

Exploratory and discriminative studies of commercial processed Brazilian coffees with different degrees of roasting and decaffeinated

Estudos exploratórios e discriminantes de cafés brasileiros processados com diferentes graus de torra e descafeinados

Autores | Authors

✉ Juliano Souza RIBEIRO

Federal Institute of Espirito Santo (IFES)
Av. Ministro Salgado Filho, s/n, Soteco
CEP: 29106-010
Vila Velha/ES - Brazil
e-mail: julianoribeiro@ifes.edu.br

Reinaldo Francisco TEÓFILO

Federal University of Viçosa (UFV)
Departament of Chemistry
Viçosa/MG - Brazil
e-mail: reinaldo.teofilo@ufv.br

Terezinha de Jesus Garcia SALVA

Agronomic Institute of Campinas (IAC)
Coffee Centre
Campinas/SP - Brazil
e-mail: tsalva@iac.sp.gov.br

Fabio AUGUSTO

University of Campinas (UNICAMP)
Chemistry Institute
Department of Analytical Chemistry
Gas Chromatography Laboratory
Campinas/SP - Brazil
e-mail: augusto@iqm.unicamp.br

Marcia Miguel Castro FERREIRA

University of Campinas (UNICAMP)
Chemistry Institute
Department of Physical Chemistry
Theoretical and Applied Chemometrics
Laboratory
Campinas/SP - Brazil
e-mail: marcia@iqm.unicamp.br

✉ Autor Correspondente | Corresponding Author

Recebido | Received: 31/07/2012
Aprovado | Approved: 12/07/2013
Publicado | Published: set./2013

Summary

The fingerprints of the volatile compounds of 21 commercial Brazilian coffee samples submitted to different industrial processing *i.e.* decaffeinated or different roasting degrees (traditional and dark) were studied. The volatiles were collected by headspace solid phase microextraction (HS-SPME) and analyzed by GC-FID and GC-MS. The chromatographic data matrices (fingerprints) obtained were explored by the principal component analysis (PCA) and partial least squares – discriminative analysis (PLS-DA). Initially the chromatographic profiles were aligned by the algorithm correlation optimized warping (COW). The PCA showed the discrimination of the decaffeinated coffees from the others with both the SPME fibres used. This separation probably occurred due to the loss of some volatile precursors during the decaffeination process, such as sucrose. For both the fibres tested, PDMS/DVB and CX / PDMS SPME, the PLS-DA models correctly classified 100% of the samples according to their roasting degree: (medium and dark), the main differences being the concentrations of some of the volatile compounds such as 2-methyl furan, 2-methylbutanal, 2,3-pentanedione, pyrazine, 2-carboxyaldehyde pyrrole, furfural and 2-furanmethanol.

Key words: *Processed coffee; Solid Phase Microextraction (SPME); Correlation optimized warping; Principal component analysis; Partial least squares – discriminative analysis.*

Resumo

Neste trabalho, foram estudados os perfis cromatográficos dos compostos voláteis de 21 amostras de cafés comerciais brasileiros submetidos a diferentes processos industriais, *i.e.*, descafeinados e com diferentes graus de torra (médio e extraforte). Os voláteis foram coletados por microextração em fase sólida (HS-SPME) e analisados por GC-FID e GC-MS. As matrizes de dados cromatográficos (perfis) obtidas, a partir dessas análises, foram exploradas por análise de componentes principais (PCA) e análise discriminante por quadrados mínimos parciais (PLS-DA). Inicialmente, os dados cromatográficos foram alinhados utilizando-se o algoritmo *correlation optimized warping* (COW). A PCA demonstrou a diferenciação dos cafés descafeinados em relação às outras classes de café, com ambas as fibras de SPME utilizadas. Essa separação, provavelmente, ocorreu em razão do desaparecimento de algumas espécies precursoras de voláteis durante o processo de descafeinação, tal como sacarose. Os modelos PLS-DA, para ambas as fibras testadas, PDMS/DVB e CX/PDMS, classificaram corretamente 100% das amostras, de acordo com o grau de torra (médio e extraforte); a principal diferença entre as classes foi a diferença de concentração de alguns compostos voláteis, tais como 2-metilfurano, 2-metilbutanal, 2,3-pentanodiona, pirazina, 2-pirrol carboxaldeído, furfural e 2-furan metanol.

Palavras-chave: *Cafés processados; Microextração em fase sólida (SPME); Correlation optimized warping (COW); Análise por componentes principais; Análise discriminante por quadrados mínimos parciais.*

Exploratory and discriminative studies of commercial processed Brazilian coffees with different degrees of roasting and decaffeinated

RIBEIRO, J. S. et al.

1 Introduction

Nowadays, world coffee consumers are looking for specific coffee tastes, and hence the food industries have had to invest into and create a greater variety of processed products from the same raw material. Amongst these processed coffee products originating from different coffee species (*Coffea arabica* and *Coffea canephora*), coffees with differentiated roasting degrees, decaffeination and freeze-drying, amongst others, are already present on the market.

Due to the obvious importance of flavour in consumer acceptance and the quality perception of coffee, the chemical composition and their key odorant compounds have been intensively studied (MAYER et al., 2000; MAYER and GROSCH, 2001; BUFFO and CARDELLI-FREIRE, 2004).

According to De Maria et al. (1999) the majority of the volatile compounds are formed during the roasting process, but by different mechanisms, such as the Maillard reactions, Strecker degradation and several other breakdown and degradation processes. Trigonelline, for example, forms pyridine and pyrrole derivatives by degradation. The pyrazines are formed by degradation of carbohydrates and chlorogenic acids are responsible for the formation of phenolic derivatives. Furan derivatives are formed by glycosides and sucrose and the lipids are the main class forming aldehydes, ketones, aliphatic alcohols and aromatic compounds.

During the roasting processes, the colour of the beans is directly correlated with the final roasting time and temperature: the higher the temperature and the longer the exposition time, the darker the coffee, so that colour can be used to define the end of the roasting process.

The degree of roasting is usually described as being "light", "medium" or "dark". A dark roasting process, for example, implies dark brown bitter beans and the lack of typical coffee aromas, whereas a light roasting process may be insufficient to complete all the pyrolytic reactions, resulting in a light brown coffee with underdeveloped organoleptic characteristics (BUFFO and CARDELLI-FREIRE, 2004).

When studying the roasting of Arabica coffee, Toci et al. (2006) found that the process degraded a considerable quantity of sucrose, and hence depending on the degree of roasting, the concentration of sucrose can be lower than the detection limit of the analytical method used. Significant degradation of the chlorogenic acids also occurs during the roasting process (83 and 97 % for the *C. arabica* and *C. canephora* samples, respectively). The reduction in protein content is around 10 % in light roasted coffee and up to 20 % in dark roasted coffee, and the trigonelline concentration is also reduced with the degree of roasting (TOCI et al., 2006).

Besides roasting, decaffeination is another process commonly applied to green coffee beans. Most of the decaffeination methods use organic solvents for caffeine extraction, such as dichloromethane, chloroform and others, dichloromethane being the most commonly used in Brazil. Two other methods are also known for the extraction of this alkaloid: water decaffeination and supercritical fluid (CO₂) decaffeination (CLARKE and VITZTHUM, 2001).

The influence of the decaffeination process on the bean composition and on the final product quality depends on the extraction method used, since other classes of compounds are also lost during the extraction process. In a study about the decaffeination process using dichloromethane, the results indicated a considerable decrease in sucrose content for the *C. canephora* and *C. arabica* varieties of green coffee (20 and 60 %, respectively). The decaffeination process also extracted 11 % of the chlorogenic acids from the *C. canephora* beans and 16 % from the *C. arabica* beans. The proteins, lipids and trigonelline were little extracted during the decaffeination process (TOCI et al., 2006).

There are currently several methods that can be combined with gas chromatography in order to analyze the volatile fraction, and the use of solid-phase microextraction (SPME) appears to be an attractive approach. Since its introduction, SPME has been shown to be an excellent sampling method, allowing for the simultaneous extraction and concentration of analytes from sample matrices (PAWLISZYN, 1998).

The use of chemometric methods is more or less essential to process data from instruments, since the technological advances in instrumentation have greatly increased their capability and the amount of data that can be generated and collected. In the past, most chemometric analyses were performed on reduced data sets, only using the areas of selected peaks detected on the chromatograms. Currently this limitation can be overcome by using the entire chromatographic profiles via chemometric analyses (RIBEIRO et al., 2009, 2010, 2012).

On the other hand, with the use of the chromatogram, small unavoidable differences in the experimental conditions are more apparent (minor changes in retention time and drift in the chromatograms caused, for instance, by column ageing or flow-rate variations). To minimize these differences, the algorithm for the correlation optimized warping (COW) of chromatographic profiles, introduced by Nielsen et al. (1998), is used.

The aligned chromatographic profiles were analyzed by the principal component analysis (PCA) to differentiate decaffeinated coffees from medium and dark roasted coffees. The partial least squares-discriminative analysis (PLS-DA) (BARKER and RAYENS, 2003) was then used to discriminate coffee samples according to

Exploratory and discriminative studies of commercial processed Brazilian coffees with different degrees of roasting and decaffeinated

RIBEIRO, J. S. *et al.*

their roasting degree (medium and dark roasted) and to compare the use of the two SPME fibres.

The identification of volatile compounds and differentiation of coffees according to their roasting degree has been carried out before (BICCHI *et al.*, 1997, 2002; LOPEZ-GALILEA *et al.*, 2006), but the data analysis proposed in this work, using the chromatographic fingerprint of roasted coffees instead of peak areas is a significant innovation. In addition, the discrimination of decaffeinated coffees is new and an alternative to the formation of coffee flavour, since the decaffeination process extracts other compounds besides caffeine.

The aim of this study was to differentiate commercial Brazilian coffee samples by their volatile compounds using modern chemometric methods.

2 Material and methods

2.1 Coffee samples

Twenty-one samples of commercial Brazilian coffees were obtained from local stores. Seven of them were decaffeinated and light roasted (**DC**), another seven were traditional or medium roasted (**TR**), and the last seven were dark roasted (**DR**). All samples were from different production batches.

2.2 SPME devices

The SPME fibres and the holder were purchased from Supelco (Bellefonte, PA). The fibres were conditioned according to the Supelco SPME data sheet (T7941231) in the GC injector port. The following types of SPME fibre were used: (i) polydimethylsiloxane /divinylbenzene (PDMS/DBV) with 65 μm thickness ($V_f = 0.357 \text{ mm}^3$) and (ii) Carboxen/polydimethylsiloxane (CX/PDMS) with 75 μm thickness ($V_f = 0.436 \text{ mm}^3$).

2.3 GC/FID parameters

The analyses were carried out using a G-6850 GC-FID system (Agilent, Wilmington, DE) fitted with a HP-5 capillary column (30 m \times 0.25 mm \times 0.25 μm), using helium (1 mL min^{-1}) as the carrier gas. The oven temperature was programmed as follows: 40 $^\circ\text{C} \rightarrow 7 \text{ }^\circ\text{C} / \text{min} \rightarrow 230 \text{ }^\circ\text{C} \rightarrow 30 \text{ }^\circ\text{C} / \text{min} \rightarrow 280 \text{ }^\circ\text{C}$. The injection port was equipped with a 0.75 mm i.d. liner and the injector was maintained at 220 $^\circ\text{C}$ in the splitless mode. Under these conditions, no sample carry-over was observed on the blank runs carried out between extractions.

2.4 GC-MS

The extracted analytes were identified in a Saturn 2000 Ion Trap GC-MS (Varian, Walnut Creek – CA, USA) fitted with the same column and operating under the

same conditions as the GC-FID. The GC-MS data were treated using the Automated Mass Spectral Deconvolution and Identification System (AMDIS) v. 2.61 software and the NIST Mass Spectral Search Program v. 1.6d (NIST, Washington, DC, USA), as well as making comparisons with previous reports on the volatile compounds of roasted coffee (RIBEIRO *et al.*, 2009).

2.5 General SPME procedure for sampling and injection

The volatile compounds were collected using different fibres but under the same conditions, in order to choose the most suitable fibre for sampling coffee volatiles. Ground coffee (250 mg) and 2 mL of saturated aqueous sodium chloride solution (salting-out effect) were transferred to a septum-sealed glass sample vial (5 mL). After 10 min of sample/headspace equilibration under agitation at 1200 rpm and 40 $^\circ\text{C}$, the fibres were exposed to the sample headspace for 10 min. After sampling, the fibres were immediately exposed in the injection port of the GC, and the analytes thermally desorbed for 10 minutes at 220 $^\circ\text{C}$. All the analyses were carried out in triplicate.

2.6 Chemometric data treatment

The original chromatographic profiles were organized into an X matrix format ($I \times J$), where each i replicate was used as a sample and the J variables were the electrical signal received by the FID detector every 0.05 seconds. The data was analysed by the Matlab 6.5 software (The MathWorks, Co., Natick, MA, USA) using the computational package PLS_Toolbox (Eigenvector Research, Inc. – PLS_Toolbox version 3.02.) (WISE *et al.*, 2004).

The chromatograms were aligned using a correlation optimized warping (COW) algorithm obtained from www.models.kvl.dk/source/. After alignment, the data matrix was smoothed by the Savitzky-Golay algorithm using a window size of five points (SAVITZKY and GOLAY, 1964) and column-wise autoscaling.

In chromatography, peak selection not only enhances the stability of the classification model but also helps in interpreting the relationship between the model and the sample compositions. In this work, the new method denominated ordered predictors selection (OPS) (TEÓFILO *et al.*, 2009) was applied. This method uses an intuitive vector formed by a combination of vectors such as the regression vector, correlation vector and others. With this intuitive vector, the independent variables are ordered according to their importance to the model. The ordered variables are then tested using increments over a previously defined window and the RMSECV and the

Exploratory and discriminative studies of commercial processed Brazilian coffees with different degrees of roasting and decaffeinated

RIBEIRO, J. S. et al.

correlation cross-validation coefficient (r_{cv}) values stored for each window analyzed. The best set of variables is indicated by the lowest RMSECV and the highest r_{cv} .

3 Results and discussion

3.1 HS-SPME-GC-MS analysis

The volatile compounds obtained from the roasted coffee headspaces, extracted with the two SPME fibres, were detected by GC-MS analysis and identified by comparing their fragmentations with the NIST data bank and some recently published papers (RIBEIRO et al., 2009, 2010). This procedure was capable of extracting and detecting a large number of volatile compounds. Pyrazine, furan and pyrrole derivatives were some of the organic classes found in the coffee headspace. Some of these compounds are described in Table 1 and used in the followed discussions.

3.2 Principal component analysis and partial least squares-discriminative analysis

3.2.1 PDMS/DVB fibre

The principal component analysis (PCA) was applied to the pre-treated data set in order to get an insight into which peaks could be responsible for the discrimination of the decaffeinated samples from the others (medium or dark roasted). From the scores plot in Figure 1A, a distinct visual clustering distinguished the decaffeinated samples from the other two classes when the data were displayed with respect to the first three principal components (PC). The decaffeinated coffees are located on the left, with negative scores in PC1 (which describes 57.56 % of the original information), well separated from the medium and dark roasted samples on the right side with positive scores (Figure 1A). PC1 also showed a tendency to separate the medium from the dark roasted coffee samples.

Table 1. Main compounds identified from the chromatographic peaks in Figure 1 by comparison of their MS spectra with those of the NIST MS data base and the literature.

Peak	Retention Time (minutes)	Compounds identified ^a	Main <i>m/z</i> ions observed in MS spectra ^b
1	2.90	2-Pentanone	43(B)
2	3.25	Methylpropanal	43(B), 72
3	3.45	3-Methyl -1-butenol	43(B), 72
4	3.50	2-Methylfuran	82(B), 53
5	3.90	2-Methylbutanal	41(B), 57, 29
6	4.00	2,3-Pentanedione	43(B), 57
7	4.45	2,6-Dimethylfuran	96(B), 43, 45
8	4.50	2,4-Dimethylfuran	96(B), 63, 53
9	4.95	<i>N</i> -Methylpyrrole	81(B), 53
10	5.10	Pyrazine	80(B)
11	5.90	1-Ethylpyrrole	83(B), 98, 55
12	6.10	4-Methylthiazole	71(B), 99, 45
13	6.30	2- Carboxyaldehyde pyrrole	95(B)
14	6.45	Furfural	96(B)
15	7.10	2-Furanmethanol	81(B), 98
16	7.80	Furfuryl formate	81(B), 126
17	8.00	Acetylfuran	95(B), 110
18	8.10	2,5-Dimethylpyrazine	108(B), 42
19	8.15	Ethylpyrazine	107(B)
20	9.05	5-Methyl-2-furancarboxyaldehyde	110(B), 53
21	9.55	3-Acetyl-1-methylpyrrole	108(B), 123
22	9.70	2-Furanmethanol acetate	81(B), 140
23	9.80	2-Ethyl-5-methylpyrazine	121(B)
24	9.90	2-Furanmethanethiol	81(B), 53
25	10.00	Trimethylpyrazine	122(B), 42
26	11.50	3-Ethyl-2,5-dimethylpyrazine	135(B)
27	11.60	2-Furfurylfuran	148(B), 91

a. Match of all compounds higher than 800. b. B denotes the base peak observed in the analyte library mass spectrum.

Exploratory and discriminative studies of commercial processed Brazilian coffees with different degrees of roasting and decaffeinated

RIBEIRO, J. S. et al.

Analyzing the loadings of the 10 selected peaks (Figure 1B), the most important compounds identified by GC-MS analysis as responsible for the discrimination of the decaffeinated coffee from the other two classes were 2-methylbutanal (**5**), 2,3-pentanedione (**6**), *N*-methylpyrrole (**9**), pyrazine (**10**), furfural (**14**), 5-methyl-2-furancarboxyaldehyde (**20**), 2-furanmethanol acetate (**22**), 2-ethyl-5-methylpyrazine (**23**), 2-furanmethanethiol (**24**) and 3-ethyl-2,5-dimethylpyrazine (**26**).

Figure 1C shows a typical overlap of two mean chromatograms, one for traditional and dark roasted coffees and the other for decaffeinated coffees, for comparison. The compounds mentioned above are enumerated and the differences between the samples can be identified. It can be seen that *N*-methylpyrrole (**9**), pyrazine (**10**), 2-furanmethanol acetate (**22**), 2-ethyl-5-methylpyrazine (**23**), 2-furanmethanethiol (**24**) and 3-ethyl-2,5-dimethylpyrazine (**26**) tended to be in higher relative concentrations in the traditional and dark roasted samples (positive loadings). The other compounds appeared in greater relative concentrations in the decaffeinated samples (negative loadings).

The decaffeination process causes a significant loss of sucrose (TOCI et al., 2006), and the principal products of sucrose pyrolysis are furan, pyrrole and pyrazine derivatives. The chromatographic results indicated that the concentration of pyrazines, pyrroles and certain furan derivatives were reduced in the decaffeinated samples.

The partial least squares – discriminative analysis (PLS-DA) was applied to classify the samples according to their roasting degree. For this, the decaffeinated coffees were excluded from the original data matrix and the data set randomly split into a training set consisting of ten

samples (thirty replicates – fifteen for each class), while the other four samples (twelve replicates – three for each class) were used as the external validation set.

To construct the training data set, the pre-treated chromatogram profiles were divided into five regions: region one ranging from retention times of 2.5 to 3.8 minutes, region two from 3.8 to 6.8 minutes, region three from 6.8 to 8.6 minutes, region four from 8.8 to 10.2 minutes and the last region from 10.2 to 15 minutes. The variables (peaks) in each region were selected by the OPS method (see materials and methods) and the original variables reduced from 7500 to 150 (14 peaks).

Figure 2C shows the main peaks selected for the construction of the PLS-DA. The peaks were identified as being: 2-methylfuran (**4**), 2-methylbutanal (**5**), 2,3-pentanedione (**6**), pyrazine (**10**), 4-methylthiazole (**12**), 2-carboxyaldehyde pyrrole (**13**), furfural (**14**), 2-furanmethanol (**15**), acetylfuran (**17**), 2,5-dimethylpyrazine (**18**), ethylpyrazine (**19**), 5-methyl-2-furancarboxyaldehyde (**20**), 3-acetyl-1-methylpyrrole (**21**) and 2-furanmethanol acetate (**22**).

A good PLS-DA model was obtained for the PDMS/DVB fibre. The measured value ($y = 1$) corresponded to the traditional class and $y = 0$ was used for the dark roasted samples. Using three latent variables (LV), the statistical parameters indicated a low root mean square error of cross validation ($RMSECV = 0.15$) and a high correlation coefficient ($r_{cv} = 0.97$). The samples from the training and external validation set were all classified correctly (all samples in triplicate).

The scores for LV1 plotted against LV2 and LV3 (Figure 2A) show how the samples are clustered in the subspace defined by the first three components of the

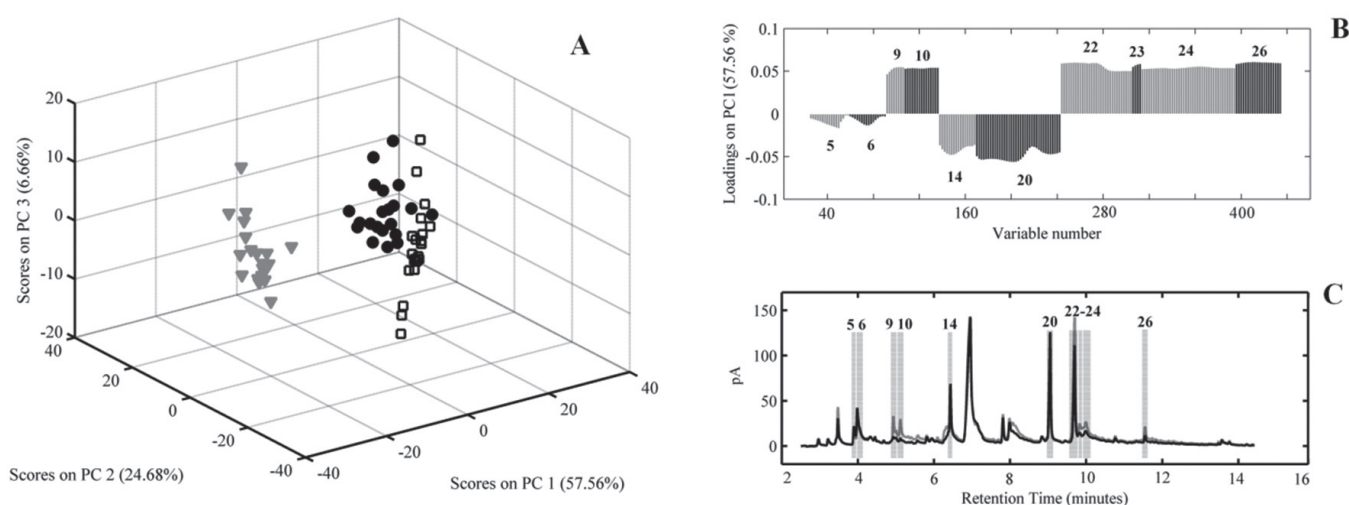


Figure 1. (A) PC1 x PC2 x PC3 scores plot for the PDMS/DVB fibre - (∇) Decaffeinated samples, (\square) dark roasted and (\bullet) medium roasted. (B) PC1 loadings for the ten peaks selected for PCA. (C) Mean chromatograms of medium and dark roasted (—) and decaffeinated (---) coffees obtained using the PDMS/DVB fibre.

Exploratory and discriminative studies of commercial processed Brazilian coffees with different degrees of roasting and decaffeinated

RIBEIRO, J. S. et al.

PLS-DA model. LV1 contains 43.72 % of the total variance and is important for a good separation between medium and dark roasted coffees. The training and external validation samples are included in Figure 2A.

The different concentrations of the selected peaks determine the sample position inside the classes. The compound 4-methylthiazole (**12**) appears in higher concentrations in dark roasted coffees than in medium roasted coffees (negative loadings). For the other selected peaks, the concentrations are higher in medium roasted coffees (positive loadings). A loading plot of latent variable 1 is shown in Figure 2B. These results differ from Franca et al. (2006) in the case of furfural.

The high concentrations of almost all the compounds in traditional coffee can be explained by the walls of the beans. All the pyrolytic reactions occur in the beans, and their walls act as autoclaves under high pressure. When the roasting process continues to produce dark roasted beans, the walls do not support the pressure and “explode”, and lots of volatile compounds are lost to the air.

This analysis shows that it is possible to discriminate medium and dark roasted coffees by applying variable selection directly to the chromatographic profiles, instead of peak areas. The discrimination was masked (see Figure 1A) when all useful and non useful retention times were taken into account.

3.2.2 CX/PDMS fibre

A similar data analysis was performed on the data matrix from the CX/PDMS fibre. PCA was applied to the eight peaks visually selected from the aligned chromatograms

and revealed basically the same information as with the PDMS/DVB fibre. PC1 described 77.94 % of the total variance and indicated a good separation between the decaffeinated coffees and the other classes. The plot of the scores is illustrated in Figure 3A.

The volatile compounds identified by the PC1 loadings (Figure 3B) considered most influential for the separation of decaffeinated coffees from the other two classes were: *N*-methylpyrrole (**9**), pyrazine (**10**), 2-carboxyaldehyde pyrrole (**13**), furfural (**14**), acetylfuran (**17**), 2,5-dimethylpyrazine (**18**), ethylpyrazine (**19**) and 5-methyl-2-furancarboxyaldehyde (**20**).

In this case a difference in concentration was also the principal factor of discrimination. The concentrations of the compounds *N*-methylpyrrole (**9**), pyrazine (**10**), 2-carboxyaldehyde pyrrole (**13**), acetylfuran (**17**), 2,5-dimethylpyrazine (**18**) and ethylpyrazine (**19**) were higher in the traditional and dark roasted coffees (positive loadings), while furfural (**14**) and 5-methyl-2-furancarboxyaldehyde (**20**) (negative loadings) were more abundant in decaffeinated coffees (Figure 3C). The results obtained with the CX/PDMS fibre indicated the same tendency as the PDMS/DVB fibre, showing pyrazines, pyrroles and certain furan derivatives as more abundant in traditional coffees.

Using the same procedure used for the PDMS/DVB fibre, a PLS-DA model was also built to classify the coffee samples according to their roasting degree. The decaffeinated coffee samples were again removed in this case and the remaining data set was randomly split into a training set (ten samples) and external validation set (four samples). The important variables (peaks) for the training

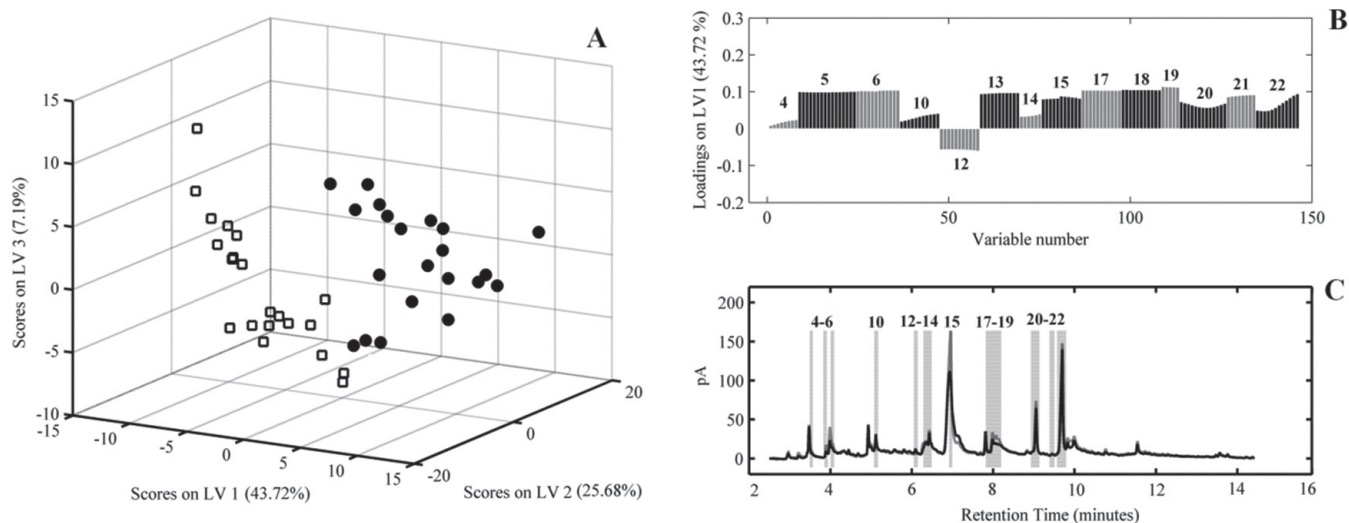


Figure 2. (A) LV1 x LV2 x LV3 scores plot for the PDMS/DVB fibre – (□) Dark roasted and (●) medium roasted peaks selected by OPS and used in the PLS-DA. (B) LV1 loadings plot of the main peaks selected by OPS for the PLS-DA model (C) Chromatograms of medium roasted (—) and dark roasted (—) coffees obtained using the PDMS/DVB fibre.

Exploratory and discriminative studies of commercial processed Brazilian coffees with different degrees of roasting and decaffeinated

RIBEIRO, J. S. *et al.*

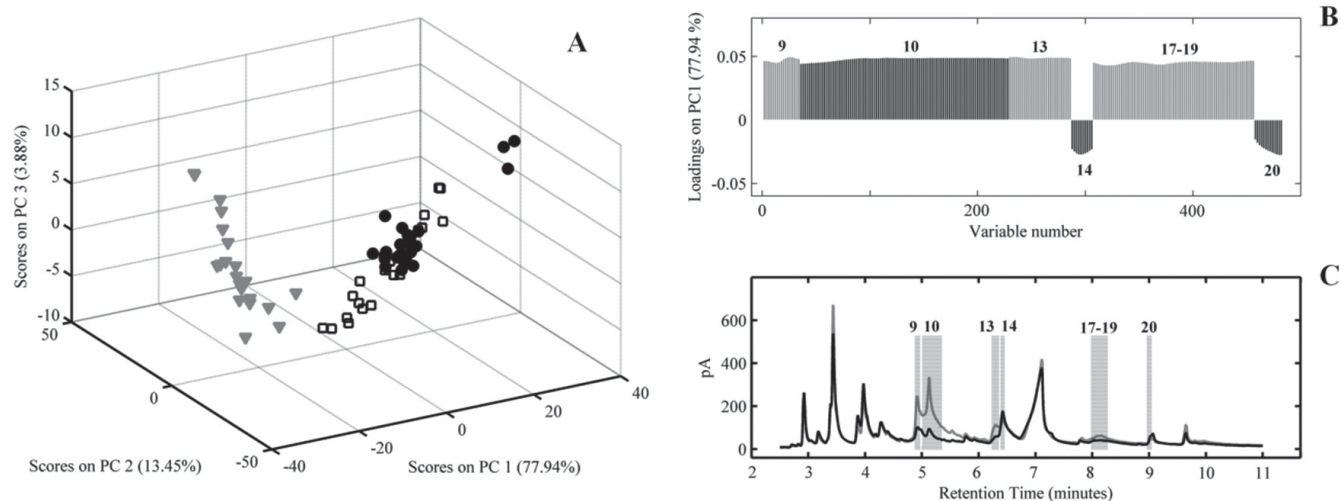


Figure 3. Plots of the PC1 x PC2 x PC3 scores for the CX/PDMS fibre (A) – (▼) Decaffeinated samples, (□) dark roasted and (●) medium roasted. PC1 loadings for the eight peaks selected for PCA (B). Mean chromatogram of medium and dark roasted (—) and decaffeinated (—) coffees obtained using the CX/PDMS fibre (C).

set were selected by the OPS method and the original data variables were reduced from 5100 to 284 (17 peaks).

The main compounds selected as the best discriminators for the roasting degree classification were 2-pentanone (**1**), methylpropanal (**2**), 3-methyl-1-butenol (**3**), 2-methylfuran (**4**), 2-methylbutanal (**5**), 2,3-pentanedione (**6**), 2,6-dimethylfuran (**7**), *N*-methylpyrrole (**9**), pyrazine (**10**), 2-carboxyaldehyde pyrrole (**13**), furfural (**14**), 2-furanmethanol (**15**), furfuryl formate (**16**), acetylfuran (**17**), 2,5-dimethylpyrazine (**18**), ethylpyrazine (**19**) and 5-methyl-2-furancarboxyaldehyde (**20**) (Figure 4C).

By applying PLS-DA, a model describing the discrimination of the predefined classes was obtained. Using three latent variables, the statistical parameters indicated a low root mean square error of cross validation (RMSECV = 0.22) and a high correlation coefficient ($r_{cv} = 0.97$). The results obtained from the PLS-DA indicated that all the samples from the external validation set were correctly classified.

The scores for LV1 plotted against LV2 and LV3 in Figure 4A show how the samples were clustered in the subspace defined by the first three components of the PLS-DA model. LV1 accounts for 53.94 % of the total variance and is important for a good separation of the medium from the dark roasted coffees. As with the PDMS/DVB fibre, after the variable selection, the volatiles not useful for discrimination amongst the samples were excluded.

It can be seen in Figure 4C, which shows the chromatograms from each class that, except for 2-pentanone, the concentrations of all the selected compounds were higher in medium roasted than in dark

roasted samples. The corresponding values for the loadings of the compounds selected are indicated in Figure 4B.

3.3 Comparison between the results of the two SPME fibres

The results obtained using the PCA demonstrated that both fibres tested in this work were useful for the identification of decaffeinated coffee samples. Table 1 shows the main compounds detected and identified by GC-MS analyses which are important for the PCA analyses.

The loadings in Figures 1B and 3B indicated that the compounds **9**, **10**, **14** and **20** contributed to the clustering in the PCA analyses for the PDMS/DVB and CX/PDMS fibres. The other volatile compounds important for the PCA analyses differed from one fibre to the other. This occurs because the coverings of the SPME fibres are made from different materials and so the equilibrium adsorption of certain compounds does not work equally in both cases.

Another perceptible difference occurred in the relative concentrations of the compounds adsorbed by both fibres. The CX/PDMS fibre was more effective for the highly volatile compounds, non-retained analytes. This means that the CX/PDMS fingerprint shows higher concentrations of the compounds than the PDMS/DVB fingerprint. This can be seen mainly with the light compounds. However, the PDMS/DVB fibre showed greater applicability for a larger range of compounds (higher overall sampling capability).

Visually, comparing Figures 1C and 3C, the CX/PDMS fibre demonstrated significant differences between the relative concentrations of some peaks used

Exploratory and discriminative studies of commercial processed Brazilian coffees with different degrees of roasting and decaffeinated

RIBEIRO, J. S. et al.

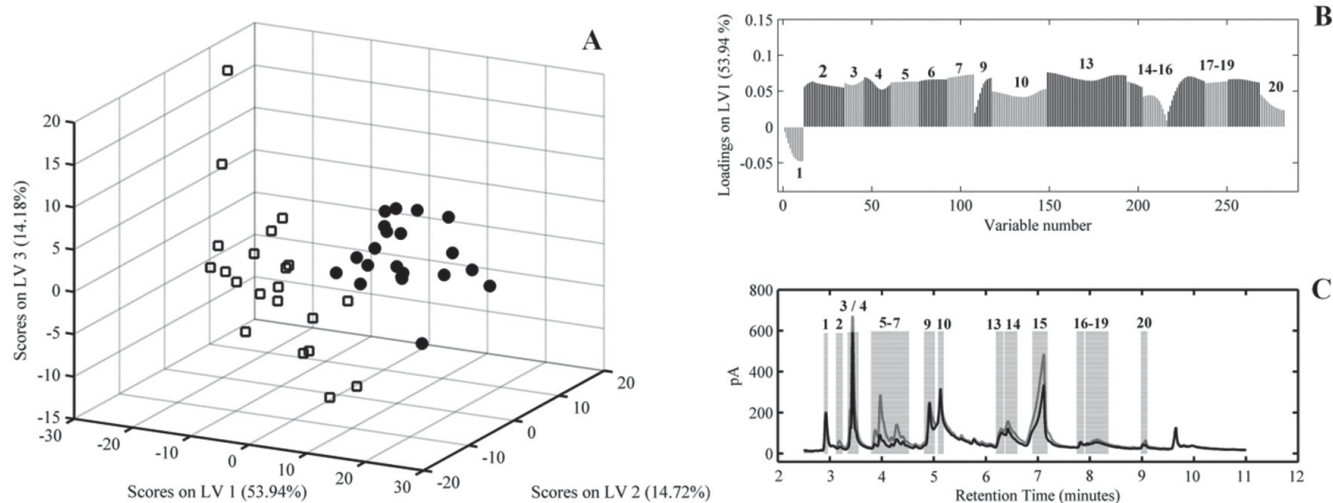


Figure 4. Plots of the CX/PDMS scores – LV1 x LV2 x LV3 - (□) Dark roasted and (●) medium roasted (A). LV1 loadings plot of the main peaks selected by OPS for the PLS-DA model (B). Peaks selected by OPS and used in the PLS-DA. Chromatograms of medium (—) and dark roasted (---) coffees obtained using the CX/PDMS fibre (C).

for discriminations in the PCA analyses. This difference was not an important parameter in the chemometric analyses because the extraction systems always work under equilibrium. However, the PDMS/DVB fibre showed excellent reproducibility and an extended lifetime as compared to the CX/PDMS fibre.

When applying PLS-DA to predict the sample roasting degree, the models constructed for both fibres for the chromatographic profile, instead of peak areas, using three latent variables, correctly classified all the samples in the training and external validation sets.

4 Conclusions

The results reported in this work demonstrated that the analysis of the chromatographic fingerprints of volatile compounds from roasted coffee as spectroscopic profiles, contrary to the traditional integrated areas of the detected peaks, gave a new vision to the treatment of the chromatographic data. This occurred because relevant additional information could be obtained and interpreted by chemometrics. The chromatographic data alignment and pretreatment algorithms have opened new ways to treat chromatographic data in general, and are winning a lot of adepts in the real world.

The PCA analyses for both fibres indicated that the compounds N-methylpyrrole, pyrazine, furfural and 5-methyl-2-furancarboxyaldehyde appeared in higher concentrations in traditional coffees as compared to decaffeinated ones.

In the PLS-DA models for both SPME fibres, the compounds 2-methylfuran, 2-methylbutanal, 2,3-pentanedione, pyrazine, 2-carboxyaldehyde pyrrole, furfural, 2-furanmethanol, acetylfuran,

2,5-dimethylpyrazine, ethylpyrazine and 5-methyl-2-furancarboxyaldehyde were found to be possible roasting markers. The relative concentrations of all the selected compounds were higher in medium roasted than in dark roasted samples.

Acknowledgements

This work was supported by grants from CAPES and FAPESP.

References

- BARKER, M.; RAYENS, W. Partial least squares for discrimination. **Journal of Chemometrics**, New Jersey, v. 17, p. 166-173, 2003. <http://dx.doi.org/10.1002/cem.785>
- BICCHI, C. P.; PANERO, O. M.; PELLEGRINO, G. M.; VANNI, A. C. Characterization of roasted coffee and coffee beverages by solid phase microextraction-gas chromatography and principal component analysis. **Journal of Agricultural and Food Chemistry**, Washington, v. 45, n. 12, p. 4680-4686, 1997. <http://dx.doi.org/10.1021/jf9704827>
- BICCHI, C.; IORI, C.; RUBIOLLO, P.; SANDRA, P. Headspace sorptive extraction (HSSE), stir bar sorptive extraction (SBSE), and solid phase microextraction (SPME) applied to the analysis of roasted arabica coffee and coffee brew. **Journal of Agricultural and Food Chemistry**, Washington, v. 50, n. 3, p. 449-459, 2002. <http://dx.doi.org/10.1021/jf010877x>
- BUFFO, R.; CARDELLI-FREIRE, C. Coffee flavour: an overview. **Flavour and Fragrance Journal**, Geneva, v. 19, n. 2, p. 99-104, 2004. <http://dx.doi.org/10.1002/ffj.1325>
- CLARKE, R. J.; VITZTHUM, O. G. **Coffee: Recent Developments**. Boston: World Agriculture Series - Blackwell Science, 2001.

Exploratory and discriminative studies of commercial processed Brazilian coffees with different degrees of roasting and decaffeinated

RIBEIRO, J. S. *et al.*

- DE MARIA, C. A. B.; MOREIRA, R. F. A.; TRUGO, L. C. Volatile components in roasted coffee, Part I: Heterocyclic compounds. **Química Nova**, São Paulo, v. 22, n. 2, p. 209-217, 1999.
- FRANCA, A. S.; OLIVEIRA, L. S.; OLIVEIRA, R. C. S.; AGRETI, P. C. M.; AUGUSTI, R. A. A preliminary evaluation of the effect of processing temperature on coffee roasting degree assessment. **Journal of Food Engineering**, Davis, v. 92, p. 345-352, 2009.
- LOPEZ-GALILEA, I.; FOURNIER, N.; CID, C.; GUICHARD, E. Changes in headspace volatile concentrations of coffee brews caused by the roasting process and the brewing procedure. **Journal of Agricultural and Food Chemistry**, Washington, v. 54, n. 22, p. 8560-8566, 2006. PMID:17061834. <http://dx.doi.org/10.1021/jf061178t>
- MAYER, F.; CZERNY, M.; GROSCH, W. Sensory study of the character impact aroma compounds of a coffee beverage. **European Food Research and Technology**, New York, v. 211, n. 4, p. 272-276, 2000. <http://dx.doi.org/10.1007/s002170000169>
- MAYER, F.; GROSCH, W. Aroma simulation on the basis of the odourant composition of roasted coffee headspace. **Flavour and Fragrance Journal**, Geneva, v. 16, n. 3, p. 180-190, 2001. <http://dx.doi.org/10.1002/ffj.975>
- NIELSEN, N. P. V.; CARSTENSEN, J. M.; SMEDSGAARD, J. Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimised warping. **Journal of Chromatography A**, Amsterdam, v. 805, n. 1-2, p. 17-35, 1998.
- PAWLISZYN, J. **Solid Phase Microextraction: Theory and Practice**. New York: Wiley-VCH, 1998.
- RIBEIRO, J. S.; AUGUSTO, F.; SALVA, T. J. G.; THOMAZIELLO, R. A.; FERREIRA, M. M. C. Prediction of sensory properties of Brazilian Arabica roasted coffees by headspace solid phase microextraction - gas chromatography and partial least squares. **Analytica Chimica Acta**, Amsterdam, v. 634, n. 2, p. 172-179, 2009.
- RIBEIRO, J. S.; AUGUSTO, F.; SALVA, T. J. G.; FERREIRA, M. M. C. Uso de perfis cromatográficos de voláteis de cafés arabicas torrados para a diferenciação das amostras segundo o sabor, o aroma e a qualidade global da bebida. **Química Nova**, São Paulo, v. 33, n. 9, p. 1897-1904, 2010. <http://dx.doi.org/10.1590/S0100-40422010000900015>
- RIBEIRO, J. S.; AUGUSTO, F.; SALVA, T. J. G.; FERREIRA, M. M. C. Prediction models for Arabica coffee beverage quality based on aroma analyses and chemometrics. **Talanta**, Seattle, v. 101, p. 253-260, 2012. PMID:23158320. <http://dx.doi.org/10.1016/j.talanta.2012.09.022>
- SAVITZKY, A.; GOLAY, M. J. E. Smoothing + differentiation of data by simplified least squares procedures. **Analytical Chemistry**, Washington, v. 36, n. 8, p. 1627-1639, 1964. <http://dx.doi.org/10.1021/ac60214a047>
- TEÓFILO, R. F.; MARTINS, J. P. A.; FERREIRA, M. M. C. Sorting variables by using informative vectors as a strategy for featuring selection in multivariate regression. **Journal of Chemometrics**, New Jersey, v. 23, p. 32-41, 2009. <http://dx.doi.org/10.1002/cem.1192>
- TOCI, A.; FARAH, A.; TRUGO, L. C. Effect of decaffeination using dichloromethane on the chemical composition of arabica and robusta raw and roasted coffees. **Química Nova**, São Paulo, v. 29, n. 5, p. 965-971, 2006. <http://dx.doi.org/10.1590/S0100-40422006000500015>
- WISE, B. M.; GALLAGHER, N. B.; BRO, R.; SHAVER, J. M.; WINDIG, W.; KOCH, R. S. **PLS_Toolbox 3.5, for Use with Matlab™**. Manson: Eigenvector Research, 2004.