Compliance with EU vs. extra-EU labelled geographical provenance in virgin olive oils: A rapid untargeted chromatographic approach based on volatile compounds

Published Version:
Compliance with EU vs. extra-EU labelled geographical provenance in virgin olive oils: A rapid untargeted chromatographic approach based on volatile compounds / Rosa Palagano, Enrico Valli, Chiara Cevoli, Alessandra Bendini, Tullia Gallina Toschi. - In: LEBENSMITTEL-WISSENSCHAFT + TECHNOLOGIE. - ISSN 0023-6438. - STAMPA. - 130:(2020), pp. 109566-109572. [10.1016/j.lwt.2020.109566]

This version is available at: https://hdl.handle.net/11585/761698 since: 2020-11-06

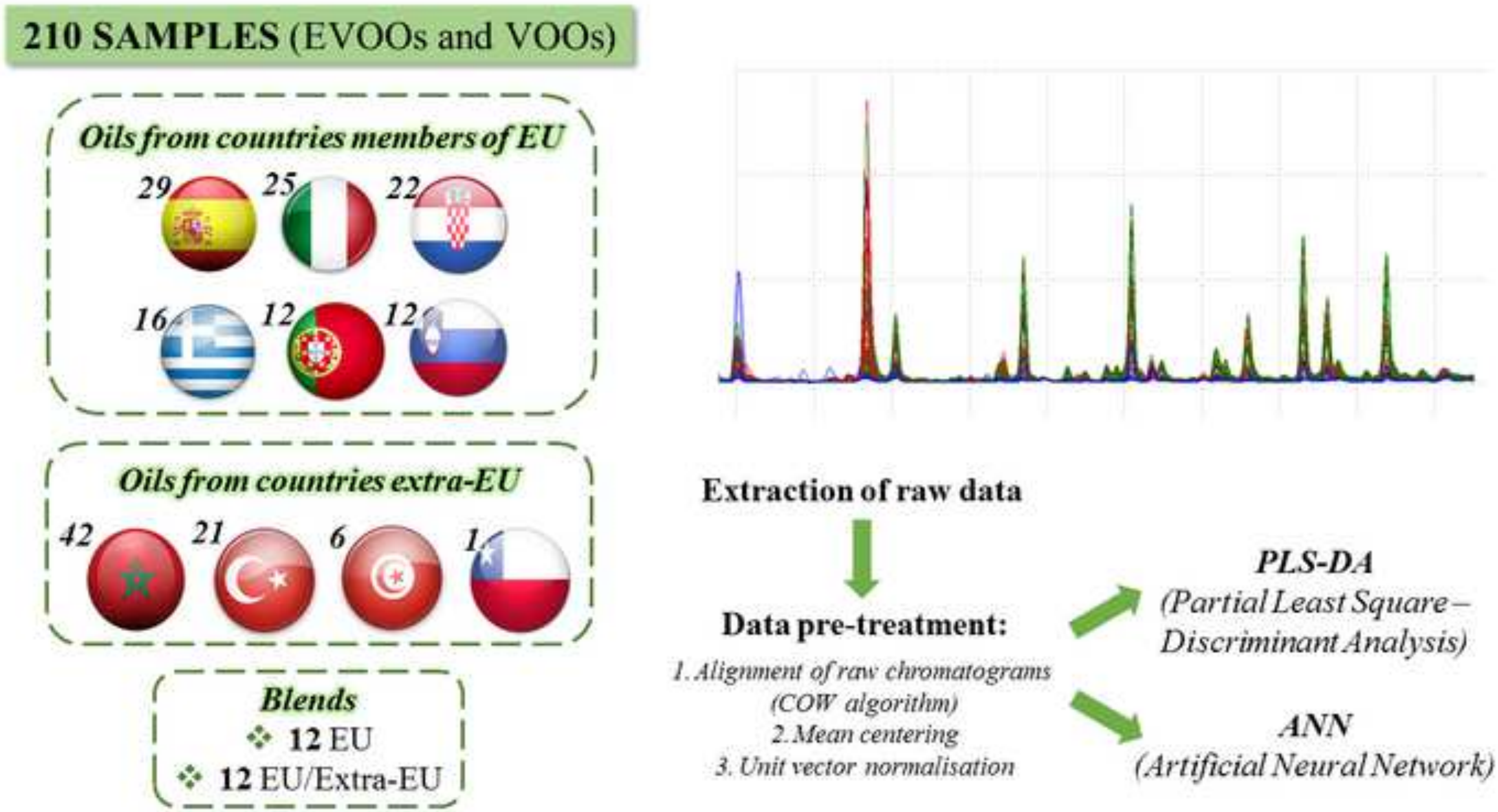
Published:

DOI: http://doi.org/10.1016/j.lwt.2020.109566

(Article begins on next page)

Highlights

- Flash-GC was applied to discriminate VOOs with different geographical origin
- Chemometric data elaborations with untargeted approach were realized
- PLS-DA and ANN classification models were built and validated
- Satisfactory percentages of samples correctly classified were obtained


210 SAMPLES (EVOOs and VOOs)
Oils from countries members of EU
29
25
22
16
12
12
Oils from countries extra-EU
42
21
6
Blends
12 EU
12 EU/Extra-EU
Extraction of raw data
Data pre-treatment:
1. Alignment of raw chromatograms (COW algorithm)
2. Mean centering
3. Unit vector normalisation
PLS-DA
(Partial Least Square – Discriminant Analysis)
ANN
(Artificial Neural Network)

Discrimination of virgin olive oils with different geographical origin: a rapid untargeted chromatographic approach based on volatile compounds

Rosa Palagano[a], Enrico Valli[a], Chiara Cevoli[a], Alessandra Bendini[a,*], Tullia Gallina Toschi[a]

[a] Department of Agricultural and Food Sciences (DISTAL), University of Bologna, Cesena, 47521, Italy

[*] Corresponding author: alessandra.bendini@unibo.it (Alessandra Bendini), Department of Agricultural and Food Sciences (DISTAL) - University of Bologna, Piazza Goidanich 60, Cesena (FC), 47521, Italy, Tel +39 0547 338121.

Abstract

Many studies have shown that the geographic origin is one of the most influencing factors in consumers' choice of olive oil. To avoid misleading, European regulation has established specific rules to report the geographical origin of extra virgin (EVOOs) and virgin olive oils (VOOs) on the product label, but an official analytical procedure to verify this information has not been yet defined. In this work, a flash gas chromatography untargeted approach for determination of volatile compounds, followed by a chemometric data elaboration, is proposed for discrimination of EVOOs and VOOs according to their geographical origin (EU and Extra-EU). A set of 210 samples was analyzed and two different classification techniques were used, one linear (Partial Least Square-Discriminant Analysis, PLS-DA) and one non-linear (Artificial Neural Network, ANN). The two models were also validated using an external data set. Satisfactory results were obtained for both chemometric approaches: considering the PLS-DA, 89% and 81% of EU and Extra-EU samples, respectively, were correctly classified; for ANN the percentages were 93% and 89%, respectively. These results confirm the reliability of the method as a rapid approach to discriminate EVOOs and VOOs according to their geographical provenance. The methodology is fast (around 200 sec for each chromatographic run), easy to use (no sample treatment is required), and the data were effectively elaborated by applying both linear and non-linear chemometric techniques.

Key words

Virgin olive oil; Geographical origin; PLS-DA; Untargeted approach; Volatile compounds; ANN.

Abbreviations

ANN: Artificial Neural Network; EVOO: Extra Virgin Olive Oil; FGC: Flash Gas Chromatography; PLS-DA: Partial Least Square - Discriminant Analysis; VOO: Virgin Olive Oil.

1. Introduction

Over the last 40 years many investigations have been focused on understanding what attributes are important determinants in consumer choice, which have highlighted that the geographic origin is one of the most influencing factors for olive oil (Dekhili, Sirieix, & Cohen, 2011; Del Giudice, Cavallo, Caracciolo, & Cicia, 2015).

In order to ensure that consumers are not misled, the fourth article of the EU Reg. 29/2012 establishes that *"Extra virgin and virgin olive oil shall bear a designation of origin on the labelling"*. This means that for extra virgin (EVOOs) and virgin olive oils (VOOs) commercialized within the EU, it is mandatory to specify the geographical provenance on the label of the product following specific rules. If an oil comes from an EU Member State or third country, a reference to the EU Member State, to the EU, or to the third country must to be reported. In the case of blends of oils originating from more than one EU Member State or third country, one of the following mentions must be used: 'blend of olive oils of European Union origin' or a reference to the EU; 'blend of olive oils not of European Union origin' or a reference to origin outside the EU; 'blend of olive oils of European Union origin and not of European Union origin' or a reference to origin within the EU and outside the EU. An exception is the case where the olives were harvested in an EU Member State or third country other than that in which the mill where the oil was extracted is located. In this case, the designation of origin shall contain the following wording: *'(extra) virgin olive oil obtained in (the Union or the name of the Member State concerned) from olives harvested in (the Union or the name of the Member State or third country concerned)'*.

However, the regulation does not specify an official analytical procedure to verify the conformity of the label-declared geographical origin, and this has raised the interest of researchers to develop a reliable and effective method for purposes of authentication (Conte et al., 2019). During the last years, different analytical techniques have been applied in order to find potentially useful markers and efficient instrumental approaches that are able to discriminate olive oils according to their geographical origin.

In this regard, traditional chromatographic techniques, analyzing both major and minor compounds either individually or in a combined way, coupled or not with specific statistical chemometric data elaboration, have been investigated. A study in 2009 (García-González et al., 2009) proposed the application of artificial neural network (ANN) models for different levels of geographical classification (country, region, province, PDO) on a set of 687 EVOOs and VOOs from Spain, Italy, and Portugal, which were chemically characterized for the content of fatty acids, hydrocarbons, sterols, and alcohols. Other researchers evaluated the triacylglycerol (TAG) content and composition to discriminate Moroccan oils (Bajoub et al., 2016) and Croatian samples (Peršurić, Saftić, Mašek, & Kraljević Pavelić, 2018). In addition, the stereospecific distribution of fatty acids in TAGs was reported to be useful in discriminating olive oils from different areas of North-Eastern Italy (Vichi, Pizzale, & Conte, 2007). Specific metabolites such as sterols and phenolic compounds have been investigated to identify the optimal markers, and may be a promising approach to discriminate oils according to geographical origin (Giacalone, Giuliano, Gulotta, Monfreda, & Presti, 2015; Ben Mohamed et al., 2018; Ghisoni et al., 2019). Interesting findings have also been recently reported on sesquiterpene hydrocarbons as geographical markers (Quintanilla-Casas et al., 2020). Moreover, volatile compounds have been amply studied by applying different instrumental techniques combined with chemometric data elaborations (Kosma et al., 2017; Bajoub et al., 2018; Lukić, Carlin, Horvat, & Vrhovsek, 2019).

Furthermore, rapid and innovative instrumental approaches have been developed and tested in order to deal with the need for simple, rapid, and environmentally friendly techniques (Valli et al., 2016). This critical review (Valli et al., 2016) reports an overview of the principal applications of optical techniques (UV-Vis, NIR, MIR, RAMAN, NMR, and fluorescence spectroscopy), methods based on electrical characteristics, and instruments equipped with electronic chemical sensors (electronic nose and tongue) for discrimination of EVOOs and VOOs according to their geographical provenance. In addition to these approaches, other promising techniques include stable isotopes analysis (Angerosa et al., 1999; Chiocchini, Portarena, Ciolfi, Brugnoli, & Lauteri, 2016; Bontempo et al., 2019), multi-element fingerprint (Sayago, González-Domínguez, Beltrán, & Fernández-Recamales, 2018), differential scanning calorimetry (Mallamace et al., 2017), and GC-IMS (Gerhardt, Birkenmeier, Sanders, Rohn, & Weller, 2017).

Melucci and co-workers (Melucci et al., 2016) proposed the application of a Flash Gas Chromatography Electronic Nose (Heracles II) and a multivariate data analysis to control the compliance of information on geographic origin declared in the label ("100% Italian" vs "non-100% Italian") for the first time. This instrumental approach allows to realize the headspace analysis in short time and the results are processed by chemometric tools following an untargeted approach. For this reason, it can be considered as a fingerprint method, since the data can be elaborated for sample classification that is not aimed towards identification and quantification of specific analytes. Following these preliminary results and the actual need for a rapid and effective method for geographical authentication of VOOs, the aim of this work was the application of flash gas chromatography (Heracles II) for rapid discrimination of 210 EVOOs and VOOs according to geographical provenance. In this case, the categories considered for samples classification were EU member states vs third countries, and the data obtained were elaborated by applying two different classification techniques, one linear (Partial Least Square-Discriminant Analysis, PLS-DA) and one non-linear (Artificial Neural Network, ANN).

2. Materials and methods

2.1 Samples

A total of 210 EVOOs and VOOs with a different geographical origin were collected directly from companies that were also asked to provide, when available, information about location of the mill, type of plant used, olive variety, and commercial category (Table S1, Supplementary material). Considering that the indication of the geographical origin on the product label is mandatory for EVOOs and VOOs, samples belonging to both these two categories were included in this study.

According to geographical provenance, samples were distributed in 3 classes (Table 1): “EU” for oils coming from EU member states; “Extra-EU” for oils coming from third countries (outside EU); “Blends” for samples obtained by mixing oils coming from different EU state members or oils coming from EU state members and third countries.

Aliquots of each sample (50 mL) were stored at -18 °C in plastic dark bottles. Oil were defrosted for at least 12 h and stored at 12°C before analysis.

2.2 Volatile compounds analysis by Flash Gas Chromatography

The analysis of volatile compounds was carried out using the Flash Gas Chromatography Electronic Nose Heracles II (Alpha MOS, Toulouse, France). The instrument was equipped with two metal capillary columns (MXT-5: 5% diphenyl, 95% methylpolysiloxane, and MXT-1701: 14% cyanopropylphenyl, 86% methylpolysiloxane, for both columns: 10 m length, 180 µm internal diameter, 0.4 µm film thickness) working in parallel mode and different in polarity of the stationary phase. This permits slight differences in the separation capability of molecules detected by a FID applied at the end of each column.

Each sample was analysed in triplicate, weighing 2 ± 0.1 g of oil in a 20 mL vial sealed with a magnetic plug. For analysis, the vial was placed in a shaker oven for 20 min at 40 °C and 500 rpm. Next, 5 mL of the headspace were collected, introduced in a splitless injector (injector temperature 200 °C, injection speed 100 µL/sec, carrier gas flow, to ensure a fast transfer of the sample from the inlet to the trap, 30mL/min), and adsorbed on a Tenax® TA trap maintained at 40 °C for 60 sec to concentrate the analytes. The syringe temperature was set at 70 °C. Subsequently, desorption was obtained by increasing the trap temperature to 240 °C in 93 sec and the sample was injected (pressure of the carrier gas at columns' head 40 kPa.) and split (split flow 5 mL/min) into the two columns. The thermal program started at 40 °C (held for 2 sec), increased up to 80 °C at 1 °C/sec, and then to 250 °C at 3 °C/sec. Hydrogen was used as the carrier gas with a pressure from 40 kPa to 64 kPa, increasing with a rate of 0.2 kPa/sec. At the end of each column, a FID detector (detector temperature 260 °C) was placed and the acquired signal was digitalized every 0.01 sec. The software used to control the instrument was AlphaSoft version 14.5.

2.3 Data processing

For the data analysis, the full chromatograms were processed by applying chemometric elaborations with an untargeted approach. The raw data of each chromatogram (intensity values for each point of the chromatogram considering that the signal was digitalized every 0.01 sec, Palagano et al., 2019 [Dataset]) were exported from the software of the instrument and the data set with all the samples was imported into MatlabR2018a®. As data pre-treatment, chromatograms were aligned by COW (Correlation Optimized Warping) algorithm (Tomasi, Van Den Berg, & Andersson, 2004) and autoscaled (mean-centering followed by division of each column (variable) by the standard deviation of that column). Preliminary tests showed that chromatograms obtained from the MXT-5 column had a discriminant power higher than the other one (MXT-1701) and for this reason the classification models were developed considering only this column. Considering the reduced number of samples for the classes “Blend EU” and “Blend EU-Extra EU”, these oils were grouped together with “EU” and “Extra-EU” samples, respectively. This means that for the data elaboration only two sample categories were considered: “EU” and “Extra-EU”.

Two different statistical techniques were used to classify samples according to their geographical origin, the first (PLS-DA) based on a linear approach, and the second (ANN) on a non-linear approach.

In particular, the PLS-DA model was built using the PLS Toolbox for Matlab2018a®: intensity values of each point of the chromatogram, for a total of 19,900 data points, were used as variables X (matrix X), while the origin (“EU” and “Extra-EU”) was implemented as variable Y (binary variables, 0 - 1). The sample data set was split into a calibration/full-cross validation set (75% of the sample) and an external validation set (25% of the sample) using the Kennard-Stone method (selects samples that best span the same range as the original data, but with an even distribution of samples across the same range) (Daszykowski, Walczak, & Massart, 2002). The threshold value useful to define the category of each sample was defined using a probabilistic approach based on Bayes's rule.

The ANN model was performed by using the Neural Net Pattern Recognition tool for Matlab2018a®. Specifically, a Multi-Layer Perceptron (MLP) neural network was built to predict the specific class to which samples belong using a non-linear method. For input and hidden layers, linear and logistic activation functions, respectively, were used, while for output layer the SoftMax function was applied. From a statistical point of view, with the SoftMax activation function and cross entropy error, the output is interpretable as posterior probabilities for categorical target variables (Bishop, 1995). One nominal output variable is returned, assuming that the target output is 1.0 in the correct class output, and 0.0 in the non-correct class. Looking for the best classification ability, different node numbers in the hidden layer and combinations were tested. The convergence of ANN was ruled by a back propagation algorithm. The original data set was randomly divided into a training set (60%), verification set (20%), and test set (20%). The training set was used to calculate the transfer function parameters of the network, the verification set to indicate possible over-learning, and the test set was treated as an unknown, the correct classification of which indicates that the neural network is performing well. It was checked that samples from both classes were contained in the test set.

3. Results and discussion

A set of 210 EVOOs and VOOs were analyzed for their volatile profile by flash gas chromatography. Considering the large amount of data and aim of this work, chemometric elaborations following an untargeted approach were carried out.

For elaborations, samples were grouped into two categories: "EU", that included oils from single EU state members and blends of oils from different EU countries, "Extra-EU" that consisted of oils from single countries outside the European Union and blends of oils from the EU and third countries.

In Figure 1-a the mean chromatogram of “EU” and “Extra-EU” categories, obtained averaging the intensity of each variable for all “EU” or “Extra-EU” samples, is reported: even if almost all peaks are concentrated in the initial part of the chromatogram (between 2000 and 10000 variables), a clear difference, in terms of variable intensities, exists between the two groups, thus confirming the discriminating power of the volatile profile with respect to the geographical origin (Melucci et al., 2016; Lukić, Carlin, Horvat, & Vrhovsek, 2019).

Concerning the PLS-DA results, the values of the estimated Y variable (geographical category) by the model in cross and external validations are shown in Figure 2. The dotted line identifies the threshold value used to define the attribution of samples to different classes. Regarding the location of each sample, a greater distance from the threshold line can be interpreted as a better classification capacity of the model.

The results, in terms of percentage and number of samples correctly classified, are reported in Table 2. The percentage ranged from 80.8% to 91.2%. The values obtained for the “EU” category were higher, likely because of the greater number and variability of samples used to build the model. The external validation percentages were lower compared to those obtained for the cross-validation as expected, but the results can be considered more robust since they were obtained considering the 25% of samples that were not used to build the model.

The VIP (Variable Importance in Projection) score obtained by the PLS-DA confirmed that the section of the chromatogram ranging from 2000 to 10000 variables has a major contribution to sample discrimination (VIP values greater than 1) according to geographical origin (Figure 1-b).

Focusing on those incorrectly classified samples, a specific trend as a function of characteristics that could usually affect the volatile profile of the oil (such as the commercial category, olive cultivar, or country of origin) was not seen.

Results related to the probabilistic approach are shown in Figure 3. The graph refers to the category "EU": this means that higher a sample is located, the higher the probability for which it is classified as member of the "EU" category. As a consequence, oils classified as members of the other category (Extra-EU) are located in the bottom area of the graph. In this case, the threshold value is fixed at 0.5, corresponding to a probability of 50%: a sample classified with a probability lower than this is considered as not correctly grouped. It is also interesting to note that most of samples were correctly classified with a probability between 90% and 100%.

Regarding ANN, an early stopping technique was used to select the number of training cycles (epochs) to avoid over-fitting, using the test set to monitor the prediction error. An example of this procedure is reported in Figure 4, where the best ANN training was characterized by 18 epochs. Above this point, the error increased further indicating that the ANN tends to overfit. Consequently, the results of ANN are related to these iterations.

Training was repeated 5 times and the network's predictions were averaged, since with ANNs convergence is influenced by the initial weight value and the randomized split of data in training, validation, and test sets. The best prediction results were obtained with a three layers network, having 5 nodes; a larger number of nodes did not increase the network performance.

The classification results, in terms of percentage of samples correctly classified, are summarized in Table 3. Means and standard deviations (in brackets) were taken into account.

As reported for the PLS-DA model, even in this case higher percentages (from 93.2% to 98.7%) were achieved for the "EU" category in all the three data sets.

Comparing the results of the external validation (PLS-DA) and testing (ANN), it is possible to note that higher percentages were obtained in the second case for both the "EU" and "Extra-EU" categories. In particular, an increment of 4.7% and 8% of samples correctly classified was obtained. This is probably due to the fact that the ANN model is based on a non-linear approach.

In general, the percentages obtained were slightly lower than those reported by other studies based on volatile compounds and chemometric untargeted data elaboration (Gerhardt, Birkenmeier, Sanders, Rohn, & Weller, 2017; Bajoub et al., 2018; Lukić, Carlin, Horvat, & Vrhovsek, 2019). This aspect can be explained by the great variability, in terms of geographical origin, olive variety, commercial category, of the samples analyzed, which represents a strong point of this work.

The results described herein confirm the suitability of flash gas chromatography for checking geographical traceability of EVOOs and VOOs, even using untargeted chromatographic signals of the volatile fraction as variables for multivariate analysis (Melucci et al., 2016). An in-house validation of this analytical method, carried out to verify that a repeatable and reproducible signal, with sufficient sensitivity to collect the valuable information from the samples, has been carried out which underlined the good performance of the technique; this will be discussed in more detail in a subsequent publication.

4. Conclusions

In this work, the application of flash gas chromatography for volatile compounds analysis combined with untargeted chemometric data elaborations (PLS-DA and ANN) to discriminate EVOOs and VOOs with different geographical origin was presented.

For both elaborations, satisfactory results, in terms of percentages of samples correctly classified, were obtained: PLS-DA (external validation) allowed classification of around 89% and 81% of "EU" and "Extra-EU" samples, respectively; for ANN (testing set) the percentages were equal 93.2% and 88.8%, respectively.

It is important to highlight that these promising results were achieved by analyzing a set of samples that are representative of the large variety of parameters (olive cultivar, country of origin, commercial category) that can describe olive oil product and affect its chemical characteristics. The results obtained herein sustained the use of multivariate chemometrics with untargeted detection of volatile compounds as a powerful tool to discriminate EVOOs and VOOs of different origin. Other studies have already reported that the analysis of volatile compounds is suitable for tracing the geographical origin of VOOs. However, the methodology proposed herein presents some advantages in comparison with other techniques generally applied for this analysis, as it is very rapid (only 200 sec are needed for each chromatographic run) and easy to use since no sample treatment is required.

Funding

This work was supported by the Horizon 2020 European Research project OLEUM *"Advanced solutions for assuring the authenticity and quality of olive oil at a global scale",* which has received funding from the European Commission within the Horizon 2020 Programme (2014–2020), grant agreement no. 635690. The information expressed in this article reflects the authors' views; the European Commission is not liable for the information contained herein.

Acknowledgments

The authors would like to thank the colleagues involved in the OLEUM project, from EUROFINS, IPTPO, ITERG, UZZK, and ZRS for collection and shipment of the samples.

Declaration of interest

None.

References

Angerosa, F., Bréas, O., Contento, S., Guillou, C., Reniero, F., & Sada, E. (1999). Application of stable isotope ratio analysis to the characterization of the geographical origin of olive oils. *Journal of Agricultural and Food Chemistry, 47,* 1013-1017.

Bajoub, A., Medina-Rodríguez, S., Ajal, E. A., Cuadros-Rodríguez, L., Monasterio, R. P., Vercammen, J., Fernández-Gutiérrez, A., & Carrasco-Pancorbo, A. (2018). A metabolic fingerprinting approach based on selected ion flow tube mass spectrometry (SIFT-MS) and chemometrics: A reliable tool for Mediterranean origin-labeled olive oils authentication. *Food Research International, 106,* 233-242.

Bajoub, A., Medina-Rodríguez, S., Hurtado-Fernández, E., Ajal E. A., Ouazzani, N., Fernández-Gutiérrez, & Carrasco-Pancorbo A. (2016). A first approach towards the development of geographical origin tracing models for North Moroccan olive oils based on triacylglycerols profiles. *European Journal of Lipid Science and Technology, 118,* 1223-1235.

Ben Mohamed, M., Rocchetti, G., Montesano, G., Ben Ali, S., Guasmi, F., Grati-Kamoun, G., & Lucini, L. (2018). Discrimination of Tunisian and Italian extra-virgin olive oils according to their phenolic and sterolic fingerprints. *Food Research International, 106,* 920-927.

Bishop, C. M. (1995). Neural networks for pattern recognition. Oxford, UK: Oxford University Press.

Bontempo, L., Paolini, M., Franceschi, P., Ziller, L., García-González, D. L., & Camin, F. (2019). Characterisation and attempted differentiation of European and extra-European olive oils using stable isotope ratio analysis. *Food Chemistry, 276*, 782-789.

Brereton, R. G., & Lloyd, G. R. (2014). Partial least squares discriminant analysis: taking the magic away. *Journal of Chemometrics, 28,* 213-225.

Chiocchini, F., Portarena, S., Ciolfi, M., Brugnoli, E., & Lauteri, M. (2016). Isoscapes of carbon and oxygen stable isotope compositions in tracing authenticity and geographical origin of Italian extra-virgin olive oils. *Food Chemistry, 202,* 291-301.

Commission implementing regulation (EU) No 29/2012 of 13 January 2012 on marketing standards for olive oil. *Official Journal of European Journal, L 12,* 14-21.

Conte, L., Bendini, A., Valli, E., Lucci, P., Moret, S., Maquet, A., Lacoste, F., Brereton, P., García-González, D. L., Moreda, W., & Gallina Toschi, T. (2019). 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, in press. doi.org/10.1016/j.tifs.2019.02.025

Daszykowski, M., Walczak, B., Massart, D. L. (2002). Representative subset selection. *Analytica Chimica Acta, 468,* 91-103.

Dekhili, S., Sirieix, L., & Cohen, E. (2011). How consumers choose olive oil: the importance of origin cues. *Food Quality and Preference, 22,* 757-762.

Del Giudice, T., Cavallo, C., Caracciolo, F., & Cicia, G. (2015). What attributes of extra virgin olive oil are really important for consumers: a meta-analysis of consumers' stated preferences. *Agricultural and Food Economics, 3,* article number 20.

García-González D. L., Luna, G., Morales, M. T., & Aparicio, R. (2009). Stepwise geographical traceability of virgin olive oils by chemical profiles using artificial neural network models. *European Journal of Lipid Science and Technology, 111*, 1003-1013.

Gerhardt, N., Birkenmeier, M., Sanders, D., Rohn, S., & Weller P. (2017). Resolution-optimized headspace gas chromatography-ion mobility spectrometry (HS-GC-IMS) for non-targeted olive oil profiling. *Analytical and Bioanalytical Chemistry, 409,* 3933-3942.

Ghisoni, S., Lucini, L., Angilletta, F., Rocchetti, G., Farinelli, D., Tombesi, S., & Trevisan, M. (2019). Discrimination of extra-virgin-olive oils from different cultivars and geographical origins by untargeted metabolomics. *Food Research International, 121,* 746-753.

Giacalone, R., Giuliano, S., Gulotta, E., Monfreda, M., & Presti, G. (2015). Origin assessment of EV olive oils by esterified sterols analysis. *Food Chemistry, 188,* 279-285.

Kosma, I., Vatavali, K., Kontakos, S., Kontominas, M., Kiritsakis, A., & Badeka, A. (2017). Geographical differentiation of Greek extra virgin olive oil from late- harvested Koroneiki cultivar fruits. *Journal of the American Oil Chemists' Society, 94*, 1373-1384.

Lukić, I., Carlin, S., Horvat, I., & Vrhovsek, U. (2019). Combined targeted and untargeted profiling of volatile aroma compounds with comprehensive two-dimensional gas chromatography for differentiation of virgin olive oils according to variety and geographical origin. *Food Chemistry, 270,* 403-414.

Mallamace, D., Vasi, S., Corsaro, C., Naccari, C., Clodoveo, M. L., Dugo, G., & Cicero, N. (2017). Calorimetric analysis points out the physical-chemistry of organic olive oils and reveals the geographical origin. *Physica A, 486,* 925-932.

Melucci, D., Bendini, A., Tesini, F., Barbieri, S., Zappi, A., Vichi, S., Conte, L., & Gallina Toschi, T. (2016). Rapid direct analysis to discriminate geographic origin of extra virgin olive oils by flash gas chromatography electronic nose and chemometrics. *Food Chemistry, 204,* 263-273.

Palagano, R., Bendini, A., Valli, E., Cevoli, C., Gallina Toschi, T. (2019). OLEUM Project. Analysis of volatile compounds of extra virgin and virgin olive oil with a different geographical origin by flash gas chromatography. http://doi.org/10.6092/unibo/amsacta/6243 [Dataset]

Peršurić, Ž., Saftić, L., Mašek, T., & Kraljević Pavelić, S. (2018). Comparison of triacylglycerol analysis by MALDI-TOF/MS, fatty acid analysis by GC-MS and non-selective analysis by NIRS in combination with chemometrics for determination of extra virgin olive oil geographical origin. A case study. *LWT - Food Science and Technology, 95,* 326-332.

Quintanilla-Casas, B., Bertin, S., Leik, K., Bustamante, J., Guardiola, F., Valli, E., Bendini, A., Gallina Toschi, T., Tres, A., & Vichi, S. (2020). Profiling versus fingerprinting analysis of sesquiterpene hydrocarbons for the geographical authentication of extra virgin olive oils. *Food Chemistry, 307,* Article number 125556.

Sayago, A., González-Domínguez, R., Beltrán, R., & Fernández-Recamales, A. (2018). Combination of complementary data mining methods for geographical characterization of extra virgin olive oils based on mineral composition. *Food Chemistry, 261,* 42-50.

Tomasi, G., Van Den Berg, F., & Andersson, C. (2011). Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. *Journal of Chemometrics, 18,* 231-241.

Valli, E., Bendini, A., Berardinelli, A., Ragni, L., Riccò, B., Grossi, M., & Gallina Toschi, T. (2016). Rapid and innovative instrumental approaches for quality and authenticity of olive oils. *European Journal of Lipid Science and Technology, 118,* 1601-1619.

Vichi, S., Pizzale, L., & Conte, L. (2007). Stereospecific distribution of fatty acids in triacylglycerols of olive oils. *European Journal of Lipid Science and Technology, 109,* 72-78.

Origin class	N	Country of origin
EU	116	29 Spain, 25 Italy, 22 Croatia, 16 Greece, 12 Portugal, 12 Slovenia
Extra-EU	70	42 Morocco, 21 Turkey, 6 Tunisia, 1 Chile
Blends	24	12 EU blends, 12 EU/Extra-EU blends

Table 1. Number of samples for each origin class considered and geographical origin. EU: oils from EU state members; Extra-EU: oils from countries outside the European Union; Blends: oils obtained mixing EU oils or EU and Extra-EU oils.

Category	Cross validation	External validation
EU	91.2% (93/102)	88.5% (23/26)
Extra-EU	91.1% (51/56)	80.8% (21/26)

Table 2. Percentages and number (in parentheses) of correctly classified samples for each category using the PLS-DA model. EU: oils from a single state member of European Union and oils obtained by mixing EU oils; Extra-EU: oils from a single country outside the European Union and oils obtained by mixing EU and Extra-EU oils.

Category	Training (%)	Validation (%)	Testing (%)
EU	98.7 (1.1)	95.4 (3.9)	93.2 (3.2)
Extra-EU	94.4 (7.0)	88.7 (7.8)	88.8 (5.4)

Table 3. Percentages (mean of 5 training of the model and standard deviation in parentheses) of samples correctly classified for each category using the ANN model. EU: oils from a single state member of European Union and oils obtained by mixing EU oils; Extra-EU: oils from a single country outside the European Union and oils obtained by mixing EU and Extra-EU oils.

Figure 1. a) Mean chromatogram obtained averaging the intensity of each variable for all "EU" (blue line) or "Extra-EU" (violet line) samples; b) VIP score for each variable (signal digitalized every 0.01 sec).

Figure 2. Graph of the values of the Y variable estimated by the PLS-DA model, in cross and external validations (grey area). Blue squares: EU; violet triangle: Extra-EU.

Figure 3. Probability values of belonging to the "EU" category. 1 = probability of 100%; 0 = probability of 0%. Gray area: test set used for external validation of the PLS-DA model. Blue squares: EU; violet triangles: Extra-EU.

Figure 4. Error graph of validation (green line) and test (red line) set used for the ANN model.

Table S1. Information about country of origin, olive variety and commercial category for all the samples collected and analyzed.

Code	Origin class	Country of origin	Olive variety	Commercial category
1	Blend EU	GRC - ITA	Manaki - Coratina	EVOO [b]
2	Blend EU	GRC - ITA	Manaki - Moraiolo	EVOO [b]
3	Blend EU	GRC - ITA	Manaki - NA	EVOO [b]
4	Blend EU	GRC - ITA - ESP	Manaki - Arbequina - NA	EVOO [b]
5	Blend EU	GRC - ITA - ESP	Manaki - Arbequina - Coratina	EVOO [b]
6	Blend EU	GRC - ITA - ESP	Manaki - Arbequina - Moraiolo	EVOO [b]
7	Blend EU	GRC - ESP	Manaki - Arbequina	EVOO [b]
8	Blend EU	GRC - ESP	Koroneiki - Arbequina - Picual - Cornicabra	EVOO [a]
9	Blend EU	ESP - ITA	Arbequina - NA	EVOO [b]
10	Blend EU	ESP - ITA	Arbequina - Coratina	EVOO [b]
11	Blend EU	ESP - ITA	Arbequina - Moraiolo	EVOO [b]
12	Blend EU	ESP - PRT	Arbequina - Arbosana	EVOO [b]
13	Blend Extra-EU	GRC - TUN	Manaki - Sahli	VOO [b]
14	Blend Extra-EU	ITA - TUN	NA - Sahli	VOO [b]
15	Blend Extra-EU	ITA - TUN	Sahli - Coratina	VOO [b]
16	Blend Extra-EU	ITA - TUN	Sahli - Moraiolo	VOO [b]
17	Blend Extra-EU	ESP - MAR	Picual - Moroccan Picholine	EVOO [b]
18	Blend Extra-EU	ESP - MAR	Picual - Languedoc Picholine	EVOO [b]
19	Blend Extra-EU	ESP - MAR	Picual - Moroccan Picholine - Koroneiki	EVOO [b]
20	Blend Extra-EU	ESP - TUN	Arbequina - Sahli	VOO [b]
21	Blend Extra-EU	ESP - TUN	Picual - Chemlali - Chetoui	EVOO [b]
22	Blend Extra-EU	ESP - TUN	Picual - Chemlali	EVOO [b]
23	Blend Extra-EU	ESP - TUN	Picual - Chemlali	EVOO [b]
24	Blend Extra-EU	ESP - TUN	Picual - Chetoui	EVOO [b]
25	EU	HRV	Picholine	EVOO [a]
26	EU	HRV	Leccio del Corno	EVOO [a]
27	EU	HRV	Istarska Bjelica	EVOO [a]
28	EU	HRV	Rosinjola	EVOO [a]
29	EU	HRV	Leccino - Pendolino	EVOO [a]
30	EU	HRV	Leccino - Pendolino	EVOO [a]
31	EU	HRV	Picholine - Leccio del Corno	EVOO [a]
32	EU	HRV	Istarska Bjelica	EVOO [a]
33	EU	HRV	Oblica	EVOO [a]
34	EU	HRV	Istarska Bjelica - Leccino - Buža	EVOO [a]
35	EU	HRV	Istarska Bjelica - Leccino - Buža	EVOO [a]
36	EU	HRV	Leccino - Pendolino	EVOO [a]

37	EU	HRV	Picholine - Leccio del Corno	EVOO [a]
38	EU	HRV	Ascolana Tenera - Itrana - Frantoio	EVOO [a]
39	EU	HRV	Buža Puntoža - Rosinjola - Bova	EVOO [a]
40	EU	HRV	Istarska Bjelica	EVOO [a]
41	EU	HRV	Ascolana Tenera - Itrana - Frantoio	EVOO [a]
42	EU	HRV	Buža Puntoža	EVOO [a]
43	EU	HRV	Buža Puntoža	EVOO [a]
44	EU	HRV	Picholine	EVOO [a]
45	EU	HRV	Plominka - Simjaca	EVOO [a]
46	EU	HRV	Oblica	EVOO [a]
47	EU	GRC	Koroneiki	VOO [a]
48	EU	GRC	Koroneiki	EVOO [a]
49	EU	GRC	Koroneiki	EVOO [b]
50	EU	GRC	Manaki	EVOO [b]
51	EU	GRC	Koroneiki	EVOO [b]
52	EU	GRC	Koroneiki	EVOO [b]
53	EU	GRC	Manaki	EVOO [a]
54	EU	GRC	Koroneiki	EVOO [a]
55	EU	GRC	Koroneiki	EVOO [a]
56	EU	GRC	NA	VOO [a]
57	EU	GRC	Koroneiki	EVOO [a]
58	EU	GRC	Koroneiki	EVOO [a]
59	EU	GRC	Koroneiki	EVOO [b]
60	EU	GRC	Koroneiki	EVOO [a]
61	EU	GRC	Koroneiki	EVOO [a]
62	EU	GRC	NA	EVOO [a]
63	EU	ITA	Coratina	EVOO [a]
64	EU	ITA	Coratina	VOO [a]
65	EU	ITA	Frantoio	EVOO [a]
66	EU	ITA	Castiglionese	VOO [a]
67	EU	ITA	Leccino - Frantoio - Pendolino	EVOO [a]
68	EU	ITA	Leccino - Frantoio - Pendolino	VOO [a]
69	EU	ITA	Arbequina	EVOO [a]
70	EU	ITA	Coratina - Ogliarola	EVOO [a]
71	EU	ITA	Nocellara del Belice	EVOO [a]
72	EU	ITA	Biancolilla	EVOO [a]
73	EU	ITA	Nocellara del Belice	EVOO [a]
74	EU	ITA	Leccino - Frantoio - Moraiolo	EVOO [a]
75	EU	ITA	Coratina	EVOO [a]
76	EU	ITA	Coratina	VOO [a]
77	EU	ITA	Nostrana di Brisighella	EVOO [a]
78	EU	ITA	Leccino - Frantoio - Moraiolo	EVOO [a]

79	EU	ITA	Nostrana di Brisighella	EVOO [a]
80	EU	ITA	NA	EVOO [a]
81	EU	ITA	Coratina	EVOO [a]
82	EU	ITA	Moraiolo	EVOO [a]
83	EU	ITA	Carolea	EVOO [a]
84	EU	ITA	Dritta - Leccino	EVOO [a]
85	EU	ITA	Frantoio	EVOO [a]
86	EU	ITA	Peranzana	EVOO [a]
87	EU	ITA	Peranzana	EVOO [a]
88	EU	PRT	Arbequina - Koroneiki	EVOO [b]
89	EU	PRT	Arbosana	EVOO [b]
90	EU	PRT	Arbosana	EVOO [b]
91	EU	PRT	Arbosana	EVOO [b]
92	EU	PRT	Koroneiki	EVOO [b]
93	EU	PRT	Arbequina	EVOO [b]
94	EU	PRT	Arbequina	EVOO [b]
95	EU	PRT	Arbequina	EVOO [b]
96	EU	PRT	Sikitita	EVOO [b]
97	EU	PRT	Arbosana	EVOO [b]
98	EU	PRT	NA	EVOO [b]
99	EU	PRT	NA	EVOO [b]
100	EU	SVN	Istarska Bjelica - Leccino - Others	EVOO [a]
101	EU	SVN	Istarska Bjelica - Leccino - Others	EVOO [a]
102	EU	SVN	Istarska Bjelica - Leccino - Others	EVOO [a]
103	EU	SVN	Istarska Bjelica - Leccino - Others	EVOO [a]
104	EU	SVN	Istarska Bjelica - Leccino - Others	EVOO [a]
105	EU	SVN	Istarska Bjelica - Leccino - Maurino	EVOO [a]
106	EU	SVN	Istarska Bjelica - Leccino - Maurino	EVOO [a]
107	EU	SVN	Istarska Bjelica	EVOO [a]
108	EU	SVN	Istarska Bjelica - Leccino - Others	EVOO [a]
109	EU	SVN	Istarska Bjelica	EVOO [a]
110	EU	SVN	Istarska Bjelica - Leccino - Others	EVOO [a]
111	EU	SVN	Istarska Bjelica - Leccino - Others	EVOO [a]
112	EU	ESP	NA	EVOO [a]
113	EU	ESP	NA	VOO [a]
114	EU	ESP	NA	VOO [a]
115	EU	ESP	NA	VOO [a]
116	EU	ESP	NA	EVOO [a]
117	EU	ESP	NA	VOO [a]
118	EU	ESP	NA	EVOO [a]
119	EU	ESP	NA	VOO [a]
120	EU	ESP	NA	VOO [a]

121	EU	ESP	NA	EVOO [a]
122	EU	ESP	Hojiblanca	EVOO [a]
123	EU	ESP	Arbequina	EVOO [a]
124	EU	ESP	Picual	EVOO [a]
125	EU	ESP	Arbequina - Hojiblanca	EVOO [a]
126	EU	ESP	Arbequina - Hojiblanca	VOO [a]
127	EU	ESP	Manzanilla	EVOO [a]
128	EU	ESP	Manzanilla	EVOO [a]
129	EU	ESP	Arbequina	EVOO [a]
130	EU	ESP	Hojiblanca	EVOO [a]
131	EU	ESP	Koroneiki	EVOO [b]
132	EU	ESP	Hojiblanca	EVOO [b]
133	EU	ESP	Manzanilla - Hojiblanca - Picual	EVOO [a]
134	EU	ESP	NA	EVOO [a]
135	EU	ESP	NA	VOO [a]
136	EU	ESP	Hojiblanca	EVOO [a]
137	EU	ESP	Hojiblanca	EVOO [a]
138	EU	ESP	Picual	VOO [a]
139	EU	ESP	NA	EVOO [a]
140	EU	ESP	Arbequina	EVOO [a]
141	Extra-EU	CHL	NA	EVOO [a]
142	Extra-EU	MAR	Arbequina	EVOO [a]
143	Extra-EU	MAR	Arbequina	EVOO [b]
144	Extra-EU	MAR	Koroneiki	EVOO [b]
145	Extra-EU	MAR	Arbosana	EVOO [b]
146	Extra-EU	MAR	Arbequina	EVOO [b]
147	Extra-EU	MAR	Arbequina	EVOO [b]
148	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
149	Extra-EU	MAR	Arbosana	EVOO [b]
150	Extra-EU	MAR	Koroneiki	EVOO [b]
151	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
152	Extra-EU	MAR	Arbequina	EVOO [b]
153	Extra-EU	MAR	Arbosana	EVOO [b]
154	Extra-EU	MAR	Koroneiki	EVOO [b]
155	Extra-EU	MAR	Moroccan Picholine - Hojiblanca	EVOO [b]
156	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
157	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
158	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
159	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
160	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
161	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
162	Extra-EU	MAR	Moroccan Picholine	EVOO [b]

163	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
164	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
165	Extra-EU	MAR	Arbequina	EVOO [b]
166	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
167	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
168	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
169	Extra-EU	MAR	Picholine - Arbosana	EVOO [b]
170	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
171	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
172	Extra-EU	MAR	Arbequina - Koroneiki	EVOO [b]
173	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
174	Extra-EU	MAR	Moroccan Picholine - Koroneiki	EVOO [b]
175	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
176	Extra-EU	MAR	Arbequina	EVOO [b]
177	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
178	Extra-EU	MAR	Arbequina - Koroneiki	VOO [b]
179	Extra-EU	MAR	Arbequina	EVOO [b]
180	Extra-EU	MAR	Moroccan Picholine - Arbequina	EVOO [b]
181	Extra-EU	MAR	Moroccan Picholine	VOO [b]
182	Extra-EU	MAR	Moroccan Picholine	EVOO [b]
183	Extra-EU	MAR	Arbequina	VOO [b]
184	Extra-EU	TUN	Chetoui - Chemlali	VOO [a]
185	Extra-EU	TUN	Sahli	VOO [a]
186	Extra-EU	TUN	Sahli - Chemlali	EVOO [b]
187	Extra-EU	TUN	Chemlali	EVOO [b]
188	Extra-EU	TUN	Chemlali	EVOO [b]
189	Extra-EU	TUN	Chetoui	EVOO [b]
190	Extra-EU	TUR	Ayvalik - Domat	EVOO [a]
191	Extra-EU	TUR	Memecik - Gemlik	EVOO [a]
192	Extra-EU	TUR	Memecik	EVOO [a]
193	Extra-EU	TUR	Ayvalik	VOO [a]
194	Extra-EU	TUR	Ayvalik	VOO [a]
195	Extra-EU	TUR	Ayvalik	EVOO [a]
196	Extra-EU	TUR	Ayvalik	VOO [a]
197	Extra-EU	TUR	Domat	EVOO [a]
198	Extra-EU	TUR	Memecik	EVOO [a]
199	Extra-EU	TUR	Karamani - Hasebi	VOO [a]
200	Extra-EU	TUR	Memecik	EVOO [a]
201	Extra-EU	TUR	Gemlik - Memecik	VOO [a]
202	Extra-EU	TUR	Memecik	EVOO [a]
203	Extra-EU	TUR	Gemlik	EVOO [a]
204	Extra-EU	TUR	Memecik	VOO [a]

205	Extra-EU	TUR	Memecik	EVOO [a]
206	Extra-EU	TUR	Saurani - Halhali - Karamani	VOO [a]
207	Extra-EU	TUR	Edremit - Domat - Gemlik	VOO [a]
208	Extra-EU	TUR	Memecik	VOO [a]
209	Extra-EU	TUR	Ayvalik - Edremit	VOO [a]
210	Extra-EU	TUR	NA	EVOO [a]

NA: information not available; EVOO: extra virgin olive oil; VOO: virgin olive oil.

EU: oils from state members of European Union; Extra-EU: oils from countries outside European Union; Blends: oils obtained mixing EU oils or EU and Extra-EU oils.

CHL: Chile; ESP: Spain; GRC: Greece; HRV: Croatia; ITA: Italy; MAR: Morocco; PRT: Portugal; SVN: Slovenia; TUN: Tunisia; TUR: Turkey.

[a] Commercial category defined by Panel Test realized in the framework of the OLEUM project.

[b] Commercial category declared by the company that provided the sample.

Figure 1

Click here to download high resolution image

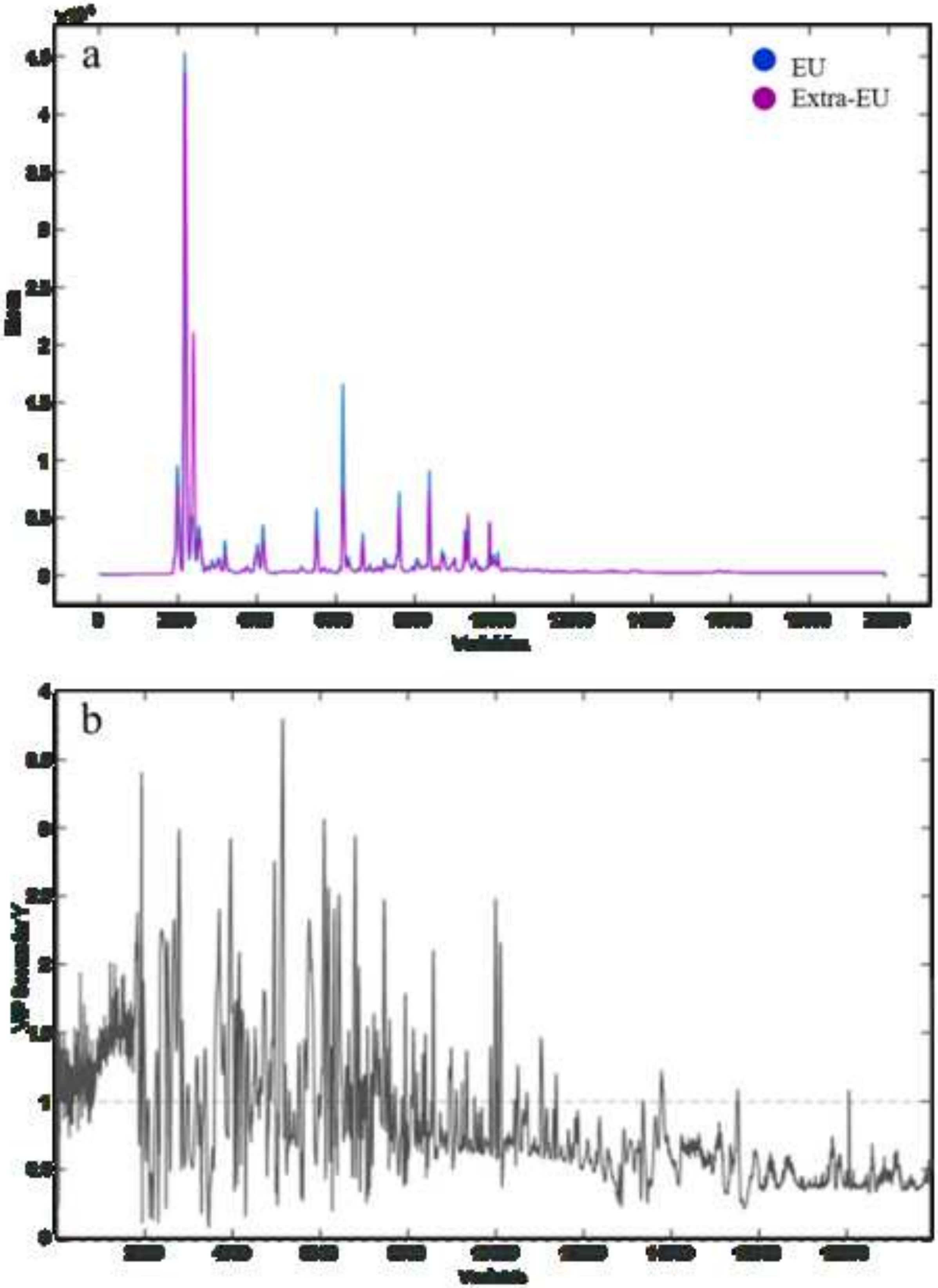


Figure 2

Click here to download high resolution image

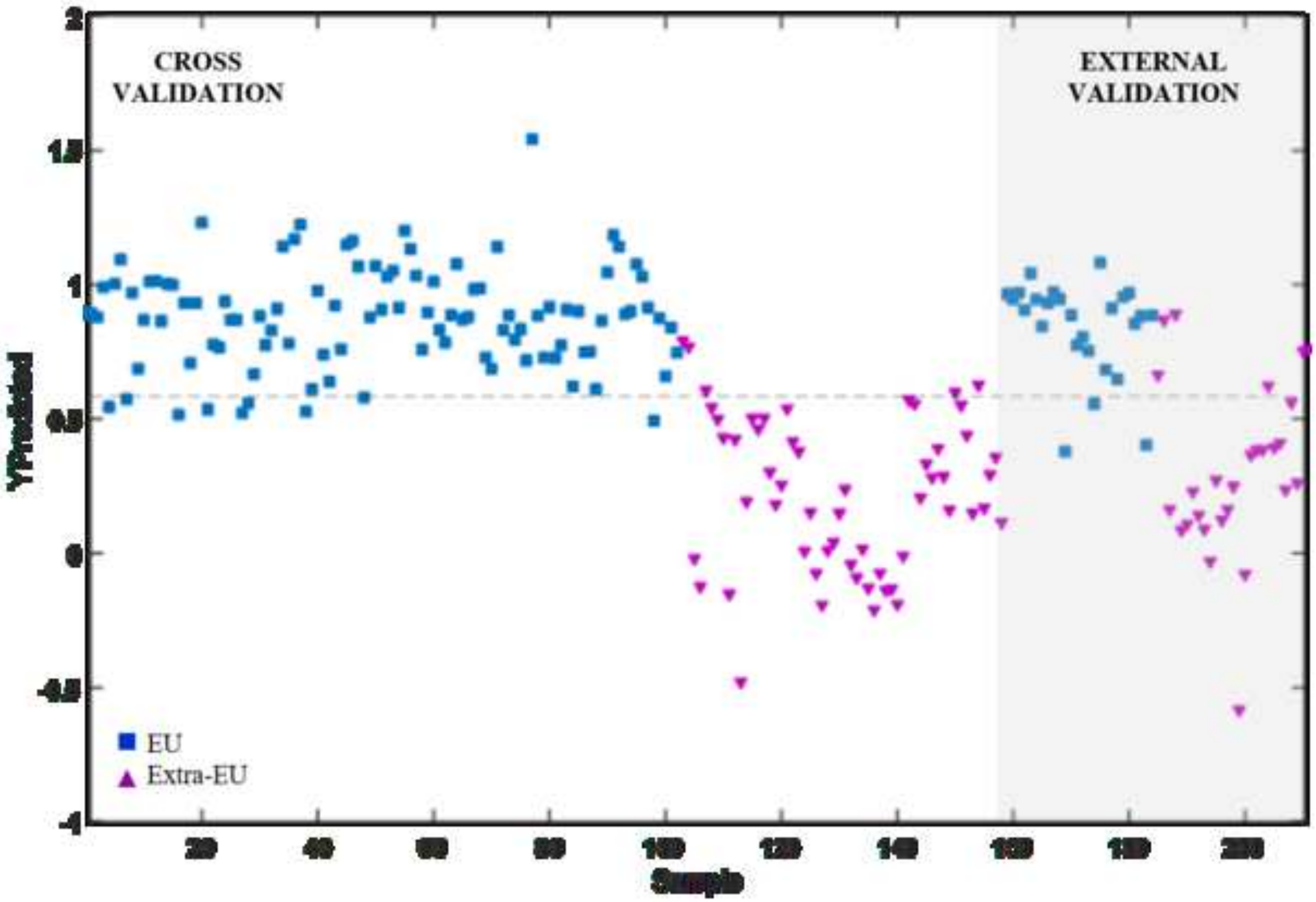


Figure 3

Click here to download high resolution image

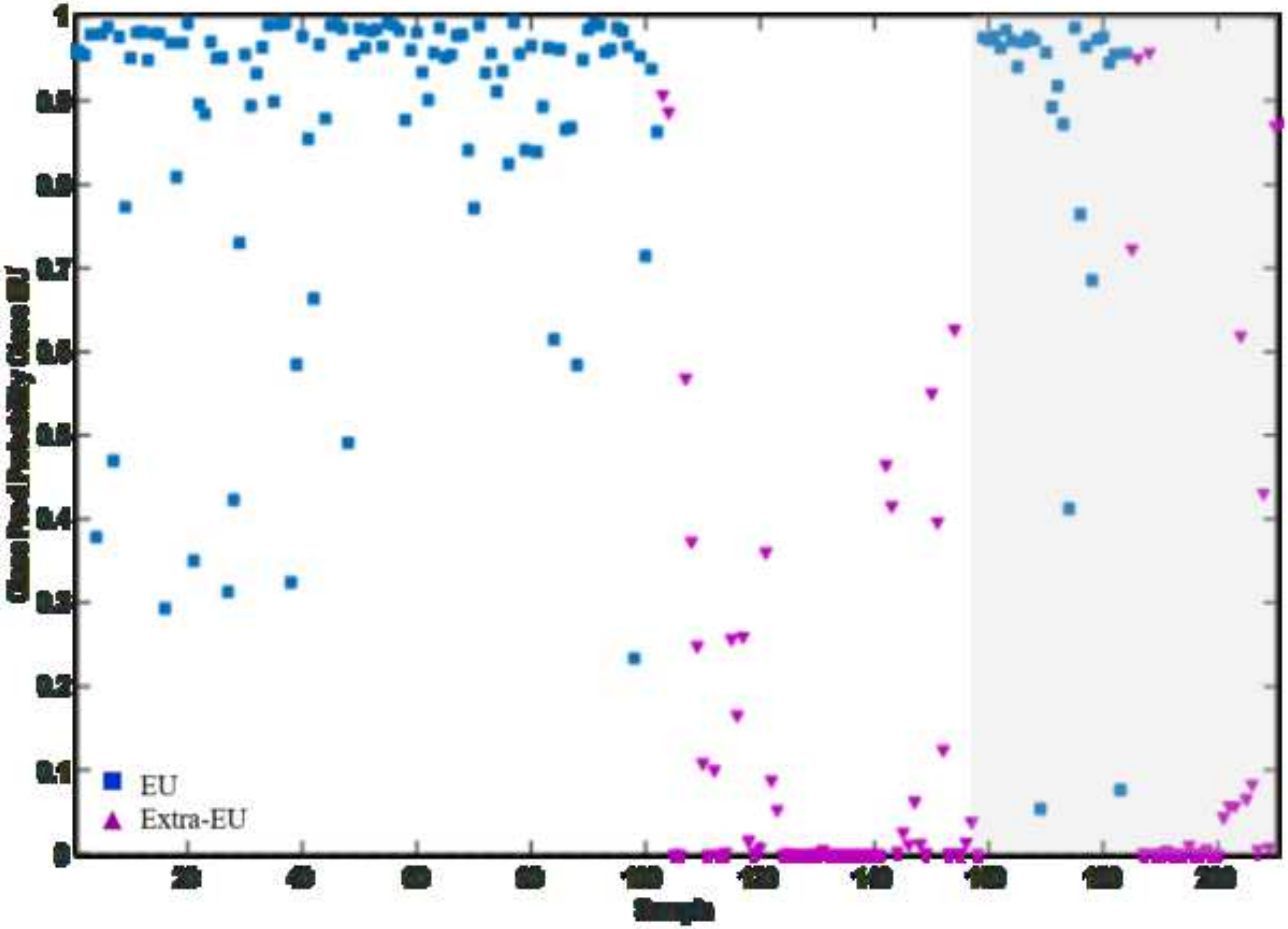
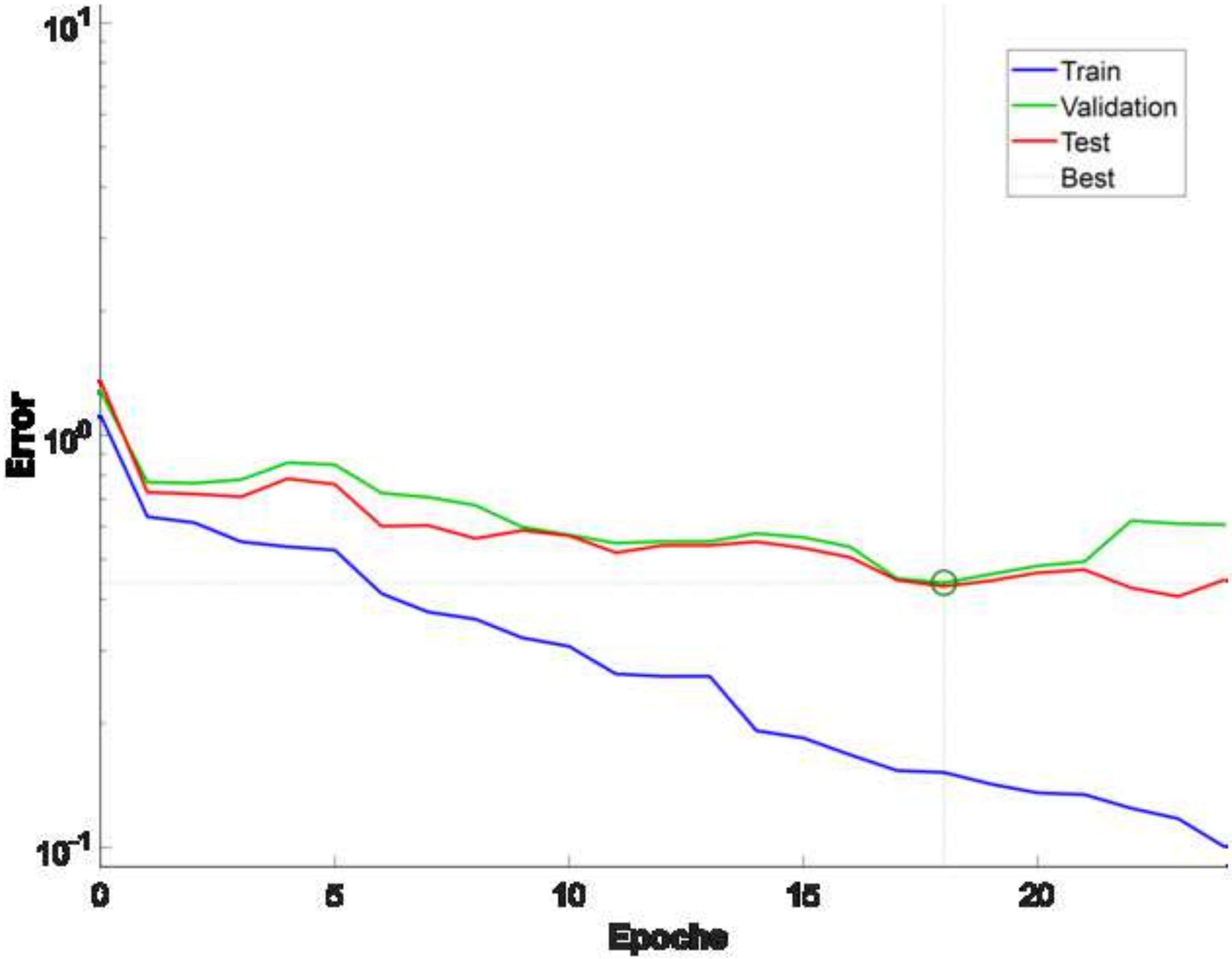


Figure 4

Click here to download high resolution image



Conflict of Interest and Authorship Conformation Form

Please check the following as appropriate:

- ○ All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version.

- ○ This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue.

- ○ The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript