JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

J. Agric. Food Chem. **2004,** *52,* 6673−6679 **6673**

Cluster Analysis Applied to the Exploratory Analysis of Commercial Spanish Olive Oils by Means of Excitation−Emission Fluorescence Spectroscopy

FRANCESCA GUIMET,* RICARD BOQUÉ, AND JOAN FERRÉ

Department of Analytical Chemistry and Organic Chemistry, Rovira i Virgili University, Plaça Imperial Tàrraco 1, E-43005 Tarragona, Catalonia, Spain

Olive oil fluorescence is related to oil composition. Here it is shown that the natural clustering of different types of commercial Spanish olive oils depends on their fluorescence excitation−emission matrices (EEMs). Fifty-six commercial samples of olive oil (29 virgin olive oils, 20 pure olive oils, and 7 olive-pomace oils) were used. The clustering method was hierarchical agglomerative clustering using the Euclidean distance as a similarity measure and the average linkage. Two spectral ranges were considered (which either contained the fluorescence peak of the chlorophylls or did not), and various methods for preprocessing the fluorescence spectra were compared. The oils were clearly distinguished using the unfolded EEMs measured between $\lambda_{ex} = 300-400$ nm and $\lambda_{em} = 400-600$ nm. The optimal preprocessing was normalization of the unfolded spectra followed by column autoscaling. Also shown are the advantages of using second-order data (EEMs) instead of first-order data (a single fluorescence spectrum) for each sample.

KEYWORDS: Fluorescence spectroscopy; olive oil; cluster analysis; second-order data

INTRODUCTION

Food science is receiving more and more attention because of its close relationship with health. Chemometric techniques applied to analytical data have proved to be important tools in food analysis, because they can be used for exploratory analysis and classification (*1*−*4*). Olive oil is an economically important product. It is obtained from the fruit of the olive tree (*Olea europaea* L.). There are different types of olive oils, with virgin olive oil being the one with the best quality. The characteristic odor and flavor of virgin olive oil are due to the olives being mechanically pressed and the lack of any refining processes. Refined olive oil is obtained from virgin olive oil with refining methods that do not lead to alterations in the initial glyceridic structure. This oil does not have adequate organoleptic properties. Hence, it is blended with virgin olive oil to form pure olive oil (or simply olive oil). By extracting the olive-pomace (i.e., the olive residue remaining from previous pressings) with authorized solvents, refined olive-pomace oil is obtained. This oil is improved with edible virgin oil to obtain the oil known as olive-pomace oil (*4*, *5*).

Olive oil has been analyzed with such techniques as chromatography (*6*, *7*), mass spectrometry (MS) (*4*), and a variety of spectroscopic techniques: infrared (IR) and Fourier transform (FT)-Raman (*8*−*11*), nuclear magnetic resonance (NMR) (*9*, *12*, *13*), fluorescence (*14*−*17*), and chemiluminescence (*18*). The advantages of fluorescence spectroscopy are its speed of analysis and the fact that solvents and reagents are not required, because olive oil exhibits natural fluorescence (*14*, *15*). Other interesting advantages are that only a small amount of sample is needed and that it is a nondestructive technique. However, fluorescence applied to olive oil has been explored only very little, basically using a fluorescence emission spectrum recorded at one excitation wavelength (first-order data) (*14*, *15*). Nevertheless, it is also possible to record entire fluorescence excitation−emission matrices (EEMs), which consist of emission spectra measured at different excitation wavelengths (second-order data). Wolbeis and Leiner (*16*) used EEMs to characterize four types of edible oils. Scott et al. (*17*) applied pattern recognition methods to fluorescence EEMs to discriminate between four types of vegetable oils and to detect adulterations in extra virgin olive oils.

Cluster analysis (CA) is a pattern recognition technique used to form groups of objects having variables of similar values. Several similarity measures can be applied to form such groups (*19*, *20*). The main advantage of CA over visualization techniques such as principal component analysis (PCA) is that it provides numerical values of the similarity between objects. As a result, the information is more objective (*19*, *20*). In addition, when a large number of principal components (PCs) are required to visualize the information, CA has the advantage of reducing dimensionality while keeping the information. In many cases, the joint use of both visualization and clustering techniques is recommended (*19*). In the field of olive oil analysis, clustering has been applied to many different forms of data for many different reasons. It has been used with

* Corresponding author (e-mail pgv@quimica.urv.es; telephone +34-977-558-155; fax +34-977-559-563).

10.1021/jf040169m CCC: $27.50
Published on Web 09/30/2004

chromatographic data to characterize the stage of ripeness of virgin olive oil by determining the content of volatile compounds (*3*). With IR and FT-Raman spectroscopic data, it has been used to detect the adulteration of virgin olive oil (*10*, *11*), and with visible and near-IR spectroscopic data, it has made a geographic classification of Mediterranean extra virgin olive oils (*21*). With NMR spectroscopic data it has been used to study the effects of climatic conditions on olive oil (*12*) and to geographically categorize virgin olive oils (*13*). Finally, it has been used with sensory, chromatographic, and MS data to differentiate virgin olive oils on the basis of the extraction methodology adopted during industrial olive oil processing (*22*). However, there are no references to applications of CA to discriminate between different types of oils using fluorescence data.

The objective of this paper is to test the ability of CA to discriminate between the three main types of commercial Spanish olive oils used for human consumption (virgin, pure, and olive-pomace oil) using EEMs as fingerprints. We applied the hierarchical agglomerative clustering (HAC) method with the Euclidean distance as a similarity measure and the average linkage method (*19*, *20*) to the unfolded EEMs. Different preprocessing methods and EEM ranges were tested to find the most appropriate way of handling data and optimizing the sample grouping into clusters. Then we compare the results obtained from the EEMs to those obtained from a single fluorescence spectrum. This spectrum is selected as the one that maximizes the differences among samples. We show that the oils are best grouped into types when EEMs are used instead of single fluorescence spectra.

MATERIALS AND METHODS

Samples. A set of 56 olive oil bottles containing three different types of edible Spanish olive oils (29 virgin, 20 pure, and 7 olive-pomace oils) (**Table 1**) were purchased in a shopping center. Although they were not reference samples, they were suitable for the exploratory purposes of this study. Most of them were well-known Spanish brands, and some even prestigious. The samples were stored in the dark at room temperature until the moment of analysis. The samples were analyzed without any prior treatment.

Instrumentation and Software. Oil EEMs were measured with an Aminco Bowman series 2 luminescence spectrometer equipped with a 150 W xenon lamp and 10 mm quartz cells. The instrument detector was operated using the EmL/Ref channel and applying a 600 V voltage for virgin and pure olive oil samples. However, P32 and olive-pomace oil samples were measured at 580 and 560 V, respectively, to avoid detector saturation. Excitation and emission ranges were $\lambda_{ex} = 300-400$ nm and $\lambda_{em} = 400-700$ nm, respectively. Measuring emission wavelengths above excitation wavelengths prevented Rayleigh scatter. The step size and band-pass of both monochromators were 5 and 4 nm, respectively. The scan rate was 30 nm s^{-1}. The instrument software was used to correct the EEMs for deviations in the ideality of the lamp, monochromators, and detector (*23*, *24*). As a result, the last emission wavelength was lost and the final ranges were $\lambda_{ex} = 300-400$ nm and $\lambda_{em} = 400-695$ nm.

Data were processed with Matlab software (version 6.0) (*25*), and dendrograms were constructed from the Matlab statistics toolbox (*25*). PCA models were validated using Unscrambler software (version 8.0) (*26*).

Algorithm for Selecting the Most Discriminatory Wavelengths. Olive oil can be characterized by means of fluorescence spectroscopy. This has usually been done with emission spectra measured at one excitation wavelength (*14*, *15*). Here we checked whether EEMs would make the clustering better than when a single fluorescence spectrum was used for each sample. Because we had recorded the EEMs, we applied the following algorithm to select the most discriminatory excitation wavelength (λ_{ex}):

1. Take the fluorescence emission spectrum at $\lambda_{ex} = 300$ nm of every sample from the EEMs and make them the rows in a matrix **Q**.

Table 1. Samples

sample[a]	olive variety	origin
V1	Arbequina	La Palma d'Ebre (Tarragona)
V2	Arbequina	Reus (Tarragona)
V3	Arbequina	La Serra d'Almos (Tarragona)
V4	Arbequina	Llorenç del Penedès (Tarragona)
V5		Jaén
V6		Granada
V7		Jaén
V8	Hojiblanca, Arbequina	Córdoba
V9	Hojiblanca	Málaga
V10	Picual, Hojiblanca	Córdoba
V11	Picual, Hojiblanca, Picuda	Córdoba
V12		Sevilla
V13	Arbequina	Tàrrega (Lleida)
V14	Arbequina	Tàrrega (Lleida)
V15	Arbequina	Reus (Tarragona)
V16	Arbequina	Córdoba
V17	Arbequina	Les Borges Blanques (Lleida)
V18	Arbequina, Cornicabra, Hojiblanca	Jaén
V19	Hojiblanca	Tàrrega (Lleida)
V20	Hojiblanca	Tàrrega (Lleida)
V21	Picual	Tàrrega (Lleida)
V22	Picual	Tàrrega (Lleida)
V23		Jaén
V24		Córdoba
V25		Málaga
V26		Sevilla
V27		Tortosa (Tarragona)
V28		Córdoba
V29	Morrut, Farga, Sevillano	Montsià (Tarragona)
P30–49		different areas of Spain
OP50–56		different areas of Spain

[a] V, virgin olive oil; P, pure olive oil; OP, olive-pomace oil.

2. Normalize each row of **Q** ($\mathbf{q}_i$) to length one (*27*) to obtain $\mathbf{Q}_N$, that is, $\mathbf{q}_{Ni} = (\mathbf{q}_i/(\mathbf{q}_i\mathbf{q}_i^T)^{1/2})$, where T means transposed and $\mathbf{q}_{Ni}$ is the ith row of $\mathbf{Q}_N$.

3. Column mean-center $\mathbf{Q}_N$ to obtain $\mathbf{Q}_{NC}$, that is, calculate the mean of each column of $\mathbf{Q}_N$ and subtract it from every value in the column.

4. Calculate the sum of the squares (SS) of all the elements of $\mathbf{Q}_{NC}$. This SS value represents the differences between each fluorescence emission spectrum at $\lambda_{ex} = 300$ nm and the average spectrum.

5. Repeat steps 1−4 for the rest of λ_{ex}.

Hence, for each λ_{ex}, a value of SS was obtained. The λ_{ex} giving the highest SS was considered to be the most discriminatory of all the excitation wavelengths tested. To find the emission wavelength (λ_{em}) that best distinguished the types of oils, a similar procedure was used, but this time **Q** contained the fluorescence excitation spectrum of each sample at one λ_{em}.

RESULTS AND DISCUSSION

EEMs. Figure 1 shows the average EEMs of the three types of oils studied in the ranges $\lambda_{ex} = 300-400$ nm and $\lambda_{em} = 400-695$ nm. Virgin and pure olive oil spectra look similar due to their high peak between $\lambda_{ex} = 300-400$ nm and $\lambda_{em} = 650-695$ nm, which is attributed to chlorophylls (*14*, *15*). In the olive-pomace oil EEMs, this peak is much less intense and an intense peak appears between $\lambda_{ex} = 340-400$ nm and $\lambda_{em} = 400-550$ nm, which is attributed to oxidation products (*14*). Oxidation products are formed when oil comes in contact with oxygen. The process of oxidation of olive oil involves radical reactions between oxygen and the double bonds of unsaturated fatty acids. Light accelerates this process. As a result, conjugated hydroperoxides are formed. These compounds are unstable, and they quickly decompose into aldehydes and ketones (*5*).

Because the chlorophyll peak may hamper oil differentiation (*28*), the EEMs without this peak ($\lambda_{ex} = 300-400$ nm, $\lambda_{em} = 400-600$ nm) were also considered (**Figure 2**). In this range,

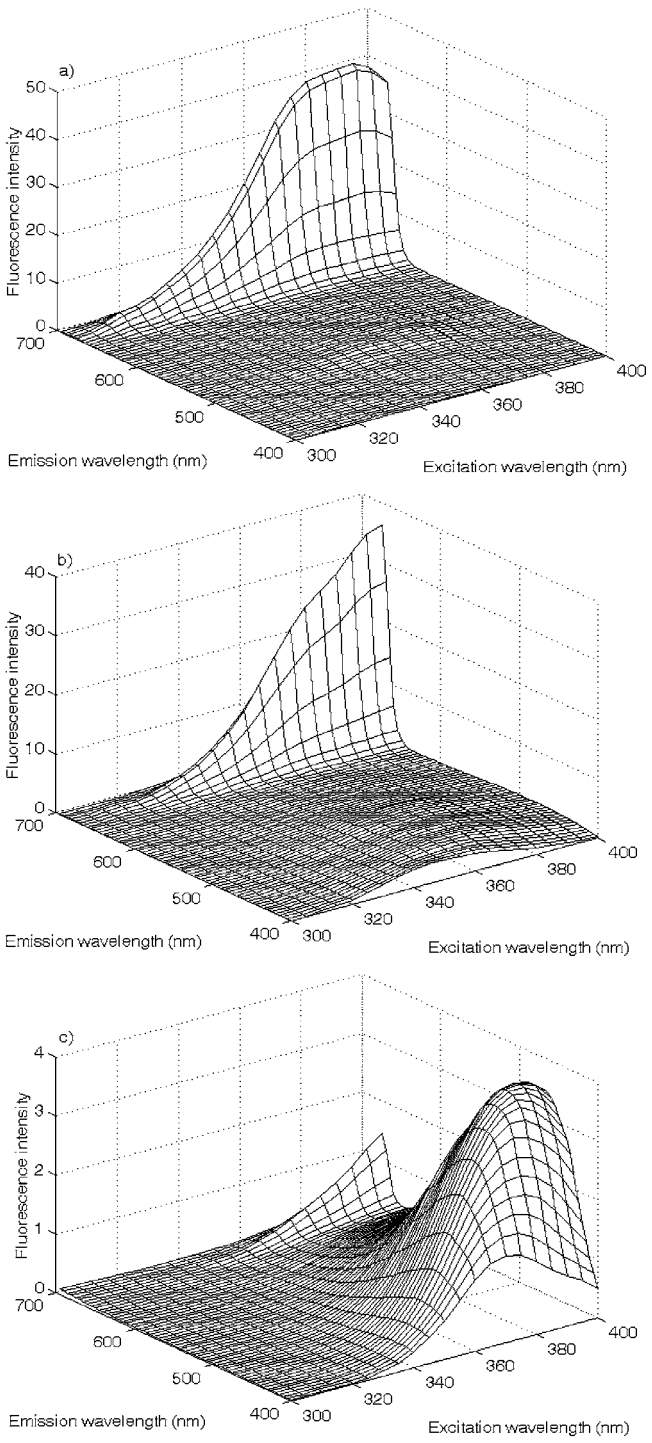

Figure 1. Average EEMs between $\lambda_{ex} = 300–400$ nm and $\lambda_{em} = 400–695$ nm: (**a**) virgin olive oil; (**b**) pure olive oil; (**c**) olive-pomace oil.

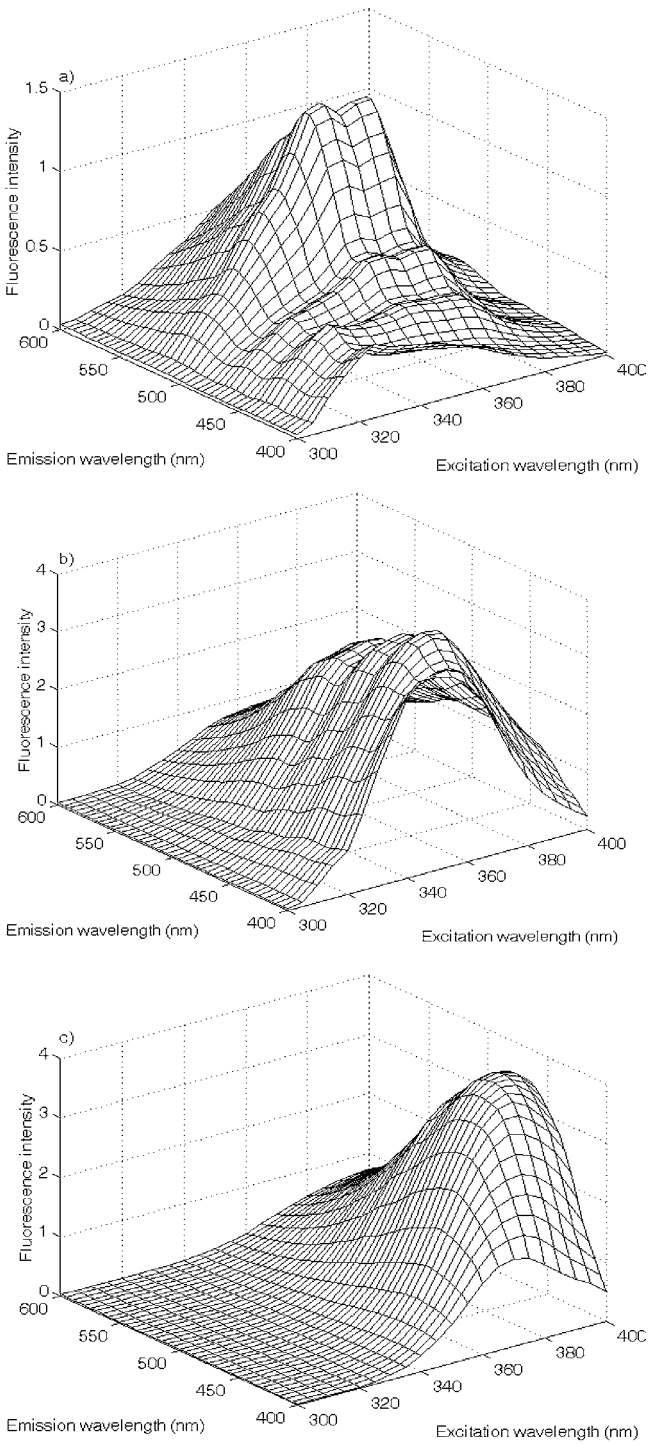

Figure 2. Average EEMs between $\lambda_{ex} = 300–400$ nm and $\lambda_{em} = 400–600$ nm: (**a**) virgin olive oil; (**b**) pure olive oil; (**c**) olive-pomace oil.

the EEMs of the studied oils have considerable differences (**Figure 2**). The shape of virgin olive oil EEMs (**Figure 2a**) is mainly due to vitamin E, which emits between $\lambda_{ex} = 300-400$ nm and $\lambda_{em} = 500-600$ nm, whereas pure and olive-pomace oil samples have a larger content of oxidation products, which give rise to a broad peak at lower emission wavelengths (*14*). The different position of this peak on pure and olive-pomace EEMs enables us to distinguish between them (**Figure 2b**,**c**).

To perform CA on the full range of EEMs, the matrices were stacked in a three-way array of $56 \times 21 \times 60$ (samples $\times$ $\lambda_{ex} \times \lambda_{em}$) and the EEMs without the chlorophylls in a three-way array of $56 \times 21 \times 41$ (samples $\times$ $\lambda_{ex} \times \lambda_{em}$).

Cluster Analysis of Unfolded EEMs. The HAC method was applied to the EEMs in the two ranges indicated above (with and without the chlorophyll peak). In both cases, each three-way array was first unfolded to a matrix of size [samples $\times$ ($\lambda_{ex} \times \lambda_{em}$)], where each row contained the unfolded EEM of a sample. These rows were then normalized to length one (*27*) to avoid variations due to differences in intensity, and the resulting matrix was column autoscaled (i.e., every column of the matrix was set to zero mean and unit variance).

EEMs Containing the Chlorophyll Peak. After unfolding of the full-range three-way array, a 56×1260 matrix was obtained; the rows were samples and the columns were the excitation and emission wavelengths. Neither the raw matrix nor the matrix

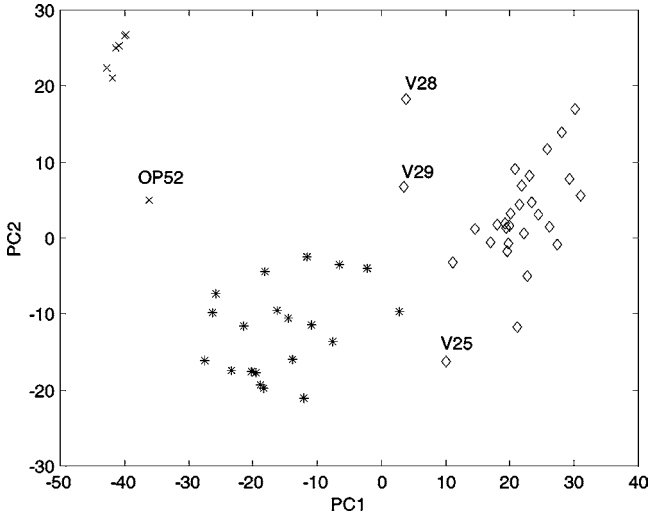

Figure 3. Score plot from PCA on the 56 × 861 unfolded matrix [samples × ($\lambda_{ex} \times \lambda_{em}$)], containing the spectra measured between λ_{ex} = 300–400 nm and λ_{em} = 400–600 nm: (◇) virgin olive oil: (∗) pure olive oil; (×) olive-pomace oil (normalized and autoscaled spectra).

preprocessed as indicated above provided a good distinction between the three types of oils. Two other preprocessing methods were also tested: row normalization to length one only (*27*) and normalization to length one followed by scaling within emission mode (*29*). However, the results did not improve. The reason was that the chlorophyll peak had large variations, even between samples of the same type. A previous study based on unfolded principal component analysis and parallel factor analysis (*28*) had already shown that the chlorophyll peak hampered oil separation. The inclusion of the chlorophyll peak caused a larger overlap between the groups. For this reason, the results for the EEMs will henceforth be shown in the ranges λ_{ex} = 300−400 nm and λ_{em} = 400−600 nm, without the chlorophyll peak.

EEMs without the Chlorophyll Peak. The unfolding of the three-way array not containing the chlorophyll peak led to a 56 × 861 matrix. The same three preprocessing methods used for the full-range matrices were tested. Then HAC was applied. Results were best when the matrix was normalized and column autoscaled. PCA was also calculated for this matrix to determine the number of variability sources in the data and to find what wavelengths caused the main differences between the types of olive oils. The PCA model was validated by leave-one-out cross-validation, and seven PCs (98.9% of explained variance) were found to be significant. However, the greatest differences between the types of oils could be seen from the score plot using the first two PCs (84.7% of explained variance (**Figure 3**). The three types of oils are separated mainly along PC1. Olive-pomace oils have the most negative scores and virgin olive oils the most positive scores. The loadings in **Figure 4** indicate that a high score in PC1 is related to a high content in vitamin E and a low content in oxidation products. This is concluded from the two main regions in **Figure 4a**. The region with positive values (region 1, around λ_{ex} = 300−380 nm and λ_{em} = 500−600 nm) is related to vitamin E, which emits fluorescence in this range (see **Figure 2a**). The region with negative loadings (region 2, around λ_{ex} = 340−400 and λ_{em} = 400−480 nm) is related to oxidation products, mainly those present in olive-pomace oils (see **Figure 2c**). Thus, the positive scores of virgin olive oils on PC1 indicate that they have a high vitamin E content and fewer oxidation products than the other oils. PC2 does not separate the groups as well as PC1, but pure olive oils

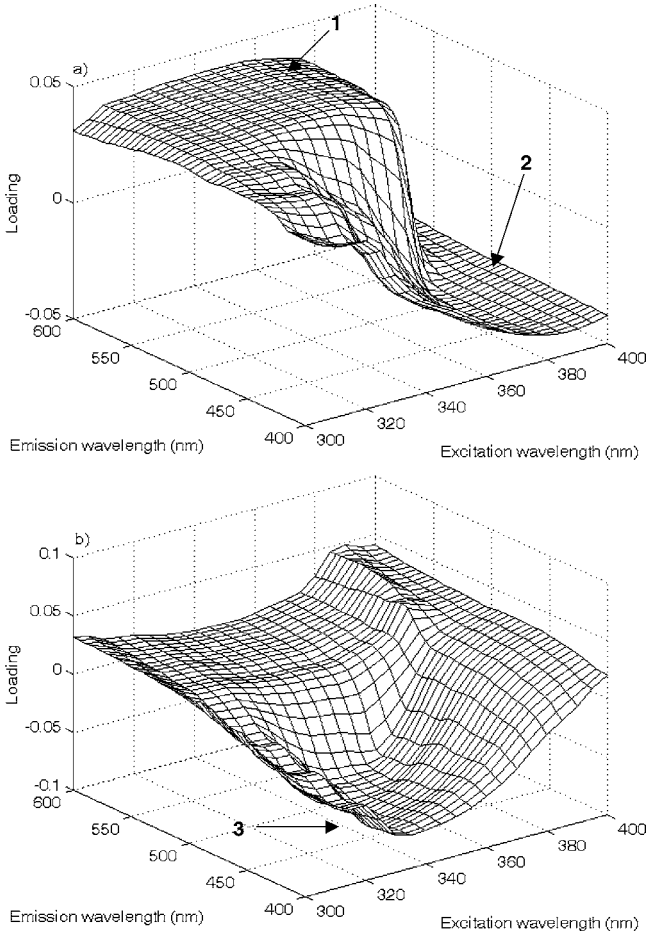

Figure 4. Loading plot from PCA on the 56 × 861 unfolded matrix [samples × ($\lambda_{ex} \times \lambda_{em}$)], containing the spectra measured between λ_{ex} = 300–400 nm and λ_{em} = 400–600 nm: (**a**) PC1; (**b**) PC2 (normalized and autoscaled spectra).

tend to have negative scores, whereas the rest of the oils tend to have positive scores. The wavelengths that most influence PC2 are in region 3, around λ_{ex} = 300−380 nm and λ_{em} = 400−550 nm (**Figure 4b**). This region has negative loadings and is related to the oxidation products of pure olive oils (see **Figure 2b**). Hence, a sample with negative scores on PC2 has a higher content on this type of products. Olive-pomace oil samples are very different from the rest on both PCs, and in general their scores are the lowest on PC1 and the highest on PC2. Only a few samples (V25, V28, V29, and OP52) appear separated from their oil group. This indicates that they are probably extreme examples of their sample type.

The CA dendrogram showed the three types of oils perfectly separated (**Figure 5**). The cophenetic correlation coefficient (Coph. *r*) was used as an indication of the cluster validity (*25*, *30*). It measures the correlation between the linked objects in the cluster tree and the distances between objects. A Coph. *r* close to one indicates that the cluster solution reflects the similarity between objects before the tree is built. In this case, Coph. *r* = 0.7. Thus, the solution was quite good. As in PCA, the largest differences are between olive-pomace oils and the rest. These two clusters merge at a dissimilarity level of 75%. Slightly below this level (dotted line), the dendrogram shows three well-differentiated clusters, each of which contains all of the samples of one of the types of oils studied. Even samples V25, V28, and V29 are grouped in the expected cluster. However, because the number of samples of olive-pomace oil is much smaller than the others, this result must be seen as a

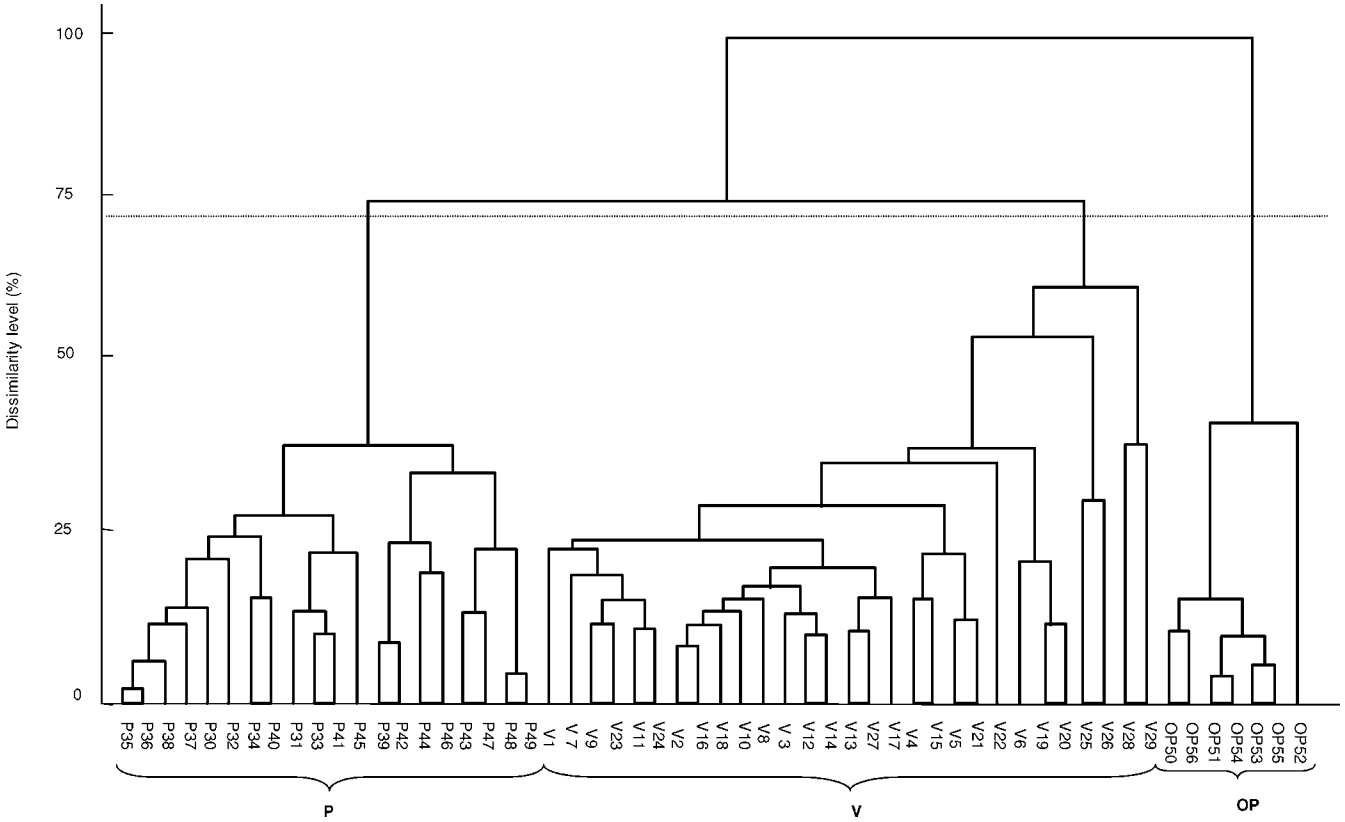

Figure 5. Dendrogram of the 56 × 861 unfolded matrix (λ_{ex} = 300–400 nm, λ_{em} = 400–600 nm) using the Euclidean distance as similarity measure and the average linkage method. The distance is expressed as a percentage of dissimilarity (normalized and autoscaled spectra).

trend and not as a general conclusion about the application of this method to olive-pomace oils. The sample source is also a limitation. Hence, the relevance of these clustering results must be considered, taking into account the facts that the samples were purchased in shopping centers and that they are not reference samples. Thus, we have only the information given by the suppliers on the label of the sample bottles.

Clustering Based on Selected Fluorescence Spectra. To determine whether the same information could be obtained by using a single fluorescence spectrum for each sample instead of the whole EEM, we applied the variable selection algorithm described under Materials and Methods to find the most discriminatory fluorescence excitation and emission spectra.

The most discriminatory variables were selected from the ranges λ_{ex} = 300−400 nm and λ_{em} = 400−600 nm. In all cases, after the variables had been selected, two preprocessing methods were applied to spectra taken from raw EEMs: normalization to length one (*27*) and normalization to length one plus column autoscaling. Then CA was applied.

Clustering Based on Emission Spectra. To find the excitation wavelength that provided the main differences between fluorescence emission spectra, the variable selection algorithm was applied for each λ_{ex}. Hence, for each λ_{ex} a 56 × 41 (samples × λ_{em}) matrix was built, and the algorithm was run on each matrix. The SS value was maximum for λ_{ex} = 345 nm. Hence, CA was applied to the emission spectra between λ_{em} = 400 and 600 nm at λ_{ex} = 345 nm (**Figure 6a**). The results were worse than when the 56 × 861 unfolded matrix was used. In both cases, the dendrograms displayed two large clusters, one of which contained all pure and olive-pomace oil samples and V25. Therefore, it was not possible to distinguish the oil types on the basis of the selected wavelengths. To test whether fluorescence emission spectra measured at other excitation wavelengths could improve oil clustering, the HAC method was again applied to the emission spectra at λ_{ex} = 365 nm [see Kyriakidis et al. (*14*)] and λ_{ex} = 390 nm, which was close to the excitation wavelength proposed by Marini et al. (λ_{ex} = 392 nm) (*15*). However, both preprocessing methods applied to these sets of spectra led to dendrograms that contained clusters with mixtures of samples of different types.

Clustering Based on Excitation Spectra. To find the emission wavelength that provided the main differences between fluorescence excitation spectra, the variable selection algorithm was applied again, this time for each λ_{em}. Hence, for each λ_{em} a 56 × 21 (samples × λ_{ex}) matrix was built, and the algorithm was run on each matrix. The difference was maximum for λ_{em} = 410 nm. Hence, CA was applied to the excitation spectra between λ_{ex} = 300 and 400 nm at λ_{ex} = 410 nm (**Figure 6b**). When only normalization was carried out, virgin olive oil samples V12, V14, and V29 could not be distinguished from the pure olive oil cluster and samples V28 and OP52 were very different from the rest of samples, joining the olive-pomace oil cluster. Autoscaling provided a dendrogram similar to the previous one. The most noteworthy difference is that V29 was the only virgin olive oil sample that remained in the pure olive oil cluster.

Conclusions. This work has shown the potential of fluorescence spectroscopy to distinguish three types of commercial Spanish olive oils (virgin, pure, and olive-pomace). Visual inspection of the EEMs of samples makes it possible to assign a large number of samples to the expected type. However, special characteristics, such as different olive oil varieties or the high deterioration of some virgin olive oil samples, may cause considerable variations in the shape of EEMs and thus hamper type differentiation. We used samples whose membership is stated in order to study the ability of the HAC method to cluster olive oil samples of different types. Different preprocessing methods and wavelength ranges were compared,

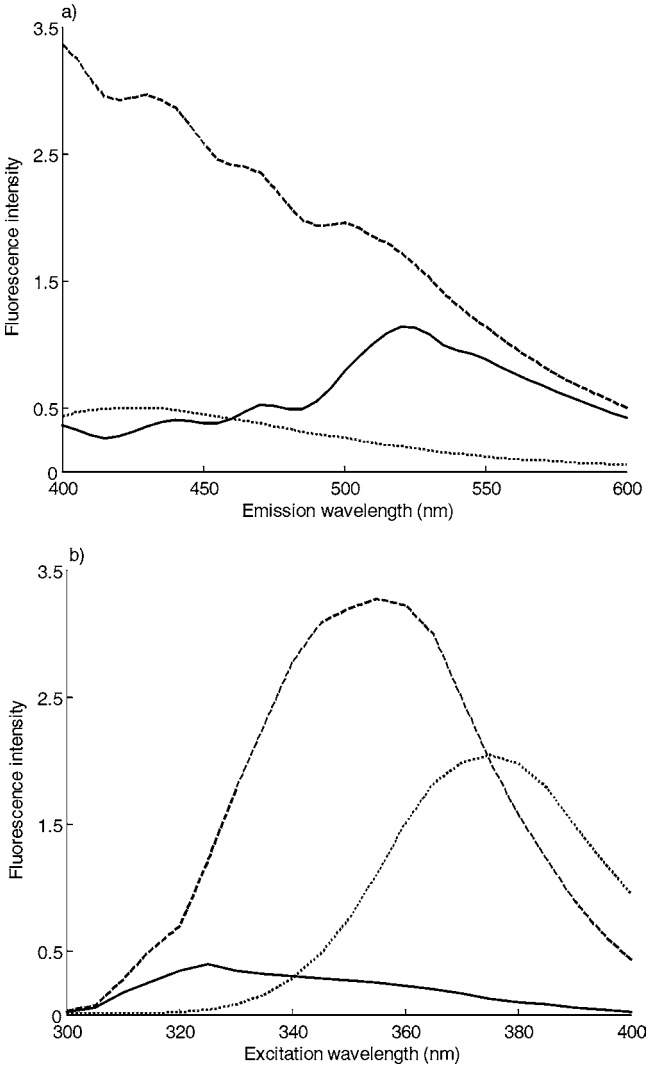

Figure 6. Average fluorescence spectra of each type of oil: (**a**) emission spectra between $\lambda_{em} = 400$ and 600 nm at $\lambda_{ex} = 345$ nm; (**b**) excitation spectra between $\lambda_{ex} = 300$ and 400 nm at $\lambda_{em} = 410$ nm; (–) virgin olive oil; (- - -) pure olive oil; (•••) olive-pomace oil.

and the results were best with the normalized and autoscaled unfolded matrix obtained from the EEMs between $\lambda_{ex} = 300-400$ nm and $\lambda_{em} = 400-600$ nm. Under these conditions, the samples were perfectly grouped. We also developed an algorithm for variable selection to find the most discriminatory emission and excitation spectra. However, the use of a single fluorescence spectrum worsened the grouping. These results indicate that working with second-order data is more advantageous than working with first-order data, because the larger number of variables used contains additional information that increases discrimination power.

The results obtained here might lay the foundations for developing and applying classification methods to the EEMs of olive oils. Possible applications are olive oil characterization or the detection of frauds.

ABBREVIATIONS USED

CA, cluster analysis; EEM, excitation−emission matrix; FT, Fourier transform; HAC, hierarchical agglomerative clustering; IR, infrared; MS, mass spectrometry; PC, principal component; PCA principal component analysis.

LITERATURE CITED

(1) Munck, L.; Norgaard, L.; Engelsen, S. B.; Bro, R.; Andersson, C. A. Chemometrics in food science—a demonstration of the feasibility of a highly exploratory, inductive evaluation strategy of fundamental scientific significance. *Chemom. Intell. Lab. Syst.* **1998**, *44*, 31−60.

(2) Forina, M.; Armanino, C.; Raggio, V. Clustering with dendrograms on interpretation variables. *Anal. Chim. Acta* **2002**, *454*, 13−19.

(3) Aparicio, R.; Morales, M. T. Characterization of olive ripeness by green aroma compounds of virgin olive oil. *J. Agric. Food Chem.* **1998**, *46*, 1116−1122.

(4) Marcos, I.; Pérez, J. L.; Fernández, M. E.; García, C.; Moreno, B. Detection of adulterants in olive oil by headspace-mass spectrometry. *J. Chromatogr. A* **2002**, *945*, 221−230.

(5) Kiritsakis, A. K. International Olive Oil Council—Quality Criteria and Classification of Olive Oil. In *Olive Oil: from the Tree to the Table*, 2nd ed.; Food and Nutrition Press: Trumbull, CT, 1998; pp 237−260.

(6) Cert, A.; Moreda, W.; Pérez-Camino, M. C. Chromatographic analysis of minor constituents in vegetable oils, review. *J. Chromatogr. A* **2000**, *881*, 131−148.

(7) Gandul-Rojas, B.; Cepero, M. R. L.; Mínguez-Mosquera, M. I. Use of chlorophyll and carotenoid pigment composition to determine authenticity of virgin olive oil. *J. Am. Oil Chem. Soc.* **2000**, *77*, 853−858.

(8) Lai, Y. W.; Kemsley, E. K.; Wilson, R. H. Quantitative analysis of potential adulterants of extra virgin olive oil using infrared spectroscopy. *Food Chem.* **1995**, *53*, 95−98.

(9) Dennis, M. J. Recent developments in food authentication. *Analyst* **1998**, *123*, 151R−156R.

(10) Downey, G.; McIntyre, P.; Davies, A. N. Detecting and quantifying sunflower oil adulteration in extra virgin olive oils from the eastern Mediterranean by visible and near-infrared spectroscopy. *J. Agric. Food Chem.* **2002**, *50*, 5520−5525.

(11) Baeten, V.; Meurens, M. Detection of virgin olive oil adulteration by Fourier transform Raman spectroscopy. *J. Agric. Food Chem.* **1996**, *44*, 2225−2230.

(12) Mannina, L.; Sobolev, A. P.; Segre, A. Olive oil as seen by NMR and chemometrics. *Spectrosc. Eur.* **2003**, *15*, 6−14.

(13) Sacchi, R.; Mannina, L.; Fiordiponti, P.; Barone, P.; Paolillo, L.; Patumi, M.; Segre, A. Characterization of Italian extra virgin olive oils using ^{1}H-NMR spectroscopy. *J. Agric. Food Chem.* **1998**, *46*, 3947−3951.

(14) Kyriakidis, N. B.; Skarkalis, P. Fluorescence spectra measurement of olive oil and other vegetable oils. *J. AOAC Int.* **2000**, *83*, 1435−1438.

(15) Marini, D.; Grassi, L.; Balestrieri, F.; Pascucci, E. Analisi spettrofotofluorimetrica dell'olio di oliva. Possibilità di applicazione. *Riv. Ital. Sostanze Grasse* **1990**, *67*, 95−99.

(16) Wolfbeis, O. S.; Leiner, M. Characterization of edible oils via 2D-fluorescence. *Mikrochim. Acta* **1984**, *1*, 221−233.

(17) Scott, S. M.; James, D.; Ali, Z.; O'Hare, W. T.; Rowell, F. J. Total luminescence spectroscopy with pattern recognition for classification of edible oils. *Analyst* **2003**, *128*, 966−973.

(18) Papadopoulos, K.; Triantis, T.; Tzikis, C. H.; Nikokavoura, A.; Dimotikali, D. Investigations of the adulteration of extra virgin olive oils with seed oils using their weak chemiluminescence. *Anal. Chim. Acta* **2002**, *464*, 135−140.

(19) Massart, D. L.; Kaufman, L. Hierarchical clustering methods. In *The Interpretation of Analytical Chemical Data by the Use of Cluster Analysis*; Wiley: New York, 1983; pp 75−99.

(20) Vandeginste, B. G. M.; Massart, D. L.; Buydens, L. M. C.; de Jong, S.; Lewi, P. J.; Smeyers-Verbeke, J. Cluster analysis. In *Handbook of Chemometrics and Qualimetrics*; Elsevier: Amsterdam, The Netherlands, 1998; Vol. 20B, Part B, pp 57−86.

(21) Downey, G.; McIntyre, P.; Davies, A. N. Geographic classification of extra virgin olive oils from the eastern Mediterranean by chemometric analysis of visible and near-infrared spectroscopic data. *Appl. Spectrosc.* **2003**, *57*, 158−163.

(22) Angerosa, F.; Mostallino, R.; Basti, C.; Vito, R.; Serraiocco, A. Virgin olive oil differentiation in relation to extraction methodologies. *J. Sci. Food Agric*. **2000**, *80*, 2190−2195.
(23) SLM AMINCO, Technical Note 101, Urbana, Italy.
(24) Lakowicz, J. R. Instrumentation for fluorescence spectroscopy. In *Principles of Fluorescence Spectroscopy*, 2nd ed.; Kluwer Academic/Plenum Publishers: New York, 1999; pp 25−61.
(25) Matlab. The Mathworks, South Natick, MA; http://www.mathworks.com; last access July 17, 2003.
(26) http://www.camo.com/rt/products/uns/, Website of the CAMO Co.; last access Dec 18, 2003.
(27) Beebe, K. R.; Pell, R. J.; Seasholtz, M. B. Preprocessing. In *Chemometrics. A Practical Guide*; Wiley: New York, 1998; pp 26−55.
(28) Guimet, F.; Ferré, J.; Boqué, R.; Rius, F. X. Application of unfold principal component analysis and parallel factor analysis to the exploratory analysis of olive oils by means of excitation−emission matrix fluorescence spectroscopy. *Anal. Chim. Acta* **2004**, *515*, 75−85.
(29) Bro, R. Validation. In Multi-way Analysis in the Food Industry. Models, Algorithms and Applications. Ph.D. Thesis, University of Amsterdam, Amsterdam, The Netherlands, 1998; pp 99−134.
(30) Halkidi, M.; Batistakis, Y. B.; Vazirgiannis, M. On clustering validation techniques. *J. Intell. Informat. Syst*. **2001**, *17*, 107−145.

Received for review April 5, 2004. Revised manuscript received August 4, 2004. Accepted August 6, 2004. We thank the Spanish Ministry of Science and Technology (Project BQU2003-01142) for financial support and the Rovira i Virgili University for a doctoral fellowship.

JF040169M