

## Virgin olive oil volatile fingerprint and chemometrics: Towards an instrumental screening tool to grade the sensory quality



Beatriz Quintanilla-Casas<sup>a,b</sup>, Julen Bustamante<sup>a,b</sup>, Francesc Guardiola<sup>a,b</sup>,  
Diego Luís García-González<sup>c</sup>, Sara Barbieri<sup>d</sup>, Alessandra Bendini<sup>d</sup>, Tullia Gallina Toschi<sup>d</sup>,  
Stefania Vichi<sup>a,b,\*</sup>, Alba Tres<sup>a,b</sup>

<sup>a</sup> *Department de Nutrició, Ciències de l'Alimentació i Gastronomia, Campus De l'Alimentació Torribera, Facultat de Farmàcia i Ciències de l'Alimentació, Universitat de Barcelona, Santa Coloma de Gramenet, Spain*

<sup>b</sup> *Institut de Recerca en Nutrició i Seguretat Alimentària (INSA-UB), Universitat de Barcelona (UB), Santa Coloma de Gramenet, Spain*

<sup>c</sup> *Instituto de la Grasa, Sevilla, Spain*

<sup>d</sup> *Department of Agricultural and Food Sciences, Alma Mater Studiorum - Università di Bologna, Bologna, Italy*

### ARTICLE INFO

#### Keywords:

Olive oil  
Sensory quality  
Volatile compounds  
HS-SPME-GC-MS  
Fingerprint

### ABSTRACT

Sensory quality, assessed following a standardized method, is one of the parameters defining the commercial category of virgin olive oil. Considering the difficulties linked to the organoleptic evaluation, especially the high number of samples to be assessed, setting up instrumental methods to support sensory panels becomes a need for the olive oil sector. Volatile fingerprint by Headspace Solid Phase Microextraction-Gas Chromatography-Mass Spectrometry can be an excellent fit-for-purpose tool as the volatile fraction is responsible for virgin olive oil sensory attributes. A fingerprinting approach was applied to the volatile profile of 176 virgin olive oils previously graded by six official sensory panels. The classification strategy consisted in two sequential Partial Least Square-Discriminant Analysis models built with the aligned chromatograms: the first discriminated extra virgin and non-extra virgin samples; the second classified the latter into virgin or lampante categories. Results were satisfactory in the cross-validation by leave 10%-out (97% of correct classification). For external validation, an uncertainty range was set for the prediction models to detect boundary samples, which would be further assessed by the sensory panels. By doing this, a considerable decrease of the panel workload (around 80%) was achieved, while maintaining a highly reliable classification of samples (error rate < 10%).

### 1. Introduction

According to the European regulation (EEC No 2568/91 and its amendments), the commercial classification of virgin olive oil into extra virgin (EVOO), virgin (VOO) and lampante (LOO-not suitable for human intake) categories is based on its quality, which is determined by physicochemical (such as acidity and peroxide index, among others) as well as sensory parameters. While performing the analysis of the former is relatively simple, the organoleptic assessment of virgin olive oils still presents some drawbacks despite trained tasters conforming the certified panels follow the guidelines detailed in the current official method (EEC No 2568/91 and its amendments; IOC/T.20/Doc. No 15/Rev. 10 2018). In particular, the high number of samples that the panel must analyse and the fact that it is a time-consuming methodology may restrict the performance of repetitions just in case of contradictory

classifications (Regulation EU No 2016/1227), harming the robustness of the results (Conte et al., 2019). For these reasons, setting up predictive models based on instrumental methods for the assessment of virgin olive oil sensory quality could be helpful to reduce the number of samples to be assessed by the sensory panels, and be supportive in case of contradictory panel assessments. In particular, instrumental screening tools could be especially useful to discern samples away from the boundaries (EVOO/VOO and VOO/LOO) (Conte et al., 2019).

Virgin olive oil positive attributes and sensory defects are determined by the composition of its volatile fraction (Angerosa et al., 2004; Aparicio, Morales, & García-González, 2012; Kalua et al., 2007). High quality virgin olive oil aroma is attributed to biogenic molecules produced by endogenous plant enzymes, specially belonging to the lipoxygenase (LOX) pathway; whereas volatiles present in low quality virgin olive oils are produced by exogenous enzymes, generally related

\* Corresponding author. Department de Nutrició, Ciències de l'Alimentació i Gastronomia, Campus de l'Alimentació Torribera, Facultat de Farmàcia i Ciències de l'Alimentació, Universitat de Barcelona, Santa Coloma de Gramenet, Spain.

E-mail address: [stefaniavichi@ub.edu](mailto:stefaniavichi@ub.edu) (S. Vichi).

<https://doi.org/10.1016/j.lwt.2019.108936>

Received 13 September 2019; Received in revised form 29 November 2019; Accepted 6 December 2019

Available online 07 December 2019

0023-6438/ © 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

to microbial metabolism, or by chemical oxidation (Angerosa et al., 2004; Kalua et al., 2007; Morales, Luna, & Aparicio, 2005). Thus, the volatile profile analysis can be the basis of the fit-for-purpose tool to support sensory panels. Indeed, several studies have been addressed to predict sensory defects and attributes in virgin olive oils according to their volatile profile analysed by different instrumental techniques (Lerma-García et al., 2010; Morales, Luna, & Aparicio, 2005; Procida, Giomo, Cichelli, & Conte, 2005). In particular, some chemometric models to discriminate virgin olive oils of distinct commercial categories have been proposed with encouraging results (Borràs et al., 2016; Cecchi et al., 2019; Contreras, Jurado-Campos, & Arce, 2019; López-Ferri, Cárdenas, García-Mesa, & Valcárcel, 2007; Sales et al., 2017; Sales, Portolés, Johnsen, Danielsen, & Beltran, 2019). They are based on the multivariate analysis of volatile profile data obtained by specific pre-concentration techniques or on data fusion approaches using distinct instrumental responses. The use of such technologies is not widespread in conventional control laboratories, so in some cases it would be hardly applicable for routine screening analysis. Among the common procedures used for volatile assessment, Headspace Solid Phase Micro Extraction (HS-SPME) coupled to gas chromatography-mass spectrometry (GC-MS) has revealed to be faster, completely automatable, affordable and highly sensitive (Angerosa, Mostallino, Basti, & Vito, 2000; Benelli et al., 2015; Cecchi et al., 2019; Morales, Luna, & Aparicio, 2005; Morales, Rios, & Aparicio, 1997; Oliver-Pozo, Aparicio-Ruiz, Romero, & García-González, 2015; Vichi et al., 2003; Vichi, Guadayol, Caixach, López-Tamames, & Buxaderas, 2007). On the other hand, most of them point out the higher capabilities of the untargeted over the traditional targeted approach. The latter implies the identification, integration and quantification of pre-defined volatile compounds. Even though this strategy is robust, easy to validate and suitable to be transferred, it becomes time-consuming when it is performed manually, especially when several target compounds are to be quantified in a high number of samples. Moreover, some valuable information, such as minor or not well-resolved peaks may be missed. This drawback can be overcome following an untargeted approach because it includes the full analytical information. Within untargeted approach in food analysis, a state-of-the-art method to find specific patterns in highly dimensional data is using the raw analytical data from chromatographic or spectroscopic techniques, known as fingerprints (Berrueta, Alonso-Salces, & Herberger, 2007). In this way, the corresponding authentication models are developed by considering the measurement at each acquisition point as a variable. Under a fingerprinting approach, deconvolution procedures generally associated to untargeted chromatography-mass spectrometry are not required.

The objective of the current work has been to develop and validate a classification model based on the volatile fingerprint by HS-SPME-GC-MS for the assignment of the commercial category (EVOO, VOO or LOO) of virgin olive oil samples. This could be used as screening tool that supports the sensory panels by reducing the number of samples to be analysed, lightning the workload and improving the panel performance.

## 2. Material and methods

### 2.1. Olive oil samples

Virgin olive oil ( $n = 176$ ) samples from several production areas (EU and non-EU countries) were evaluated and graded into EVOO, VOO AND LOO according to the International Olive Council procedure (IOCT.20/Doc. No 15/Rev. 10 2018) by six panels from different countries (Eurofins, Germany; ITERG, France; IPTPO, Croatia; UNIBO, Italia; UZZK, Turkey; ZRS, Slovenia) recognized by National or/and International authorities. They were partners in the OLEUM project ("Advanced solutions for assuring authenticity and quality of olive oil at global scale", funded by the European Commission within the Horizon 2020 Programme 2014–2020, GA no. 635690). The sample set included

54 EVOO, 74 VOO and 48 LOO; VOO and LOO presented different types and intensities of sensory defects (Supplementary Information Tables S1 and S2). The sample set was split randomly into a training set: 75% of the samples from each category that were used to develop and optimize the models by a leave 10%-out cross-validation, and an external validation set: conformed by the remaining 25% of the samples that were used to test the authentication strategy performed.

### 2.2. Headspace-solid phase microextraction (HS-SPME)

Virgin olive oil volatile compounds were extracted as reported by Vichi et al. (2003). Here, a Triplus autosampler (Thermo Fischer Scientific, Bremen, Germany) was used. Once 2 g of oil sample were weighted into a 10 mL vial fitted with a silicone septum, it was kept at 40 °C under agitation. After 10 min of sample conditioning, a divinylbenzene/carboxen/polymethylsiloxane (DVB/CAR/PDMS) fibre (2 cm length, 50/30 thickness) from Supelco (Bellefonte, PA) was exposed for 30 min at the sample headspace. Immediately after that, the fibre was desorbed during 10 min in the gas chromatograph injector port (260 °C), which was maintained in split-less mode for the first 5 min.

### 2.3. Gas chromatography-mass spectrometry (GC/MS)

Analytes separation was carried out by a ThermoFinnigan Trace GC equipped with a Supelcowax-10 column of 60 m  $\times$  0.25 mm i.d., 0.25  $\mu$ m film thickness (Supelco, Bellefonte, PA). The GC oven temperature was held at 40 °C for the first 10 min, then increased to 180 °C at 3 °C/min and finally to 250 °C at 15 °C/min, holding 2 min. Helium was the carrier gas at flow rate of 1.2 mL/min. An ion trap mass selective spectrometer ITQ90 MS (Thermo Fischer Scientific, Waltham, MA) coupled to the GC was used for the detection of volatiles. The temperature of the ion source was 200 °C and the transfer line, 275 °C. Mass spectra were acquired in full scan mode from  $m/z$  40 to 300 with a scan event time of 0.37 s; electron energy and emission current were 70 eV and 250  $\mu$ A, respectively.

### 2.4. Data processing and chemometrics

The total ion chromatogram (TIC) from 5.5 to 59 min was extracted for all virgin olive oil samples, obtaining a total of 8784 variables corresponding to the intensities at each scan acquired. As fibre interferences were detected in the chromatogram, fibre-related ions were excluded when extracting TICs, resulting a  $m/z$  range from 40 to 140 (without  $m/z$  73, main fibre-related ion).

In order to solve retention time shifting between samples, all TICs were aligned by means of *icoshift* algorithm (Tomasi, Savorani, & Engelsen, 2011) in Matlab®. Once aligned, the matrix obtained (8784 variables and 176 samples) was imported to SIMCA software v13.0© (Umetrics AB, Sweden) to develop and optimize the classification models. First, mean centering and unit variance scaling were applied as data pre-processing; then, a Principal Component Analysis (PCA) was carried out to explore the data and to detect potential outliers. Partial Least Square – Discriminant Analysis (PLS-DA) might find the maximum separation between classes by minimising the influence of the common features in different categories, in favour of increasing the influence of the, even subtle, differences on the volatile profile between classes (Tres, Ruiz-Samblas, van der Veer, & van Ruth, 2013). Thus, PLS-DA was applied to the training set of samples for developing the model that would allow the classification of the samples into EVOO, VOO and LOO. Two different approaches were tested: i) single PLS-DA model to classify olive oil as EVOO, VOO and LOO; ii) sequential approach consisting in two consecutive binary PLS-DA models: the first to discriminate EVOO vs non-EVOO and the second to classify the non-EVOO samples into VOO vs LOO. In all cases, models were optimized and internally validated by leave 10%-out cross-validation and the

number of latent variables (LVs) was selected according to the lowest Root Mean Square Error of cross-validation (RMSE<sub>cv</sub>) value obtained. Model performance was assessed by the % of correct classifications. Unless otherwise specified, the cut-off thresholds for the class assignment in PLS-DA models were set at a predicted value equal to 0.5, and the samples were assigned to the class leading to the highest predicted value (provided it was above 0.5). Undesirable features, such as model overfitting, random behaviour and low prediction power were assessed through permutation tests (with 20 permutations), ANOVA (on the cross-validated residuals) and  $Q^2$  score. Finally, the optimized calibration model was externally validated applying it to predict the quality grade of the samples in the validation set (25% of the total sampling). In this step, in addition to the single cut-off value (predicted value = 0.5) for class assignment, an uncertainty range was applied (predicted values from 0.35 to 0.65) in order to detect boundary samples that strictly need the organoleptic assessment by the sensory panel.

Model exploration was carried out by assessing the Variable Importance in Projection (VIP) scores, which are known to be suitable estimators of the importance of each variable in the projection used in a PLS based model (Wold, Sjöström, & Eriksson, 2001). Here, variables with a VIP score higher than 1.5 were considered as relevant for the model.

### 3. Results and discussion

#### 3.1. Data exploration

To explore the presence of outlier samples, a PCA with 5 principal components (45% of explained variance) was performed after data alignment and pre-processing as explained in section 2.4. According to Hotelling's  $T^2$  and  $Q^2$  residuals tests, two lampante samples were detected as outliers and removed from the original dataset; thereby remaining 174 oil samples. This sample set was then split randomly into a training set with the 75% of the samples from each category ( $n = 129$ ; 40 EVOO, 55 VOO and 34 LOO), which was used for the classification model development and optimization while the remaining 25% ( $n = 45$ ; 14 EVOO, 19 VOO and 12 LOO) conformed the validation set.

#### 3.2. Development of the classification strategy

The first attempt to develop the classification model was performing a PLS-DA model with the three classes (virgin olive oil commercial categories): EVOO, VOO and LOO. The model with 5 LVs led to the lowest RMSE<sub>cv</sub>. The prediction power of this model ( $Q^2$ ) and the % of correct classifications were not satisfactory enough, especially for the discrimination between EVOO and VOO classes (Table 1). It might be because EVOO category is characterized by the absence of sensory defects, while both VOO and LOO categories might present defects but in a different degree. Therefore, the discrimination of EVOO from the oil categories with defects (VOO and LOO) as a group seemed more feasible. On the other hand, considering only VOO and LOO samples, the

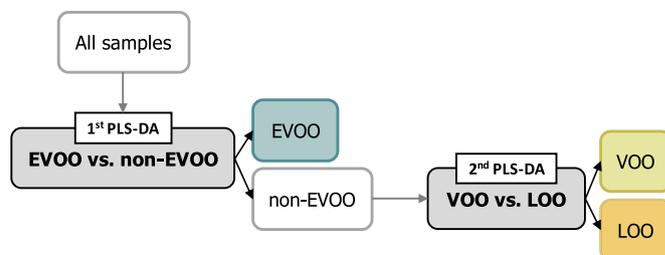
**Table 1**

Cross-validation (leave 10%-out) prediction results of the single classification model (PLS-DA) with the three commercial categories of olive oil.

	n	EVOO	VOO	LOO	Not classified (Pred. value < 0.5)	Correct class.	RMSE <sub>cv</sub>
<b>EVOO</b>	40	36	0	2	2	90.0%	0.43
<b>VOO</b>	34	0	21	5	8	61.8%	0.33
<b>LOO</b>	55	5	0	47	3	85.5%	0.33
<b>Total</b>	129	41	21	54	13	80.6%	

5LVs;  $Q^2$ : 0.38; ANOVA p-value:  $1.75 \cdot 10^{-22}$

EVOO: extra virgin olive oil, non-EVOO: non-extra virgin olive oil, VOO: virgin olive oil, LOO: lampante olive oil. RMSE<sub>cv</sub>: Root Mean Square Error of cross-validation, LV: latent variables.



**Fig. 1.** Hierarchical classification model diagram based on two consecutive binary PLS-DAs to predict virgin olive oil quality grade.

discrimination might be focused on more subtle differences between these categories. On this basis, a hierarchical classification model that combines two consecutive binary PLS-DA models was developed as shown in Fig. 1. The first PLS-DA was developed with the whole sample set ( $n = 129$ , EVOO, VOO and LOO samples), with the aim of discriminating between EVOO and non-EVOO; the second PLS-DA was built up with the observed non-EVOO samples ( $n = 89$ ) in order to classify them as VOO or LOO. The models with 4 and 5 LVs, respectively, were the ones leading to the lowest RMSE<sub>cv</sub>. As it is shown in Table 2, the classification results from the leave 10%-out cross-validation were satisfactory for both PLS-DAs, especially for the second one in which a fully correct classification was reached. The discrimination threshold in both binary models was set at 0.5. In both cases, ANOVA results ( $p < 0.05$ ) indicated that the models did not classify at random. The  $Q^2$  values obtained from the permutation tests ( $n = 20$ ) allowed to check that both models were not overfitted, because in all cases  $Q^2$  for random models resulted below 0.3.

#### 3.3. External validation of the classification strategy

The external validation was carried out by predicting the category of the 45 samples (validation set) that had not been used in the model development. As shown by the prediction plot of the validation of the first binary PLS-DA ( $n = 45$ ) (Fig. 2a), setting a single threshold at 0.5, 2 EVOO and 2 non-EVOO samples were incorrectly classified. It is remarkable that these 2 non-EVOO samples, wrongly classified as EVOO (samples 142 and 143 from Supplementary Table S2), were indeed VOO that needed to be reassessed because of a misalignment between the sensory panels. Then, in the validation of the second binary PLS-DA, samples classified as non-EVOO ( $n = 31$ ) were discriminated into VOO and LOO. The two EVOO samples that had been misclassified in the first binary PLS-DA were here assigned to VOO. Regarding the real non-EVOO samples, just 3 LOO samples were wrongly classified as VOO (Fig. 2b). Of them, one had been reassessed by the panel (sample 158 from Supplementary Table S2) and another one (sample 174 from Supplementary Table S2), presented defect intensities (fusty-muddy 3.5; winey-vinegary 3.8) very close to the limit for the VOO category ( $\leq 3.5$ ). As a global result, an 84.4% (38 out of 45) of correct classification was obtained (Table 3).

Despite of the satisfactory classification results, a more detailed observation of the predicted values (Fig. 2) evidenced that various samples had been correctly classified but presented predicted values very close to the classification threshold (0.5) (Supplementary information, Table S2). Bearing in mind that the aim of this study was to provide a screening tool to support the sensory panel rather than a method to completely replace the organoleptic assessment, we decided to set an uncertainty range to each binary PLS-DA model instead of a single cut-off threshold as a decision criterion. This uncertainty range covered the predicted values between 0.35 and 0.65, thus samples above 0.65 and below 0.35 would be directly classified as EVOO, VOO or LOO (depending on the binary PLS-DA). On the contrary, no definitive category assignments would be raised for samples that resulted boundary when applying the single threshold and now fall within the

**Table 2**

Cross-validation (leave 10%-out) prediction results of the hierarchical classification model for olive oil commercial category assignment, based on two sequential binary PLS-DAs (1<sup>st</sup>: EVOO vs. Non-EVOO; 2<sup>nd</sup>: VOO vs. LOO) with a discrimination threshold set at 0.5.

1st PLS-DA (EVOO vs non-EVOO)				2nd PLS-DA (VOO vs LOO)				
	n	EVOO	Non-EVOO	Correct class		VOO	LOO	Correct class
<b>EVOO</b>	40	37	3	92.5%	<b>VOO</b>	55	55	100%
<b>Non-EVOO</b>	89	5	84	94.4%	<b>LOO</b>	34	0	100%
<b>Total</b>	129	42	87	<b>93.8%</b>	<b>Total</b>	89	55	<b>100%</b>

4LVs; RMSEcv: 0.30; Q<sup>2</sup>: 0.55; ANOVA p-value: 1.53·10<sup>-19</sup>      5LVs; RMSEcv: 0.36; Q<sup>2</sup>: 0.48; ANOVA p-value: 5.56·10<sup>-07</sup>

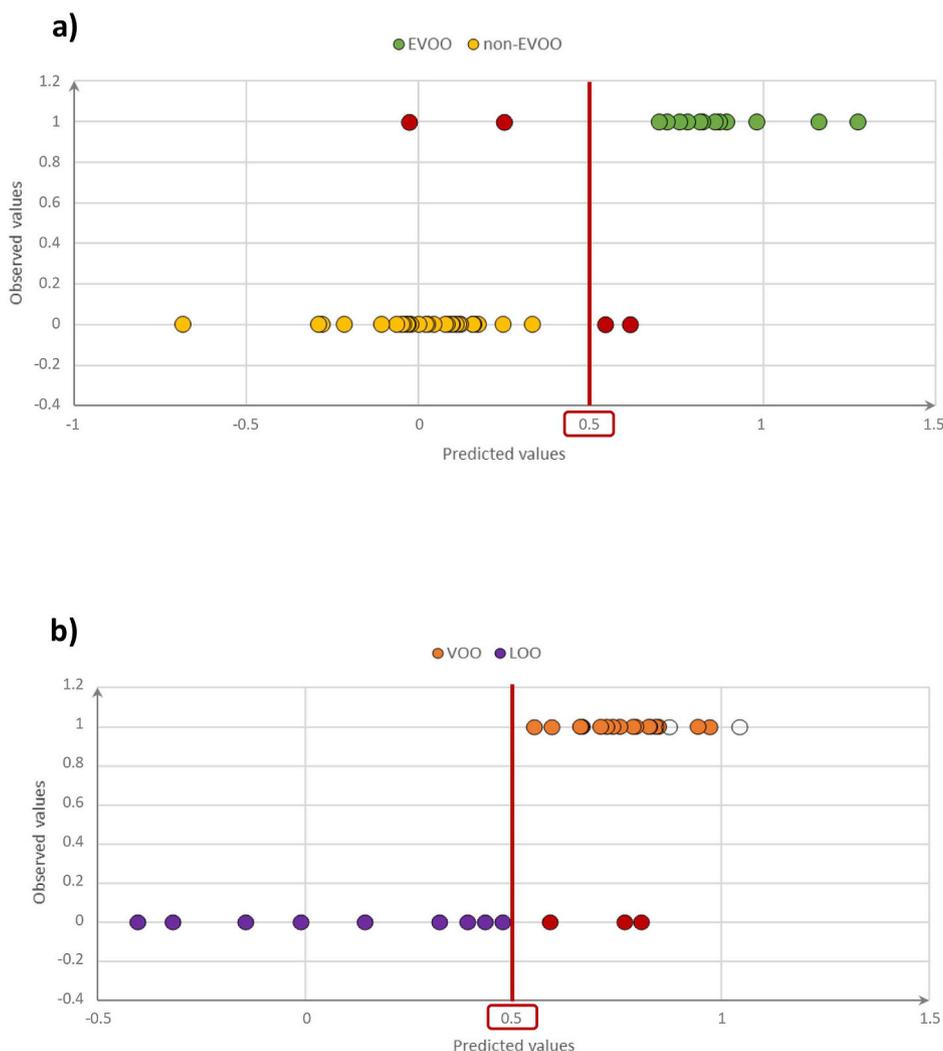
EVOO: extra virgin olive oil, non-EVOO: non-extra virgin olive oil, VOO: virgin olive oil, LOO: lampante olive oil. RMSEcv: Root Mean Square Error of cross-validation, LV: latent variable.

uncertainty range (Supplementary information, Table S2). The category of these uncertain samples would then be determined by the sensory panel. In this way, the number of samples to be sensory assessed would be reduced focusing the evaluation on the most critical ones, increasing the panel efficiency and reducing its workload.

Applying the aforementioned range to the validation samples, a total of 8 out of 45 samples were classified as uncertain (Fig. 3a–b). 3 of these uncertain samples had been wrongly classified by the single threshold approach (samples 142, 143 and 153 from Supplementary information, Table S2), and included the 2 VOO samples that needed

panel reassessment. In this way, only 4 samples were misclassified instead of 7, increasing the classification rate of the screening method from 84.4%, obtained by applying a single threshold, to 89.2%. Moreover, if uncertain samples (17.8%) would be graded by the sensory panel, which is the reference method, the percent of samples with a reliable assignment to a commercial category would be 91.1% (Table 3).

Combining the instrumental method to the reference sensory method accomplishes a considerable decrease of the panel workload (around 80% of reduction) while maintaining a highly reliable



**Fig. 2.** Observed vs. predicted values of external validation set with a classification threshold of 0.5: a) 1<sup>st</sup> classification model: “extra virgin (EVOO) vs. non-extra virgin (non-EVOO)”; b) 2<sup>nd</sup> classification model: “virgin (VOO) vs. lampante (LOO)”. Misclassified samples in red (EVOOs misclassified as non-EVOOs in Fig. 2a are not coloured in Fig. 2b). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

**Table 3**  
Global outcome of external validation of the classification approach based on the hierarchical model by combining two consecutive binary PLS-DAs.

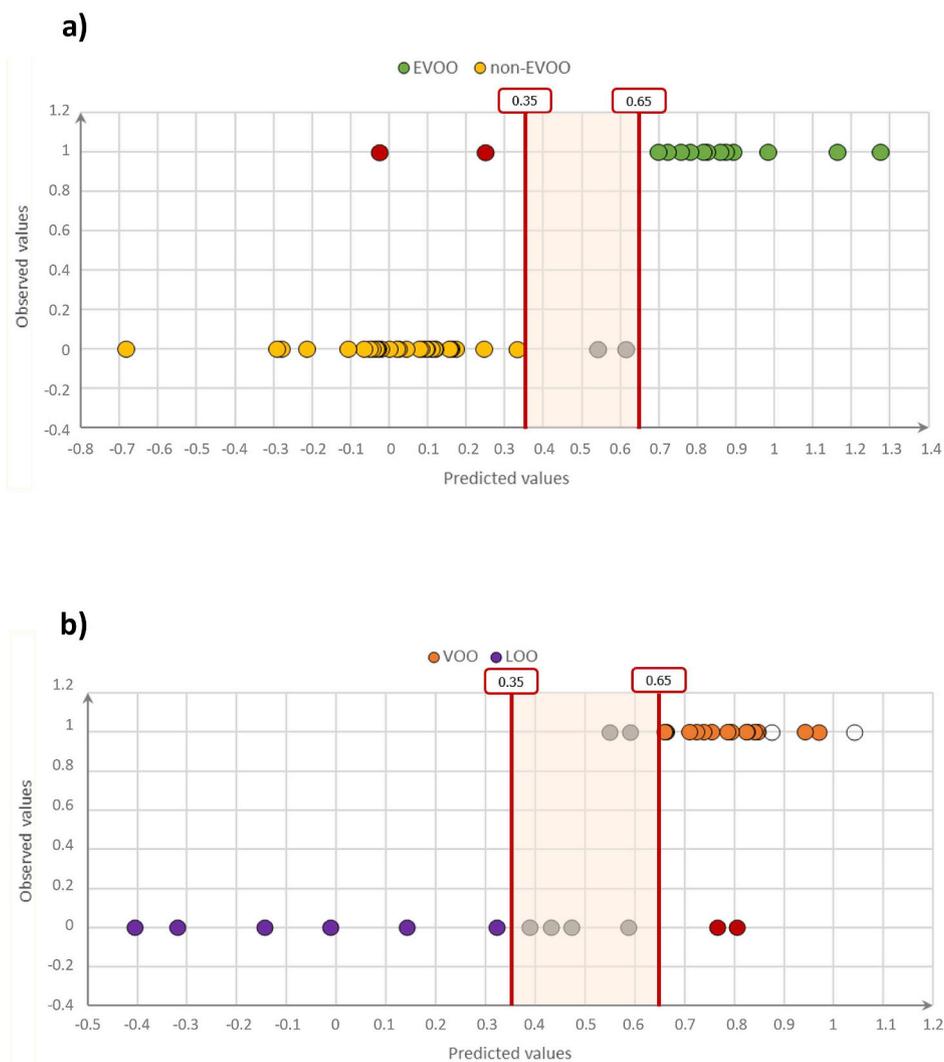
Single classification threshold (Prediction value = 0.5)					
	n	Predicted category			Correct class.
		EVOO	VOO	LOO	
EVOO	14	12	2	0	85.7% (12/14)
VOO	19	2	17	0	89.5% (17/19)
LOO	12	0	3	9	75% (9/12)
Total	45	14	22	9	84.4% (38/45)

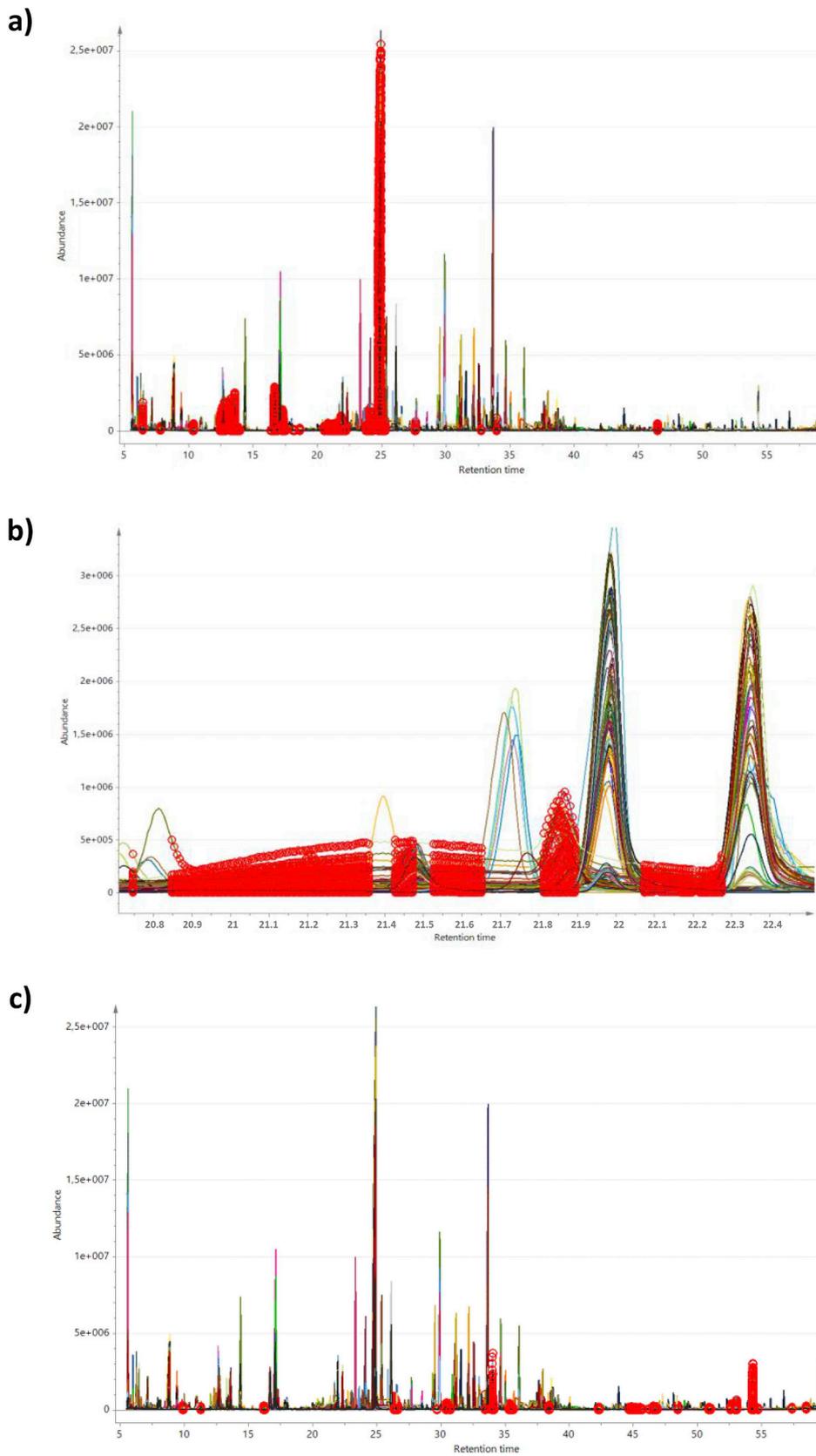
Uncertainty range (Prediction values from 0.35 to 0.65)									
	n	Predicted category				Instrumental screening tool		Screening tool + reference method	
		EVOO	Uncertain (EVOO vs. VOO)	VOO	Uncertain (VOO vs. LOO)	LOO	Uncertainty <sup>a</sup>	Correct class <sup>b</sup>	Reliable assignment <sup>c</sup>
EVOO	14	12	0	2	0	0%	85.7% (12/14)	85.7% (12/14)	
VOO	19	0	2	15	2	21.1% (4/19)	100% (15/15)	100% (19/19)	
LOO	12	0	0	2	4	33.3% (4/12)	75% (6/8)	83.3% (10/12)	
Total	45	12	2	19	6	17.8% (8/45)	89.2% (33/37)	91.1% (41/45)	

EVOO: extra virgin olive oil, non-EVOO: non-extra virgin olive oil, VOO: virgin olive oil, LOO: lampante olive oil.

- <sup>a</sup> Samples that fall into the uncertainty range are not classified into a category by the screening tool, to be further assessed by the sensory panel.
- <sup>b</sup> Samples correctly classified by the screening tool out of the total of samples classified (excluding uncertain samples).
- <sup>c</sup> Sum of samples correctly classified by the screening tool and uncertain samples to be assessed by the sensory panel (reference method).



**Fig. 3.** Observed vs. predicted values of external validation set with a classification range from 0.35 to 0.65: **a)** 1<sup>st</sup> classification model: “extra virgin (EVOO) vs. non-extra virgin (non-EVOO)”; **b)** 2<sup>nd</sup> classification model: “virgin (VOO) vs. lampante (LOO)”. Misclassified samples in red (EVOOs misclassified as non-EVOOs in Fig. 2a are not coloured in Fig. 2b). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)



**Fig. 4.** Comparison between the most relevant variables for extra virgin and lampante olive oil class: **a)** Raw chromatograms from oil samples aligned: variables with a VIP score > 1.5 for the extra virgin class (from the 1<sup>st</sup> model “extra virgin vs. non-extra virgin”) spotted in red; **b)** Fig. 4a scaled up in the region where Z-3-hexenal elutes (Approx. RT: 20.8–22.2 min); **c)** Raw chromatograms from oil samples aligned: variables with a VIP score > 1.5 for the lampante virgin class (from the 1<sup>st</sup> model “extra virgin vs. non-extra virgin”) spotted in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

classification of samples (error rate under 10%). In fact, this strategy allows achieving misclassification rates in external validation comparable to those obtained by other approaches (Cecchi et al., 2019; Contreras et al., 2019; López-Feria et al., 2007; Sales et al., 2017, 2019), but by a simple, automatable and low-expensive technique available in most of the control laboratories.

### 3.4. Exploration of classification models

To investigate the reliability of the models, we studied which variables were relevant for them and if they agreed with known volatile compounds in virgin olive oils. For this purpose, we assessed the variables with a strong weight for each class in each binary discrimination models through the VIP parameter. We took into consideration the variables (acquisition points in the volatile fraction fingerprint) with VIP scores above 1.5 for each class, considering that these were the ones that better describe the classes. Fig. 4 show the aligned TICs with the most important variables for the EVOO (Fig. 4a and b) and LOO class (Fig. 4c), spotted in red. We can observe that some of the relevant variables corresponded to minor chromatographic peaks and/or to peaks whose shape (Fig. 4b) would entail a hard integration under a targeted approach. This is one of the advantages of the fingerprinting approach because by using the whole analytical signal it allows to exploit the full information, without the need of peak identification and integration.

Then, we related these variables with the chromatographic peaks corresponding to volatile compounds, identified by comparing their mass spectra and retention indices to the ones available in the NIST mass spectra library and in the literature. Finally, we verified if these compounds agree with the previous knowledge about the sensory impact of certain volatile compounds in virgin olive oils (Angerosa et al., 2004; Kalua et al., 2007). Volatile compounds previously associated to virgin olive oil positive attributes (Angerosa et al., 2000) were here important for the EVOO identification, such as the ones coming from the lipoxygenase pathway, both C5 compounds (pentene dimers, 1-penten-3-one and 1-penten-3-ol) and C6 compounds (E-2-hexenal, hexyl acetate, Z-3-hexenyl acetate and 1-hexanol) and secondary plant metabolites (terpene hydrocarbons). Conversely, volatile compounds derived from autoxidation process (nonanal and 2,4-hexadienal) as well as microbial metabolites (organic acids and fermentative alcohols) were identified as important for the classification of LOO samples (Fig. 4c). These compounds were previously reported in several studies for being related with virgin olive oil defects (Morales, Luna, & Aparicio, 2005). Identification data of the above-mentioned compounds are listed in Table S3 (Supplementary information) together with previous reports of their presence in EVOO or defective olive oils. The VIPs were also evaluated for the VOO class (data not shown) and the most important variables for this class were like those observed for the EVOO class in the first binary PLS-DA. It is indeed a reasonable finding that agrees with the unsatisfactory result obtained when developing the unique PLS-DA model with the three quality grades.

## 4. Conclusions

Volatile profile of virgin olive oil is highly correlated to its organoleptic quality, given that volatile compounds are responsible for both positive and negative olfactory attributes. In pursuit of a suitable analytical method to analyse olive oil's volatile profile, HS-SPME-GC-MS proved to be appropriate because it is an automatable solvent-free technique and low amount of sample is required. Following a fingerprinting approach, where each chromatographic data point is considered as one variable, and the subsequent chemometric analysis based on a hierarchical classification model allowed to obtain a successful classification rate for both internal and external validation. The usefulness of applying an uncertainty range instead of a single classification threshold in the classification strategy has been verified. By doing this,

this classification strategy would serve as instrumental screening method, providing satisfactory results (89.2% of correct classification) by itself, and supporting the sensory panels by lighten their workload (almost 80% of reduction) while maintaining a highly reliable classification of samples (error rate under 10%). Nevertheless, even though these are promising results, further samples from different crop-years are necessary in order to develop a more consistent classification strategy, especially to increase the representation of the LOO class.

## Declaration of competing interest

Authors declare no conflict of interest.

## Acknowledgements

This work was developed in the context of the project OLEUM “Advanced solutions for assuring authenticity and quality of olive oil at global scale”, funded by the European Commission within the Horizon 2020 Program (2014–2020, grant agreement no. 635690). The information and views set out in this article are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein. B. Quintanilla-Casas thanks the Spanish Ministry of Science, Innovation and Universities pre-doctoral fellowship FPU16/01744. A. Tres thanks the Spanish Ministry of Economy, Industry and Competitiveness “Juan de la Cierva” post-doctoral fellowship (JCI-2012\_13412) and the Ministry of Science, Innovation and Universities “Ramón y Cajal” postdoctoral fellowship (RYC-2017-23601).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.lwt.2019.108936>.

## References

- Angerosa, F., Mostallino, R., Basti, C., & Vito, R. (2000). Virgin olive oil odour notes: Their relationships with volatile compounds from the lipoxygenase pathway and secoiridoid compounds. *Food Chemistry*, 68, 283–287. [https://doi.org/10.1016/S0308-8146\(99\)00189-2](https://doi.org/10.1016/S0308-8146(99)00189-2).
- Angerosa, F., Servili, M., Selvaggini, R., Taticchi, A., Esposto, S., & Montedoro, G. (2004). Volatile compounds in virgin olive oil: Occurrence and their relationship with the quality. *Journal of Chromatography A*, 1054, 17–31. <https://doi.org/10.1016/j.chroma.2004.07.093>.
- Aparicio, R., Morales, M. T., & García-González, D. L. (2012). Towards new analyses of aroma and volatiles to understand sensory perception of olive oil. *European Journal of Lipid Science and Technology*, 114, 1114–1125. <https://doi.org/10.1002/ejlt.201200193>.
- Benelli, G., Caruso, G., Giunti, G., Cuzzola, A., Saba, A., & Gucci, R. (2015). Changes in olive oil volatile organic compounds induced by water status and light environment in canopies of *Olea europaea* L. trees. *Journal of the Science of Food and Agriculture*, 95, 2475–2481. <https://doi.org/10.1002/jsfa.6977>.
- Berrueta, L. A., Alonso-Salces, R. M., & Herberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatography A*, 1158, 196–214. <https://doi.org/10.1016/j.chroma.2007.05.024>.
- Borràs, E., Ferré, J., Boqué, R., Aceña, L., Calvo, A., & Busto, O. (2016). Prediction of olive oil sensory descriptors using instrumental data fusion and partial least squares (PLS) regression. *Talanta*, 155, 116–123. <https://doi.org/10.1016/j.talanta.2016.04.040>.
- Cecchi, L., Migliorini, M., Giambanelli, E., Rossetti, A., Cane, A., Melani, F., et al. (2019). Headspace solid-phase microextraction–gas chromatography–mass spectrometry quantification of the volatile profile of more than 1200 virgin olive oils for supporting the panel test in their classification: Comparison of different chemometric approaches. *Journal of Agricultural and Food Chemistry*, 67, 9112–9120. <https://doi.org/10.1021/acs.jafc.9b03346>.
- Commission Regulation (ECC) No 2591/91 of 11 July 1991 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis. Consolidated version with amendments.
- Commission Implementing Regulation (EU) No 2016/1227 of 27 July 2016 amending regulation (ECC) No 2568/91 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis.
- Conte, L., Bendini, A., Valli, E., Lucci, P., Moret, S., Maquet, A., et al. (2019). Trends in food science & technology olive oil quality and authenticity: A review of current EU

- legislation, standards, relevant methods of analyses, their drawbacks and recommendations for the future. *Trends in Food Science & Technology*. <https://doi.org/10.1016/j.tifs.2019.02.025>.
- Contreras, M., Jurado-Campos, N., & Arce, L. (2019). A robustness study of calibration models for olive oil classification: Targeted and non-targeted fingerprint approaches based on GC-IMS. *Food Chemistry*, 288, 315–324. <https://doi.org/10.1016/j.foodchem.2019.02.104>.
- International Olive Council (IOC) T.20 Doc. No 15/Rev. 10 (2018). Sensory analysis of olive oil - method for the organoleptic assessment of virgin olive oil.
- Kalua, C. M., Allen, M. S., Bedgood, D. R., Jr., Bishop, A. G., Prenzler, P. D., & Robards, K. (2007). Olive oil volatile compounds, flavour development and quality: A critical review. *Food Chemistry*, 100, 273–286. <https://doi.org/10.1016/j.foodchem.2005.09.059>.
- Lerma-García, M. J., Cerretani, L., Cevoli, C., Simó-Alfonso, E. F., Bendini, A., & Gallina Toschi, T. (2010). Use of electronic nose to determine defect percentage in oils. Comparison with sensory panel results. *Sensors and Actuators B: Chemical*, 147, 283–289. <https://doi.org/10.1016/j.snb.2010.03.058>.
- López-Feria, S., Cárdenas, S., García-Mesa, J. A., & Valcárcel, M. (2007). Usefulness of the direct coupling headspace – mass spectrometry for sensory quality characterization of virgin olive oil samples. *Analytica Chimica Acta*, 583, 411–417. <https://doi.org/10.1016/j.aca.2006.10.027>.
- Morales, M. T., Luna, G., & Aparicio, R. (2005). Comparative study of virgin olive oil sensory defects. *Food Chemistry*, 91, 293–301. <https://doi.org/10.1016/j.foodchem.2004.06.011>.
- Morales, M. T., Rios, J. J., & Aparicio, R. (1997). Changes in the volatile composition of virgin olive oil during oxidation: Flavors and off-flavors. *Journal of Agricultural and Food Chemistry*, 45(7), 2666–2673. <https://doi.org/10.1021/jf960585>.
- Oliver-Pozo, C., Aparicio-Ruiz, R., Romero, I., & García-González, D. L. (2015). Analysis of volatile markers for virgin olive oil aroma defects by SPME-GC/FID: Possible sources of incorrect data. *Journal of Agricultural and Food Chemistry*, 63, 10477–10483. <https://doi.org/10.1021/acs.jafc.5b03986>.
- Procida, G., Giomo, A., Cichelli, A., & Conte, L. S. (2005). Study of volatile compounds of defective virgin olive oils and sensory evaluation: A chemometric approach. *Journal of the Science of Food and Agriculture*, 85, 2175–2183. <https://doi.org/10.1002/jsfa.2122>.
- Sales, C., Cervera, M. I., Gil, R., Portolés, T., Pitarch, E., & Beltran, J. (2017). Quality classification of Spanish olive oils by untargeted gas chromatography coupled to hybrid quadrupole-time of flight mass spectrometry with atmospheric pressure chemical ionization and metabolomics-based statistical approach. *Food Chemistry*, 216, 365–373.
- Sales, C., Portolés, T., Johnsen, L. G., Danielsen, M., & Beltran, J. (2019). Olive oil quality classification and measurement of its organoleptic attributes by untargeted GC – MS and multivariate statistical-based approach. *Food Chemistry*, 271, 488–496. <https://doi.org/10.1016/j.foodchem.2018.07.200>.
- Tomasi, G., Savorani, F., & Engelsen, S. B. (2011). icoshift : An effective tool for the alignment of chromatographic data. *Journal of Chromatography A*, 1218, 7832–7840. <https://doi.org/10.1016/j.chroma.2011.08.086>.
- Tres, A., Ruiz-Samblas, C., van der Veer, G., & van Ruth, S. M. (2013). Geographical provenance of palm oil by fatty acid and volatile compound fingerprinting techniques. *Food Chemistry*, 137, 142–150. <https://doi.org/10.1016/j.foodchem.2012.09.094>.
- Vichi, S., Castellote, A. I., Pizzale, L., Conte, L. S., Buxaderas, S., & Lopez-Tamames, E. (2003). Analysis of virgin olive oil volatile compounds by headspace solid-phase microextraction coupled to gas chromatography with mass spectrometric and flame ionization detection. *Journal of Chromatography A*, 983, 19–33. [https://doi.org/10.1016/S0021-9673\(02\)01691-6](https://doi.org/10.1016/S0021-9673(02)01691-6).
- Vichi, S., Guadayol, J. M., Caixach, J., López-Tamames, E., & Buxaderas, S. (2007). Comparative study of different extraction techniques for the analysis of virgin olive oil aroma. *Food Chemistry*, 105, 1171–1178. <https://doi.org/10.1016/j.foodchem.2007.02.018>.
- Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58, 109–130. [https://doi.org/10.1016/S0169-7439\(01\)00155-1](https://doi.org/10.1016/S0169-7439(01)00155-1).