JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Discrimination of Olives According to Fruit Quality Using Fourier Transform Raman Spectroscopy and Pattern Recognition Techniques

BARBARA MUIK,[†,§] BERNHARD LENDL,[§] ANTONIO MOLINA-DÍAZ,[†] DOMINGO ORTEGA-CALDERÓN,[#] AND MARÍA JOSÉ AYORA-CAÑADA*[,†]

Department of Physical and Analytical Chemistry, University of Jaén, Paraje las Lagunillas s/n, E-23071 Jaén, Spain; Institute of Chemical Technologies and Analytics, Vienna University of Technology, Getreidemarkt 9/164, A-1060 Wien, Austria; and CIFA Venta del Llano, IFAPA, Ctra. Bailén-Motril km 18.5, E-23620 Mengíbar, Jaén, Spain

Fourier transform Raman spectroscopy combined with pattern recognition has been used to discriminate olives of different qualities. They included samples of sound olives, olives with frostbite, olives that have been collected from the ground, fermented olives, and olive samples with diseases. Milled olives were measured in a dedicated sample cup, which was rotated during spectrum acquisition. A preliminary study of the data set structure was performed using hierarchical cluster analysis and principal component analysis. Two supervised pattern recognition techniques, *K*-nearest neighbors and soft independent modeling of class analogy (SIMCA), were tested using a "leave-a-fourth-out" cross-validation procedure. SIMCA provided the best results, with prediction abilities of 95% for sound, 93% for frostbite, 96% for ground, and 92% for fermented olives. The olive samples with diseases (too few to define a class) were included in the validation and recognized as not belonging to any class. None of the damaged olive samples was wrongly predicted to the class of sound olives. With this approach a selection of sound olives for the production of high-quality virgin olive oil can be achieved.

KEYWORDS: Fourier transform Raman spectroscopy; pattern recognition; olives; olive oil quality

INTRODUCTION

One of the essential characteristics that differentiates virgin olive oil from other vegetable oils is that it is edible at the moment of production, because solely mechanical or other physical means are used. The fact that the oil extraction is solvent-free and natural antioxidants are maintained in the oil is reflected in the higher nutritional and economic value of this product. However, virgin olive oil comes in different grades. According to quality parameters such as low acidity (<0.8%) and good organoleptic characteristics, which are established by the International Olive Oil Council, only the highest quality oil can be sold as "extra virgin". Poor quality oil, called *lampante*, is not directly edible and needs refinement prior to consumption (*1*).

The quality of olive oil depends on various factors, which are all part of the entire production cycle. Under ideal conditions, sound, ripe fruits give extra virgin olive oil. However, during the growth phase, diseases, pests, and freeze damage can lead to decreased olive oil quality. Among them, freeze damage (frostbite) is found to affect the oil quality less. In a recent study the quality indices of the oil were not found to be affected, but lower oil stability and sensorial changes were observed (*2*). Diseases and pests always lead to decreased olive oil quality, generally reflected in high acidity values. During harvest, the collection of overripe olives, which have already fallen to the ground, is the major cause of poor quality (*3*). The fact that olives lie on the ground for several days leads to the deterioration of the fruit flesh with inevitable negative consequences for the quality of the oil that is subsequently produced. The action of microbial lipases increases acidity, and the contact with the soil leads to off-flavors in the oil. The contents of several components that give the olive oil its aroma, as well as the content of antioxidants, such as polyphenols, decrease. Finally, storing of olives prior to processing also deteriorates the quality of the extracted oil. Acidity and peroxide values in the oil increase. Furthermore, when olives are stored in huge piles, the drupes get crushed and consequently liquid seeps out. Such conditions favor enzymic activities that lead to defects and off-flavors in the oil. Anaerobic fermentative processes cause an increase in other products such as alcohols. These alcohols, together with

* Corresponding author (telephone ++34953212259; fax ++34953212141; e-mail mjayora@ujaen.es).
[†] University of Jaén.
[§] Vienna University of Technology.
[#] IFAPA.

10.1021/jf049240e CCC: $27.50
Published on Web 09/09/2004

n-octane and lactic acid, are responsible for the defective "fusty" note in the organoleptic evaluation (*4*).

The quality of the produced oil will strongly depend on the quality of the olives. Therefore, in a quality-oriented production plant, olives with defects should be separated from sound fruits, because a very small portion of bad fruits can ruin the whole batch. At the moment, this separation is done only in some production plants, by visual inspection or communication with the farmer. This is not very reliable, but measuring conventional analytical parameters used to characterize the oil, such as acidity, would be far too time-consuming.

To our best knowledge there is no work reported that deals with the separation of olives in different classes according to their condition. On the contrary, the classification of olive oil has been approached, mainly focusing on geographical origin (*5*, *6*), cultivar (*7*), varietyies (*8*), and adulteration (*9*). Only two works (*10*, *11*) treated the discrimination of the different olive oils according to their quality, using metal oxide sensors and mass spectrometry, respectively.

Fourier transform (FT) Raman spectroscopy is a very promising tool in process analytical chemistry because many samples can be examined nondestructively in a short time without any sample preparation. The sample is irradiated by a near-infrared laser beam. Frequency shifts of the scattered light can be analyzed and presented as spectra. Spectral bands represent vibrations characteristic for chemical bonds and structural units within the molecules of the examined sample. In combination with chemometric data evaluation FT-Raman spectroscopy is a powerful tool capable of extracting chemical information also from complex matrices. In this context FT-Raman spectroscopy was applied in different classification tasks to matrices such as rice (*12*), wood types (*13*), corals (*14*), honey (*15*), and epithelial cancers (*16*).

In previous works we reported applications of this technique to olives in the quantitative analysis of humidity, oil content (*17*), and free fatty acid content (*18*). In the present work FT-Raman spectroscopy, combined with multivariate data analysis, was applied to olive analysis with the aim of discriminating among olives of different quality. The capability of this technique was studied in the separation of sound olives, olives with freeze damage, olives with diseases (*Verticillium* wilt and anthracnose), olives collected from the ground, and fermented olives. The pattern recognition techniques hierarchical cluster analysis (HCA), principal component analysis (PCA), *K*-nearest neighbor (KNN), and soft independent modeling of class analogy (SIMCA) were used.

EXPERIMENTAL PROCEDURES

Olive Samples. One hundred and forty-six olive samples were collected in the period January−March 2004 in the olive culture and process research station at CIFA Venta del Llano (Jaén, Spain). They included 59 samples of sound olives (of 8 different varieties) collected from the tree, 28 of olives with frostbite, 7 samples with diseases (*Verticillium* wilt and anthracnose), 26 fermented olive samples, and 26 samples of olives collected from the ground. The details of the olive samples are shown in **Table 1**. Throughout this paper samples will be referred to as belonging to the following five classes: sound, frostbite, disease, fermented, and ground.

Table 1. Characteristics of the Olive Samples

variety	sound	disease	frostbite	ground	fermented
Picual	11	3	4	8	unknown
Arbequina	10		4	6	unknown
Frantoio	10		4		unknown
Cornicabra	7	1	4	6	unknown
Picudo	6		4	6	unknown
Picholine	6		4		unknown
Picholine Marrocaine	6		4		unknown
Blanqueta	2	2			unknown
Razzola	1	1			unknown
sum	59	7	28	26	26

Spectra Acquisition. A Bruker RFS (FT)-Raman spectrometer fitted with a liquid nitrogen-cooled Ge detector was used to record the Raman spectra. The radiation of excitation was the 1064 nm line from a Nd:YAG laser (coherent). The Raman scattered radiation was collected at 180° geometry. The samples were investigated in a homemade cell. This consisted of a hollow cylindrical magnet, which was covered inside with a Teflon layer. The dimensions of the cell were 13 mm inner diameter and 5 mm depth. The cell was attached to a synchronous motor mounted in the sample compartment. The motor rotated the sample cell around the horizontal axis of the laser beam at 5 rounds per minute. In this way, a circumference of 3 mm radius was illuminated by the laser instead of a spot of few micrometers, thus compensating for sample heterogeneity. All spectra were obtained with 500 mW laser power, at a resolution of 4 cm^{-1}, and were the average of 300 scans, resulting in an acquisition time of 8 min. No sample pretreatment except milling using a hammer mill was necessary.

Data Analysis. *Data Pretreatment.* Data pretreatment by vector normalization of Raman spectra was done in the range of 2880−2950 cm^{-1}. For data analysis the spectral ranges of 300−1800 and 2600−3500 cm^{-1} were used. No baseline correction and smoothing were needed.

PCA, HCA, and SIMCA were performed using the software packages PLS_toolbox (Eigenvector Research, Inc., Manson, WA) and Statistics Toolbox [version 2.2 (R11), 1998] for Matlab (The Mathswork, Inc., Natick, MA). KNN analysis was performed using the software package V-Parvus-2003 (M. Forina, Genoa, Italy).

HCA is an unsupervised classification procedure that involves a measurement of the similarity between objects based on their measured properties (variables). Objects are grouped in clusters in terms of their nearness in the multidimensional space. In this work the distance matrix was calculated using Euclidean distances. From the distance matrices the dendrograms were created using the Ward algorithm. This agglomeration method is generally recommended for large data sets. The elements or clusters are joined with the criterion that the sum of heterogeneities of all clusters should increase as little as possible

PCA transforms the original data matrix into a product of two matrices: the scores matrix, which contains the information about the objects, and the loadings matrix, which contains the information about the variables. It is an unsupervised technique that reduces the dimensionality of the original data matrix retaining the maximum amount of variability. It allows the relationship between variables and observations to be studied, as well as recognizes the data structure.

KNN, a supervised classification method, is based on the distance of the objects in the multidimensional space. To classify an unknown, the distance is calculated between it and a set of samples with known class membership. The predicted class of an unknown sample is then assigned as the class of the *K* samples nearest to it.

SIMCA, another supervised classification method, constructs a model for each class independently by PCA. The number of significant components is defined by means of a cross-validation procedure. In the original SIMCA formulation developed by Wold (*19*), the distance of a point from a class was determined by the out-of-space distance, that is, by the Euclidean distance of the point from the subspace spanned by the *k* principal components used to model the class. In this work we used a more advanced formulation in which the SIMCA distance includes an in-space distance, as well as an out-of-space distance (*20*). The in-space distance is a measure of how well the projection of the point into the principal component subspace agrees with the projections of the known class data. These two distances can be calculated for the class d_0^2 and for the objects d_i^2. For objects belonging to the class, the ratio $F = d_i^2/d_0^2$ then follows an *F* distribution. This results in a critical distance F_{crit} with probabilistic meaning (in this work a level of

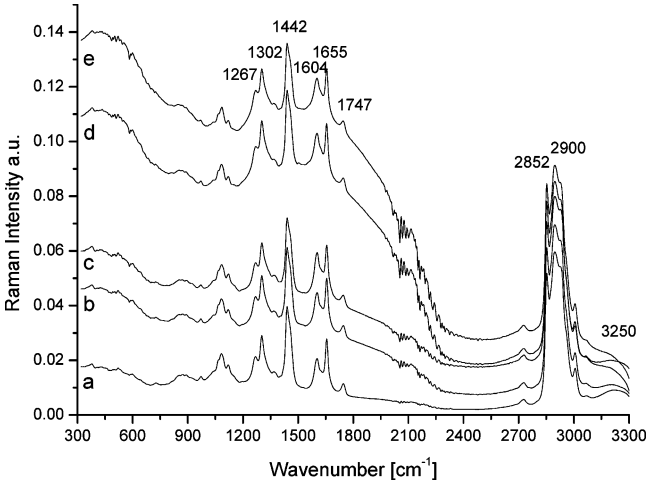

Figure 1. Raman spectra of the milled olives: (**a**) mean sound; (**b**) mean frostbite; (**c**) mean disease; (**d**) mean fermented; (**e**) mean ground.

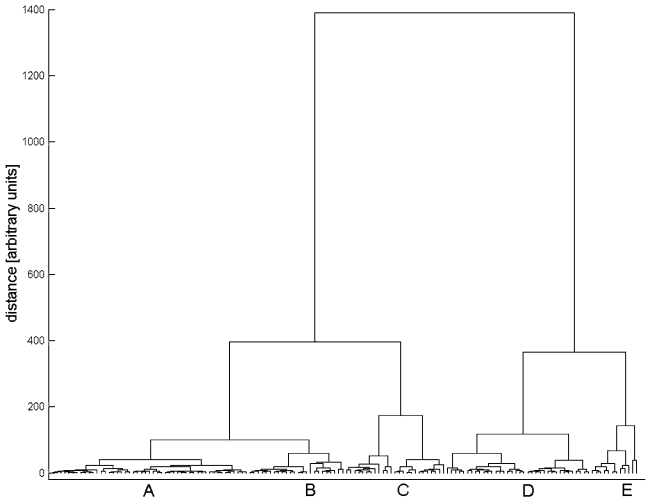

Figure 2. Dendrogram of the mean-centered and normalized data using Euclidean distances and the Ward algorithm.

significance of 95%). The in-space and out-of-space distances are then combined, and the unknown test point is assigned to one or more classes if it falls into the statistical limits.

RESULTS AND DISCUSSION

Raman Spectra of Milled Olives. Figure 1 shows characteristic FT-Raman spectra of milled olives in the region 300−3300 cm^{-1}. Presented are the mean spectra of the corresponding classes sound, frostbite, disease, ground, and fermented. Generally, the FT-Raman spectra of the milled olives present a series of bands with various Raman scattering intensities and shapes. The major bands of the pure virgin olive oil at 1267 cm^{-1} [in-plane δ(=C−H) deformation in unconjugated cis double bond], 1302 cm^{-1} (in-phase methylene twisting motion), 1442 cm^{-1} [$\delta(CH_2)$], 1655 cm^{-1} [ν(C=C)], 1747 cm^{-1} [ν(C=O)], 2852 cm^{-1} [ν_{sym} (CH_2)], and 2900 cm^{-1} [ν_{sym} (CH_3)] dominate the spectrum. A broad band centered at 3250 cm^{-1} corresponds to hydrogen-bonded OH vibration of water. Furthermore, bands from the different parts of the olive fruit contribute to the spectra; the band at 1604 cm^{-1} can be assigned to the aromatic ring stretch of lignin, which is a major compound of the olive kernel. A detailed discussion of the different features in the olive spectra and a tentative band assignment can be found elsewhere (*17*).

Visually, the most noticeable difference between the spectra of the different class-means is the fluorescence background, which is nearly absent in the sound class and increases with the damages the olive suffers. This is probably caused by destruction of cell membranes and alterations in the vegetable matter, such as oxidation and enzymic reactions. Apart from this, there are various relative intensity variations, minor shifting in band positions, and small changes in the spectral contours between classes.

Unsupervised Pattern Recognition. The goal of unsupervised pattern recognition is to evaluate whether clustering exists in a data set without using class membership information in the calculations. Even if the class memberships of the samples are a priori known, a preliminary study based on unsupervised pattern recognition methods, such as HCA and PCA, is advisable to characterize the structure of the data set.

HCA was applied to the mean-centered and normalized data set. Results are shown as a dendrogram plot in **Figure 2**. Five major clusters can be identified. Cluster A contains the majority (45 samples) of the samples from the sound class and 4 samples of the frostbite class. The first part of cluster B contains the rest (14 samples) of the sound class, the second part samples from the frostbite class (6 samples), and some of the disease class (4 samples). In clusters A and B the samples are generally closer to each other than in the other clusters, which is expressed as a smaller distance in the *y*-axis and indicates a higher similarity. Cluster C consists mainly of samples from the frostbite class (18 samples) mixed with some samples from the other classes (2 disease, 2 ground, and 3 fermented). Cluster D is a mixed cluster between the ground (12 samples) and fermented (23 samples) classes and contains the last sample from the disease class. Finally, cluster E consists solely of the remaining 12 samples from the ground class. The dendrogram shows that the majority of the samples from the classes ground and fermented are clearly separated from the rest of the classes. Also, the class sound is well separated, but slightly overlaps with the class frostbite.

PCA. As a first step PCA on the mean-centered data without spectral pretreatment was performed. Whereas scores plots facilitate visualization of the clustering of samples, loadings plots reveal the underlying variances by showing the influence of the different spectral regions on the principal components. An inspection of the loadings plots showed that the first principal component (PC) was mainly influenced by the fluorescence background and the second PC captured variation of the oil content (not shown). As the variation in oil content did not contribute to the class separation, which was revealed in the inspection of the corresponding scores plot, it was decided to perform a vector normalization in the range of 2880−2950 cm^{-1}, where the major oil bands can be found, to compensate for differences in oil content in the samples.

After this normalization, a second PCA showed that the data pretreatment successfully diminished this variation. The inspection of the second PC showed that it contained variance due to vegetable matter, mainly the kernel. This variance did not contribute to the class separation, but it was not diminished with any data pretreatment. **Figure 3a** shows the two-dimensional plot of the samples in the space defined by the first and third PCs of the normalized data. The class sound forms a tight cluster, which is only slightly overlapped by samples from the frostbite and disease classes. The samples from the frostbite class form a cluster between sound and the other two classes, ground and fermented, which are overlapped. A further class separation, especially between ground and fermented, is achieved

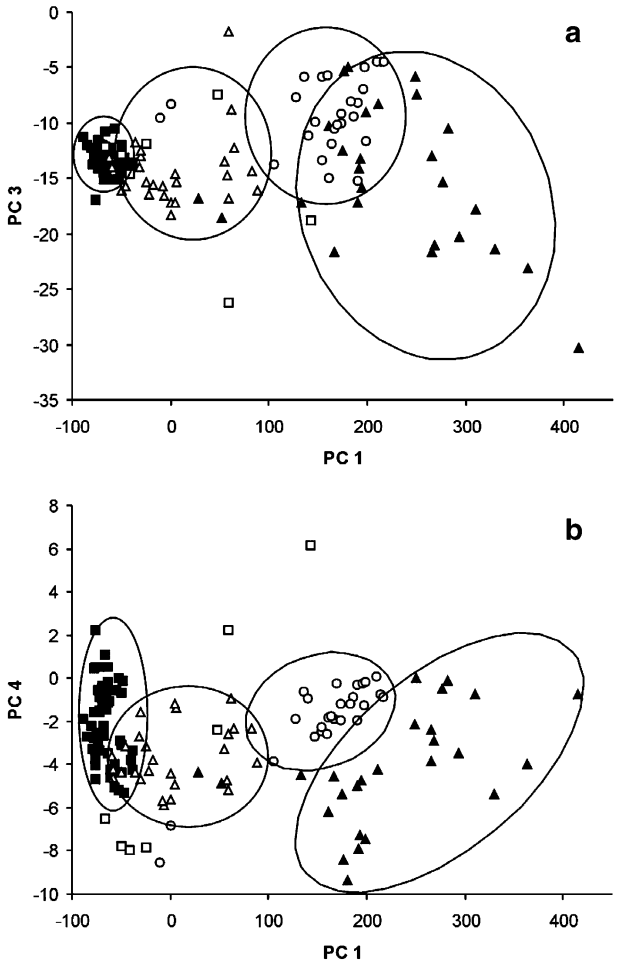

Figure 3. Samples in the space defined by the (**a**) first and third PCs and by the (**b**) first and fourth PCs of the mean-centered and normalized data: (■) sound, (△) frostbite, (○) ground, (▲) fermented, and (□) disease olives.

by taking into account the fourth PC (**Figure 3b**) and the fifth PC (not shown).

The preliminary study of the data set, using HCA and PCA, led to the following results.

Normalizing the data helps to diminish the variation of oil content in the samples and gives therefore a better class separation. The classes sound, frostbite, ground, and fermented form clusters, which partly overlap, but the samples of the class disease do not form a group. The last is not surprising, because the olive samples suffered from two diseases, and there were too few samples to identify grouping. The samples from class sound form a tighter group than the other classes. The largest group is formed by the samples from class ground. Due to the manner in which these olives were collected vegetable matter and soil can be incorporated. This leads, together with the different alteration processes that olives suffer on the ground, to more diversity.

In addition to the knowledge about the structure residing in the data set, two outliers have been detected, one of the class sound and the other of the class ground, and removed. Finally, it can be stated that the natural groupings among the Raman spectra of the samples coincide largely with the known class memberships.

Supervised Pattern Recognition. Because of the promising results of the unsupervised techniques, the supervised techniques, KNN and SIMCA, were applied to the Raman spectra. These classification techniques use a training data set to establish classification rules that then allow samples of unknown origin to be classified. The goodness of the classification rules needs to be validated. For this a "leave-a-fourth-out" cross-validation procedure was applied. Three-fourths of the samples were assigned to the training set and the other fourth to the test set. This assignment was repeated four times, so that each sample was predicted once. All classification results presented in this section refer to the validation results. The reliability of the classification models was studied in terms of prediction ability, which is characterized by the percentage of the test set members adequately classified by the rules developed in the training step. Further attention was paid to the classification errors, which were divided into false positive, that is, how many samples have been wrongly classified to class *x*, and false negative, that is, how many samples of class *x* have been wrongly assigned to other classes.

Table 2. Validation Results of KNN Classification (Prediction Ability Is Given in Percentage)

	sound	frostbite	ground	fermented	disease	false neg
sound	58 (100%)					0/58
frostbite	2	26 (93%)				2/28
ground		2	16 (64%)	7		9/25
fermented		3	1	22 (85%)		4/26
disease	1	4		1	1 (14%)	6/7
false pos	3/86	9/116	1/119	8/118	0/137	

After elimination of the outliers, the final data set consisted of a total of 144 samples: 58 samples of the class sound, 26 of the class frostbite, 25 of the class ground, 26 of the class fermented, and 7 of the class disease. Data pretreatment consisted of normalizing and mean-centering.

KNN was applied to the data set using the decision criteria of three nearest neighbors and a majority vote. Results are summarized in **Table 2**. The prediction abilities were good for the classes sound (100%) and frostbite (93%), less accurate for the classes ground (64%) and fermented (85%), and very poor for the class disease (14%). The seven ground samples, which have been classified in the fermented class, indicate that this classification technique is not capable of separating these two classes well. In the case of the seven disease samples, only one was correctly classified, four were classified as frostbite, and one each as ground and sound. A detailed inspection of the results showed that distances of these samples to the nearest neighbors were longer than the distances for the other samples. This is in agreement with the results from the unsupervised techniques and indicates that these samples do not form a class. Finally, although all samples of the class sound have been correctly classified, three samples of other classes gave false positive responses. This means that 3.5% of the altered olive samples were accepted as sound.

SIMCA constructs a model for each class independently by principal component analysis. In the PCA step of the SIMCA model development a sufficient number of samples is necessary to obtain a good representation of the class. The samples with diseases were too few to define a particular class. Furthermore, they showed no tendency to form a group, as seen in the unsupervised techniques and confirmed using KNN. Nevertheless, they were included in the validation procedure to check if they would be wrongly classified in any of the classes or recognized as not belonging to any class.

The numbers of PCs used for the class models were seven for the class sound and five each for the classes frostbite, ground, and fermented. Results, which are presented in **Table 3**, show good prediction abilities for all classes, with 95, 93, 96, and

Table 3. Validation Results of SIMCA Classification (Prediction Ability Is Given in Percentage)

	sound	frostbite	ground	fermented	false neg
sound	55 (95%)	3			3/58
frostbite		26 (93%)	2		2/28
ground		1	24 (96%)		1/25
fermented		1	1	24 (92%)	2/26
disease					
false pos	0/86	5/116	3/119	0/118	

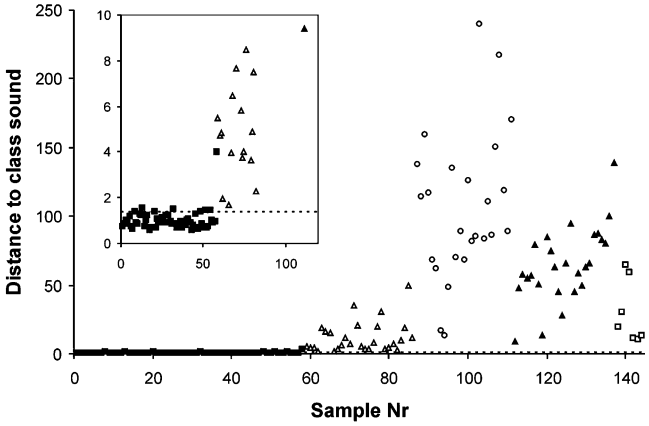

Figure 4. SIMCA distances of all samples to the class sound model: (■) sound, (△) frostbite, (○) ground, (▲) fermented, and (□) disease olives. The horizontal dashed line marks the statistical limit.

92% for the classes sound, frostbite, ground, and fermented, respectively. In contrast to KNN, a very good separation between the classes ground and fermented was achieved. There were also few cases of double classification of samples. Eight sound samples were additionally classified to the frostbite class and three frostbite samples additionally to the ground class. This result indicates a certain overlap between the models of these classes, but is not worrying because the final decision based on the distance was for all of these samples their correct class. One of the advantages of SIMCA is that it is able to recognize when a sample does not belong to any of the established classes. In our case, the seven disease samples were correctly rejected by all class models. They were found to be closest to frostbite or ground, but far outside the statistical limits.

The distances of the damaged olive samples to the sound class model were sufficiently long, so that no sample was falsely classified as sound. All samples with their corresponding distances to the class sound model are shown in **Figure 4**. A zoom to the distances between 0 and 10 is inserted in this figure, for a better visualization of the samples that fall close to the statistical limit (i.e., the critical distance). The statistical limit is marked with a horizontal dashed line. Only one sound sample is found clearly outside this limit. Samples from other classes, which can be found close, are from the class frostbite and from the class fermented. No samples with diseases and from the class ground are situated close.

According to the obtained results SIMCA was shown to be the more appropriate classification technique for this task. The application of SIMCA to the Raman spectra of the milled olives gave prediction abilities >90% for all defined classes. It should be emphasized that with this classification technique no sample was falsely classified in the class sound. This selectivity of the sound class model is especially important if the discrimination is done to improve the production of high-quality virgin olive oil.

The presented work shows the possible implementation of Raman spectroscopy in the olive oil production process. This measurement works directly on the olives and therefore enables a screening of olives according to their quality in order to optimize the production of high-quality virgin olive oil.

ACKNOWLEDGMENT

We thank Dr. J. A. García Mesa for providing the olive samples.

LITERATURE CITED

(1) IOOC. *Trade Standard Applying to Olive Oil and Olive Pomace Oil*; COI/T15 3; International Olive Oil Council: Madrid, Spain, 2003.

(2) Morelló, J. R.; Motilva, M. J.; Ramo, T.; Romero, M. P. Effect of freeze injuries in olive fruit on virgin olive oil composition. *Food Chem.* **2003**, *81*, 547−553.

(3) Uceda, M.; Hermoso, M. La calidad del aceite de oliva. In *El Cultivo del Olivo*, 2nd ed.; Barranco, R., Fernández-Escobar, Rallo, L., Eds.; Junta de Andalucia, Grupo Mundi-Prensa: Madrid, Spain, 1997; pp 562−567.

(4) Boskou, D. *Olive Oil: Chemistry and Technology*; AOCS Press: Champaign, IL, 1996; pp 14−19.

(5) Vlahov, G.; Del Re, P.; Simone, N. Determination of geographical origin of olive oils using 13C nuclear magnetic resonance spectroscopy. I—Classification of olive oils of the Puglia region with denomination of protected origin. *J. Agric. Food Chem.* **2003**, *51*, 5612−5615.

(6) Tapp, H. S.; Defernez, M.; Kemsley, E. K. FTIR spectroscopy and multivariate analysis can distinguish the geographic origin of extra virgin olive oils. *J. Agric. Food Chem.* **2003**, *51*, 6110−6115.

(7) Brescia, M. A.; Alviti, G.; Liuzzi, V.; Sacco, A. Chemometric classification of olive cultivars based on compositional data of oils. *J. Am. Oil Chem. Soc.* **2003**, *80*, 945−950.

(8) Giansante, L.; Di Vincenzo, D.; Bianchi, G. Classification of monovarietal Italian olive oils by unsupervised (PCA) and supervised (LDA) chemometrics. *J. Sci. Food Agric.* **2003**, *83*, 905−911.

(9) Vigli, G.; Philippidis, A.; Spyros, A.; Dais, P. Classification of edible oils by employing ^{31}P and ^{1}H NMR spectroscopy in combination with multivariate statistical analysis. A proposal for the detection of seed oil adulteration in virgin olive oils. *J. Agric. Food Chem.* **2003**, *51*, 5715−5722.

(10) Pena, F.; Cardenas, S.; Gallego, M.; Valcarcel, M. Characterization of olive oil classes using a chemsensor and pattern recognition techniques. *J. Am. Oil Chem. Soc.* **2002**, *79*, 1103−1108.

(11) Garcia-Gonzales, D. L.; Aparicio, R. Virgen olive oil quality classification combining neural network and MOS sensors. *J.* Agric. Food Chem. **2003**, *51*, 3515−3519.

(12) Ootake, Y.; Kokot, S. Discrimination between glutinous and nonglutinous rice by vibrational spectroscopy. I: Comparison of FT-NIR DRIFT, PAS and Raman spectroscopy. *J. Near Infrared Spectrosc.* **1998**, *6*, 241−249.

(13) Lavine, B. K.; Davidson, C. E.; Moores, A. J.; Griffiths, P. R. Raman spectroscopy and genetic algorithms for the classification of wood types. *Appl. Spectrosc.* **2001**, *55*, 960−966.

(14) Kaczorowska, B.; Hacura, A.; Kupka, T.; Wrzalik, R. Spectroscopic characterization of natural corals. *Anal. Bioanal. Chem* **2003**, *377*, 1032−1037.

(15) Paradkar, M. M.; Irudayaraj, J. Discrimination and classification of beet and cane inverts in honey by FT-Raman spectroscopy. *Food Chem.* **2001**, *76*, 231−239.

(16) Stone, N.; Kendall, C.; Shepherd, N.; Crow, P.; Barr, H. Near-infrared Raman spectroscopy for the classification of epithelial pre-cancers and cancers. *J. Raman Spectrosc.* **2002**, *33*, 564−573.

(17) Muik, B.; Lendl, B.; Molina−Díaz, A.; Ayora-Cañada, M. J. Fourier transform Raman spectrometry for the quantitative analysis of oil content and humidity in olives. *Appl. Spectrosc.* **2003**, *57*, 233−237.

(18) Muik, B.; Lendl, B.; Molina-Díaz, A.; Ayora-Cañada, M. A. Direct, reagent-free determination of free fatty acid content in olive oil and in olives by Fourier transform Raman spectrometry. *Anal. Chim. Acta* **2003**, *487* (2), 211−220.

(19) Wold, S. Pattern recognition by means of disjoint principal components models. *Pattern Recognition* **1976**, *8*, 127−139.

(20) Wold, S.; Sjöström, M. SIMCA: a method for analyzing chemical data in terms of similarity and analogy. In *Chemometrics: Theory and Application*; Kowalski, B. R., Ed.; ACS Symposium Series 52; American Chemical Society: Washington, DC, 1977; pp 243−282.

Received for review May 12, 2004. Revised manuscript received July 28, 2004. Accepted July 29, 2004. B.M. thanks the Spanish Ministerio de Asuntos Exteriores for a Ph.D. fellowship.

JF049240E