



Chemometric strategies for authenticating extra virgin olive oils from two geographically adjacent Catalan protected designations of origin

Aina Fort, Itziar Ruisánchez^{*}, M. Pilar Callao

Chemometrics, Qualimetric and Nanosensors Grup, Department of Analytical and Organic Chemistry, Rovira i Virgili University, Marcel·lí Domingo s/n, 43007 Tarragona, Spain

ARTICLE INFO

Keywords:

Food authentication
Extra-virgin olive oil
FT-Raman
Fluorescence
PLS-DA
Data fusion

ABSTRACT

This paper proposes using multivariate methodology to authenticate the geographical origin of monovarietal Arbequina extra virgin olive oils from two geographically adjacent Catalan PDOs (protected designations of origin). Two spectroscopic techniques, fluorescence and FT-Raman were used to obtain multivariate data. The results obtained by principal component analysis (PCA), partial least squares–discriminant analysis (PLS–DA) and low and high-data fusion are discussed. The training and test samples were randomly chosen. To obtain greater representativeness in the results, the data set was independently divided three times. Thus, three different training and test sets were obtained and used to build independent classification models and carry out data fusion strategies. When the two fusion strategies were applied, the performance parameters (sensitivity and specificity) were better than when they were applied individually, which indicates that synergies between FT-Raman and fluorescence expand the information about the samples. It was also observed that the influence on the results of the type of data set obtained by the random division of the data set was minimized.

1. Introduction

The extra-virgin olive oil (EVOO) is a basic ingredient in the Mediterranean diet and is becoming increasingly more popular worldwide because of its nutritional and sensory qualities as well as for reducing the risk factor of some pathologies [1,2]. Therefore, considerable effort has been made to ensure both authentication and detection of adulteration of EVOOs, becoming an area of great interest. Authentication includes ensuring aspects such as its botanical variety and origin of the product [3]. To encourage the use of quality-differenced labelling to identify agricultural products and foodstuffs, a European regulation was published in 1992 [4] which has been updated several times [5].

Because of both legislative and consumer interest, considerable research has been carried out into developing analytical methods for assuring the quality of EVOOs [6]. Recently, Meenu et al. [7] have published a critical review on analytical techniques for detecting the adulteration of EVOOs. Regarding authentication, in general, it is very difficult to find a marker that solves the problem by itself, so the usual approach is multivariate. Since authentication requires a qualitative response, the appropriate techniques are multivariate classification techniques [8–10].

In the bibliography it can be found that EVOOs have been analyzed

using any instrumental technique that gives a multivariate response [11–15]. Spectroscopic techniques [16,17] are worthy of special mention because the spectrum is easy and quick to obtain, it does not require sample pretreatment (or little), and as a result is a cheaper and more environmentally friendly measurement. Raman spectroscopy is being applied for both adulteration [18–20] and authentication problems [21–23]. Fluorescence techniques are gaining attention as a method of food analysis. As Meenu et al. [7] point out, molecular fluorescence spectroscopy has been used to analyse the quality of EVOOs, although its applications to authentication and the detection of adulteration are limited. Some recent papers describe the application of fluorescence spectroscopy both, for identifying the adulteration of EVOOs [24–26] and for the authentication of cultivar and geographical origin [27]. There are several ways to work, such as fixing the excitation wavelength and recording the emission spectra [25,28], recording the synchronous emission spectra [24,26] and using front face mode [26,29].

If multivariate models are to be improved and complementary information obtained from various analytical techniques is available, the application of data fusion techniques would be a good strategy. In the literature, there is an increasing number of studies on the application of data fusion techniques to food analysis [30]. In the particular case of the

^{*} Corresponding author.

quality analysis of EVOOs, techniques such as liquid and gas chromatography, UV, IR and other spectroscopies, physico-chemical parameters and sensorial data have been fused [31,32]. Of these, IR spectroscopies are by far the most widely used [33]. To our knowledge, just one paper has used confocal Raman and fluorescence spectroscopy to study the quantitative adulteration of EVOO [34], but no one has applied it to study authentication by means of qualitative approach.

In this context, the aim of the present study is to establish a methodology to authenticate the geographical origin of monovarietal Arbequina extra-virgin olive oils from two geographically adjacent Catalan PDOs. The main challenges are, in the first place that the two types of EVOO's are from the same olive variety which implies that they would have similar instrumental response. In the second place, belonging to adjacent geographical origin allows to extrapolate that for no adjacent PDO the results would be better or equal.

Two methods, fluorescence and FT-Raman spectroscopies were used. Those techniques have been selected based in its own characteristics previously mentioned. In addition, they can present some synergies and the synergies study have been not deeply explored with qualitative multivariate purpose. As a result, we expect to obtain valuable information about their combination.

In relation to the classification technique, two main approaches can be followed: discriminant or modeling. Among them, the partial least squares–discriminant analysis (PLS-DA) has been selected, since it focuses on the search for a delimiter between the classes. Principal component analysis (PCA) has been applied to visualize sample distribution, possible presence of trends and outliers. To improve the classification results obtained from the individual models, low and high-data fusion are applied and discussed in detail.

2. Materials and methods

2.1. Samples

A total of 84 single-variety (Arbequina) EVOO samples from the 2019 harvest in two geographical origins in Catalonia (Spain) were analysed: 47 samples from Siurana (S) and 37 from Les Garrigues (LG). The Catalan Government's Official Tasting Panel of Virgin Olive Oils of Catalonia supplied the samples and confirms the status of the oils.

Approximately 65% of the samples were used for training and the rest were used for validation. Hence, the training set consisted of 54 samples (22 LG and 32 S) and the test set of 30 samples (15 LG and 15 S).

The training and test samples were randomly chosen. For greater representativeness in the results, the overall sample was randomly and independently divided three times, thus obtaining three independent random data sets which were used for the independent classification models and the data fusion strategies.

2.2. Instrumentation and software

Thermo Nicolet 5700 FT-IR (Wisconsin, USA) equipped with the FT-Raman module NXR, an InGaAs detector, a CaF₂ beam splitter and a 1064 nm Nd-YAG laser was used. The spectrum was recorded by filling a 48-well plate with an oil's drop. Measurement conditions were: room temperature, 1064 nm laser, 0.25 W power, number of scans 124, 4 cm⁻¹ resolution, and range measurement from 100 to 3700 cm⁻¹. No sample pretreatment was applied.

Fluorescence spectra were recorded using a Shimadzu RF-5301PC instrument (Shimadzu Corporation, Kyoto, Japan). The slit width was 5 nm for excitation and for emission. The fluorescence emission spectra were collected between 360 and 800 nm using an excitation wavelength at 350 nm. The integration time was 0.1 s, and the sampling interval was 10 nm. No sample pretreatment was applied.

For chemometric treatment, MATLAB software, version 8.0.0.783 – R2012b (Natick, MA, USA), and PLS Toolbox 7.0.2 (Eigenvector Research Inc., Wenatchee, WA, USA) were used.

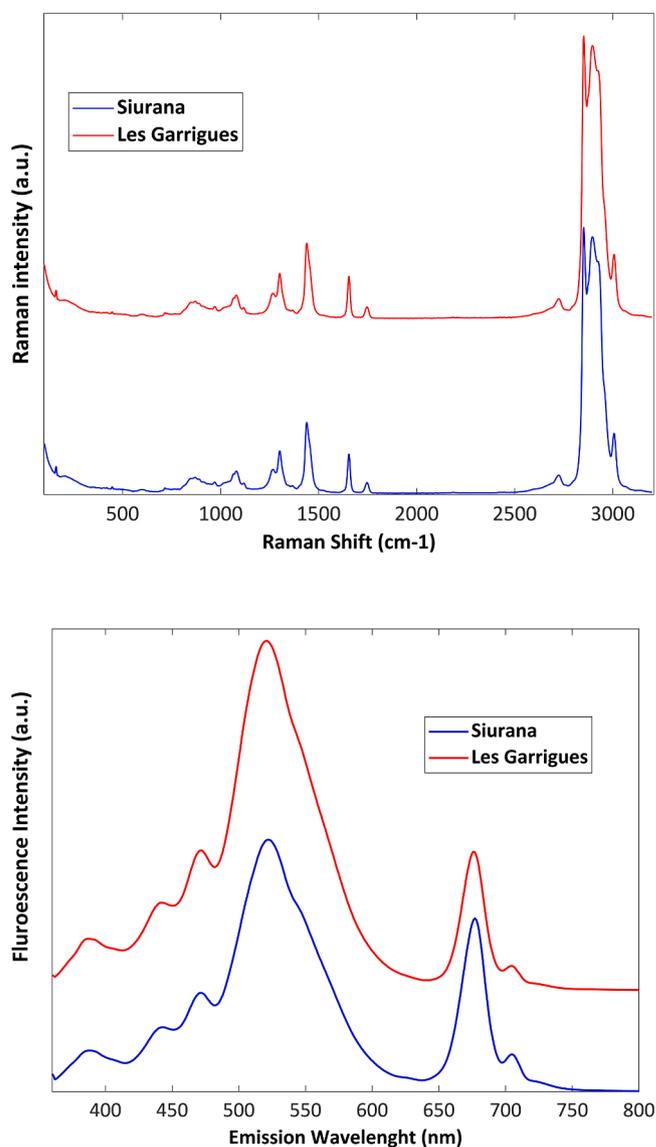


Fig. 1. Raw spectra of, a) mean FT-Raman spectrum of the two classes, b) mean Fluorescence spectrum of the two classes. In both, the spectrum of Les Garrigues (red) is at the top and Siurana (blue) at the bottom.

2.3. Data analysis

The VIP method [35–36] was applied to detect and select the most important variable in Raman data. They provide information about the significance of each variable on the latent variables (LV). The greater the VIP scores (usually greater than 1), the more important the corresponding variable is.

Principal component analysis (PCA) compresses the data, and reduce it from the high dimensional variable space to a lower dimensional principal component space. It was applied to explore sample distribution in the multivariate space, check any natural clustering and identify possible outliers.

PLS-DA [37] establishes a PLS linear regression between a matrix of independent variables (fluorescence or FT-Raman spectra) and a matrix of dependent variables (class code). For instance, the class code can be one if the sample belongs to the class and zero if it does not belong. Once the model was established, the results obtained for each sample will have values in an interval between zero and one. A threshold based on minimizing false assignments is set, which allows the samples to be assigned to one or another class, [31,38].

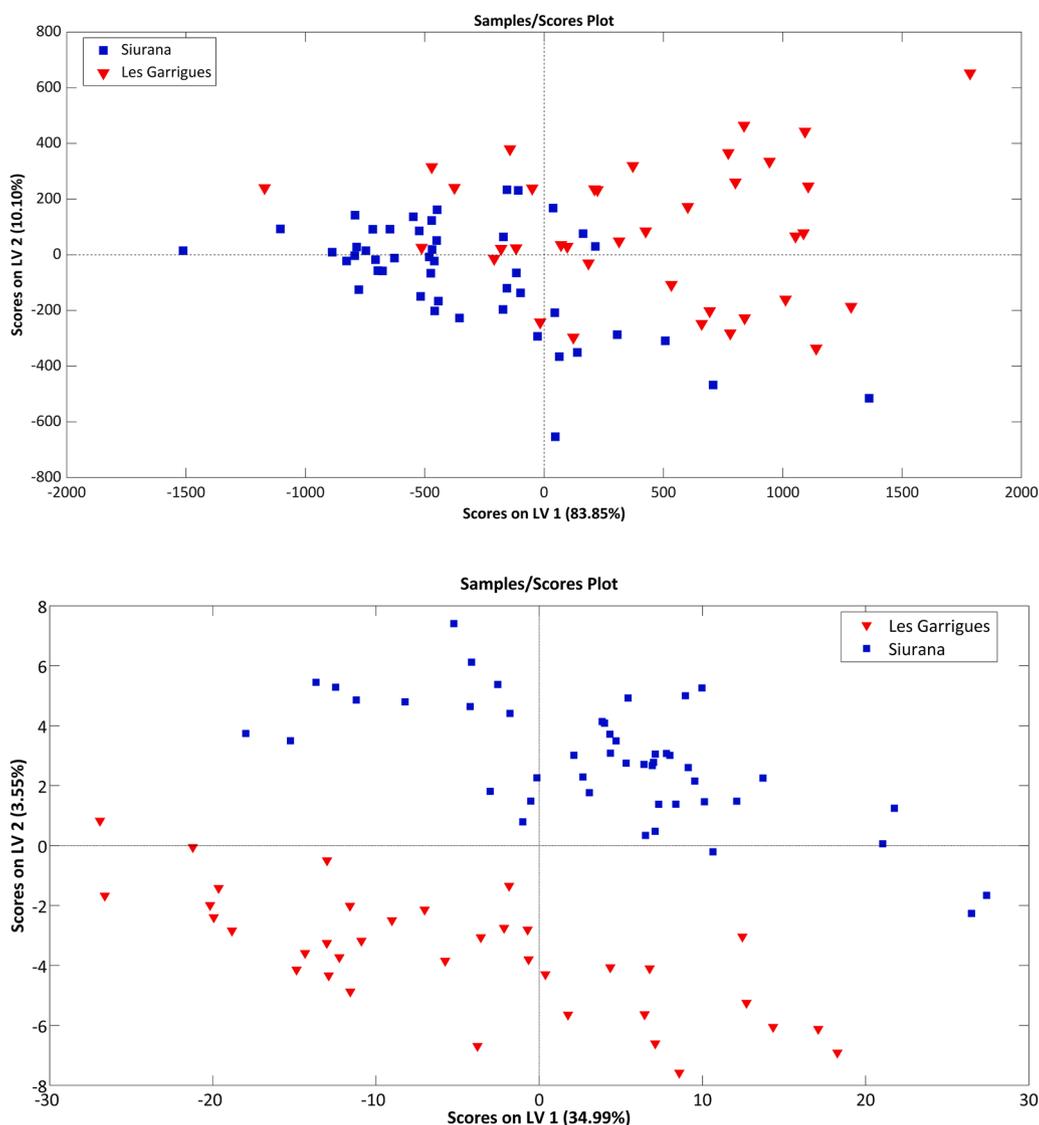


Fig. 2. PLS-DA score plot of the first two LVs for Les Garrigues PDO (red triangle) and Siurana PDO (blue square): a) of the fluorescence data and b) of the FT-Raman data.

High-level data fusion [39,40], uses the results obtained for each sample by the FT-Raman and fluorescence individual PLS-DA models. Several operators are applied to the individual results keeping just the highest value and assigning the sample to that class. The final sample assignment (ensemble decision) was fixed considering the majority of the decisions from each operator (majority vote). Of the operators available, the minimum, maximum, average and product operators were applied.

In the low-level fusion approach data from VIP-Raman and fluorescence were concatenated (joined together in a single matrix). Before the model was calculated, data were balanced to have all variables in the same scale [39–41].

To assess the performance of the classification model, sensitivity, specificity and efficiency were calculated [9,40,42] on the basis of true assignments (true positive and true negative) and false assignments (false positive and false negative).

3. Results and discussion

Fig. 1 shows the mean FT-Raman (Fig. 1a) and fluorescence (Fig. 1b) spectra of the EVOOs for each of the geographical PDOs considered. At first glance, no differences are observed between the spectra

corresponding to each of the classes (LG and S). The bands and their intensities appear as expected when EVOOs are analyzed, both for FT-Raman [43] and fluorescence [44].

Prior to multivariate analysis, a 15-point smoothing was applied to the FT-Raman spectra and variables were selected by considering the VIP scores, which reduced the number of variables from 6431 to 1681. Fluorescence spectra were not pre-treated.

Principal component analysis (PCA) was applied to the centered VIP-FT-Raman spectra and centered fluorescence spectra of the data set (84 samples). The score plot PC1 vs PC2 of the fluorescence spectra (Supplementary figure 1) showed not a clear separation between the two classes but a trend of separation was observed. Similarly, the separation between classes was not clear in the score plot of the FT-Raman spectra, although a trend was observed when PC1 was plotted against PC3 (Fig. 2S). In both cases, based on the score plot, no outliers were detected.

First, independent two-class models were built for each data source (FT-Raman and fluorescence) and for each of the three random data sets. Each model was validated by leave-one-out cross-validation. The optimal number of LVs used to build the PLS-DA models was selected on the basis of the RMSECV. For FT-Raman and fluorescence models, the first two and four PCs were considered, respectively. Looking at the PLS-

Table 1
Model performance parameters obtained from individual techniques and low-level data fusion.

| | TRAINING | | | TEST | | |
|---------------------|---------------|-------------|------------|-------------|-------------|------------|
| | Sensitivity | Specificity | Efficiency | Sensitivity | Specificity | Efficiency |
| FT-Raman | LES GARRIGUES | | | | | |
| Random 1 | 95 | 100 | 98 | 53 | 93 | 73 |
| Random 2 | 100 | 100 | 100 | 80 | 93 | 97 |
| Random 3 | 100 | 97 | 98 | 47 | 100 | 73 |
| | SIURANA | | | | | |
| Random 1 | 100 | 95 | 98 | 93 | 53 | 73 |
| Random 2 | 100 | 100 | 100 | 93 | 80 | 97 |
| Random 3 | 97 | 100 | 98 | 100 | 47 | 73 |
| Fluorescence | LES GARRIGUES | | | | | |
| Random 1 | 77 | 91 | 85 | 93 | 93 | 93 |
| Random 2 | 91 | 91 | 91 | 67 | 93 | 80 |
| Random 3 | 86 | 94 | 91 | 67 | 93 | 80 |
| | SIURANA | | | | | |
| Random 1 | 91 | 77 | 85 | 93 | 93 | 93 |
| Random 2 | 91 | 91 | 91 | 93 | 67 | 80 |
| Random 3 | 94 | 86 | 91 | 93 | 67 | 80 |
| low-level | LES GARRIGUES | | | | | |
| Random 1 | 100 | 100 | 100 | 100 | 93 | 97 |
| Random 2 | 91 | 84 | 97 | 80 | 80 | 80 |
| Random 3 | 100 | 100 | 100 | 67 | 100 | 83 |
| | SIURANA | | | | | |
| Random 1 | 100 | 100 | 100 | 93 | 100 | 97 |
| Random 2 | 84 | 91 | 87 | 80 | 80 | 80 |
| Random 3 | 100 | 100 | 100 | 100 | 67 | 83 |

Table 2
Number of samples submitted to high-level fusion.

| Samples type | | Random 1 | Random 2 | Random 3 |
|---------------|----------|----------|----------|----------|
| Les Garrigues | Training | 6 | 2 | 3 |
| | Test | 8 | 8 | 13 |
| Siurana | Training | 3 | 3 | 3 |
| | Test | 2 | 2 | 1 |

DA score plots, it can be seen that with the first two LV's, with the fluorescence spectra the trend observed in the PCA score plot was kept (Fig. 2a), while with the FT-Raman, the trend is highlighted and a separation between the two classes was observed (Fig. 2b).

Table 1 shows the performance parameters of the models developed from the data obtained by the individual techniques and by the low-level data fusion strategy. In the training set, it can be observed that FT-Raman presents excellent results with values higher than 95% for both sensitivity and specificity, regardless of the random subset considered. In the case of the models obtained with fluorescence data, the values are slightly lower (from 77% to 94%). The influence of the random set was also observed. For instance, with random set 1 approximately 23% of the Les Garrigues samples were wrongly assigned as Siurana while with random sets 2 and 3, the erroneous assignment dropped to 9 and 14%.

The performance parameters obtained with the test samples were generally worse and, with both techniques, the random set used was also observed to have an influence. This was more evident in the case of FT-Raman and Les Garrigues. It is to be expected that the more samples there are, the lower the influence of random choice will be. However, in a real application, it is not always easy to find a large number of samples of guaranteed origin.

Since the results are susceptible to improvement, we applied two data fusion strategies: low- and high-level. The low-level PLS DA models built by concatenating and normalizing the 1681 FT-Raman with the 440 fluorescence variables required 3 or 4 LVs depending on the random data set considered. Table 1 shows that the performance parameters of the test set improved for both techniques and also for the fluorescence training set (there was little change in the FT-Raman set) which in a way shows the synergy between both techniques.

To apply high-level fusion, given the nature of the operators, the

Table 3
Representative examples of class assignment using high-level data fusion.

| real sample type | | y predicted value | | |
|-------------------|---------------|-------------------|--------------|-------------------|
| | | LG class | S class | Ensemble decision |
| S training (2116) | FT-Raman | 0,325 | 0,675 | |
| | Fluorescence | 0,698 | 0,302 | |
| | Minimum | 0,325 | 0,302 | LG |
| | Maximum | 0,698 | 0,675 | LG |
| | Product | 0,227 | 0,204 | LG |
| | Mean | 0,512 | 0,488 | LG |
| | majority vote | | | LG |
| S training(0336) | FT-Raman | 0,312 | 0,688 | |
| | Fluorescence | 0,525 | 0,475 | |
| | Minimum | 0,312 | 0,475 | S |
| | Maximum | 0,525 | 0,688 | S |
| | Product | 0,164 | 0,327 | S |
| | Mean | 0,419 | 0,581 | S |
| | majority vote | | | S |
| LGtest(3148) | FT-Raman | 0,633 | 0,367 | |
| | Fluorescence | 0,318 | 0,682 | |
| | Minimum | 0,318 | 0,367 | S |
| | Maximum | 0,633 | 0,682 | S |
| | Product | 0,201 | 0,250 | S |
| | Mean | 0,476 | 0,524 | S |
| | majority vote | | | S |
| LGtest(0355) | FT-Raman | 0,489 | 0,511 | |
| | Fluorescence | 0,842 | 0,158 | |
| | Minimum | 0,489 | 0,158 | LG |
| | Maximum | 0,842 | 0,511 | LG |
| | Product | 0,412 | 0,081 | LG |
| | Mean | 0,666 | 0,334 | LG |
| | majority vote | | | LG |

results that were the same for both techniques will still be the same after the fusion so, for practical purposes, only non-concordant results were fused. Table 2 shows the number of non-concordant samples and the group to which they belong. It can be observed that most of them belonged to the Les Garrigues class. The percentage of non-concordant samples oscillates around 20%. It must be emphasized that all concordant samples were positive (properly assigned to their own class).

The fusion operators were applied to the values predicted by the PLS DA models obtained with the individual techniques, and the higher

Table 4
Model performance parameters obtained with high-level data fusion.

| | TRAINING | | | TEST | | |
|----------|---------------|-------------|------------|-------------|-------------|------------|
| | Sensitivity | Specificity | Efficiency | Sensitivity | Specificity | Efficiency |
| | LES GARRIGUES | | | | | |
| Random 1 | 95 | 100 | 98 | 93 | 100 | 97 |
| Random 2 | 100 | 97 | 98 | 93 | 93 | 93 |
| Random 3 | 94 | 100 | 98 | 73 | 100 | 87 |
| | SIURANA | | | | | |
| Random 1 | 100 | 95 | 98 | 100 | 93 | 97 |
| Random 2 | 97 | 100 | 98 | 93 | 93 | 93 |
| Random 3 | 100 | 95 | 98 | 100 | 73 | 87 |

values (indicated in bold, Table 3) were considered for the final ensemble decision. As an example, Table 3 presents the result of the ensemble decision of four samples selected to show different fusion outcomes: that is, samples of both classes that were properly assigned (Straining0336 and LGtest0355) and samples not solved (Straining2116 and LGtest3148).

Most of the non-concordant samples submitted to high-level fusion were favourably resolved. Table 4 shows the values of the performance parameters obtained and that the results were slightly better than after low-level data fusion, although both data fusion strategies improved the classification of EVOOSs.

A comparison of the three random data sets shows that the results are worst for selection 3 in all cases studied (individual models, and low- and high-level strategies). Even if selection 3 is included, it can be concluded that the Siurana test samples are properly recognized by their own model and not assigned to the Les Garrigues class (values higher than 93% or 100%). Les Garrigues test samples were mistakenly assigned to the Siurana class in approximately 25% (random 3) and 7% (random 1 and 2) of the cases.

In comparison with other reported methods, the classification models based on the individual techniques (FT-Raman and Fluorescence) gave similar performance as the referenced at the introduction section [21–23,27]. An added difficulty in this study is that all the samples belong to the same monovarietal type (Arbequina) and the two geographical areas are adjacent, although they have their own PDO. As for the data fusion strategy, the most widely applied according to the references, is the mid-level data fusion, even with other types of samples. Few references deal with high-level data fusion. The performance parameters of the high-level data fusion were comparable to the mid-level. Furthermore, high-level is really simple to implement since it is based on the combination of the results obtained from each individual model, which means that it is not necessary to develop another multivariate model.

4. Conclusions

FT-Raman and fluorescence with PLS-DA were used to authenticate EVOOs from two geographically adjacent Catalan PDOs. The data obtained for each technique were randomly divided into two subsets: the training set and the test set. This division was carried out in triplicate to obtain three datasets (random 1, 2 and 3) so that the performance parameters were more representative.

When the classification models were established with the data of each technique individually, the influence of the random set was observed. By way of example, sensitivities for the Les Garrigues oils varied significantly when they were evaluated with the three random test sets. Values were 47, 53 and 80% when FT-Raman was used and 67, 67 and 93% with fluorescence. It is to be expected that the more samples there are, the lower the influence of random choice will be. However, in a real application, it is not always easy to find a large number of samples of guaranteed origin.

After applying the two fusion strategies the performance parameters clearly improved, which indicates that there are synergies between FT-

Raman and fluorescence when characterizing samples from both PDOs. Further improvement was observed when the samples were randomly divided into training and test sets. In these conditions the sensitivities of the Les Garrigues class from the test set were 73, 93 and 93%. The variation was clearly lower.

Although the proposed strategy implies the use of more than one instrument, nowadays, many laboratories have a wide range of analytical equipment, which means that the data fusion strategy is a feasible way to obtain complementary information in a multivariate approach. Both FT-Raman and in fluorescence measurements are fast, simple and do not require the sample to be pre-treated, so we consider that the methodology presented is highly suitable to address the problem presented.

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.

Acknowledgements

This work was supported by the research program “Program of research activity (2016PFR-URV-B2)” at the Rovira i Virgili University, Tarragona, Spain. Special mention to M. Angels Calvo, head of the Official Tasting Panel of Virgin Olive Oils of Catalonia Governmentss.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.microc.2021.106611>.

References

- [1] European Food Safety Authority, EFSA J. 9, 2033 (2011).
- [2] S. Portarena, C.h. Anselmi, C. Zadra, D. Farinelli, F. Famiani, C.h. Baldacchini, E. Brugnoli, Cultivar discrimination, fatty acid profile and carotenoid characterization of monovarietal olive oils by Raman spectroscopy at a single glance, *Food Control* 96 (2019) 137–145, <https://doi.org/10.1016/j.foodcont.2018.09.011>.
- [3] S. Vichi, A. Tres, B. Quintanilla-Casas, J. Bustamante, F. Guardiola, E. Martí, J. F. Hermoso, A. Ninot, A. Romero, Catalan virgin olive oil protected designations of origin: physicochemical and major sensory attributes, *J. Lipid Sci. Technol.* 121 (3) (2019) 1800130, <https://doi.org/10.1002/ejlt.v121.310.1002/ejlt.201800130>.
- [4] Regulation (EEC) No 2081/92 of 14 July 1992 on the protection of geographical indications and designations of origin for agricultural products and foodstuffs, *Official Journal of the European Communities*, L208/1–8.
- [5] Regulation (EU) No 29/2012 of 13 January 2012 on marketing standards for olive oil, *Official Journal of the European Communities*, L12/21.
- [6] M. De la Guardia, A. González, *Food Protected Designation of Origin – Methodologies and Applications*, Elsevier, Amsterdam, 2013.
- [7] M. Meenu, Q. Cai, Xu Ba, A critical review on analytical techniques to detect adulteration of extra virgin olive oil, *Trends Food Sci. Technol.* 91 (2019) 391–408, <https://doi.org/10.1016/j.tifs.2019.07.045>.
- [8] M.P. Callao, I. Ruisánchez, An overview of multivariate qualitative methods for food fraud detection, *Food Control* 86 (2018) 283–293, <https://doi.org/10.1016/j.foodcont.2017.11.034>.

- [9] M.I. López, M.P. Callao, I. Ruisánchez, Tutorial on the validation of qualitative methods: from the univariate to the multivariate approach, *Anal. Chim. Acta* 891 (2015) 62–72, <https://doi.org/10.1016/j.aca.2015.06.032>.
- [10] E. Szymanska, J. Gerretzen, J. Engel, B. Geurts, L. Blanchet, L.M.C. Buydens, Chemometrics and qualitative analysis have a vibrant relationship, *Trends Anal. Chem.* 69 (2015) 34–51, <https://doi.org/10.1016/j.trac.2015.02.015>.
- [11] M. Pan, S. Sun, Q. Zhou, J. Chen, A Simple and Portable Screening Method for Adulterated Olive Oils Using the Hand-Held FTIR Spectrometer and Chemometrics Tools, *J. Food Sci.* 83 (6) (2018) 1605–1612, <https://doi.org/10.1111/1750-3841.14190>.
- [12] H. Miho, C.M. Díez, A. Mena-Bravo, V. Sánchez de Medina, J. Morala, E. Melliou, P. Magiatis, L. Rallo, D. Barranco, F. Priego-Capote, Cultivar influence on variability in olive oil phenolic profiles determined through an extensive germplasm survey, *Food Chem.* 266 (2018) 192–199, <https://doi.org/10.1016/j.foodchem.2018.06.002>.
- [13] İ.S. Özdemir, C. Dağ, D. Makuc, E. Ertaş, J. Plavec, S. Bekiroğlu, Characterisation of the Turkish and Slovenian extra virgin olive oils by chemometric analysis of the presaturation ¹H NMR spectra, *LWT - Food Sci. Technol.* 92 (2018) 10–15, <https://doi.org/10.1016/j.lwt.2018.02.015>.
- [14] A.M. Gómez-Caravaca, R.M. Maggio, L. Cerretani, Chemometric applications to assess quality and critical parameters of virgin and extra-virgin olive oil. A review, *Analytica Chimica Acta* 913 (2016) 1–21, <https://doi.org/10.1016/j.aca.2016.01.025>.
- [15] S. Guyader, F. Thomas, V. Portaluri, E. Jamin, S. Akoka, V. Silvestre, G. Remaud, Authentication of edible fats and oils by non-targeted ¹³C INEPT NMR spectroscopy, *Food Control* 91 (2018) 216–224, <https://doi.org/10.1016/j.foodcont.2018.03.046>.
- [16] N. Nenadis, M.Z. Tsimidou, Perspective of vibrational spectroscopy analytical methods in on-field/official control of olives and virgin olive oil, *Eur. J. Lipid Sci. Technol.* 119 (2017) 1600148, <https://doi.org/10.1002/ejlt.201600148>.
- [17] F.T. Borghi, P.C. Santos, F.D. Santos, M.H.C. Nascimento, T. Corrêa, M. Cesconetto, A.A. Pires, A.V.F.N. Ribeiro, V. Lacerda, W. Romão, P.R. Filgueiras, Quantification and classification of vegetable oils in extra virgin olive oil samples using a portable near-infrared spectrometer associated with chemometrics, *Microchem. J.* 159 (2020) 105544, <https://doi.org/10.1016/j.microc.2020.105544>.
- [18] T.O. Mendes, R.A. da Rocha, B.L.S. Porto, M.A.L. de Oliveira, V.d.C. dos Anjos, M.J. V. Bell, Quantification of Extra-virgin Olive Oil Adulteration with Soybean Oil: a Comparative Study of NIR, MIR, and Raman Spectroscopy Associated with Chemometric Approaches, *Food Anal. Methods* 8 (9) (2015) 2339–2346, <https://doi.org/10.1007/s12161-015-0121-y>.
- [19] E. Sánchez-López, M.I. Sánchez-Rodríguez, A. Marinas, J.M. Marinas, F.J. Urbano, J.M. Caridad, M. Moalem, Chemometric study of Andalusian extra virgin olive oils Raman spectra: Qualitative and quantitative information, *Talanta* 156–157 (2016) 180–190, <https://doi.org/10.1016/j.talanta.2016.05.014>.
- [20] K. Georgouli, J. Martínez del Rincon, A. Koidis, Continuous statistical modelling for rapid detection of adulteration of extra virgin olive oil using mid-infrared and Raman spectroscopic data, *Food Chem.* 217 (2017) 735–742, <https://doi.org/10.1016/j.foodchem.2016.09.011>.
- [21] R. Korifi, Y. Le Dreau, J. Molinet, J. Artaud, N. Dupuy, Composition and authentication of virgin olive oil from French PDO regions by chemometric treatment of Raman spectra, *J. Raman Spectroscopy* 42 (2011) 1540–1547, <https://doi.org/10.1002/jrs.2891>.
- [22] A.M. Jiménez-Carvelo, M.T. Osorio, A. Koidis, A. González-Casado, L. Cuadros-Rodríguez, Chemometric classification and quantification of olive oil in blends with any edible vegetable oils using FTIR-ATR and Raman spectroscopy, *LWT - Food Sci. Technol.* 86 (2017) 174–184, <https://doi.org/10.1016/j.lwt.2017.07.050>.
- [23] N. McReynolds, J.M. Aunón García, Z. Guengerich, T.K. Smith, K. Dholakia, Optical Spectroscopic Analysis for the Discrimination of Extra-Virgin Olive Oil, *Appl. Spectrosc.* 70 (11) (2016) 1872–1882, <https://doi.org/10.1177/0003702816645931>.
- [24] T. Mu, S. Chen, Y. Zhang, H.e. Chen, P. Guo, F. Meng, Portable detection and quantification of olive oil adulteration by 473-nm laser-induced fluorescence, *Food Anal. Methods* 9 (1) (2016) 275–279, <https://doi.org/10.1007/s12161-015-0199-2>.
- [25] K.D. Tavares Melo Milanez, T.C. Araújo Nóbrega, D. Nascimento, M. Insausti, B. S. Fernández Band, M.J. Coelho Pontes, Multivariate modeling for detecting adulteration of extra virgin olive oil with soybean oil using fluorescence and UV-Vis spectroscopies: A preliminary approach, *LWT - Food Sci. Technol.* 85 (Part A) (2017) 9–15, <https://doi.org/10.1016/j.lwt.2017.06.060>.
- [26] K. Abamba Omwange, D.F. Al Riza, Y. Saito, T. Suzuki, Y. Ogawa, K. Shiraga, F. Giametta, N. Kondo, Potential of front face fluorescence spectroscopy and fluorescence imaging in discriminating adulterated extra-virgin olive oil with virgin olive oil, *Food Control* 124 (2021) 107906, <https://doi.org/10.1016/j.foodcont.2021.107906>.
- [27] A.M. Jiménez-Carvelo, V.A. Lozano, A.C. Olivieri, Comparative chemometric analysis of fluorescence and near infrared spectroscopies for authenticity confirmation and geographical origin of Argentinean extra virgin olive oils, *Food Control* 96 (2019) 22–28, <https://doi.org/10.1016/j.foodcont.2018.08.024>.
- [28] O. Uncu, A. Napiórkowska, T.K. Szajna, B. Ozen, Evaluation of three spectroscopic techniques in determination of adulteration of cold pressed pomegranate seed oils, *Microchem. J.* 158 (2020) 105128, <https://doi.org/10.1016/j.microc.2020.105128>.
- [29] D.F. Al Riza, N. Kondo, V.K. Rotich, C. Perone, F. Giametta, Cultivar and geographical origin authentication of Italian extra virgin olive oil using front-face fluorescence spectroscopy and chemometrics, *Food Control* 121 (2021) 107604, <https://doi.org/10.1016/j.foodcont.2020.107604>.
- [30] J. Biancolillo A., Boqué R., Cocchi M. & Marini, F., (2019) Data handling in science and technology, Data fusion methodology and applications, vol.31, Chapter 10 - Data Fusion Strategies in Food Analysis, Elsevier B.V. Amsterdam, Netherlands. ISBN: 978-0-444-63984-4.
- [31] D.N. Vera, A.M. Jiménez-Carvelo, L. Cuadros-Rodríguez, I. Ruisánchez, M. P. Callao, Authentication of the geographical origin of extra-virgin olive oil of the Arbequina cultivar by chromatographic fingerprinting and chemometrics, *Talanta* 203 (2019) 194–202, <https://doi.org/10.1016/j.talanta.2019.05.064>.
- [32] T.R. Gonçalves, L.N. Rosa, A.S. Torquato, L.F.O. da Silva, P.H. Marçó, S.T. M. Gomes, M. Matsushita, P. Valderrama, Assessment of Brazilian Monovarietal Olive Oil in Two Different Package Systems by Using Data Fusion and Chemometrics, *Food Anal. Methods* 13 (1) (2020) 86–96, <https://doi.org/10.1007/s12161-019-01511-w>.
- [33] A. Maléchaux, Y. Le Dréau, J. Artaud, N. Dupuy, Control chart and data fusion for varietal origin discrimination: Application to olive oil, *Talanta* 217 (2020) 121115, <https://doi.org/10.1016/j.talanta.2020.121115>.
- [34] H. Wang, X. Wan, Effect of chlorophyll fluorescence quenching on quantitative analysis of adulteration in extra virgin olive oil, *Spectrochimica Acta Part A: Mol. Biomolecular Spectroscopy* 248 (2021), <https://doi.org/10.1016/j.saa.2020.119183>.
- [35] I.G. Chong, C.H. Jun, Performance of some variable selection methods when multicollinearity is present, *Chemom. Intell. Lab. Syst.* 78 (2005) 103–112, <https://doi.org/10.1016/j.chemolab.2004.12.011>.
- [36] C.S.W. Miaw, M.M. Sena, S.V.C. de Souza, I. Ruisánchez, M.P. Callao, Variable selection for multivariate classification aiming to detect individual adulterants and their blends in grape nectars, *Talanta* 190 (2018) 55–61, <https://doi.org/10.1016/j.talanta.2018.07.078>.
- [37] R.G. Brereton, G.R. Lloyd, Partial least squares discriminant analysis: taking the magic away, *J. Chemom.* 28 (2014) 213–225, <https://doi.org/10.1002/cem.2609>.
- [38] C. Márquez, I. Ruisánchez, M.P. Callao, Qualitative and quantitative multivariate strategies for determining paprika adulteration with Sudan I and II dyes, *Microchem. J.* 145 (2019) 686–692, <https://doi.org/10.1016/j.microc.2018.11.034>.
- [39] C.V. Di Anibal, M.P. Callao, I. Ruisánchez, ¹H NMR and UV-visible data fusion for determining Sudan dyes in culinary spices, *Talanta* 84 (3) (2011) 829–833, <https://doi.org/10.1016/j.talanta.2011.02.014>.
- [40] C. Márquez, M.I. López, I. Ruisánchez, M.P. Callao, FT-Raman and NIR spectroscopy data fusion strategy for multivariate qualitative analysis of food fraud, *Talanta* 161 (2016) 80–86, <https://doi.org/10.1016/j.talanta.2016.08.003>.
- [41] E. Borrás, J. Ferré, R. Boqué, M. Mestres, L. Aceña, A. Calvo, O. Busto, Olive oil sensory defects classification with data fusion of instrumental techniques and multivariate analysis (PLS-DA), *Food Chemistry* 203 (2016) 314–322, <https://doi.org/10.1016/j.foodchem.2016.02.038>.
- [42] C.S.W. Miaw, M.M. Sena, S.V.C. de Souza, M.P. Callao, I. Ruisánchez, Detection of adulterants in grape nectars by attenuated total reflectance Fourier-transform mid-infrared spectroscopy and multivariate classification strategies, *Food Chem.* 266 (2018) 254–261, <https://doi.org/10.1016/j.foodchem.2018.06.006>.
- [43] I. Ruisánchez, A.M. Jiménez-Carvelo, M.P. Callao, ROC curves for the optimization of one-class model parameters. A case study: Authenticating extra virgin olive oil from a Catalan protected designation of origin, *Talanta* 222 (2021) 121564, <https://doi.org/10.1016/j.talanta.2020.121564>.
- [44] K.D. Tavares Melo Milanez, T.C. Araújo Nóbrega, N.D. Silva, M. Insausti, M. J. Coelho Pontes, Transfer of multivariate classification models applied to digital images and fluorescence spectroscopy data, *Microchem. J.* 133 (2017) 669–675, <https://doi.org/10.1016/j.microc.2017.03.004>.