %)	0 (0 %)	0 (0 %)
French	0 (0 %)	1 (16.7 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	3 (100 %)	3 (100 %)

The reference compound for the chromatograms normalization is pyridine
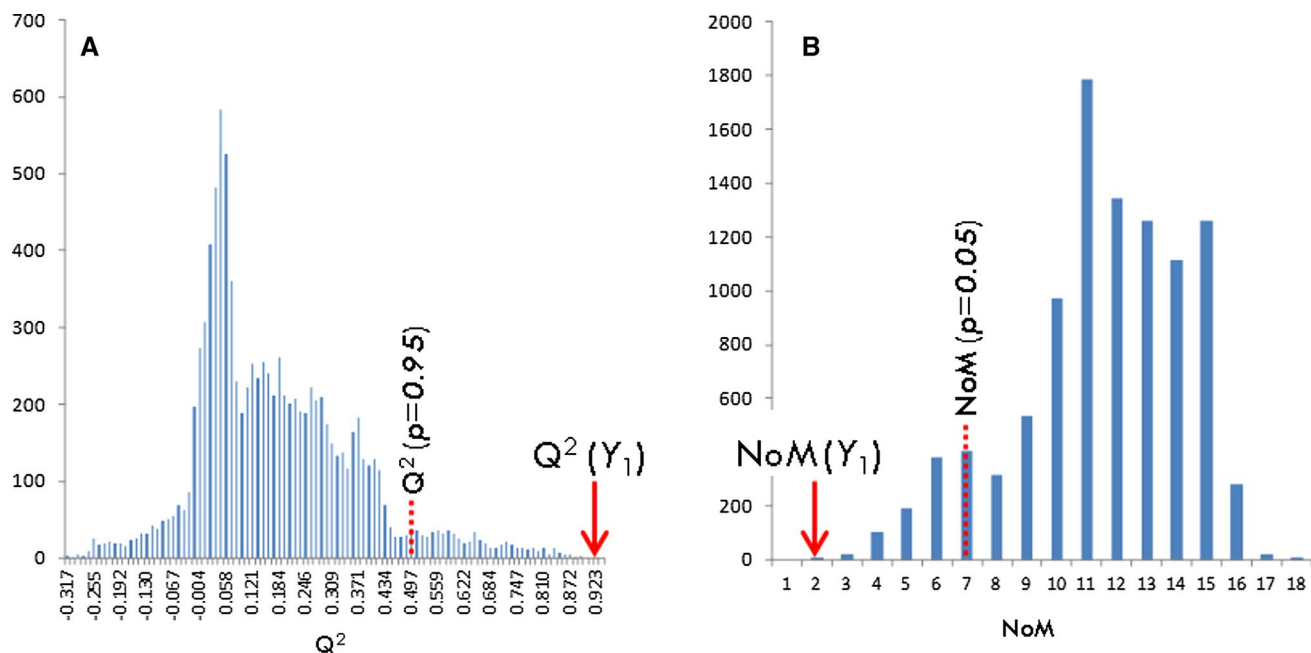
Fig. 2 Permutation test results of a total of 10,000 permutations of the original data: **a** individual prediction error (Q^2) and **b** number of misclassifications (NoM)

In conclusion, this essay reveals that DA can be an effective method for differentiating the roasting of coffee samples, which demonstrates its usefulness as potential quality control tool for coffee.

Conclusions

In this work, 21 coffee samples with different roasting degrees were studied by discriminant analysis (DA), through the modeling of chromatographic results obtained by HS-SPME-GC/MS and/or GC–MS. The goal was to attain a differentiation of these samples through discriminant chemical markers.

A previous normalization of the chromatograms was accomplished, where the peak area of pyridine is proposed as reference. Using three discriminant functions relying on only 10 chemical markers, it was possible to fully explain the variance of the experimental data. Half of these markers belong to the pyrazines family, which are known to grant sensorial notes related to roasted hazelnut and peanuts. Moreover, the main function of the model (Y_1), which explains 97.5 % of the data variance, establishes a direct relation with the corresponding roast degree; namely, the higher the global score of the function, the softer the roast degree.

The main function passed a permutation test aiming at a complementary statistical validation of the DA model, showing that it does not suffer from overfitting. Moreover, the proposed model was also cross-validated, being concluded that it is robust in the classification of Dark and French classes, but sensitive to coffee samples of Medium and Light Roast classes.

In the whole, the study demonstrates the usefulness of DA as a tool for the quality control of roasting treatment of coffee samples. In the future, the proposed DA tool can be extended with advantage to the differentiation of the eight roasting degrees of the AGTRON Roasting Classification, as soon as a large database is available.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Compliance with ethics requirements This article does not contain any studies with human or animal subjects.

References

1. Akiyama M, Murakami K, Hirano Y, Ikeda M, Iwatsuki K, Wada A, Tokuno K, Onishi M, Iwabuchi H (2008) Characterization of headspace aroma compounds of freshly brewed arabica coffees and studies on a characteristic aroma compound of Ethiopian coffee. J Food Sci 73(5):335–346
2. Barbosa HMA, de Melo MMR, Coimbra MA, Passos CP, Silva CM (2014) Optimization of the supercritical fluid coextraction of oil and diterpenes from spent coffee grounds using experimental design and response surface methodology. J Supercrit Fluids 85:165–172
3. Campos-Vega R, Loarca-Piña G, Vergara-Castañeda HA, Oomah BD (2015) Spent coffee grounds: a review on current research and future prospects. Trends Food Sci Technol 45:24–36
4. Costa LL, Toci AT, Silveira CLP, Herszkowicz N, Pinto M, Farah A (2010) Discrimination of Brazilian *C. Canephora* by region using mineral composition. In: Proceedings of the 23rd International Coll on the Chemistry of Coffee, Bali
5. Damoradam S (2007) Amino acids, peptides and proteins. In: Fennema OR (ed) Food chemistry. Marcel Dekker, New York, pp 412–416
6. de Melo MMR, Barbosa HMA, Passos CP, Silva CM (2014) Supercritical fluid extraction of spent coffee grounds: measurement of extraction curves, oil characterization and economic analysis. J Supercrit Fluids 86:150–159
7. de Toledo PRAB, de Melo MMR, Pezza HR, Toci AT, Pezza L, Silva CM (2016) Discriminant analysis for unveiling the origin of roasted coffee samples: a tool for quality control of coffee related products. Food Control. doi:10.1016/j.foodcont.2016.08.001
8. Eggers R (2005) Roasting techniques. In: Illy A, Viani R (eds) Espresso coffee: the science of quality, 2nd edn. Elsevier Academic Press, London, pp 184–191
9. Franca AS, Oliveira LS, Oliveira RCS, Agresti PCM, Augusti R (2009) A preliminary evaluation of the effect of processing temperature on coffee roasting degree assessment. J Food Eng 92:345–352
10. Low JH, Rahman WAWA, Jamaluddin J (2015) The influence of extraction parameters on spent coffee grounds as a renewable tannin resource. J Clean Prod 101:222–228
11. Maeztu L, Sanz C, Andueza S, Paz De Peña M, Bello J, Cid C (2001) Characterization of espresso coffee aroma by static headspace GC-MS and sensory flavor profile. J Agric Food Chem 49:5437–5444
12. Agresti PCM, Franca AS, Oliveira LS, Augusti R (2008) Discrimination between defective and non-defective Brazilian coffee beans by their volatile profile. Food Chem 106:787–796
13. McLachlan G (2004) Discriminant analysis and statistical pattern recognition, vol 544. Wiley, Chicago, pp 168–211
14. Moeenfard M, Silva JA, Borges N, Santos A, Alves A (2014) Diterpenes in espresso coffee: impact of preparation parameters. Eur Food Res Technol 240:763–773
15. Moon JK, Shibamoto T (2009) Role of roasting conditions in the profile of volatile flavor chemicals formed from coffee beans. J Agric Food Chem 57:5823–5831
16. Morais SAL, Aquino FJT, Chang R, Nascimento EA, Oliveira GA, Santos NC (2007) Chemical analysis of Arabica coffee

(*Coffea arabica* L.) and defective beans submitted to different degrees of roasting. Coffee Sci 2:97–111

17. Murota A (1993) Canonical discriminant analysis applied to the headspace GC profiles of coffee cultivars. Biosci Biotechnol Biochem 57:1043–1048

18. Nascimento EAD, Morais SALD, Rocha RS (2003) Constituintes voláteis de cafés "gourmet" e mole do cerrado do triângulo mineiro em função da torra. Ciência e Tecnologia de Alimentos 23:282–284

19. Powers JJ, Keith ES (1968) Stepwise discriminant analysis of gas chromatographic data as an aid in classifying the flavor quality of foods. J Food Sci 33:207–213

20. Ribeiro JS, Teófilo RF, Salva TDJG, Augusto F, Ferreira MMC (2013) Exploratory and discriminative studies of commercial processed Brazilian coffees with different degrees of roasting and decaffeinated. Braz J Food Technol 16:198–206

21. Rizzi GP (2006) Formation of strecker aldehydes from polyphenol-derived quinones and alpha-amino acids in a nonenzymic model system. J Agric Food Chem 54:1893–1897

22. Salinas-Vargas ME, Cañizares-Macías MP (2014) On-line solid-phase extraction using a C_{18} minicolumn coupled to a flow injection system for determination of caffeine in green and roasted coffee beans. Food Chem 147:182–188

23. Somporn C, Kamtuo A, Theerakulpisut P, Siriamornpun S (2011) Effects of roasting degree on radical scavenging activity, phenolics and volatile compounds of Arabica coffee beans (*Coffea arabica* L. cv. Catimor). Int J Food Sci Technol 46:2287–2296

24. Toci AT, Farah A (2014) Volatile fingerprint of Brazilian defective coffee seeds: corroboration of potential marker compounds and identification of new low quality indicators. Food Chem 153:298–314

25. Toci AT, Silva CM, Fernandes F (2014) Effect of roasting speed on the volatile composition of different quality coffee blends roasted in an industrial semi-fluidized bed roaster and in a small scale fluidized bed roaster. In: 25nd international conference on coffee science

26. Tsukui A, Santos Júnior HM, Oigman SS, de Souza ROMA, Bizzo HR, Rezende CM (2014) Microwave-assisted extraction of green coffee oil and quantification of diterpenes by HPLC. Food Chem 164:266–271

27. Vasconcelos ALS, Franca AS, Glória MBA, Mendonça JCF (2007) A comparative study of chemical attributes and levels of amines in defective green and roasted coffee beans. Food Chem 101:26–32

28. Westerhuis JA, Hoefsloot HCJ, Smit S, Vis DJ, Smilde AK, van Velzen EJJ, van Duijnhoven JPM, van Dorsten FA (2008) Assessment of PLSDA cross validation. Metabolomics 4:81–89

29. Whistler RL, BeMiller JN (1997) Carbohydrate chemistry for food scientists. Marcel Dekker, New York, pp 171–174