CHEMOMETRIC STUDIES FOR CLASSIFICATION OF OLIVE OILS AND DETECTION OF ADULTERATION

A Thesis Submitted to
The Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in Food Engineering

by
Gözde GÜRDENİZ

July, 2008

We approve the thesis of **Gözde GÜRDENİZ**

_____

Assist. Prof. Dr. Banu ÖZEN
Supervisor

_____

Assist. Prof. Dr. Figen TOKATLI
Co-Supervisor

_____

Assist. Prof. Dr. Figen KOREL
Co-Supervisor

_____

Assoc. Prof. Dr. Durmuş ÖZDEMİR
Committee Member

_____

Assist. Prof. Dr. Fahri YEMİŞÇİOĞLU
Committee Member

15 July 2008
Date

_____

Prof. Dr. Şebnem HARSA
Head of the Department of
Food Engineering

_____

Prof. Dr. Hasan BÖKE
Dean of the Graduate School of
Engineering and Science

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to my advisor, Assist. Prof. Dr. Banu ÖZEN for his supervision, guidance, support, encouragement and patience during my studies.

I would like to appreciate deeply to my co-supervisors Assist. Prof. Dr. Figen TOKATLI and Assist. Prof. Dr. KOREL for their all kind of support and valuable comments.

I express my special thanks to Olive Research Institute (Izmir, Turkey) and Olive Nursery (Edremit, Turkey) for providing olive samples. Tariş Olive and Olive Oil Agricultural Sales Cooperatives Union in Izmir for providing olive oil samples. This study was performed as a part of CODA project (MIRG-CT-029134) supported by EU Marie Curie Reintegration Grant.

Also, technical support of IZTECH R&D Center for Environmental Studies in this research is gratefully acknowledged.

I would like to thank to my friends Pınar KADİROĞLU, Derya OCAKOĞLU and Gaye YILDIRIM for their help.

ABSTRACT

CHEMOMETRIC STUDIES FOR CLASSIFICATION OF OLIVE OILS AND DETECTION OF ADULTERATION

The aim of this study is to classify extra-virgin olive oils according to variety, geographical origin and harvest year and also to detect and quantify olive oil adulteration. In order to classify extra virgin olive oils, principal component analysis was applied on both fatty acid composition and middle infrared spectra. Spectral data was manipulated with a wavelet function prior to principal component analysis. Results revealed more successful classification of oils according geographical origin and variety using fatty acid composition than spectral data. However, each method has quite good ability to differentiate olive oil samples with respect to harvest year.

Middle infrared spectra of all olive oil samples were related with fatty acid profile and free fatty acidity using partial least square analysis. Orthogonal signal correction and wavelet compression were applied before partial least square analysis. Correlation coefficient and relative error of prediction for oleic acid (highest amount fatty acid) were determined as 0.93 and 1.38, respectively. Also, partial least square regression resulted in 0.85 as R^2 value and 0.085 as standard error of prediction value for free fatty acidity quantification.

In adulteration part, spectral data manipulated with principal component and partial least square analysis, to distinguish adulterated and pure olive oil samples, and to quantify level of adulteration, respectively. The detection limit of monovarietal adulteration varied between 5 and 10% and R^2 value of partial least square was determined as higher than 0.95. Hazelnut, corn-sunflower binary mixture, cottonseed and rapeseed oils can be detected in olive oil at levels higher than 10%, 5%, 5% and 5%, respectively.

ÖZET

ZEYTİNYAĞLARININ SINIFLANDIRILMASI VE TAĞŞİŞİN BELİRLENMESİ İÇİN KEMOMETRİK ÇALIŞMALAR

Bu çalışmada naturel sızma zeytinyağlarının elde edildiği zeytinin türüne, yetiştirildiği coğrafi bölgeye ve hasat yılına göre sınıflandırması ve zeytinyağında tağşişin nitelik ve nicelik yönünden araştırılması amaçlanmıştır. Asal bileşenler analiz yöntemi yağların sınıflandırılması için yağ asitleri kompozisyonuna ve orta bölge kızıl ötesi spektra verilerine uygulanmıştır. Spektral veri, asal bileşenler analizinden önce dalga analizi ile işlenerek sıkıştırılmıştır. Sonuçlar, coğrafi bölgeye ve türe göre sınıflandırmada, yağ asitleri kompozisyonunun spektra verisinden daha başarılı olduğunu göstermektedir. Her iki analitik yöntemin zeytinyağı örneklerini hasat yılına göre ayırma kabiliyeti vardır.

Bütün yağ örneklerinin spektra verileri ile yağ asidi profilleri ve serbest yağ asitliği arasında bağlantı kurmak için kısmi en küçük kareler analizi kullanılmıştır. En küçük kareler analizinden önce, ortogonal sinyal düzeltme filtresi ve dalga analizi sıkıştırma kullanılmıştır. Oleik asitin korrelasyon katsayısı (R^2) ve bağıl hata tahmin değeri sırasıyla 0.93 ve 1.38 olarak belirlenmiştir. Diğer taraftan serbest yağ asitliğinin belirlenmesi için en küçük kareler regresyonunda R^2 değeri 0.85 ve hata tahmin değeri 0.085 olarak belirlenmiştir.

Bu çalışmanın son kısmında, spektral veri asal bileşenler analizi ile işlenerek tağşişin nicelik bakımından tespitinde kullanılmıştır. Yapılan tağşiş miktarının belirlenmesi içinse, spektral veri en küçük kareler analizine tabi tutulmuştur. Tağşiş yapmak için seçilen türe göre, tek tip zeytinyağlarının tağşiş miktarının belirlenmesindeki limit 5% ile 10% arasında değişirken, R^2 değerinin 0.95'ten daha yüksek olduğu bulunmuştur. Diğer taraftan, tağşiş için kullanılan fındık, mısır-ayçiçek ikili karışımı, pamuk tohumu ve kolza tohumu gibi yağların zeytinyağının içinde belirlenebildikleri miktarlar sırasıyla %10, %5, %5 ve %2 olarak tespit edilmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

PDO	Protected Denomination of Origin
PGI	Protected Geographical Indication
PCA	Principal component analysis
FT-IR	Fourier transform infrared
PLS	Partial least square
FFA	Free fatty acidity
EC	European Community
GC	Gas chromatography
IOOC	International Olive Oil Commission
EU	European Union
LDA	Linear discriminant analysis
HS-MS	Headspace mass spectrometry
PDO	Protected Designation of Origin
NIR	Near infrared reflectance
ATR	Attenuated total reflection
GA	Genetic algorithm
GI	Geographical indicators
HPLC	High performance liquid chromatography
ECN42	Equivalent carbon number 42
MIR	Middle infrared
SIMCA	Soft independent modelling of class analogy
OSC	Orthogonal signal correction
SEC	Standard error of calibration
SEP	Standard error of prediction
REP	Relative error of prediction

CHAPTER 1

INTRODUCTION

Olive oil`s recent popularity could be attributed to its sensorial characteristics as well as its potential health benefits. These benefits have been related to its well-balanced fatty acid composition, where oleic acid is the main component, and to the presence of minor biomolecules, such as vitamins and natural antioxidants (Matos, et al. 2007).

Olive oil constitutes various chemical components including triacylglycerols, free fatty acids, phosphatides as the major components and also minor components such as phenolic compounds, hydrocarbons etc. With increasing consumer demand for high quality olive oil, oil produced from olives of just one variety (monovarietal) or one geographical region have been appeared on the market. Therefore, it has become important to characterize each monovarietal olive oil by its chemical and sensorial properties. Chemical composition of olive oils might also differ due to the influence of geographical, agronomic and technological factors (Aparicio and Luna 2002). Differences in composition depending on geographic origin or variety are the basis of the legislations such as Protected Denomination of Origin (PDO) and Protected Geographical Indication (PGI). PDO and PGI certifications allow labelling of food products with growing areas and provide extra economical benefits for producers of designated areas. Consequently, there is a need to develop reliable analytical methods for geographical and varietal classification and adulteration determination of olive oils (Ulberth and Buchgraber 2000, Babcock and Clemens 2004).

To characterize each olive oil variety few series of chemical compounds or a univariate statistics is not adequate. Instead multivariate analysis techniques should be applied to a number of variables (chemical compounds and/or sensory descriptors). The multivariate data analysis enables the extraction of meaningful information from the large amount of data such as chemical and sensorial properties of olive oil (Aparicio and Luna 2002). Multivariate data analysis can be used for both classification and regression issues. It is common to employ principal component analysis (PCA) which shows the relation between observations to classify olive oil with respect to variety or geographical origin.

Chromatographic methods have been generally preferred in classification and adulteration studies. Although chromatographic methods supply high degree of precision, there is an increasing demand for rapid, inexpensive and effective techniques for determination of authenticity of olive oils. Infrared spectroscopy combined with chemometric techniques is one of the promising rapid methods (Downey 1998). FT-IR (fourier transform infrared spectroscopy) is a quite suitable analysis tool for oil and fat analysis because it could be applied directly to samples without any chemical treatment (Bendini, et al. 2007). High number of data generated as a result of IR measurements makes it necessary to use multivariate data analysis tools. Therefore, FT-IR spectroscopy combined with PCA could be performed for varietal and geographical characterization whereas its combination with partial least square (PLS) has been widely used for quantification of adulteration.

In mid-infrared spectra the intensity and the exact frequency at which the maximum absorbance of the bands appears imply differences among complex samples of similar nature (Guillén and Cabo 1999). Fatty acids are one of the major ingredients that olive oil contains in its chemical structure. Thus, a relationship between spectra and the quantity of each individual fatty acid can be introduced with PLS analysis. In addition, a relationship could also exist between spectral data and oil quality parameters such as free fatty acidity (FFA) value.

The aim of this study is to apply chemometric techniques for the classification of extra-virgin olive oil samples according to variety, geographical origin and harvest year using two different data sets (1) fatty acid profile obtained from gas chromatography (GC) analysis and (2) spectral data obtained from FT-IR. Discrimination ability of these two methods was also compared and discussed. Another aim is to detect and quantify olive oil adulteration with other vegetable or seed oils using chemometric techniques. Furthermore, the relation between FT-IR spectra versus fatty acid profile and FFA was studied in accordance with chemometric techniques.

CHAPTER 2

OLIVE OIL

2.1. Brief History

As it is known today, olive tree was domesticated about 6,000 years ago in Mediterranean shores of Syria and Palestine. It expanded to Anatolia via Cyprus and to Egypt via Crete (Luchetti 2002).

Phoenicians were responsible for the spread of the olive tree to western regions, since they traded with other maritime centers. In the 14th and 12th centuries B.C., they introduced it to the Greek mainland. Olive culture reached Spain, Italy and Northern Africa, and then spread into Southern France with the contributions of Greek colonies (Luchetti 2002). Olive trees were planted in the entire Mediterranean basin under Roman rule Romans used oil as a food beside its application as ointment, pharmaceuticals and in lighting (lampante oil). In ancient Egypt, olive plant was cultivated in order to obtain its oil and it was used in religious ceremonies (Harwood and Aparicio 2000).

Major progress in olive processing started with invention of screw press by Greeks. The equipment was improved and disseminated by Romans. The fall of Roman Empire caused a reduction in olive cultivation until Middle Ages. During the 1900s, mechanical extraction systems were emerged as a result of studies on percolation and centrifugation. The Centriolive plant, the first industrial decanter based on the continuous centrifugation of the olive paste, was founded toward the end of 1960s. Despite the improvement in pressing systems, some countries still use the same pressing system today as they did in the past centuries (Harwood and Aparicio 2000).

2.2. World Olive Oil Production

Mediterranean countries which produce 98 percent of world's olive oil are the leaders of olive oil production. European Community (EC) accounts for 79% world's

olive oil production and within EC Spain, Italy and Greece supply more than 98% of EC production.

The world olive oil consumption is also concentrated in the producing countries. European Union (EU) consumes 71% of world's olive oil and Italy, Spain and Greece are responsible from 67% of this consumption. Table 2.1 presents geographical distribution of world olive oil production and consumption.

Table 2.4. Geographical distribution of world olive oil consumption and production (average percentages between crop seasons 2000 and 2006) (Source: International Olive Oil Council (IOOC) 2007)

	Production (%)	Consumption (%)
Spain	39.54	24.0
Italy	24.57	32.1
Greece	14.49	11.1
Tunisia	5.20	1.6
Syria	4.80	4.3
Turkey	4.30	2.0
Morocco	2.20	2.0
France	< 0.2	3.9
others	4.90	19.1

Turkey produces 4.3% of world's olive oil and is sixth biggest world olive oil producer. In world olive oil consumption, Turkey shares also the sixth place in world olive oil consumption with Morocco.

According to IOOC (2007), Spain has produced 1,108,700 t followed by Italy with 591,700 t in 2006/07 crop year. The estimates for Greek olive oil production stand at approximately 370,000 t while the figure for Tunisia comes to around 170,000 t. Turkey is in the fifth place and has produced 166 000 t followed by Syria (154 000 t) and Morocco (75 000 t) (IOOC 2007).

2.3. Production of Olive Oil

2.3.1. Olive Harvesting and Washing

Harvesting generally takes place at the end of the autumn or in the beginning of the winter. The purpose of preliminary washing is to remove any foreign material that could damage machinery or contaminate the oil (Vossen 1998).

2.3.2. Milling and Olive Paste Mixing (Malaxation)

Olive crushing, the first step of olive oil production is employed to produce a paste with easily extractable oil droplets. Two types of machines are used to crush olives: stone mills and stainless steel hammermills.

The olive paste is slowly and continuously mixed to bring small oil droplets in contact with each other to form larger droplets by breaking up the oil/water emulsion. The mixing time could be 20-30 minutes and temperature of the olive paste does not exceed 22-25 ^{0}C (Aparicio and Harwood 2000, Vossen 1998).

2.3.3. Oil Extraction from the Paste

The next step is the extraction of the oil from the paste and fruit water (water of vegetation). Oil can be extracted by pressing, centrifugation, percolation, or through combinations of different methods (Vossen 1997).

2.3.3.1. Traditional Press

Traditional pressing is a discontinuous oil extraction method. In this method the ground olives are pressed in cloth bags then the liquid mix is rested in a series of tanks to separate the oil (Aktas, et al. 2001).

2.3.3.2. Continuous System (by Centrifugation)

The centrifugation method is a continuous or on-line process by which the oil is separated from the solids and water in the same process as in a decantation. The efficiency of the oil extraction increases with savings in time. However, the contact time between the oil and the fermenting fruit water decreases (Vossen 1998).

2.3.3.2.1. Three-phase System Decanter

In the three-phase system decanter, water is added to the system. As the centrifuge rotated at a high speed (3500-3600 rpm), non-miscible liquids (olive oil and vegetation water) are separated by proper nozzles from oil pomace due to specific weights differences This liquid is then taken to a vertical centrifuge where the olive oil is separated from the fruit vegetable water (Harwood and Aparicio 2000).

2.3.3.2.2. Two-phase System Decanter

Two-phase system decanter functions under the same principle as three-phase system decanters but only two phases (oil, sludge) are obtained. If fresh olives are processed, no additional water is required for the separation process. Compared to three-phase decanting, this process is more advantageous. Firstly the throughput rate of produced oil is higher because no additional water is required to produce the pulp; then energy and water consumption is also reduced as a result of the lower processing quantity. Moreover wastewater production decreases considerably (Coputa, et al. 2003).

2.4. Definitions of Olive Oil

Olive oil is a vegetable oil obtained solely from the fruit of the olive tree (*Olea europaea L.*). It is produced by processes that do not alter its natural state (without the use of solvents or re-esterification processes) and does not include other kind of oils (IOOC 2007).

Extra virgin (extra natural leaky) olive oil: The oils that have free fatty acidity, expressed as oleic acid, of not more than 1.0 gram per 100 gram (EU 1991; Turkish

Food Codex 2000). On the other hand, IOOC (2007) defines olive oils with the acidity not more than 0.8 gram per 100 gram as extra virgin olive oil.

Virgin (natural first) olive oil: Free fatty acidity of this class should not be more than 2.0 gram per 100 gram (EU 1991; Turkish Food Codex 2000; IOOC 2007).

Ordinary virgin (natural second) olive oil: Free fatty acidity should not be more than 3.3 gram per 100 gram (EU 1991; Turkish Food Codex 2000; IOOC 2007).

Refined Olive Oil: The oil that is obtained from virgin oil by refining methods without causing any change in triglyceride structure of raw olive oil. The free fatty acidity should not be more than 0.3 gram per 100 gram (Turkish Food Codex, 2000; IOOC 2007). On the other hand, maximum free fatty acidity of refined olive oil is given not to be more than 0.5 gram per 100 gram in EU (European Union) regulations (1991).

Riviera olive oil: The oil that is obtained by mixing refined olive oil with natural olive oil that can directly be consumed as a food. The free fatty acidity should not be more than 1.5 gram per 100 gram (Turkish Food Codex, 2000). According to EU (1991) and IOOC (2007), the mixture of refined and virgin olive oil is named as olive oil with free fatty acidity not more than 1 gram per 100 gram.

Refined Pomace oil: Oil that is obtained by refining methods not causing any change in triglyceride structure of raw pomace oil. Refined pomace oil can be marketed directly or by mixing with natural olive oil. The free fatty acidity should not be more than 0.3 gram per 100 gram (Turkish Food Codex 2000; IOOC 2007).

Mixed (olive) pomace oil: The oil that is obtained by mixing refined pomace oil and virgin olive oil and can be consumed directly as a food. The free fatty acidity should not be more than 1.5 gram per 100 gram (Turkish Food Codex 2000). IOOC (2007) stated maximum free fatty acidity as 1 gram per 100 gram.

2.5. Composition of Olive Oil

Olive oil is composed of two main fractions which are saponifiables and unsaponifiables. Saponifiable constituents (triacylglycerols, free fatty acids, phosphatides) constitute 95-98% of olive oil whereas 2-5% of olive oil are unsaponifiables (fatty alcohols, wax esters, hydrocarbons, tocopherols and tocotrienols,

phenolic compounds, volatiles, pigments, minor glyceridic compounds, phospholipids and triterpenic acids) (Aparicio and Harwood 2000, Cert, et al. 2000).

Saponifiable fraction mainly composed of triacylglycerols. Also, most of the fatty acids in olive oil exist as triacylglcerols. Turkish Food Codex (2000), IOOC (2007) and EU (1991) restricted the amount of the individual fatty acids within the specified limits listed in Table 2.2.

Table 2.5. Fatty acid compositions of cooking olive oil, cooking pomace oil and extra virgin olive oil (Source: Turkish Food Codex 2000, IOOC 2007, EU 1991)

	Olive and olive-pomace oil (Turkish Food Codex 2000, IOOC 2007)	Extra virgin oil (EU 1991)
Myristic acid	≤ 0.05	≤ 0.1
Palmitic acid	7.5-20	-
Palmitoleic acid	0.3-3.5	-
Palmitoleic acid	≤ 0.3	-
Heptadecenoic acid	≤ 0.3	-
Stearic acid	0.5-5.0	-
Oleic acid	55.0-83.0	-
Linoleic acid	3.5-21.0	-
Linolenic acid	≤ 0.91	≤ 0.9
Arachidic acid	≤ 0.6	≤ 0.7
Gadoleic acid	≤ 0.4	-
Behenic acid	≤ 0.22	≤ 0.3
Lignoceric acid	≤ 0.2	≤ 0.5

[1] IOOC states this value ≤ 1
[2] This value of olive-pomace oil should be ≤ 0.3

Oleic acid, as the characteristic monounsaturated fatty acid of olive oil, constitutes 55-83% of total fatty acids. High intake of oleic acid in the Mediterranean region was reported to be the reason for decreases in the rates of coronary artery disease and also total mortality (Grundy 1997). Also, olive oil improves the lipid profile;

therefore, decreases cardiovascular risk by reducing the ratio of low density lipoprotein/high density lipoprotein (Martínez-González and Sánchez-Villegas 2004).

2.6. Monovarietal Characterization

Monovarietal olive oil is the oil produced from certain olive cultivars of certain region. To characterize each monovariety, first chemical and sensory properties of olive oils have to be determined. Then, the effect of external parameters (climate, ripeness and extraction systems) on sensory and chemical characteristics of the same monovarietal virgin olive oils has to be studied.

Each monovarietal olive oil is identified with its own chemical characteristic. Mean concentration of some chemical compounds of diverse European monovarietal virgin olive oils are given in Table 2.3. In the studies given in Table 2.3, olive ripeness was normal and extraction system was three-phase centrifugation system.

Table 2.6. Chemical characteristics of European monovarietal virgin olive oils (mean values) obtained from olive trees cultivated in different geographical origins (Source: Aparicio and Luna 2002)

Varieties of cultivar†	Country	P [%]	S [%]	O [%]	L [%]	Ln [%]	PH [mg/kg]	ER [mg/kg]	ST [mg/kg]	MST [mg/kg]	AA [mg/kg]	TA [mg/kg]
Arbequina	Spain	14.3	2.1	75.3	8.5	0.6	564.6	22.6	1432.6	154.7	204.9	1247.8
Cornicabra	Spain	8.4	2.6	78.7	7.5	0.6	237.2	45.5	1519.3	152.4	181.0	715.4
Farga	Spain	9.4	2.4	71.7	8.6	0.5	97.0	43.1	1551.6	177.1	360.8	841.0
Hojiblanca	Spain	9.0	3.3	75.2	9.5	0.7	44.9	29.6	1946.2	260.5	175.8	1146.5
Imperial	Spain	9.7	2.3	80.1	4.6	0.9	90.7	29.8	1481.0	180.6	188.3	945.7
Lechin	Spain	10.5	1.4	70.8	13.9	1.1	108.1	26.6	1806.9	327.4	135.9	734.7
Morrut	Spain	7.6	2.0	73.0	13.7	0.1	113.4	47.2	1463.7	120.3	135.8	639.8
Nevadillo	Spain	11.0	3.0	75.3	7.3	0.7	50.1	30.7	1596.1	186.1	205.1	1069.7
Picual	Spain	9.9	3.2	77.8	5.0	0.7	24.5	18.5	1310.2	160.1	227.1	1162.7
Redondilla	Spain	12.5	1.7	68.3	14.1	1.1	118.6	37.2	2032.3	387.4	121.4	1066.6
Sevillenca	Spain	11.1	1.7	66.5	17.1	1.0	215.0	32.1	2002.3	300.3	89.3	998.2
Verdial de Huevar	Spain	9.0	2.2	66.7	8.4	0.8	228.7	69.1	1606.3	337.4	162.4	878.3
Coratina	Italy	9.7	3.2	76.9	7.5	0.6	81.4	26.2	1192.3	–	65.0	712.2
Frantoio	Italy	9.5	2.9	78.2	7.4	0.6	192.3	26.2	1325.3	245.3	177.2	872.3
Leccino	Italy	14.3	1.8	77.7	7.7	0.8	177.6	18.9	1271.5	117.4	121.8	1071.5
Moraiolo	Italy	10.5	2.2	77.7	7.5	0.6	101.2	26.1	1184.7	148.3	134.3	836.4
Koroneiki	Greece	13.3	2.0	71.9	8.8	0.7	147.2	44.9	1965.2	–	249.3	593.8

† P - palmitic acid; S - stearic acid; O - oleic acid; L - linoleic acid; Ln - linolenic acid; PH - phytol; ER - erythrodiol; ST - total content of sterols; MST - total content of methyl sterols; AA - total content of aliphatic alcohols; TA - total content of triterpenic alcohols.

The differences among the chemical compositions of each variety allow the discrimination. Application of univariate approach to each of the chemical compounds

neglects any interaction between several of the factors that determine the compositional profile of the classes of oil compounds. It was pointed out that the classes of compounds that make up olive oil are biosynthesised through independent and genetically controlled pathways that determine the ratios of all the homologous components of each chemical class (Bianchi, et al. 2001). Multivariate statistical analysis allows the use of chemical variables taking into account the interactions between the chemical compounds. There are numerous studies involving the classification of olive oil samples using chemical and sensorial properties with chemometrics. Stefanoudaki, et al. (2000) was able to characterize three European olive oil varieties with sensory data described by a trained panel. Also fatty acids, fatty alcohols, polycyclic triterpenes and squalene were used to differentiate monovarietal olive oils. Firstly, PCA was applied to the analytical data to reveal the compounds (variables) with the highest weights (loadings). Then using these most influential variables linear discriminant analysis (LDA) was performed and 96% correct group classification was achieved (Giasante, et al. 2003). Analytical parameters such as fatty acid profile, triacylglycerols and sterols were also used in other studies. Brescia, et al. (2003) described how cultivar differences can be established between Italian oils, obtained from single varieties, based on acid, sterol, and triacylglycerols differences determined by chemometrics. Both PCA and DA indicated that triacylglycerols and fatty acid composition provided the best basis for differentiation of olive oils among cultivars. Also, applicability of triacylglycerols and fatty acid for classification of French olive oils (Aix-en-Provence, Haute-Provence, Nyons, Nice and Valleé des Baux de Provence) was studied. A linear discriminant analysis was applied to the samples. In this study 37 parameters were used in differentiation of registered designations of origin: Nyons, Nice and Haute-Provence (Ollivier, et al. 2006).

Recently, there is an increasing interest on rapid measurements for classification studies instead of chromatographic techniques. Casale, et al. (2007) employed combined information from headspace mass spectrometry (HS-MS) and visible spectroscopy for the classification of Ligurian olive oils. Application of LDA, after feature selection, was sufficient to differentiate the three geographical denominations of Liguria (Riviera dei Fiori, Riviera del Ponente Savonese and Riviera di Levante). The success was 100% in classification and close to 100% in prediction. Similarly, HS-MS designed for the sensory characterization and classification of extra virgin olive oil on the basis of its protected designation of origin, olive variety and geographical origin is reported in

another study. Results revealed that an average of ca. 87% of samples correctly classified and a specificity of ca. 97% was obtained with HS-MS in combination with appropriate chemometric treatment (López-Feria, et al. 2008). Also, electronic nose and an electronic tongue, in combination with multivariate analysis, have been used to verify the geographical origin and the uniqueness of specific extra virgin olive oils. The results of this study stated that neural networks have provided very satisfactory for the analysis of results and have indicated that the electronic nose as the most appropriate tool for the characterization of the analyzed oils (Cosio, et al. 2006). Forina, et al. (2007) focused on usefulness of visible spectra with chemometrics in the extra virgin olive oil geographical characterization. The developed class models for West Liguria Protected Designation of Origin (PDO) olive oils can be differentiated from other olive oil samples (Greece, Spain and Tunisia) with 100% sensitivity and specificity. In a further study, chemometric treatment of near infrared (NIR) reflectance spectra was applied for the classification of five geographically close registered designations of origin (RDOs) of French virgin olive oils (Aix-en-Provence, Haute-Provence, Nice, Nyons and Vallée des Baux). Partial least square-discriminant analysis (PLS-DA) was used as multivariate statistical tool for the classification of French RDOs. The results were quite satisfactory, in spite of the similarity of cultivar compositions between two denominations of origin (Aix-en-Provence and Vallée des Baux) (Galtier, et al. 2007). Downey, et al. (2003) examined visible and NIR spectra for their ability to classify extra virgin olive oils from the eastern Mediterranean on the basis of their geographic origin. A correct classification rate of 93.9% on the prediction set was achieved using factorial discriminant analysis on raw spectral data over the combined wavelength range.

2.6.1. Factors Affecting Olive Oil Composition

The major challenge to characterize olive oil according to cultivar and geographical origin is variation of its composition with respect to factors such as: environmental (soil, climate), agronomic (irrigation, fertilization), cultivation (harvesting, ripeness), and technological (post-harvest storage and extraction system). There are numerous studies involving the effects of these factors on sensory properties and chemical composition of olive oil (Aparicio and Luna 2002). Torres, et al. (2005)

stated that the amount of total phenols was strongly affected by the extraction method (two-phase centrifugation and pressure). Gutiérrez, et al. (1999) pointed out the correlation between chemical components and the ripeness index of olive oil. Results of the 1998 season indicated that the irrigation treatments had a significant effect in the oxidative stability, polyphenols and the composition of fatty acids of olive oil (Faci, et al. 2002).

Climate has a great influence on the chemical composition of vegetable oils and its effect on monovarietal characterization was widely studied by the scientists. Boggia, et al. (2002) studied with olive oils from the three geographical areas which were mentioned in the PDO regulation and obtained in 1998/99 and 1999/2000 harvest years. 31 variables determined by chemical-physical analyses were used to classify oils on the basis of their geographical origin. For 1998/99 harvest year, class-models of the three geographical areas were obtained with good predictive ability. However, the oil samples obtained from the 1999/2000 crop season were different from similar samples obtained in the previous year. These years showed clearly different climatic conditions. In 1998-99, high summer temperatures and poor autumn rainfalls contributed to limiting *Dacus oleae* infestation and improving oil quality. On the contrary, in the following year lower summer temperatures favored the spreading of the infestation, and strong autumn winds and rainstorms further worsened olive oil quality. Also, the effect of climate on fatty acid composition of Sicilian cultivars collected in nine years of cultivars was studied. PCA revealed that only olive oils samples of 1999-2000 year are always grouped together and slightly separated from the others. Similar to previously mentioned study, this year was characterized by bad climatic conditions and a widespread infestation which leads to differentiation of this harvest year (D'Imperio, et al. 2007). Salvador, et al. (2002) detected significant statistical differences in quality indices and major fatty acid and sterol compositions with respect to the year of production, with the exception of total phenols.

2.6.2. Application of Fourier Transform Infrared Spectrometry for Monovarietal Olive Oil Characterization

As a rapid analysis technique, mid-infrared spectroscopy supplies high speed measurement, moderate instrument cost and relative ease of sample presentation,

especially for liquid and paste material (Downey 1998). Infrared spectroscopy uses interaction of electromagnetic radiation with a sample to obtain both qualitative and quantitative chemical information (Harwood and Aparicio 2002). MIR radiation is defined as the infrared radiation between 4000 and 400 cm^{-1} wavenumber. The interaction of infrared radiation with matter causes it to be absorbed and also the chemical bond in a sample vibrates. Functional groups, chemical structural fragments within molecules, tend to absorb infrared radiation in the same wavenumber range, without regard to the structure of the rest of the molecule (Smith 1996).

Attenuated total reflection (ATR) sample accessories have greatly eased the analysis of solids, liquids, semisolids, and thin films. A schematic diagram of an ATR accessory is shown in Figure 2.1. ATR cell includes a crystal of highly-refractive material (e.g. ZnSe). The test sample is layered on the top surface of the crystal. Infrared radiation enters at one end undergoes total internal reflection at the top and bottom faces, and then exit from the other end to the detector. A small evanescent wave is generated at each point of reflection, which goes a short distance inside any sample in contact with the top face. This interaction results in the absorption of radiation by the sample and the consequent attenuation of the input signal at a number of wavelengths, thus producing an absorption spectrum (Downey 1998).

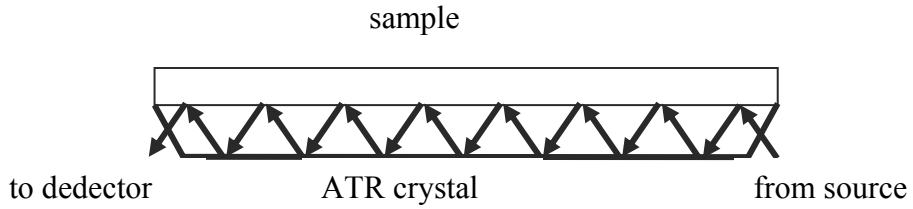


Figure 2.1. A schematic diagram of an attenuated total reflectance accessory (Source: Downey 1998)

MIR spectra contain information about the complete chemical composition and physical state of the material under analysis. Therefore, high number of data is generated (Downey 1998). Multivariate data analysis is required to extract the relevant information from these spectra.

Despite availability of FT-IR with ATR and chemometrics for monovarietal classification of olive oil, there are not many studies involving classification of

monovarietal olive oils using FT-IR and chemometric analysis. Bendini, et al. (2007) collected FT-IR spectra of 84 monovarietal virgin olive oil samples from eight Italian regions and manipulated fingerprint region data by PCA. Discrimination of majority of samples coming from the Emilia Romagna, Sardinian and Sicilian region was achieved. In another study, applicability of FTIR to distinguish extra virgin olive oils from four European countries, in combination with multivariate analysis, was investigated. Both PLS-DA and genetic algorithm (GA)-PLS was applied independently and results confirmed that PLS-LDA approach produced a cross-validation success rate of 96%, whereas the GA-PLS approach achieved a 100% cross-validation success rate. Also, Caetano, et al. (2007) reported application of FT-IR spectroscopy to discriminate between Italian and non-Italian and between Ligurian and non-Ligurian olive oils. Two chemometric techniques, classification and regression trees and support vector machines based on the Gaussian kernel and the recently introduced Euclidean distance-based Pearson VII Universal Kernel, successfully achieved to discriminate olive oil samples between various geographical origins.

2.7. Geographical Indications

In order to protect high-quality agricultural products based on their geographical origin, EU established legislations to designate geographical indicators (GIs). GIs are special signs that can add economic value to the agricultural products in markets. They symbolize cultural identity of region of origin by representing the skills of its population and natural resources. Also, they create indistinguishable characteristics for the products (Addor and Grazioli 2002).

Two types of GI designations were established by EU Council Regulation: Protection of Designations of Origin (PDO) and Protection of Geographical Indication (PGI). PDO labeling means that the product is produced, processed, and prepared within the specified geographical region. 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, products quality and reputation could still be ascribed to that geographical region. These designations could be used in labelling of olive oil, produced in particular geographical region, with typical

characteristics linked to natural factors, to the environment, and to the traditions of that region (Babcock and Clemens 2004, Ozen, et al. 2005).

As a high-value agricultural product, GIs were applied to olive oil in European Union. Therefore, there is an economic basis for identifying characteristics that distinguish PDO olive oils.

2.8. Adulteration

Agricultural products having PDO or PGI labels might have high market price so they are attractive targets for adulteration. Adulteration of food products involves the replacement of high-cost ingredients with lower grade and cheaper substitutes (Tay, et al. 2002). Actually, blend edible oils can be prepared only for suitable products, but if the resulting blend deviates from the mixture proportions given on the label, or if the blend is traded as genuine, it means the oil is adulterated (Ulberth and Buchgraber 2000).

Olive oil is also one of the agricultural products that is often adulterated with cheaper oils. Consumers generally prefer olive oil for its health benefits and they could feel cheated to receive oil that does not provide their expectations. Adulteration received much more attention after the toxic oil syndrome resulting from consumption of olive oil spiked with aniline-denatured rapeseed oil that affected more than 20,000 people. Oils widely used for olive oil adulteration include olive pomace oil, corn oil, peanut oil, cottonseed oil, sunflower oil, soybean oil, and poppy seed oil (Harwood and Aparicio 2002).

2.8.1. Methods to Detect Adulteration of Olive Oil

Monitoring authenticity of edible oils is carried out using instrumental techniques that provide data about their qualitative and quantitative composition. There exist numerous methodologies to qualify and quantify vegetable or seed oils in olive oil. Examples of some standard techniques involving the application chromatographic methods are provided in this section.

The limits of total and desmethyl sterol compositions of olive oil are given in EC (1991). Sterol composition significantly varies in between vegetable oils and this

variation can be used to detect adulteration (Aparicio and Aparicio-Ruíz 2000). Saponification method and analysis of trimethylsilylether with capillary column GC was applied to detect seed oil adulteration (IOOC 2007). Another method is determination of saturated fatty acids at 2-position of triacylglycerols. Main saturated fatty acids of olive oil are palmitic and stearic acids and they can not be more than 1.5 percent at 2-position of triacylglycerols for virgin olive oil. The methyl esters can be prepared and then analyzed with thin-layer chromatography. (EC 1991, Aparicio and Harwood 2000). Triacylglcerides of vegetable oils can also be analysed to confirm the purity of vegetable oils. During triacylglcerides analysis by high performance liquid chromatography (HPLC), out of the entire chromatogram, the only peaks which are taken into consideration are trilinolein and Equivalent Carbon Number 42 (ECN42) (Christopoulou, et al. 2004). Determination of trilinolein content was used for the detection of adulteration of olive oil with other vegetable oils (EC 1991). Nowadays, the trilinolein content has been replaced by the ΔECN42. Christopoulou, et al. (2004) stated that the limit for the ΔECN42 in olive oil is satisfactory for the detection of adulteration of an olive oil with oils: sunflower, soybean, cotton, corn, walnut, sesame, safflower, canola and rapeseed. The established limit for the ΔECN42 is not satisfactory for detecting percentages lower than or equal to 5% of hazelnut, almond, peanut and mustard oils in mixtures with olive oil. Wax esters are other components of olive oil that can be used for adulteration detection. Wax determination involves the use of capillary column gas-liquid chromatography (EC 1991). They are very useful to distinguish refined olive oil and olive–pomace oils from virgin olive oil because virgin olive oil has a higher content of C_{36} and C_{38} waxes than of C_{40}, C_{42} and C_{44} whereas the other oils have an inverse relation (Aparicio and Aparicio-Ruíz 2000).

2.8.2. Emerging Techniques to Detect Adulteration of Olive Oils with an Emphasis on IR Spectroscopy

Methods of food adulteration have become more sophisticated due to its economic profits. There is an increasing demand for the development of new rapid and sensitive methods instead of traditional time-consuming and expensive analysis techniques. There exist several new emerging methods mainly focusing on this subject.

Headspace autosampler directly coupled to a mass spectrometer (ChemSensor) was employed to detect hazelnut oil adulteration. Using PLS and principal component regression analysis, minimum adulteration levels of 7 and 15% can be detected in refined and virgin olive oils, respectively (Peña, et al. 2005). Poulli, et al. (2007) described the differentiation of virgin olive from olive-pomace, corn, sunflower, soybean, rapeseed and walnut oils using total synchronous fluorescence spectra with PLS. Olive-pomace, corn, sunflower, soybean, rapeseed and walnut oil were detected in virgin olive oil at levels of 2.6%, 3.8%, 4.3%, 4.2%, 3.6%, and 13.8% (w/w), respectively. Also, differential scanning calorimeter was employed to detect adulteration of extra virgin olive oil with refined hazelnut oil and results revealed that thermal properties of cooling thermograms were affected by hazelnut oil addition at a concentration as low as 5% (Chiavaro, et. al 2008).

Also, use of spectroscopy which includes IR and Raman techniques combined with chemometric methods is a relatively new approach to determine authenticity of olive oil. NIR, MIR, and Raman spectroscopic techniques were used to quantify the amount of olive pomace oil adulteration in extra virgin olive oil. Developed PLS model using data by Raman spectroscopy had R^2 of 0.997 and standard error of 1.72%. In addition, NIR and middle infrared (MIR) techniques also provided good predictions with a R^2 value greater than 0.99 (Yang and Irudayaraj 2001). Christy, et al. (2004) studied NIR spectroscopy to detect and quantify adulteration of olive oil with soy, sunflower, corn, walnut and hazelnut oils. PCA of all spectral data revealed 100% classification of mixtures according to adulterant. Quantification of adulterant in olive oil was performed by PLS which was resulted in error limits of ± 0.57% (corn oil), ±1.32% (sunflower oil), ±0.96% (soy oil), ±0.56% (walnut oil) and ±0.57% (hazelnut oil).

Detection of olive oil adulteration with hazelnut oil is a real challenge due to its very similar chemical composition to olive oil. There are studies involving the application of FT-IR with chemometrics to detect hazelnut oil adulteration. FT-IR equipped with a ZnSe-ATR accessory was used to detect the adulteration of extra-virgin olive oil with hazelnut oil. FT-IR data was manipulated with discriminant analysis. However, adulteration of virgin olive oil with hazelnut oil could be detected only at levels of 25% and higher (Ozen, et al. 2002). Baeten, et al. (2005) obtained better results in their study which involves application of Raman and MIR spectroscopies to determine the level of hazelnut oil in olive oil. In this study, MIR spectra worked better

than Raman spectra. Results of stepwise linear discriminant analysis indicated that MIR spectra of unsaponifiables matter of samples allowed detection of adulteration at a level of 8% for blends obtained by mixing Turkish hazelnut and olive oils.

Other vegetable or seed oils were also used to adulterate olive oil and FT-IR was also suggested as the possible adulteration detection method. Olive oil and sunflower adulterated olive oil samples (20ml/l) were analyzed by FT-IR equipped with ZnSe-ATR accessory. According to this study, DA was able to classify the samples as pure and adulterated. PLS model developed to determine the level of mixing resulted in R^2 value of validation set of 0.974 which indicated success of the model (Tay, et al. 2002). Vlachos, et al. (2006) also studied the determination of olive oil adulteration with vegetable oils (sunflower oil, soybean oil, sesame oil, corn oil) using FT-IR. A linear relation was obtained between percent adulteration and height of the band shift at 3009 cm^{-1} of FT-IR spectra which can be used to determine extra virgin olive oil adulteration with different types of vegetable oils. In this study, the detection limit for olive oil adulteration is 9% if the adulterant is corn oil or sesame seed oil while it is lower (6%) if the adulterant is sunflower or soybean oil.

2.9. Use of Chemometric Techniques to Determine Authenticity of Olive Oil

The authenticity determination of olive oil can be differentiated into 2 main scopes which are:

- characterization and denomination of geographical origin and
- identification and quantification of economic adulteration (Ulberth and Buchgraber 2000).

In authentication of food products numerous chemical compounds create multivariate data. Chemometric techniques deal with the multivariate data to extract relevant information and to organize them into meaningful structures (Harwood and Aparicio 2000). The multivariate data analysis aims to display multidimensional data in a lower space without loss of any significant information to group samples with similar properties and to classify observations on the basis of their similarities (Massart, et al. 1988).

As an example fatty acid profile of olive oils from different varieties or geographical regions could result in generation of multivariate data sets for olive oil authentication. Also, spectroscopic techniques are generally applied for authentication because spectra contain information about the complete chemical composition of the olive oil. Therefore, multivariate data analysis is required to extract the relevant information from the spectra or fatty acid profile of olive oil (Downey 1998).

If a data set includes N observations of *K* variables, they can be arranged in a [N X K] matrix X. The relative sizes of K and N *is* important to be able to employ multivariate methods. Spectroscopic or chromatographic data sets containing higher number of K than N prevents direct application of many multivariate methods (because the product matrix X^TX, required in the calculations, cannot be inverted). PCA and PLS, data reduction methods, are most useful for this kind of data sets (Defernez and Kemsley 1997).

2.9.1. Principal Component Analysis (PCA)

PCA, being one of the commonly applied chemometric tools among the techniques to model complex data sets, aims to represent the n-dimensional data structure in a smaller number of dimensions, usually two or three. So, the groupings of observations, outliers, etc., which define the structure of the data set can be observed (Massart, et al. 1988).

PCA is given as an abstract mathematical transformation of the original data matrix which can be presented as

$$X = T.P+E \qquad (2.1)$$

where *X* is the original data matrix; *T* are the scores, and have as many rows as the original data matrix; *P* are the loadings, and have as many columns as the original data matrix; and *E* is the error matrix (Brereton 2003).

In order to apply PCA, X-variables are expected to be collinear. The collinearity means that the variability exists in X matrix and carries most of the available information. Therefore, redundancy and smaller variabilities can be removed. Then, PCA aims to express the main information in the variables by the so-called PCs of X

(Marten and Naes 1989). The PCs are oriented so that the first PC describes as much as possible of the original variation between the objects.

Two PCs are two lines orthogonal to each other and together they define a plane in K-dimensional variable space. All the observations are projected onto this plane and structure of the investigated data set can be visualized. The co-ordinate values of the observations on this plane are called *scores* and this configuration is called scores plot (Eriksson, et al. 2001).

Scores plot displays the similarities and differences between observations. On the other hand, loadings plot can be constructed for two PCs to observe the most important original variables. The important variables are placed in the longest distance from the origin and are responsible for the pattern seen among observations. In addition, correlated variables which are in the same or opposite directions on a straight line through the origin of loading plot can be determined (Euerby and Petersson 2003).

PCA analysis can also be used to develop a modelling technique called soft independent modelling of class analogy (SIMCA) which is one of the most commonly used class-modelling tools in chemometrics. In SIMCA, PCA is performed for each class separately and this results in a PC model for each class. The class distance can be calculated as the geometric distance from the PCs models. Each group may be bounded by a region of space representing 95% confidence that a particular object belongs to a class. The residual variance of a variable of a class is used for estimating the modelling power of a particular variable (Brereton 2003). On the other hand, discriminatory power measures how well a variable discriminates between two classes. This differs from the modelling power in the sense that a variable being able to model one class well, it does not necessarily imply being able to discriminate two groups effectively (Berrueta, et al. 2007).

2.9.2. Partial Least Square Analysis (PLS)

The regression extension of PCA is named as PLS. The technique is a two-block regression method that is also based on principal component models. If the independent variables is called X matrix, the dependent or response variables constitute Y matrix, response variables are called the Y matrix (Bye 1995). Both X and Y matrices are decomposed into smaller matrices as given below:

$$X=T.P+E \quad (2.2)$$

$$Y=T.Q+F \quad (2.3)$$

where Y is the matrix of dependent variables; X is the matrix of independent variables; T and U are the scores matrices; P are the loadings of X matrix; E is the error matrix of X-Matrix; Q is loading matrix of Y-Matrix; and F is the error matrix for Y-Matrix (Otto 1997).

The PLS analysis involves the determination of PLS components that describe the variation of Y-data. The first PLS component is a line in X variable space which should approximate the observations and correlate with Y vector. First PLS usually is not enough to describe the variation in the Y-data. So, the second PLS component describe the remaining variation as much as possible. Second PLS component is also a line orthogonal to first PLS component and it improves the description of X data and provides good correlation with Y remained after first component. Also, the vectors of two PLS components have the ability to predict y data (Eriksson, et al. 2001).

CHAPTER 3

EXPERIMENTAL STUDY

3.1. Materials

3.1.1. Olive and Olive Oil Samples

Two independent sets of olive oil samples were used in the analysis. In the first set, olive varieties used in oil production were provided by Olive Research Institute (Izmir, Turkey) and Olive Nursery (Edremit, Turkey). Each variety as listed in Table 3.1 was obtained from Izmir but the varieties Ayvalik and Gemlik were also obtained from Edremit. 15-25 kg of olives from each cultivar were picked from trees and then milled with a maximum 5 kg capacity laboratory scale olive oil mill (TEM Spremoliva, Italy). At least two different batches of oil were obtained from each cultivar. Samples were obtained both in 2005/06 (1^{st}) and 2006/07 (2^{nd}) harvest years. The samples were listed in Table 3.1.

Table 3.1. Olive varieties and olives of different geographical regions

Sample name	Sample code
Ayvalik	A
Gemlik	G
Memecik	M
Erkence	E
Nizip	N
Ayvalik (Edremit)	AE
Gemlik (Edremit)	GE

Second class of oil includes extra-virgin olive oil samples, which belong to 30 different locations of Aegean region of Turkey, and they were obtained from the same olive oil producer for two consecutive harvest years. Figure 3.1 represents north and

south of Aegean region where the olives were obtained from. There are 25 and 22 olive oil samples belonging to 2005/06 and 2006/07 harvest years, respectively. The olive oil samples were coded either as north (N) or south (S) according to the region where olives came from (Table 3.2). Also, numbers besides N or S represent the first or second harvest year. While mainly Ayvalik variety is cultivated in north Aegean, Memecik is the dominant variety in south part of Aegean region. The samples were kept in dark glass bottles and stored at 8 ^{o}C until further analysis. Analysis was performed within 2-3 months following extraction and receival of the oils.

Figure 3.1. Commercial olive oil samples belonging to North (N) and South (S) of Aegean region (Source: World Sites Atlas 2008)

Table 3.2. Commercial olive oil samples from north and south parts of Aegean Region of Turkey belonging to 2005/06 and 2006/07 harvest years

Sample No	Geographic origin	Sample code	Harvest year	Sample No	Geographic origin	Sample code	Harvest year
1	Ezine (N)	Ez	1st-2nd	16	Tepekoy (S)	Tep	1st-2nd
2	Ezine Gulpinar (N)	Ez-org	1st	17	Bayindir (S)	Bay	1st
3	Kucukkuyu1 (N)	Kuckuy1	1st-2nd	18	Odemis (S)	Ode	2nd
4	Kucukkuyu2 (N)	Kuckuy2	1st	19	Tire (S)	Tire	2nd
5	Altinoluk (N)	Altol	1st-2nd	20	Selcuk (S)	Sel	1st-2nd
6	Altinoluk-sulubaski (N)	Altol-sb	1st	21	Kusadasi (S)	Kus	2nd
7	Edremit (N)	Edr	1st-2nd	22	Germencik (S)	Ger	2nd
8	Havran (N)	Hav	1st-2nd	23	Aydin (S)	Ayd	1st-2nd
9	Burhaniye (N)	Bur	1st-2nd	24	Ortaklar (S)	Ort	1st-2nd
10	Gomec (N)	Gom	1st-2nd	25	Kosk (S)	Kosk	2nd
11	Ayvalik (N)	Ayv	1st-2nd	26	Dalaman (S)	Dal	2nd
12	Altinova (N)	Altov	1st-2nd	27	Kocarli (S)	Koc	1st-2nd
13	Zeytindag (N)	Zey	1st-2nd	28	Erbeyli (S)	Erb	2nd
14	Akhisar (N)	Akh	1st	29	Çine (S)	Cine	2nd
15	Menemen (N)	Men	1st	30	Milas (S)	Mil	1st-2nd

3.1.2. Chemical Reagents

Reagents used in chemical analysis were obtained from Riedel-de Haën and Sigma-Aldrich and they are either high performance HPLC or analytical grade. In chromatographic analysis, fatty acid methyl esters containing C8-C24 (2%-11% relative concentration) was used as reference standard (Supelco # 18918).

3.2. Methods

3.2.1. Determination of Percent Free Fatty Acids

European Official Methods of Analysis (EEC 1991) was used for the determination of free fatty acid value in terms of % oleic acid. 1 g potassium hydrogen phthalate ($KHC_8H_4O_4$) was dried in oven at 110 ^{0}C for 2 hours. 0.4 g of potassium hydrogen phthalate was put into a flask, and 75 ml distilled water and few drops of phenolphthalein (0.5 g phenolphthalein in 50 ml 95% ethanol (v/v)) were added. 1 mol/L potassium hydroxide (KOH) was prepared with distilled water and it was standardized with previously prepared potassium hydrogen phthalate. 150 mL diethyl ether-ethanol (1:1) mixture was neutralized with KOH with the addition of phenolphthalein. Since the expected acidity value is <1, 20 g oil sample was dissolved in diethyl ether-ethanol solution. The sample solution was titrated with 0.1 mol/L solution of KOH until the indicator changes color (the pink color of the phenolphtalein persists for at least 10 seconds). Acidity is expressed as percentage of oleic acid with the equation given below:

$$V \times c \times \frac{M}{1000} \times \frac{100}{m} = \frac{V \times c \times M}{10 \times m} \tag{3.1}$$

where:

V = the volume of titrated KOH solution used in milliliters;

c = the exact concentration in moles per liter of the titrated solution of KOH used;

M = the molar weight in grams per mole of the acid used to express the result (=282);

m = the weight in grams of the sample.

3.2.2. Gas Chromatography Analysis

3.2.2.1. Sample Preparation

European Official Methods of Analysis (EEC, 1991) was used for the preparation of methyl esters. 100 mg oil sample was weighed in 20 mL centrifuge tubes. The samples were dissolved in 10 mL n-hexane and saponified to their methyl esters with the addition of methanolic potassium hydroxide solution (11.2 g in 100 mL). The sample solution was vortexed for 30 s and centrifuged for 15 min. Clear supernatant was transferred into 2 mL autosampler vial for chromatographic analysis.

3.2.2.2. Analytical Conditions

Chromatographic analyses were performed on an Agilent 6890 GC (Agilent Technologies, Santa Clara, USA) equipped with Agilent 7683 autosampler. The instrumental configuration and analytical conditions were presented in Table 3.3.

Table 3.3. Chromatographic method for the analysis of fatty acid methyl esters

Instrumentation	
Chromatographic system	Agilent 6890 GC
Inlet	Split/splitless
Detector	FID
Automatic sampler	Agilent 7683
Liner	Split liner (p/n 5183-4647)
Column	100 m x 0,25 mm ID, 0.2 μm HP-88 (J&W 112-88A7)

(cont. on next page)

Table 3.3. (cont.) Chromatographic method for the analysis of fatty acid methyl esters

Experimental Conditions GC-FID	
Inlet temperature	250 ^{o}C
Injection volume	1μL
Split ratio	1/50
Carrier Gas	Helium
Head pressure	2 mL/min constant flow
Oven temperature	175^{o}C, 10 min, 3^{o}C/min, 220^{o}C, 5 min
Detector temperature	280 ^{o}C
Detector gases	Hydrogen:40mL/min; Air:450mL/min; Helium make-up gas:30mL/min

Fatty acids used in the analysis were myristic acid (C14:0), palmitic acid (C16:0), palmitoleic acid (C16:1), stearic acid (C18:0), oleic acid (C18:1), linoleic acid (C18:2), arachidic acid (C20:0), linolenic acid (C18:3) and behenic acid (C22:0). Each sample was analysed at least two times with GC.

3.2.3. FT-IR Analysis

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

Horizontal ATR accessory was used to collect the spectral data of oil. The resolution was set at 2 cm^{-1} and the number of scans collected for each spectrum was 128. ZnSe crystal was cleaned with hexane in between sample runs. Measurements were conducted more at least two times.

3.3. Preparation of Adulterated Samples and Binary Olive Oil Mixtures

Monovarietal olive oil adulteration was performed by mixing E and AE with N. The percentage of N added into E and AE varied between 2-20% (vol/vol). On the other

hand, vegetable and seed oil adulteration was implemented with the addition of hazelnut, sunflower-corn binary mixture, cottonseed and rapeseed oil into commercial olive oil samples. Each four commercial olive oil samples were randomly selected from north and south regions and mixtures of oils from north and south were prepared speretaly. Then each mixture was blended with hazelnut oil to obtain two sets of 2-50% (vol/vol) adulterated samples. For cottonseed and rapeseed oil adulteration, four commercial olive oil samples belonging to north were mixed with rapeseed and cottonseed oils. For sunflower-corn oil adulteration, sunflower-corn oil binary mixtures were prepared at different concentrations (0-100% vol/vol) and mixed with blend of four commercial olive oil samples belonging to north at 2-20% (vol/vol). Composition of each oil sample in mixture was presented in Table 3.5.

Table 3.5. Percentages of sunflower, corn and olive oil in 2-20% adulterated mixtures

Sunflower oil (%)	Corn oil (%)	Olive oil (%)
1	1	98
2	0	98
0	2	98
2.5	2.5	95
5	0	95
0	5	95
5	5	90
10	0	90
0	10	90
7.5	7.5	85
10	5	85
5	10	85
10	10	80
15	5	80
5	15	80
20	0	80
0	20	80

3.4. Statistical Analysis

Data analysis was performed using multivariate statistical methods with SIMCA software (Umetrics, Sweden). Both fatty acid profile and FT-IR spectral were statistically analysed. Whole FT-IR spectra of the samples were not used in the data analysis. Two portions of whole spectra (3620-2520 and 1875.5-675 cm^{-1}) where there are most significant differences in the peak intensities, were selected to be employed.

3.4.1. Pre-treatment of Data

In order to transform the data into a form suitable for PCA and PLS, the data is often pre-treated. Within this concept, both fatty acid composition and spectral data were scaled and mean centered.

Signal correction and compression techniques are quite applicable to MIR spectra. Wavelets analysis is a widely used spectral compression technique preferred in classification studies. Its main abilities are compressing and de-noising complicated signals. Wavelets are small oscillating waves having the capability of probing a signal according to scale, that is, bandpass of frequencies. There are different wavelet functions such as Daubechies, Symmlet and Coiflet. Among wavelet functions Daubechies-10 was chosen and wavelet analysis used as a spectral data pre-treatment method before classification approaches (PCA) (Eriksson, et al. 2000). Orthogonal signal correction (OSC) is a signal correction technique and it is used to construct a filter that removes the part from the spectral matrix X that is definitely unrelated to Y. The removed part should be mathematically orthogonal to Y (Wold, et al. 1998). In this study, OSC in combination with wavelet analysis was applied on spectral data for quantification issues (PLS).

3.4.2. Classification

The discrimination of olive oil samples was achieved with 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. Degree of fit increases as the number of components increases but predictive ability does not necessarily increase after a certain model complexity. Therefore, it is important to reach an optimal balance between fit and predictive ability.

Results of PCA are visualized by scores, loadings and Coomans' plots. Score plots were constructed to observe principal groupings among observations. To determine the variables responsible from the groupings among observations loading plots were used. Also, Coomans' plot is used to interpret results of SIMCA and to discriminate two classes. In this plot, the distance from the model for class 1 is plotted against that from model 2. The critical distances (usually at 95% of confidence level) are indicated on both axes. So, four regions are defined on the plot: class 1, class 2, overlap of classes 1 and 2, and outer region (far from both classes). By plotting objects in this plot it is easy to visualize how certain a classification is. Scores and Coomans' plots were employed for discrimination of samples according to variety, geographical origin, and harvest year and also for differentiation of adulterated samples from pure samples. Classification studies were performed by setting 9 GC variables and FT-IR spectra as independent variable data set.

3.4.3. Quantification

Quantification of fatty acid profile and free fatty acidity using MIR spectra was implemented. The whole observation data set was divided into validation and calibration sets. PLS was employed on calibration set relating MIR spectra (X block) with fatty acid profile and free fatty acidity value (Y block). The ability of the PLS model was inspected with validation set. The predictability of the models was tested by computing the standard error of calibration (SEC) for the calibration data set, the standard error of prediction (SEP) for the validation data set and the relative error of prediction (REP):

$$SEC = \sqrt{\frac{\sum_{i=1}^{m} (\hat{Y}_i - Y_i)^2}{m-2}} \qquad (3.2)$$

$$SEP = \sqrt{\frac{\sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2}{n}} \tag{3.3}$$

$$REP = \frac{SEP}{\bar{Y}} x100 \tag{3.4}$$

where Y_i, $\hat{Y}_i$ and $\bar{Y}$ are observed, predicted percentage of each fatty acid and average observed fatty acid of observations, respectively; *m* and *n* are the number of samples in the calibration and prediction sets, respectively.

Quantification of hazelnut, sunflower-corn, cottonseed and rapeseed oil in olive oil and also N in blend of E-N and AE-N was performed by PLS regression analysis, which relates the FT-IR absorbance of each adulterated samples and monovarietal mixture (X block) with the percentages of adulterant oil and a monovarietal in that mixture (Y block). For hazelnut, cottonseed and rapeseed oil, adulterated olive oil samples randomly divided into two as the calibration and validation data.

On the other hand, due to the low number of samples, the data belonging to mixture of two monovarieties were not divided into two as the calibration and validation data sets. Instead, cross-validation technique was used to asses 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. 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 error criterion; (SEP) as given in equation 3.3.

CHAPTER 4

RESULTS and DISCUSSIONS

4.1. Classification of Lab Scale Extracted Olive Oils

Olive varieties used in this study are commercially important cultivars in olive oil production. All varieties except Nizip are mainly cultivated in west part of Turkey. Nizip, on the other hand, originates from southeast part of Turkey. Five varieties (A, G, M, E and N) used in this study were obtained from the same orchard in Izmir which is in the middle part of Aegean region and two of these varieties (A and G) were also supplied from another orchard in north part of the same region (Edremit) for two consecutive harvest years. Same extraction system was used in obtaining the oils. Data of FFA values of the olive oil samples were presented in Table 4.1. The acidity values varied between 0.94 and 0.13 belonging to 1st harvest year of the N and M varieties, respectively. None of the olive oils has the acidity value higher than 1% which is the maximum accepted value of EU (1991) for extra virgin olive oil classification.

Table 4.1. Free fatty acid values of lab scale extracted olive oils expressed as % oleic acid

Variety	Harvest year	
	2005	2006
Ayvalik (A)	0.38±0.07^{c}	0.31±0.1^{bc}
Gemlik (G)	0.13±0.03^{a}	0.16±0.02^{ab}
Memecik (M)	0.24±0.02^{bc}	0.33±0.1^{c}
Erkence (E)	0.16±0.03^{a}	0.14±0.01^{a}
Nizip (N)	0.94±0.18^{d}	0.45±0.3^{db}
Ayvalik-Edremit (AE)	0.32±0.02^{c}	0.80±0.04^{d}
Gemlik-Edremit (GE)	0.17±0.02^{ab}	0.25±0.06^{bc}

^{a,d} Values in each column with different superscript letters present significant differences ($P< 0.05$).

Data from GC and FT-IR analysis of oils in combination with chemometrics were employed for the classification of samples. The variety effect was investigated using the varieties grown in Izmir. On the other hand, comparison of two varieties (A and G) from two different areas (Izmir and Edremit) provided information about the effect of growing location.

4.1.1. Classification using GC Data

Fatty acid profile of olive oil samples, determined by GC analysis, presented in Table 4.2. Fatty acid content of each variety was the average of at least two GC measurements for each batch of oil. The monounsaturated fatty acids have great importance because of their nutritional implication and their effect on oxidative stability of oils (Aparicio, et al. 1999). As the most abundant fatty acid in olive oil, the amount of oleic acid varied remarkably between the olive oil samples. The oleic acid content was at the lowest concentration (66.32%) for M variety of 1st harvest year and (63.57%) for E variety of 2nd harvest year. On the other hand, the highest oleic acid concentration was observed for GE variety for two consecutive harvest years (72.05% for 1st and 72.95% for 2nd harvest years). The content of linoleic acid varied between 8.88% for GE and 14.95% for E varieties for 1st harvest year whereas its content was between 16.89% for E and 7.41% for GE varieties belonging to 2nd harvest year. Also, fatty acid composition of each variety changed with respect to year of harvesting. In other studies, similar fatty acid profile for Ayvalık variety and flesh oil of Memecik variety was determined (Ozkaya, et al. 2008, Nergiz and Engez 2000). On the other hand, Tanılgan, et al. (2007) observed higher oleic and lower linoleic acid contents for Ayvalık (C18:1 is 79% and C18:1 is 6.8%) and Gemlik (C18:1 is 81.1%, C18:2 is 4.9%) varieties. The difference could be due to extraction method, olive ripeness or harvest year differences.

The 9 GC variables of 122 observations were used as the multivariate data set. The data set were submitted to PCA to visualize the presence of principal groupings. Figure 4.1 shows the score plot of lab scale extracted olive oil samples.

Table 4.2. Fatty acid profiles (% of total fatty acids) of lab scale extracted olive oil samples

	Variety	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3	C20:0	C22:0
2005										
	A	0.03±0.00	14.24±0.37	0.84±0.01	2.77±0.09	69.58±0.44	11.22±0.11	0.45±0.02	0.75±0.01	0.13±0.02
	G	0.01±0.00	14.01±0.75	0.93±0.07	2.92±0.06	70.67±1.21	10.13±0.39	0.39±0.03	0.78±0.02	0.15±0.06
	M	0.02±0.00	14.95±0.73	0.89±0.05	2±0.08	66.32±0.96	14.5±0.72	0.35±0.04	0.86±0.01	0.11±0.03
	E	0.01±0.00	14.09±0.88	0.75±0.07	2.39±0.09	66.44±1.59	14.95±0.85	0.38±0.02	0.86±0.03	0.12±0.02
	N	0.01±0.01	15.19±0.52	0.81±0.47	4.57±0.37	68.35±0.97	9.95±0.22	0.49±0.02	0.53±0.05	0.09±0.02
	AE	0.02±0.00	14.16±0.25	1.03±0.02	2.29±0.02	69.03±0.23	11.33±0.09	0.42±0.04	0.67±0.02	0.13±0.03
	GE	0.01±0.00	13.72±0.31	1.06±0.02	3.19±0.17	72.05±0.31	8.88±0.21	0.37±0.04	0.59±0.01	0.14±0.06
2006										
	A	0.02±0.00	16.51±0.37	2.65±0.03	2.08±0.03	65±0.26	12.69±0.16	0.49±0.02	0.42±0.01	0.13±0.02
	G	0.01±0.00	14.45±0.45	2.07±0.21	3.37±0.39	71.2±1.65	7.82±0.66	0.49±0.02	0.48±0.02	0.12±0.01
	M	0.02±0.00	12.71±0.62	1.29±0.03	2.13±0.10	72.88±0.61	10.01±0.16	0.46±0.01	0.4±0.02	0.12±0.01
	E	0.02±0.00	14.62±0.31	1.6±0.04	2.28±0.07	63.57±0.46	16.89±0.43	0.52±0.01	0.4±0.01	0.12±0.01
	N	0.01±0.00	14.98±0.65	1.54±0.02	5.1±0.19	67.21±0.66	10.14±0.14	0.32±0.01	0.56±0.02	0.13±0.02
	AE	0.02±0.00	15.03±0.64	1.87±0.06	2.11±0.09	66.29±0.45	13.77±0.33	0.39±0.01	0.4±0.02	0.13±0.02
	GE	0.01±0.00	13.48±0.40	2.08±0.09	3.13±0.12	72.95±0.53	7.41±0.12	0.39±0.01	0.43±0.02	0.11±0.01

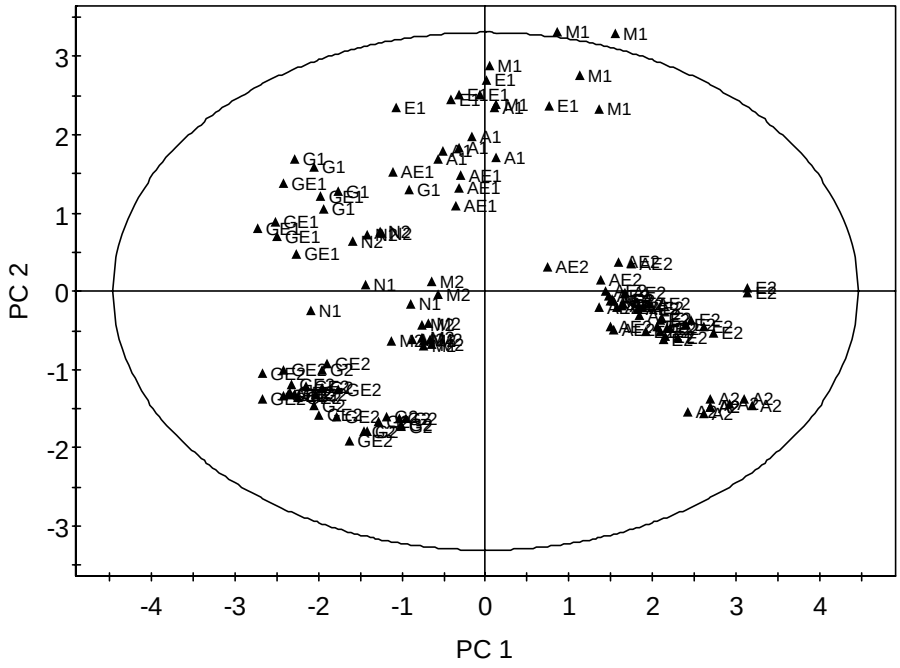


Figure 4.1. PCA score plot of the fatty acid profile of lab scale extracted olive oil samples for two harvest years

In Figure 4.1, first PC accounted for 35.6% of total variation whereas second PC explained much less 19.6%. In score plot the area inside the ellipse presents 95% confidence. Some varieties were grouped together and differentiation of cultivars with respect to harvest years was apparent.

Loading plot obtained from PCA of fatty acid data is represented in Figure 4.2. Variables C18:1, C18:2, C18:3 and C16:1 were far from the origin and this means that they have an important effect on the classification of oil samples with respect to cultivar, geographical origin and harvest year. The high discrimination power of C16:0, C18:1 and C18:2 were also stated in the classification of Sicilian cultivars with respect to cultivar and geographical effects (D`Imperio, et al. 2007).

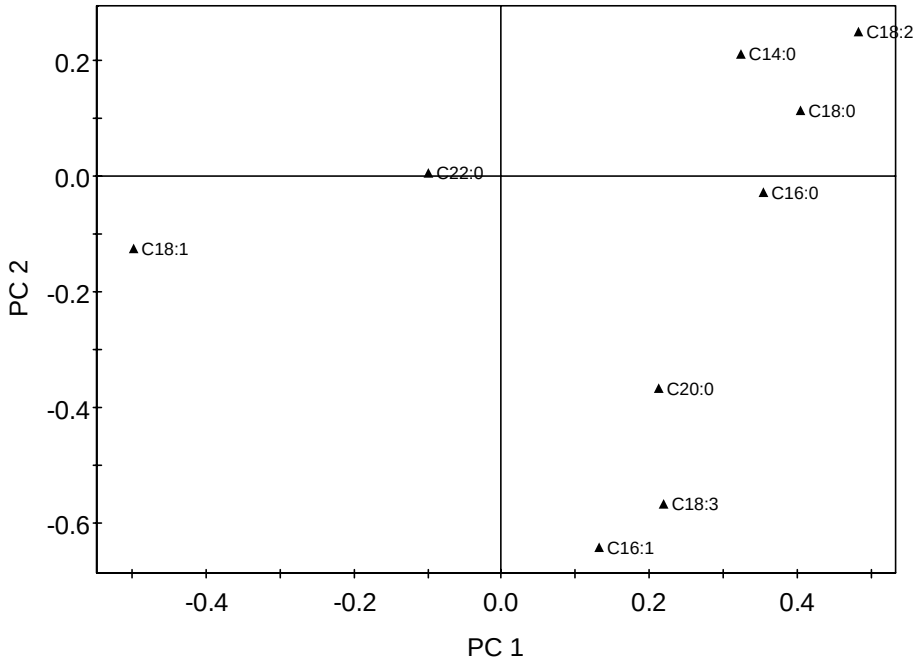


Figure 4.2. PCA loading plot of the fatty acid profile of lab scale extracted olive oil samples

Coomans' plot is an extremely useful tool to visualize principal groupings, in which the two axes represent the distance of individual model. In Coomans' plot; PCA model of each class were built separately and two class models are plotted against each other with the critical levels as straight lines displaying the boundaries. Any sample having a distance to the corresponding centroid greater than the critical distance is considered as being outside the class model and, as a consequence, rejected as an outlier for the specific category. Coomans' plot using fatty acid compositions was constructed for discrimination of olive oil samples (Figure 4.3).

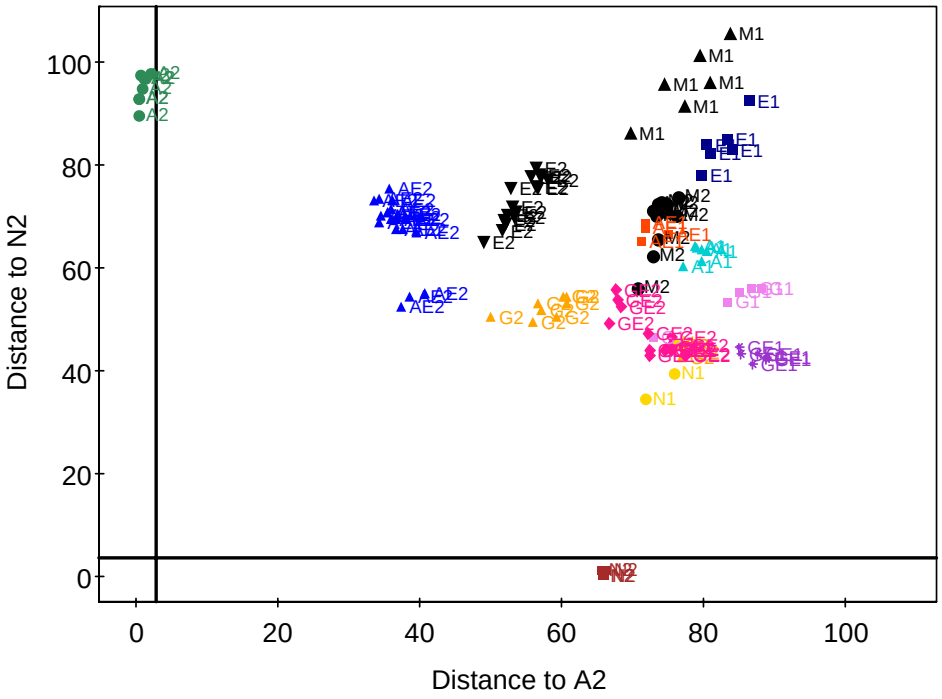


Figure 4.3. Coomans' plot for the classification of olive oil samples with respect to variety and harvest year using fatty acid profile

PCs of 14 class models belonging to two harvest years were calculated independently and general statistics of each model was listed in Table 4.3. Among them, A2 and N2 models were preferred to be plotted against each other since they have different fatty acid profile. Inspection of Figure 4.3 revealed that using fatty acid profiles, each sample of A2 and N2 were correctly plotted in its critical limits apart from the origin which indicates quite good discrimination of two models. Also, most of the other models were identified as separate classes and placed in the outer space of A2 and N2 models. Only N1 plotted in the GE2 class model. Also M2 and AE1, having similar fatty acid compositions, can not be differentiated from each other. These observations revealed that harvest year and cultivar significantly influence fatty acid composition of olive oil. Usefulness of fatty acid composition in differentiation of Calabrian (Lanteri, et al. 2002) and Sicilian (D`Imperio, et al. 2007) cultivars was also demonstrated previously.

Table 4.3. PCA class models developed with fatty acid profile and general statistics of each class model

PCA class models	Number of PCs	R^2X(cum) (%)	PCA class models	Number of PCs	R^2X(cum) (%)
A1	3	94.9	E2	3	78.6
A2	4	97.4	N1	1	75.8
G1	2	88.9	N2	2	96.8
G2	3	92.7	AE1	2	84.2
M1	3	89.7	AE2	4	94.4
M2	3	92.8	GE1	3	94.7
E1	4	98.3	GE2	2	61.5

R^2X(cum): Cumulative sum of squares of all the X's explained by all extracted components

Also, effect of cultivar was studied by constructing A and G classes. Each class includes samples belonging to two different growing locations (Izmir and Edremit) and harvest years (2005/06 and 2006/07). PCA analysis was applied to each class. A class was described with 2 PCs accounting for 74.6% of total variation whereas G class was described with 7 PCs accounting for 99.6% of total variation. For A and G classes Coomans' plot is given in Figure 4.4. General examination of the Figure 4.4 reveals successful discrimination of A and G classes except some samples of A. So, regardless to growing location and harvest year, cultivar seems to have the ability in differentiation of olive oil samples.

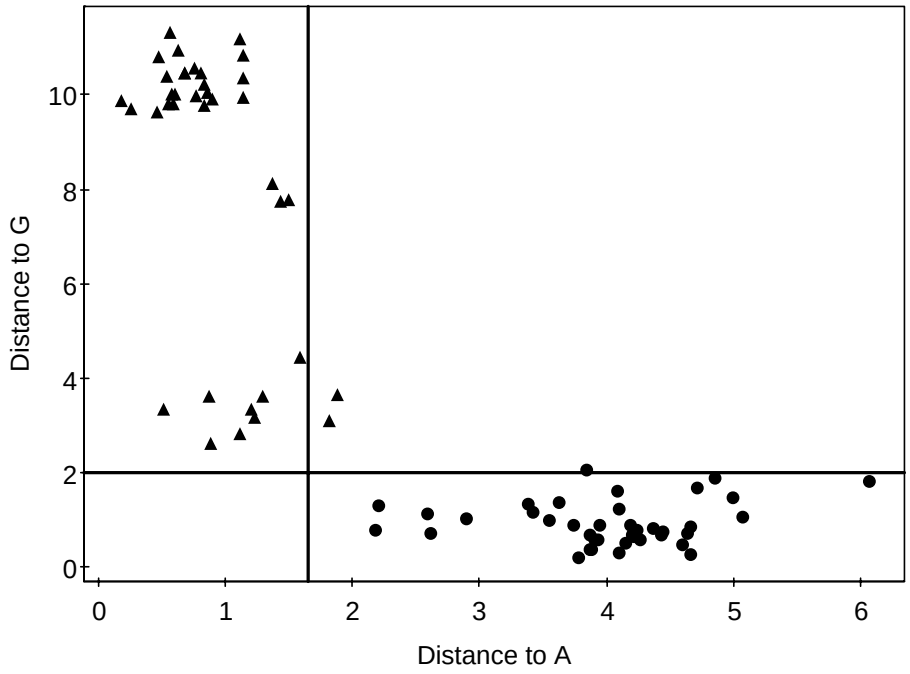


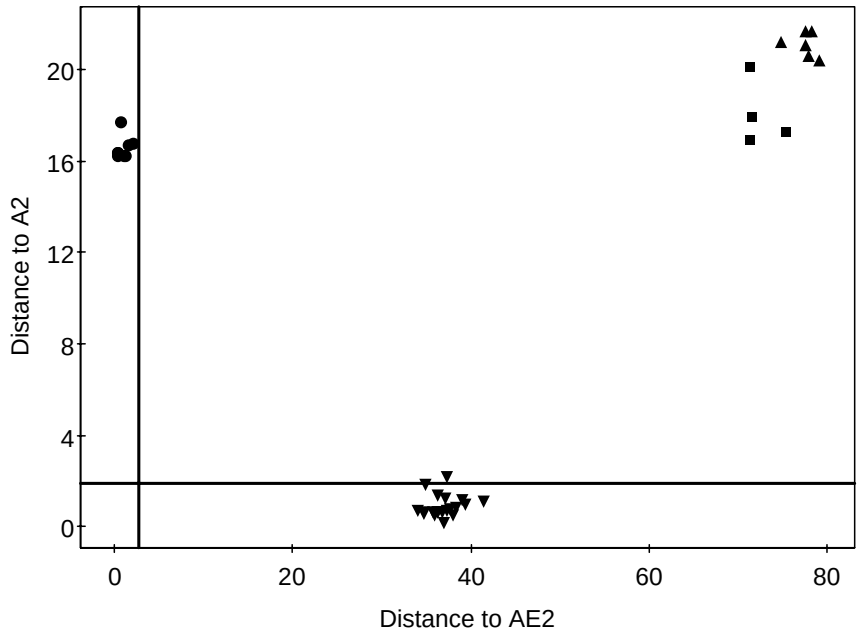
Figure 4.4. Coomans' plot for classification with respect to cultivar using fatty acid composition (▲: A, ●: G)

In order to investigate the effect of growing location on fatty acid profile, PCA was performed separately on varieties A and G from two different areas (Izmir and Edremit). The general statistics of each class models are presented in Table 4.4. Coomans' plot was constructed for both A-AE and G-GE including two consecutive harvest years (Figure 4.5).

Table 4.14. PCA class models developed for location effect with fatty acid profile and general statistics of each class model

PCA class models of A-AE	Number of PCs	R^2X(cum) (%)	PCA class models of G-GE	Number of PCs	R^2X(cum) (%)
A1	3	94.6	G1	3	96.3
A2	4	97.5	G2	3	93.4
AE1	2	84	GE1	4	87
AE2	4	92.7	GE2	2	84.2

(a)



(b)

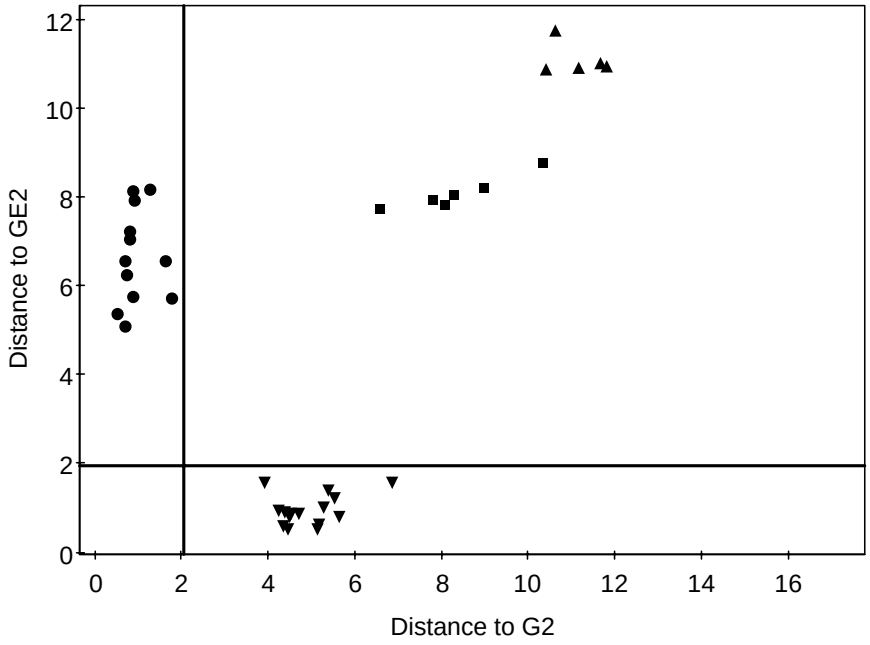


Figure 4.5. Coomans' plot for classification with respect to geographical origin and harvest year for (a) Ayvalık (b) Gemlik olive oil varieties using fatty acid composition (▲:A1-G1, ■:AE1-GE1, ●:A2-G2, ▼:AE2-GE2)

Figure 4.5(a) displays Coomans' plot in which A2 and AE2 plotted against each other using fatty acid profiles. Samples of 2nd harvest year, A2 and AE2, were accepted by their own class models. On the other hand, samples of 1st harvest year, A1 and AE1 were plotted among the critical limits of A2 and AE2 model. Similarly, quite successful differentiation of G2 and GE2 is observed in Figure 4.5(b). According to these observations, geographical origin and harvest year have significant influence on fatty acid profile of the olive oil. A previous study on Sicilian olive oils suggested that although the effect of the cultivar is predominant in the olive oil classification based on the fatty acid composition, a minor but well defined geographic effect is also present (D`Imperio, et al. 2007). In another study, Stefanoudaki, et al. (1999) was able to discriminate olive oil samples with respect to growing location based on altitude and also stated that at low-altitude areas, other environmental factors such as relative humidity and rainfall significantly differ.

4.1.2. Classification using FT-IR Data

FT-IR has become an emerging tool with the advantages of short analysis time and easy sample preparation for the authentication and classification of olive oil. Figure 4.6 illustrates typical olive oil spectra obtained in this study.

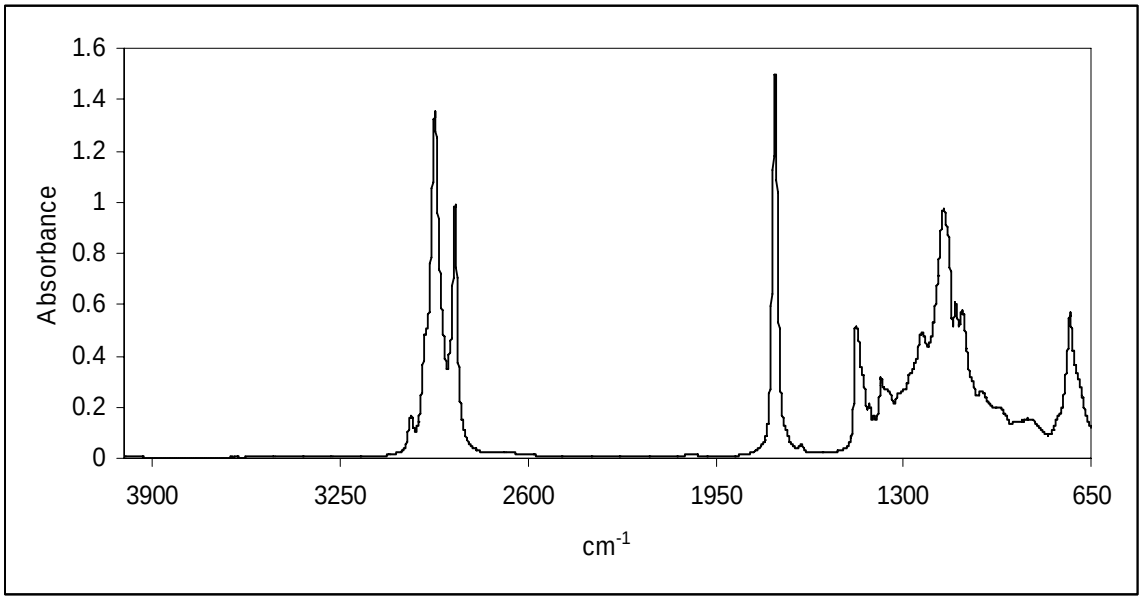


Figure 4.6. Typical olive oil FT-IR spectrum

In mid-IR spectrum, the peaks around 2950-2800 cm^{-1} region are due to C-H stretching vibrations of $–CH_3$ and $–CH_2$ groups. The large peak around 1745 cm^{-1} results from C=O double bond stretching vibration of carbonyl groups. Peaks around 1470-1200 cm^{-1} region corresponds to CH bending of $–CH_3$ and $–CH_2$. Fingerprint region lay between 1250-700 cm^{-1} which is due to stretching vibration of C-O ester group and CH_2 rocking vibration (Harwood & Aparicio 2000). The entire spectral profiles of each olive oil sample used in this study were similar. Thus, 3520-2520 cm^{-1} and 1875.5-675 cm^{-1} spectral ranges, where some differences were observed in the absorbance units, were used in data analysis. Wavelet analysis was applied as a compression technique to spectral data to increase the computational efficiency and also to enhance the classification studies. 125 observations were manipulated with PCA and score plot of first two PCs was illustrated in Figure 4.7. The first and second PCs described 49% and 18.5% of total variation, respectively.

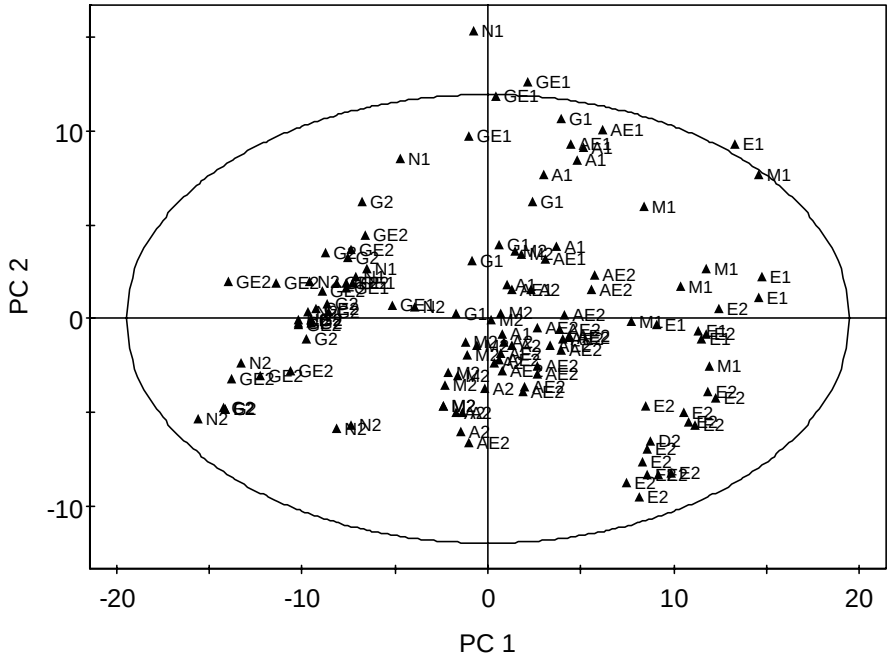


Figure 4.7. PCA score plot of the spectral data of lab scale extracted olive oil samples

Although samples of same varieties of 1st or 2nd harvest year such as E, GE and AE were grouped separately, most varieties could not be differentiated from each other. So, Coomans' plot was constructed to more apparently visualize the discrimination of samples with respect to varieties or years (Figure 4.8). 14 class models were developed

and overall statistics of these models was given in Table 4.5. Actually, each sample of A2 and N2 were placed on its own model region. However, in the outer space of A2 and N2 models, the only classification was observed for sample groups of AE2 and GE2, and most of the other samples were completely spread not showing any groupings.In an another study, Bendini, et al. (2007) used fingerprint region of FT-IR spectra with PCA to classify monovarietal olive oil samples from different regions of Italy belonging to only 2004 harvest year and was able to discriminate most of monovarietal virgin olive oil samples. To demonstrate the discrimination ability of harvest year, each variety belonging to two consecutive harvest years were also plotted against each other on the same graph and results revealed that all cultivars could be differentiated according to harvest year. Further application of PCA on spectral data of olive varieties for each harvest year separately reveals better differentiation of cultivars using Coomans' plot. Boggia, et al. (2002) stated that variation of climatic conditions from one year to the next affects the olive oil quality.

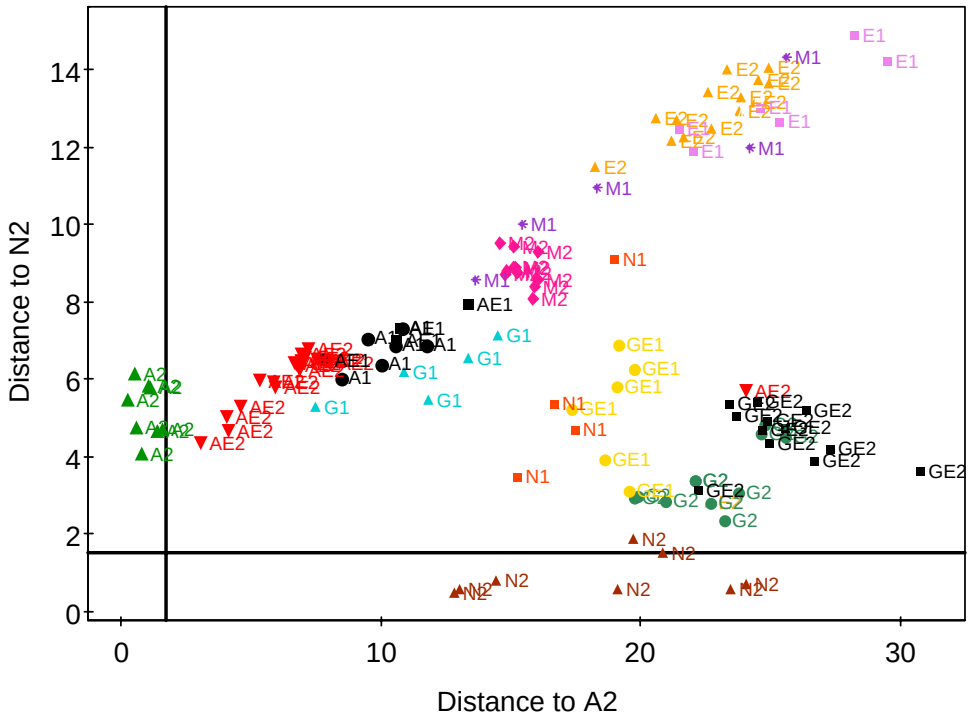


Figure 4.8. Coomans' plot for the classification of olive oil samples with respect to variety and harvest year using spectral data

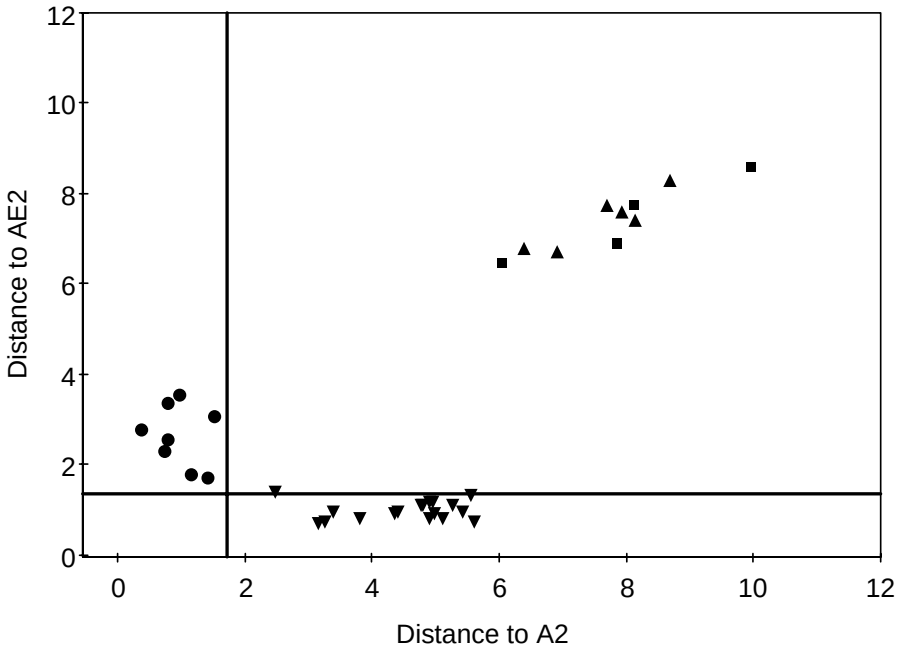
Table 4.5. PCA class models developed for FT-IR classification data and general statistics of each class model

PCA class models	Number of PCs	R^2X(cum) (%)	PCA class models	Number of PCs	R^2X(cum) (%)
A1	3	91.8	E2	5	95.5
A2	5	97	N1	3	93.1
G1	2	74.5	N2	2	96.8
G2	4	93.3	AE1	2	88.6
M1	3	94.1	AE2	5	91.8
M2	2	72.8	GE1	2	84.8
E1	4	96.9	GE2	3	85.6

In order to observe the effect of cultivar, A and G classed were created regardless to the growing location and harvest year. PCA was applied to each class and Coomans' plot was constructed. Since two classes could not be discriminated from each other, the plot was not given. In conclusion FT-IR spectra manipulated with PCA do not have the ability to differentiate the cultivars A and G regardless to the growing location and harvest year.

Figure 4.9 represents Coomans' plot constructed using mid-IR spectra to differentiate A & G varieties grown in two different regions. Each plot employed 4 class models and general statistics was given in Table 4.6.

(a)



(b)

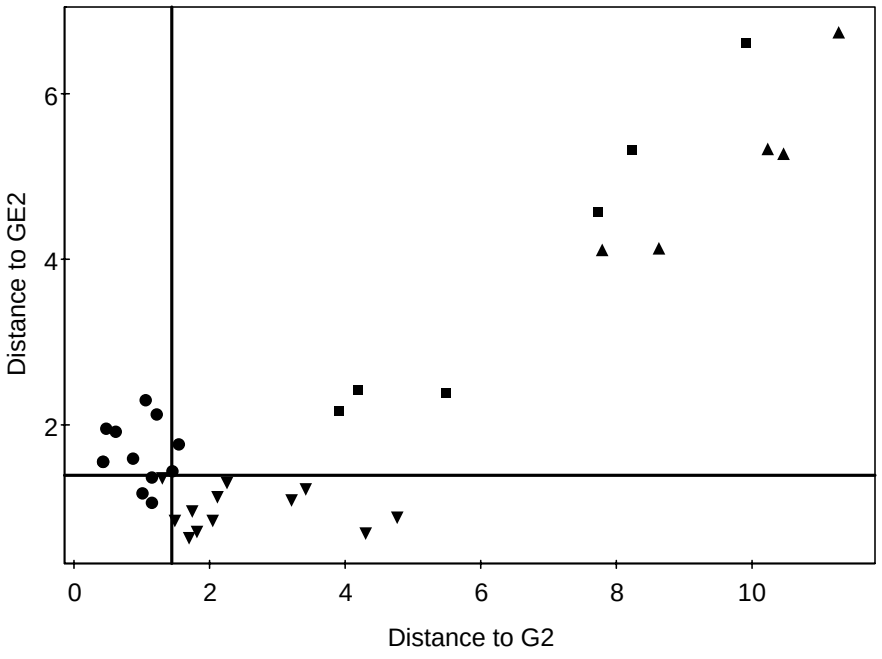


Figure 4.9. Coomans' plot for classification with respect to geographical origin and harvest year for (a) Ayvalık (b) Gemlik olive oil varieties using spectral data (▲: A1-G1, ■: AE1-GE1, ●: A2-G2, ▼: AE2-GE2)

Table 4.6. PCA class models developed for location effect with FT-IR data and general statistics of each class model

PCA class models of A-AE	Number of PCs	R^2X(cum) (%)	PCA class models of G-GE	Number of PCs	R^2X(cum) (%)
A1	3	89.5	G1	2	75.9
A2	5	94.6	G2	5	95.3
AE1	2	86.3	GE1	2	83.3
AE2	6	86.2	GE2	3	84.7

Although, most samples of the A2 and AE2 classified correctly (Figure 4.9(a)), some samples of G2 were plotted beyond its critical limits (Figure 4.9(b)). A clear discrimination can not be observed between A and AE, G and GE for the 1st harvest year. In another study (Tapp, et al. 2003), more distinct differentiation compared to our study was achieved using FT-IR, however olive oil samples from different countries which possessed more variable chemical properties were used.

Results show that application of PCA to fatty acid composition is quite successful for the classification of olive oil samples with respect to variety, geographical origin and harvest year. On the other hand, mid-IR spectra can not supply distinct varietal, geographical or seasonal grouping as much as fatty acid composition does. Nevertheless with the advantage of rapid and easy analysis ability, FT-IR can be used for the more general authentication issues with respect to harvest year and variety.

4.2. Classification of Commercial Olive Oils

Commercial olive oil samples belonging to North (N) and South (S) Aegean regions were obtained in 2005/06 (1st) and 2006/07 (2nd) harvest years. Table 4.7 presented free fatty acidity value of the samples. The acidity of the analysed samples was not higher than 1 except Men variety of the 1st harvest year whose acidity was determined as 1.2. On the other hand, fatty acid profile and spectral data implied differences among not only N and S geographical origins but also 1st and 2nd harvest years. Thus, chemometric analysis was applied on GC and FT-IR data to differentiate samples according to geographical origin and harvest year.

Table 4.7. Free fatty acid values of commercial olive oils expressed as % oleic acid

	2005-2006				2006-2007			
	Sample code	Acidity (% oleic)	Sample code	Acidity (% oleic)	Sample code	Acidity (% oleic)	Sample code	Acidity (% oleic)
North	Ez	0.44	Hav	0.52	Ez	0.42	Ayv	0.40
	Ez-org1	0.33	Bur	0.43	KucKuy	0.33	Altova	0.32
	KucKuy1	0.48	Gom	0.40	Altol	0.40	Zey	0.55
	KucKuy2	0.56	Ayv	0.60	Edr	0.41		
	Altol	0.38	Altova	0.57	Hav	0.36		
	Altol-sulbas	0.37	Zey	0.60	Bur	0.38		
	Edr	0.38			Gom	0.44		
South	Akh	0.86	Mil	0.86	Tep	0.22	Ort	0.25
	Men	1.20			Bay	0.83	Kosk	0.36
	Tep	0.31			Ode	0.34	Dal	0.28
	Bay	0.77			Tire	0.31	Koc	0.34
	Sel	0.38			Sel	0.46	Erb	0.39
	Ayd	0.84			Kus	0.59	Cine	0.26
	Ort	0.72			Ger	0.35	Mil	0.52
	Koc	0.84			Ayd	0.26		

4.2.1. Classification using GC Data

The fatty acid profiles of commercial olive oil samples belonging to 1st and 2nd harvest years were given in Table 4.8 and 4.9, respectively. Fatty acid content of each sample was the average of at least two GC measurements. For the 1st harvest year, content of oleic acid was the highest for Kuckuy1 (72.18%) and lowest for Men (69.82 %) while for 2nd harvest year it was at its highest amount for Tire (75.29%) and lowest amount for Kuckuy (70.20%). Inspection of Table 4.8 and 4.9 revealed that oleic acid was at higher amounts for the oils obtained from South than North for both 1st and 2nd harvest years. As the most abundant saturated fatty acid, palmitic acid was higher for North than South in both years. Also, content of linoleic acid was higher for 1st harvest year compared to 2nd year.

Table 4.8 . Fatty acid profiles (% of total fatty acids) of commercial olive oil samples belonging to 2005/06 harvest year

	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C20:0	C18:3	C22:0
North									
Ez	0.02±0.00	13.26±0.01	0.70±0.00	2.77±0.01	71.99±0.09	10.09±0.05	0.58±0.02	0.46±0.03	0.14±0.02
Ez-org	0.02±0.01	12.75±0.03	0.57±0.00	2.87±0.03	70.97±0.08	11.57±0.01	0.67±0.02	0.47±0.00	0.11±0.02
Kuckuy1	0.02±0.00	13.28±0.03	0.68±0.01	2.81±0.01	72.18±0.10	9.84±0.05	0.60±0.02	0.43±0.03	0.15±0.02
Kuckuy2	0.02±0.00	12.79±0.06	0.59±0.01	3.08±0.03	71.38±0.07	10.82±0.04	0.68±0.02	0.46±0.04	0.17±0.05
Altol	0.02±0.00	13.35±0.02	0.69±0.01	2.78±0.01	72.11±0.02	9.89±0.03	0.62±0.02	0.44±0.05	0.10±0.02
Altol-sb	0.02±0.01	12.40±0.03	0.60±0.01	3.00±0.03	71.82±0.05	10.87±0.02	0.63±0.01	0.48±0.02	0.17±0.03
Edr	0.02±0.00	13.32±0.02	0.68±0.00	2.80±0.01	72.04±0.09	9.92±0.04	0.61±0.01	0.45±0.00	0.17±0.04
Hav	0.02±0.00	12.92±0.01	0.59±0.00	3.19±0.01	71.45±0.06	10.50±0.06	0.69±0.00	0.50±0.03	0.14±0.01
Bur	0.02±0.00	13.14±0.05	0.62±0.01	2.93±0.02	71.87±0.02	10.11±0.02	0.65±0.02	0.49±0.03	0.17±0.05
Gom	0.02±0.01	13.16±0.02	0.63±0.01	2.95±0.02	71.86±0.14	10.12±0.07	0.65±0.02	0.46±0.02	0.14±0.01
Ayv	0.02±0.00	13.28±0.09	0.64±0.01	2.93±0.02	71.69±0.70	10.18±0.05	0.67±0.01	0.46±0.00	0.13±0.01
Altov	0.02±0.00	13.15±0.03	0.66±0.01	2.91±0.02	71.58±0.13	10.48±0.02	0.60±0.01	0.45±0.03	0.14±0.03
Zey	0.04±0.03	12.72±0.10	0.61±0.01	3.07±0.03	71.33±0.15	10.86±0.04	0.68±0.02	0.51±0.01	0.18±0.08
South									
Akh	0.02±0.00	13.19±0.22	0.68±0.04	3.28±0.52	71.81±0.20	9.75±0.55	0.68±0.03	0.45±0.00	0.13±0.01
Men	0.02±0.00	12.80±0.04	0.75±0.00	3.65±0.02	69.82±0.10	11.69±0.04	0.71±0.01	0.45±0.04	0.11±0.02
Tep	0.02±0.00	12.68±0.04	0.83±0.00	3.25±0.01	71.60±0.21	10.40±0.05	0.65±0.32	0.46±0.03	0.11±0.01
Bay	0.02±0.00	12.2±0.03	0.65±0.01	2.92±0.01	74.39±0.01	8.42±0.05	0.78±0.00	0.44±0.02	0.12±0.01
Sel	0.02±0.00	12.48±0.01	0.73±0.00	3.33±0.02	70.36±0.04	11.76±0.02	0.78±0.00	0.41±0.02	0.14±0.05
Ayd	0.02±0.00	12.63±0.13	0.70±0.02	3.06±0.04	72.60±0.27	9.63±0.04	0.76±0.01	0.47±0.02	0.14±0.03
Ort	0.02±0.00	12.31±0.07	0.76±0.01	3.28±0.01	73.03±0.15	9.34±0.07	0.72±0.02	0.42±0.01	0.12±0.06
Koc	0.02±0.00	12.26±0.30	0.69±0.01	3.39±0.18	73.51±0.38	8.87±0.12	0.67±0.01	0.46±0.03	0.12±0.04
Mil	0.02±0.00	12.06±0.05	0.69±0.01	2.70±0.02	72.99±0.11	10.13±0.02	0.83±0.01	0.44±0.01	0.13±0.02

Table 4.9. Fatty acid profiles (% of total fatty acids) of commercial olive oil samples belonging to 2006/07 harvest year

	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C20:0	C18:3	C22:0
North									
Ez	0.01±0.00	14.08±0.26	1.36±0.00	2.57±0.06	70.33±0.22	10.72±0.10	0.34±0.01	0.46±0.01	0.14±0.01
Kuckuy	0.02±0.00	12.99±0.01	1.08±0.00	2.77±0.01	70.20±0.06	11.93±0.05	0.37±0.00	0.48±0.01	0.16±0.03
Altol	0.02±0.00	13.48±0.12	1.17±0.01	2.72±0.02	71.03±0.16	10.47±0.03	0.43±0.00	0.51±0.01	0.16±0.01
Edr	0.02±0.00	13.09±0.25	1.10±0.01	2.76±0.09	71.65±0.29	10.34±0.06	0.39±0.00	0.50±0.00	0.15±0.01
Hav	0.02±0.00	12.62±0.05	0.99±0.01	2.87±0.03	71.74±0.05	10.71±0.07	0.39±0.00	0.50±0.01	0.15±0.00
Bur	0.02±0.00	12.92±0.06	1.03±0.00	2.85±0.02	70.92±0.11	11.22±0.05	0.39±0.00	0.50±0.01	0.15±0.01
Gom	0.02±0.00	13.18±0.22	1.04±0.01	2.92±0.09	70.51±0.26	11.26±0.07	0.39±0.00	0.51±0.02	0.16±0.02
Ayv	0.02±0.00	13.49±0.04	1.17±0.02	2.74±0.03	70.73±0.12	10.84±0.01	0.39±0.00	0.49±0.00	0.14±0.01
Altov	0.02±0.00	13.16±0.26	1.17±0.02	2.78±0.08	71.71±0.52	10.10±0.32	0.42±0.01	0.49±0.01	0.15±0.02
Zey	0.02±0.00	12.61±0.01	1.06±0.02	2.69±0.01	71.99±0.02	10.62±0.01	0.38±0.00	0.48±0.01	0.15±0.00
South									
Tep	0.02±0.00	12.76±0.50	1.19±0.01	2.70±0.12	74.30±0.46	7.89±0.15	0.48±0.01	0.50±0.00	0.16±0.00
Bay	0.02±0.00	11.49±0.04	1.04±0.00	2.46±0.02	75.17±0.03	8.79±0.04	0.47±0.00	0.44±0.00	0.12±0.00
Ode	0.02±0.00	11.50±0.08	1.13±0.01	2.75±0.04	74.13±0.12	9.41±0.04	0.48±0.00	0.46±0.01	0.13±0.01
Tire	0.02±0.00	12.03±0.07	1.09±0.01	2.35±0.02	75.29±0.10	8.16±0.04	0.49±0.01	0.44±0.01	0.13±0.01
Sel	0.02±0.00	12.36±0.23	1.14±0.03	2.76±0.12	74.21±0.60	8.46±0.69	0.45±0.02	0.47±0.01	0.15±0.03
Kus	0.02±0.00	11.59±0.03	0.90±0.00	2.81±0.01	74.61±0.05	9.07±0.04	0.43±0.00	0.46±0.00	0.12±0.01
Ger	0.02±0.00	12.77±0.13	1.26±0.01	2.72±0.05	72.86±0.20	9.25±0.06	0.51±0.00	0.48±0.01	0.13±0.01
Ayd	0.02±0.00	12.73±0.81	1.11±0.03	2.99±0.23	73.24±0.89	8.78±0.18	0.47±0.02	0.50±0.03	0.15±0.03
Ort	0.02±0.00	12.62±0.40	1.27±0.15	2.67±0.21	73.76±0.29	8.56±0.61	0.49±0.01	0.47±0.01	0.13±0.01
Kosk	0.02±0.00	11.51±0.03	1.12±0.01	2.74±0.02	74.14±0.06	9.40±0.03	0.48±0.02	0.46±0.00	0.12±0.01
Dal	0.02±0.00	11.50±0.02	1.06±0.01	2.83±0.01	74.63±0.06	8.89±0.05	0.48±0.00	0.47±0.01	0.13±0.00
Koc	0.02±0.00	12.10±0.03	1.09±0.01	2.76±0.00	74.18±0.05	8.80±0.05	0.46±0.01	0.47±0.02	0.12±0.01
Erb	0.02±0.00	11.52±0.07	1.07±0.03	2.75±0.02	74.36±0.1	9.22±0.06	0.48±0.00	0.45±0.02	0.13±0.01
Cine	0.02±0.00	12.08±0.57	1.09±0.03	2.73±0.13	74.28±0.50	8.75±0.02	0.46±0.01	0.47±0.02	0.12±0.00
Mil	0.02±0.00	12.81±0.82	1.15±0.01	2.69±0.17	72.19±0.81	10.06±0.19	0.48±0.02	0.47±0.03	0.12±0.02

The data including 9 GC variables were submitted to PCA to visualize the presence of principal groupings. The PC modelling resulted in a model with 2 PCs where first and second PCs explaining, 33.8% and 26.7% of total variation, respectively. Figure 4.10 is the score plot of olive oil samples belonging to N and S regions obtained at two consecutive harvest years. The discrimination of olive oil samples with respect to geographical origin was very clear for 2nd harvest year. However, some samples of S1 can not be differentiated from sample group of N1 in score plot. On the other hand, harvest year seemed to have quite good ability on differentiation. Variation of chemical composition in two consecutive years can be due to changes in climatic conditions (rainfall, temperature and humidity) (Aparicio and Luna 2002).

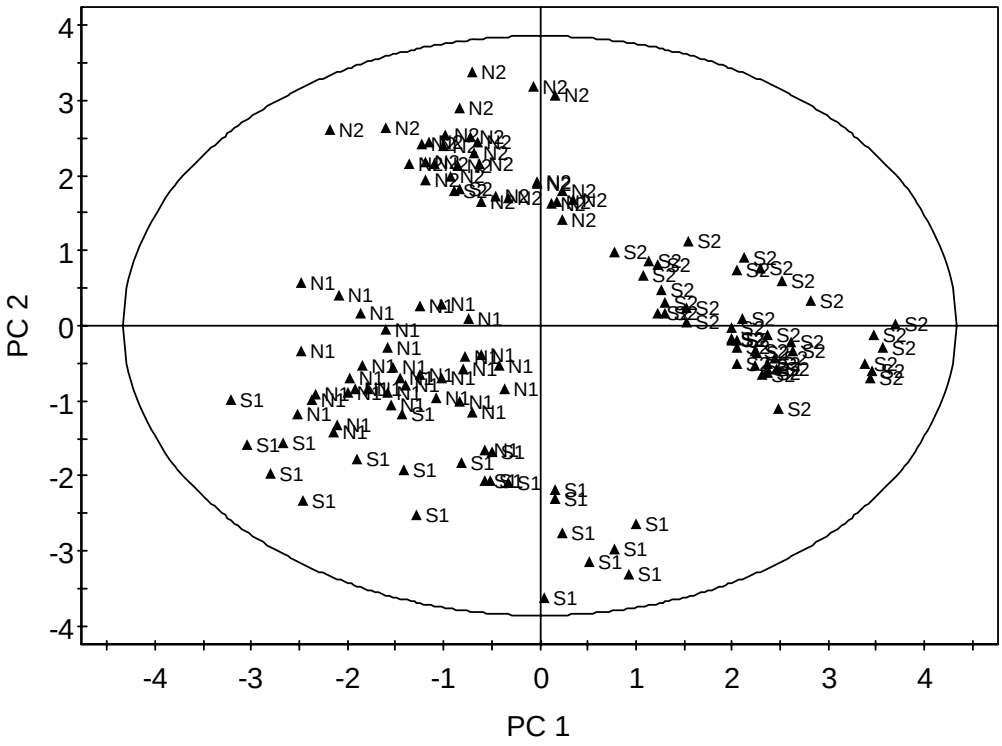


Figure 4.10. PCA score plot of the fatty acid profile of commercial olive oil samples (N1: olive oils from north in 2005/06, S1: olive oils from south in 2005/06, N2: olive oils from north in 2006/07, S2: olive oils from south in 2006/07)

Figure 4.11 showed the loading plot of first 2 PCs. In the first component, dominating fatty acids were C18:1 and C18:2 which were responsible from the

separation of N from S. First PC with the contribution of second PCs brings the information of variables C16:1 and C20:0 and these fatty acids were encoded for harvest year discrimination. Thus, the model interpretation suggested that samples of S had higher C18:1 content compared to samples of N. Also, C16:1 was lower for 2nd harvest year, whereas C20:0 was higher for the same year.

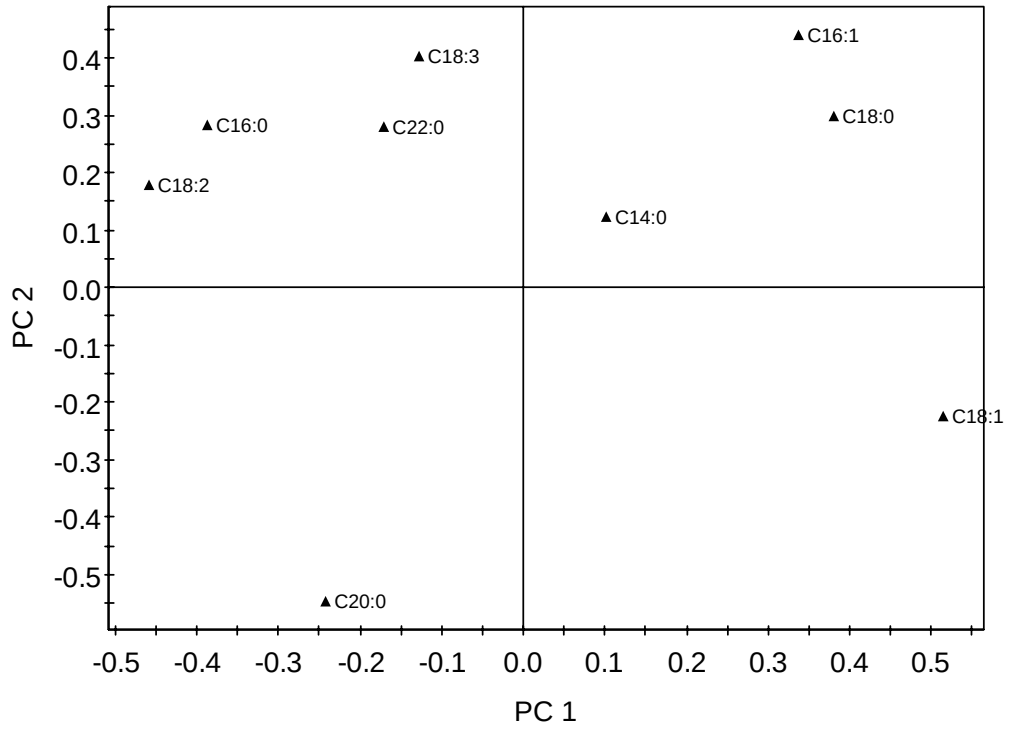


Figure 4.11. PCA loading plot of the fatty acid profile of commercial olive oil samples

Coomans' plot was constructed for the classification of samples using fatty acid profiles (Figure 4.12). Four PCA classes (N1, S1, N2 and S2) were calculated independently and N2 and S2 as the class were plotted against each other. The number of PCs and R^2X (cum) of the each PCA class model were given in Table 4.10. Examination of Coomans' plot revealed that only few samples belonging to N2 and S2 were drawn in the outside of their critical limits. Samples of N1 and S1 were identified as two different classes correctly placed in the outer space of N2 and S2 models. According to this plot, analysis of fatty acids with PCA has the ability to discriminate oil samples with respect to geographical origin and harvest year. A study on Cornicabra virgin olive oil also revealed significant statistical differences in quality indices, major

fatty acids and sterol compositions with respect to the year of production, with the exception of total phenols (Salvador, et al. 2003).

Table 4.10. PCA class models for fatty acid profile of commercial olive oils and general statistics of each class model

PCA class models	Number of PCs	R^2X(cum) (%)
N1	3	79.9
N2	3	83.9
S1	2	59
S2	3	81.9

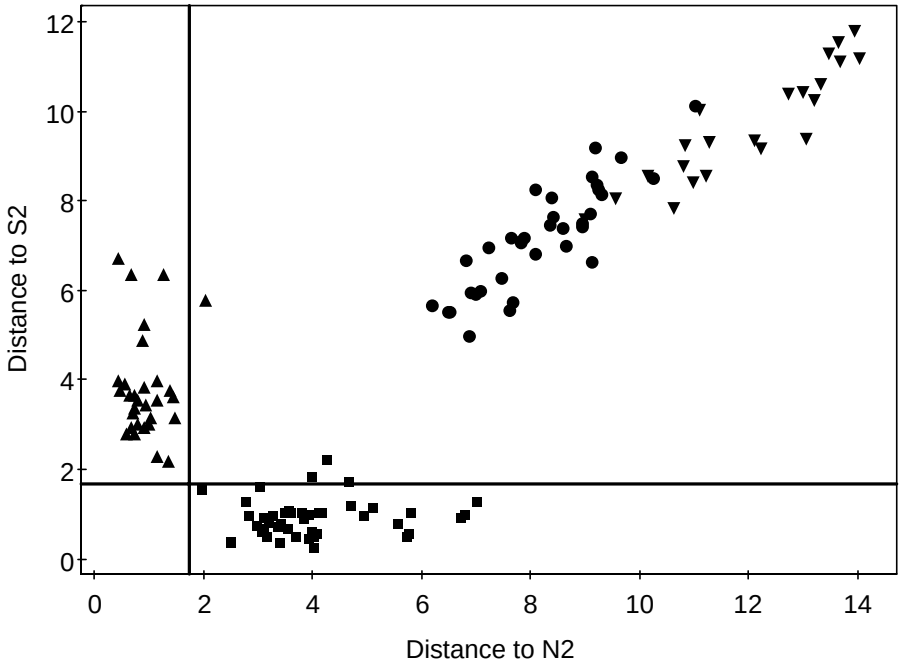


Figure 4.12. Coomans' plot of the fatty acid profile of commercial olive oils for the discrimination with respect to geographical origin and harvest year (▲: N2, ■: S2, ●: N1, ▼: S2)

4.2.2. Classification using FT-IR Data

The entire range of spectra looks almost similar for each olive oil sample unless one observes very closely. Multivariate data analysis is therefore required to extract the relevant information from these spectra. Spectral compression (Wavelet analysis) was applied to reduce the size of spectral data which increase the efficiency of the model. PCA was performed with data set containing 115 observations and score plot was constructed (Figure 4.13).

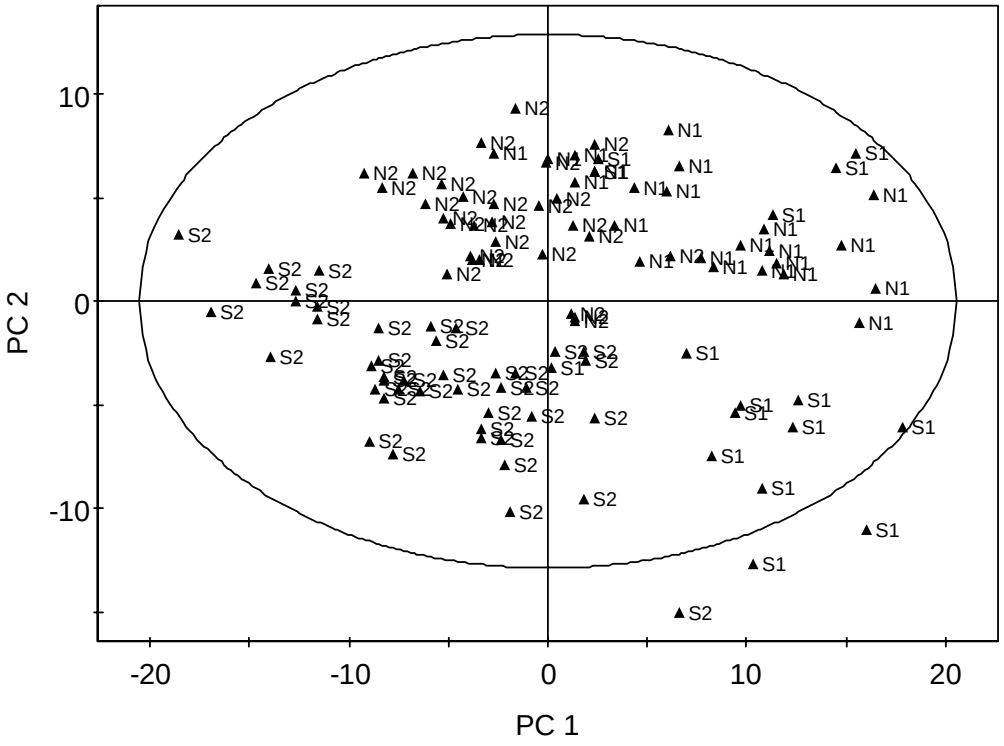


Figure 4.13. PCA score plot of spectral data of commercial olive oil samples (N1: olive oils from north in 2005/06, S1: olive oils from south in 2005/06, N2: olive oils from north in 2006/07, S2: olive oils from south in 2006/07)

First and second PCs described 37.8% and 13.3% of total variation, respectively. A careful analysis of the graph allowed separation of N from S with some exceptions of S placed in between the samples of N. On the other hand samples of 1st harvest year were generally plotted in the left whereas 2nd harvest year placed in the right side of the scores plot. Coomans' plot was constructed to represent the discrimination of samples

with respect to variety and harvest year (Figure 4.14). The plot consisted of N1, S1, N2 and S2 PCA class models where N2 and S2 as the class models placed against each other. The number of PCs and R^2X (cum) of the each PCA class model were given in Table 4.11.

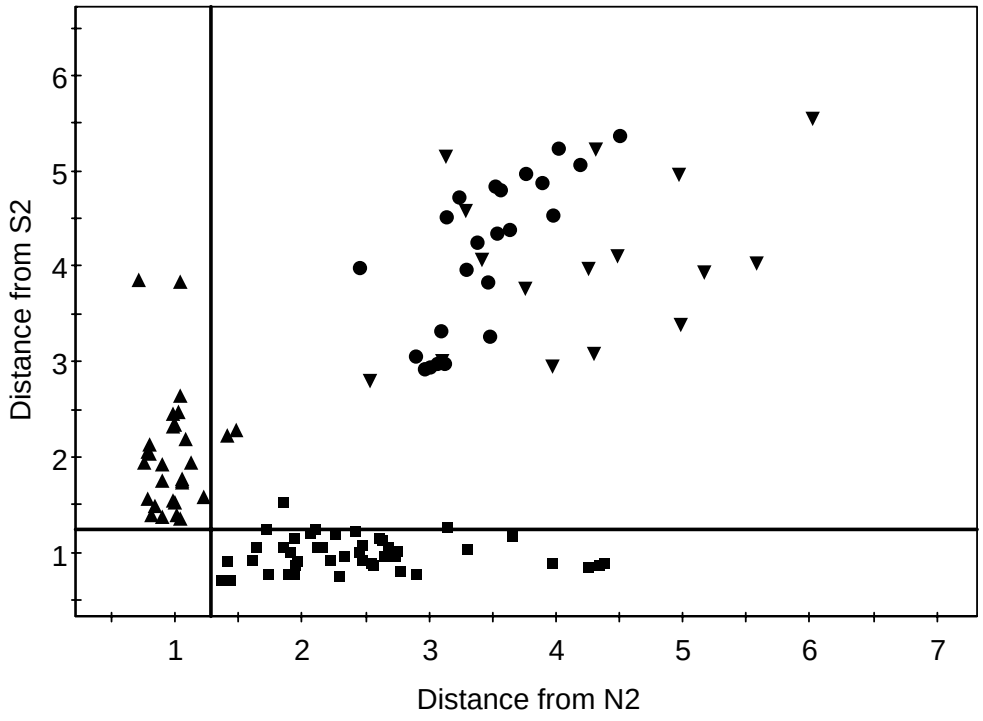


Figure 4.14. Coomans' plot of the FT-IR spectra for the discrimination with respect to geographical origin and harvest year (▲:N2, ■:S2, ●:N1, ▼:S2)

Analysis of Figure 4.14 revealed that discrimination of N2 and S2 was achieved except some samples. The samples of N1 and S1 were correctly plotted beyond the critical limits of N2 and S2 models but they can not be differentiated from each other very well.

Table 4.11. PCA class models and general statistics of each class model

PCA class models	Number of PCs	R^2X (cum) (%)
N1	6	82.9
N2	6	81.7
S1	6	85.7
S2	6	81.7

Although the ability of FT-IR for the differentiation of olive oil samples with respect to geographical origin is low, different harvest years can be sufficiently identified. When compared with spectral data fatty acid profile obtained from GC supplied more discrete chemical information for the classification of commercial olive oils with respect to geographical origin.

4.3. Relation between FT-IR Profile-Free Fatty Acid Value and FT-IR Profile-Fatty Acid Profile Obtained with GC Measurement

Each olive oil type has its own fatty acid profile and spectral property. It is stated that spectral features of oils vary with the degree of unsaturation (Harwood and Aparicio 2000). In addition, FFA content is one of the most important factors for the determination of quality and economic value of olive oils. The importance of providing fast and accurate methods to identify quality parameters such as fatty acid profile and FFA content has been recently stressed (Iñón, et al. 2003). FT-IR is a rapid analysis technique that implies the differences between chemical composition and structures of oil samples. Therefore, manipulation of FT-IR data with chemometrics was studied for the determination of FFA values and fatty acid profile of olive oil samples.

PLS analysis was applied to data to check whether there is a correlation between FT-IR spectra vs. fatty acid profile and FFA value of olive oils obtained with traditional analysis techniques. The spectral data vs. fatty acid profile and FFA value of both extracted and commercial olive oil were included in the analysis. The data set was divided into calibration and validation sets with 60 and 27 observations, respectively. In order to enhance the predictive power of the PLS models, OSC was applied to remove systematic variation in X (spectral data) that is not related with Y (fatty acid

percentage). Wavelength compression was also selected to increase the efficiency and speed. Using PLS algorithm, fatty acid profile and FFA of each olive oil sample was predicted relating spectral data (X variable) with fatty acid data and FFA (Y variables) obtained with analytical methods. The PLS regression analysis was resulted in 6 PCs explaining 79.2% of total variation (Y) with a predictive ability of 70%. As the major fatty acid of olive oil ranging between 66-75% in our samples predicted and actual percentages of C18:1 for calibration and validation sets was illustrated in Figure 4.15. The R^2 (cal) is 0.98 which indicated good prediction of C18:1 percentage from spectral data. In order to test the predictive ability, the values of validation set were predicted using the proposed model. Although R^2 value (0.93) of the validation set was lower, quite good prediction of oleic acid amount (%) was achieved with the proposed PLS model.

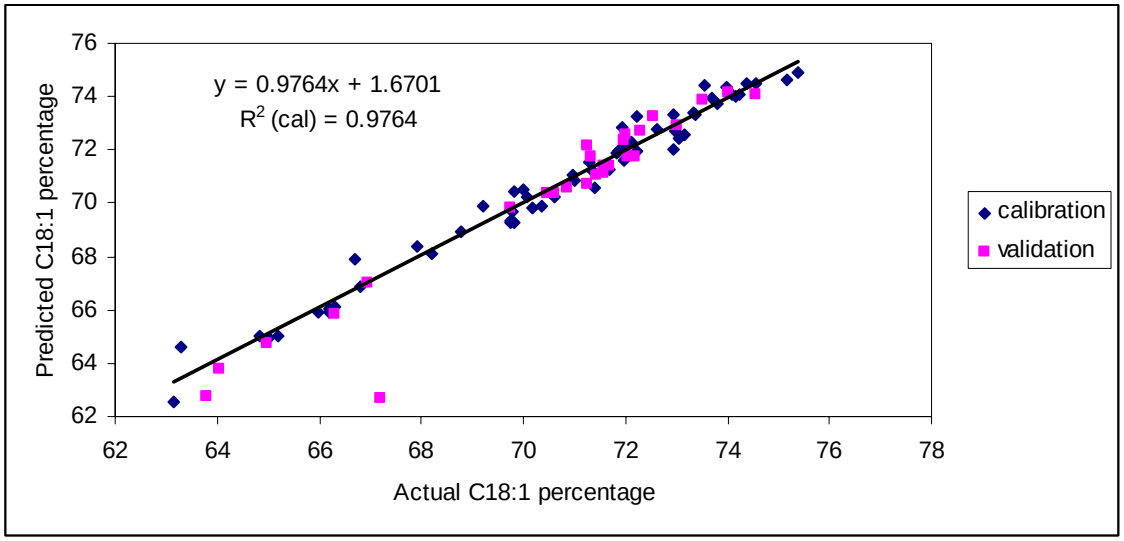


Figure 4.15. PLS regression of actual vs. predicted oleic acid percentage of the calibration set and validation sets (R^2 (cal): regression coefficient of the calibration set)

C16:0 constitutes the highest proportion of saturated fatty acids ranging from 11.5 to 16.5% of total fatty acids of the analysed samples. The observed vs. predicted C16:0 percentage of the calibration determined using PLS regression, was shown in Figure 4.16. The predictive ability of PLS model for C16:0 was lower than C18:1. Also,

obtained PLS regression for C16:0 was tested with validation set (Figure 4.16) and resulted in R^2 value of 0.71.

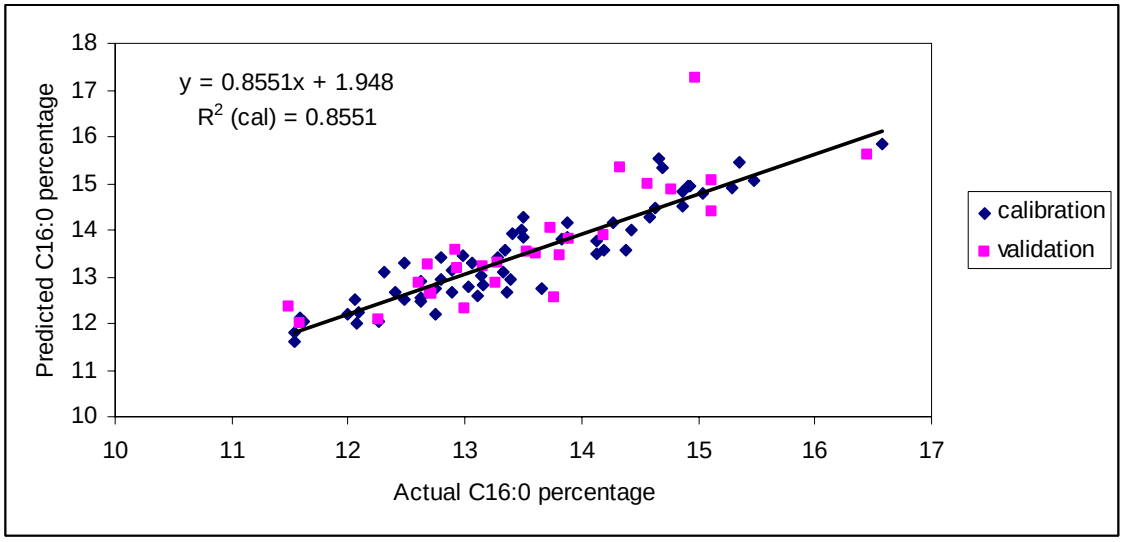


Figure 4.16. PLS regression of actual vs. predicted palmitic acid percentage of the calibration and validation sets

The regression curve was plotted and validation step was applied for each fatty acid. The R^2 values of calibration set and error criterions (SEC, SEP and REP) were given in Table 4.12. REP value was selected as the error criterion to decide applicability of spectral data in determination of fatty acid content. Best results were obtained for the MUFAs and C18:1 with REP value lower than 2%. MUFAs are important because of their nutritional implication (Aguilera, et al. 2005) and C18:1 is the characteristic MUFA of olive oil. Also, REP was lower than 5% for C16:0 which is the highest amount of saturated fatty acid of olive oil. Good predictions were achieved for C18:2 and PUFA with REP values around 5%. However REP values were quite high for other fatty acids. Higher amount fatty acids seemed to have high R^2, SEC and SEP values whereas their REP values were quite low. So, they performed better in terms of PLS model efficiency compared to lower amount fatty acids. In overall consideration, MUFAs constituting highest proportion can be predicted best from the MIR data followed by SFAs and PUFAs. Galtier and Dupuy (2007) reported similar results using NIR data for fatty acid profile determination of virgin olive oils.

Table 4.12. Summary of proposed PLS model of the regression between FT-IR spectra and fatty acid profile for both calibration and validation sets

Fatty acid	Fatty acid range %	R^2 (cal)[a]	R^2 (val)[b]	SEC[c]	SEP[d]	REP[e]
C14:0	0.01-0.03	0.25	0.22	0.0033	0.0085	22.23
C16:0	11.49-16.51	0.85	0.71	0.43	0.66	4.87
C16:1	0.57-2.65	0.93	0.81	0.13	0.23	19.32
C18:0	2-4.57	0.83	0.55	0.23	0.41	16.63
C18:1	63.57-75.29	0.97	0.93	0.48	0.97	1.39
C18:2	7.41-16.89	0.97	0.93	0.38	0.66	6.04
C20:0	0.34-0.86	0.69	0.64	0.07	0.07	14.87
C18:3	0.32-0.52	0.83	0.56	0.06	0.07	14.89
C22:0	0.11-0.18	0.11	0.18	0.02	0.02	13.63
SFA[f]	14.4-20.55	0.97	0.79	0.37	0.6	3.53
MUFA[g]	64.72-76.2	0.97	0.94	0.49	0.86	1.21
PUFA[h]	7.68-17.74	0.98	0.94	0.37	0.63	5.53

[a]PLS regression correlation coefficient for calibration samples, [b]PLS regression correlation coefficient for validation samples, [c]standard error of calibration, [d] standard error of prediction, [e]relative error of prediction, [f]saturated fatty acids, [g]monounsaturated fatty acids, [h]polyunsaturated fatty acids

The PLS regression model of FFA values ranging from 0.3 to 1.2 was given in Figure 4.17. Using the calibration set, PLS model resulted in R^2 and SEC values given as 0.89 and 0.08, respectively. These values indicated good prediction of FFA value from spectral data. The validation of the proposed PLS regression for FFA was accomplished by evaluation of the validation data set. The results of the PLS regression of the validation set was also illustrated in Figure 4.17. R^2 and SEP values were determined as 0.84 and 0.085, respectively. These values indicated successful presentation of FFA values using FT-IR spectra.

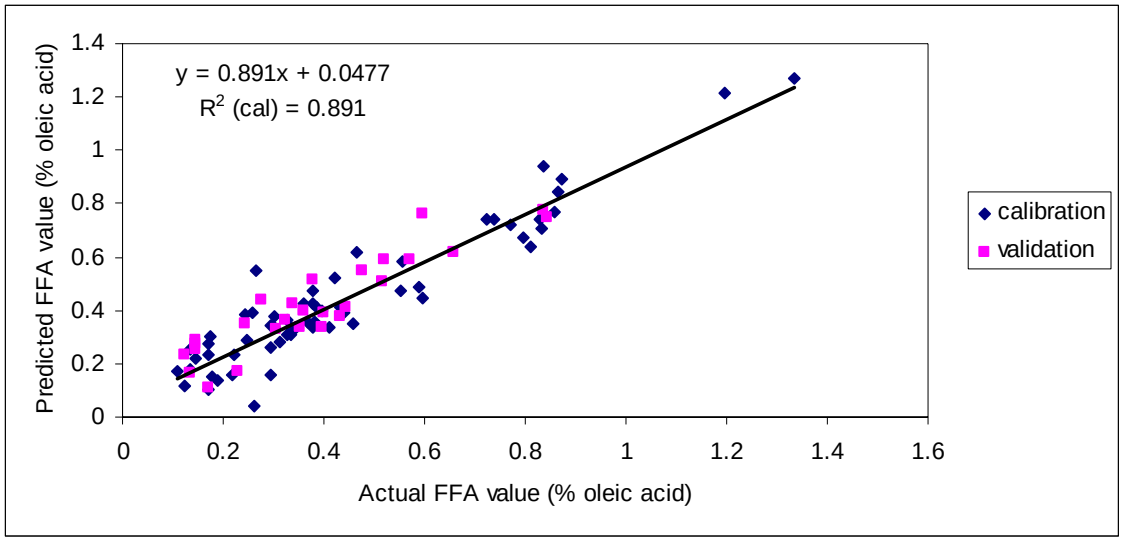


Figure 4.17. PLS regression of actual vs. predicted FFA of the calibration and validation sets

In other studies specific regions of FT-IR spectra were used for the FFA determination and similar results were obtained. In the previous studies, wider % free fatty acid range was employed by adding oleic acid to olive oil samples (Iñón, et al. 2003; Bertran, et al. 1999, Bendini, et al. 2007, Ismail, et al. 1993). According to the results reported in this study and obtained by other authors also, use of FT-IR spectroscopy with chemometrics is determined as a good alternative to standard methods for FFA percentage analysis of olive oils belonging to different categories and origin.

Multivariate analysis of FT-IR data has the potential to predict some of the fatty acids and FFA value of the virgin olive oil samples. Therefore, with the advantage of simple and rapid sample analysis ability, FT-IR is applicable for quick determination of characteristic fatty acids and FFA value.

4.4. Determination of Adulteration

Adulteration was performed either by mixing N monovariety with AE and E monovarieties or by mixing hazelnut, sunflower-corn binary mixture, cottonseed and rapeseed oil with olive oil. PCA was applied to differentiate non-adulterated olive oil from adulterated oils. Scores and Coomans' plot were constructed to visualize

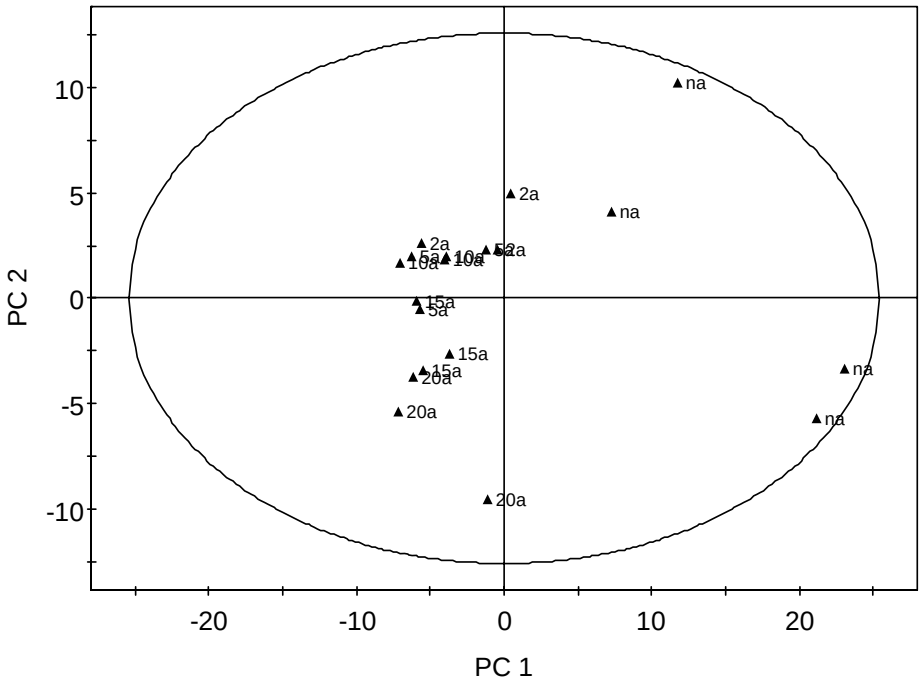
differentiation. Moreover, quantification of adulterant in the mixture was determined by PLS. To test the performance of PLS, cross validation was conducted for monovarietal olive oil adulteration due to low number of samples. On the other hand, calibration model was developed and tested by validation set for hazelnut, sunflower-corn, cottonseed and rapeseed adulteration.

4.4.1. Monovarietal Olive Oil Adulteration

For monovarietal adulteration, blends of varieties AE-N and E-N were prepared. Each cultivar yields high amounts of oil (monovarietal) and AE variety has an important economic potential with its high quality oil in Turkish market and is preferred by consumers due to its sensory characteristics. E has also a promising economic value. While AE and E are widely cultivated in West part of Turkey, N is a cultivar grown in Southeast region of Turkey and has different sensory characteristics than AE and E.

PCA was performed initially to extract information and to examine the qualitative differences of samples. There were two distinct data sets containing 19 and 20 observations belonging to spectral data of AE-N and E-N mixtures, respectively. Wavelet analysis was performed for compression of the spectral data containing more than 4000 variables to increase the efficiency and computational speed. Scores plots for the first two PCs explaining 72.1% of total spectral variation for AE-N model and 77.9% of total spectral variation for E-N model were constructed (Figure 4.18). According to Figure 4.18(a), samples of pure AE were distributed in the right while AE-N mixtures were placed in the left side. Similar grouping was also observed for E-N model except for samples containing 2% N which overlap with pure samples (Figure 4.18(b)). Thus in the further analysis 2% N was excluded from observations of E-N mixture.

(a)



(b)

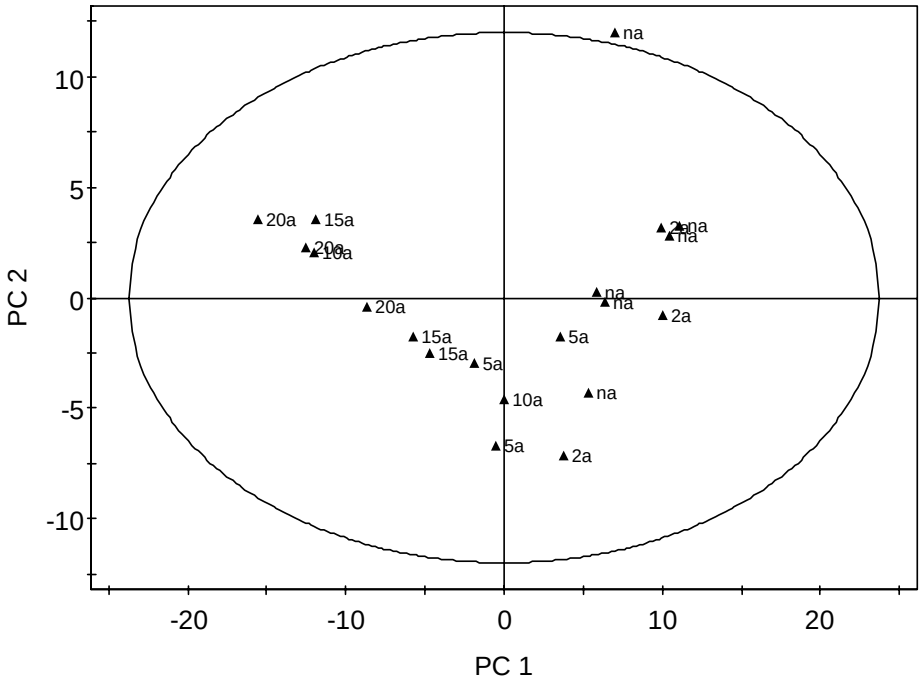


Figure 4.18. PCA scores plot of (a) AE-N and (b) E-N mixtures (na indicates non-adulterated samples, numbers near a notation indicates adulteration %)

Figure 4.19(a) represents Coomans' plot of AE-N model with 4 PCs for adulterated and 2 PCs for non-adulterated classes describing 79% and 86.9% of total variation, respectively. AE samples were successfully discriminated from samples of AE-N mixtures. On the other hand, PCA of E and E-N class models yielded 3 PCs explaining 92% and 4 PCs explaining 96.7% of total spectral variation, respectively. Differentiation of E from E-N mixture was represented in Figure 4.19(b). Samples of AE-N mixture were placed in its region far from the samples of AE which exhibits better differentiation than E-N model. This could be due to a more pronounced difference in chemical structure composition of AE and N.

(a)

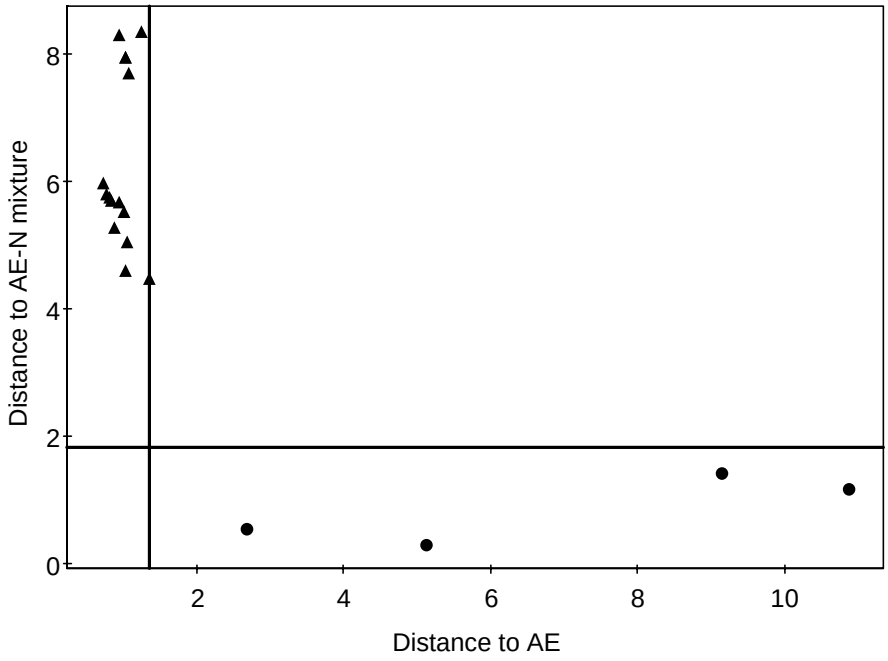


Figure 4.19. Coomans' plot for the discrimination of (a) AE from AE-N mixture (b) E from E-N mixture (● AE and E ▲AE-N and E-N mixture)

(cont. on next page)

(b)

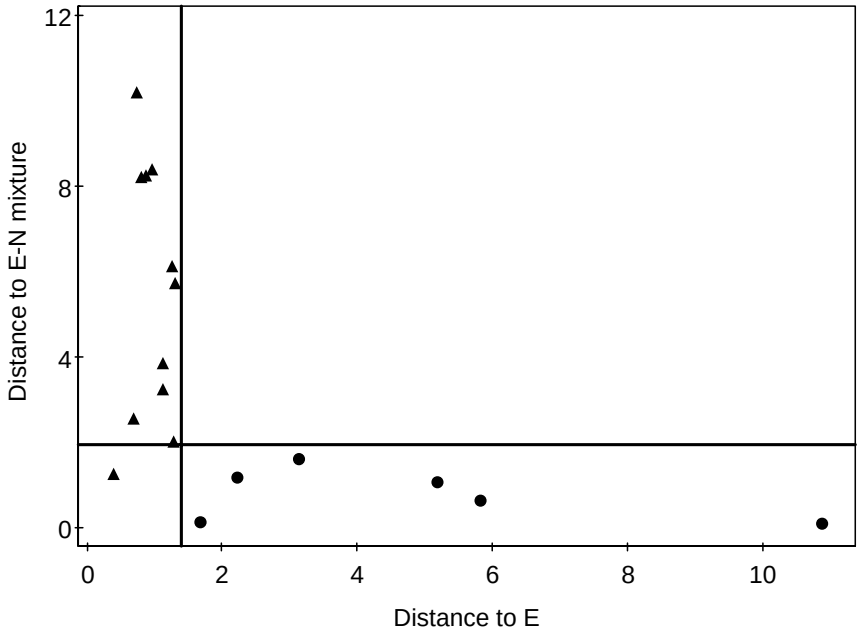
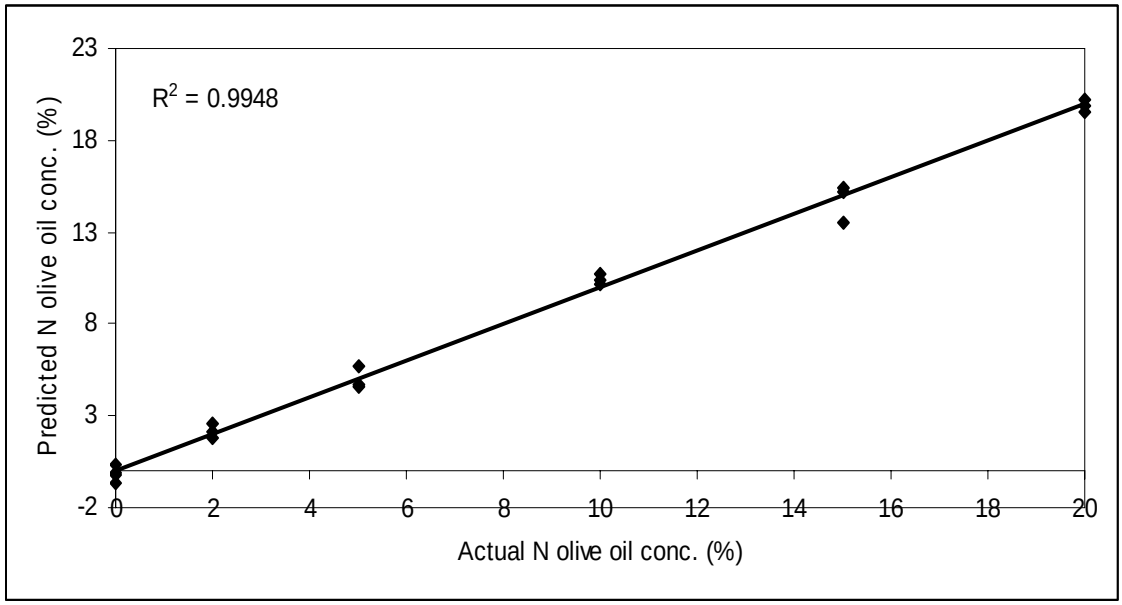


Figure 4.19. (cont.) Coomans' plot for the discrimination of (a) AE from AE-N mixture (b) E from E-N mixture (● AE and E ▲AE-N and E-N mixture)

Quantification of the percentage of N in AE-N and E-N oil mixtures was performed using PLS algorithm. The data set to be analysed contains 19 and 17 observations belonging to AE-N and E-N models, respectively. The predictive ability and R^2 values of the models were sufficient. However, to enhance the predictive power of the PLS models, OSC was applied to remove systematic variation in X (spectral data) that is not related with Y (percent adulteration), and also wavelength compression was selected to increase the efficiency and speed. The PLS regression analysis resulted in 2 PCs explaining 99.5% and 99.4% of variation (Y) with a predictive ability of 98.6% and 96.9% for AE-N and E-N data sets, respectively. The Figure 4.20 shows the concentration values obtained from PLS models versus actual concentration of N in E and AE.

(a)



(b)

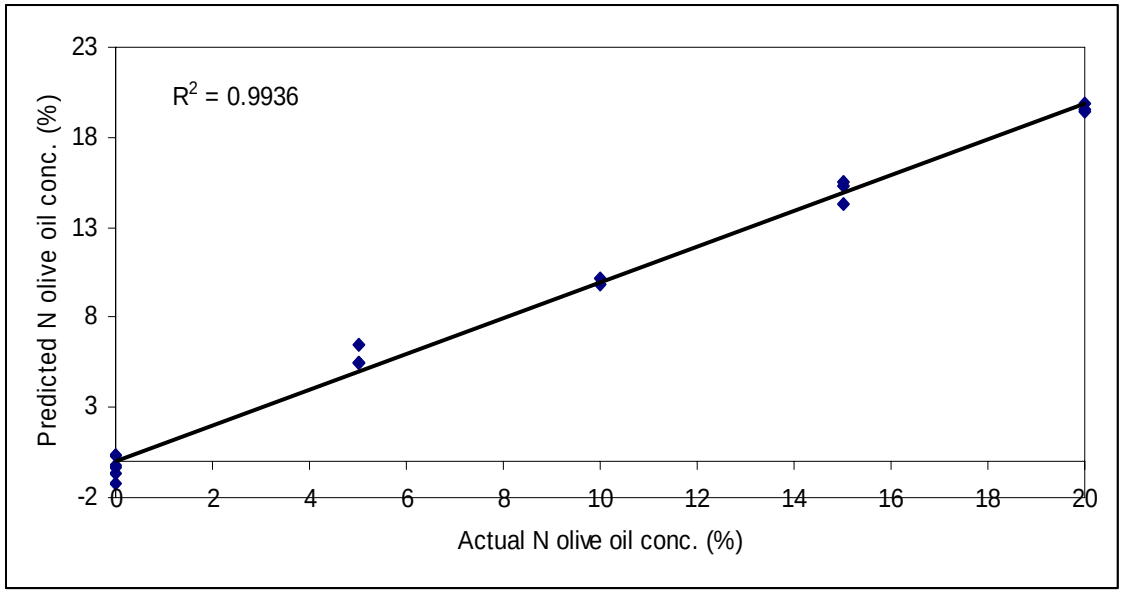


Figure 4.20. PLS regression of predicted vs. actual N olive oil content in (a) AE-N and (b) E-N mixtures

The difference between the actual concentration and the predicted N concentration is small and correlation coefficients, R^2, for the observed vs. predicted curves are 0.99 for N in AE and 0.99 for N in E. In order to validate the developed models, cross validation was applied by removing one sample at a time. The removed sample was predicted with model created using the remaining observations and the procedure was repeated until each sample was excluded once. N content (%) in AE and E predicted from PLS model using cross validation was given in Table 4.13. SEP and R^2 parameters were employed to evaluate the goodness of fit of the cross validation data set, and these values were calculated with respect to the results of cross validation. Cross validation results of AE-N model yielded 0.9 as SEP value and 0.98 as R^2 value whereas E-N model resulted in 1.22 as SEP and 0.97 as R^2 values. High coefficient of determination and low SEP values indicate success of the PLS regression models. The PLS regression model appears to have a reasonable ability to estimate the N percentage in AE even at percentages about 5. Also N amount in E could be determined at a level ≥ 5%.

Table 4.13. Predicted (obtained by cross validation) and actual N content in AE and E samples using PLS model

Actual N content [%] in AE	Predicted N content [%] in AE	Actual N content [%] in E	Predicted N content [%] in E
0	-0.55	0	-1.86
0	-0.46	0	-1.03
0	4.05	0	0.51
0	3.9	0	-0.13
2	2.63	0	4.12
2	2.83	0	-0.64
2	1.79	5	5.53
5	5.94	5	5.57
5	4.32	5	6.93
5	5.50	10	9.85
10	10.19	10	10.22
10	10.42	15	15.60

(cont. on next page)

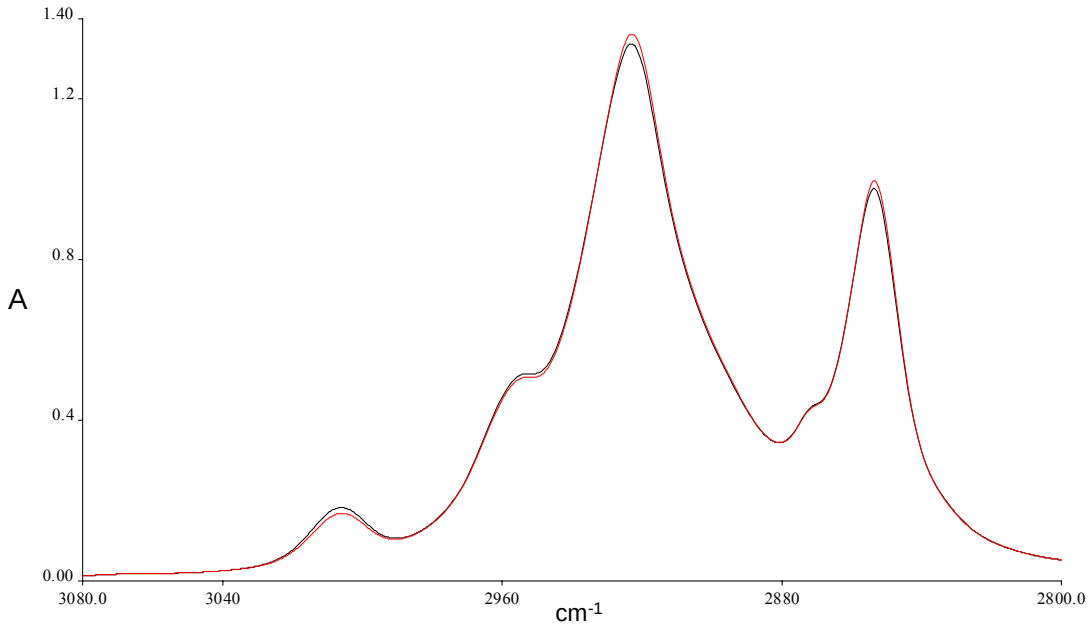
Table 4.13. (cont.) Predicted (obtained by cross validation) and actual N content in AE and E samples using PLS model

Actual N content [%] in AE	Predicted N content [%] in AE	Actual N content [%] in E	Predicted N content [%] in E
10	11.45	15	14.22
15	15.17	15	15.33
15	14.78	20	19.32
15	12.75	20	19.85
20	19.42	20	19.26
20	19.46		
20	19.16		

4.4.2. Hazelnut Oil Adulteration of Olive Oil

The olive oil and hazelnut oil spectra were given in Figure 4.21. As olive oil and hazelnut oil sample looks similar along the entire wavelength range, distinct spectral zones were shown in Figure 4.21 to better visualize the differences of two spectra. However, any notable difference between olive and hazelnut oil spectra can not be observed because they have quite similar chemical profile. As the spectra of hazelnut and olive oil can not be differentiated to the naked eye, PCA was employed to be able extract relevant information from spectral data.

(a)



(b)

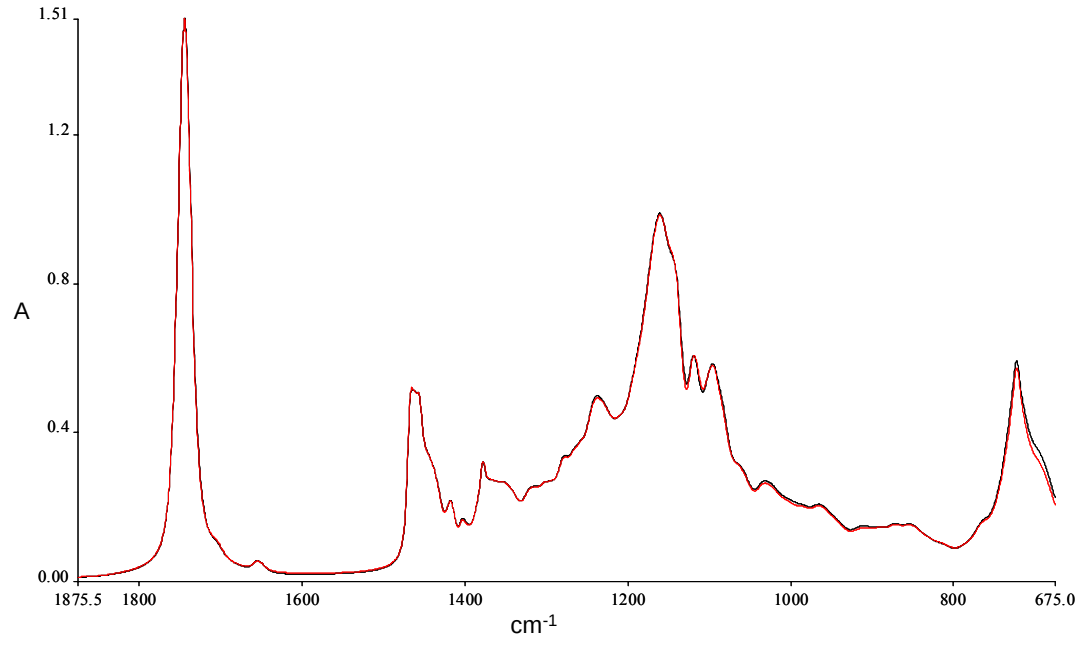


Figure 4.21. The spectra of olive oil and hazelnut oil around (a) 3080-2800 cm^{-1} (b) 2875.5-675 cm^{-1} regions (— olive oil, — hazelnut oil)

PCA was applied to data set containing 148 observations including both non-adulterated and hazelnut oil adulterated olive oil samples. Non-adulterated samples included

commercial olive oil samples of both North and South Aegean region. Adulterated samples were obtained by adulterating North and South olive oil between 2-50%. Wavelet analysis was applied for the compression of spectral data. Scores plot with 2 PCs describing 75% of total variance was constructed (Figure 4.22).

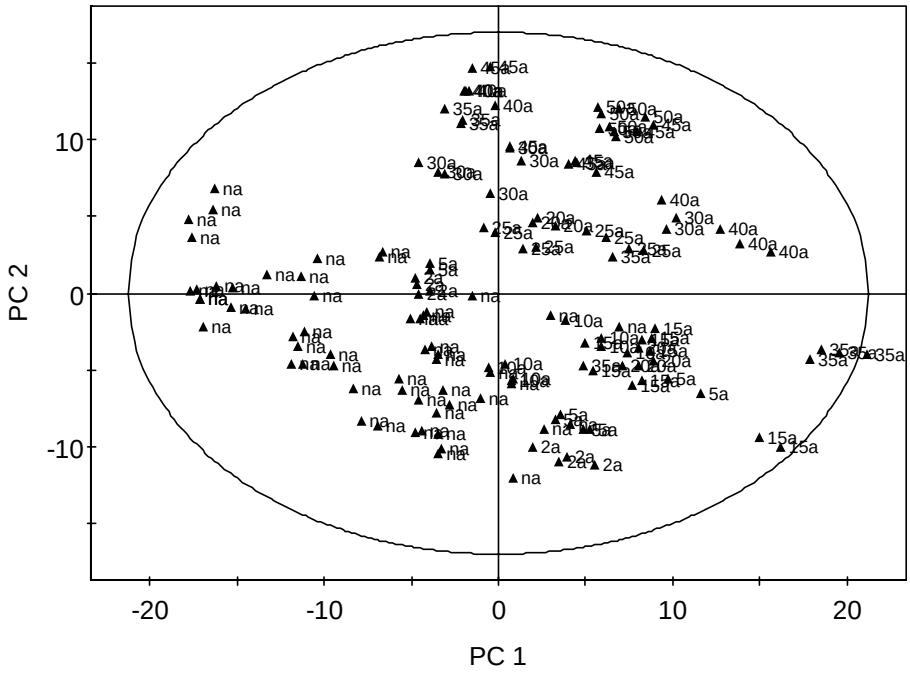


Figure 4.22. PCA scores plot of non-adulterated olive oil (na) and adulterated olive oil (a) samples (numbers near a notation indicates adulteration %)

A careful examination of Figure 4.22 revealed distinct grouping of non-adulterated and adulterated samples except for the samples containing hazelnut oil lower than 10%. Therefore, 2 and 5% adulterated samples were excluded and Coomans' plot was constructed to observe the differentiation of non-adulterated olive oil class with 4 PCs accounting for 0.79 of total variance from adulterated olive oil samples with 4 PCs accounting for 0.83 of total variance (Figure 4.23).

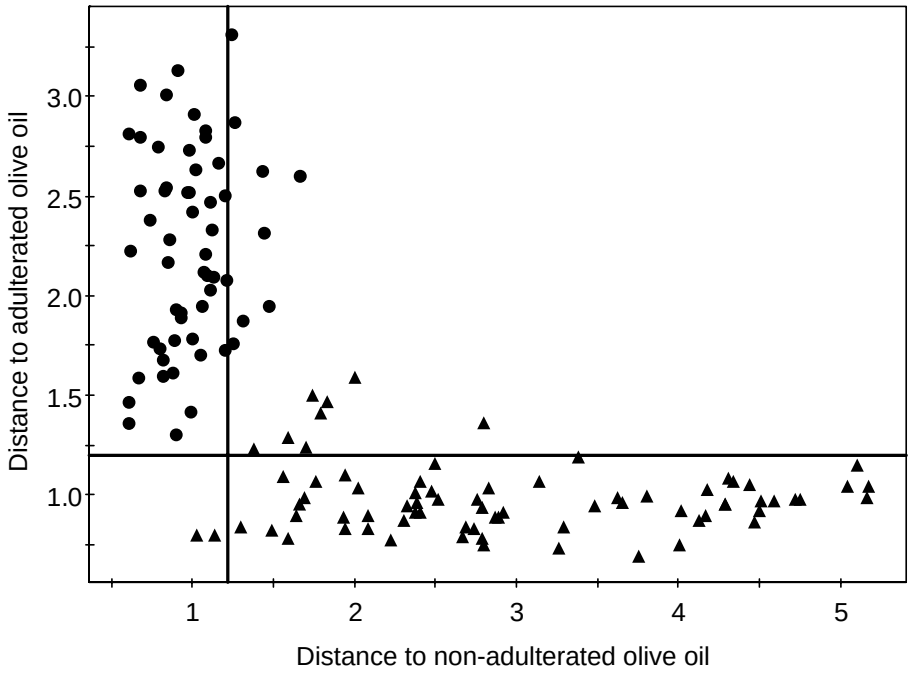


Figure 4.23. Coomans' plot for the classification of olive oil and hazelnut adulterated olive oil samples (● olive oil ▲ hazelnut oil adulterated olive oil)

Most samples of pure and adulterated olive oil were correctly plotted in their critical limits. However some samples were placed in the region where they could be classed as neither pure nor adulterated oil. Ozen and Mauer (2002) also studied hazelnut oil adulteration of olive oil with FT-IR spectra in combination with chemometrics but could not detect olive oil adulteration with 5-20% hazelnut oil. In another study, hazelnut adulteration at a level >8% could be detected analysing unsaponifiable matter with FT-MIR in combination with stepwise linear discriminant analysis (Beaten, et al. 2005).

Quantification of hazelnut oil content in adulterated oil samples was performed using PLS algorithm with the same data set. However, as pointed in PCA part, samples containing 2 and 5% hazelnut oil were excluded from the observations. The samples of all the adulterated and pure olive oils were randomly divided into a calibration and a validation set. OSC was applied to remove systematic variation in X (spectral data) that is not related with Y (percent adulteration) and also wavelength compression was selected to increase the efficiency and speed. Calibration set consisted of 104 samples while the validation set included 31 samples. The model was fitted successfully using 2

PCs which explained 98.6% variation of Y (% hazelnut oil) with a predictive ability of 98.4%. The predicted and actual hazelnut oil amount was presented in Figure 4.24. The calculated R^2 and SEC values were 0.99 and 2.1, respectively. These values indicated good prediction of hazelnut oil percentage from spectral data. The validation of the proposed PLS regression was accomplished by evaluation of the validation data set (Figure 4.24). R^2 and SEP values, as an indication of goodness of the validation, were determined as 0.97 and 3.1. Peña, et al. (2005) was able to detect and quantify the hazelnut oil adulteration at a level higher than 15% in commercial olive oil by direct coupling of HS-MS with multivariate regression techniques and error criterion of the prediction set was determined as 1.3.

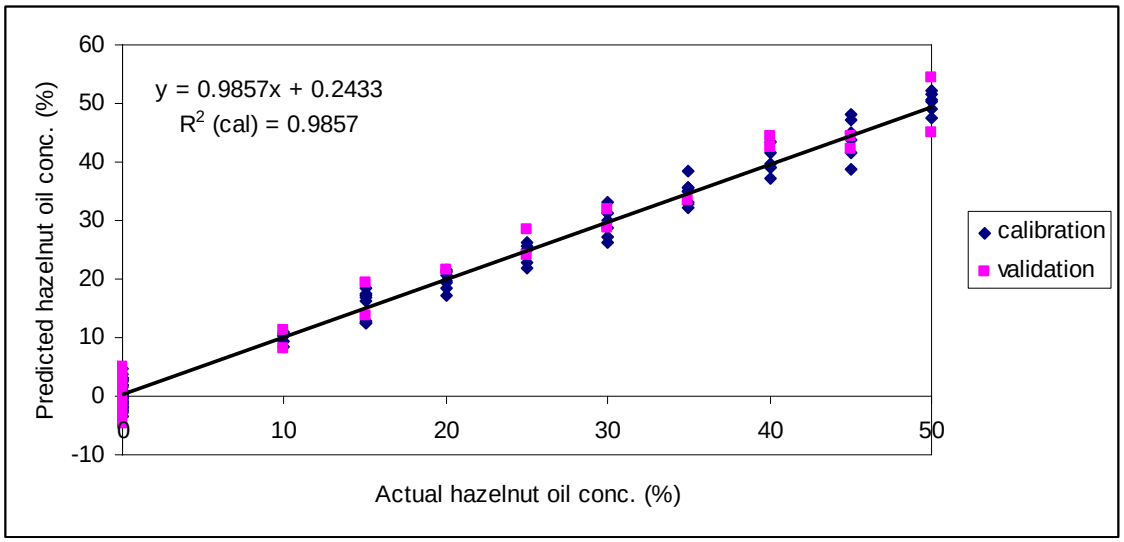


Figure 4.24. PLS regression of actual vs. predicted hazelnut oil concentration in olive oil for calibration and validation sets

4.4.3. Adulteration of Olive Oil with Binary Oil Mixture

The spectra of olive oil, corn oil and sunflower oil are provided in Figure 4.25. Whole spectra are shown in two distinct wavelength intervals to visualize the spectral differences. It is possible to observe differences in the spectra of three oil samples. Especially olive oil spectra can easily be differentiated by naked eye from corn and sunflower oil samples.

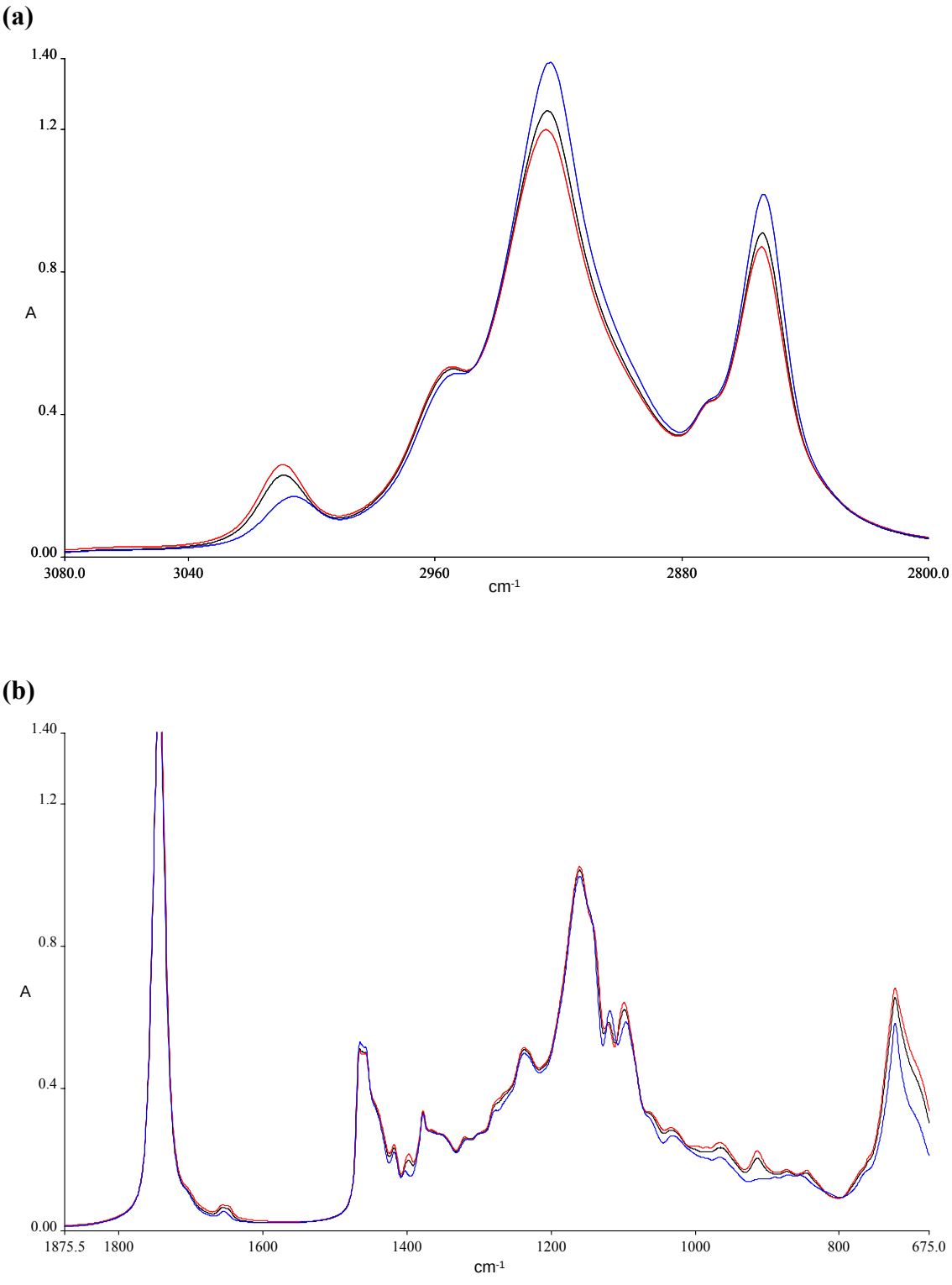


Figure 4.25. The spectra of olive oil, corn oil and sunflower oil around a) 3080-2800 cm^{-1} b) 2875.5-675 cm^{-1} regions (— olive oil, — corn oil, — sunflower oil)

PCA was applied to spectral data to discriminate sunflower-corn oil adulterated olive oil samples from non-adulterated olive oil samples. Non-adulterated samples included commercial olive oil samples belonging to North and South Aegean regions whereas for adulteration only some of North olive oil samples were used. Both discrimination and quantification studies were performed using total percentage of sunflower-corn oil mixture as a single adulterant. As the first step, wavelet analysis was performed on spectral data to increase the efficiency of the PCA model. The data set contains 167 observations. Score plot for the first 2 PCs explaining 84.7% of total spectral variance was illustrated in Figure 4.26.

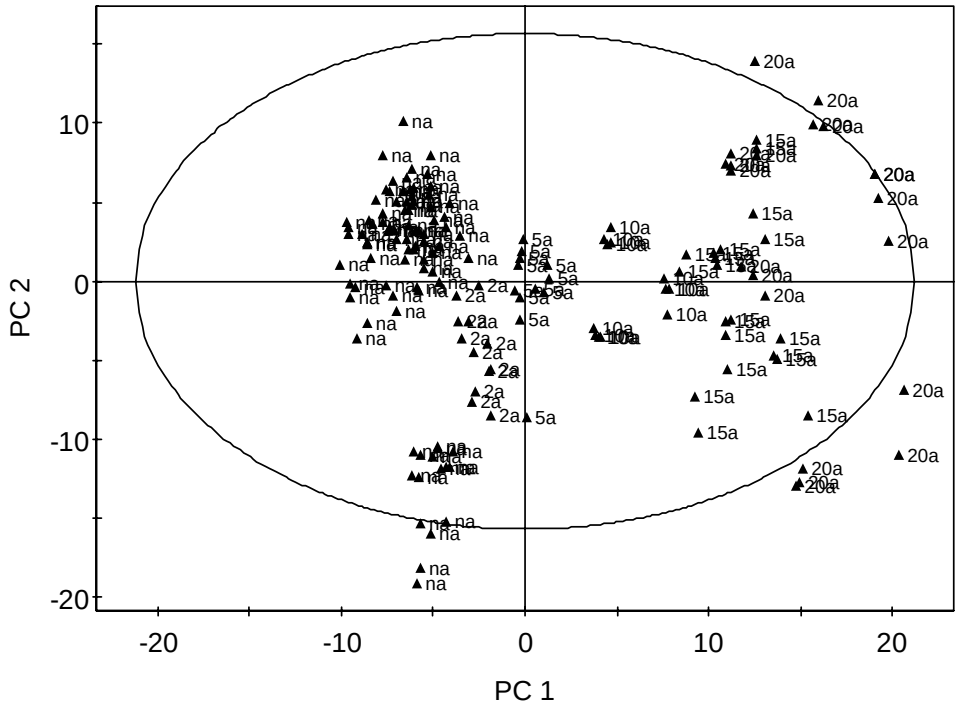


Figure 4.26. PCA scores plot of non-adulterated olive oil (na) and sunflower-corn adulterated olive oil (a) samples (numbers near a notation indicates adulteration %)

The scores of non-adulterated and 2% adulterated samples were placed together in the left whereas scores of ≥5% adulterated samples were located in the right side, representing a visual discrimination except 2% adulterated oil samples. Thus, Coomans' plot was constructed excluding 2% adulterated samples. Developed model had 4 PCs

for both non-adulterated and adulterated models describing 88.3% and 94.3% of variation, respectively (Figure 4.27).

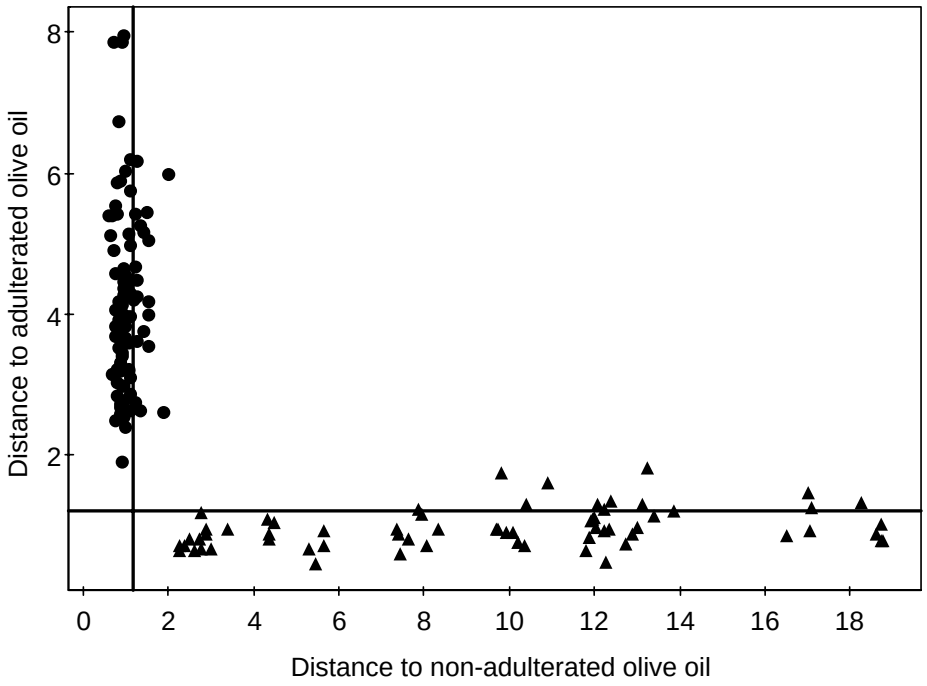


Figure 4.27. Coomans' plot for the classification of olive oil and sunflower-corn adulterated olive oil samples (●: olive oil, ▲: adulterated olive oil)

Coomans' plot revealed that most samples of each class were placed in their region. According to scores and Coomans' plot, a general discrimination of adulterated samples from non-adulterated oil samples were achieved except 2% adulterated olive oil. There were other studies involving the use of sunflower-olive oil and corn-olive oil binary blends for the adulteration detection with FT-IR. Tay, et al. (2002) was able discriminate olive oil samples from sunflower adulterated samples at levels between 2-10%. Also, results of FT-IR analysis indicated that the detection limit for olive oil adulteration was 9% if the adulterant is corn oil while it was lower (6%) if the adulterant is sunflower oil (Vlachos, et al. 2006).

PLS analysis was also applied to same data set used in PCA for quantification of sunflower-corn oil mixture in olive oil. As previously mentioned, clear differentiation of samples containing 2% sunflower-corn mixture could not be achieved. Therefore, these samples were excluded from the PLS model. The data set was divided into calibration

and validation sets containing 112 and 43 observations, respectively. The spectral data was manipulated with OSC and wavelength compression before PLS analysis. The PLS regression model of calibration set with 2 PCs can explain 99% of variation with a predictive ability of 98.9%. The R^2 value of PLS regression curve (Figure 4.28) was 0.99, and SEC value was determined as 0.78. These results indicated good predictive ability of the calibration model. Also, using the PLS regression equation the samples of validation set were predicted (Figure 4.28). R^2 and SEP values were calculated as 0.98 and 1.04 exhibiting high predictive ability of the proposed model.

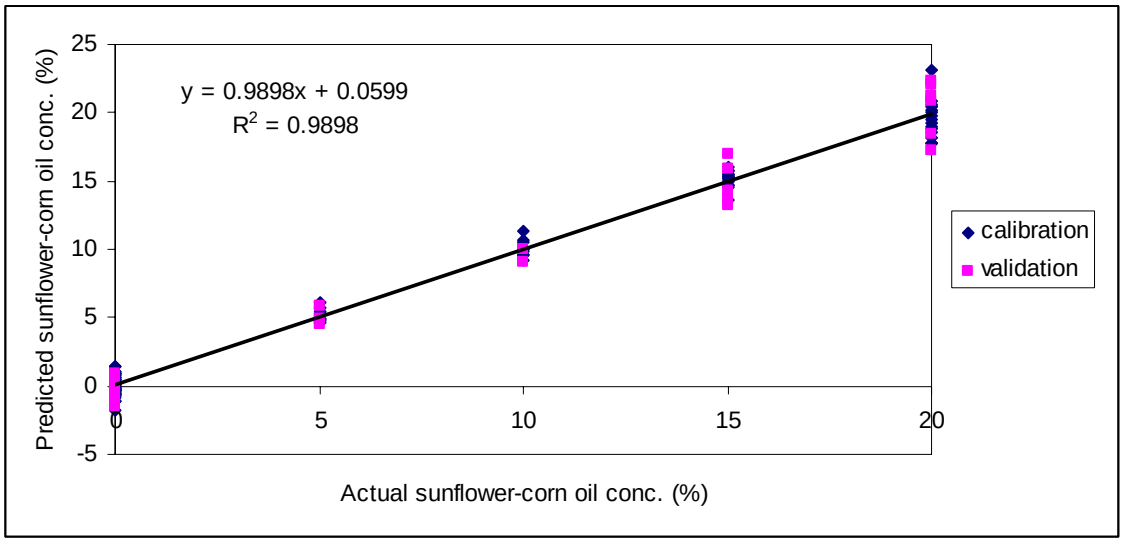


Figure 4.28. PLS regression of actual vs. predicted sunflower-corn mixture concentration in olive oil for calibration and validation sets

Quantification of sunflower and corn oil was also performed by other researchers. Özdemir and Öztürk (2007) focused on quantification of binary and tertiary adulteration of olive oil with sunflower and corn oil using NIR in conjunction with genetic inverse least square. Overall, SEP ranged between 2.49 and 2.88% (v/v) for the binary mixtures of olive and sunflower oil whereas it was between 1.42 and 6.38% (v/v) for the ternary mixtures of olive, sunflower and corn oil. In another study, NIR spectra of binary mixtures of corn-olive oil and sunflower-olive oil were manipulated with PLS to quantify corn and sunflower oil in binary olive oil mixture. The error terms were determined as 1.32 and 0.57 for sunflower and corn oils at level of 0-100% (v/v) (Christy, et al. 2004). Also, Paulli, et al. (2007) used total synchronous fluorescence and

calculated detection limit of adulterant in sunflower-olive and corn-olive oil binary mixtures as 4.3% and 3.8%, respectively.

4.4.4. Cottonseed Oil Adulteration of Olive Oil

The spectra of olive oil and cottonseed oil are illustrated in Figure 4.29. To observe the visual differences of two spectra, whole spectra were divided into two regions as shown in Figure 4.29. The olive oil can be differentiated from cottonseed oil by naked eye.

(a)

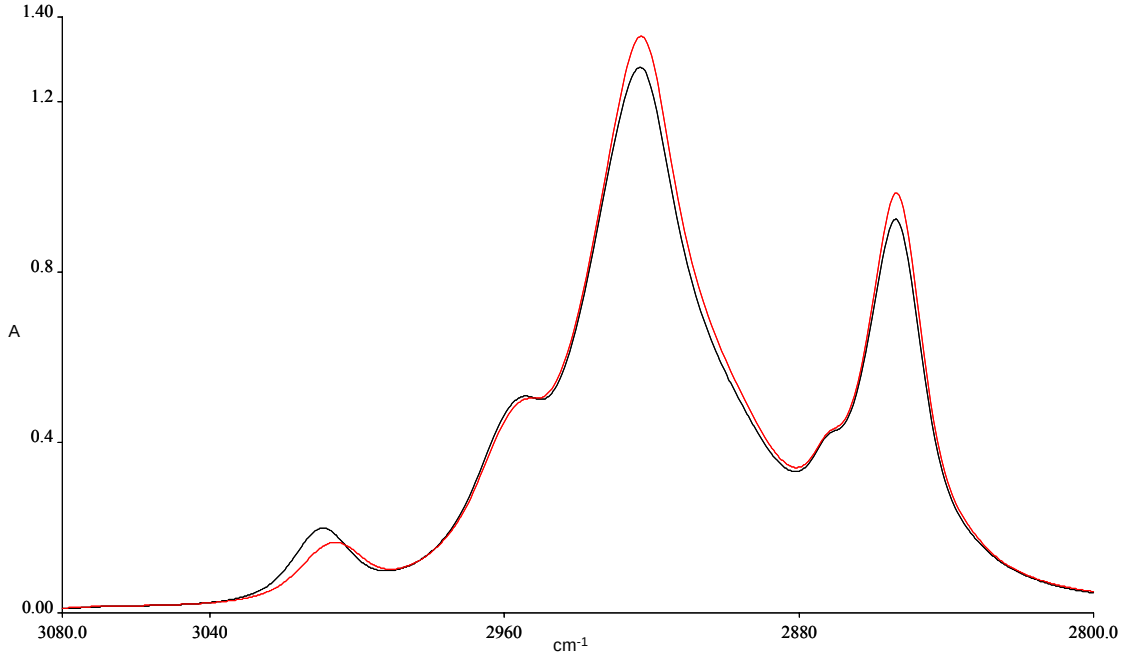


Figure 4.29. The spectra of olive oil and cottonseed oil around (a) 3080-2800 cm^{-1} (b) 2875.5-675 cm^{-1} regions (— olive oil, ═ cottonseed oil)

(cont. on next page)

(b)

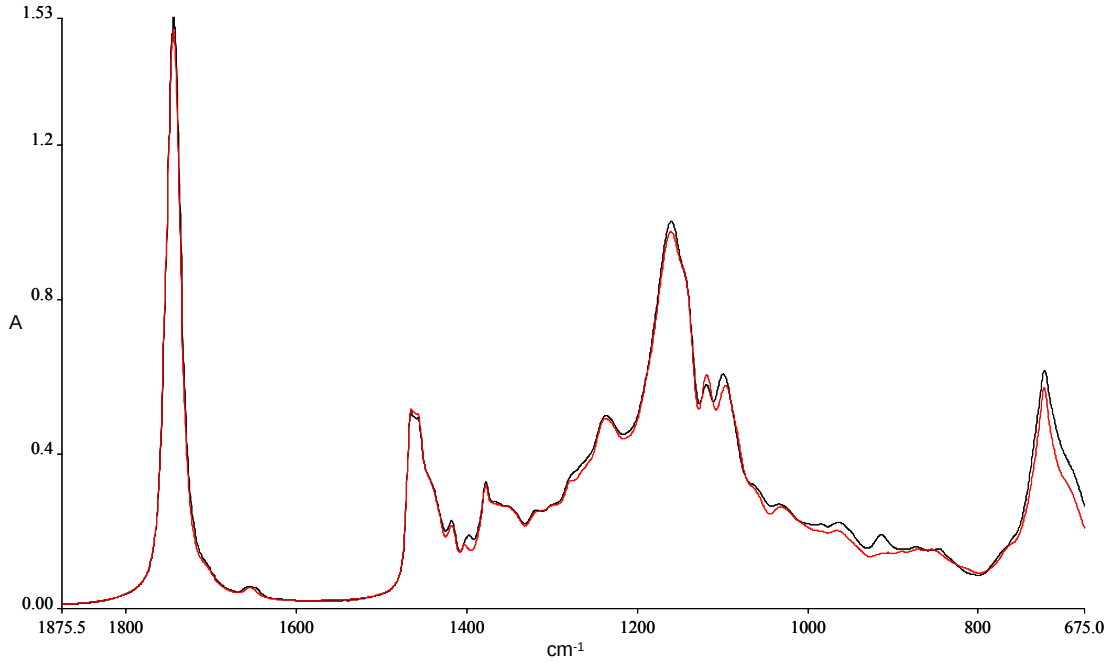


Figure 4.29. (cont.) The spectra of olive oil and cottonseed oil around (a) 3080-2800 cm^{-1} (b) 2875.5-675 cm^{-1} regions (— olive oil, ═ cottonseed oil)

Spectral data set containing cottonseed adulterated and non-adulterated olive oil samples were subjected to PCA. Non-adulterated samples were consisted of commercial olive oil samples belonging to North and South Aegean regions. Adulterated samples included olive oil including cottonseed oil between 2-20%. After applying wavelet compression to the original data, PCA model was fitted for 99 observations with 5 PCs adequately describing 83.9% of the variance. Scores plot for the first 2 PCs, where first component describes 59% of total variation and second 10.5%, is illustrated in Figure 4.30.

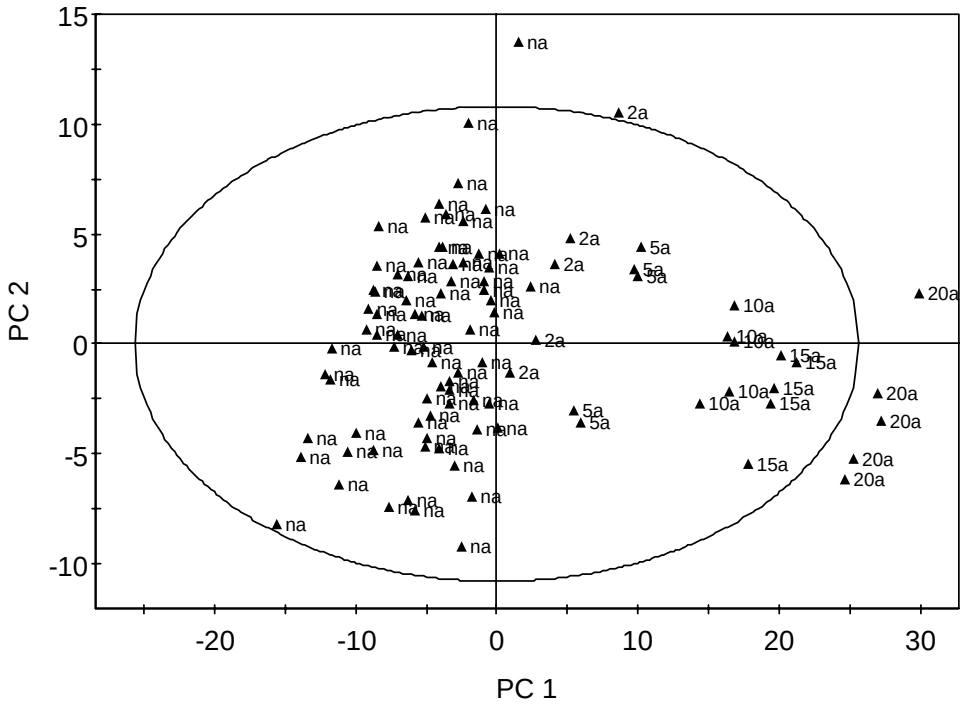


Figure 4.30. PCA scores plot of non-adulterated olive oil (na) and cottonseed adulterated olive oil (a) samples (numbers besides a notation indicates adulteration %)

Most of the non-adulterated samples were placed in the left hemisphere. These samples were discriminated from the rest whose scores were plotted in the right hemisphere. However, adulterated samples containing 2% cottonseed oil were drawn close to non-adulterated samples. Therefore, 2% adulterated samples were excluded and Coomans' plot was constructed (Figure 4.31). Non-adulterated class was described with 5 PCs describing 79.3% variation whereas adulterated class model explained 79% of total variation with 3 PCs.

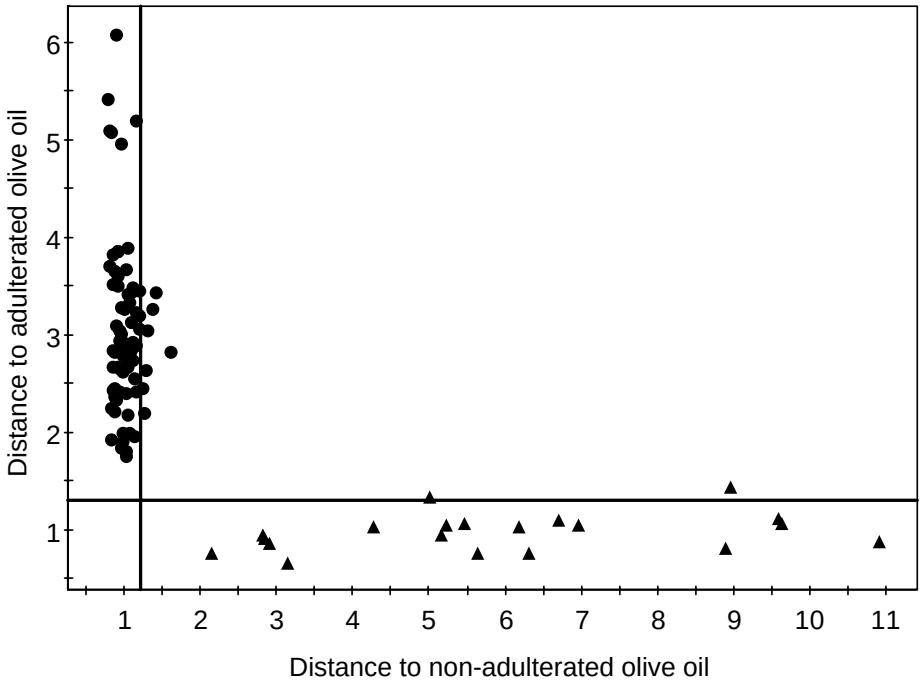


Figure 4.31. Coomans' plot for the classification of olive oil and cottonseed adulterated olive oil samples (●: olive oil ▲: adulterated olive oil)

According to Figure 4.31, only few samples were plotted beyond its critical limits. It could be concluded that quite good discrimination of olive oil samples from cottonseed adulterated olive oil samples was achieved.

PLS analysis was also applied to the same data set used in PCA for quantification of cottonseed oil in olive oil. The spectral data was manipulated with OSC wavelength compression before PLS analysis. The PLS analysis yielded a 3 PCs model explaining 99.2% variation of Y with a predictive ability of 97.9%. The R^2 value of PLS regression curve (Figure 4.32) was 0.99, and SEC value was determined as 0.49. Hence, calibration model had good predictive ability. In addition, using the PLS regression equation the samples of validation set were predicted (Figure 4.32). R^2 and SEP values were determined as 0.95 and 1.4 exhibiting high predictive ability of the proposed model. In a previous study, detection of cottonseed oil adulteration at a level of 1% was achieved using ΔECN42 (Christopoulou, et al. 2004).

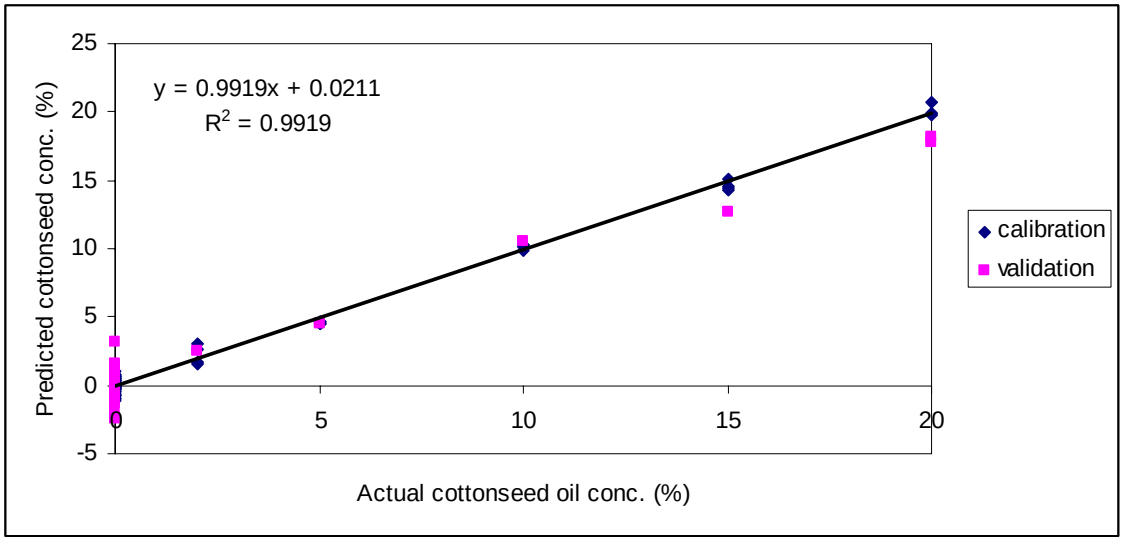


Figure 4.32. PLS regression of actual vs. predicted cottonseed oil concentration in olive oil of calibration and validation sets

4.4.5. Rapeseed Oil Adulteration of Olive Oil

Figure 4.33 displays the spectra of olive oil and rapeseed oil. The whole spectra are differentiated into two regions to observe the differences between olive oil and rapeseed oil.

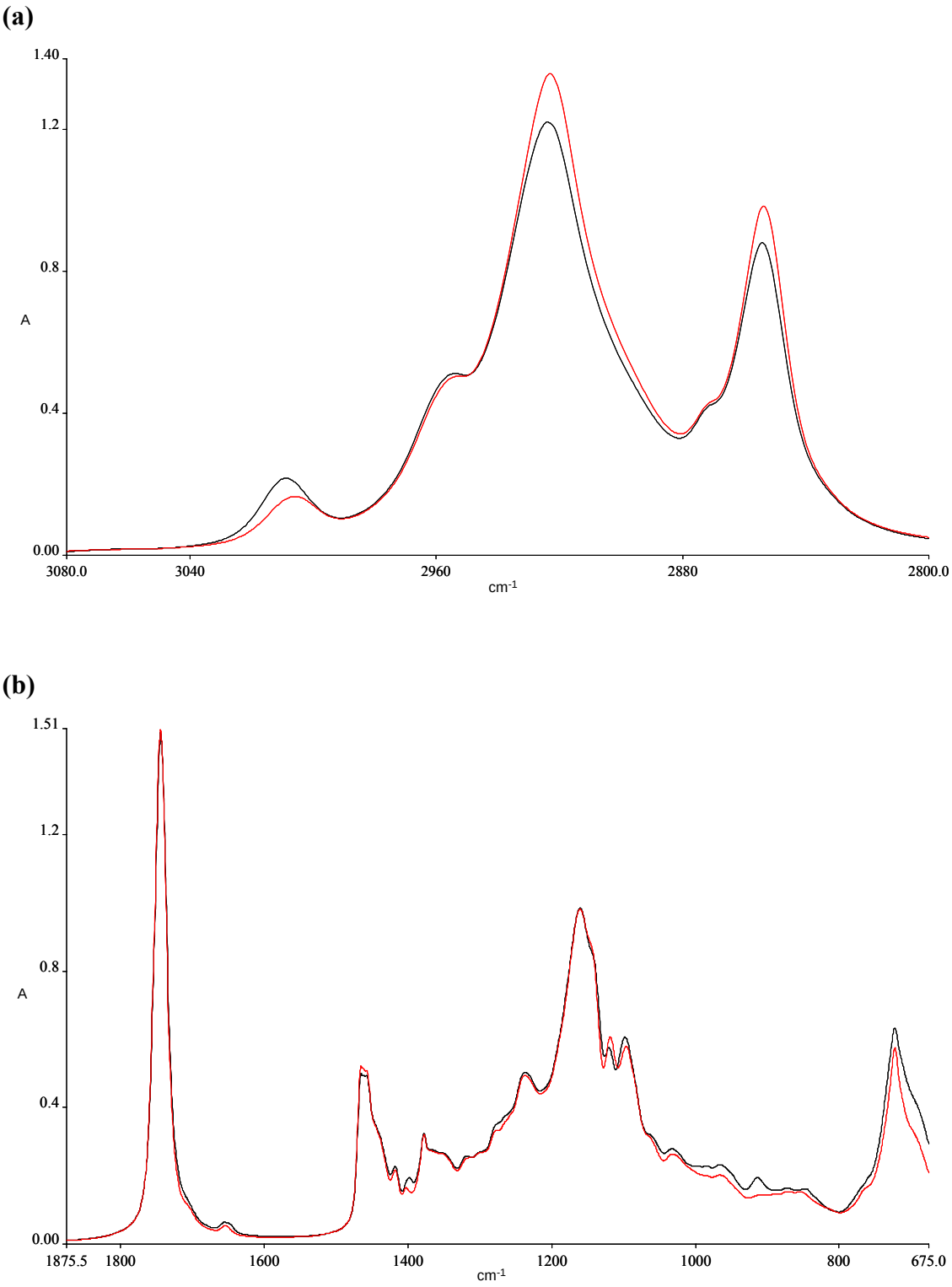


Figure 4.33. The spectra of olive oil and rapeseed oil around (a) 3080-2800 cm^{-1} (b) 2875.5-675 cm^{-1} regions (— olive oil, — rapeseed oil)

Spectral data set containing rapeseed adulterated and non-adulterated olive oil samples were subjected to PCA. After applying wavelet compression to the original data, PCA model of the 98 observations resulted in 4 PCs adequately describing 84.2% of spectral variation. Figure 4.34 represented scores plot for first 2 PCs where the first component describes 61.5% of total variation and second 9.6%. Non-adulterated samples, placed in the left side of the scores plot together, differentiated from other adulterated samples plotted in the right side. However 2% adulterated samples were placed close to non-adulterated samples so they were excluded in further analysis.

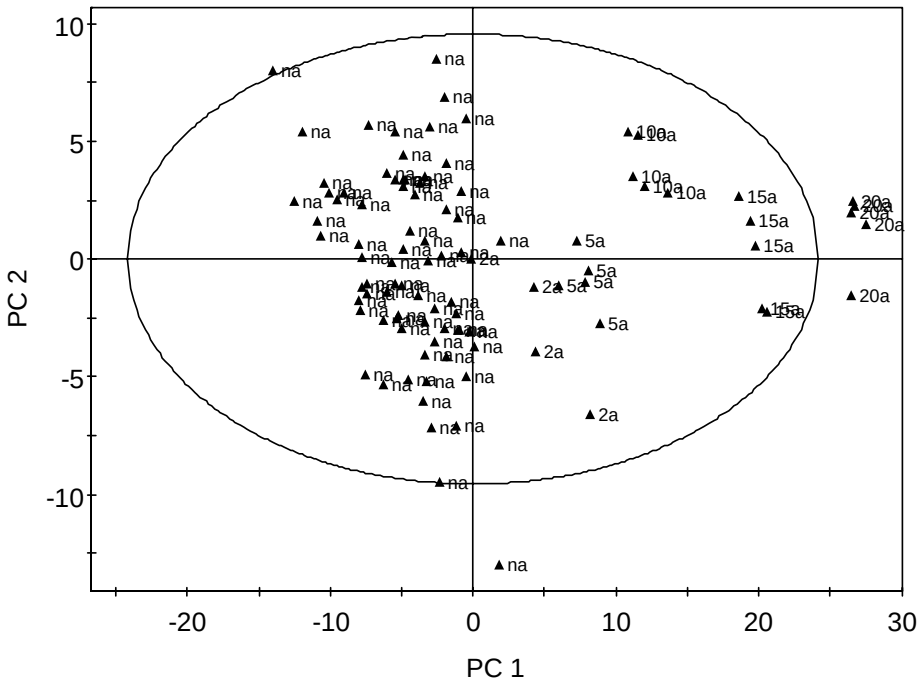


Figure 4.34. PCA scores plot of non-adulterated olive oil (na) and rapeseed oil adulterated olive oil (a) samples (numbers near a notation indicates adulteration %)

To discriminate non-adulterated samples from adulterated samples Coomans' plot was constructed (Figure 4.35). Non-adulterated class was explained with 5 PCs describing 82.1% total variation and adulterated class was described with 4 PCs describing 88.7% of total variation. Quite successful discrimination of olive oil samples from adulterated samples was performed. Besides, each grouping of sample points in

non-adulterated class model represents different adulteration percentages in ascending order from left to right.

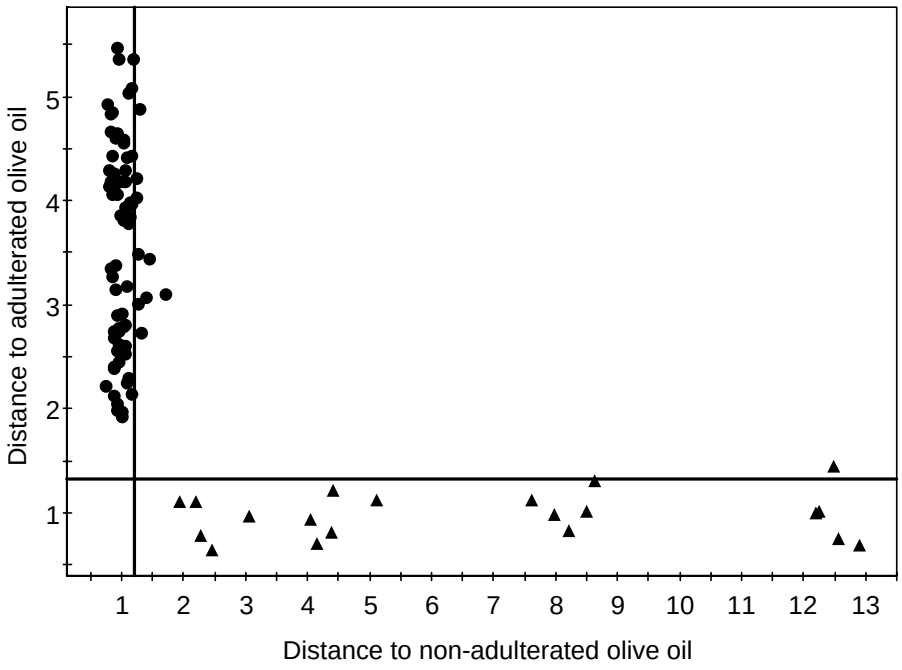


Figure 4.35. Coomans' plot for the classification of olive oil and rapeseed adulterated olive oil samples (●: olive oil, ▲: adulterated olive oil)

Quantification of rapeseed oil in olive oil was performed with PLS analysis using data set used in PCA excluding 2% adulterated samples. The data set was divided into calibration and validation sets containing 68 and 26 observations, respectively. The spectral data was manipulated with OSC and wavelength compression before PLS analysis. The PLS regression analysis on calibration set was resulted in 2 PCs explaining 99.4% of variation (Y) with a predictive ability of 99.1%. The R^2 value of regression curve, represented in Figure 4.36 was 0.99, and SEC value was determined as 0.48, respectively. Calibration model with good predictive ability was obtained. Also, using the PLS regression equation the samples of validation set were predicted (Figure 4.36). R^2 and SEP values were calculated as 0.93 and 1.32 exhibiting high predictive ability of the proposed model. As another approach, the parameter ΔECN42 was employed as a parameter for the detection of fraud of olive oils containing 4% rapeseed

oil (Christopoulou, et al. 2004). Also, total synchronous fluorescence spectra could discriminate rapeseed oil down to a level of 3.6% (Poulli, et al. 2007).

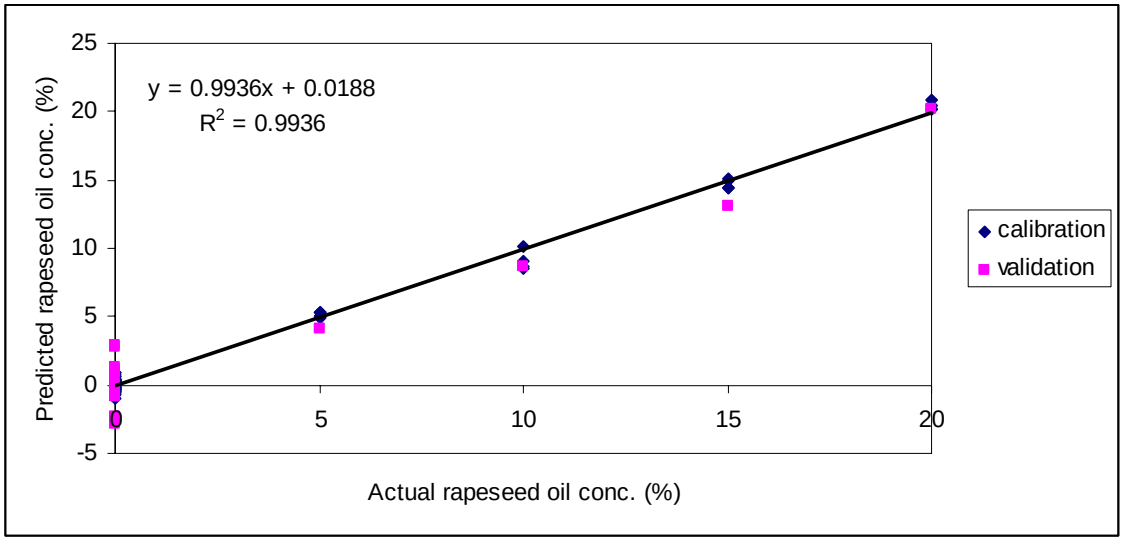


Figure 4.36. PLS regression of actual vs. predicted rapeseed oil concentration in olive oil for calibration and validation sets

4.4.6. Determination of Adulteration Regardless of the Type of Adulterant

In the last part of this study adulterated oils regardless of the type of adulterant were grouped as one adulterant class and it was tried to be determined if PCA still could discriminate pure samples from adulterated ones. Observations were classified as adulterated including spectra of rapeseed, cottonseed oils and corn-sunflower binary oil mixture and non-adulterated olive oil samples. To construct Coomans' plot (Figure 4.37) PCA was performed on adulterated and pure classes separately. Non-adulterated and adulterated classes were modelled with 4 PCs describing 87.9% variation. Modeling adulterated oil samples together means more variable chemical information making them difficult to be placed in the same group. Thus, in addition to 2% adulterated samples 5% adulterated olive oil samples could not be separated from non-adulterated olive oil samples according to scores plot which is not displayed here. So, 2-5% adulterated samples were excluded in the further analysis. According to Coomans' plot most samples of adulterated and non-adulterated samples are correctly placed in its

region. However, few samples are plotted in the region where they can not be classified as non-adulterated or adulterated.

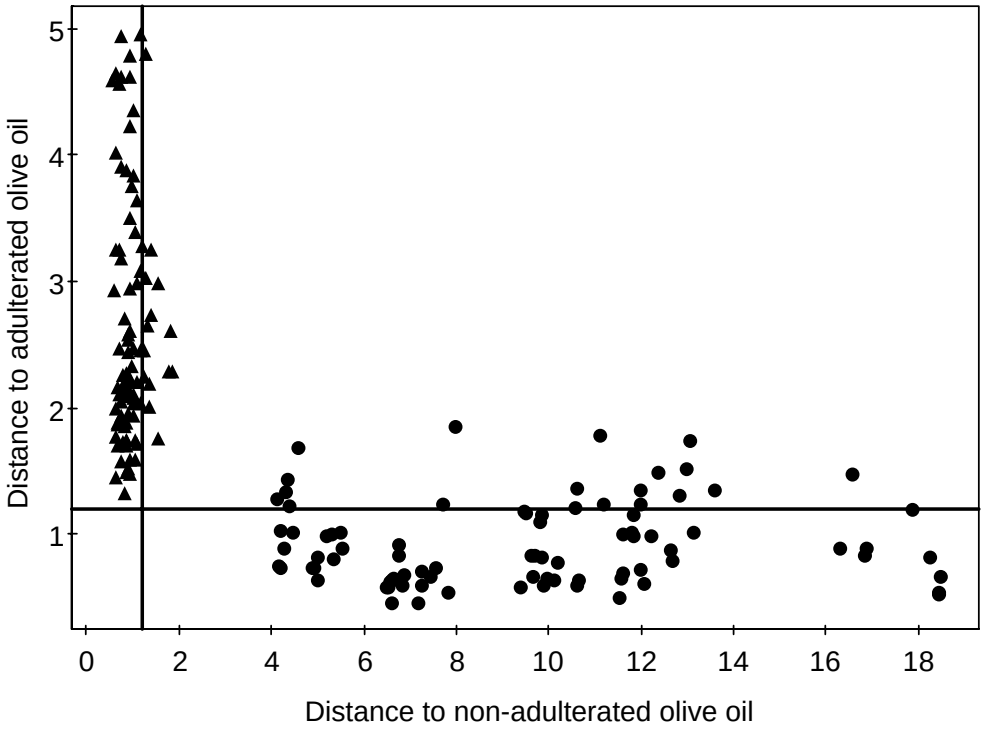


Figure 4.37. Coomans' plot for corn-sunflower binary mixture, cottonseed and rapeseed adulterated versus non-adulterated olive oil (●: olive oil, ▲: adulterated olive oil)

There are studies involving the application of other analytical and chemometric methods to discriminate olive oil from set of different adulterant oils. Availability of chromatographic profiles with SIMCA model to distinguish between non-adulterated olive oil samples and those adulterated with one of the vegetable oils (sunflower, corn, peanut and coconut oils) was illustrated (Capote, et al. 2007). Also, NIR spectra manipulated with PCA was able to classify adulterated olive oil samples with respect to type of adulterant oil (Christy, et al. 2004).

To sum up the adulteration determination results, Table 4.14 was constructed including results of individual and overall adulteration models. Results of individual modeling revealed that monovarietal, sunflower-corn, cottonseed and rapeseed oil adulteration can be detected at a level of 5% whereas detection limit is 10% for hazelnut oil adulteration. On the other hand, lower detection limit is achieved according to

overall adulteration model due to more variable chemical information of overall adulteration model

Table 4.14. General results of adulteration determination including individual and overall adulteration modeling

	Results of individual adulteration modelling						Results of overall adulteration model
	Detection limit	R^2 (cv)[a]	R^2 (cal)	SEC	R^2 (val)	SEP	Detection limit
N in AE	5%	0.98	-	-	-	0.9	-
N in E	5%	0.97	-	-	-	1.22	-
Hazelnut oil	10%	-	0.99	2.1	0.97	3.1	-
Sunflower-corn mixture	5%	-	0.99	0.78	0.98	1.04	10%
Cottonseed	5%	-	0.99	0.49	0.95	1.4	10%
Rapeseed	5%	-	0.99	0.48	0.93	1.32	10%

[a]R^2 (cv): correlation coefficient of cross validation

CHAPTER 5

CONCLUSION

In this study, the efficiency of analysis of fatty acid composition and mid-IR spectra data with a chemometric tool, PCA, for the differentiation of olive oil samples was demonstrated and compared. Results show that application of PCA to fatty acid composition is quite successful for the classification of olive oil samples with respect to variety, geographical origin and harvest year. On the other hand, mid-IR spectra can not supply distinct varietal, geographical or seasonal grouping as much as fatty acid composition does. As one of the factors affecting the olive oil composition harvest year seems to be significant in discrimination. With the advantage of rapid and easy analysis ability, FT-IR still could be used for the more general authentication issues with respect to variety, geographical region and harvest year.

FT-IR was employed to detect adulteration of olive oil with other vegetable and seed oils (hazelnut, corn-sunflower binary mixture, cottonseed and rapeseed) and also to detect mixture of a monovariety with another variety (N in AE and E). The qualification of adulteration was performed by PCA analysis. Scores and Coomans' plots were constructed to visualize differentiation of adulterated from non-adulterated olive oil samples. The results of PLS analysis of MIR spectra indicated that detection limit of adulteration was determined to be higher than 5% for Nizip (in Ayvalik-Edremit) and cottonseed oils, 10% for Nizip (in Erkence), 15% for hazelnut oil and 2% for binary mixture of sunflower-corn and rapeseed oils. In conclusion, MIR analysis with chemometric techniques could be employed as an efficient tool to detect adulteration of extra-virgin olive oil with different edible oils.

Moreover, a correlation was tried to be established between fatty acid compositions and MIR spectra with PLS. The error values obtained for oleic acid (constitutes around 99% of MUFA of olive oil) and MUFAs (constitutes around 70% of fatty acids) are lower than 2% indicating success of the PLS model. Also, REP was lower than 5% for C16:0 whereas good predictions were achieved for C18:2 and PUFA with REP values around 5%. Another PLS model was tried to be constructed between free fatty acidity and MIR spectra. SEP value was determined as 0.085. Both fatty acid

composition and free fatty acidity value analysis require chemical treatments meaning excess labor, cost and time. However, multivariate analysis of FT-IR data has the potential to predict some of the fatty acids and free fatty acidity of the virgin olive oil samples with the advantage of direct analysis opportunity.

REFERENCES

Addor, F. and A. Grazioli. 2002. Geographical Indications beyond Wines and Spirits. *The Journal of World Intellectual Property* 5(6):865-897.

Aguilera, M.P., B. Beltrán, D. Ortega, A. Fernández, A. Jiménez, M. Uceda. 2005. Characterisation of virgin olive oil of Italian olive cultivars: 'Frantoio' and 'Leccino', grown in Andalusia. *Food chemistry* 89:387–391.

Aktas, E.S., S. Imre, L. Ersoy. 2001 Characterization and lime treatment of olive mill wastewater. *Water Research* 35:2336–2340.

Aparicio, R. and R. Aparicio-Ruíz. 2000. Review. Authentication of vegetable oils by chromatographic techniques. *Journal of Chromatography* 881:93–104.

Aparicio, R. and G. Luna. 2002. Characterization of monovarietal virgin olive oils. *European Journal of Lipid Science Technology* 104:614–627.

Aparicio, R., L. Roda, M.A. Albi, G. Francisca. 1999. Effect of Various Compounds on Virgin Olive Oil Stability Measured by Rancimat. *Journal of Agricultural Food Chemistry* 47:4150-4155.

Babcock, B.A. and R. Clemens. 2004. Geographical indications and property rights: Protecting value-added agricultural products. *MATRIC Briefing Paper 04-MBP 7*, Ames, Iowa.

Baeten, V., J.A.F. Pierna, P. Dardenne, M. Meurens, D.L. García-González, R. Aparicio-Ruiz. 2005. Detection of the presence of hazelnut oil in olive oil by FT-Raman and FT-MIR spectroscopy. *Journal of Agricultural and Food Chemistry* 53:6201-6206.

Bendini, A., L. Cerretani, F. Di Virgilio, P. Belloni, M. Bonoli-Carbognin, G. Lercker, 2007. Preliminary evaluation of the application of the FT-IR spectroscopy to control the geographic origin and quality of virgin olive oils. *Journal of Food Quality* 30:424–437.

Bertran, E., M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, I. Montoliu. 1999. Determination of Olive Oil Free Fatty Acid by Fourier Transform Infrared Spectroscopy. *Journal of the American Oil Chemists' Society* 76:611–616.

Berrueta, L.A., R.M. Alonso-Salces, K. Héberger. 2007. Review. Supervised pattern recognition in food analysis. *Journal of Chromatography A* 1158:196–214.

Bianchi, G., L. Giansante, A. Shaw, D.B. Kell. 2001. Chemometric criteria for the characterisation of Italian Protected Denomination of Origin (DOP) olive oils from their metabolic profiles. *European Journal of Lipid Science and Technology* 103:141–150.

Boggia, R., P. Zunin, S. Lanteri, N. Rossi, F. Evangelisti. 2002. Classification and class-modeling of "Riviera Ligure" extra-virgin olive oil using chemical-physical parameters. *Journal of Agricultural Food Chemistry* 50:2444-2449.

Brereton, Richard G. 2003. *Chemometrics. Data Analysis for the Laboratory and Chemical Plant*. Chichester:John Wiley & Sons.

Brescia, M.A., G. Alviti, V. Liuzzi, A. Sacco. 2003. Chemometric classification of olive cultivars cased on compositional cata of oils. *Journal of the American Oil Chemists' Society* 80:945-950.

Bye, E. 1996. Chemometrics in occupational hygiene—how and why a picture can tell more than a thousand words and figures! *Annals of Occupational Hygiene* 40:145-169.

Caetano, S., B. Üstün, S. Hennessy, J. Smeyers-Verbeke, W. Melssen, G. Downey, L. Buydens, Y.V. Heyden. 2007. Geographical classification of olive oils by the application of CART and SVM to their FT-IR. *Journal of Chemometrics* 21:324–334.

Capote, F.P., J.R. Jiménez, M.D. Luque de Castro. 2007. Sequential (step-by-step) detection, identification and quantitation of extra virgin olive oil adulteration by chemometric treatment of chromatographic profiles. *Analytical and Bioanalytical Chemistry* 388:1859-1865.

Casale, M., C. Armanino, C. Casolino, M. Forina. 2007. Combining information from headspace mass spectrometry and visible spectroscopy in the classification of the Ligurian olive oils. *Analytica Chimica Acta* 589:89–95.

Cert, A., W. Moreda, M.C. Pérez-Camino. 2000. Review: Chromatographic analysis of minor constituents in vegetable oils. *Journal of Chromatography Analysis* 881:131-148.

Chiavaro, E., E. Vittadini, M.T. Rodriguez-Estrada, L. Cerretani, A. Bendini, 2008 Differential scanning calorimeter application to the detection of refined hazelnut oil in extra virgin olive oil. *Food Chemistry* 110:248–256.

Christopoulou, E., M. Lazaraki, M. Komaitis, K. Kaselimis, 2004. Effectiveness of determinations of fatty acids and triglycerides for the detection of adulteration of olive oils with vegetable oils. *Food Chemistry* 84:463–474.

Christy, A.A., S. Kasemsumran, Y. Du, Y. Ozaki. 2004. The detection and quantification of adulteration in olive oil by near-infrared spectroscopy and chemometrics. *Analytical Sciences* 20:935-940.

Caputo, A.C., F. Scacchia, P.M. Pelagagge. 2003. Disposal of by-products in olive oil industry: Waste-to-energy solutions. *Applied Thermal Engineering* 23:197–214.

Cosio, M.S., D. Ballabio, S. Benedetti, C. Gigliotti. 2006. Geographical origin and authentication of extra virgin olive oils by an electronic nose in combination with artificial neural networks. *Analytica Chimica Acta* 567:202–210.

D'Imperio, M., G. Dugo, M. Alfa, L. Mannina, A.L. Segre. 2007. Statistical analysis on Sicilian oils. *Food Chemistry* 102:956-965.

Defernez, M. and E.K. Kemsley. 1997. The use and misuse of chemometrics for treating classification problems. *Trends in analytical chemistry* 16:216-221.

Downey, G., P. McIntyre, A.N. Davies. 2003. Geographic classification of extra virgin olive oils from the eastern Mediterranean by chemometric analysis of visible and near-infrared spectroscopic data. *Applied Spectroscopy* 57:158-163.

Eriksson, Lennart, Erik Johansson and Nauna Kettaneh-Wold. 2001. *Multivariate data analysis. Principals and applications*. Sweden:Umetrics AB.

Eriksson, L., J. Trygg, E. Johansson, R. Bro, S. Wold. 2000. Orthogonal signal correction, wavelet analysis, and multivariate calibration of complicated process fluorescence data. *Analytica Chimica Acta* 420:181–195.

Euerby, M.R. and P. Petersson. 2003. Chromatographic classification and comparison of commercially available reversed-phase liquid chromatographic columns using principal component analysis. *Journal of Chromatography A* 994:13-36.

European Union Commission. 1991. Regulation EEC 2568/91 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis. *Official Journal of European Communities L248.*

Faci, J.M., M.J. Berenguer, J.L. Espada, S. Gracia. 2000. Effect of variable water irrigation supply in olive (olea europaea l.) cv. arbequina in aragon (Spain). Extra virgin oil quality parameters. *IV International Symposium on Olive Growing,* Valenzano, Italy.

Forina, M., R. Boggia, M. Casale. 2007. The information content of visible spectra of extra virgin olive oil in the characterization of its origin. *Annali di Chimica* 97:515-633.

Downey, G. 1998. Food and food ingredient authentication by mid-infrared spectroscopy and chemometrics. *Trends Analytical Chem*istry 17:418-424.

Galtier, O., N. Dupuy, Y. Le Dréau , D. Ollivier, C. Pinatel, J. Kister, J. Artaud. 2007. Geographic origins and compositions of virgin olive oils determinated by chemometric analysis of NIR spectra. *Analytica Chimica Acta* 595:136-144.

Giansante, L., D. Di Vincenzo, G. Bianchi. 2003. Classification of monovarietal Italian olive oils by unsupervised (PCA) and supervised (LDA) chemometrics. *Journal of the Science of Food and Agriculture* 83:05–911.

Grund, S.M. 1997. What is the desirable ratio of saturated, polyunsaturated, and monounsaturated fatty acids in the diet? *American Journal of Clinical Nutrition* 66:988-990.

Gutiérrez, F., B. Jímenez, A. Ruíz, M.A. Albi. 1999. Effect of Olive Ripeness on the Oxidative Stability of Virgin Olive Oil Extracted from the Varieties Picual and Hojiblanca and on the Different Components Involved. *Journal of Agricultural Food Chemistry* 47:121-127.

Harwood John and Ramón Aparicio. 2000. *Handbook of olive oil. Analysis and Properties.* Gaithersburg: Aspen publications.

Iñón, F.A., J.M. Garrigues, S. Garrigues, A. Molina, M. de la Guardia. 2003. Selection of calibration set samples in determination of olive oil acidity by partial least squares–attenuated total reflectance–Fourier transform infrared spectroscopy *Analytica Chimica Acta* 489:59–75.

International Olive Oil Council. 2007. www.internationaloliveoil.org (accessed November 10, 2007)

Ismail, A.A., F.R. van de Voort, G. Emo, J. Sedman. 1993. Rapid quantitative determination of free fatty acids in fats and oils by Fourier transform infrared spectroscopy. *Journal of the American Oil Chemists' Society* 70:335-341.

Lanteri, S., C. Armanino, E. Perri, A. Palopoli. 2002. Study of oils from Calabrian olive cultivars by chemometric methods. *Food Chemistry* 76:501-507.

López-Feria, S., C. Soledad, J.A. García-Mesa, M. Valcárcel. 2008. Classification of extra virgin olive oils according to the protected designation of origin, olive variety and geographical origin. *Talanta* 75:937–943.

Luchetti, F. 2002. Importance and future of olive oil in the world market – an introduction to olive oil. *European Journal of Lipid Science and Technology* 104:559–563.

Martens, Harald and Tormod Naes. 1989. *Multivariate Calibration*. Chichester: John Wiley & Sons.

Martínez-González, M.G. and A. Sánchez-Villegas. 2004. Review. The emerging role of Mediterranean diets in cardiovascular epidemiology: Monounsaturated fats, olive oil, red wine or the whole pattern? *European Journal of Epidemiology* 19:9–13.

Massart, D.L., B.G.M. Vanbeginste, S.N. Deming, Y. Michotte, L. Kaufman. 1988. *Chemometrics: a textbook*. Amsterdam: Elsevier Science Publishers B.V.

Matos, L.C., Cunha, S.C., Amaral J.S., José, A.P., Andrade, P.B., Seabra, R.M., Oliveira, B.P.P. 2007. Chemometric characterization of three varietal olive oils (Cvs. Cobrançosa, Madural and Verdeal Transmontana) extracted from olives with different maturation indices. *Food Chemistry* 102:406-414.

Nergiz, C. and Y. Engez. 2000. Compositional variation of olive fruit during ripening. *Food Chemistry* 69:55-59.

Ollivier, D., J. Artaud, C. Pinatel, J. Durbec, M. Guérère. 2006. Differentiation of French virgin olive oil RDOs by sensory characteristics, fatty acid and triacylglycerol compositions and chemometrics. *Food Chemistry* 97:382–393.

Otto, Matthias. 1997. *Chemometrics. Statistics and computer application in analytical chemistry* Weinheim: Wiley-VCH.

Özdemir, D. and B. Öztürk. 2007. Near infrared spectroscopic determination of olive oil adulteration with sunflower and corn Oil. *Journal of Food and Drug Analysis* 15(1):40-47.

Ozen, B.F. and L.J. Mauer. 2002. Detection of hazelnut oil adulteration using FT-IR spectroscopy. *Journal of Agricultural Food Chem*istry 50:3898-3901.

Ozen, B.F., F. Tokatli, and F. Korel. 2005. Emerging topics in olive oil research: determination of geographical origin and adulteration. *Olive Oil and Olive-Pomace Oil Symposium & Exhibition* 10-12 November, Izmir, Turkey.

Ozkaya, M.T., E. Ergulen, S. Ulger, N. Ozilbey. Molecular, morphological and oil composition variability within olive (Olea europaea L.) at semi-arid conditions. *Biotechnology and Biotechnological Equipment* 22:699-704.

Peňa, F., S. Cárdenas, M. Gallego, M. Valcárcel. 2005. 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. *Journal of Chromatography A* 1074:215-221.

Poulli, K.I., G.A. Mousdis, C.A. Georgiou. 2007 Rapid synchronous fluorescence method for virgin olive oil adulteration assessment. *Food Chemistry* 150:369-375.

Salvador, M.D., F. Aranda, S. Gomez-Alonso, G. Fregapane. 2003. Influence of extraction system, production year and area on Cornicabro virgin olive oil: a study of five crop seasons. *Food Chemistry* 80:359-366.

Smith, Brian C. 1996. *Fundamentals of Fourier transform infrared spectroscopy*. Boca Raton: CRC Press.

Stefanoudaki, E., F. Kotsifaki, A. Koutsaftakis. 1999. Classification of virgin olive oils of the two major Cretan cultivars based on their fatty acid composition. *Journal of the American Oil Chemists' Society* 76:623–626.

Tanılgan, K., M.M. Özcan, A. Ünver. 2007. Physical and chemical characteristics of five Turkish olive (Olea europea L.) varieties and their oils. *Grasas Y Aceites* 58(2):142-147.

Tapp, H.S., M. Defernez, E.K. Kemsley. 2003. FTIR spectroscopy and multivariate analysis can distinguish the geographic origin of extra virgin olive oils. *Journal of Agricultural and Food Chemistry* 51:6110-6115.

Tay, A., R.K. Singh, S.S. Krishnan, J.P. Gore. 2002. Authentication of olive oil adulterated with vegetable oils using Fourier Transform Infrared Spectroscopy. *Food Science and Technology* 35:99–103.

Torres, M.M. and D.M. Maestri. 2006. The effects of genotype and extraction methods on chemical composition of virgin olive oils from Traslasierra Valley (Córdoba, Argentina). *Food Chemistry* 96:507-511.

Turkish Food Codex. 2000. Communiqué on Cooking Olive Oil and Cooking Pomace Oil. *The Official Gazette.*

Ulberth, F. and M. Buchgraber 2000. Authenticity of fats and oils. *European Journal of Lipid Science Technology* 102:687–694.

Vlachos, N., Y. Skopelitis, M. Psaroudaki, V. Konstantinidou, A. Chatzilazarou, E. Tegou. 2006. Application of Fourier transform-infrared spectroscopy to edible oils. *Analytica Chimica Acta* 573-574:459-465.

Vossen, Paul. 1998. Olive Oil Production.
http://ucce.ucdavis.edu/files/filelibrary/1271/23975.pdf .
(accessed November 2007)

Vossen, Paul. 1998. Spanish Olive Oil Production.
http://cesonoma.ucdavis.edu/HORTIC/spain_olive.pdf
(accessed November 2007)

World Sites Atlas. 2008. Political Map of Turkey.
http://www.sitesatlas.com/Maps/Maps/601.htm
(accessed June 2008)

Wold, S., H. Antti, F. Lindgren, J. Öhman. 1998. Orthogonal signal correction of near-infrared spectra. *Chemometrics and Intelligent Laboratory Systems*. 44:175–185.

Yang, H. and J. Irudayaraj. 2001. Comparison of Near-Infrared, Fourier Transform-Infrared, and Fourier Transform-Raman Methods for Determining Olive Pomace Oil Adulteration in Extra Virgin Olive Oil. *Journal of the American Oil Chemists' Society* 78:889–895.