

Gozde Gurdeniz
Figen Tokatli
Banu Ozen

Izmir Institute of Technology,
Department of Food Engineering,
Urla-Izmir, Turkey

Differentiation of mixtures of monovarietal olive oils by mid-infrared spectroscopy and chemometrics

Fourier transform infrared (FT-IR) spectroscopy in combination with chemometric techniques has become a useful tool for authenticity determination of extra-virgin olive oils. Spectroscopic analysis of monovarietal extra-virgin olive oils obtained from three different olive cultivars (Erkence, Ayvalik and Nizip) and mixtures (Erkence-Nizip and Ayvalik-Nizip) of monovarietal olive oils was performed with an FT-IR spectrometer equipped with a ZnSe attenuated total reflection sample accessory and a deuterated tri-glycine sulfate detector. Using spectral data, principal component analysis successfully classified each cultivar and differentiated the mixtures from pure monovarietal oils. Quantification of two different monovarietal oil mixtures (2–20%) is achieved using partial least square (PLS) regression models. Correlation coefficients (R^2) of the proposed PLS regression models are 0.94 and 0.96 for the Erkence-Nizip and Ayvalik-Nizip mixtures, respectively. Cross-validation was applied to check the goodness of fit for the PLS regression models, and R^2 of the cross-validation was determined as 0.84 and 0.91, respectively, for the two mixtures.

Keywords: Olive oil, FTIR, adulteration, chemometrics.

1 Introduction

Olive oil has gained popularity in recent years because of its health benefits and sensory qualities, and has become a very important agricultural product for most of the countries of the Mediterranean basin [1]. Extra-virgin olive oil has a highly variable chemical composition. This variability mostly depends on four major factors: environment (climate, soil, and altitude), agronomic practices (irrigation, fertilization), cultivation (harvesting, ripeness, and cultivar) and technological (processing) factors [2]. With increasing consumer demand for high-quality olive oil, oils produced from olives of just one variety (monovarietal) or of one geographical region have appeared on the market. Since authenticity and quality issues can often be associated with a specific region, there has been an increasing interest in the certification of the geographical origin of extra-virgin olive oils. In Europe, certification of extra-virgin olive oils according to the geographical origin is assured by the Denominations of Origin [Protected Denomination of Origin (PDO) and Protected Geographical Indication (PGI)]. These designations were established to protect high-quality agricultural products. PDO labeling means that the product is produced, processed, and prepared within the specified geographical area. On the other hand, PGI designation means that the

product is produced, processed, or prepared in a certain geographical area, and the quality, reputation, or other characteristics are attributable to that area. Even though part of the production process is carried out outside that area, product quality and reputation could still be ascribed to that geographical area. These designations could be used in labeling of olive oils produced in particular geographical regions, with typical characteristics linked to natural factors, to the environment, and to the traditions of these regions [3, 4].

Consequently, there is a need to develop reliable analytical methods for geographical classification and to determine the authenticity of olive oils. There are studies involving the use of chemical and sensorial properties of olive oil in combination with multivariate statistical analysis for the classification of olive oil varieties according to the geographical region. Lanteri et al. [5] was able to discriminate different olive oil varieties using measured parameters (free acidity, peroxide index and fatty acids). Classical analytical parameters such as long-chain fatty alcohols, fatty acids and polycyclic triterpenes were used to determine the varietal origin of olive oil [6]. In addition, it was reported that classification of olive oils could be achieved with an electronic nose combined with neural networks [7]. ^1H NMR fingerprints and ^{13}C NMR also supplied appropriate information for the classification of the geographic origin [1, 8].

Olive oils having PDO or PGI status have a much higher market price; therefore, they might be subjected to adul-

Correspondence: Banu Ozen, Izmir Institute of Technology, Department of Food Engineering, Urla-Izmir, Turkey. Phone: +90-233-7506319, Fax: +90-233-7506196, e-mail: banuozen@iyte.edu.tr

teration. Adulteration of high-priced oils such as olive oil is of concern to the food industry due to economical and safety (toxicological) concerns. Several methodologies have been proposed for monitoring adulteration of virgin olive oils with other edible oils [9–13]. The established limits for fatty acids and triacylglycerols are useful for the detection of adulteration of an olive oil with the other vegetable oils at levels as low as 5%, using chromatographic methods [9]. Also, sterol component determination with solid-phase extraction–gas chromatography was demonstrated as a possible tool for the detection of hazelnut oil in olive-hazelnut oil mixtures [10]. Moreover, an electronic nose based on metal oxide semiconductor sensors was employed for the determination of olive oil adulteration and supplied >95% precision [11]. Pena et al. [12] was able to quantify adulteration of virgin olive oil and olive oil with hazelnut oil (7 and 15% for refined and virgin olive oils, respectively) through direct analysis of oil samples by headspace–mass spectrometry and using various multivariate pattern recognition and regression techniques. The differentiation of virgin olive oil from binary oil mixtures was accomplished by using total synchronous fluorescence spectra [13].

However, some of these analyses are time consuming and labor intensive. An alternative method is Fourier transform infrared (FT-IR) spectroscopy, which is a rapid analysis technique with minimum sample preparation. FT-IR in combination with chemometrics has been successfully used for the detection of adulteration of high-value edible oils with various kinds of other edible oils and for the geographical discrimination of olive oils [14, 15]. Adulteration of dietary supplement oils with less expensive common food oils (down to 2% vol/vol) was detected with the application of FT-IR data to partial least square (PLS) analysis [16]. The presence of sunflower oil (as adulterant) in extra-virgin olive oil can be detected through FT-IR spectroscopy down to a level of 2% [17]. Beaten et al. [18] was able to detect the presence of hazelnut oil in olive oil at low percentages (>8%) using mid-infrared (MIR) spectroscopy. The usefulness of FT-IR for evaluating the adulteration of olive oil with other types of vegetable oils at percentages down to 9% was also indicated [19]. However, we did not encounter any study in the literature that involves the detection of a monovarietal extra-virgin olive oil in a mixture of two different monovarietal oils.

In this study, FT-IR spectroscopy in combination with chemometric techniques is used to discriminate different types of monovarietal olive oils and to differentiate a monovarietal olive oil from a mixture of two different varieties. Moreover, the quantification of a monovarietal olive oil in a mixture of two varieties is performed.

2 Materials and methods

2.1 Oil samples

Two olive cultivars, Erkence (Erk) and Nizip (Niz), were obtained from the Olive Research Institute (Izmir, Turkey), and the Ayvalik (Ayv) cultivar was provided by Olive Nursery (Edremit, Turkey). From each cultivar, 15–25 kg of olives were picked from trees and then milled with a maximum 5-kg capacity laboratory-scale olive oil mill (TEM Spemoliva, Italy). At least two different batches of oil were obtained from each cultivar. The samples were stored in dark brown bottles at a temperature of 8 °C.

For adulteration studies, olive oil mixtures were obtained by mixing Erk and Ayv with Niz. The percentage of Niz added to Erk and Ayv varied between 2 and 20% (vol/vol).

2.2 FT-MIR analysis

All infrared spectra (4000–650 cm^{-1}) were acquired with a Perkin Elmer Spectrum 100 FT-IR spectrometer (Perkin Elmer Inc., Wellesley, MA, USA). This instrument was equipped with a horizontal attenuated total reflectance (HATR) sampling accessory (ZnSe crystal) and a deuterated tri-glycine sulfate (DTGS) detector.

The HATR accessory was used to collect the spectral data of the oils. The resolution was set at 2 cm^{-1} and the number of scans collected for each spectrum was 128. The ZnSe crystal was cleaned with hexane in between sample runs.

2.3 Statistical analysis

Data analysis was performed using multivariate statistical methods with Soft Independent Modeling of Class Analogy (SIMCA) software (Umetrics, Sweden). The 3120–2520 and 1875.5–675 cm^{-1} regions of the FT-IR spectra were used in the analysis. The discrimination of olive oil samples was achieved with principal component analysis (PCA), which is a multivariate projection method designed to extract and display the systematic variation in a data matrix X . It is important to accurately determine the number of components that should be included in the model since it is linked to the difference between the degree of fit and the predictive ability. The degree of fit increases as the number of components increases, but predictive ability does not increase after a certain model complexity. So, it is important to reach an optimal balance between fit and predictive ability. Significant principal components (PC) of each category build the class model, which is computed after a separate scaling for each category. Using the PCA class model, Coomans' plot is con-

structured for the classification of each cultivar (Erk vs. Niz and Ayv vs. Niz) and for the differentiation of mixtures from pure varieties (Erk vs. Erk-Niz mixture and Ayv vs. Ayv-Niz mixture).

Quantification of Niz in Erk-Niz and Ayv-Niz mixtures was performed by PLS regression analysis, which relates the FT-IR absorbance of each monovarietal mixture (*X* block) with the percentages of Niz in that mixture (*Y* block). The spectral regions used for PCA were also used for PLS models. In order to build PLS models, 19 and 17 observations were used for Erk-Niz and Ayv-Niz, respectively.

Due to the low number of samples, the data were not divided into two as the modeling data and the model validation data. Instead, the cross-validation technique was used to assess the model performance. Cross-validation evaluates the data by excluding selected samples in the PLS regression model and then building a model for the remaining samples. The model is tested using the samples excluded from the model and the error values for the predicted observations are calculated. New samples are then excluded from the model set and a new model is built. This procedure is repeated until all samples in the PLS model have been excluded once. After predicting all the observations once by the cross-validation technique, the error values between predicted and calculated response (% of adulterant in this case) were used to calculate two error criteria, the root mean square of error (*RMSE*) and the standard error of prediction (*SEP*).

The equations for *RMSE* and *SEP* are given below [20, 21]:

$$RMSE = 100 \sqrt{\frac{\sum_{i=1}^n (\hat{Y}_i - Y_i)^2}{N}}$$

$$SEP = \frac{100}{\bar{Y}_i} \sqrt{\frac{\sum_{i=1}^n (\hat{Y}_i - Y_i)^2}{N}}$$

where Y_i and \hat{Y}_i are the observed and predicted percentages of Niz in Erk and Ayv, respectively, \bar{Y}_i is its average value, and N is the number of samples.

3 Results and discussion

The first part of this study involves the classification of extra-virgin olive oils from different varieties (monovarietal olive oils). Erk, Ayv and Niz oil cultivars yield high amounts of oil (monovarietal), and the Ayv variety has an important economic potential with its high-quality oil in the Turkish market and is preferred by consumers due to its sensory characteristics. Erk has also a promising economic value.

While Ayv and Erk are widely cultivated in the western part of Turkey, Niz is a cultivar grown in the southeast region of Turkey and has different sensory characteristics than Ayv and Erk. The FT-IR spectra between 4000 and 650 cm^{-1} of all varieties are similar, and the Erk spectrum is represented in Fig. 1. To differentiate these olive oil varieties, the spectrum of Niz is compared with the spectra of Erk and Ayv. Due to the high similarity in their compositions, each olive oil sample looks similar along the entire wavelength range, while differences in the spectra are more visible in studies involving differentiation of other edible oils from olive oil [17, 22]. There is no shift in the positions of the absorption bands, but some changes in the intensity of the bands are observed as shown in Fig. 2. These changes in the intensities are related to the composition of molecular bonds absorbing at those regions. Tapp *et al.* [23] also obtained indistinguishable spectra for olive oils from different countries. The most notable spectral differences are observed in the 3120–2520 and 1875.5–675 cm^{-1} regions. The peaks in the 3050–2800 cm^{-1} region result from C-H stretching vibrations and the large peak around 1740 cm^{-1} is due to C = O double bond stretching vibration. Deformation and bending of C-H and stretching vibration of C-O result in peaks in the 1500–650 cm^{-1} region. A visible difference in the intensities of the peaks around 3006 cm^{-1} is due to the C-H stretching vibration of the *cis*-double bond (=CH). Symmetric and asymmetric stretching vibration of the aliphatic CH_2 and CH_3 groups causes distinction in the band around 2925–2872 cm^{-1} . Shifting of peaks in the band around 1238–1163 cm^{-1} is due to the stretching vibration of C-O ester groups. These differences in the spectra are the result of compositional differences between the olive oil varieties [15, 18].

PCA was performed initially to extract information and to examine the qualitative differences between samples. PCA is a data reconstruction and reduction method, and each PC represents an independent source of spectral variation in the data. Coomans' plot is a useful method to visualize the classification results. It is created by calculating two independent PCA models and plotting the residual distances of samples of each of the two models. Using the 3120–2520 and 1875.5–675 cm^{-1} regions, PCA class models for Erk-Niz and Ayv-Niz with two PCs for both classes were constructed separately (Fig. 3). PCA successfully discriminated Erk and Ayv from Niz. Similarly, Tapp *et al.* [23] was able to classify olive oils from different countries using FT-IR with chemometrics. There are also studies that aim to discriminate different olive oil cultivars using fatty acid compositions [5, 24], and these studies showed that the use of the fatty acid composition was successful in the classification of olive oils with respect to geographic origin and cultivar. Discrimination of olive oils

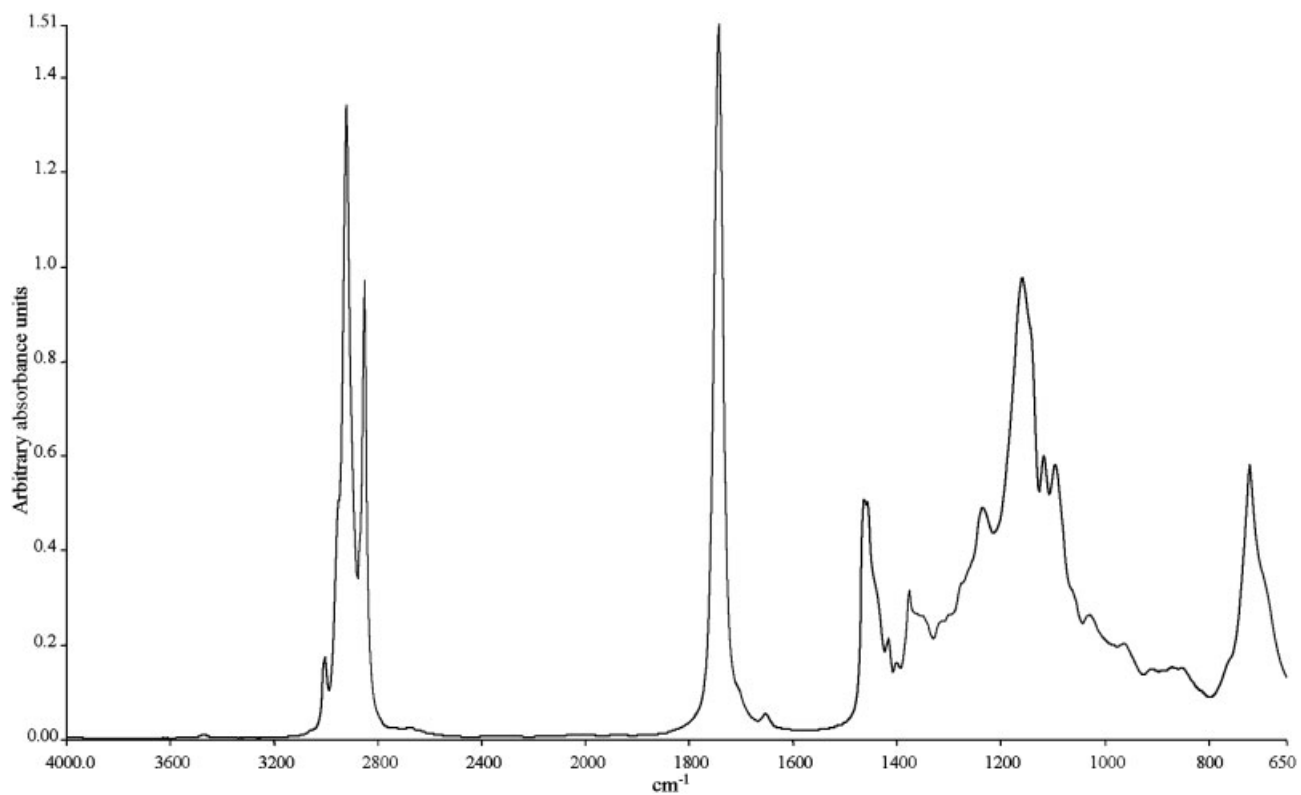


Fig. 1. FT-IR spectrum of the Erk variety between 4000 and 650 cm^{-1} .

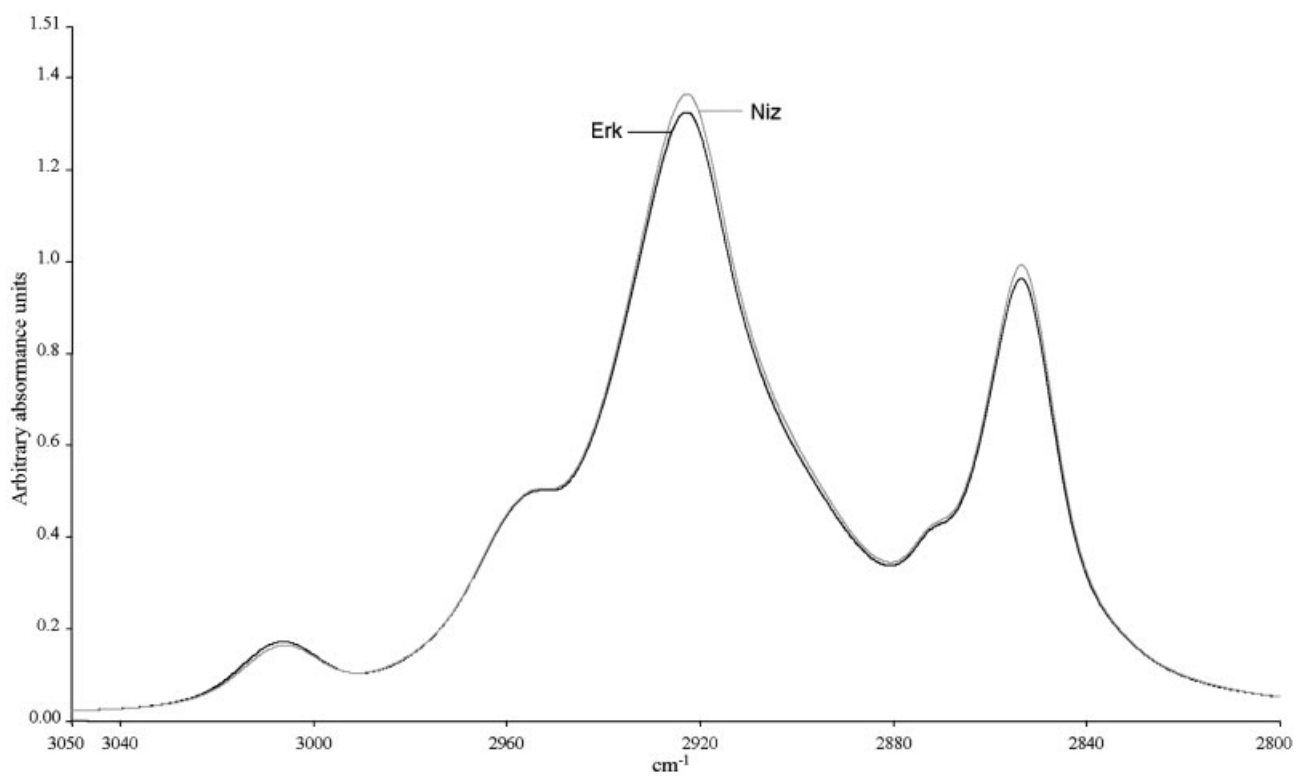


Fig. 2. The spectra of the Erk and Niz varieties around the 3050–2800 cm^{-1} region.

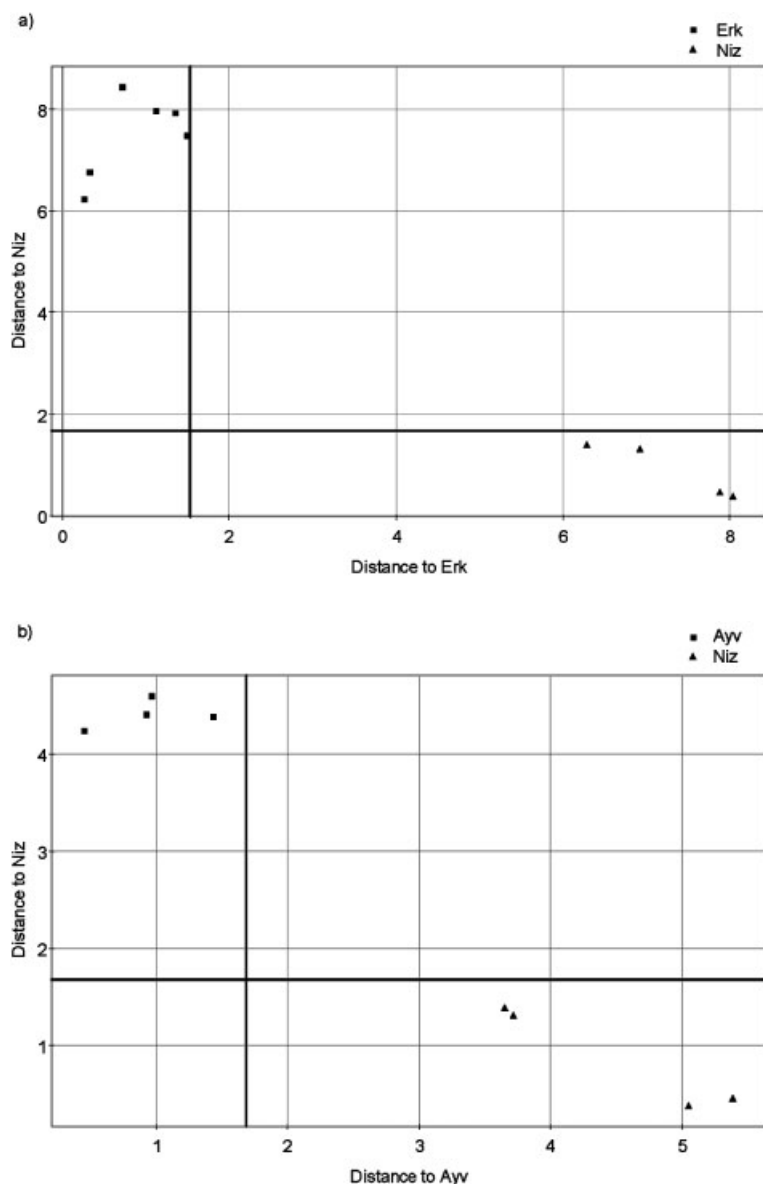


Fig. 3. Coomans' plot for the classification of (a) Erk (two PCs) and Niz (two PCs), (b) Ayv (two PCs) and Niz (two PCs). Vertical and horizontal lines show the 95% confidence interval; the 3120–2520 and 1875.5–675 cm^{-1} regions were used in modeling.

based on varieties has not been studied using FT-IR spectra with chemometrics before. These results indicate that FT-IR could be a useful analysis method for the differentiation of monovarietal olive oils.

After the successful discrimination of monovarietal oils, Erk and Ayv are mixed separately with Niz at various concentrations (2–20% vol/vol) to determine the detection limit of a variety from a mixture of two cultivars using FT-IR. PCA class models and Coomans' plots are constructed both for the Erk-Niz mixtures with pure Erk and also for the Ayv-Niz mixtures with pure Ayv. The differentiation between the Erk and Erk-Niz olive oil mixture is clear, except for one of the samples containing 5% Niz,

which is misclassified as pure Erk. Each pure Ayv sample is accepted by its class model and most of the Ayv-Niz mixtures are classified correctly (Fig. 4).

Quantification of the percentage of Niz in Erk-Niz and Ayv-Niz oil mixtures was performed using a PLS algorithm. The spectral regions used for PCA were also used for PLS models, and four PLS components sufficiently described the model. Fig. 5 shows the concentration values obtained from the PLS model *versus* the actual concentration of Niz in Erk and Ayv. The difference between the actual concentration and the predicted Niz concentration is small, and correlation coefficients, R^2 , for the observed vs. predicted curves are 0.938 for Niz in Erk

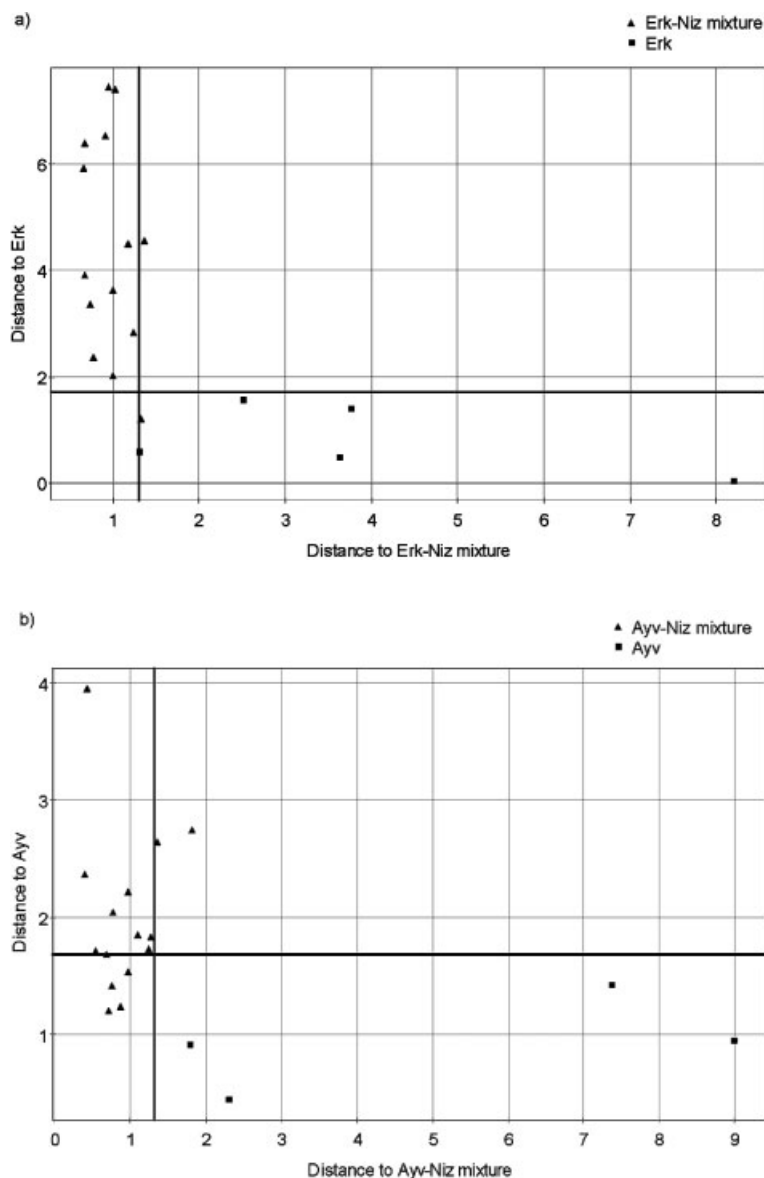


Fig. 4. Coomans' plot for the classification of (a) Erk (three PCs) and Erk-Niz mixture (three PCs), (b) Ayv (two PCs) and Ayv-Niz mixture (four PCs). Vertical and horizontal lines show the 95% confidence interval; the 3120–2520 and 1875.5–675 cm^{-1} regions were used in modeling.

and 0.963 for Niz in Ayv. In order to validate the developed models, cross-validation is applied by removing one sample at a time. The removed sample was predicted with a model created with the remaining observations and the procedure was repeated until each sample was excluded once. Niz content (%) in Erk and Ayv predicted from the PLS model using cross-validation is given in Tab. 1. The performance statistics of cross-validation are shown in Tab. 2, and it summarizes the relationship between the actual and the predicted values modeled. The standard errors of prediction (*SEP*), the root mean square error (*RMSE*), and the correlation coefficient (R^2) parameters were employed to evaluate the goodness of fit of the validation data set, which were calculated with respect to the results of cross-validation. A high coefficient of deter-

mination, R^2 , and low *RMSE* and *SEP* values indicate success of the PLS regression model. The PLS regression model appears to have a reasonable ability to estimate the Niz percentage in both Ayv and Erk, based on the high R^2 , low *RMSE* and *SEP* results (Tab. 2). According to these analyses, the developed FT-IR method and chemometric analysis were useful for quantifying Niz olive oil added to Erk and Ayv at 2–20% vol/vol. IR analysis is generally successful in detecting adulteration of olive oils with other edible oils at levels as low as 2% [14–18]. However there is not any comparable study that involves the discrimination of one variety in olive oil mixtures. Although the use of the fatty acid composition in classification was successful [5, 24], the ability of GC analysis to differentiate olive oil mixtures was not tested.

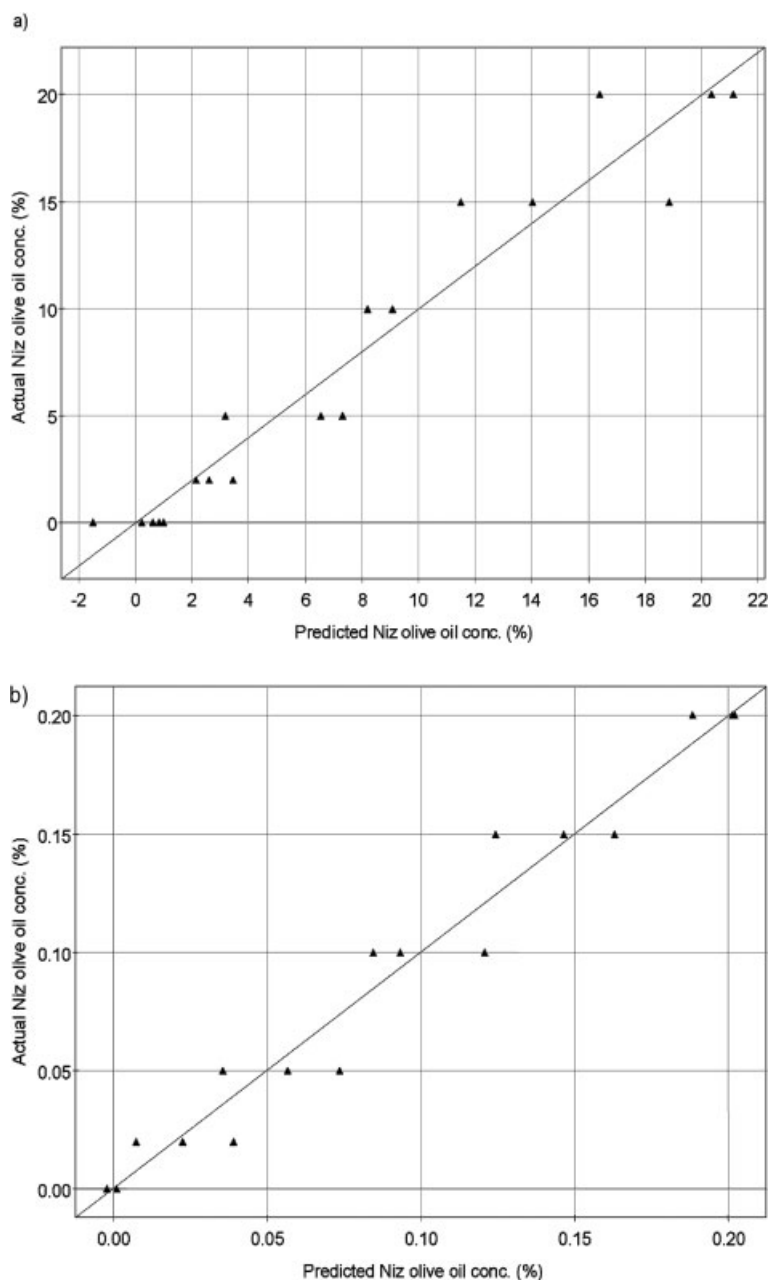


Fig. 5. PLS regression of predicted vs. actual Niz olive oil content in (a) Erk (four PLS components) and (b) Ayv (four PLS components) samples. The 3120–2520 and 1875.5–675 cm^{-1} regions were used in modeling.

In summary, the performance of FT-IR is examined in the identification of olive oils coming from different cultivars and in the differentiation of their mixtures. The presence of a variety in a binary mixture of monovarietal oils could be detected at low levels by manipulating FT-IR data with multivariate statistics. Quantification of one cultivar in a mixture of two cultivars is also accomplished with chemometric techniques. Although three varieties were used in this study, the same methods could be applied to other cultivars. To determine the presence of one monovarietal type in another, calibration curves need to be re-established for each mixture.

Acknowledgments

The authors wish to thank the Olive Research Institute (Izmir) and Olive Nursery (Edremit) for providing olive samples. This study was performed as part of the CODA (MIRG-CT-2005-029134) project supported by an EU Marie Curie Reintegration Grant.

Tab. 1. Actual and predicted (obtained by cross-validation) Niz contents in Erk and Ayv samples using PLS model.

Actual Niz content [%] in Erk	Predicted Niz content [%] in Erk	Actual Niz content [%] in Ayv	Predicted Niz content [%] in Ayv
0	1.47	0	−0.55
0	0.16	0	−0.46
0	0.96	2	4.05
0	1.48	2	3.9
0	−6.14	2	4.53
2	−2.35	5	2.7
2	2.05	5	5.91
2	4.7	5	9.42
5	2.36	10	9.46
5	8.69	10	12.40
5	7.33	10	7.92
10	7.83	15	14.76
10	9.12	15	16.52
15	10.88	15	10.94
15	20.74	20	17.9
15	11.77	20	19.43
20	15.18	20	19.41
20	20.51		
20	21.63		

Tab. 2. Summary of the proposed PLS regression model using cross-validation.

	R^2	RMSE	SEP
Erk-Niz mixture	0.84	3.14	40.85
Ayv-Niz mixture	0.91	2.09	22.83

R^2 , correlation coefficient; RMSE, root mean square error (%); SEP, standard error of prediction.

References

- [1] S. Rezzi, D. E. Axelson, K. Héberger, F. Reniero, C. Mariani, C. Guillou: Classification of olive oils using high throughput flow ^1H NMR fingerprinting with principal component analysis, linear discriminant analysis and probabilistic neural networks. *Anal Chim Acta* 2005, **552**, 13–24.
- [2] R. Aparicio, G. Luna: Characterisation of monovarietal virgin olive oil. *Eur J Lipid Sci Technol.* 2002, **104**, 614–627.
- [3] B. A. Babcock, R. Clemens: Geographical indications and property rights: Protecting value-added agricultural products. *MATRIC Briefing Paper 04-MBP 7*, Iowa State University, Ames, IA (USA) 2004.
- [4] B. Ozen, F. Tokatli, F. Korel: Emerging topics in olive oil research: Determination of geographical origin and adulteration. *Olive Oil and Olive-Pomace Oil Symposium & Exhibition*, Izmir (Turkey) 2005.
- [5] S. Lanteri, C. Armanino, E. Perri, A. Palopoli: Study of oils from Calabrian olive cultivars by chemometric methods. *Food Chem.* 2002, **76**, 501–507.
- [6] L. Giansante, D. Di Vincenzo, G. Bianchi: Classification of monovarietal Italian olive oils by unsupervised (PCA) and supervised (LDA) chemometrics. *J Sci Food Agric.* 2003, **83**, 905–911.
- [7] M. S. Cosio, D. Ballabio, S. Benedetti, C. Gigliotti: Geographical origin and authentication of extra virgin olive oils by an electronic nose in combination with artificial neural networks. *Anal Chim Acta* 2006, **567**, 202–210.
- [8] A. D. Shaw, A. Camillo, G. Vlahov, A. Jonesa, G. Bianchi, J. Rowland, D. B. Kell: Discrimination of the variety and region of origin of extra virgin olive oils using ^{13}C NMR and multivariate calibration with variable reduction. *Anal Chim Acta* 1997, **348**, 357–374.
- [9] E. Christopoulou, M. Lazaraki, M. Komaitis, K. Kaselimis: Effectiveness of determinations of fatty acids and triglycerides for the detection of adulteration of olive oils with vegetable oils. *Food Chem.* 2004, **84**, 463–474.
- [10] L. Cercaci, M. T. Rodriguez-Estrada, G. Lercker: Solid-phase extraction–thin-layer chromatography–gas chromatography method for the detection of hazelnut oil in olive oils by determination of esterified sterols. *J Chromatogr A* 2003, **985**, 211–220.
- [11] M. C. C. Oliveros, J. L. P. Pavón, C. G. Pinto, M. E. F. Laespada, B. M. Cordero, M. Forina: Electronic nose based on metal oxide semiconductor sensors as a fast alternative for the detection of adulteration of virgin olive oils. *Anal Chim Acta* 2002, **459**, 219–228.
- [12] F. Peña, S. Cárdenas, M. Gallego, M. Valcárcel: Direct olive oil authentication: Detection of adulteration of olive oil with hazelnut oil by direct coupling of headspace and mass spectrometry, and multivariate regression techniques. *J Chromatogr A* 2005, **112**, 215–221.
- [13] K. I. Poulli, G. A. Mousdis, C. A. Georgiou: Rapid synchronous fluorescence method for virgin olive oil adulteration assessment. *Food Chem.* 2007, in press.
- [14] A. A. Christy, S. Kasemsumran, Y. Du, Y. Ozaki: The detection and quantification of adulteration in olive oil by near-infrared spectroscopy and chemometrics. *Anal Sci.* 2004, **20**, 935–940.
- [15] G. Downey: Food and food ingredient authentication by mid-infrared spectroscopy and chemometrics. *Trends Anal Chem.* 1998, **17**, 418–424.
- [16] B. F. Ozen, I. Weiss, L. J. Mauer: Dietary supplement oil classification and detection of adulteration using Fourier transform infrared spectroscopy. *J Agric Food Chem.* 2003, **51**, 5871–5876.
- [17] A. Tay, R. K. Singh, S. S. Krishnan, J. P. Gore: Authentication of olive oil adulterated with vegetable oils using Fourier transform infrared spectroscopy. *Lebensm Wiss Technol.* 2002, **35**, 99–103.
- [18] V. Baeten, J. A. F. Pierna, P. Dardenne, M. Meurens, D. L. García-González, R. Aparicio-Ruiz: Detection of the presence of hazelnut oil in olive oil by FT-Raman and FT-MIR spectroscopy. *J Agric Food Chem.* 2005, **53**, 6201–6206.
- [19] N. Vlachos, Y. Skopelitis, M. Psaroudaki, V. Konstantinidou, A. Chatzilazarou, E. Tegou: Application of Fourier transform-infrared spectroscopy to edible oils. *Anal Chim Acta* 2006, **573/574**, 459–465.

- [20] A. O. Aptula, N. G. Jeliaskova, T. W. Schultz, M. T. D. Cronin: The better predictive model: High q^2 for the training set or low root mean square error of prediction for the test set. *QSAR Comb Sci.* 2005, **24**, 385–396.
- [21] C. Hervás, G. Zurera, J. A. Martínez: Optimization of computational neural network for its application in the prediction of microbial growth in foods. *Food Sci Tech Int.* 2001, **7**, 159–163.
- [22] B. F. Ozen, L. J. Mauer: Detection of hazelnut oil adulteration using FT-IR spectroscopy. *J Agric Food Chem.* 2002, **50**, 3898–3901.
- [23] H. S. Tapp, M. Defernez, E. K. Kemsley: FTIR spectroscopy and multivariate analysis can distinguish the geographic origin of extra virgin olive oils. *J Agric Food Chem.* 2003, **51**, 6110–6115.
- [24] M. D'Imperio, G. Dugo, M. Alfa, L. Mannina, A. L. Segre: Statistical analysis on Sicilian oils. *Food Chem.* 2007, **102**, 956–965.

[Received: March 19, 2007; accepted: August 31, 2007]