Talanta 101 (2012) 253–260

Contents lists available at SciVerse ScienceDirect

Talanta

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

Prediction models for Arabica coffee beverage quality based on aroma analyses and chemometrics

J.S. Ribeiro [a,*], F. Augusto [b], T.J.G. Salva [c], M.M.C. Ferreira [d]

[a] Federal Institute of Espírito Santo, Campus Vila Velha, CET, Av. Ministro Salgado Filho, 1000, Soteco, CEP: 29106-010, Vila Velha, ES, Brazil
[b] Gas Chromatography Laboratory, Chemistry Institute, University of Campinas, P.O. Box 6154, 13083-970 Campinas, SP, Brazil
[c] Agronomic Institute of Campinas-Coffee Centre, P.O. Box 28, 13012-970 Campinas, SP, Brazil
[d] Theoretical and Applied Chemometrics Laboratory, Chemistry Institute, University of Campinas, P.O. Box 6154, 13083-970 Campinas, SP, Brazil

ARTICLE INFO

Article history:
Received 27 July 2012
Received in revised form
12 September 2012
Accepted 14 September 2012
Available online 23 September 2012

Keywords:
Chemometrics
Sensorial data
Overall quality
Flavor
Bitterness
SPME

ABSTRACT

In this work, soft modeling based on chemometric analyses of coffee beverage sensory data and the chromatographic profiles of volatile roasted coffee compounds is proposed to predict the scores of acidity, bitterness, flavor, cleanliness, body, and overall quality of the coffee beverage. A partial least squares (PLS) regression method was used to construct the models. The ordered predictor selection (OPS) algorithm was applied to select the compounds for the regression model of each sensory attribute in order to take only significant chromatographic peaks into account.

The prediction errors of these models, using 4 or 5 latent variables, were equal to 0.28, 0.33, 0.35, 0.33, 0.34 and 0.41, for each of the attributes and compatible with the errors of the mean scores of the experts. Thus, the results proved the feasibility of using a similar methodology in on-line or routine applications to predict the sensory quality of Brazilian Arabica coffee.

1. Introduction

In Brazil, the criteria used to classify coffee beans includes physical aspects, such as size, color, and the presence of defects in the beans, and the sensorial characteristics of the beverage itself [1]. The International Organization for Standardization (ISO) presents some norms assigned to composition, defects, bean type and format, and color, as well as characteristics after roasting and the sensorial cup profile [2].

Despite physical aspects, the most important quality criteria used for assessing Brazilian Arabica coffees is cup tasting or classification by beverage type. However, such tests are often criticized due to their subjectivity, besides the long time consumed to perform the sensorial analysis and the scarcity of experts [3,4]. In order to circumvent these limitations, specialists and scientists are searching worldwide for alternatives using different analytical techniques whose results could be related to coffee cup quality [5–9].

In this regard, the identification of chemical markers associated with sensory attributes of the coffee beverage becomes a viable alternative in the search for an objective and less time-consuming analytical method [8,9]. Furthermore, it is well known that aroma is one of the most important attributes of the coffee beverage. Thus, it has recently been highly valued and widely used to discriminate high-quality coffees [6,8,10].

In this paper, several sensorial properties of Arabica coffee are predicted by the proposed PLS models, based on analytical data, obtained by solid phase microextraction–gas chromatographic analysis (SPME–GC), and sensory evaluations by experts. Initial results relating chromatographic profiles to sensorial data have recently been published [8] and were so encouraging that the work was extended to improve the correlation between chromatographic data and sensory analysis, to introduce more attributes, such as acidity and bitterness, and to identify the significant volatile compounds related to each PLS model.

2. Materials and methods

2.1. Coffee samples

Fifty-three green Arabica coffee samples with different beverage characteristics were roasted separately in a gas-fired drum roaster (Pinhalense S/A Máquinas Agrícolas, Brazil) to the medium roasting point (Agtron #55—according to the SCAA roast color classification system). The roasted coffee samples were packed in special plastic films (polystyrene and polyethylene) and aluminum foil to avoid loss of aroma and contamination from external substances.

* Corresponding author. Tel.: +55 27 31490825.
E-mail address: julianoribeiro@ifes.edu.br (J.S. Ribeiro).

0039-9140/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2012.09.022

Table 1
Main compounds identified from the mass analyses by comparison of their MS spectra with those of the NIST MS database and literature.

No.	R.T.	Compounds	Fragmentation	Math	Models
1	2.01	Acetaldehyde	44 (B), 43, 42	932	D
2	2.055	Methanethiol	47 (B), 45, 60	986	C, D, E
3	2.2	Acetone	43 (B), 58	870	B, C, D, F
4	2.305	Methyl acetate	43 (B), 74	944	B, C, D, E
5	2.422	2-Methylpropanal	43 (B), 39, 72	911	D
6	2.57	2,3-Butanedione	43 (B), 86	961	D
7	2.625	2-Butanone	43 (B), 73	959	C, D
8	3.065	3-Methylbutanal	41 (B), 39, 58, 71	962	A–F
9	3.158	2-Methylbutanal	41 (B), 57, 39	962	A, F
10	3.41	3-Methyl 2-butanone	43 (B), 86	893	D
11	3.477	2,3-Pentanedione	43 (B), 57	945	A, D, E, F
12	3.516	Acetic acid	45 (B), 60	952	B, D
13	3.88	2,5-Dimethylfuran	43 (B), 53, 96	922	D, F
14	4	3-Methyl pyridazine	39 (B), 65, 94	891	A, D
15	4.05	Pyrazine	80 (B), 53, 96	910	B
16	4.145	1-Methyl pyrrole	81 (B), 39, 42, 53	954	A, C, D, E, F
17	4.21	Pyridine	79 (B), 52	936	A–F
18	4.364	Pyrrole	67 (B), 39	942	A–F
19	4.865	2,3-Hexanedione	43 (B), 71, 114	911	A, F
20	5.032	2,4-Hexanedione	57 (B), 114	924	B, C, D, F
21	5.16	3,3-Dimethyl 2-butanone	57 (B), 43, 100	800	A, E
22	5.325	Dihydro-2-methyl 3(2H) furanone	43 (B), 72, 100	928	A–F
23	5.485	1-Ethyl 1H-pyrrole	80 (B), 95, 67, 53, 78	806	C
24	5.55	4-Methyl thiazole	71 (B), 99, 45	863	C
25	5.677	Methyl pyrazine	94 (B), 67, 39	965	C, E, F
26	5.794	Furfuryl methyl ether	81 (B), 53, 112	876	A, B
27	5.845	3-Methyl phenol	108 (B), 43, 65, 79	768	B, C, D, E, F
28	5.92	Furfural	39 (B), 95	953	A, C, D, E, F
29	5.99	2,n-Dimethyl 1H-pirrole	94 (B), 42	800	C
30	6.25	Trimethyl oxazole	111 (B), 42, 55, 68	882	C
31	6.66	2-Furanmethanol	98 (B), 41, 53, 81, 69	951	C, D, E, F
32	6.745	3-Methyl butanoic acid	60 (B), 45, 87, 99	822	D
33	7.675	Furfuryl formiate	81 (B), 53, 39, 126	882	B, F
34	7.74	2-Furanmethanethiol	81 (B), 53, 114	864	C, E
35	7.865	2,5-Dimethyl pyrazine	42 (B), 108, 39	902	A, E, F
36	7.953	Ethyl pyrazine	107 (B)	848	B, C, E
37	8.062	2,3-Dimethyl pyrazine	67 (B), 40, 108	886	D, F
38	8.335	Butyrolactone	42 (B), 39, 56, 86	885	D
39	8.58	Ethenyl pyrazine	106(B), 52, 79	820	A, C, D
40	8.78	N/I	–	–	A
41	8.83	2-n-Butyl furan	81 (B), 124	755	C
42	9	3-Ethyl pyridine	92 (B), 107, 65, 39	918	A, B, C, F
43	9.19	Benzaldehyde	77 (B), 105, 51	850	F
44	9.34	5-Methyl-2-furancarboxaldehyde	53 (B), 110, 81	938	A, C, D, E, F
45	9.542	1-Acetyloxy 2-butanone	43 (B), 57, 130	902	D, F
46	9.64	N/I	–	–	B, C
47	9.95	Phenol	94 (B), 66, 39	898	B, C, D, E, F
48	10.23	2-Furanmethanol acetate	81 (B), 43, 98, 140	949	A–F
49	10.29	2-Ethyl-6-methyl-pyrazine	121 (B), 94, 128	830	A, D
50	10.37	2-Ethyl-5-methyl-pyrazine	121 (B), 39, 58	793	A, B, C, D, F
51	10.42	Trimethyl pyrazine	42 (B), 122, 39	832	A, C
52	10.51	1-Methyl-1H-pyrrole 2-carboxaldehyde	109 (B), 53, 39, 80		B, C, D, E, F
53	10.572	2-Propionylfuran	95 (B), 39, 124	889	B, D, E, F
54	10.77	2-Etenyl 6-methyl pyrazine	52 (B), 120, 39	778	B, C, D, F
55	10.88	2-Pyrrolycarboxaldehyde	95 (B), 66, 39	886	D, E
56	10.973	N/I	–	–	B
57	11.15	Limonene	68 (B), 93, 136	784	D, F
58	11.25	2-Acetylpyridine	79 (B), 43, 121	847	F
59	11.37	N/I	–	–	A
60	11.444	2,2′-Bifuryl	134 (B), 78, 105	805	A, E
61	11.57	Benzeneacetaldehyde	91 (B), 120, 65, 39	893	B, C, D, E
62	11.67	4-Pyridazinamide	95 (B), 43	748	B, F
63	11.76	1-(2′-furyl)-2-butanone	57 (B), 81, 138	840	A, C, D
64	11.9	n-Methyl phenol	108 (B), 79, 91	800	B, C
65	11.95	N/I	-	–	C, D, E, F
66	12.14	3-Acetoxypyridine	95 (B), 43, 137	785	F
67	12.21	2-Acetyl pyrrole	94 (B), 109, 66	800	A, F
68	12.42	N/I	–	–	C, D, F
69	12.48	2-Acetyl N-methylpirrole	108 (B), 123	909	F
70	12.6	3-Athyl 2,5-dimethyl pyrazine	42 (B), 39, 135	905	E
71	12.91	p-Guaiacol	81 (B), 109, 53, 39	901	A, B, C, D
72	13	N/I	–	–	A, C, D, E, F
73	13.17	N/I	–	–	C
74	13.27	N/I	–	–	D
75	13.37	N/I	–	–	B, C, E

Table 1 (*continued*)

No.	R.T.	Compounds	Fragmentation	Math	Models
76	13.6	Maltol	43 (B), 71, 126, 55, 97	882	C, E, F
77	13.83	N/I	–	–	C, E, F
78	14.06	5-Methyl 2-propionylfuran	109 (B), 53, 138	876	A, C, D, E, F
79	14.85	N/I	–		A, C, D, E, F
80	14.925	N/I	–		C, E
81	15.01	N/I	–		C, D, E
82	15.1	N/I	–		D, E, F
83	15.2	N/I	–		C, D, E, F
84	15.32	N/I	–		D, E, F
85	15.545	N-Furfuryl pirrole	81 (B), 53, 147	894	D, F
86	15.61	Coelution—N/I	–	–	C, D, E, F
87	15.76	Coelution—N/I	–	–	B, C, D, E, F
88	15.9	N/I	–	–	E, F
89	16.01	N/I	–	–	F
90	16.14	N/I	–	–	A, B, D, F
91	16.245	N/I	–	–	A, B, C, D, F
92	16.385	Furfuryl methyl disulfide	81 (B), 53, 160	891	F
93	16.49	N/I	–	–	D, F
94	16.56	N/I	–	–	F
95	16.63	Furfuryl pentanoate	81 (B), 98, 182	784	C, D, F
96	16.845	N/I	–	–	A, C, D, E, F
97	17.41	N/I	–	–	C, D, E
98	17.56	N/I	–	–	C
99	17.72	N/I	–	–	D, F
100	17.82	N/I	–	–	A, C, D, F
101	17.965	N/I	–	–	C, D
102	18.055	Furfuryl disulfide	81 (B), 53, 161	793	A–F
103	18.18	4-Ethyl guaiacol	137 (B), 152	906	B, C, E, F
104	18.315	N/I	–	–	C
105	18.522	N/I	–	–	C, D, F
106	18.59	N/I	–	–	B, C, D, F
107	18.77	Difurfuryl ether	81 (B), 53, 39, 95, 69	901	F
108	18.87	N/I	–	–	B, C, E, F
109	19.115	4-Vinylguaiacol	77 (B), 135, 150, 107, 51	907	C, E
110	19.19	N/I	–	–	C
111	19.39	N/I	–	–	B
112	20.57	N/I	–	–	C
113	20.96	N/I	–	–	A–E
114	21.22	N/I	–	–	C
115	21.7	N/I	–	–	A–F

A. Acidity; B. Bitterness; C. Flavor; D. Cleanliness; E. Body; F. Overall quality; N/I. Not identified.

The samples were stored at −5 °C for a maximum of 3 h before being ground in a mortar and analyzed.

2.2. Sensory analysis

The coffee samples for cupping were prepared using 10 g of roasted and ground coffee in 100 mL of hot water [11]. The cup quality was assessed according to acidity, bitterness, flavor, cleanliness (clean cup), body (mouthfeel) and overall quality. The quality and intensity of each attribute were evaluated simultaneously using a scale varying from 1 to 5, as presented in Ribeiro et al. [9].

All the samples were submitted to a sensory evaluation by five experts. A 2-way ANOVA with all the scores of the five experts and a paired *t*-test between each expert (10 tests for each attribute) was able to verify the experts who did or did not clash for each attribute under consideration. Experts who had high variance in relation to others were removed and averages were used as "main scores" for the six attributes that were built using only the attributes from the experts who had similarity.

Thus, for each attribute score, 1 was conferred for a very low quality (meaning a more intense perception in the case of an undesirable sensation or a less intense perception if it was a desirable sensation). When acidity is analyzed, for example, 1 may refer either to low acidity or to high and undesirable acidity, such as that due to microbial fermentation. Thus, acidity with score 1 is worse than acidity with score 5. On the other hand, when bitterness was evaluated, a score of 1 referred to a bitterer beverage than one that received a score of 5. In this case, score 5 was awarded to a beverage with normal and pleasant bitterness [9].

The scales were arbitrarily defined, and the diversity of the samples was considered regarding each attribute in order to define them. Thus, according to the experts, the scores for body, acidity, and bitterness of the samples varied between low, normal, and high, while the diversity for overall quality, flavor, and cleanliness was much higher. This happens because small differences in the intensity of the latter attributes could be easily perceived by the experts. On the other hand, acidity, bitterness, and body possess only three distinct points (low, normal, high-…weak, regular, strong).

2.3. SPME devices and CG-FID parameters

SPME fibers coated with 65 μm thick polydimethylsiloxane/divinylbenzene (PDMS/DBV) and the manual holder were purchased from Supelco (Bellefonte, PA). The fibers were conditioned according to the SPME data Sheet (T7941231) from Supelco in the GC injector port. The analyses were performed on a G-6850 GC-FID system (Agilent, Wilmington, DE) fitted with a HP-5 capillary column (30 m × 0.25 mm × 0.25 μm). Helium (1 mL min^{-1}) was the carrier gas. The limit of detection (LD) of the FID was 10^{-12} g. The oven temperature was programmed as follows: 40 °C→5 °C/min→150 °C→30 °C/min→260 °C. The injection port was equipped with a

0.75 mm i.d. liner, and the injector was maintained at 220 °C in the splitless mode. Under these conditions, no sample carryover was observed on blank runs conducted between extractions. Identification of the extracted analytes was performed on an HP-5890 gas chromatographer (Hewlett-Packard, Wilmington, DE, USA) equipped with a HP-5973 mass-selective detector (scan mode—LD 10^{-12} g) fitted with the same column and operated under the same conditions as the GC-FID. The interface and detector (ion source) temperatures used were 240 °C and 200 °C, respectively. GC–MS data treatment was carried out 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). Also, comparisons were made with earlier reports on the volatile compounds of roasted coffee [8,12].

2.4. General SPME procedure for sampling and injection

The conditions adopted for the SPME extractions were chosen according to the optimization found in Ribeiro et al. [12]. Ground coffee (250 mg) and 2 ml of saturated aqueous sodium chloride solution were transferred to a septum-sealed glass sample vial (5 mL). After 10 min of sample/headspace equilibration under agitation of 900 rpm at 42.5 °C, the fibers were exposed to the sample headspace for 22 min. After sampling, the fiber was immediately placed in the injection port of the GC, and the analytes were thermally desorbed at 220 °C. All analyses were carried out in triplicate.

2.5. Chemometric data treatment

The original chromatographic profiles were organized into a matrix format **X** ($I \times J$), where each replicate represented one sample. Data analysis was carried out using Matlab 6.5 software (The MathWorks, Co., Natick, MA, USA) using the computational package PLS_Toolbox (Eigenvector Research, Inc.—PLS_Toolbox version 3.02) [13].

The data analysis was performed using the entire chromatogram, as with an infrared spectrum, and not using relative peak areas. Five different pretreatments were applied to the matrix **X** of chromatograms; these pretreatments are described in Ribeiro et al. [8]. Variable selection was carried out by the ordered predictors selection (OPS) method [14].

Partial least squares (PLS) was the regression method used for modeling. More information on the regression method can be found in Ferreira et al. [15] and Ribeiro et al. [8].

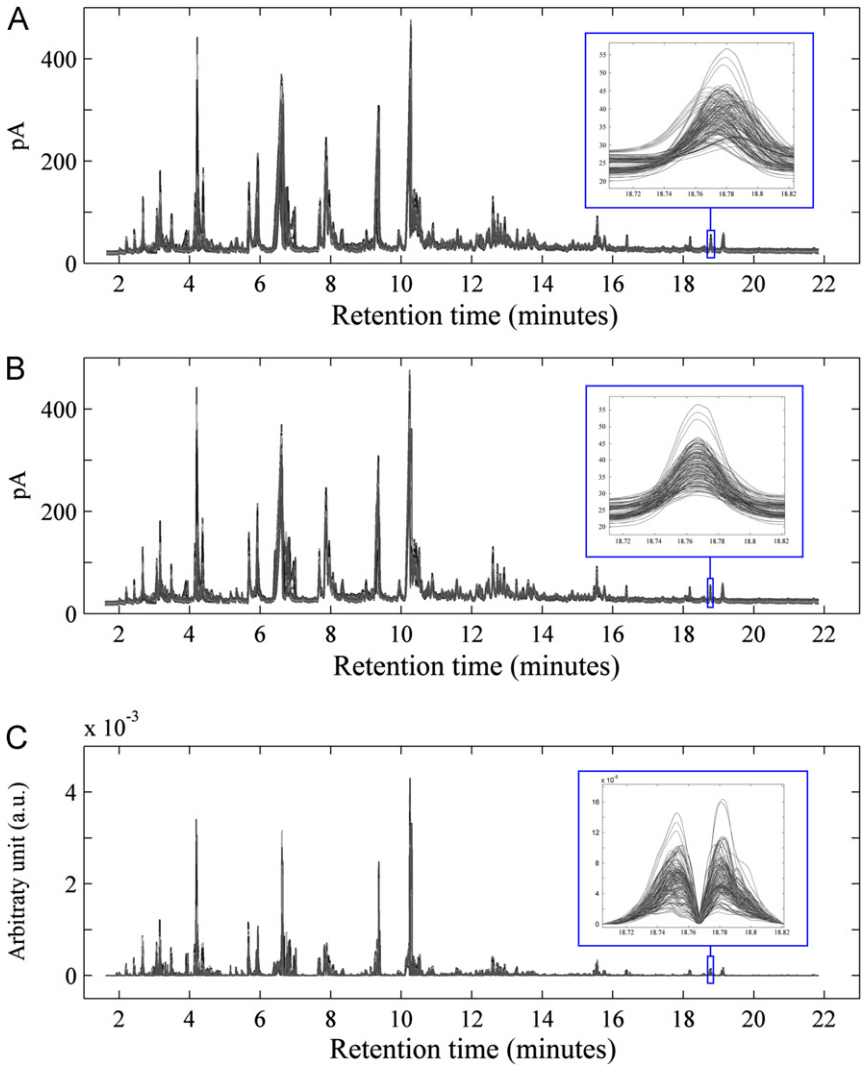

Fig. 1. Original chromatograms (A), after alignment (B) and after being fully pre-treated (C). The regions selected and expanded show the alignment of the peak before and after the alignment algorithm COW (B) and then smoothed and with the first derivative function (C).

3. Results and discussions

3.1. Mass detections of volatile compounds

Four different samples with distinct sensorial characteristics were analyzed by mass spectrometry for volatile compound identifications. In these analyses, more than 260 compounds were detected. Table 1 presents only the 115 used in one or more PLS regression models as important peaks for predicting the sensory attributes. Table 1 also shows the fragmentation of the compounds, the math number, and the model used.

3.2. Pre-treatment of the chromatographic data

Initially, the chromatograms of the original Arabica coffee samples were obtained in triplicate and overlaid (Fig. 1). In Fig. 1, and in its expanded region, it can be seen that the original data needed pre-treatment, such as peaks alignment (B) and baseline correction (C) before the construction of the PLS models. Thus, subsequent calculations were performed with the pre-treated data matrix ($\mathbf{X}_p$).

3.3. Regression models

To build the regression models for the six sensory attributes (acidity, bitterness, flavor, cleanliness, body, and overall quality), the mean values of the notes indicated by the cuppers were used as the dependent variables ($\mathbf{y}$) and 159 chromatograms referring to 53 Arabica coffee samples were used as independent variables (matrix $\mathbf{X}$).

The calibration sets consisted of 43 randomly selected samples (129 chromatograms). The 10 remaining samples, corresponding to 30 chromatograms, were used to form the external validation set. Leave five out cross-validation was the method used to select the number of components in the models. In this case, three replicates of five samples were left out at a time.

The variable selection for the construction of the models was carried out by the OPS method in the pretreated data matrix (159×20640) without baseline regions. In this way, from an initial set of 20,640 variables, 1732 were selected for the construction of the acidity model (A), 1515 for bitterness (B), 2783 for flavor (C), 1902 for cleanliness (D), 2223 for body (E), and 2179 for overall quality (F). These variables are indicated as vertical lines in Fig. 2.

When building the PLS models, the leverage vs. Student residuals plot was examined for outlier detection. For all the six sensory attributes investigated, a few replicates presented high values of leverage, while others had high values of residue. However, since no replicates presented high values of leverage and residue simultaneously, no sample was considered atypical.

Table 2 shows the number of latent variables selected for each sensory attribute prediction model and the respective statistical parameters of root mean square error of calibration, root mean square error of cross validation (RMSECV) and correlation coefficient of cross validation (r_{cv}).

Using the number of latent variables (Table 2) for all the calibration models, it was possible, in general, to describe 95% and 45% of the variance used in blocks **Y** and **X**, respectively.

Table 2
Latent variable numbers, RMSECV and r_{cv}, for PLS models.

Models	No. LV	RMSECV	r_{cv}
Acidity	5	0.27 ± 0.01	0.83 ± 0.01
Bitterness	4	0.33 ± 0.02	0.89 ± 0.02
Flavor	5	0.26 ± 0.01	0.95 ± 0.00
Cleanliness	5	0.36 ± 0.01	0.92 ± 0.01
Body	4	0.26 ± 0.01	0.89 ± 0.01
Overall quality	5	0.38 ± 0.01	0.92 ± 0.00

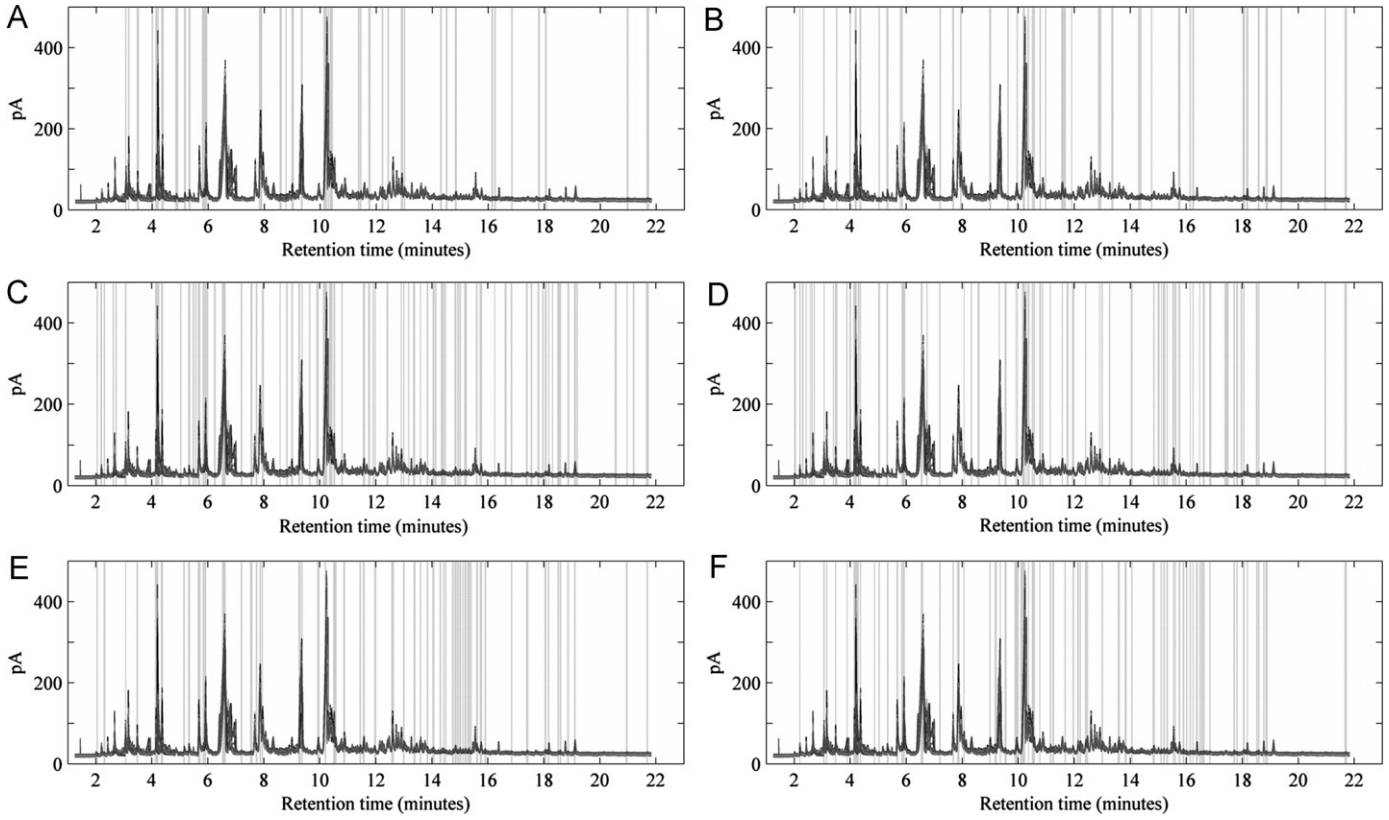

Fig. 2. Peaks selected by the OPS method for the regression models. Acidity (A), bitterness (B), flavor (C), cleanliness (D), body (E), and overall quality (F).

The models were validated by the external data set (10 samples). Fig. 3 shows the experimental values for each sensory attribute vs. the respective values estimated from cross validation. The predicted values for the external validation samples were also included in this figure to show that they are in the same range as the other samples. The values supplied by the experts, compared to those predicted by the PLS models for the 10 samples used in the external validation step, are shown in Table 3. The RMSEP (root mean square error of prediction) values were 0.27 for acidity, 0.33 for bitterness, 0.33 for flavor, 0.41 for cleanliness, 0.34 for body and 0.35 for overall quality.

The literature describes approximately 900 volatile compounds found in coffee [16]. These descriptions have been made since 1880, when Bernheimer identified the first volatile compound from roasted coffee [17].

In a particular volatile chromatographic analysis from a roasted coffee sample, the amount of compound may be smaller or larger than in another sample. This is due to several factors, mainly coffee species, region in which it was grown, processing method, roasting degree, extraction time, and extraction method.

Since the SPME technique used for volatile extraction is based on fibers with different sorbent materials (polar and nonpolar), it is expected that the compounds extracted by a given fiber are those that have more affinity for it.

According to the literature [18], the fiber used in this work, PDMS/DVB, is optimal for the efficient extraction of volatile

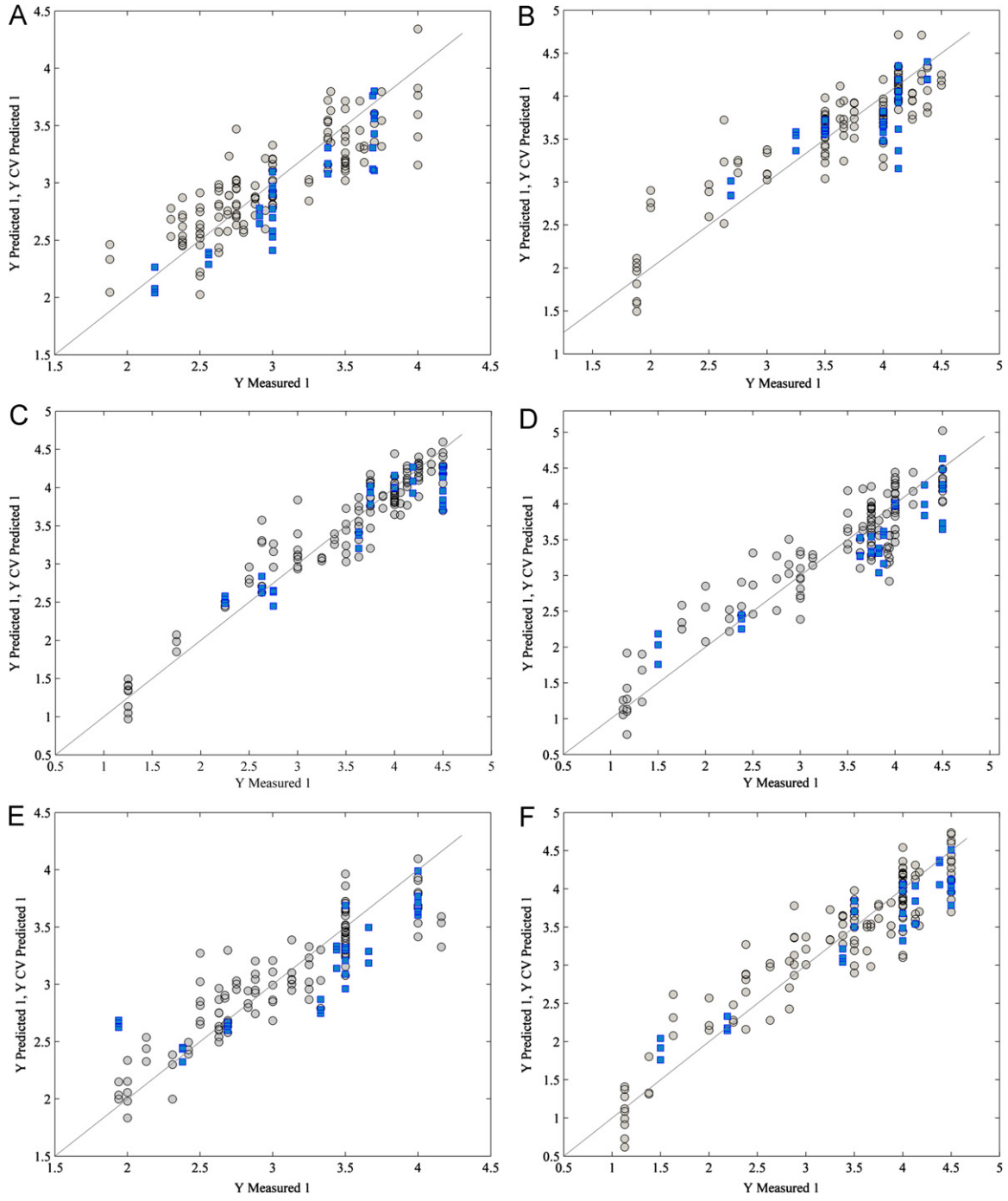

Fig. 3. Plots of measured versus predicted samples in calibration (○) and prediction (■) sets. Acidity (A), bitterness (B), flavor (C), cleanliness (D), body (E) and overall quality (F).

Table 3
Measured values given by the experts and predicted values from the regression models.

	Acidity		Bitterness		Flavor	
Samples	**Measured**	**Predicted**	**Measured**	**Predicted**	**Measured**	**Predicted**
1	3.00 ± 0.43	2.75 ± 0.16	4.13 ± 0.63	3.38 ± 0.23	2.75 ± 0.29	2.58 ± 0.11
2	3.70 ± 0.39	3.65 ± 0.13	4.38 ± 0.75	4.26 ± 0.12	4.19 ± 0.52	4.09 ± 0.17
3	3.69 ± 0.39	3.40 ± 0.32	4.13 ± 0.63	3.99 ± 0.06	3.63 ± 0.48	3.33 ± 0.11
4	3.38 ± 0.50	3.18 ± 0.11	4.13 ± 0.63	4.30 ± 0.09	3.75 ± 0.29	3.91 ± 0.12
5	2.56 ± 0.43	2.35 ± 0.05	2.69 ± 0.63	2.90 ± 0.10	2.25 ± 0.00	2.53 ± 0.05
6	2.19 ± 0.50	2.13 ± 0.11	3.25 ± 0.50	3.50 ± 0.12	2.63 ± 0.48	2.71 ± 0.11
7	2.91 ± 0.39	2.71 ± 0.06	3.50 ± 0.50	3.57 ± 0.02	4.50 ± 0.48	3.89 ± 0.28
8	3.00 ± 0.25	2.54 ± 0.14	3.50 ± 0.63	3.65 ± 0.06	4.00 ± 0.21	4.05 ± 0.09
9	3.70 ± 0.30	3.38 ± 0.25	4.00 ± 0.75	3.68 ± 0.13	4.50 ± 0.52	4.11 ± 0.14
10	3.00 ± 0.25	3.00 ± 0.08	4.00 ± 0.63	3.61 ± 0.12	4.50 ± 0.65	4.12 ± 0.24
	Cleanliness		**Body**		**Overall quality**	
	Measured	**Predicted**	**Measured**	**Predicted**	**Measured**	**Predicted**
1	3.83 ± 0.29	3.24 ± 0.18	3.33 ± 0.58	2.79 ± 0.08	3.38 ± 0.48	3.12 ± 0.09
2	4.31 ± 0.38	4.03 ± 0.22	3.44 ± 0.38	3.33 ± 0.08	4.38 ± 0.25	4.26 ± 0.18
3	3.63 ± 0.48	3.36 ± 0.15	3.66 ± 0.58	3.32 ± 0.16	4.13 ± 0.25	3.81 ± 0.25
4	3.88 ± 0.14	3.45 ± 0.25	2.69 ± 0.31	2.71 ± 0.03	4.00 ± 0.00	3.49 ± 0.18
5	1.50 ± 0.20	1.99 ± 0.22	1.94 ± 0.24	2.61 ± 0.05	1.50 ± 0.20	1.91 ± 0.14
6	2.38 ± 0.43	2.37 ± 0.10	2.38 ± 0.43	2.36 ± 0.05	2.19 ± 0.13	2.22 ± 0.10
7	3.75 ± 0.29	3.40 ± 0.13	3.50 ± 0.48	3.12 ± 0.12	3.50 ± 0.41	3.68 ± 0.17
8	4.00 ± 0.31	3.99 ± 0.02	3.50 ± 0.25	3.45 ± 0.23	4.00 ± 0.25	4.03 ± 0.05
9	4.50 ± 0.24	3.88 ± 0.34	4.00 ± 0.50	3.53 ± 0.03	4.50 ± 0.25	3.92 ± 0.12
10	4.50 ± 0.41	4.44 ± 0.22	4.00 ± 0.41	3.85 ± 0.09	4.50 ± 0.50	4.25 ± 0.23

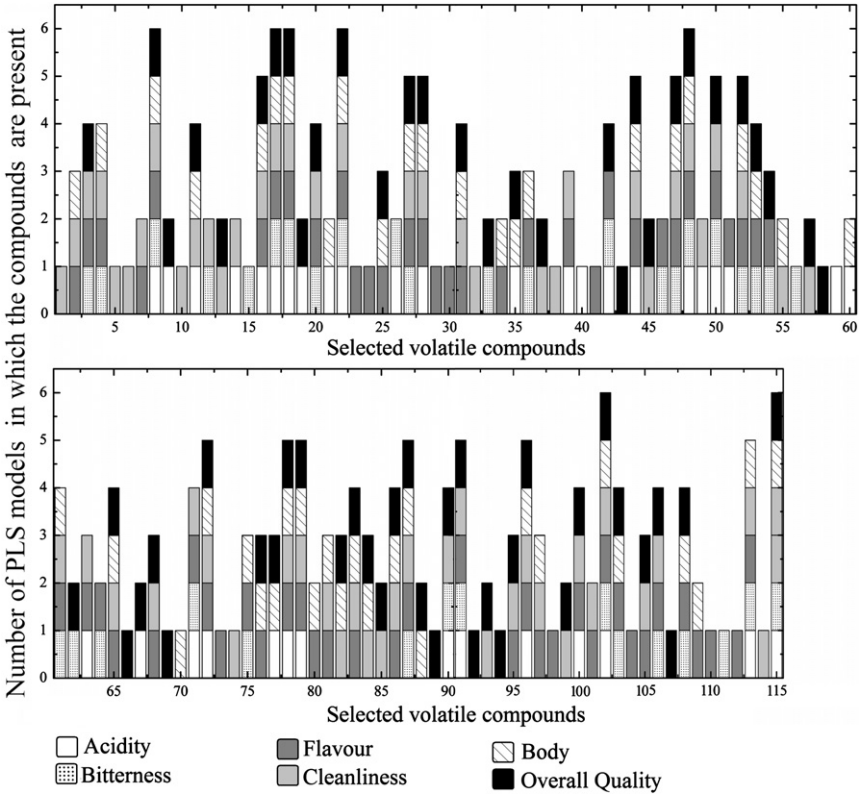

Fig. 4. Distribution of volatile compounds listed in Table 1, in accordance with the PLS model(s).

organic compounds (in the range of C2–C12) with intermediate polarity, such as nitro/amines, and some aldehydes and ketones, among others. Thus, hydrocarbons, carboxylic acids and alcohols are less extracted by this coating. Based on these aspects, the discussion of the compounds that were important for the construction of the regression models is restricted to volatile compounds extracted by the fiber used and detected by mass spectrometry.

Given that the overall quality of the roasted coffee depends on various sensory attributes (such as flavor, acidity, and body, among others) it is expected that some compounds indicated in Table 1 are important for most sensory attributes evaluated.

Fig. 4 summarizes the influential volatile compounds distributed between the PLS models. Seven volatile compounds (6.1%) are important in the construction of all models (**8**, **17**, **18**, **22**, **48**, **102** and **115**). Another fourteen (12.2%) were used in five models (**16**, **27**, **28**, **44**, **47**, **50**, **52**, **72**, **78**, **79**, **87**, **91**, **96** and **113**), and 18 more compounds (15.6%) were important for the construction of 4 regression models (**3**, **4**, **11**, **20**, **31**, **42**, **53**, **54**, **61**, **65**, **71**, **83**, **86**, **90**, **100**, **103**, **106** and **108**). Therefore, 34% of the peaks were selected for at least four regression models, showing that most of these compounds are important in providing these sensorial features.

However, sensorial analysis, according to experts, describes a much closer relationship between the attributes flavor, cleanliness, and overall quality. Thus, it is expected that a coffee well evaluated in terms of overall quality generally presents a high evaluation of cleanliness and flavor. This can be seen in Fig. 2, where 33 selected volatile compounds are part of the regression models of the three attributes.

Most of the compounds indicated in Table 1 for overall quality, flavor, and cleanliness in PLS models are cited in Ribeiro et al. [10] as important markers for discriminating coffee samples by Principal Component Analysis. Among them are pyrrole, 1-methyl pyrrole, 2,4-hexanedione, dihydro-2-methyl-3(2H)-furanone, furfural, 2-ethyl-5-methylpyrazine, 2-ethenyl N-methylpyrazine, and 5-methyl 2-propionylfuran.

Among the compounds selected to calibrate flavor, five of them are cited in the literature as potent odorants for roasted coffee (3-methylbutanal, 2,3-pentanedione, 4-vinylguaiacol, methanethiol and 1-methyl pyrrole) [19,20]. The compounds 3-methylbutanal, 2,3-pentanedione and 4-vinylguaiacol are found in high quality roasted coffees [10], while a greater amount of methanethiol is found in low-graded roasted coffees [10]. According to Agresti et al.[6] and Ribeiro et al. [10] 1-methyl pyrrole was found in coffees containing defective beans, bringing negative notes to the flavor attribute.

The PLS model constructed for overall quality confirmed some of the results presented in Ribeiro et al. [8]. Wherein the compounds furfural and 5-methyl-2-furancarboxyaldehyde appear as potential positive markers of this attribute, pyridine and ethenyl pyrazine indicate low quality beverages.

The four PLS models constructed in Ribeiro et al. [8] were reconstructed in this work with new samples and experts and the similar results obtained have increased the reliability of the correlation between coffee volatiles and quality attributes.

Comparing the PLS models based on NIR spectroscopy [9] with those created by chromatographic data, it can be noted that the PLS models from chromatographic data have better prediction power (low RMSECV and RMSEP). However, it is known that NIRS and SPME-GC are different analytical techniques that provide complementary chemical information. The advantage of SPME-GC is the fact that one can focus on pre-selected volatile compounds and, by using their intensity ratios, infer about the quality of new coffee samples.

4. Conclusions

The PLS regression models generated from the chromatographic profiles of roasted Arabica coffee volatiles adequately predicted the notes of acidity, cleanliness, overall quality, bitterness, body, and flavor of the beverage. This work confirmed previous results, showing a high linear relationship between the marks awarded by judges for sensory attributes and volatile compounds found in certain flavor profiles of roasted Arabica coffees. These results, together with the previous ones, show the reliability of using gas chromatography in the monitoring of several important sensory attributes related to quality of coffees from Brazil and other coffee-producing countries.

The prediction errors of these models, using 4 or 5 latent variables, were equal to 0.28, 0.33, 0.35, 0.33, 0.34, and 0.41, for each of the attributes and are better than those published and compatible with the scoring errors among experts.

In the case of the PLS models for the two new attributes studied, it should be mentioned that the compounds 2-methyl butanal, 3-methyl pyridazine, 2,3-hexanedione, and furfuryl methyl ether are important for coffee acidity, while compounds such as furfuryl formiate, ethyl pyrazine, 4-pyridazinamide, and n-methyl phenol are important in bitterness prediction.

Acknowledgments

This work was supported by Grants from CAPES and FAPESP. The authors acknowledge Prof. Dr. Carol H. Collins for English revision.

References

[1] 〈http://www.abic.com.br〉.
[2] 〈http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_tc_browse.htm?commid=47950〉.
[3] J.G. Cortez, Cafeicult. Mod. 1 (1988) 31.
[4] A.M. Feria-Morales, Food Qual. Pref. 13 (2002) 355.
[5] I. Esteban-Diez, J.M. Gonzalez-Saiz, C. Pizarro, J. Near Infrared Spectrosc. 12 (2004) 287.
[6] P.D.C. Agresti, A.S. Franca, L.S. Oliveira, R. Augusti, Food Chem. 106 (2008) 787.
[7] A.T. Toci, A. Farah, Food Chem. 108 (2008) 1133.
[8] J.S. Ribeiro, R.F. Teófilo, F. Augusto, T.J.G. Salva, M.M.C. Ferreira, Anal. Chim. Acta 634 (2009) 172.
[9] J.S. Ribeiro, M.M.C. Ferreira, T.J.G. Salva, Talanta 83 (2011) 1352.
[10] J.S. Ribeiro, F. Augusto, T.J.G. Salva, M.M.C. Ferreira, Quím. Nova 33 (2010) 1897.
[11] 〈http://www.agricultura.gov.br〉.
[12] J.S. Ribeiro, R.F. Teófilo, F. Augusto, M.M.C. Ferreira, Chemom. Intell. Lab. Syst. 102 (2010) 45.
[13] B.M. Wise, N.B. Gallagher, R. Bro, J.M. Shaver, W. Windig, R.S. Koch, PLS Toolbox 3.5, for Use With MatlabTM, Eigenvector Research, 2004.
[14] R.F. Teófilo, J.P.A. Martins, M.M.C. Ferreira, J. Chemom. 22 (2009) 32.
[15] M.M.C. Ferreira, A.M. Antunes, M.S. Melgo, Quim. Nova 22 (1999) 724.
[16] L.M. Nijssen, C.A. Visscher, H. Maarse, L.C. Willemsens, M.M. Boelens, Volatile Compounds in Food: Qualitative and Quantitative Data, 7th ed., TNO Nutrition and Food Research Institute, Zeist, The Netherlands, 1996.
[17] O. Bernheimer, Monatsh. Chem. 1 (1880) 456.
[18] H. Kataoka, H.L. Lord, J. Pawliszyn, J. Chromatogr. A 880 (2000) 35.
[19] P. Semmelroch, W. Grosch, J. Agric. Food Chem. 44 (1996) 537.
[20] M. Czerny, W. Grosch, J. Agric. Food Chem. 48 (2000) 868.