Microchemical Journal 213 (2025) 113771



Contents lists available at ScienceDirect

Microchemical Journal

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





SPME-GC–MS and chemometrics for coffee characterization, classification and authentication

Nerea Núñez [a,b,*], Erica Moret [c], Paolo Lucci [d], Sabrina Moret [c], Javier Saurina [a,b], Oscar Núñez [a,b,e]

[a] Department of Chemical Engineering and Analytical Chemistry, University of Barcelona. Martí i Franquès 1-11, E08028 Barcelona, Spain
[b] Research Institute in Food Nutrition and Food Safety, University of Barcelona, Recinte Torribera, Av. Prat de la Riba 171, Edifici de Recerca (Gaudí), Santa Coloma de Gramenet, E08921 Barcelona, Spain
[c] Department of Agricultural, Food, Environmental and Animal Sciences, University of Udine, Via delle Scienze 206, 33100 Udine, Italy
[d] Department of Agricultural, Food and Environmental Sciences, Polytechnic University of Marche, Via Brecce Bianche, 60131 Ancona, Italy
[e] Serra Húnter Fellow, Departament de Recerca i Universitats, Generalitat de Catalunya, Via Laietana 2, E08003 Barcelona, Spain

ARTICLE INFO

Keywords:
Coffee authentication
SPME
GC–MS fingerprinting
Non-targeted method
Chemometrics

ABSTRACT

In recent decades, the complexity of the food chain has contributed to a surge of food adulteration issues, resulting in numerous instances of food fraud. For this reason, ensuring the authenticity of food is crucial for society as a whole. In this context, beverages are particularly vulnerable to adulteration by adding flavors and aromas or incorporating unspecified substances to enhance volume, among other deceptive practices.

This work focuses on the detection of fraud in coffee, one of the world's most popular beverages, which is a product easily prone to manipulation. Fingerprinting studies of volatile compounds in 185 samples were performed by gas chromatography (with polar and non-polar columns) coupled to mass spectrometry (GC–MS) and in combination with chemometrics for data analysis. In this group of samples, 42 were chicory, 96 were coffee of different species and geographical production regions, and 47 were soluble coffees. Headspace-solid phase microextraction (HS-SPME) was employed to obtain the volatile compounds in the samples directly from the solid coffee. The GC–MS fingerprints served as reliablechemical descriptors for the classification of coffee samples using chemometrics. Moreover, some compounds found in samples were tentatively identified using NIST Research Libraries.

Furthermore, two adulteration coffee studies were performed using partial least squares (PLS) regression, which demonstrated the feasibility of the proposed methodology for the quantification of adulterant levels up to 15%, with calibration and prediction errors below 2.9% and 7.4%, respectively.

1. Introduction

The quality of natural products is a collective concern within society, including governments, food manufacturers, and researchers. Food quality control aims to protect the consumer and ensure that all food products are healthy and suitable for human consumption. Food products are highly complex matrices composed of a wide variety of natural compounds, in addition to substances derived from their processing or packaging, which makes the quality of these products important in terms of nutrition and food safety. Considering the complexity of the food chain and the involvement of various factors both in food production and consumption, food handling and adulteration practices are increasing, leading to cases of food fraud. In this context, beverages are food products that can be easily adulterated through incorrect labeling, supplementation with flavors or aromas, or addition of unspecified substances, among other practices [1–3].

In this field, coffee, one of the most popular beverages in the world, is a very easily adulterable drink. Coffee is an infusion made from roasted and ground beans of the coffee plant, possessing a distinctive taste and aroma [4–6]. The coffee plant is classified as *Coffea* with over seventy varieties and belongs to the *Rubiaceae* family, with *Coffea arabica* (arabica variety) and *Coffea canephora* (robusta variety) being the principal species. The arabica variety is most highly esteemed by consumers and considered superior to the robusta variety due to its superior

* Corresponding author.
E-mail address: nereant7@gmail.com (N. Núñez).

https://doi.org/10.1016/j.microc.2025.113771
Received 12 January 2025; Received in revised form 5 April 2025; Accepted 24 April 2025
Available online 24 April 2025
0026-265X/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

sensory properties. For this reason, arabica coffee normally has a higher price in the international market [7].

Coffee contains a large number of bioactive substances with antioxidant activity, recognized for its beneficial effects on health. Some studies link its consumption with the reduction of the development of some serious diseases such as type II diabetes, cancer and some cardiovascular diseases. Furthermore, coffee contains volatile compounds responsible for its characteristic aroma and significantly contribute to its sensory profile. The content of these substances and, consequently, the characteristics of coffee may vary depending on factors such as variety, origin, and climate, among others [4,6,8,9].

Unfortunately, coffee adulteration is on the rise, leading to cases of food fraud due to the coffee vulnerability and the complexity of the food chain. For that reason, determining the authenticity of coffees by analytical methodologies is crucial to ensure their quality [10].

Several studies in the literature have analyzed volatile compounds using analytical techniques. For instance, Claro et al. [11] employed spectrophotometry and chromatography techniques (Liquid Chromatography coupled to Ultraviolet–Visible, LC-UV–Vis, and Headspace Gas Chromatography coupled to Flame Ionization Detection, HS-GC-FID) combined with chemometrics to determine the compound profiles in green coffee beans and differentiate them into special or traditional categories. Similarly, Mannino et al. [12] utilized high-throughput metabolomics, including High-Performance Liquid Chromatography coupled to Diode-Array Detection and Tandem Mass Spectrometry (HPLC-DAD-MS/MS), Gas Chromatography coupled to Mass Spectrometry (GC–MS), and Polymerase Chain Reaction coupled to Restriction Fragment Length Polymorphism (PCR-RFLP) fingerprinting, to distinguish between different coffee species and origins, providing insights into the volatile profiles and genetic markers of *Coffea arabica* and *Coffea canephora*. In another study, Pasias et al. [13] developed a rapid GC-FID method to determine caffeine levels in coffee grains. Additionally, Gamal et al. [14] optimized GC–MS/MS and LC-MS/MS methods to detect organic contaminants in green and roasted coffee, identifying various pesticides and PCBs. In another work, the variability of volatile compounds in roasted coffee was explored by Caporaso et al. [15], who used SPME-GC–MS to analyze the flavor profiles of Arabica and Robusta beans, revealing significant intra-batch and inter-batch variability and the potential for classification based on geographical origin and species. Lastly, Zhang et al. [16] analyzed the flavor characteristics of cold brew coffee using GC–MS and electronic nose and tongue techniques, focusing on how roasting degrees and freeze-drying processes influence the aroma.

In this study, a non-targeted method of headspace-solid phase microextraction with gas chromatography coupled to mass spectrometry (HS-SPME-GC–MS) has been developed to achieve the characterization, classification, and authentication of different coffee samples according to geographical production region, and variety (arabica/robusta). Moreover, decaffeinated and non-decaffeinated instant coffee samples were analyzed. Some samples of chicory, a potential coffee adulterant, have also been considered. The GC–MS fingerprints have been utilized as descriptors of chemical information to address the characterization and classification of the analyzed coffees using principal component analysis (PCA), partial least squares-discriminant analysis (PLS-DA) and partial least squares (PLS) regression.

2. Materials and methods

2.1. Samples

A total of 185 commercially available samples (see Table S1 in supplementary material) were analyzed by gas chromatography (with polar and non-polar column) coupled to mass spectrometry (GC–MS). Out of these 185 samples, 42 were chicory (a typical coffee substitute or adulterant), 96 were coffee from three geographical production regions (Vietnam, Cambodia, and Costa Rica) and species (Arabica, Robusta, and Arabica-Robusta mixture), and 47 samples were soluble coffee (decaffeinated and non-decaffeinated). The chicory samples were purchased from Barcelona supermarkets (Spain) and the coffee samples were from Vietnam, Cambodia and Costa Rica local supermarkets.

Furthermore, a Cambodian Coffee sample was injected as the Quality Control (QC) at the beginning, the end, and three times throughout the corresponding sequence to assess the reproducibility of the method.

Moreover, some samples were also employed for adulteration studies. The adulteration cases proposed were Vietnamese Arabica coffee with Vietnamese Robusta coffee, and Vietnamese Robusta coffee with chicory. For both adulteration cases, two sample sets were prepared for calibration (to build the PLS regression model) and prediction (to validate the model), with various adulteration percentages as described in Fig. S1 (supplementary material). The calibration set included adulterant percentages of 0 % (pure coffee), 20 %, 40 %, 60 %, 80 %, and 100 % (i.e., chicory or Robusta coffee). The validation set included adulteration percentages of 15 %, 25 %, 50 %, 75 % and 85 %. Each blended adulteration level was prepared in triplicate, thus resulting in 33 samples for every proposed adulteration case.

2.2. Sample treatment: Headspace-solid phase microextraction (HS-SPME)

The sample treatment consisted of a headspace-solid phase microextraction (HS-SPME) using a divinylbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) fiber (Supelco, Merck KgaA, Darmstadt, Germany). In accordance with Kim et al. [17] no pretreatment was performed prior to the microextraction. Thus, the analysis was performed directly from solid samples. The solid coffee or chicory (ca. 0.5 g) was placed in a sealed vial of 20 mL tightly capped. The sample equilibration temperature was 50 °C, with an agitation time of 10 min at 500 rpm. Then, the fiber was exposed to the headspace of the sample vial at 50 °C for 10 min. The extracted volatiles were introduced into the gas chromatograph by fiber desorption at 250 °C for 1 min.

2.3. Instrumentation

Two GC–MS methods were employed to obtain the chromatographic fingerprints with a GC–MS instrument (Agilent Technologies, Santa Clara, CA, USA) equipped with gas chromatograph (model 7890B) and a selective mass quadripolar detector (model 5977A).

For the first methodology, the chromatographic separation was carried out in a polar UF-WAX ms (60 m × 0.25 mm, 0.25 μm) column from Agilent. For the other method, the chromatographic separation was performed in a non-polar DB-5 ms (30 m × 0.25 mm, 0.5 μm) column from Agilent. The polar column (UF-WAX ms) was used to separate polar compounds, while the non-polar column (DB-5 ms) captured non-polar compounds. This complementarity allowed for a more thorough evaluation of the volatile profiles under different chromatographic conditions and provided insight into how each column influences the detection and classification of coffee samples. In both cases, helium was the carrier gas. The temperature was initially maintained at 40 °C for 2 min, then increased to 200 °C at a rate of 3 °C/min until min 60; subsequently, it raised to 250 °C at 10 °C/min until min 70, and finally held at 250 °C for 5 min. An electron ionization (EI) source at 70 eV was employed. The temperatures of the ionization source and the quadrupole were 230 °C and 150 °C, respectively. Full scan data (m/z: 50–650) were acquired in positive acquisition mode.

The GC–MS conditions were based on a previously established methodology in the literature by Kim et al., optimized for the analysis of volatile compounds in food matrices. Preliminary tests were conducted to compare the split and splitless injection modes, with the splitless mode being ultimately selected due to its higher sensitivity for detecting minor volatile compounds. Once optimized for the UF-WAX column, the same conditions were applied to the DB-5 column to ensure consistency

between the methods.

2.4. Data analysis

Coffee and chicory samples were randomly analyzed with the proposed non-targeted HS-SPME-GC–MS methods. The obtained sample fingerprints were employed as chemical descriptors to build different data matrices to be analyzed by PCA, PLS-DA and PLS using the SOLO 8.6 chemometric software from Eigenvector Research (Manson, WA, USA) [18]. Details of the theoretical foundation of these statistical methodologies are discussed elsewhere [19]. The X-data matrices consisted of the GC–MS fingerprints, i.e., ion intensities at each *m/z* and retention time. The Y data matrices defined the sample classes in the PLS-DA and adulterant percentages in the PLS regression. Autoscaling was applied to GC–MS fingerprints to provide the same weight to each variable by suppressing differences in magnitude and amplitude. The optimal number of principal components (PCs) or latent variables (LVs) was estimated from the first significant minimum point of the cross-validation (CV) error through a Venetian blind approach. For validation of PLS-DA models, 70 % of the samples were utilized as the calibration set and the remaining 30 % as validation set. In the case of PLS regression, models were validated with the prediction sets (see Fig. S1).

3. Results and discussion

3.1. Non-targeted HS-SPME-GC–MS fingerprints

The main objective of this study was to obtain GC–MS fingerprints of different coffee and chicory samples for their classification and authentication. Two methodologies were proposed to obtain GC–MS fingerprints rich in volatile compounds.

Chromatographic separation was carried out using a polar UF-WAX ms column in the first method and a non-polar DB-5 column in the second method, both employing helium as a carrier gas. The optimal temperature conditions of the method were selected considering the detected signals for each sample type and the total analysis time.

The obtained GC–MS fingerprints, for both employed columns, constitute an intricate system, giving place to ca. 10,000 features per sample, reflecting the chemical diversity of the analyzed samples. Notable differences are observed (Fig. 1) between the chromatograms generated using the polar and non-polar column, highlighting how the selectivity of each column influences compound retention and resolution. Chromatograms from the polar column tend to show a more uniform distribution of peaks, with a higher number of signals spread across the retention time. In contrast, for the non-polar column, compounds are primarily observed between 2 and 40 min. In the Vietnamese Arabica coffee samples, Vietnamese Robusta coffee samples, Cambodian coffee samples, and Costa Rica coffee samples, high-intensity peaks are observed in similar intervals and small peak signals shown differences between samples. On the other hand, decaffeinated and regular soluble coffees exhibit simpler profiles with a lower density of peaks, likely reflecting the loss of volatile compounds during industrial processing. Chicory, in contrast, presents a distinctive chromatographic profile with specific signals that may be associated with its unique compounds. It is important to highlight that the GC–MS fingerprints are reproducible within samples of the same category, making them reliable chemical descriptors to address sample classification using chemometric methods.

3.2. PCA exploration

The potential of non-targeted GC–MS fingerprints as chemical descriptors was initially assessed through PCA. Data matrices (X-data) were constructed using the intensity signals at each *m/z* and retention time for all samples under analysis. Furthermore, an autoscaling preprocessing method was employed to ensure equal weighting for all variables.

Fig. 2 depicts the PCA score plots from non-targeted GC–MS fingerprints using both UF-WAX ms and DB-5 ms columns. For the polar column, Fig. a.1 shows all analyzed coffees vs. chicory samples, Fig. a.2 shows non-soluble coffees vs. soluble coffees, and Fig. a.3 shows soluble decaffeinated coffees vs. non-decaffeinated soluble coffees. Figs. b.1, b.2 and b.3 show the equivalent information for the non-polar column. Similar PCA information is provided for exploring coffees regarding the variety (Arabica, Robusta and Arabica-Robusta mixture) and the geographical origin region in Figs. S2 and S3 (supplementary material), respectively.

Regarding the distribution of the samples, the plots of scores (Fig. 2) show that chicory samples are satisfactorily discriminated from coffee samples when both GC methods are employed. As shown in Fig. 2.a.1 and 2.b.1, chicory and coffee samples are grouped according to their class, achieving better separation with the non-polar column (DB5). This separation is further emphasized by the fact that chicory and coffee samples are located in different regions of scores' plot, suggesting a significant chemical difference between the two categories. For coffee typology, Fig. 2.a.2 and 2.b.2 show a perfect discrimination when comparing soluble and non-soluble coffees for both GC–MS methods, indicating a clear separation based on volatolomics.

For the specific case of coffee samples analyzed with the WAX column, the samples are distributed along PC1, while for the DB5 column, the samples are mainly grouped according to PC2, in both cases located in the positive value region. However, in the case of soluble coffees (both regular and decaffeinated), the discrimination of the samples is less clear, as some overlap between the sample groups is observed, regardless of the GC–MS method used. This overlap suggests that the variability in soluble coffee is higher than in the ground (normal) counterpart, which could indicate greater heterogeneity in soluble coffee due to the manufacturing process.

Regarding the coffee classifying based on variety, Fig. S2 shows that coffee samples containing a mix of Arabica and Robusta varieties are located approximately at the center of the plot, positioned between pure Arabica and Robusta types as expected. The 100 % Arabica and 100 % Robusta samples show clear grouping, with a distribution reflecting opposite trends between them, particularly in the principal components PC1 and PC2. This distribution pattern confirms the chemical and profile differences between the two coffee varieties.

Finally, Fig. S3 shows that, although there is some overlap between samples, particularly between coffees from Vietnam and Cambodia, the samples are correctly grouped according to their geographic region of origin. This result supports the hypothesis that geographic origin significantly influences the chemical characteristics of the samples, which is reflected in the PCA's ability to group the samples effectively based on their origin. The presence of overlaps may be attributed to the possible similarity in the growing and processing conditions of coffees from these two regions, which could make their chemical profiles more similar.

3.3. PLS-DA classification

The non-targeted GC–MS fingerprints of the analyzed coffee samples were also subjected to a supervised PLS-DA to deal with classification and authentication. The same X-data matrices in PCA were used as the source of information; Y-data matrices were designed specifically for each case under study by coding each sample class attribute.

Fig. 3 depicts PLS-DA score plots obtained for both GC–MS methodologies, categorized by (i) coffee vs. chicory, (ii) non-soluble coffee vs. soluble coffee, and (iii) soluble decaffeinated coffee vs. soluble non-decaffeinated coffee. The PLS-DA results obtained for classifications based on coffee variety and the geographical origin region were shown in Fig. S4 and Fig. S5, respectively, in supplementary material. In addition, sensitivity, specificity and classification prediction error values by cross-validation for multiclass PLS-DA models are shown in Table S2 (supplementary material).

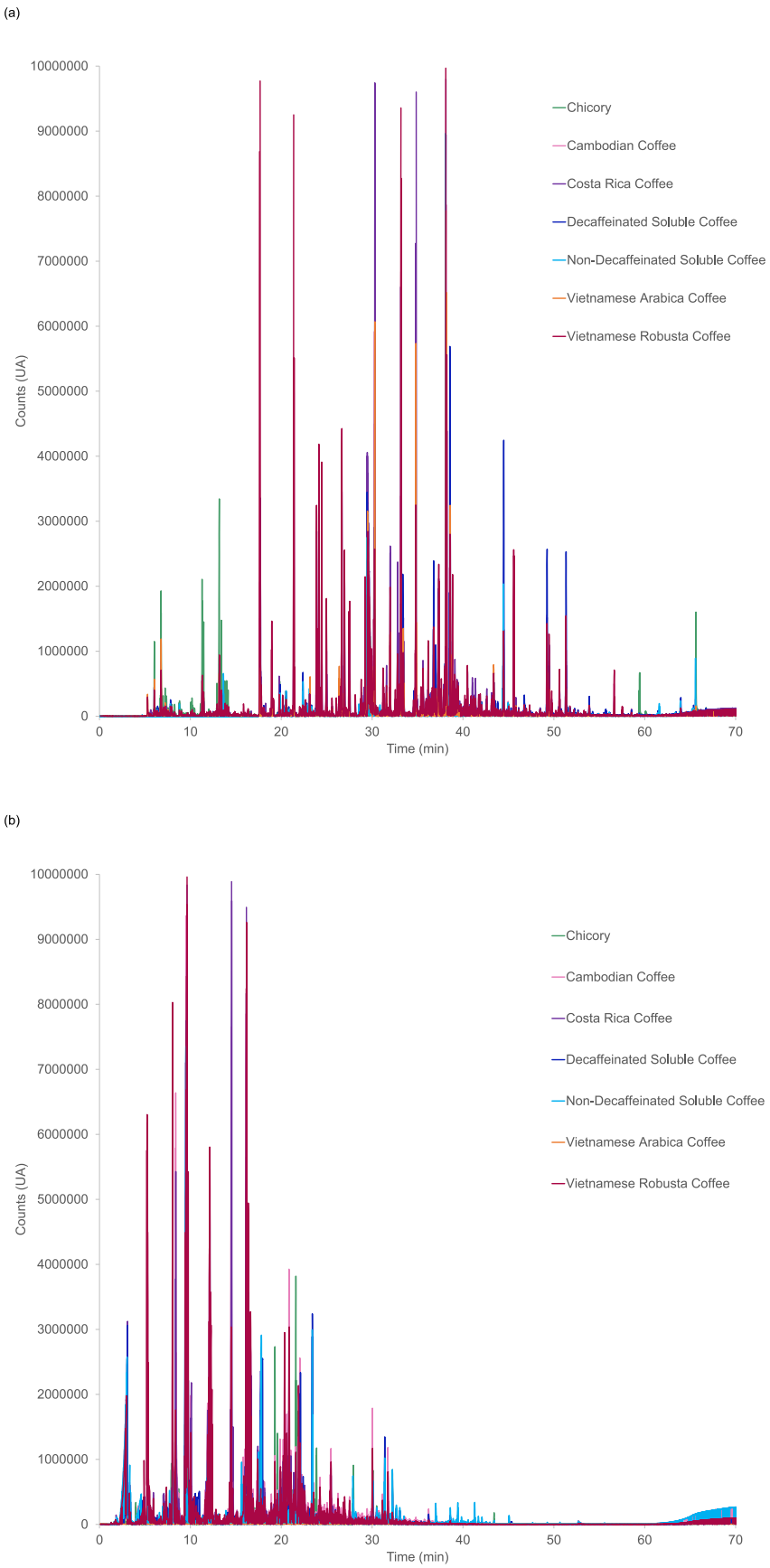


Fig. 1. GC–MS total ion chromatograms (a) polar column and (b) non-polar column for chicory, Vietnamese Arabica coffee, Vietnamese Robusta coffee, Cambodian Coffee, Costa Rica coffee, soluble non-decaffeinated coffee and soluble decaffeinated coffee.

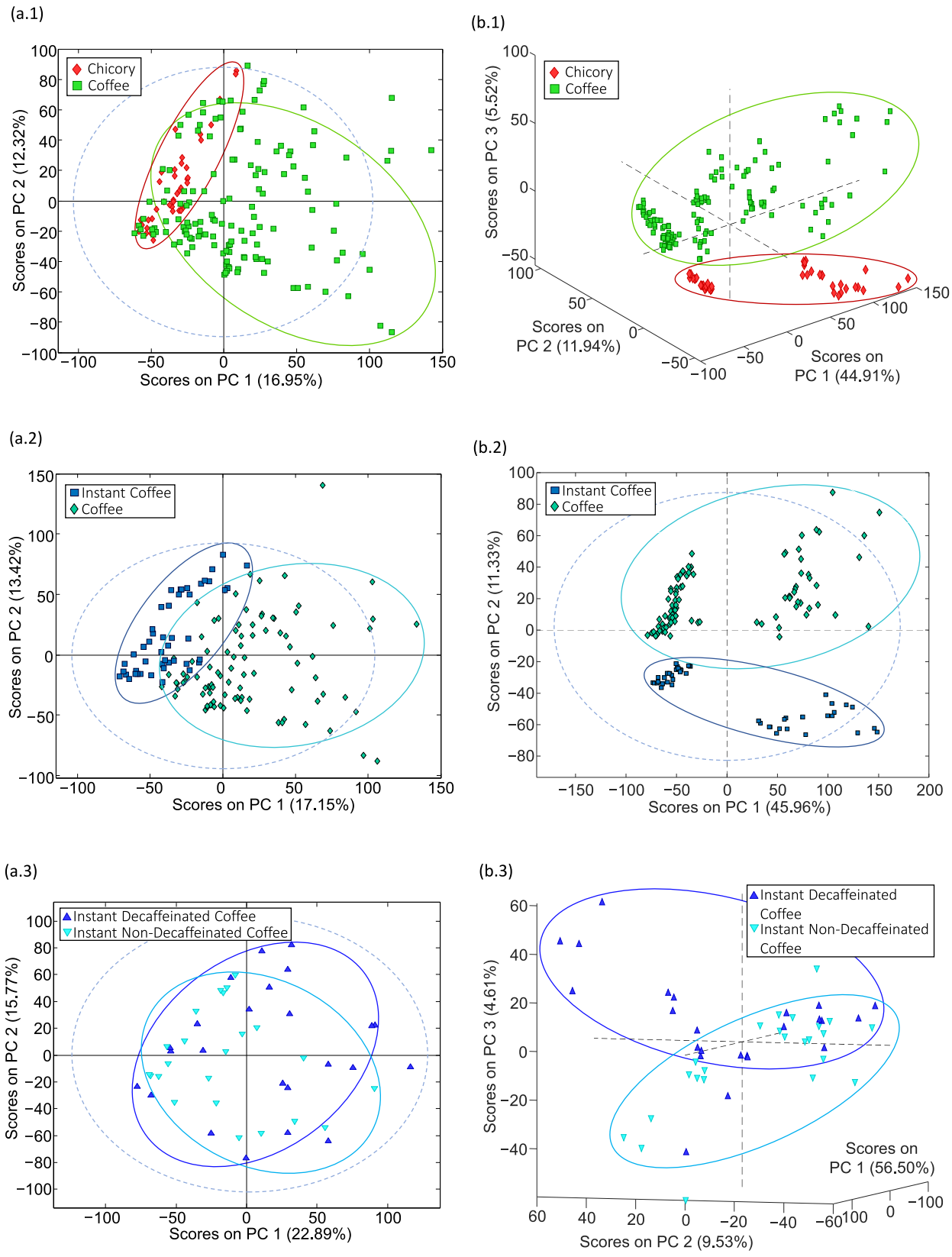


Fig. 2. PCA score plots using GC–MS fingerprints from (a) polar column and (b) non-polar columns. (a.1) Coffees vs. Chicory (score plot of PC1 vs. PC2), (a.2) Coffee vs. soluble Coffee (score plot of PC1 vs. PC2), and (a.3) Soluble non-decaffeinated Coffee vs. soluble decaffeinated coffee (score plot of PC1 vs. PC2); (b.1) All Coffee vs. Chicory (score plot of PC1 vs. PC2 vs. PC3), (b.2) Coffee vs. soluble Coffee (score plot of PC1 vs. PC2), and (b.3) Soluble non-decaffeinated Coffee vs. soluble decaffeinated coffee (score plot of PC1 vs. PC2 vs. PC3).

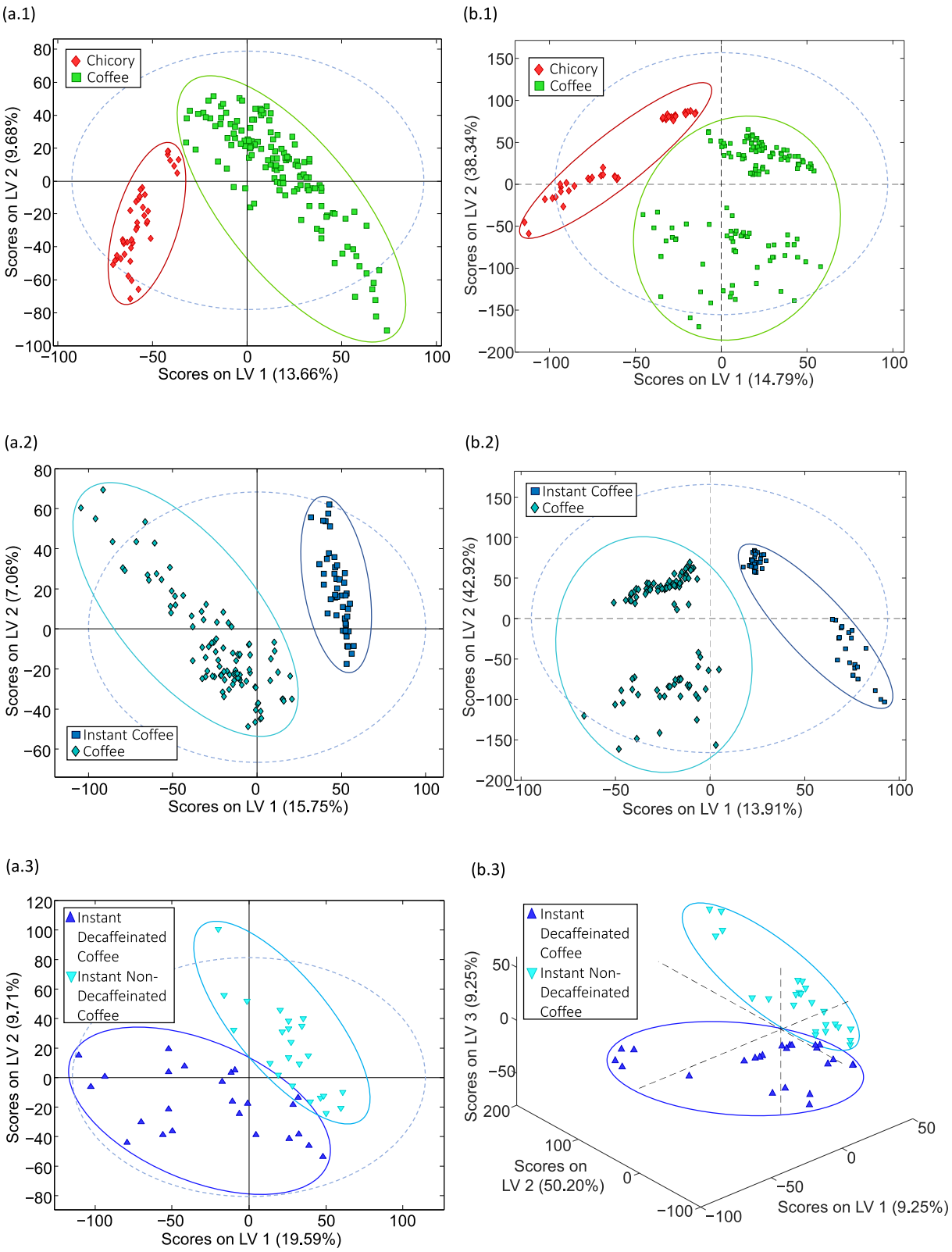


Fig. 3. PLS-DA score plots using GC–MS fingerprints from (a) polar column and (b) non-polar columns. (a.1) Coffees vs. Chicory (score plot of LV1 vs. LV2), (a.2) Coffee vs. soluble Coffee (score plot of LV1 vs. LV2), and (a.3) Soluble non-decaffeinated Coffee vs. soluble decaffeinated coffee (score plot of LV1 vs. LV2); (b.1) All Coffee vs. Chicory (score plot of LV1 vs. LV2), (b.2) Coffee vs. soluble Coffee (score plot of LV1 vs. LV2), and (b.3) Soluble non-decaffeinated Coffee vs. soluble decaffeinated coffee (score plot of LV1 vs. LV2 vs. LV3).

Perfect discrimination among chicory versus coffee was achieved with both GC–MS fingerprints; this is an interesting result as chicory is often employed as a coffee adulterant (Fig. 3.a.1 and Fig. 3.b.1). The chicory samples are distributed in the negative value region of LV1, while the coffee samples are located in the positive value region of LV1. Results were also quite satisfactory for ground vs soluble coffee (see Fig. 3.a.2 and Fig. 3.b.2), showing a good discrimination among both sample groups. For both methodologies, the ground coffees remain in the negative value region of LV1, while the instant coffees cluster in the positive value region of LV1. This separation suggests that the chemical profiles of instant and non-instant coffees are sufficiently distinct to be identified through PLS-DA analysis. Excellent results were also obtained when soluble decaffeinated and non-decaffeinated coffees were considered (Fig. 3a.3 and Fig. 3.b.3). This finding highlights the ability of the method to capture key chemical differences related to the decaffeination process, which is an important factor in the classification of the samples. This excellent sample discrimination performance is summarized in Table S2, with sensitivity and specificity values of 100 % for all the classifications studied in Fig. 3. The classification errors were 0 % for both polar and non-polar columns.

Furthermore, as shown in Fig. S4.a and Fig. S4.b (supplementary material), the classification based on varieties was also very satisfactory. As observed in the scores' plot, coffee samples containing a blend of Arabica and Robusta are located between the 100 % Arabica and 100 % Robusta samples for both methodologies employed. Robusta samples were positioned in the negative value region of LV2 and the Arabica samples in the positive value region of LV2. Sensitivity and specificity values were higher than 95 % and 95.8 %, respectively, with the polar column, and values higher than 93.3 % and 95.9 % with the non-polar column. Classification errors obtained were below 2.5 % for the polar column and below 3.3 % for the non-polar column.

Regarding the geographical production region, an excellent classification was also accomplished (see Fig. S5.a and Fig. S5.b, supplementary material). The sample groups clustered according to their studied geographical origin, with each group differentiated from the others for both methodologies. In this case, for the polar column, sensitivity and specificity values of 100 % were obtained. For the non-polar column, sensitivity values of 100 % and specificity values higher than 94.5 % were observed, with a classification error below 2.7 %.

Furthermore, to explore the potential of a multiclass classification model, as an example, a PLS-DA was conducted to group the coffee samples based on multiple attributes, including coffee variety (Arabica, Robusta) and geographic origin (Vietnam, Costa Rica, Cambodia). The results, shown in a PLS-DA score plot (Fig. S6 in the supplementary material), suggest that clustering based on these attributes is feasible, and such an approach could be considered for more comprehensive future studies.

For the polar column, the PLS-DA score plot (LV1 vs. LV2 vs. LV3) shows that the Costa Rican samples are positioned in the lower area of the plot, whereas those from Vietnam are located in the upper region (see Fig. S6.a). Additionally, Arabica coffee samples from Costa Rica and Vietnam are clustered on the right side of the plot, while Vietnamese Robusta samples are positioned on the left. The Vietnam Arabica-Robusta mixture samples are distributed between the 100 % Arabica and 100 % Robusta clusters. For the non-polar column (Fig. S6.b), the PLS-DA score plot (LV1 vs. LV2 vs. LV3) reveals that Costa Rican samples are situated on the left side of the plot, while Vietnamese samples are found on the right. Similarly, Arabica coffee samples from Costa Rica and Vietnam are grouped in the lower-left region, whereas Vietnamese Robusta samples are positioned in the upper-right region. The Vietnam Arabica-Robusta mixture samples are located in the lower-right area of the plot.

3.4. PLS-DA validation

Paired PLS-DA models, where a single sample class was compared with all the others, were evaluated to prove the viability of the proposed methodology to classify and predict coffees according to the sample type, origin region and variety. Each paired PLS-DA model was built using 70 % of samples randomly selected for each group as the calibration set while the remaining 30 % of the samples were employed as the prediction set.

Fig. 4 shows the class prediction plots of both calibration and prediction steps for the type of coffee sample analyzed. Figs. S7 and S8, in supplementary material, show similar information regarding the coffee varieties and the geographical production regions, respectively. Moreover, the optimal number of LVs, as well as the sensitivity, specificity and classification error for both calibration and prediction steps, for each paired PLS-DA classification model, were summarized in Table 1.

As shown in Table 1, satisfactory validation results were obtained for the classification of coffee samples by paired PLS-DA models for the two GC–MS methodologies.

When coffees were compared with chicory, sensitivity and specificity values of 100 % for calibration and prediction were obtained, regardless of the GC–MS methodology employed. The same results were attained when coffees were compared with instant coffees. In contrast, when decaffeinated and non-decaffeinated instant coffees were addressed, classification errors for prediction of 0 % and 7.1 % for the polar and non-polar columns, respectively, were obtained.

When studying the coffee variety, sensitivity and specificity values of 100 % were obtained with the polar column for both calibration and prediction, while with the non-polar column, values higher than 93.3 % were observed. For both calibration and prediction, classification errors were 0 % for the polar column. For the non-polar column, the classification errors for calibration were 0 %, while for the prediction, the classification errors were below 3.3 %.

Finally, sensitivity and specificity values for the geographical production regions were 100 % in both calibration and prediction when the Vietnam and Cambodia samples were evaluated against the other samples. In the case of Costa Rica, sensitivity was 100 % while specificity was higher than 98 % and 95.7 % for calibration and prediction, respectively. Classification errors were always below 0.9 % for the paired PLS-DA calibration and lower than 2.2 % for prediction.

The results obtained indicate that the GC–MS fingerprinting methods proposed appear to be suitable chemical descriptors for the characterization, classification and authentication of the analyzed coffee and chicory samples.

3.5. Volatile compound identification

The composition of volatile compounds of coffee is intricate and plays a significant role in the coffee aroma, flavor and overall sensory profile. Volatile compounds are abundant in coffee regardless of its variety, geographical origin region, or other features. The peaks separated by HS-SPME-GC/MS were identified using the NIST Research Library, and the experimental retention index (RI) of the compounds were compared with the theoretical RI values from the NIST database. The volatile compounds detected through HS-SPME-GC/MS using polar and non-polar columns are summarized in Table 2. Vietnamese Arabica Coffee exhibited 24 identified compounds using the WAX (polar) column and 20 with the DB5 (non-polar) column, while Vietnamese Robusta Coffee showed 25 and 22 compounds, respectively. When analyzing both Vietnamese Arabica and Robusta Coffee mixtures, 29 compounds were identified with the WAX column and 28 with the DB5 column. Cambodian Coffee presented 27 compounds with WAX and 21 with DB5, whereas Costa Rica Coffee had 30 and 22 compounds, respectively. Costa Rica Arabica Coffee contained 23 compounds identified with WAX and 22 with DB5. In the case of Soluble Non-Decaffeinated Coffee, 21 compounds were found using WAX and 20 with DB5. Finally, Decaffeinated Coffee exhibited 25 compounds with WAX and 20 with DB5. The results reveal significant differences in the aromatic profile and the compounds identified tentatively using the

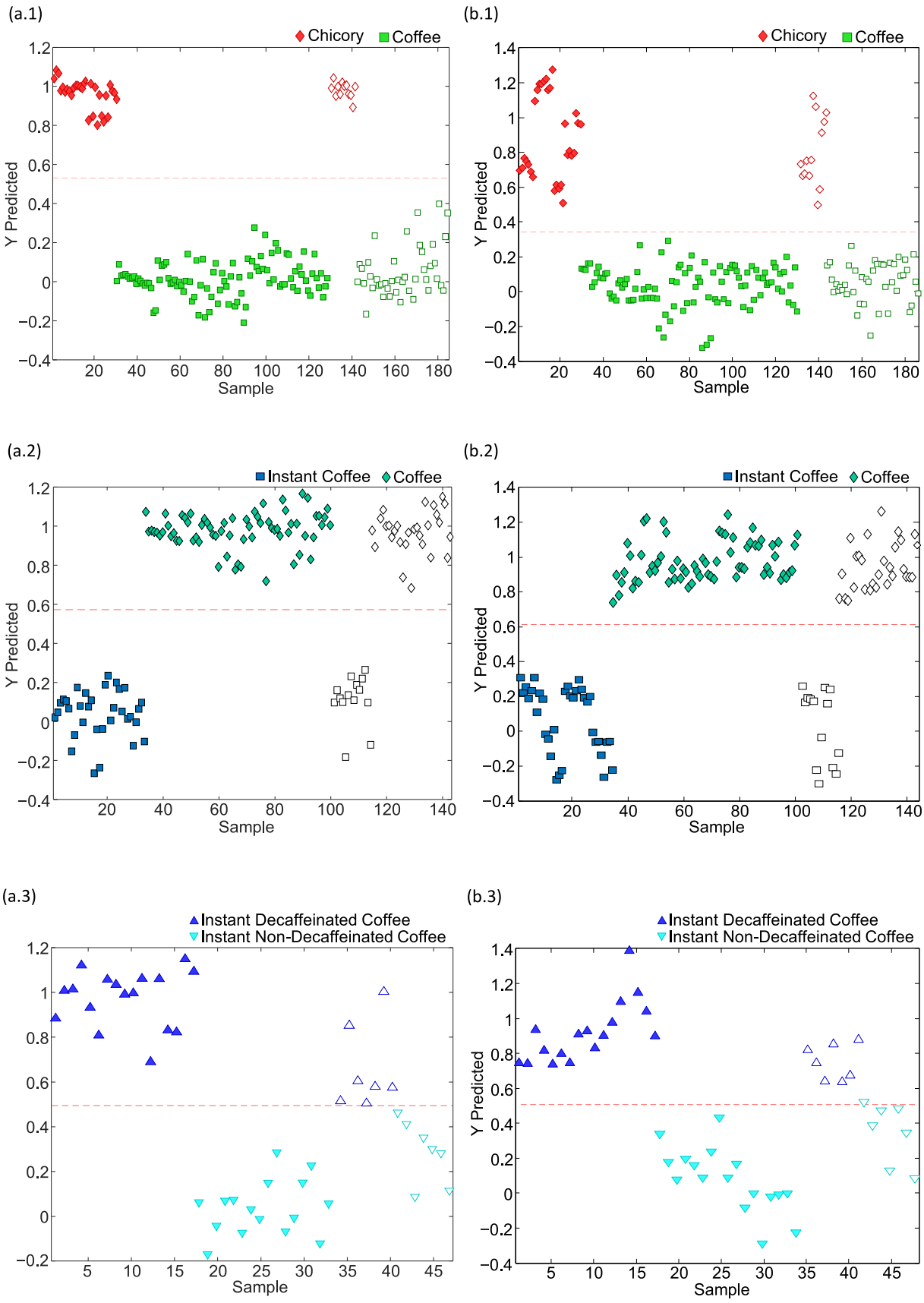


Fig. 4. Prediction plots from paired PLS-DA using fingerprints from the (a) polar and (b) non-polar columns: (1) Coffee vs. Chicory, (2) Coffee vs. soluble Coffee, and (3) Soluble non-decaffeinated Coffee vs. soluble decaffeinated coffee. Filled and empty symbols correspond to calibration and prediction sets, respectively. Red lines represent the threshold between classes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Sensitivity, specificity and classification error for calibration and prediction on paired PLS-DA models according to the sample type, coffee variety and geographical production region with fingerprints obtained from both GC–MS methods were employed.

	LVs	Class	Calibration Sensitivity (%)	Specificity (%)	Classification error (%)	Prediction Sensitivity (%)	Specificity (%)	Classification error (%)
All Coffees and chicory classification								
Polar column	4	Coffee	100	100	0	100	100	0
	4	Chicory	100	100	0	100	100	0
Non-polar column	3	Coffee	100	100	0	100	100	0
	3	Chicory	100	100	0	100	100	0
Coffee and soluble Coffee								
Polar column	3	Coffee	100	100	0	100	100	0
	3	Soluble Coffee	100	100	0	100	100	0
Non-polar column	3	Coffee	100	100	0	100	100	0
	3	Soluble Coffee	100	100	0	100	100	0
Soluble non-decaffeinated Coffee and soluble decaffeinated coffee classification								
Polar column	4	Soluble non-decaffeinated Coffee	100	100	0	100	100	0
	4	Soluble decaffeinated Coffee	100	100	0	100	100	0
Non-polar column	4	Soluble non-decaffeinated Coffee	100	100	0	100	85.7	7.1
	4	Soluble decaffeinated Coffee	100	100	0	85.7	100	7.1
Coffee variety classification								
Polar column	4	Arabica	100	100	0	100	100	0
	2	Arabica-Robusta mixture	100	100	0	100	100	0
	2	Robusta	100	100	0	100	100	0
Non-polar column	3	Arabica	100	100	0	100	100	0
	3	Arabica-Robusta mixture	100	100	0	100	93.3	3.3
	4	Robusta	100	100	0	100	100	0
Coffee geographical production region classification								
Polar column	3	Vietnam	100	100	0	100	100	0
	5	Cambodia	100	100	0	100	100	0
	5	Costa Rica	100	100	0	100	95.7	2.2
Non-polar column	3	Vietnam	100	100	0	100	100	0
	3	Cambodia	100	100	0	100	100	0
	2	Costa Rica	100	98	0.9	100	95.7	2.2

NIST database in the diverse sample typologies, distributed based on coffee variety, geographic origin, and coffee solubility.

In this context, the aromatic profile of Arabica coffee included a high content of phenolic compounds, esters, and aldehydes, particularly in the polar column. Compounds such as furfural and maltol were identified with a 90 % probability in the NIST database. These compounds are known for their contributions to sweet and fruity aromas, providing characteristic notes of caramel, fruits, and nuts [15,20]. Furthermore, compounds such as acetic acid and hexanoic acid, which were identified with a medium probability (ca. 75–80 %), contributed to the freshness and acidity of this variety's overall profile [21]. Acetic acid, in particular, is fundamental for the perception of the sour aroma [20,21]. The non-polar column also reflected the presence of aliphatic hydrocarbons and aldehydes, such as hexanal and nonanal, with identification probabilities exceeding 85 %. These compounds contribute to a slightly citrus profile, providing balance and complexity to the sensory profile of Arabica coffee [21]. In contrast, Robusta coffee exhibited a high concentration of pyrazines, such as 2-methylpyrazine and 2,5-dimethylpyrazine, detected with identification probabilities above 90 %. These compounds are typical of the roasting process [21]. The results also indicate that Robusta has a lower proportion of esters and fruity compounds compared to Arabica. In addition, sulfur-containing compounds, such as dimethyl sulfide, were detected in Robusta coffee using the polar column.

Regarding geographical origin, Vietnamese coffee was characterized by a high concentration of pyrazines and aldehydes detected in both columns, with the most prominent being 2,3-butanedione and methylpyrazine, both with identification probabilities exceeding 90 %. These compounds are responsible for buttery and earthy aromatic notes [20,21]. The profile from the polar column contained acetic acid, identified with an 85 % probability, which adds moderate acidity to Vietnamese coffee, balancing the sweetness of other compounds [21]. On the other hand, Cambodian coffee showed a higher content of volatile acids and phenolics than Vietnamese coffee. Compounds such as benzoic acid and guaiacol, with identification probabilities close to 80 %, contribute to spicy, acidity and burnt perceptions [15,20]. The non-polar column also detected phenolic derivatives. Costa Rica coffee presented profiles rich in esters and aldehydes, being remarkable compounds such as 1-hexanol and hexanal, both with high identification probabilities in the NIST database (85–90 %). These compounds render Costa Rica coffee a sweet and fruity flavor [15]. Esters, such as ethyl acetate, contributed to a more balanced perception of the unique characteristics of this coffee-growing region [22].

In the soluble coffee samples, Table 2 shows a predominance of pyrazine and aldehyde compounds, such as furfural and acetaldehyde, with high identification probabilities exceeding 90 %. These compounds generate a profile of sweet, almond, caramel and toasted notes, commonly perceived in soluble coffee [20,21]. The low presence of acids and esters suggests a less acidic and fruity aroma typical of soluble coffee due to the industrial processing that eliminate some volatiles [21].

Finally, it is important to emphasize that the previous chemometric analysis of the coffee samples relied on only the HS-SPME-GC–MS fingerprints, consisting of ion intensities at each mass-to-charge ratio (m/z) and retention time. Further compound identification, including compounds with high identification probabilities (above 90 %), was not utilized for classification purposes in this study but for providing insight into the aromatic profiles of the coffee samples.

3.6. Detection and quantitation of coffee adulteration by PLS regression

Finally, the proposed methodologies were employed to quantify adulteration levels in blended coffee samples using PLS regression. Two

Table 2
Volatile compounds identified in coffee samples using WAX column (polar column) and DB5 column (non-polar column).

UF-WAX ms column (polar column) N°	Compound	RT (min)	Exact Mass	RI value	RI theoretical value	Probability (%) Vietnamese Arabica Coffee	Vietnamese Robusta Coffee	Vietnamese Arabica and Robusta Coffee	Cambodian Coffee	Costa Rica Coffee	Costa Rica Arabica Coffee	Soluble Non-Decaffeinated Coffee	Soluble Decaffeinated Coffee
1	Hexamethylcyclotrisiloxane	6.353	222.0564	830	830 ± 2	–	–	–	–	–	–	70.8 %-78.9 %	75.2 %-79.2 %
2	2,3-Dihydro-5-methylfuran	7.800	84.0575	640	640 ± 0	–	–	–	–	–	–	69.2 %-81.5 %	74.3 %-80.8 %
3	Pyridine	17.724	79.0422	1185	1185 ± 10	79.6 %-88.5 %	79.8 %-87.8 %	86.2 %-88.4 %	83.7 %-89.3 %	82.5 %-88.8 %	81.4 %-88.1 %	79.6 %-88.7 %	82.7 %-88.0 %
4	2-Pentylfuran	19.754	138.1045	1231	1231 ± 9	–	–	–	70.5 %-77.9 %	69.4 %-82.4 %	80.1 %-85.8 %		–
5	Methylpyrazine	21.397	94.0531	1266	1266 ± 10	63.3 %-70.2 %	52.6 %-73.5 %	58.3 %-66.3 %	58.7 %-73.5 %	62.8 %-71.0 %	59.4 %-67.1 %	51.4 %-72.5 %	50.1 %-68.5 %
6	2,5-Dimethylpyrazine	23.889	108,0687	1320	1320 ± 11	82.0 %-88.5 %	76.9 %-85.4 %	80.3 %-86.3 %	79.0 %-84.5 %	82.0 %-85.5 %	83.7 %-86.8 %	73.2 %-78.8 %	70.4 %-73.9 %
7	Ethylpyrazine	24.445	108,0687	1337	1337 ± 12	75.4 %-81.4 %	72.8 %-85.1 %	79.3 %-81.3 %	75.9 %-85.5 %	75.6 %-83.4 %	79.1 %-82.7 %		77.2 %-84.7 %
8	2,3-Dimethylpyrazine	24.982	108,0688	1343	1343 ± 10	–	–	–	–	88.1 %-95.1 %	–		–
9	2-Ethyl-6-methylpyrazine	26.659	122,0844	1386	1386 ± 11	–	51.5 %-61.4 %	50.9 %-52.0 %	51.6 %-66.6 %	–	–		–
10	2-Ethyl-5-methylpyrazine	26.950	122.0844	1387	1387 ± 10	50.1 %-62.7 %	50.8 %-69.1 %	55.3 %-57.1 %	52.2 %-67.2 %	56.2 %-67.3 %	54.5 %-65.5 %	54.6 %-66.3 %	59.9 %-67.9 %
11	Trimethylpyrazine	27.466	122.0844	1402	1402 ± 8	–	84.0 %-89.0 %	82.9 %-87.6 %	84.5 %-89.7 %	83.4 %-88.2 %	82.1 %-89.9 %		85.4 %-88.9 %
12	2-Ethyl-3-methylpyrazine	27.527	122.0844	1407	1407 ± 9	–	86.9 %-91.4 %	86.2 %-90.0 %	85.8 %-91.1 %	84.1 %-91.9 %	89.5 %-90.9 %		–
13	3-Ethyl-2,5-dimethylpyrazine	29.231	136.1000	1443	1443 ± 8	71.5 %-76.5 %	71.2 %-78.0 %	69.4 %-75.1 %	73.6 %-77.2 %	69.7 %-77.5 %	73.1 %-76.8 %		–
14	1-Acetoxy-2-propanone	30.195	116.0473	1474	1474 ± 9	94.5 %-96.1 %	92.8 %-95.6 %	95.2 %-96.4 %	93.0 %-96.2 %	93.7 %-96.0 %	93.8 %-95.9 %	92.7 %-96.3 %	91.4 %-94.7 %
15	Furfural	30.317	96.0211	1461	1461 ± 11	50.6 % – 62 %	58.1 %-69.9 %	54.9 %-57.6 %	57.3 %-61.1 %	50.6 %-59.0 %	50.3 %-59.3 %	57.8 %-77.7 %	53.6 %-62.7 %
16	1-(2-Furanyl)ethanone	32.001	110.0368	1499	1499 ± 10	77.6 %-86.3 %	70.4 %-84.6 %	75.0 %-84.3 %	77.3 %-87.3 %	81.6 %-88.4 %	78.1 %-86.5 %		60.4 %-72.4 %
17	2,3-Pentanedione	32.802	100.0524	1058	1058 ± 9	54.9 %-61.6 %	50.4 %-61.4 %	51.6 %-58.9 %	–	51.7 %-59.4 %	–		–
18	1-Acetoxy-2-butanone	33.033	130.0630	1536	1536 ± 17	70.3 %-73.5 %	–	71.5 %-74.0 %	–	73.5 %-83.2 %	69.7 %-74.4 %		–
19	Acetate of 2-furanmethanol	33.162	140.0473	1531	1531 ± 10	92.5 %-96.7 %	91.3 %-96.2 %	91.0 %-94.4 %	94.4 %-96.5 %	94.0 %-96.6 %	95.7 %-96.7 %	81.8 %-93.9 %	86.4 %-95.3 %
20	Propanoic acid	33.57	74.0368	1535	1535 ± 11	72.6 %-79.9 %	73.7 %-79.3 %	73.5 %-76.9 %	69.0 %-77.3 %	66.2 %-79.3 %	–	74.3 %-82.6 %	76.3 %-80.2 %
21	2-Methylpropanoic acid	34.676	88.0524	1570	1570 ± 12	–	–	–	–	–	90.8 %-94.2 %	83.8 %-88.3 %	86.1 %-90.5 %
22	5-Methyl-2-furancarboxaldehyde	34.832	110.0368	1570	1570 ± 10	90.0 %-95.6 %	92.0 %-95.5 %	93.5 %-95.1 %	93.4 %-94.5 %	93.7 %-95.0 %	–	88.6 %-94.9 %	80.0 %-85.4 %
23	Propanoate of 2-furanmethanol	35.599	154.0630	1601	1601 ± 18	70.0 %-71.8 %	71.3 %-82.2 %	70.2 %-74.2 %	–	68.3 %-72.9 %	64.0 %-71.7 %		–
24	2,2′-Methylenebis(furan)	36.169	148.0524	1632	1632 ± 5	73.5 %-83.3 %	–	72.6 %-80.5 %	78.4 %-84.7 %	50.8 %-57.6 %	61.6 %-72.7 %		–
25	1-Ethyl-1H-pyrrole-2-carboxaldehyde	36.264	123.0684	1610	1610 ± 10	–	–	–	–	–	–	60.2 %-66.7 %	55.7 %-71.1 %

(*continued on next page*)

Table 2 (*continued*)

UF-WAX ms column (polar column) N°	Compound	RT (min)	Exact Mass	RI value	RI theoretical value	Probability (%) Vietnamese Arabica Coffee	Vietnamese Robusta Coffee	Vietnamese Arabica and Robusta Coffee	Cambodian Coffee	Costa Rica Coffee	Costa Rica Arabica Coffee	Soluble Non-Decaffeinated Coffee	Soluble Decaffeinated Coffee
26	1-Methyl-1H-pyrrole-2-carboxaldehyde	36.747	109.0528	1626	1626 ± 11	78.8 %-85.7 %	–	76.4 %-79.7 %	76.6 %-82.2 %	78.2 %-85.1 %	77.6 %-82.2 %	74.8 %-87.7 %	84.2 %-88.0 %
27	Butanoic acid	36.977	88.0524	1625	1625 ± 12	–	64.4 %-74.1 %	64.8 %-72.1 %	55.5 %-63.1 %	–	–	60.4 %-82.5 %	67.7 %-75.4 %
28	2-Furanmethanol	38.240	98.0368	1660	1660 ± 9	79.9 % – 84 %	79.6 %-84.2 %	80.3 %-83.2 %	79.7 %-81.0 %	79.9 %-83.8 %	80.1 %-83.9 %	79.7 %-86.7 %	80.2 %-86.6 %
29	3-Methylbutanoic acid	38.695	102.0681	1666	1666 ± 11	61.8 %-76.9 %	63.1 %-87.1 %	69.0 %-74.2 %	75.2 %-87.9 %	76.6 %-89.2 %	70.6 %-74.5 %	73.2 %-81.8 %	68.1 %-78.8 %
30	2-Acetyl-3-methylpyrazine	39.442	136.0637	1630	1630 ± 4	–	–	–	–	52.0 %-54.8 %	–		–
31	3,4-Dimethyl-2,5-furandione	41.044	126.0317	1714	1714 ± 29	–	–	–	–	60.6 %-69.5 %	–		–
32	3-Methyl-2-butenoic acid	43.481	100.0524	1782	1782 ± 10	51.7 %-55.0 %	–	46.1 %-49.0 %	–	51.3 %-57.4 %	–		–
33	4-Methylpentanoic acid	43.603	116.0837	1803	1803 ± 12	–	–	–	–	–	–		87.4 %-91.1 %
34	1-(2-Furanylmethyl)-1H-pyrrole	44.472	147.0684	1824	1824 ± 6	–	–	–	70.3 %-79.2 %	–	–		–
35	2-Methoxyphenol	45.585	124.0524	1861	1861 ± 13	69.7 %-79.7 %	72.7 %-85.0 %	82.1 %-85.2 %	75.3 %-84.7 %	70.6 %-75.1 %	71.6 %-81.1 %		–
36	3-Ethyl-2-hydroxy-2-cyclopenten-1-one	46.793	126.150	1894	1894 ± 3	–	–	–	–	–	–		61.4 %-73.3 %
37	Maltol	49.312	126.0317	1969	1969 ± 15	92.9 %-95.5 %	91.5 %-95.3 %	93.0 %-94.8 %	94.1 %-95.6 %	93.8 %-95.1 %	92.9 %-95.2 %	91.8 %-95.0 %	91.7 %-95.7 %
38	1-(1H-Pyrrol-2-yl)ethanone	49.543	109.0528	1973	1973 ± 12	50.5 %-57.6 %	53.6 %-63.8 %	71.4 %-75.4 %	50.3 %-61.0 %	50.5 %-57.5 %	53.4 %-63.5 %	56.8 %-67.1 %	53.5 %-64.3 %
39	Phenol	50.724	94.0419	2000	2000 ± 15	–	–	–	64.9 %-74.7 %	–	–	58.2 %-70.2 %	–
40	1H-Pyrrole-2-carboxaldehyde	51.295	95.0371	2030	2030 ± 14	50.5 %-58.3 %	63.6 %-76.0 %	58.8 %-71.8 %	–	50.3 %-58.2 %	–	69.5 %-77.8 %	72.5 %-78.6 %
41	4-Ethyl-2-methoxyphenol	51.356	152.0837	2032	2032 ± 12	–	68.4 %-75.7 %	69.1 %-72,8%	61.6 %-73.3 %	–	–		–

DB5 column (non polar column) N°	Compound	RT (min)	*m/z*	RI value	RI theoretical value	Probability (%) Vietnamese Arabica Coffee	Vietnamese Robusta Coffee	Vietnamese Arabica and Robusta Coffee	Cambodian Coffee	Costa Rica Coffee	Costa Rica Arabica Coffee	Soluble Non-Decaffeinated Coffee	Soluble Decaffeinated Coffee
1	2,3-Dihydro-5-methylfuran	3.414	84.0575	670	670 ± 8	–	–	–	–	53.0 %-59.4 %	–	65.6 %-78.0 %	67.0 %-70.4 %
2	Propanoic acid	4.663	74.0368	700	700 ± 20	–	51.0 %-59.2 %	52.1 %-55.3 %	53.2 %-62.9 %	–	72.5 %-75.0 %	51.1 %-69.0 %	–
3	Pyridine	5.301	79.0422	746	746 ± 7	82.2 %-86.2 %	80.9 %-88.9 %	82.0 %-87.1 %	77.9 %-86.0 %	80.1 %-87.8 %	81.0 %-88.5 %	75.5 %-84.7 %	79.5 %-85.9 %
4	1-Methylpiperidine	6.232	99.1048	779	779 ± N/A	–	–	–	–	–	77.8 %-83.3 %	–	–
5	1-Methyl-1,2,3,6-tetrahydropyridine	6.992	97.0891	–	–	–	–	–	–	–	86.8 %-90.4 %	–	–
6	Hexamethylcyclotrisiloxane	7.705	222.0564	851	851 ± N/A	–	–	–	–	–	–	71.1 %-82.5 %	76.9 %-78.4 %

(*continued on next page*)

Table 2 (*continued*)

UF-WAX ms column (polar column) N°	Compound	RT (min)	Exact Mass	RI value	RI theoretical value	Probability (%) Vietnamese Arabica Coffee	Vietnamese Robusta Coffee	Vietnamese Arabica and Robusta Coffee	Cambodian Coffee	Costa Rica Coffee	Costa Rica Arabica Coffee	Soluble Non-Decaffeinated Coffee	Soluble Decaffeinated Coffee
7	Methylpyrazine	8.037	94.0531	831	831 ± 7	61.9 %-68.4 %	66.5 %-69.5 %	59.5 %-67.3 %	61.9 %-67.5 %	55.9 %-61.5 %	–	56.3 %-70.6 %	–
8	Furfural	8.390	96.0211	833	833 ± 4	60.0 %-64.2 %	50.0 %-55.1 %	51.7 %-56.1 %	57.4 %-63.0 %	57,7%-62.0 %	52.7 %-60.7 %	51.0 %-56.6 %	52.4 %-60.9 %
9	2-Furanmethanol	9.666	98.0368	859	859 ± 6	79.1 %-83.2 %	75.7 %-82.0 %	78.0 %-79.2 %	77.5 %-84.1 %	77.0 %-80.3 %	74.3 %-79.9 %	77.7 %-81.6 %	78.3 %-82.7 %
10	1-(Acetyloxy)-2-propanone	10.060	116.0473	870	870 ± 5	–	69.9 %-72.1 %	60.1 %-68–0 %	–	80.3 %-86,7%	75.1 %-85.6 %	76.1 %-81.9 %	–
11	3-Methylpentanoic acid	10.142	116.0837	947	947 ± 6	61.9 %-66.5 %	70.3 %-75.8 %	67.2 %-74.7 %	62.3 %-71.2 %	61.8 %-69.8 %	–	–	–
12	3-Methylbutanoic acid	10.291	102.0681	863	863 ± 16	70.7 %-75.0 %	70.1 %-78.0 %	72.5 %-74.1 %	–	72.3 %-78.2 %	73.6–75.9	–	84.4 %-86.8 %
13	2-Methylbutanoic acid	10.603	102.0681	861	861 ± 14	73.5 %-82.8 %	–	56.0 %-58.5 %	52.1 %-61.6 %	54.7 %-63.0 %	–	–	82.9 %-84.4 %
14	1-(2-Furyl)ethanone	11.913	110.0368	911	911 ± 4	66.6 %-71.5 %	–	65.85–71.1 %	51.3 %-60.9 %	–	50.5 %-58.2 %	–	–
15	Butyrolactone	12.113	86.0368	915	915 ± 6	–	–	–	–	–	–	50.3 %-53.9 %	50.1 %-52.8 %
16	Ethylpyrazine	12.239	108.0687	921	921 ± 7	61.1 %-66.9 %	63.6 %-71.9 %	64.1 %-70.7 %	63.0 %-73.5 %	67.2 %-77.9 %	68.1 %-74.1 %	58.3 %-63.9 %	61.6 %-64.2 %
17	2,3-Dimethylpyrazine	12.375	108.0687	925	925 ± 6	64.7 %-68.8 %	64.1 %-67.2 %	65.0 %-66.9 %	60.3 %-67.2 %	69.0 %-74.9 %	63.3 %-74.2 %	–	–
18	3-Ethylpyridine	14.317	107.0735	959	959 ± 5	–	67.7 %-74.3 %	68.1 %-75.1 %	50.9 %-57.8 %	–	61.4 %-75.9 %	–	–
19	5-Methyl-2-furancarboxaldehyde	14.500	110.0368	965	965 ± 5	89.2 %-93.7 %	82.8 %-92.2 %	82.7 %-87.4 %	89.8 %-93.8 %	89.1 %-92.1 %	85.5 %-89.5 %	–	64.4 %-66 %
20	1-(Acetyloxy)-2-butanone	14.649	130.0630	967	967 ± 7	52.1 %-61.9 %	–	51.8 %-61.8 %	–	62.1 %-75.8 %	55.8 %-62.6 %	–	–
21	4-Methylpentanoic acid	14.840	116.0837	949	949 ± 6	–	–	–	–	–	–	–	68.0 %-76.8 %
22	Phenol	15.79	94.0419	980	980 ± 4	–	51.3 %-57.2 %	52.8 %-57.2 %	–	–	–	–	–
23	2-Furfurylmethanol acetate	16.109	140.0473	995	995 ± 4	90.2 %-92.7 %	88.9 %-92.7 %	85.6 %-89.1 %	91.3 %-92.6 %	92.5 %-94.7 %	91.2 %-93.5 %	–	85.0 %-92.1 %
24	2-Ethyl-6-methylpyrazine	16.333	122.0844	1003	1003 ± 6	57.4 %-69.9 %	–	57.0 %-65.0 %	54.9 %-59.1 %	58.2 %-69.4 %	57.6 %-61.7 %	59.8 %-66.3 %	54.5 %-62.8 %
25	2-Ethyl-3-methylpyrazine	16.523	122.0844	1004	1004 ± 5	55.8 %-66.8 %	54.7 %-66.3 %	54.8 %-60.4 %	58.3 %-65.7 %	61.5 %-69.8 %	65.0 %-69.9 %	–	–
26	1-Methyl-1H-pyrrole-2-carboxaldehyde	16.625	109.0528	1016	1016 ± 6	71.6 %-73.9 %	–	71.4 %-73.7 %	–	70.1 %-75.4 %	67.8 %-75.9 %	51.0 %-54.8 %	54.6 %-67.1 %
27	1-(2-Furyl)propanone	16.828	124.0524	1011	1011 ± 6	–	–	–	–	52.8 %-60.2 %	–	–	–
28	1H-Pyrrole-2-carboxaldehyde	17.419	95.0371	1015	1015 ± 7	–	–	–	–	–	–	53.4 %-56.4 %	–
29	3-Methyl-1,2-cyclopentanedione	17.725	112.0524	1043	1043 ± 0	–	–	–	–	–	–	55.3 %-67.7 %	59.7 %-66.9 %
30	1-(1H-Pyrrol-2-yl)ethanone	19.822	109.0528	1064	1064 ± 5	53.4 %-68.9 %	52.3 %-60.1 %	53.1 %-54.7 %	50.2–60.2 %	50.2 %-53.9 %	50.2 %-51.7 %	72.9 %-78.3 %	62.1 %-74.9 %
31	3-Ethyl-2,5-dimethylpyrazine	20.358	136.1000	1082	1082 ± 3	57.3 %-68.2 %	66.1 %-69.9 %	59.7 %-66.4 %	51.3 %-61.1 %	67.2 %-69.5 %	60.1 %-69.7 %	–	–
32	2,2′-Methylenebisfuran	20.61	148.0524	1088	1088 ± 2	–	–	–	–	–	54.8 %-56.3 %	–	–
33	2-methoxyphenol	20.847	124.0524	1090	1090 ± 3	–	61.9 %-67.4 %	62.4 %-67.7 %	66.2 %-77.3 %	–	–	–	–

(*continued on next page*)

Table 2 (*continued*)

UF-WAX ms column (polar column) N°	Compound	RT (min)	Exact Mass	RI value	RI theoretical value	Probability (%) Vietnamese Arabica Coffee	Vietnamese Robusta Coffee	Vietnamese Arabica and Robusta Coffee	Cambodian Coffee	Costa Rica Coffee	Costa Rica Arabica Coffee	Soluble Non-Decaffeinated Coffee	Soluble Decaffeinated Coffee
34	Maltol	22.008	126.0317	1110	1110 ± 6	68.9 %-76.2 %	70.2 %-75.3 %	69.5 %-74.8 %	–	50.3 %-55.7 %	–	82.6 %-91.9 %	91.0 %-92.6 %
35	3-Ethyl-2-hydroxy-2-cyclopenten-1-one	22.382	126.0681	1091	1091 ± 17	–	–	–	–	–	–	–	60.7 %-62.4 %
36	Decamethylcyclopentasiloxane	23.407	370.0940	1175	1175 ± 40	–	–	–	–	–	–	90.7 %-94.2 %	53.1 %-55.7 %
37	5H-5-Methyl-6,7-dihydrocyclopentapyrazine	23.536	134.0844	1149	1149 ± 8	–	–	–	50.2 %-52.3 %	–	–	–	–
38	2-(2-Furanylmethyl)-5-methylfuran	25.389	162.0681	1190	1190 ± 6	53.8 %-59.2 %	–	53.8 %-60.0 %	–	–	–	–	–
39	1-(2-Furanylmethyl)-1H-pyrrole	25.477	147.0684	1187	1187 ± 7	91.2 %-94.6 %	91.6 %-96.7 %	79.4 %-82.8 %	91.2 %-95.5 %	93.8 %-96.7 %	90.6 %-96.3 %	–	–
40	5-Hydroxymethylfurfural	27.915	126.0312	1233	1233 ± 8	–	–	–	–	–	–	89.8 %-92.0 %	–
41	4-Ethyl-2-methoxyphenol	30.032	152.0837	1282	1282 ± 12	–	58.0 %-71.6 %	57.8 %-72.1 %	–	–	–	–	–
42	Dodecamethylcyclohexasiloxane	31.000	444.1127	1335	1335 ± 15	–	–	–	–	–	–	90.9 %-97.1 %	83.1 %-89.7 %
43	2-Methoxy-4-vinylphenol	31.736	150.0681	1317	1317 ± 5	–	55.8 %-64.8 %	54.9 %-65.1 %	52.4 %-59.3 %	–	–	–	–
44	1,1′-(1,4-Phenylene)bisethanone	36.991	162.0681	1461	1461 ± 10	–	–	–	–	–	–	52.6 %-66.1 %	–

coffee adulteration scenarios were considered. The first case was Vietnamese Robusta coffee adulterated with chicory, and the second case was Vietnamese Arabica coffee adulterated with Vietnamese Robusta coffee.

Independent calibration and validation sets were employed for each adulteration case. The calibration sets encompassed adulteration levels between 0 % to 100 %, whereas the validation sets comprised adulteration levels between 15 to 85 % (see Section 2.1 and Fig. S1 in supplementary material). Each level of adulteration was prepared in triplicate, resulting in 33 sample extracts for each case under study. Additionally, a quality control solution was prepared at a 50 % adulteration level.

The PLS results obtained for the two adulteration cases considered when using both GC–MS proposed methodologies are depicted in Fig. 5. In general, good results were obtained with R^2 coefficients higher than 0.993, indicating the good performance of both GC–MS methodologies. Moreover, Root Mean Square Error of Calibration (RMSEC) and Root Mean Square Error of Prediction (RMSEP) were also acceptable in all evaluated cases, with values below 2.9 % and 7.4 %, respectively.

4. Evaluation of the greenness and blueness of the method

Sustainability and practical applicability are crucial factors in modern analytical chemistry. The integration of Green Analytical Chemistry (GAC) principles into method development aims to reduce environmental impact, minimize reagent consumption, and improve safety while ensuring efficiency and reliability in real-world applications [23,24]. Alongside environmental sustainability, the practical applicability of an analytical method—often referred to as "blueness"—assesses its feasibility for routine use, including factors such as automation, sample throughput, and reagent requirements [25,26].

To provide a quantitative assessment of both aspects, the AGREE metric was used to evaluate the greenness of the method, while the Blue Applicability Grade Index (BAGI) was applied to assess its practical utility (blueness). These tools allow for an objective comparison of our methodology against existing approaches, reinforcing its suitability for large-scale applications such as coffee authentication and adulteration detection.

4.1. Greenness assessment using AGREE metric

The AGREE tool provides a comprehensive evaluation of analytical methods based on the 12 principles of Green Analytical Chemistry (GAC), incorporating aspects such as solvent use, energy consumption, sample preparation, waste generation, and reagent toxicity [23]. This metric generates a circular pictogram with a quantitative score from 0 to 1, where higher values indicate a greener method.

The employed HS-SPME-GC–MS methodology was evaluated using the AGREE framework based on the following GAC principles [23]:

1) Direct analytical techniques should be applied to avoid sample treatment: HS-SPME is a non-invasive extraction method, so the score is 0.90.
2) Minimal sample size and minimal number of samples are goals: Small sample volume (0.5 g), so the score is 0.75.
3) In situ measurements should be performed: On-line analysis with GC–MS, so the score is 0.66.
4) Integration of analytical processes and operations saves energy and reduces the use of reagents: 5 steps in total (extraction, desorption, separation, detection, data analysis), so the score is 0.60.
5) Automated and miniaturized methods should be selected: HS-SPME and GC–MS were automatic and miniaturized, so the score is 1.
6) Derivatization should be avoided: No derivatization applied, so the score is 1.

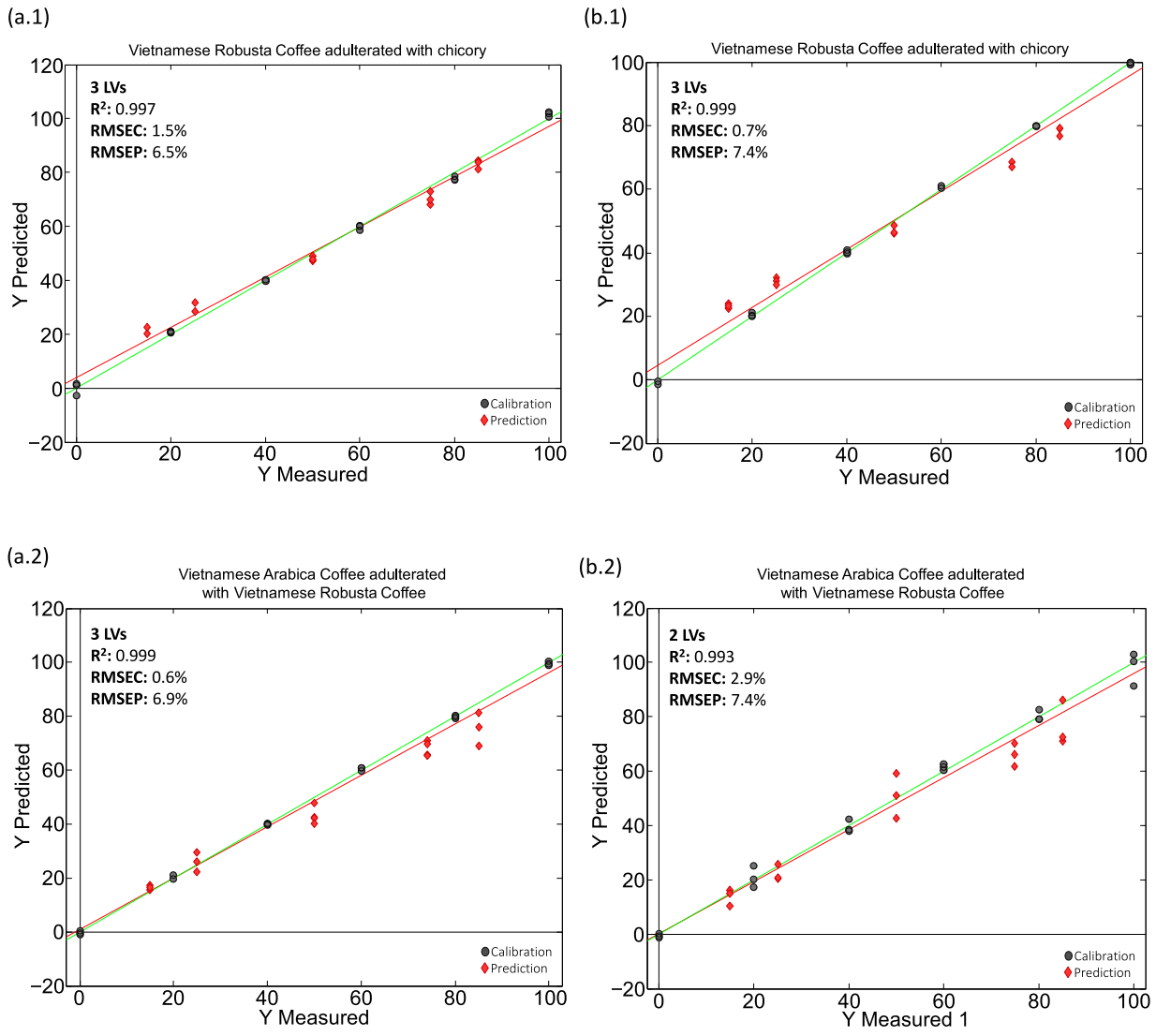


Fig. 5. PLS regression for results using fingerprints from (a) polar column and (b) non-polar column for (1) Vietnamese Robusta Coffee adulterated with Chicory and (2) Vietnamese Arabica Coffee adulterated with Vietnamese Robusta Coffee.

7) Generation of a large volume of analytical waste should be avoided and proper management of analytical waste should be provided: Minimal waste (0.5 g per sample), so the score is 1.
8) Multianalyte or multiparameter methods are preferred versus methods using one analyte at a time: Was a non-targeted approach, so multiple analytes were detected per hour, so the score is 1.
9) The use of energy should be minimized: GC–MS consumes ≥ 1.5 kWh per sample, so the score is 0.
10) Reagents obtained from renewable source should be preferred: No solvents or chemicals used, so the score is 1.
11) Toxic reagents should be eliminated or replaced: No toxic reagents are used, so the score is 1.
12) The safety of the operator should be increased: No chemical hazards present, so the score is 1.

The final AGREE score of the proposed methodology was the average of the 12 GAC principles. The obtained AGREE score was 0.83, classifying the methodology as a highly sustainable analytical approach. Compared to conventional GC–MS workflows that require derivatization, solvent extraction, or extensive sample handling, our methodology significantly reduces chemical waste, hazard exposure, and energy consumption, making it an eco-friendly alternative for coffee authentication and adulteration detection.

4.2. *Blueness assessment using BAGI metric*

We applied the Blue Applicability Grade Index (BAGI) to evaluate the practical applicability of the method, which assesses 10 key parameters related to efficiency, automation, sample processing, and reagent requirements [25]. Each category is scored based on a standardized scale, with a maximum score of 100, where higher values indicate a more practical method. To be regarded as practical, it is recommended that the method achieve at least 60 points.

The BAGI evaluation of the HS-SPME-GC–MS methodology yielded the following attributes [25]:

1) Type of analysis: Qualitative analysis scores 2.5.
2) Multi- or single- element analysis: Multi-element analysis for ≥ 15 compounds scores 10.
3) Analytical technique: Sophisticated instrumentation (GC–MS) scores 5.
4) Simultaneous sample preparation: Every sample was analysed individually, so the score is 2.5.
5) Sample preparation: Miniaturized extraction sample preparation (SPME)scores 5.

6) Samples per h (sample preparation + analysis time): Each sample takes more than 70 min, which is equivalent to a minutes that 1 sample per hour, so is punctuated with a 2.5.
7) Reagents and materials: Commercially available reagents that are non-common in QC labs scores 7.5.
8) Preconcentration: HS-SPME preconcentrates the volatile compounds present in samples, so it is scored 7.5.
9) Automation degree: The instrument equipped with an autosampler for HS-SPME and GC–MS scores 10.
10) Amount of sample: 0.5 g of coffee was employed per sample, so it is scored 10.

The final BAGI score of the proposed methodology was the sum of the scores of the 10 BAGI attributes. The obtained BAGI score was 62.5, indicating good applicability for routine and large-scale implementations. Compared to conventional GC–MS approaches, which often require sample derivatization, extensive pre-processing, or complex extraction steps, our HS-SPME-GC–MS workflow reduces sample handling time, minimizes reagent use, and enhances automation, making it an ideal choice for high-throughput authentication of coffee samples.

5. Comparison with other scientific publications

This section has been included to contextualize our findings by contrasting them with previous research in the field. A detailed summary table (Table S3) in the Supplementary Material presents key features and outcomes from relevant studies, allowing for a direct comparison with our work. The table includes information on authors, publication year, study objectives, methodologies used, and how our results align or differ from those of other researchers.

The position of the present study within the broader literature is contrasted and its contribution to the advancement of coffee authentication research is highlighted. The supplementary material also offers a comparative analysis of various GC–MS-based studies on coffee, encompassing topics such as volatile compound profiling, fraud detection, and contamination assessment. While these approaches all contribute valuable insights into coffee quality and authenticity, they differ in their analytical scope, with some emphasizing metabolic markers for sensory quality, others focusing on chemical contaminants, and some addressing adulteration detection. These distinctions underline the diverse applications of GC–MS in coffee research and demonstrate how our study complements and expands upon existing methodologies [14,15,27–37].

6. Conclusions

Two GC–MS methods have been proposed to obtain HS-SPME-GC–MS fingerprints providing chemical descriptors of the samples suitable for the characterization, classification, and authentication of coffee samples of different types, varieties, and origins, as well as for discriminating coffee with chicory (one of its possible adulterants).

Chemometric analysis by PCA using the HS-SPME-GC–MS fingerprints show good discrimination capabilities between coffee and chicory and between coffee varieties with the two GC–MS methodologies (with polar or non-polar column).

The PLS-DA models showed very good results providing sensitivity and specificity values greater than 93.3 % and 94.5 %, respectively. Furthermore, PLS-DA paired models simulating an external validation provided excellent results, with sensitivity and specificity values of 100 % when coffee is classified in front chicory and when non-soluble coffee is classified in front soluble coffee. For soluble decaffeinated and non-decaffeinated coffees, prediction values higher of 85.7 % were obtained. For coffee varieties, sensitivity values of 100 % and specificity values higher than 93.3 % were observed. Finally, for geographical production regions, sensitivity values of 100 % and specificity values higher than 95.7 % were achieved.

The capability of the proposed methods to detect and quantify coffee frauds through multivariate PLS regression was studied for two adulteration cases: Vietnamese Arabica Coffee adulterated with Vietnamese Robusta Coffee and Vietnamese Robusta Coffee adulterated with chicory. Very acceptable calibration and prediction errors were accomplished, lower than 2.9 % and 7.4 %, respectively. Hence, the HS-SPMEGC-MS methodologies combined with chemometrics are good strategies to authenticate coffee and to detect and quantify coffee frauds.

Additionally, the identification of volatile compounds in the samples provides a detailed understanding of the characteristic aromatic profiles of each coffee type. These results allow for the correlation of the complexity and diversity of coffee aromas and flavors with their chemical composition, highlighting the utility of the HS-SPME-GC/MS technique for sensory studies and authenticity in coffee. Moreover, the precise identification of volatile compounds opens the door to future studies dealing with agricultural and processing practices, thereby improving coffee quality and meeting consumer expectations in a constantly evolving market.

Furthermore, although our study does not directly measure aroma or taste, it is well established in the literature that volatile compounds such as esters, aldehydes, and pyrazines play crucial roles in the aroma and flavor profiles of coffee. Previous studies have demonstrated a direct link between volatile compounds and sensory attributes, supporting the idea that the chemical profiles obtained here could offer insights into the sensory qualities of the coffee samples. In future research, it would be valuable to combine chemical profiling with sensory analysis to further correlate these findings.

Finally, AGREE and BAGI evaluations demonstrate that the employed HS-SPME-GC–MS approach is both environmentally sustainable (AGREE score: 0.79) and highly applicable for routine analysis (BAGI score: 62.5/100). These findings reinforce its potential as a robust, scalable, and eco-friendly alternative for coffee authentication and adulteration detection, bridging the gap between sustainability and analytical efficiency.

7. Author statement

All the authors have made substantial contributions to all of the following

1) The conception and design of the study, or acquisition of data, or analysis and interpretation of data.
2) Drafting the article or revising it critically for important intellectual content.
3) Final approval of the version to be submitted.

CRediT authorship contribution statement

Nerea Núñez: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation. **Erica Moret:** Writing – review & editing, Supervision, Software, Investigation. **Paolo Lucci:** Writing – review & editing, Supervision, Investigation, Funding acquisition, Conceptualization. **Sabrina Moret:** Writing – review & editing, Resources, Investigation. **Javier Saurina:** Writing – review & editing, Supervision, Software, Investigation, Funding acquisition, Conceptualization. **Oscar Núñez:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Funding acquisition, Conceptualization.

Funding

This research was funded by the Ministerio de Ciencia e Innovación (MICINN) under the project PID2023-147160OB-C22. The authors want to acknowledge the Agency for Administration of University and Research Grants (Generalitat de Catalunya, Spain) for the project

2021SGR365 and María de Maetzu Unit of Excellence (Research Institute of Nutrition and Food Safety, INSA-UB, University of Barcelona), Grant CEX2021-001234-M, funded by MCIN/AEI/https://doi.org/10.13039/501100011033.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.microc.2025.113771.

Data availability

Data will be made available on request.

References

[1] F. Ulberth, Tools to combat food fraud – a gap analysis, Food Chem. 330 (2020) 127044, https://doi.org/10.1016/j.foodchem.2020.127044.
[2] G. Campmajó, N. Núñez, O. Núñez, The Role of Liquid Chromatography-Mass Spectrometry in Food Integrity and Authenticity, in: Kamble, G.S. (ed.) Mass Spectrometry - Future Perceptions and Applications, in: IntechOpen, London, 2019: pp. 3–20. https://doi.org/https://doi.org/10.5772/57353.
[3] V. Barrere, K. Everstine, J. Théolier, S. Godefroy, Food fraud vulnerability assessment: towards a global consensus on procedures to manage and mitigate food fraud, Trends Food Sci. Technol. 100 (2020) 131–137, https://doi.org/10.1016/j.tifs.2020.04.002.
[4] D. Giacalone, T.K. Degn, N. Yang, C. Liu, I. Fisk, M. Münchow, Common roasting defects in coffee: aroma composition, sensory characterization and consumer perception, Food Qual. Prefer. 71 (2019) 463–474, https://doi.org/10.1016/j.foodqual.2018.03.009.
[5] L. Chen, D. Li, Y. Rong, Fermentation mechanism of ginkgo rice wine using an ultra-high-performance liquid chromatography–quadrupole/time-of-flight mass spectrometry based metabolomics method, J. Food Compos. Anal. 105 (2022) 104230, https://doi.org/10.1016/j.jfca.2021.104230.
[6] X. Dong, S. Li, J. Sun, Y. Li, D. Zhang, nutrients Association of Co ff ee , Deca ff einated Co ff ee and Ca ff eine Intake from Co ff ee with Cognitive Performance in Older Adults: National Health and, (2020) 2011–2014.
[7] M. de S.G. Barbosa, M.B. dos S. Scholz, C.S.G. Kitzberger, M. de T. Benassi, Correlation between the composition of green Arabica coffee beans and the sensory quality of coffee brews Food Chem. 292 (2019) 275 280 10.1016/j.foodchem.2019.04.072.
[8] J. Montenegro, L.S. dos Santos, R.G.G. de Souza, L.G.B. Lima, D.S. Mattos, B.P.P. B. Viana, A.C.S. da Fonseca Bastos, L. Muzzi, C.A. Conte-Júnior, E.R.P. Gimba, O. Freitas-Silva, A.J. Teodoro, Bioactive compounds, antioxidant activity and antiproliferative effects in prostate cancer cells of green and roasted coffee extracts obtained by microwave-assisted extraction (MAE), Food Res. Int. 140 (2021), https://doi.org/10.1016/j.foodres.2020.110014.
[9] F. Machado, M.A. Coimbra, M.D. del Castillo, F. Coreta-Gomes, Mechanisms of action of coffee bioactive compounds–a key to unveil the coffee paradox, Crit. Rev. Food Sci. Nutr. 64 (2023) 10164–10186, https://doi.org/10.1080/10408398.2023.2221734.
[10] B. Sezer, H. Apaydin, G. Bilge, I.H. Boyaci, Coffee arabica adulteration: Detection of wheat, corn and chickpea, Food Chem. 264 (2018) 142–148, https://doi.org/10.1016/j.foodchem.2018.05.037.
[11] W.P. Claro Gomes, G. Gonçalves Bortoleto, W.R. Melchert, Spectrophotometry and chromatography analyses combined with chemometrics tools to differentiate green coffee beans into special or traditional, J. Food Sci. 88 (2023) 5012–5025, https://doi.org/10.1111/1750-3841.16807.
[12] G. Mannino, R. Kunz, M.E. Maffei, Discrimination of Green Coffee (Coffea arabica and Coffea canephora) of Different Geographical Origin Based on and DNA RFLP Fingerprinting, Antioxidants. (2023).
[13] I.N. Pasias, I. Kiriakou, C. Proestos, Development of a rapid method for the determination of caffeine in coffee grains by GC-FID—a fully validated approach, Antioxidants. 6 (2017) 4–9, https://doi.org/10.3390/antiox6030067.
[14] A. Gamal, M. Soliman, M.S. Al-Anany, F. Eissa, Optimization and validation of high throughput methods for the determination of 132 organic contaminants in green and roasted coffee using GC-QqQ-MS/MS and LC-QqQ-MS/MS, Food Chem. 449 (2024) 139223, https://doi.org/10.1016/j.foodchem.2024.139223.
[15] N. Caporaso, M.B. Whitworth, C. Cui, I.D. Fisk, Variability of single bean coffee volatile compounds of Arabica and robusta roasted coffees analysed by SPME-GC-MS, Food Res. Int. 108 (2018) 628–640, https://doi.org/10.1016/j.foodres.2018.03.077.
[16] D. Zhang, M. Gao, Y. Cai, J. Wu, F. Lao, Profiling flavor characteristics of cold brew coffee with GC-MS, electronic nose and tongue: effect of roasting degrees and freeze-drying, J. Sci. Food Agric. 104 (2024) 6139–6148, https://doi.org/10.1002/jsfa.13437.
[17] S.J. Kim, S. Lee, E. Bang, S. Lee, J.K. Rhee, Y.C. Na, Comparative evaluation of flavor compounds in fermented green and roasted coffee beans by solid phase microextraction-gas chromatography/mass spectrometry, Flavour, Fragr. J. 34 (2019) 365–376, https://doi.org/10.1002/ffj.3517.
[18] Eigenvector Research Incorporated. Powerful Resources for Intelligent Data Analysis. Available online: http://www.eigenvector.com/software/solo.htm (accessed on 15 January 2019), n.d.
[19] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. de Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of chemometrics and qualimetrics, Elsevier, Amsterdam, The Netherlands, 1997.
[20] D. Bressanello, E. Liberto, C. Cordero, P. Rubiolo, G. Pellegrino, M.R. Ruosi, C. Bicchi, Coffee aroma: Chemometric comparison of the chemical information provided by three different samplings combined with GC–MS to describe the sensory properties in cup, Food Chem. 214 (2017) 218–226, https://doi.org/10.1016/j.foodchem.2016.07.088.
[21] N. Yang, C. Liu, X. Liu, T.K. Degn, M. Munchow, I. Fisk, Determination of volatile marker compounds of common coffee roast defects, Food Chem. 211 (2016) 206–214, https://doi.org/10.1016/j.foodchem.2016.04.124.
[22] X. Wang, Y. Wang, G. Hu, D. Hong, T. Guo, J. Li, Z. Li, M. Qiu, Review on factors affecting coffee volatiles: from seed to cup, J. Sci. Food Agric. 102 (2022) 1341–1352, https://doi.org/10.1002/jsfa.11647.
[23] L. Yin, L. Yu, Y. Guo, C. Wang, Y. Ge, X. Zheng, N. Zhang, J. You, Y. Zhang, M. Shi, Green analytical chemistry metrics for evaluating the greenness of analytical procedures, J. Pharm. Anal. 14 (2024) 101013, https://doi.org/10.1016/j.jpha.2024.101013.
[24] S.F. Hammad, M.A.A. Hamid, L. Adly, S.H. Elagamy, Comprehensive review of greenness, whiteness, and blueness assessments of analytical methods, Green Anal. Chem. 12 (2025) 100209, https://doi.org/10.1016/j.greeac.2025.100209.
[25] N. Manousi, W. Wojnowski, J. Płotka-Wasylka, V. Samanidou, Blue applicability grade index (BAGI) and software: a new tool for the evaluation of method practicality, Green Chem. 25 (2023) 7598–7604, https://doi.org/10.1039/d3gc02347h.
[26] M.K. Halim, O.M. Badran, A.E.F. Abbas, Greenness, blueness and whiteness evaluated-chemometric approach enhanced by Latin hypercube technique for the analysis of lidocaine, diclofenac and carcinogenic impurity 2,6-dimethylaniline, Sustain. Chem. Pharm. 38 (2024) 101463, https://doi.org/10.1016/j.scp.2024.101463.
[27] C. de C. Couto, D.W.H. Chávez, E.M.M. Oliveira, O. Freitas-Silva, S. Casal, SPME-GC-MS untargeted metabolomics approach to identify potential volatile compounds as markers for fraud detection in roasted and ground coffee Food Chem. 446 2024 10.1016/j.foodchem.2024.138862.
[28] J.C. de Souza, C. de Carvalho Couto, A.M.G.N. Mamede, P. Valderrama, A.T. Toci, O. Freitas-Silva, Using volatile compounds for the identification of coffee adulterants: marker compounds and non-targeted analysis, Eur. Food Res. Technol. 250 (2024) 2639–2649, https://doi.org/10.1007/s00217-024-04563-3.
[29] M.M. Urugo, Y.B. Tola, B.T. Kebede, O. Onwuchekwa, D.S. Mattinson, Utilizing HS-SPME-GC-MS for regional classification of ethiopian green coffee beans: an in-depth analysis of volatile compounds, ACS Food Sci. Technol. 4 (2024) 1265–1277, https://doi.org/10.1021/acsfoodscitech.4c00101.
[30] P. Thammarat, C. Kulsing, K. Wongravee, N. Leepipatpiboon, T. Nhujak, Identification of volatile compounds and selection of discriminant markers for elephant dung coffee using static headspace gas chromatography—mass spectrometry and chemometrics, Molecules 23 (2018), https://doi.org/10.3390/molecules23081910.
[31] B. Mehari, M. Redi-Abshiro, B.S. Chandravanshi, S. Combrinck, R. McCrindle, M. Atlabachew, GC-MS profiling of fatty acids in green coffee (Coffea arabica L.) beans and chemometric modeling for tracing geographical origins from Ethiopia, J. Sci. Food Agric. 99 (2019) 3811–3823, https://doi.org/10.1002/jsfa.9603.
[32] P. Aditiawati, D.I. Astuti, J.A. Kriswantoro, S.M. Khanza, T. Kamarisima, F. Irifune, E. Amalia, S.P.P. Fukusaki, GC/MS-based metabolic profiling for the evaluation of solid state fermentation to improve quality of Arabica coffee beans, Metabolomics 16 (2020) 1–14, https://doi.org/10.1007/s11306-020-01678-y.
[33] K. Zhang, J. Cheng, Q. Hong, W. Dong, X. Chen, G. Wu, Z. Zhang, Identification of changes in the volatile compounds of robusta coffee beans during drying based on HS-SPME/GC-MS and E-nose analyses with the aid of chemometrics, Lwt. 161 (2022) 113317, https://doi.org/10.1016/j.lwt.2022.113317.
[34] M. Gancarz, B. Dobrzański, U. Malaga-Toboła, S. Tabor, M. Combrzyński, D. Ćwikła, W.R. Strobel, A. Oniszczuk, H. Karami, Y. Darvishi, A. Żytek, R. Rusinek, Impact of Coffee Bean Roasting on the Content of Pyridines Determined by Analysis of Volatile Organic Compounds, Molecules 27 (2022) 1–12, https://doi.org/10.3390/molecules27051559.
[35] F. Amalia, T. Irifune, T. Takegami, U. Yusianto, S.P. Sumirat, E.F. Putri, Identification of potential quality markers in Indonesia's Arabica specialty coffee using GC/MS-based metabolomics approach, Metabolomics 19 (2023) 1–11, https://doi.org/10.1007/s11306-023-02051-5.
[36] F.S. Aurum, M.Z. Zaman, E. Purwanto, D. Praseptiangga, K. Nakano, Coffee authentication via targeted metabolomics and machine learning: Unveiling origins and their discriminating biochemicals, Food Biosci. 56 (2023) 103122, https://doi.org/10.1016/j.fbio.2023.103122.
[37] F. Hardoyono, K. Windhani, Combination of metal oxide semiconductor gas sensor array and solid-phase microextraction gas chromatography–mass spectrometry for

odour classification of brewed coffee, Flavour Fragr. J. 38 (2023) 451–463, https://doi.org/10.1002/ffj.3759.