DEVELOPMENT OF MOLECULAR SPECTROSCOPIC MULTIVARIATE CALIBRATION MODELS FOR THE DETERMINATION OF FATTY ACID AND TRIACHYL GLYCEROL COMPOSITIONS OF OLIVE OILS

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 Chemistry

by
Esra KUDAY

March 2014
İZMİR

We approve the thesis of **Esra KUDAY**

Examining Committee Members:

_____

Prof.Dr.Durmuş ÖZDEMİR
Department of Chemistry, İzmir Institute of Technology

_____

Assoc. Prof. Dr. Figen KOREL
Department of Food Engineering, İzmir Institute of Technology

_____

Prof. Dr. Şerife Hanım YALÇIN
Department of Chemistry, İzmir Institute of Technology

17 March 2014

_____

Prof. Dr. Durmuş ÖZDEMİR
Supervisor, Department of Chemistry
İzmir Institute of Technology

_____

Prof. Dr. Ahmet E. EROĞLU
Head of the Department of Chemistry

_____

Prof. Dr. R. Tuğrul SENGER
Dean of the Graduate School of
Engineering and Sciences

ACKNOWNLEDGEMENT

I would like to express my gratitude to people who have been very helpful during my thesis study.

Firstly I would like to thank my advisor, Prof Dr. Durmuş ÖZDEMİR who gave me information about chemometry and chemometric techniques, for all his support, patience and guidance.

Next, I would like to acknowledge Doç.Dr. Harun DIRAMAN for providing samples and data for the essential parts in my thesis.

Special thanks to Deniz ÇELİK and Duygu VURAL for their help, support and constructive comments on this thesis and also for their friendship in three years period of my IYTE life.

I also want to thank in particular Begüm AKARSU, Özlem ECE and all my laboratory mates for their friendship and emotional support.

Finally, this thesis could not have been written without the support of my lovely family. I would like to express my special thanks to my mother Necla KUDAY, my father Rabil KUDAY, my brother Harun KUDAY and my fiance Mehmet SÖNMEZ for their endless love, motivation and prays all the time. I owe a lot to them.

ABSTRACT

DEVELOPMENT OF MOLECULAR SPECTROSCOPIC MULTIVARIATE CALIBRATION MODELS FOR THE DETERMINATION OF FATTY ACID AND TRIACHYLGLYCEROL COMPOSITIONS OF OLIVE OILS

The determination of fatty acid methyl esters and triachyl glycerol compositions of olive oils by chromatographic methods require not only sample pre treatment carried out but also extend the time of the analysis. Also, chromatographic methods are expensive. Therefore, there is a need for an alternative method.

In this study, it is aimed to develop molecular spectroscopic multivariate calibration models for the determination of some of the fatty acid methyl esters and tri-achyl glycerol compositions of olive oils. For this purpose, 79 olive oil samples from different regions of Turkey (Manisa and Bursa) were collected and scanned with Fourier Transform Infrared spectroscopy equipped with attenuated total reflectance (FTIR-ATR) accessory, Fourier Transform Near Infrared spectroscopy (FTNIR) and Gas Chromatography (GC), High Performance Liquid Chromatography (HPLC). Chromatographic analyses of these samples were done by Olive Oil Research Institute and spectroscopic analyses were done by our group. The data obtained from High Performance Liquid Chromatography (HPLC) was chosen as a reference method for tri-achyl glycerol compositions and also the data obtained from Gas Chromatography (GC) was chosen as a reference method for fatty acid methyl esters. Also, Genetic inverse least squares (GILS) and Partial least square methods (PLS) were used for multivariate calibration.

In conclusion, NIR and FTIR combined with multivariate calibration models can be more advantageous compare to chromatographic methods because of their simplicity and speed. When investigating the results, relatively successful calibration models were obtained from GILS method than PLS method.

ÖZET

ZEYTİN YAĞLARININ YAĞ ASİTLERİ VE TRİAÇİL GLİSEROL KOMPOZİYONLARININ BELİRLENMESİ İÇİN MOLEKÜLER SPEKTROSKOPİK ÇOK DEĞİŞKENLİ KALİBRASYON MODELLERİNİN GELİŞTİRİLMESİ

Zeytinyağlarının yağ asidi metil esterleri ve triaçil gliserol kompozisyonlarının kromotografik yöntemlerle belirlenmesi hem örneğin ön işlem yapılmasını gerektirmekte hem de analiz süresinin uzun olmasına sebep olmaktadır. Ayrıca kromotografik yöntemlerin pahalı olması alternatif yöntem arayışına yol açmaktadır.

Bu çalışmada zeytinyağlarının bazı yağ asidi metil esterleri ve triaçil gliserol kompozisyonlarını belirlemek için moleküler spektroskopik çok değişkenli kalibrasyon modellerinin geliştirilmesi amaçlanmıştır. Bu amaçla, Türkiye'nin farklı yörelerinden (Manisa ve Bursa) 79 adet zeytinyağı örneği toplanmış ve toplanan örnekler Fourier Transform Infrared spektroskopisinde zayıflatılmış toplam reflektans aparatı (FTIR-ATR),Fourier Transform Yakın Infrared spektroskopisi (FTNIR) ve Gaz Kromatografisi (GC) ayrıca Yüksek Performanslı Sıvı Kromatografisi (HPLC) ile taranmıştır.BU örneklerin kromatografik analizleri Zeytincilik Araştırma Enstitüsü tarafından ve spektroskopik analizleri grubumuz tarafından yapılmıştır. Triaçil gliserol kompozisyonları için Yüksek performans likit kromatografisi (HPLC), yağ asidi metil esterleri kompozisyonları için Gaz Kromotografisi (GC)'nden elde edilen veriler kullanılıp, referans yöntem olarak alınmıştır. Ayrıca Genetik en küçük ters kareler yöntemi ve Kısmi ters kareler yöntemi çok değişkenli kalibrasyon olarak kullanılmıştır.

Sonuç olarak, NIR ve FTIR spektroskopisiyle birleştirilmiş çok değişkenli kalibrasyon modelleri kromatografik yöntemlerle kıyaslandığında basit olması ve hızı yönünden daha avantajlı olabilmiştir. Sonuçlar incelendiğinde, GILS methodundan elde edilmiş kalibrasyon modelleri PLS e göre nispeten daha iyi sonuçlar vermiştir.

TABLE OF CONTENTS

LIST OF FIGURES

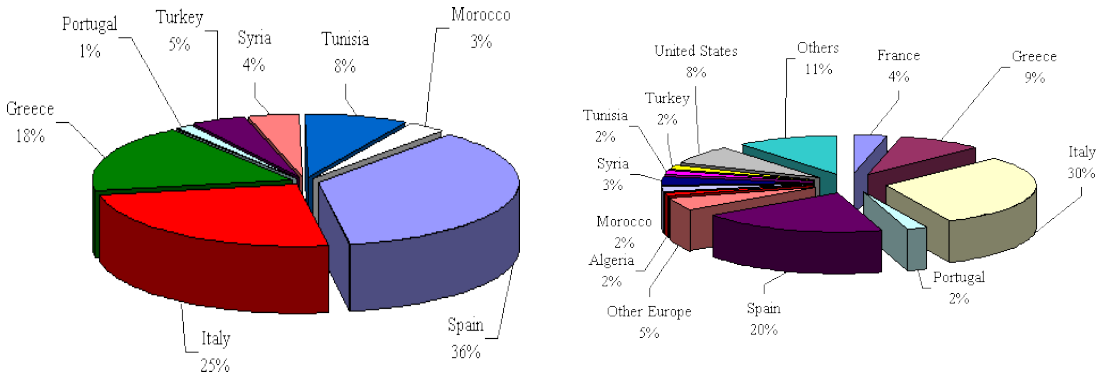
Figure **Page**

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Olive oil is an important vegetable oil in the world. A lot of countries, especially Mediterranean countries such as Spain, Italy, Portugal, Tunisia, Turkey and Morocco meet the vast majority of olive oil production (Figure1.a). The main consuming Mediterranean countries are also the main olive oil producers. These countries represent %77 of world consumption (Figure1.b). Although Turkey is in the Mediterranean region, production and consuming of olive oil is highly low. In recent years, studies on olive oil have increased. Besides, also worldwide, due to the rapid increase in production and consumption of olive oil, scientists give importance and they study intensively about this vegetable oil. Because it's importance for health and it is used as a cure for many diseases.



(a) (b)

Figure 1.1. a) Main producing countries in 2005 b) Main consuming countries in 2005 (Source: UNCTAD.org,2005 data from International Olive Oil Council (IOOC))

Turkey is the fifth country in terms of the presence of the tree in olive producing countries in the world and also is the fourth country in terms of production.

Seventy six percentage of the Aegean, 14% of the Mediterranean, 5.7% of Marmara and 4%in South Eastand0.3% of Black Sea region produce olive oil in Turkey. Aegean and Marmara region is dominated by the production of Gemlik olive varieties. Fruits of Gemlik type are rich in fat. It is a quite common olive grown region. Total olive oil produced in Turkey is divided into %70 for oil seed and % 30 for table. 70% of table oil is consumed in domestic, % 15 of table oil is exported and the rest of it is stocked. 45% of production of olive oil is used in domestic and 40% of olive oil is used to export and 15% of it is stocked. (Güler et al 2010).

Olive oil constitutes various chemical components including triacylglycerols, free fatty acids, phosphotides as the major components and also minor components such as phenolic compounds, hydrocarbons etc. With increasing consumer demand for high quality olive oil, oil was produced from olives of just one variety. The European Union (EU) has established different categories of olive oil according to the production process (Bertran et all, 1998). In order to ensure the quality within the different categories, several analyses should be made. One of the important methods to determine olive oil quality is the fatty acid content. For the determination of the fatty acid composition, the TAGs are transesterified to give the methyl esters prior to analysis because the esters are less polar than the corresponding fatty acids, and, thus, are more compatible with the various chromatographic systems (Jennings, 1999).

Chromatographic methods have been generally prefered 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 the authenticity of olive oils. Though these procedures have been successfully used in various chemical analyses, they do suffer from the disadvantages of being time, labor, and resources consuming. This can be avoided by using a spectroscopic technique which offers the potential for rapid determination of large numbers of samples by unskilled workers with minimal use and disposal of costly solvents and chemicals. Most vegetable oil spectroscopic applications focus on detecting adulteration (Lai, 1995), determining classification (Bewig, 1994), and establishing geographical origins (Tsimidou, 1994). Also, a few quantitative determinations have been done without any previous separation techniques. Thus, Fourier transform infrared (FTIR) spectroscopy has become an alternative to these

techniques because of its simplicity in sample handling and unnecessary pretreatment. The need to develop new simplified routine methods has resulted in little attention given to the selection of samples and their calibration; however, this process is essential if reliable predictions are desired. Thus, an overfitted calibration model may provide good results for the calibration set but poor predictive ability for similar samples if sufficient variability is not included in the calibration process. Therefore, proper development of analytical methods requires careful use of multivariate calibration methods to obtain reliable results.

The aim of this study is to develop multivariate calibration models to determine of fatty acid methyl esters and triachyl glycerol compositions of olive oils using spectral data. These spectral data were used to construct calibration models. Extra virgin olive oil samples which are obtained from Gemlik variety harvested Marmara and Aegean regions were used in this study. Spectroscopic and chromatographic analyses were done simultaneously.

1.1 Definitions and Composition of Olive Oil

1.1.1. Definition of Olive Oil

According to the International Olive Oil Council (IOOC) the international definition of olive oil is oil produced by extraction of the fruit of the olive tree (OleaEuropaea Sativa Hoffman et Link) to the exclusion of oils obtained using solvents or reesterification processes and of any mixture with oils of other kinds (Firestone 2005).There are two major clusters of olive oil. One of them is virgin olive oil, and the other is olive pomace olive oil.

Virgin olive oil is the oil obtained from the fruit of the olive tree solely by mechanical or other physical means under conditions, particularly thermal conditions, that do not lead to alterations in the oil, and which has not undergone any treatment other than washing, decantation, centrifugation, and filtration.

Virgin olive oil is designed as a natural oil and categorized four different sub groups:

- *Extra virgin olive oil* is the virgin olive oil that has an organoleptic rating of 6.5 or more as determined by the IOOC and a free acidity, expressed as oleic acid, of not more than 1 g per 100g.
- *Fine virgin olive oil* is the virgin olive oil that has an organoleptic rating of 5.5 or more and a free acidity, expressed as oleic acid, of not more than 1.5 g per 100g.
- *Ordinary virgin olive oil* is the virgin olive oil that has an organoleptic rating of 3.5 or more and a free acidity, expressed as oleic acid, of not more than 3.3 g per 100g.
- *Lampante virgin olive oil* is the virgin olive oil that has an organoleptic rating of less than 3.5 or more and a free aciditiy, expressed as oleic acid, of more than 3.3 g per 100g

Olive-pomace oil is an oil obtained by solvent extraction of olive–pomace and does not include any oil obtained by a reesterification procedure or any mixture with other kinds of oils. Olive pomace oil has different various categories.

- *Crude olive–pomace oil* isolive–pomace oil intended for refining to produce a refined olive pomace oil suitable for human consumption, or intended for technical purposes.
- *Refined olive–pomace oil* is is the oil obtained from "olive pomace" by extraction by means of solvents and made edible by means of refining methods which do not lead to alteration in the initial glyceridic structure.
- *Olive–pomace oil* is a blend of refined olive–pomace oil and virgin olive oil .In no case may this be called ‘olive oil.’ .These definitions about olive oil were all taken from the source of IOOC. In this study extra virgin olive oil which has the high quality was used.

1.1.2. Chemical Composition of Olive Oil

Olive oil has not a specific combination. Because, it depends on species of olive, year of manufacture, the region of olive, and the method of squeeze of olives. Olive oil is compositional so varies from year to year.

Generally olive oil is composed mainly of triacylglycerols (triglycerides or fats) and contains small quantities of free fatty acids (FFA), glycerol, phosphatides,

pigments, flavor compounds, sterols, and microscopic bits of olive. The Table 1.1 shows the chemical composition and percentages of the components of olive oil.

Table 1.1. Chemical composition of olive oil

Composition	Ratio (%)
Triglycerides	% 99.8
Saturated fatty acids (SFA)	%14
Palmitic acid(PA)	%7.5-20
Stearic acid (SA)	%0.5-5.0
Mono unsaturated fatty acids (MUFA)	%72
Oleic acid(OA)	%55-83
Palmitoleic acid(POA)	%0.3-3.5
Poly unsaturated fatty acids (PUFA)	%12
Linoleic acid(LO)	%3.5-21
Linolenic acid(LN)	%0.0-1.5
Non triglyceride compounds	%0.2
Vitamin E	150mg/kg
Polyphenols	300mg/kg
Cholesterol	0

Triacylglycerols are the major energy reserve for plants and animals. These molecules derived from the natural esterification of three fatty acid molecules with a glycerol molecule. Most prevalent in olive oil is the oleic-oleic-oleic (OOO) triacylglycerol, followed, in order of incidence, by palmitic-oleic-oleic (POO), then oleic-oleic-linoleic (OOL), then palmitic-oleic-linoleic (POL) and stearic-oleic-oleic (SOO).

Fatty acids are a class of compounds containing a long hydrocarbon chain and a terminal carboxylate group (-COOH). They have the general structure CH3(CH2)nCOOH. Fatty acids belong to a category of biological molecules called lipids, which are generally water-insoluble but highly soluble in organic solvents. Fatty acids can be either saturated or unsaturated, a distinction that has important consequences for their chemical properties as well as the properties of other lipids with

fatty acid components:

Saturated fatty acids have no double bonds between the carbon atoms of the fatty acid chain which are fully saturated with hydrogen atoms. The major saturated fatty acids in olive oil are Palmitic acid (C16:0) is a saturated fatty acid that makes up 7.5 to 20% of olive oil. Stearic acid (C18:0) is a saturated fatty acid that makes up 0.5 to 5% of olive oil.

Unsaturated fatty acids have one or more double bonds between carbon atoms. Monounsaturated fatty acids contain one double bond near the middle of the chain, creating a "kink" in the chain. One of the carbon atoms, bonded to only one hydrogen atom, forms a double bond with a neighboring carbon atom. Oleic acid and palmitoleic acid can be given an example. Oleic acid (C18:1) is a monounsaturated omega-9 fatty acid and it makes up 55 to 83% of olive oil. Palmitoleic acid (C16:1) is a monounsaturated omega-7 fatty acid that makes up 0.3 to 3.5% of olive oil.

Polyunsaturated fatty acids contain more than one double bond. Linoleic and linolenic acids can be given an example. Linoleic acid (C18:2) is a polyunsaturated omega-6 fatty acid that makes up about 3.5 to 21% of olive oil. Linolenic acid (C18:3) (especially α-linolenic acid) is a polyunsaturated omega-3 fatty acid that makes up 0 to 1.5% of olive oil.

Polyphenols in olive oil are natural antioxidants that contribute to a bitter taste, astringency and resistance to oxidation. They have been shown to have a host of beneficial effects from healing sunburn to lowering cholesterol, blood pressure, and risk of coronary disease. This molecule makes up 300mg/kg of olive oil.

Tocopherols; olive oil contains the tocopherols α-, β-, γ-, δ- (α- tocopherol covers almost 88%). The tocopherol content of olive oil depends not only on the presence of these compounds in olive fruit but also on several other factors, involved in the transportation, sorage and olive fruit processing. According to Viola et al. (1997), 6 the ratio of vitamin-E to polyunsaturated fatty acids in olive oils is better than in other edible oils.

Pigments; the colour of olive oil is mainly related to the presence of chlorophyll and pheophytin. Carotenoids are also responsible for the colour of olive oil. The presence of these constituents depends on several factors, such as cultivar, soil and climate, and fruit maturation as well as applied conditions during olive oil processing.

Phenolic Compounds; olive fruit contains simple and complex phenolic

compounds. Most of these compounds pass into the oil increase its oxidative stability and improve the taste. Hydrohtyrosol, tyrosol and some phenolic acids are mainly found in olive oil (Kiritsakis, et al. 1998). The phenol content and the specific composition of these phenols in olive oil depend on the altitude where olive trees are grown, on the harvesting time and on the processing conditions (Cinquanta, et al. 1997; Kiritsakis, et al. 1998).

Aroma Components; aroma and the taste of olive oil are its main sensory characteristics. These characteristics are attributed to a group of aroma compounds. Their formation occurs in olive fruit, via a series of enzymatic reactions (Kiritsakis, et al. 1998).

Fatty acids and tri achyl glycerol compositions which are specified in the Table 1.1 and described in detail. OLL, PoOO, POP, OOO and SOO are used in this study.

CHAPTER 2

MULTIVARIATE ANALYSIS TECHNIQUES

According to the International Chemometric Society (ICS), chemometry chemical data and the implementation of mathematical and statistical techniques can be defined as the assimilation of these data more useful information. The purpose of which is accepted as a branch of analytical chemistry, chemometrics, chemometric methods to integrate its various applications in the field of chemistry.

In recent years, computerized devices have improved due to technological advances. The evaluation of a large number of data processing and statistical information is needed in analytical chemistry. Chemometric techniques collect not only quality data used in design of experiments, provide guidance on the optimization of experimental parameters, calibration and signal processing. General information about calibration techniques and detailed information about the genetic inverse least square method are given in this chapter.

2.1 Calibration Methods

2.1.1 Overview

Calibration is a process a model is constructed to obtain a relation between the output of an instrument and chemical or physical properties of samples. Prediction is a process that the constructed model is used to estimate the properties of samples which their instrument responses are given. The model is set up by measuring instrument responses and concentration levels of certain chemical contents of the samples. Then, this model is used to predict the concentration of an unknown content sample (Beebe, et al. 1998).

Calibration is an important step in chemical analysis since a good accuracy and precision can only be achieved with a good calibration model. Calibration techniques are divided into two types: univariate and multivariate calibration. One response is

taken from an instrument and that response is related to the concentration of the chemical component of a sample in many applications. This method is called *univariate calibration* because the number of the instrumental response for each sample is just one. If more than one device is performed to reduce data for more than one sample the analyte is called *multivariate calibration* (Beebe, et al.1998).

Univariate calibration methods are defined as zero-order calibration and include such as spectrophotometer measurements obtained the use of a certain wavelength for determining the concentration of an analyte. In this method, absorption at a wavelength or a peak area is taken and its relation to the concentration of a sample is then modelled. This method is generally have been used for quantitative analysis in many spectroscopic techniques such as UV-Vis, IR and NIR spectroscopy, where the relationship between the concentration of an analyte and the instrumental response is expressed by Lambert Beer´s law.

Univariate calibration methods are divided into two types: classical and inverse calibration methods.

2.1.2 Classical Univariate Calibration

The classical univariate calibration method uses the statistical model, which assumes Beer's law. In this method, a series of experiments can be performed to relate the concentration to a single spectroscopic wavelength or chromatographic peak area. The formula of classical univariate calibration is:

$$\mathbf{a} \approx \mathbf{c} \cdot s \tag{2.1}$$

where, $\mathbf{a}$ is a vector of absorbance at one wavelength for a number of samples. $\mathbf{c}$ is of the corresponding concentrations. The scalar s relates these parameters and is determined by this equation:

$$s \approx (\mathbf{c}' \cdot \mathbf{c}) \cdot \mathbf{c}' \cdot \mathbf{a} \tag{2.2}$$

$\mathbf{c}'$ is the transpose of the concentration vector. After this step, the prediction model for the unknown sample is constructed by this equation:

$$\hat{c} \approx \hat{a} / s \quad (2.3)$$

where the hat symbol of the scalar a and s refer to the prediction.

To increase the quality of the prediction models, errors or residuals are calculated. Errors or residuals are the difference between the predicted and observed concentration values. If the residuals are less, the model can be better. (Brereton, 2000).

$$e = c - \hat{c} \quad (2.4)$$

2.1.2.1 Inverse Univariate Calibration

Although classical univariate calibration is the most widely used method in analytical chemistry, it is not always the most suitable approach in terms of two reasons. Firstly, the overall objective is to estimate concentration from a spectrum or chromatogram. The other reason depends on the error distributions. Generally, the error in response depends on instrumental response. But today, devices are more reliable due to the developments in the reproducibility of instruments. Because concentration values are generally determined gravimetrically by weighing, dilutions, the source of error is larger than instrumental error. It can be said that the source of error is due to the concentration. Classical calibration constructs a model where all errors are in the response (Figure 2.1.a). After the developments in instrumentation, the more suitable assumption indicates that errors are in the measurement of concentration (Figure 2.1.b).

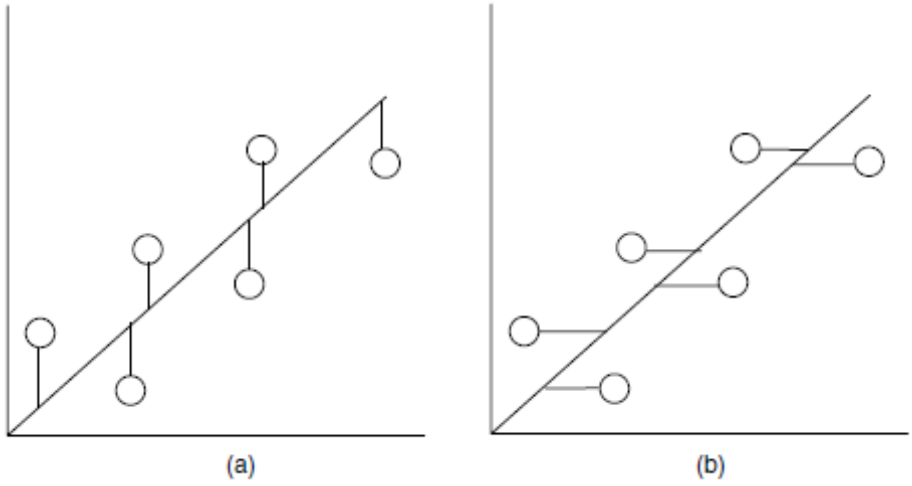


Figure 2.1. Difference between errors in a) classical and b) inverse calibration

Inverse calibration can be modeled as:

$$\boldsymbol{c} \approx \mathbf{a}.b \tag{2.5}$$

b is scalar coefficient and is inverse of *s*. *b* is only approximately the inverse of *s* (see above), because each model makes different assumptions about error distributions. *b* can be determined by this formula.

$$b \approx (\mathbf{a}^{\mathbf{1}}.\mathbf{a})^{-1}.\mathbf{a}^{\mathbf{1}}.\mathbf{c} \tag{2.6}$$

The prediction of an unknown sample can be performed by using *b* value (Brereton, 2003).

$$\hat{\mathbf{c}} \approx \hat{\mathbf{a}}.b \tag{2.7}$$

2.1.3 Multivariate Calibration

Mathematical methods applied to chemical analysis when multivariate calibration methods contain more than one instrument for each sample. In spectroscopy, generally multivariate calibration is related to data included instrument signals and calculated multiple wavelengths for a sample containing multiple components. Multivariate calibration methods find the solutions to problems in univariate calibration

methods. Due to lot of data are used to estimate the concentration of an analyte, Multivariate calibration has some advantages over univariate calibration.

There has to be one measurement for each component by univariate method. So,

It takes a long time. In multivariate calibration method, there can be measurement for one more components at the same time. So, spent time will be less. (Beebe et al.1998)

Multivariate calibration has fault-detection capabilities. That means unknown interferences in the sample can be sort out by multivariate calibration. The presence of interferences may cause wrong prediction of concentration of analyte in univariate calibration. To prevent this problem, physical separation of analyte from interfering material or using selective measurements is needed and this means necessity of more effort. Figure 2.2 demonstrates how calibration curve is affected by interferences.

By multivarite calibration, nonlinearities caused by the interferences can be reduced by selecting more variables and chance of obtaining better calibration curve can be increased. Therefore, time and effort spent to remove interferences physically is respectably decreased. (Öztürk, 2003).

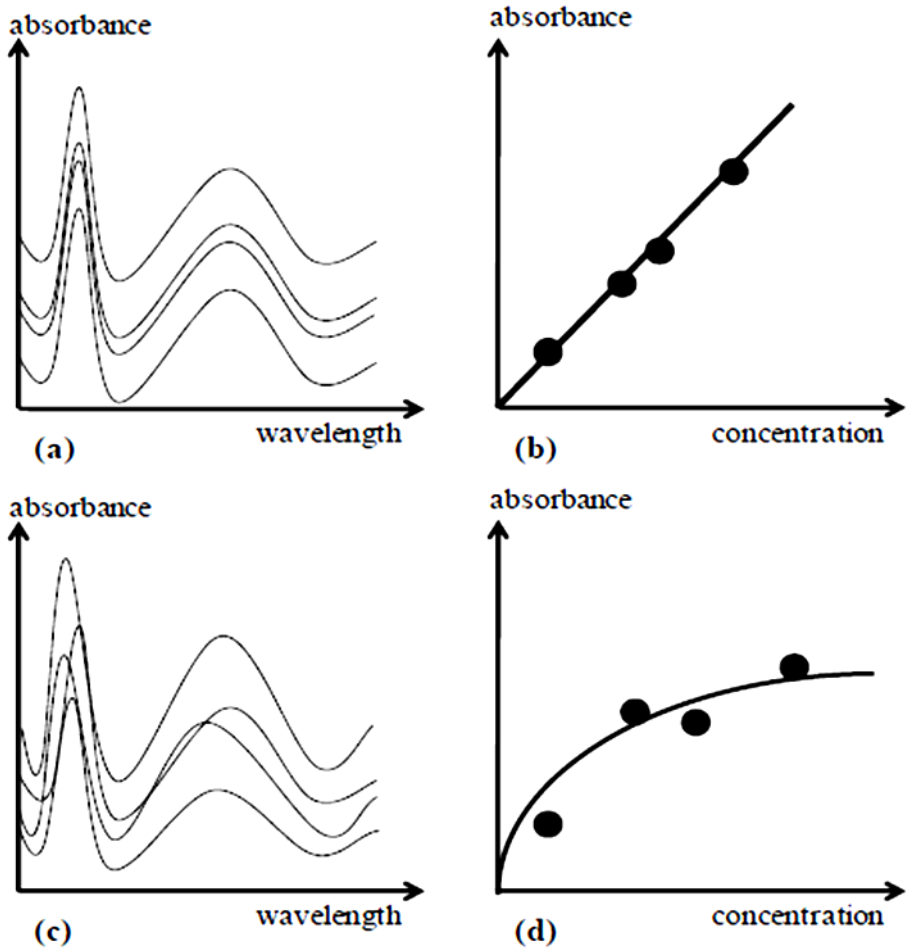


Figure 2.2. (a) Spectra of a sample in different concentrations which has no interference (b) univariate calibration curve; (c) spectra of a sample in different concentrations which has interfering materials (d) univariate calibration curve. (Source:Karaman 2008)

During the last few years, advances in chemometrics, for the analysis of complex chemical mixtures led to the development of a wide variety of calibration methods. Some of the most widely used multivariate calibration methods are Classical Least Squares (CLS), Inverse Least Squares (ILS), Partial Least Squares (PLS) and Principle Component Regression (PCR).

In this study, Genetic Inverse Least Square (GILS) method is used. Before explain this method, an overview of ILS and CLS methods will be given.

2.1.3.1 Classical Least Squares (CLS)

This method is based on Lambert Beer law and accepts absorbance values depend on the concentration. Modeling is done with this equation.

$$\mathbf{A} = \mathbf{K} \times \mathbf{C} + \mathbf{E} \tag{2.8}$$

A: is the matrix which consists of absorbance values at different wavelengths of samples.

C: is the matrix which consists of concentration values of samples which have multicomponent. If component analysis is done in sample, **c** is used as an vector..

K: is an absorptivity coefficient and also is the matrix multiplied by path length. Each member of this matrix corresponds to absorptive coefficient of an absorption value at a certain wavelength.

E: is the error matrix.

K matrix is calculated by the following formula.

$$\mathbf{K} = (\mathbf{C}' \cdot \mathbf{C})^{-1} \cdot \mathbf{C}' \cdot \mathbf{A} \tag{2.9}$$

In the prediction step, an unknown sample spectrum is measured. (**â**).Given **â** and **K**, concentration can be predicted by using simple matrix algebra:

$$\hat{\mathbf{c}} = \hat{\mathbf{a}} \cdot \mathbf{K}' \cdot (\mathbf{K} \cdot \mathbf{K}')^{-1} \tag{2.10}$$

The notations of prediction elements are vector,because there are more than one component and there are more than one absorbance value in one unknown sample. The difference between the reference and predicted concentration values is residual. Residual is calculated the following formula:

$$\mathbf{e} = \mathbf{c} - \hat{\mathbf{c}} \tag{2.11}$$

CLS method can be applied to simple systems where all of the pure-component spectra can be measured. The pure-component spectra are measured for each analyte in the sample. These are utilized to form spectral matrix and the CLS model is then

constructed. This calibration model is used to predict the concentrations of components in unknown samples.

CLS method has some advantages and disadvantages. In a wide spectral, absorbance measurements in a large number of wavelength can be used in calibration. One of the advantages of this method is that calculations are fast and the selection of wavelength is not required. Small number of samples is needed to construct the calibration model. Since many variables are used, it is possible to overcome overlapping problems (Beebe, et al. 1998). All the components in the sample must be known in this calibration method. CLS calibration is not suitable for the analysis of mixtures containing components that interact with each other.

2.1.3.2 Inverse Least Squares (ILS)

In some cases, CLS may not work because the system of interest is not simple or it may not be possible to obtain the pure spectra of all the analyte in the unknown samples. This need can be eliminated by using Inverse least squares (ILS) method. ILS method involves the application of linear systems the reverse statement of Lambert Beer law. Briefly, it is accepted that concentration varies depending on absorbance. Modeling is:

$$\mathbf{C}=\mathbf{A}\times\mathbf{P}+\mathbf{E} \tag{2.12}$$

C is the concentration matrix, **A** is the absorbance matrix **E** is the error matrix. P matrix contains the model coefficients and can be determined by:

$$\mathbf{P}=(\mathbf{A}'\cdot\mathbf{A})^{-1}\cdot\mathbf{A}'\cdot\mathbf{C} \tag{2.13}$$

A predicted concentration of a multi-component sample can be obtained by:

$$\hat{\mathbf{c}}=\hat{\mathbf{a}}\cdot\mathbf{P} \tag{2.14}$$

The residual is, as in the CLS model, the difference between the reference and predicted concentration values.

$$\mathbf{e}=\mathbf{c}-\hat{\mathbf{c}} \tag{2.15}$$

In conclusion, ILS can be used to construct accurate calibrations when just knowing the concentrations of analytes in the sample. That means there is no need to know all of the components in the sample. To use ILS, one should be selected as many variables as there are sources of variation in the system instead of using the full spectra. The weaknesses of ILS are that it has limited outlier detection and there is no efficient method for optimal wavelength selection for predictive models. Also collinearity between the absorbance values causes problems the validation of the model because it prevents stabilization of the predictions against noise in absorbances. So, it is very important to select the best set of wavelengths to use in the construction of calibration (Beebe, et al. 1998).

2.1.3.3 Genetic Inverse Least Squares (GILS)

GILS is a modified version of ILS. In this version genetic algorithms (GA) are used as a tool for wavelength selection. GA is global search and optimization method based on the principles of natural evolution and selection developed by Darwin (Wang, et al.1991). According to the Darwin's theory of evolution, individuals who fit better to the environment are more likely survive and breed, thus are able to pass their genetic information to their offspring. However, individuals who do not fit and unable to adapt will eventually be eliminated from the population. This process progresses slowly over a long period of time.

In early 1960's the studies on GA is done by Holland for the first time. He developed a genetic algorithm on adaptive systems in his research so, he was considered as a father of this field (Gilbert, et al. 1997). Over the years, GA have attracted attention and have been applied to various global optimization problems in many areas including (Fontain 1992, Cong and Li 1994, Wienke, et al. 1993, Hibbert 1993, Lucasius and Kateman 1991). In terms of calibration, there have been several applications of GA to wavelength selection (Lucasius, et al. 1994, Lucasius and 19 Kateman 1992, Paradkar and Williams 1997, Ozdemir, et al. 1998a, Ozdemir, et al. 1998b, Ozdemir and Williams 1999).

Generally genetic algorithm consists of five basic steps. These steps have taken their names from the biological foundation of the algorithm. The implementation of a typical GA is shown in Figure 2.3.

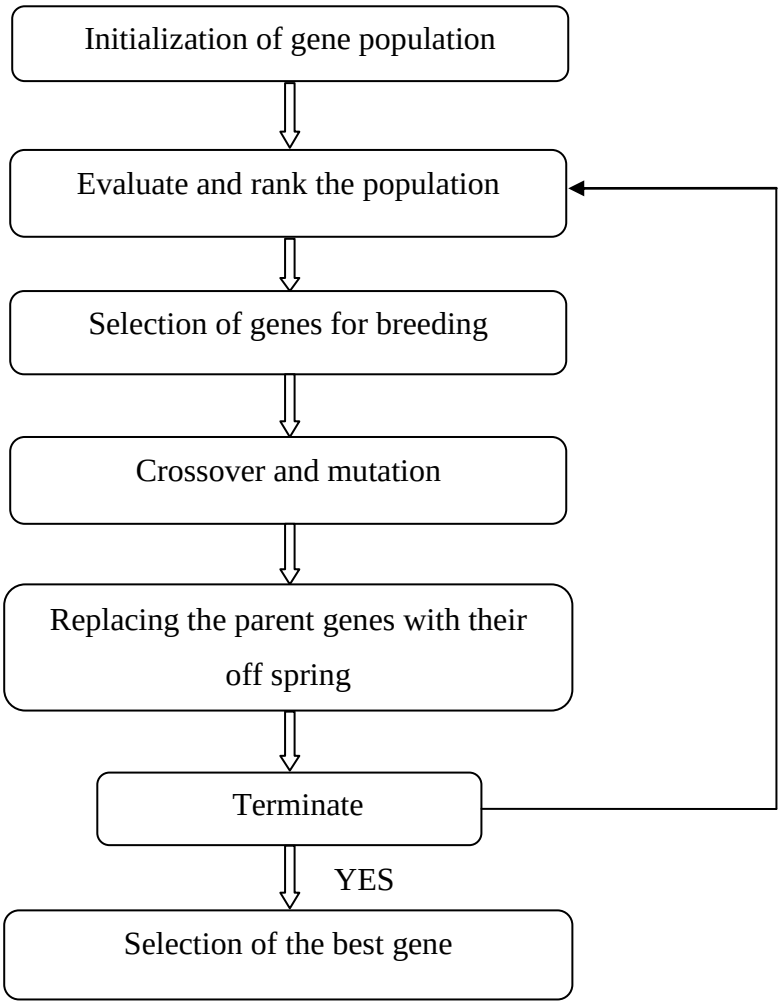


Figure 2.3. General flow chart of genetic algorithm used in GILS

2.1.3.3.1 Initialization

A gene is defined as a potential solution to a given problem which changes from application to application and depends upon the problem being investigated. In the GILS method, the term ‘gene’ is referred as the collection of instrumental response at the wavelength range of the data set. Each gene finds a value between instrument signal and components with concentration. Population is used to describe the collection of individual genes in the current generation.

In the initialization step, in the initial gene pool, a gene consists of absorbance values at randomly chosen wavelengths between predefined limit and upper limit. Because the iteration time of algorithm varies from direct proportion to the number of gene. So, this determines the operating time. An example gene is provided below:

$$S = [A_{8432} A_{6895} A_{5128}]$$

Where S defines a gene and A is the absorbance measured at the indicated wavelength. Each absorbance value is a vector of sample and these vectors forms the new absorbance matrix. This matrix is also defined as a gene. the population is formed according to the number of genes initially entered as an input of the software.

2.1.3.3.2 Evaluate and Rank the Population

In order to evaluate each gene's success in the prediction of analyte concentration, fitness function such as the reciprocal of standard error of calibration (SEC) is used. SEC is calculated from the ILS model in which absorbance values from the selected wavelengths are used to construct the model. SEC is calculated from the following equation:

$$\text{SEC} = \sqrt{\frac{\sum_{i=1}^{m} (c_i - \hat{c}_i)^2}{m - 2}} \qquad (2.16)$$

Where c_i is the reference and $\hat{c}_i$ is the predicted concentration values of i^{th} sample. m is the number of samples. The degree of freedom is $m-2$ because when a linear model is assumed, there are only two parameters to be extracted which are the slope of the actual vs. reference concentration plot and the intercept. In each step, increase in the fitness value is targeted.

2.1.3.3.3 Selection of Genes for Breeding

The third step depends on the selection of the parent genes from the current population for breeding. The selection is made by using a selection method according to their fitness values. The goal of a selection method is to give higher chance to those genes with higher fitness so that only the best performing members of the population will survive in the long run and will be able to transfer their information to next

generations. Here, it is expected that the genes better suited for the problem will generate better off-springs. The genes with low fitness values will have lower chance to breed and as a consequence most of them will be unable to survive (Wang, et al. 1991).

There are a variety of selection methods that can be used for parent (Wang, et al. 1991). The simplest selection method is the top down for parent selection. After genes are ranked in the current gene pool, they are allowed to mate in a way that the first gene mates with the second gene, third one with the forth one and so on. All the members of the current gene are given a chance to breed. The roulette wheel selection method, which is used in GILS, is the one where the chance of selecting gene is directly proportional to its fitness. In this method, each slot in the roulette wheel represents a gene. The gene with the highest fitness has the slot that has the largest area and the gene with the lowest fitness has the slot that has the smallest area. Therefore, when the wheel is rotated, there is a higher chance of selection for a gene with high fitness than for a gene with a low fitness. There will also be the genes which are selected multiple times and some of the genes will not be selected at all.

2.1.3.3.4 Crossover and Mutation

The genetic algorithm does most of its work in the breeding/mating step. After the selection of parent genes is completed, all of them mate to produce their offspring by crossing over until there is no more rest. The step involves breaking the genes at random points and cross-coupling them as illustrated in the following example:

Consider S_1 and S_2 are parent genes which are to breed; S_3 and S_4 are their corresponding off-springs.

$S_1=[A_{4135}A_{5442}\#\ A_{9217}A_{4320}]$

$S_2=[A_{5123}A_{8397}A_{9743}A_{7832}\#\ A_{8522}A_{9210}]$

$S_3=[A_{4135}A_{5442}A_{8522}A_{9210}]$

$S_4=[A_{9217}A_{4320}A_{5123}A_{8397}A_{9743}A_{7832}]$

Here, the first part of S_1 is combined with the second part of S_2 to give S_3. Likewise the second part of S_1 with the first part of S_2 to give S_4. This process is called single point crossover and it is used in GILS. The symbol # is used to indicate the separation of the genes and the place where crossover occurs. Singlepoint crossover will

not provide different off-spring if both parent genes are identical, which may happen in the roulette wheel selection, and broken at the same point. To avoid this problem, two points crossover, where each gene is broken in two points and recombined, can be used.

Mutation, which introduces random deviations into the population, can be also introduced into the algorithm during the mating step at a rate of 1% as is typical in GA.

Replacing one of the wavelengths in an existing gene with a randomly generated new wavelength usually does this. However, in this study it is not used in GILS.

2.1.3.3.5 Replacing the Parent Genes by their Offspring

After crossover, the parent genes are replaced by their off-springs. The ranking process based on their fitness values follows the evolution step. Then the selection for breeding/mating starts again. This is repeated until a predefined number of iterations are reached.

At the end, the gene with the lowest SEC (highest fitness) is selected for model building. This model is used to predict the concentrations of component being analyzed in the validation set. The success of the model in the prediction of the validation set is evaluated using standard error of prediction (SEP) which is calculated as:

$$SEP = \sqrt{\frac{\sum_{i=1}^{m} (c_i - \hat{c}_i)^2}{m}} \qquad (2.17)$$

m is the number of samples in validation sets.

2.1.3.3.6 Termination

The termination of the algorithm is done by setting the predefined iteration number for the number of breeding/mating cycles. No extensive statistical test has been done to optimize it, though it can also be optimized. Since the random processes are heavily involved in the GILS, the program is set to run predefined number of times for each component in a given multi-component mixture. The best run, i.e. the one

generating the lowest SEC for the calibration set and at the same time obtained SEP for the validation set that is in the same range with SEC, is subsequently selected for the evaluation and further analysis.

GILS has some major advantages over the classical univariate and multivariate calibration methods. First of all, it is quite simple in terms of the mathematics involved in the model building and prediction steps, but at the same time it has the advantages of the multivariate calibration methods with a reduced data set since it uses the full spectrum to extract genes. By selecting a subset of instrument responses, it is able to eliminate nonlinearities that might be present in the full spectral region.

2.1.3.4. The Eigenvector Quantitation Methods

In real samples, spectrum can be made by using different variables such as the components of sample, inter-component interactions, instrument variations (i.e., detector noise), changing environmental conditions, differences in sampling handling.

When the spectral data are taking place, there should be finite number of these variations even if with all these complex takes place. Hopefully the largest variations in the calibration set would be the concentrations of the constituents of the mixtures in the spectrum. If it was possible to calculate a set of "variation spectra" that represented the changes in the absorbances at all the wavelengths in the spectra, then this data could be used instead of the raw spectral data for building the calibration model. The "variation spectra" could be used to rebuild the spectrum of sample by multiplying each one by a different constant scaling factor and adding the results together until the new spectrum closely matches the unknown spectrum. Because of the difference between all the concentrations of constituent, each spectrum in the calibration set would have a different set of scaling constants for each variation. Therefore, the fraction of each spectrum that must be added to construct the unknown data should be related to the concentration of the constituents. The "variation spectra" are often called eigenvectors or spectral loadings or loading vectors or principle components or factors. The scaling constants used to reconstruct the spectra are generally known as scores.

The eigenvectors must relate to the concentrations of the constituents that make up the samples, since they came from the original calibration data.

The calculated scores are unique to each separate principle component and training spectrum, and can be used in place of absorbances. Since the representation of the mixture spectrum is reduced from many wavelengths to a few scores as shown in Figure 2.4. This method is combining both the CLS and ILS methods together in the same calculation. Since it is better than the classical models in the meaning of accuracy and robustness.

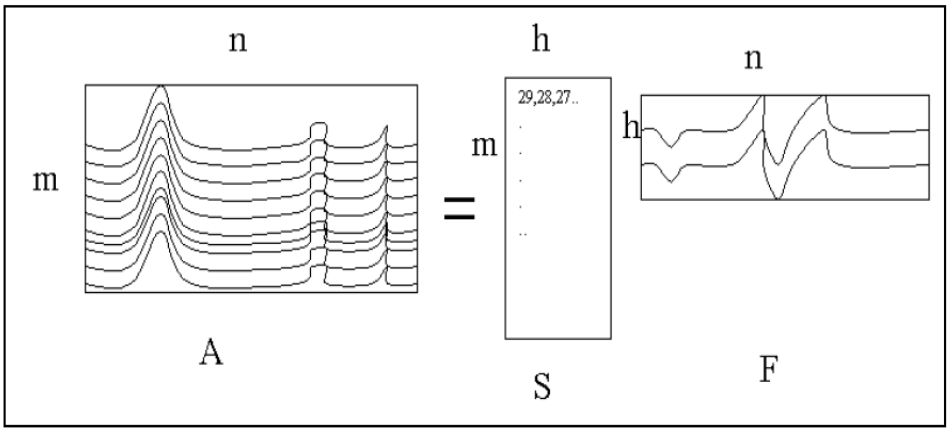


Figure 2.4. PCA breaks the spectral data into most common spectral variations (Factors, eigenvectors, loadings) and the corresponding scaling coefficients (scores)

The trick in using these models comes from the calculation of the eigenvectors. These models are based on the concentration predictions and changes in the data, not the absolute absorbance measurements that are used in all classical models.

In order to calculate the PCA model, the spectral data must change in some way. To accomplish this it is the best way to vary the concentration of the constituent. Since there can be problem with colinearity. For example, if the concentrations of the two constituents present always in same ratio, the model will detect only one constituent not two. Also not only the concentration of the constituents if the absorbance peak of the A increase or decrease when constituent B also increases or decreases, only one variation will be detected and this is the changes in the mixture of A and B. Therefore, it is very important to have randomly concentration ratios in the mixtures.

2.1.3.5. Partial Least Squares

Partial Least Squares (PLS) is another method that also calculates the variations in the spectra. It is soft modeling techniques in which the data are decomposed into new variables that are linear combinations of the original data. This new variable is named as principal components or factors and therefore, PLS is often called factor methods. The way in which the new variables are created can be visualized for a two dimensional system. If the instrument responses for a set of m samples at two wavelengths (n=2) are plotted against each other, a new axis is formed in the direction that represents maximum variability of the data. This new axis is called first principle component or first eigenvector. If all the samples fall on this new axis, then all of the variations can be described using only one eigenvector (Hartnett, 1997). Otherwise a second eigenvector can be found that is perpendicular or orthogonal to the first eigenvector. The second one describes the maximum amount of residuals, not fit by the first one, in the data set and so on. If more than two wavelengths are included in instrument response matrix, the plotting space becomes multidimensional and several eigenvector can be found, each one successfully accounting for the maximum possible amount of remaining variability and each orthogonal to others. In general, the number of principle component or factor that can be generated is less than or equal to the number of sample (Haaland, et al 2002).

PLS is full spectrum method so it retains the full spectrum advantages of CLS. However, all of the component concentrations need to be known because, both the PLS can perform the analysis one component at a time while avoiding the ILS wavelength selection problems. PLS and PCR differ in the way the matrix of the spectra decomposed into two smaller matrices. In the PCR, decomposition is performed independently of analyte concentration whereas in the PLS, the concentration information is used to extract factors. Therefore, the PLS method is expected to provide better calibration models and prediction (Mark,1986). The model for either the PLS is described as:

$$\boldsymbol{A} = \boldsymbol{TB} + \boldsymbol{E_A} \qquad (2.18)$$

where A is the same before, B is a hxn matrix of basis vectors or loading spectra. T is an mxh matrix of intensities or scores in the new coordinate system defined

by the h loading vectors. E_A is now the mxn matrix if spectral residuals not fit by the model. The difference between CLS and these factor methods is that the loading vectors in B are not pure component spectra but they are linear combinations of the original calibration spectra. Also the intensities in the new coordinate system are no logger constrained to the concentrations as were in CLS, but modeling can be done to relate the scores in T to the component concentrations. The number of basis vectors, h, to represent original calibration spectra is determined by an algorithm during the calibration step.

The spectral intensities in the new coordinate system can be related to the concentrations of the analyte with an ILS model given by:

$$\mathbf{c} = \boldsymbol{T}\boldsymbol{v} + \mathbf{e_c} \tag{2.19}$$

where c is the mxl vector of component concentrations, *v* is the hxl vector of coefficients which relate spectral intensities to the component concentration and e_c is the mxl vector of errors in reference values of the component that is being analyzed. However, since the columns of the T matrix are orthogonal, inversion of the diagonal $(T^T T)$ matrix is trivial. The estimate of v vector is given as:

$$\hat{\boldsymbol{v}}_{\boldsymbol{h}} = (\boldsymbol{T}^{\boldsymbol{T}} \boldsymbol{T})^{-1} \boldsymbol{T}^{\boldsymbol{T}} \boldsymbol{c} \tag{2.20}$$

where $\hat{\boldsymbol{v}}_{\boldsymbol{h}}$ is the least-squares estimate of *v*. The T and B matrices are calculated in a stepwise manner (one vector at a time) until the desired model has been obtained. As mentioned earlier, PLS and PCR differ in the way they generate T and B matrices. In the PCR model, NIPALS (nonlinear iterative partial least squares) algorithm developed by Wold (Wold, 1966) is used. The NIPALS algorithm extracts the full spectrum loading vectors without using concentration information in the decomposition of spectral matrix A. Therefore; the prediction of component concentrations is expected to be poorer that the results obtained by PLS which applies a modified version of NIPALS algorithm (Haaland, et al, 2002). This modified version of the algorithm uses concentration information in the process of obtaining loading vectors thereby resulting in a generator predictive ability.

There are two PLS methods that are available today in the analysis of complex chemical mixtures. These are called PLS1 and PLS2 methods. In the PLS1 method, the analysis performed one component at a time and other component concentrations not

included in the model building step. This is the most commonly used form the PLS method and it is reported that the predictions obtained with PLS1 are better that those obtained PLS2. It is suggested that PLS2 algorithm should be used for qualitative application.

Before applying the factor based methods to the data, it is common practice to do some sort of data pretreatment such as mean centering and scaling (Kowalski, et al, 1986).The mean centering is usually applied to both calibration spectra and corresponding analyte concentrations in which the average concentrations for the component of interest are subtracted from each spectrum and from given component concentrations, respectively. After the data pretreatment, a CLS calibration model is selected for the analysis of one component at a time. Then the PLS1 algorithm starts with the calculation of the estimated first weighed loading vector, $\widehat{\boldsymbol{w}}_{\boldsymbol{h}}$, by setting h to 1. This is done with the method of least squares and is given by:

$$\widehat{\boldsymbol{w}}_{\boldsymbol{h}} = \mathbf{A}^{\mathbf{T}}\,\mathbf{c}\,(\mathbf{c}^{\mathbf{T}}\,\mathbf{c})^{-1} \tag{2.21}$$

where $\widehat{\boldsymbol{w}}_{\boldsymbol{h}}$ is an *nxl* vector representing the first order approximation of the pure component spectra for the component that is being analyzed. This weighted loading vector is then used to form the score vector $\hat{t}_h$, with an ILS prediction model. The method of least squares is used to regression of A on $\widehat{\boldsymbol{w}}_{\boldsymbol{h}}$ which produces the first estimated $\hat{t}_h$ vector as given:

$$\hat{\boldsymbol{t}}_{\boldsymbol{h}} = \mathbf{A}\,\widehat{\boldsymbol{w}}_{\boldsymbol{h}} \tag{2.22}$$

With a linear least-squares regression, this score vector can be related to the component concentrations. The scalar regression coefficient, $\widehat{\boldsymbol{v}}_{\boldsymbol{h}}$, is estimated by:

$$\widehat{\mathbf{V}}_{\boldsymbol{h}} = \hat{\boldsymbol{t}}_{\boldsymbol{h}}^{\boldsymbol{T}}\,\boldsymbol{c}\,(\hat{\boldsymbol{t}}_{\boldsymbol{h}}^{\boldsymbol{T}}\,\hat{\boldsymbol{t}}_{\boldsymbol{h}})^{-1} \tag{2.23}$$

The least-square estimated regression coefficient is later used to obtain concentration residuals. In order to eliminate collinearty problems, the PLS loading vector, $\hat{b}_h$, is now calculated with a new model for A. Once again the method of least squares is used to find estimated b vector by:

$$\widehat{\boldsymbol{b}}_{\boldsymbol{h}} = \hat{\boldsymbol{t}}_{\boldsymbol{h}}^{\boldsymbol{T}}\,\mathbf{A}\,(\hat{\boldsymbol{t}}_{\boldsymbol{h}}^{\boldsymbol{T}}\,\hat{\boldsymbol{t}}_{\boldsymbol{h}})^{-1} \tag{2.24}$$

where $\widehat{b}_h$ is an nx1 vector. It is now possible to calculate the first PLS approximation to the calibration spectra by multiplying the score vector ($\hat{t}_h$) with transpose of PLS loading vector ($\hat{b}_h^T$). The first residual matrix is calculated by subtracting the PLS approximation matrix from A matrix. The residuals in concentration vector calculated in a similar manner where scalar regression coefficient ($\hat{v}_h$) is multiplied with score vector and this product is subtracted from original concentration vector. The following equations provide residuals in both A and c.

$$\mathbf{E_A = A - \hat{t}_h \hat{b}_h^T} \tag{2.25}$$

$$\mathbf{e_c = c - \hat{V}_h \hat{t}_h} \tag{2.26}$$

This is the end of the first iteration in the calibration step. This is the process is repeated for a desired number of loading vectors by incrementing *h*, substituting E_A for A and e_c for concentration in the first CLS calibration model at the beginning of the algorithm.

The prediction step of PLS1 algorithm involves the calculation of final calibration coefficients, b_f, which have the dimension of an original spectrum. Once the bf is calculated, it is possible to calculate the concentration of a new sample using the average concentration of the analyte and its spectra. The following equations show the prediction step in PLS1.

$$\mathbf{b_f = \widehat{W} (\widehat{B}\, \widehat{W}^T)^{-1} \hat{v}} \tag{2.27}$$

where $\widehat{W}$ and $\hat{B}$ contains individual $\hat{w}_h$ and $\hat{b}_h$ vectors, respectively and $\hat{v}$ is formed from individual regression coefficients ($\hat{v}_h$) The final prediction equation is then given as:

$$\mathbf{\hat{c} = a^T bf + c_0} \tag{2.28}$$

where $\hat{c}$ is the predicted unknown sample c, a is the spectrum of that sample and c_0 is the average concentration of calibration samples.

The process of determining the optimal number of PLS factors may vary from algorithm. The cross-validation approach is one of the methods for this (Malinowski, 1977). For m calibration spectra, the PLS1 algorithm is performed on m-1 spectra and

the left out spectrum is used to validate the model. This process is repeated until each spectrum is left out once in the calibration set. The predicted concentration for each left out sample is then compared with their original values and the prediction error sum of the squares (PRESS) is calculated for each added factor. The PRESS is a measure of how well a particular model fits the calibration data and given by:

$$\mathbf{PRESS}=\sum_{i=1}^{m}(\hat{c}-c_i)^2 \tag{2.29}$$

where c_i is the reference (known) concentration of the i^{th} sample and concentration is the predicted concentration of the i^{th} sample for m calibration standard.

It is not the minimum PRESS value, however, that is used for the selection of optimal number of PLS factors since this may lead to over fitting resulting in a poorer prediction. Therefore a comparison needs to be done between two models that contain h and h+1 factor. Here, the better model is the one with smaller number of factors where the difference between the two PRESS values is determined by the F test to be significant.

CHAPTER 3

INSTRUMENTATION AND EXPERIMENTATION

3.1 Instrumentation

3.1.1. Fourier Transform Infrared (FTIR) Spectrometry

In infrared spectroscopy, IR radiation is passed through a sample. Some of the infrared radiation is absorbed by the sample and some of it is passed through (transmitted). The resulting spectrum represents the molecular absorption and transmission, creating a molecular fingerprint of the sample. Like a fingerprint no two unique molecular structures produce the same infrared spectrum. This makes infrared spectroscopy useful for several types of analysis.

It can used ;

* identify unknown materials,

* determine the quality or consistency of a sample,

* determine the amount of components in a mixture.

Molecules are flexible, moving collections of atoms. The atoms in a molecule are constantly oscillating around average positions. Bond lengths and bond angles are continuously changing due to this vibration. A molecule absorbs infrared radiation when the vibration of the atoms in the molecule produces an oscillating electric field with the same frequency as the frequency of incident IR "light".

All of the motions can be described in terms of two types of molecular vibrations. One type of vibration, a stretch, produces a change of bond length. A stretch is a rhythmic movement along the line between the atoms so that the interatomic distance is either increasing or decreasing.

The second type of vibration, a bend, results in a change in bond angle. These are also sometimes called scissoring, rocking, or "wig wag" motions. Each of these two main types of vibration can have variations. A stretch can be symmetric or asymmetric. Bending can occur in the plane of the molecule or out of plane; it can be scissoring, like blades of a pair of scissors, or rocking, where two atoms move in the same direction.

(Thermo Nicolet, Introduction to FTIR, 2012).

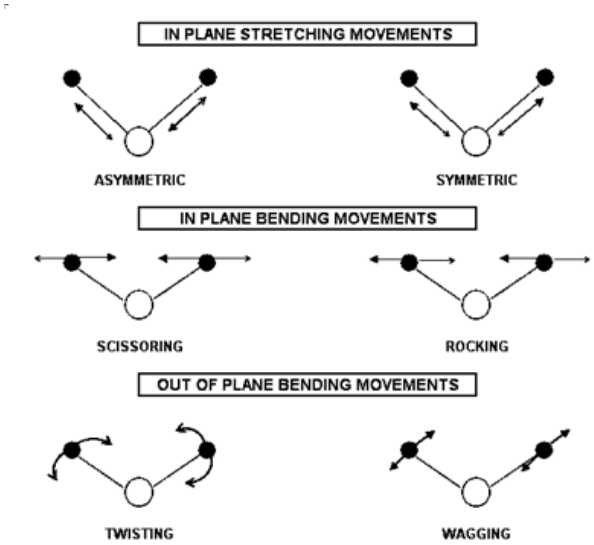


Figure 3.1.Types of molecular vibrations. + indicates motion from the page toward thereader; - indicates the motion away from the reader (Source: Skoog, et al.1998)

The normal instrumental process is as follows:

1. The Source: Infrared energy is emitted from a glowing black-body source. This beam passes through an aperture which controls the amount of energy presented to the sample (and, ultimately, to the detector).

2. The Interferometer: The beam enters the interferometer where the "spectral encoding" takes place. The resulting interferogram signal then exits the interferometer.

3. The Sample: The beam enters the sample compartment where it is transmitted through or reflected off of the surface of the sample, depending on the type of analysis being accomplished. This is where specific frequencies of energy, which are uniquely characteristic of the sample, are absorbed.

4. The Detector: The beam finally passes to the detector for final measurement. The detectors used are specially designed to measure the special interferogram signal.

5. The Computer: The measured signal is digitized and sent to the computer where the Fourier transformation takes place. The final infrared spectrum is then presented to the user for interpretation and any further manipulation.

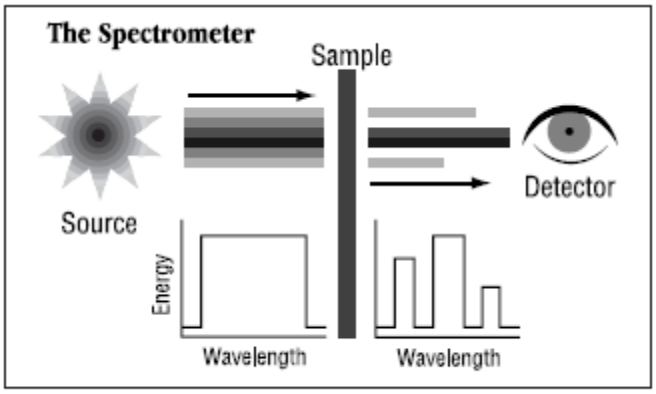


Figure 3.2. Optical diagram of Fourier Transform Infrared (FTIR) Spectrometer (Source: Thermo Nicolet, 2012)

Some of the major advantages of FT-IR over the dispersive technique include:

• Speed: Because all of the frequencies are measured simultaneously, most measurements by FT-IR are made in a matter of seconds rather than several minutes. This is sometimes referred to as the Felgett Advantage.

• Sensitivity: Sensitivity is dramatically improved with FT-IR for many reasons. The detectors employed are much more sensitive, the optical throughput is much higher which results in much lower noise levels, and the fast scans enable the coaddition of several scans in order to reduce the random measurement noise to any desired level.

• Mechanical Simplicity: The moving mirror in the interferometer is the only continuously moving part in the instrument. Thus, there is very little possibility of mechanical breakdown.

• Internally Calibrated: These instruments employ a HeNe laser as an internal wavelength calibration standard. These instruments are self-calibrating and never need to be calibrated by the user.

These advantages, along with several others, make measurements made by FT-IR extremely accurate and reproducible. Thus, it a very reliable technique for positive identification of virtually any sample. The sensitivity benefits enable identification of even the smallest of contaminants. This makes FT-IR an invaluable tool for quality control or quality assurance applications whether it is batch-to-batch comparisons to quality standards or analysis of an unknown contaminant. In addition,the sensitivity and accuracy of FT-IR detectors, along with a wide variety of software algorithms, have dramatically increased the practical use of infrared for quantitative analysis.

Quantitative method can be easily developed and calibrated and can be incorporated into simple procedures for routine analysis.

3.1.2. Near Infrared (NIR) Spectrometry

Infrared energy is the electromagnetic energy of molecular vibration. The energy band is defined for convenience as the near infrared (0.78 to 2.50 microns); the infrared (or mid-infrared) 2.50 to 40.0 microns; and the far infrared (40.0 to 1000 microns). However, even though official standards, textbooks, and the scientific literature generally state that the NIR spectral region extends from 780-2500 nanometers (12821 - 4000 cm-1), a simple set of liquid phase hydrocarbon spectra demonstrates that the vibrational information characterized by the harmonic vibrations of the C-H stretch fundamental and their corresponding combination bands occurs from approximately 690 to 3000 nm. The predominant near-infrared spectral features include: the methyl C-H stretching vibrations, methylene C-H stretching vibrations, aromatic C-H stretching vibrations, and O-H stretching vibrations. Minor but still important spectral features include: methoxy C-H stretching, carbonyl associated C-H stretching; N-H from primary amides, secondary amides (both alkyl, and aryl group associations), N-H from primary, secondary, and tertiary amines, and N-H from amine salts.

The advantages touted for NIR measurements over other vibration techniques have proven themselves true throughout the 1980s up until today, they include:

1) C-H associated vibrational information is repeated 8 times from 690 nm to 3000 nm;

2) Simple harmonics may be selected or more information rich combination regions;

3) Low cost instruments with high signal-to-noise (SNR) are simple to make and typically exhibit signal-to-noise ratios (SNR) of 25000-100000:1;

4) High NIR throughput is possible, even when employing low cost fiber optics;

5) Variable path lengths for industrial use are possible, typically 1 mm to 10 cm or more using different NIR spectral regions;

6) NIR Light penetrates plant and animal tissue easily for biomedical applications (when using 900 nm and longer).

Qualitative and quantitative near infrared (NIR) spectroscopic methods typically require the application of multivariate calibration algorithms and statistical methods (i.e. chemometrics) to model NIR spectral response to chemical or physical properties of the samples used for calibration. The NIR method relies on the spectra-structure correlations existing between a measured spectral response caused by harmonics of the fundamental vibrations occurring at infrared frequencies. These harmonic vibrations occur at unique frequencies depending upon the quantity of absorber (analyte), type of absorbing molecules present within the sample, and the sample thickness.

Quantitative methods are possible where changes in the response of the near infrared spectrometer are proportional to changes in the concentration of chemical components, or in the physical characteristics (scattering/absorptive properties) of samples undergoing analysis. Recent refinements of the NIR measurement technique include the emergence of chemometrics and the diminishing distinction between near infrared, and infrared as measurement techniques. Rather the techniques are complementary, with each spectral region providing unique advantages for the analyst. For a wide range of NIR applications, particular attention is given to the appearance of methyl, methylene, methoxy, carbonyl, and aromatic C-H groups; hydroxy O-H; and N-H from amides, amines, and amine salts.

Near infrared spectroscopy is used where multicomponent molecular vibrational analysis is required in the presence of interfering substances. The near infrared spectra consist of overtones and combination bands of the fundamental molecular absorptions found in the mid infrared region. Near infrared spectra consist of generally overlapping vibrational bands that may appear non-specific and poorly resolved. The use of chemometric mathematical data processing and multiple harmonics can be used to calibrate for qualitative of quantitative analysis despite these apparent spectroscopic limitations.(Howard and Campbell,2008)

3.1.3. Advantages and Disadvantages of Spectroscopic Techniques

The biggest advantage of spectroscopic techniques is little or no sample preparation, and real-time data. Unlike most conventional analytical methods, NIRS is rapid, non-destructive, does not use chemicals, or generate chemical wastes requiring disposal, simultaneously determines numerous constituents or parameters, and can be

transported to nearly any environment, or true portable for field work. NIR instrumentation is simple to operate by non-chemists, and operates without fume hoods, drains, or other installations. NIR is not a stand-alone technology. Its accuracy is dependent upon the accuracy of the reference method used for training; however, the data from the NIR method has better reproducibility than the primary method.

Another advantage of spectroscopy is 'thermal' noise. All internal electronic components are a source of thermal noise in the Mid-IR and Far-IR. However, internal sources of IR are either insignificant to NIR detectors or can be made insignificant by minor shielding.

Spectroscopy is not a stand-alone technology. Separate calibrations are required for each constituent or parameter and a portion of unknown samples must periodically be analyzed by the reference method to ensure that calibrations remain reliable. It may be necessary to update calibrations several times during the initial phases of use to incorporate "outlying" samples, until the calibration is acceptable. Despite the intuitive disadvantage of broad and overlapping absorption bands, sophisticated chemometric techniques can extract meaningful information from the complex NIR spectra. The information about samples in the NIR spectra could not easily be accessed until the advent of sufficiently powerful computers that allowed the development of complex statistical relationships between the spectral data and constituents or parameters (e.g. functional properties) determined by conventional techniques. These statistical relationships between the spectral data and data from reference analyses are called calibration models (Howard and Campbell, 2008).

3.1.4. High Performance Liquid Chromatography (HPLC)

High performance liquid chromatography (HPLC) is a chemistry tool for quantifiying and analyzing mixtures of chemical compounds which is used to find the amount of a chemical compounds within a mixture of other chemicals. High performance liquid chromatography (HPLC) has the ability to separate, identify and quantitate the compounds that are present in any sample that can be dissolved in a liquid.

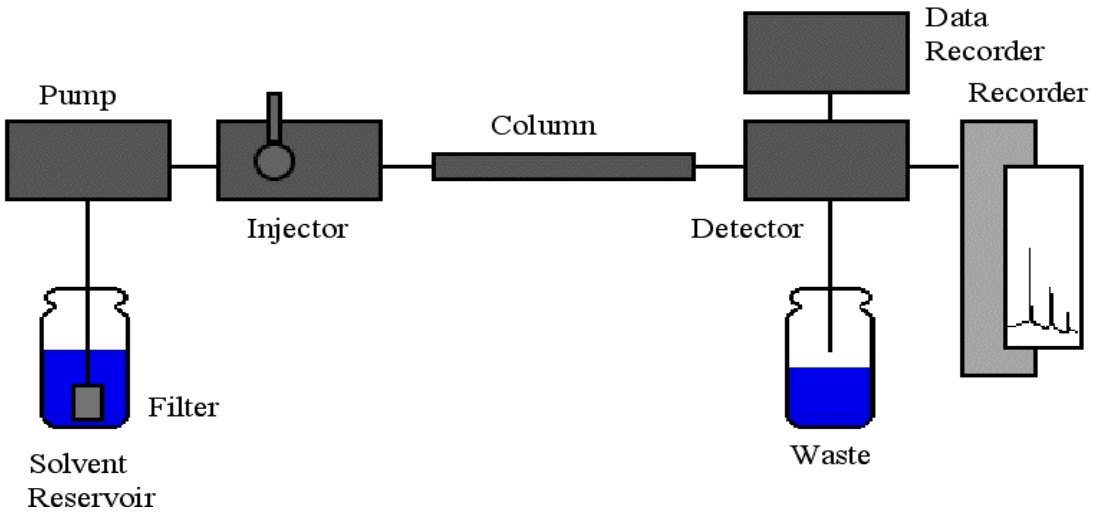


Figure 3.3. Schematic representation of a system for high performance liquid chromatography (HPLC).(Source: Boomer.org,2013)

HPLC is a separation technique that involves the injection of a small volume of liquid sample into a tube packed with tiny particles (3 to 5 micron (μm) in diameter called the stationary phase) where individual components of the sample are moved down the packed tube (column) with a liquid (mobile phase) forced through the column by high pressure delivered by a pump. These components are separated from one another by the column packing that involves various chemical and/or physical interactions between their molecules and the packing particles. These separated components are detected at the exit of this tube (column) by a flow-through device (detector) that measures their amount. An output from this detector is called a "liquid chromatogram" (Agilent Technologies, 2013).

High performance liquid chromatography (HPLC) and combined chromatographic methods has a great emphasis in olive oil analysis techniques. Several minor components of olive oil such as sterols, phenolic compounds, pigments, tocopherols and triacylglycerols can be identified and quantitated with this technique. Reversed-phase high performance liquid chromatography (RP-HPLC) currently is the most popular and reliable technique for the determination of triacylglycerols. Numerous mobile phases have been employed with different modifiers, which include methanol, acetonitrile or tetrahydrofuran (Ryan, et al. 1999). Percentage determination of the various triglycerides present in virgin olive oil or high performance liquid chromatography offers a way of detecting possible adulterations with oils which, while having a similar fatty acid composition to olive oil, have a different triglyceride composition.

3.1.5. Gas Chromatography (GC)

Gas chromatography (GC) is a powerful and widely used tool for the separation, identification and quantitation of components in a mixture. In this technique, a sample is converted to the vapor state and a flowing stream of carrier gas (often helium or nitrogen) sweeps the sample into a thermally-controlled column. In the case of gas liquid chromatography, the column is usually packed with solid particles that are coated with a non-volatile liquid, referred to as the stationary phase. As the sample mixture moves through the column, sample components that interact strongly with the stationary phase spend more time in the stationary phase vs. the moving gas phase and thus require more time to move through the column. Retention time is defined as the time from injection of the sample to the time a specific sample component is detected. Components with higher volatility (lower boiling points) tend to spend more time in the moving gas phase and therefore tend to have shorter retention times. After exiting the column the separated components are detected and a detector response is recorded (Figure 3.4).

The most application field of Gas Chromatography (GC) in olive oil analysis is the determination of methyl esters of fatty acids. The aim of this determination is to establish the percentage composition of fatty acids in olive oil, more commonly known as fatty acid composition, which is influenced by the olive variety, production zone, climate and stage of maturity of the drupes when they are collected. Determination of fatty acid composition of olive oil is not only a quality indicator but also is used for characterization of the oils.

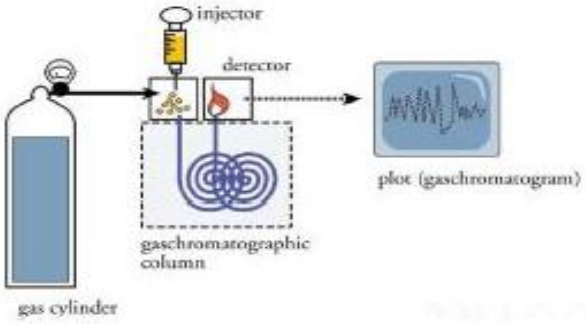


Figure 3.4. Schematic representation of a system for gas chromatography (GC)
(Source: Oliveoil, 2012)

3.2 Experimentation

3.2.1 Olive Oil Samples

In this study, Gemlik variety of olive oil samples are used harvest for the year 2011.Marmara and Manisa region samples are provided by the Olive Research Institute. Also at two locations within the Olive Research Institute (Bornova and Kemalpaşa) trees in the Gemlik variety, samples are produced a three-phase continuous system. Data of these samples will be used as reference. The samples were stored in dark brown bottles in deep freezer until they were analyzed. Details about the region of samples are given in Table 3.1. Also all FAME and TAG components, are shown in Table 3.2 and Table 3.3. The components used in this study are marked in bold.

Table 3.1 : Regions and settlements, that collected olive oils of which harvest Gemlik variety.

Olive Research Institute:1)Olive Oil Research Institute Garden/Bornova 2) Olive Oil Research Institute Garden Garden/ Kemalpaşa
Manisa region (Akhisar and Salihli): 1)Dereköy, Ballıca, Zeytinliova, Mecidiye, Beyoba, Kayalıoğlu/Akhisar 2) Dombaylı, Kestelli, Görece, Pazarköy, Borlu/Salihli
Marmara Region: 1)Mudanya, Gemlik,İznik, Orhangazi/Bursa 2)Marmara Birlik 3)Erdek,Edincik/Balıkesir 4)Mürefte/Tekirdağ

Table 3.2. : FAME components of olive oil

FATTY ACIDS	TCR	DRA	MF	TR1	KY1	İLY	BYL	KK	TR 6	İZN	SR1	KRK	AKY	ED	KST	MNV	MDN	EDN	ÇKR	GM	SR 2	BR 2	TR 5
MA	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
PA*	11.77	12.03	11.94	13.48	13.70	12.43	11.86	12.87	11.99	12.47	14.31	12.87	14.86	12.54	12.77	14.87	12.11	12.32	11.54	12.13	14.30	12.25	11.89
POA*	1.08	1.11	0.97	1.56	1.24	1.02	1.08	1.30	1.24	1.20	1.24	1.10	1.53	1.36	1.23	1.47	1.20	1.26	0.94	1.27	1.40	1.09	1.66
MG	0.14	0.14	0.15	0.12	0.12	0.15	0.14	0.12	0.13	0.15	0.12	0.13	0.13	0.15	0.15	0.15	0.13	0.15	0.17	0.12	0.13	0.13	0.11
MGO	0.25	0.25	0.24	0.22	0.25	0.25	0.25	0.24	0.24	0.26	0.22	0.24	0.23	0.26	0.26	0.21	0.24	0.26	0.29	0.23	0.23	0.26	0.24
SA*	3.09	3.02	3.39	2.89	3.06	2.94	3.06	2.85	3.08	3.07	2.41	2.62	2.81	3.32	3.04	2.65	3.10	3.27	3.18	3.05	2.66	2.60	2.67
EA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OA*	74.11	73.41	73.33	69.59	72.22	74.60	72.22	70.42	71.94	72.58	67.95	71.90	67.16	71.70	73.31	66.46	71.73	71.56	73.29	71.04	69.11	73.03	71.95
VA	1.40	1.62	1.63	2.30	1.57	1.67	1.57	1.65	1.98	1.91	2.15	1.84	2.24	2.26	1.66	2.00	1.79	1.79	1.35	2.02	2.15	1.52	2.13
LO*	6.63	6.91	6.74	8.32	8.19	5.26	8.19	9.00	7.76	6.77	9.96	7.62	9.55	6.71	5.90	10.69	8.05	7.64	7.56	8.46	8.39	7.50	7.70
TLO +TLN	0.05	0.07	0.06	0.06	0.06	0.06	0.06	0.05	0.06	0.06	0.05	0.07	0.05	0.05	0.06	0.05	0.05	0.07	0.07	0.05	0.05	0.06	0.04
LN*	0.56	0.56	0.57	0.61	0.62	0.62	0.62	0.64	0.64	0.64	0.66	0.66	0.66	0.66	0.67	0.67	0.67	0.68	0.68	0.68	0.69	0.69	0.69
AA	0.45	0.42	0.49	0.40	0.45	0.46	0.45	0.40	0.45	0.43	0.41	0.44	0.38	0.49	0.45	0.38	0.44	0.49	0.44	0.45	0.42	0.39	0.42
GA	0.28	0.27	0.29	0.25	0.30	0.31	0.30	0.27	0.29	0.27	0.30	0.31	0.25	0.29	0.29	0.26	0.29	0.29	0.29	0.30	0.29	0.29	0.30
BA	0.12	0.11	0.13	0.10	0.12	0.13	0.12	0.11	0.12	0.11	0.12	0.13	0.09	0.13	0.12	0.10	0.11	0.14	0.12	0.12	0.12	0.11	0.12
LG	0.05	0.05	0.06	0.05	0.05	0.06	0.05	0.05	0.05	0.05	0.06	0.06	0.05	0.06	0.06	0.05	0.05	0.06	0.05	0.05	0.06	0.05	0.06
SQ	0.88	0.79	0.94	1.04	0.94	1.09	0.94	0.77	1.05	0.88	0.62	0.66	0.56	0.94	0.69	0.75	0.88	0.93	0.68	1.03	0.67	0.62	0.84
TFA	0.05	0.07	0.06	0.06	0.06	0.06	0.06	0.05	0.06	0.06	0.05	0.07	0.05	0.05	0.06	0.05	0.05	0.07	0.07	0.05	0.05	0.06	0.04
SFA*	15.63	15.78	16.17	17.06	17.52	16.18	15.69	16.41	15.83	16.29	17.45	16.26	18.33	16.70	16.60	18.21	15.95	16.44	15.51	15.93	17.70	15.54	15.28
MUFA	77.12	76.66	76.46	73.92	75.58	77.85	75.42	73.88	75.69	76.22	71.86	75.39	71.41	75.87	76.75	70.40	75.25	75.16	76.16	74.86	73.18	76.19	76.28
PUFA*	7.19	7.47	7.31	8.93	8.81	5.88	8.81	9.64	8.40	7.41	10.62	8.28	10.21	7.37	6.57	11.36	8.72	8.32	8.24	9.14	9.08	8.19	8.39

(cont. on next page)

Table 3.2. (cont.)

FATTY ACIDS	GM 2	KH	TR 2	MCD2	BLN1	TR 4	MCD10	BLN4	BLN5	GM1	TR3	KY2	BLC3	KY5	ZO 2	BLN3	GRC	MLG	KY4	ZAE-KP	MCD7	DMB1
MA	0.02	0.01	0.01	0.02	0.01	0.02	0.02	0.03	0.02	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.02
PA*	**13.92**	**13.43**	**13.82**	**13.17**	**14.44**	**13.20**	**13.05**	**14.79**	**13.86**	**13.19**	**12.64**	**14.57**	**12.80**	**14.14**	**15.53**	**10.41**	**14.19**	**13.59**	**14.13**	**13.82**	**13.79**	**12.46**
POA*	**1.38**	**1.53**	**1.80**	**1.11**	**1.38**	**1.65**	**1.19**	**1.07**	**1.01**	**1.37**	**1.67**	**1.34**	**1.22**	**1.34**	**1.21**	**0.68**	**1.47**	**1.29**	**1.75**	**1.25**	**1.40**	**1.00**
MG	0.13	0.12	0.11	0.18	0.13	0.11	0.14	0.15	0.13	0.15	0.10	0.12	0.14	0.13	0.14	0.07	0.15	0.16	0.12	0.13	0.13	0.11
MGO	0.24	0.23	0.23	0.30	0.25	0.23	0.27	0.23	0.23	0.29	0.24	0.23	0.28	0.24	0.23	0.13	0.24	0.26	0.25	0.24	0.25	0.21
SA*	**2.59**	**2.86**	**2.67**	**2.93**	**2.66**	**2.69**	**2.24**	**2.58**	**2.35**	**2.62**	**2.47**	**2.36**	**2.47**	**2.64**	**2.67**	**1.95**	**3.20**	**3.14**	**2.46**	**2.56**	**2.51**	**2.38**
EA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OA*	**69.41**	**69.20**	**68.45**	**70.34**	**69.38**	**68.47**	**70.40**	**65.36**	**69.08**	**72.31**	**69.47**	**65.80**	**70.32**	**66.73**	**62.61**	**75.65**	**66.94**	**69.03**	**69.49**	**66.58**	**67.33**	**74.66**
VA	2.15	2.07	2.15	1.51	2.05	2.02	1.90	1.74	1.73	1.84	2.60	1.82	1.53	2.35	1.95	1.43	1.45	1.70	2.03	2.27	2.15	1.67
LO*	**8.49**	**8.90**	**9.15**	**8.70**	**8.03**	**9.98**	**9.07**	**12.26**	**9.81**	**6.49**	**9.12**	**11.92**	**9.41**	**10.57**	**13.87**	**7.64**	**10.52**	**8.90**	**8.01**	**11.42**	**10.62**	**5.26**
TLO +TLN	0.08	0.04	0.04	0.07	0.06	0.05	0.08	0.09	0.06	0.06	0.05	0.07	0.07	0.06	0.07	0.06	0.06	0.06	0.07	0.07	0.10	0.06
LN*	**0.71**	**0.72**	**0.72**	**0.74**	**0.74**	**0.74**	**0.75**	**0.75**	**0.75**	**0.75**	**0.76**	**0.77**	**0.77**	**0.78**	**0.78**	**0.78**	**0.78**	**0.79**	**0.79**	**0.80**	**0.80**	**0.80**
AA	0.41	0.44	0.39	0.43	0.40	0.40	0.38	0.43	0.38	0.43	0.39	0.43	0.41	0.42	0.44	0.43	0.47	0.52	0.41	0.39	0.41	0.43
GA	0.28	0.28	0.26	0.30	0.28	0.27	0.32	0.31	0.30	0.30	0.29	0.32	0.33	0.29	0.28	0.52	0.30	0.32	0.28	0.28	0.31	0.38
BA	0.12	0.12	0.10	0.12	0.11	0.11	0.11	0.12	0.11	0.12	0.11	0.13	0.12	0.11	0.12	0.14	0.12	0.15	0.11	0.11	0.12	0.13
LG	0.06	0.05	0.05	0.06	0.06	0.05	0.06	0.06	0.06	0.07	0.05	0.07	0.06	0.06	0.06	0.08	0.06	0.08	0.06	0.05	0.06	0.07
SQ	0.57	0.80	0.96	0.42	0.68	1.05	0.40	0.40	0.38	0.72	1.22	0.58	0.54	0.63	0.37	0.79	1.09	0.71	0.86	0.45	0.63	1.04
TFA	0.08	0.04	0.04	0.07	0.06	0.05	0.08	0.09	0.06	0.06	0.05	0.07	0.07	0.06	0.07	0.06	0.06	0.06	0.07	0.07	0.04	0.06
SFA*	**17.25**	**17.03**	**17.15**	**16.91**	**17.81**	**16.58**	**16.00**	**18.16**	**16.91**	**16.59**	**15.78**	**17.70**	**16.02**	**17.52**	**18.98**	**13.10**	**18.21**	**17.66**	**17.30**	**17.08**	**17.04**	**15.60**
MUFA	73.46	73.31	72.89	73.56	73.34	72.64	74.08	68.71	72.35	76.11	74.27	69.51	73.68	70.95	66.28	78.41	70.40	72.60	73.80	70.62	71.44	77.92
PUFA*	**9.20**	**9.62**	**9.87**	**9.44**	**8.77**	**10.72**	**9.82**	**13.01**	**10.56**	**7.24**	**9.88**	**12.69**	**10.18**	**11.35**	**14.65**	**8.42**	**11.30**	**9.69**	**8.80**	**12.22**	**11.42**	**6.06**

(cont. on next page)

Table 3.2. (cont.)

FATTY ACIDS	MND	DR 1	AKH2	ZO 1	AKH3	SR 3	DMB2	MCD3	AKH1	MCD9	BLK	BLN5	ZAE-B	KY3	MCD4	KYH	MCD5	BR 1	BLC1	MCD8	BLC2	BLN2
MA	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.03	0.02	0.03	0.03	0.03	0.02
PA*	**14.24**	**14.45**	**12.80**	**13.39**	**13.63**	**14.68**	**15.59**	**14.48**	**13.54**	**13.25**	**15.20**	**15.37**	**15.66**	**14.36**	**14.60**	**14.34**	**13.60**	**12.71**	**14.39**	**14.30**	**12.95**	**13.35**
POA*	**1.57**	**1.36**	**1.54**	**1.40**	**1.24**	**1.63**	**1.52**	**1.45**	**1.03**	**1.25**	**1.64**	**1.67**	**1.40**	**1.55**	**1.58**	**1.49**	**1.38**	**1.20**	**1.39**	**1.49**	**1.38**	**1.67**
MG	0.13	0.14	0.13	0.13	0.12	0.13	0.13	0.15	0.05	0.12	0.13	0.14	0.14	0.12	0.12	0.13	0.11	0.15	0.12	0.11	0.11	0.14
MGO	0.23	0.23	0.26	0.26	0.24	0.25	0.25	0.27	0.08	0.24	0.25	0.25	0.24	0.24	0.24	0.25	0.25	0.28	0.25	0.26	0.25	0.25
SA*	**2.84**	**2.65**	**3.09**	**2.47**	**2.26**	**2.29**	**2.25**	**2.61**	**2.41**	**2.04**	**2.41**	**2.32**	**2.45**	**2.36**	**2.36**	**2.52**	**2.07**	**2.55**	**2.14**	**2.01**	**1.83**	**2.59**
EA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OA*	**66.75**	**67.08**	**72.20**	**67.15**	**67.34**	**67.14**	**67.36**	**68.34**	**66.12**	**67.41**	**65.64**	**66.74**	**62.07**	**67.01**	**66.78**	**67.82**	**65.69**	**67.10**	**65.21**	**65.39**	**65.60**	**66.45**
VA	2.00	1.76	2.29	2.26	1.83	2.30	2.40	2.17	1.78	1.74	2.23	2.14	2.61	2.20	2.01	1.89	1.87	1.87	2.00	2.60	2.09	2.36
LO*	**10.44**	**10.50**	**5.74**	**11.02**	**11.35**	**9.70**	**8.62**	**8.62**	**12.97**	**12.05**	**10.60**	**9.37**	**13.50**	**10.17**	**10.36**	**9.55**	**12.80**	**12.14**	**12.55**	**12.29**	**13.80**	**9.09**
TLO +TLN	0.06	0.06	0.06	0.08	0.07	0.06	0.05	0.07	0.03	0.08	0.05	0.07	0.01	0.08	0.05	0.06	0.10	0.10	0.10	0.07	0.08	0.06
LN*	**0.82**	**0.83**	**0.85**	**0.86**	**0.87**	**0.87**	**0.88**	**0.89**	**0.90**	**0.92**	**0.92**	**0.93**	**0.94**	**0.94**	**0.95**	**0.96**	**0.97**	**0.98**	**0.99**	**0.99**	**1.02**	**1.02**
AA	0.44	0.43	0.50	0.41	0.40	0.41	0.43	0.43	0.48	0.35	0.42	0.44	0.41	0.41	0.41	0.45	0.34	0.39	0.36	0.32	0.33	0.47
GA	0.28	0.29	0.32	0.32	0.32	0.30	0.30	0.30	0.35	0.35	0.30	0.31	0.29	0.31	0.30	0.30	0.32	0.34	0.31	0.31	0.34	0.30
BA	0.12	0.11	0.14	0.12	0.12	0.12	0.12	0.12	0.15	0.10	0.12	0.12	0.11	0.12	0.11	0.13	0.09	0.11	0.10	0.12	0.09	0.13
LG	0.06	0.07	0.08	0.06	0.06	0.05	0.08	0.07	0.07	0.06	0.07	0.08	0.06	0.07	0.07	0.08	0.05	0.06	0.06	0.05	0.05	0.07
SQ	0.75	1.05	0.82	0.62	0.37	0.81	1.72	0.68	0.69	0.45	1.04	1.15	0.47	0.86	0.89	1.03	0.55	0.39	0.55	0.35	0.44	0.84
TFA	0.06	0.06	0.06	0.08	0.07	0.06	0.05	0.07	0.03	0.08	0.05	0.07	0.01	0.08	0.05	0.06	0.10	0.10	0.10	0.07	0.08	0.06
SFA*	**17.85**	**17.87**	**16.75**	**16.60**	**16.61**	**17.70**	**18.62**	**17.88**	**16.73**	**15.94**	**18.37**	**18.49**	**18.85**	**17.46**	**17.69**	**17.66**	**16.29**	**15.99**	**17.20**	**16.94**	**15.39**	**16.77**
MUFA	70.83	70.72	76.61	71.39	70.97	71.62	71.83	72.53	69.36	70.99	70.06	71.11	66.61	71.31	2.06	71.75	1.97	70.79	69.16	70.05	69.66	71.03
PUFA*	**11.26**	**11.33**	**6.59**	**11.88**	**12.22**	**10.57**	**9.50**	**9.51**	**13.87**	**12.97**	**11.52**	**10.30**	**14.44**	**11.11**	**11.31**	**10.51**	**13.77**	**13.12**	**13.54**	**13.28**	**14.82**	**10.11**

(cont. on next page)

Table 3.2. (cont.)

FATTY ACIDS	KY 6	DK	MCD6	DR 2	AH	MCD1	KRK	KD	DNG	BY
MA	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
PA*	**15.46**	**14.51**	**14.82**	**14.86**	**15.92**	**15.05**	**16.11**	**14.98**	**15.92**	**15.37**
POA*	**1.57**	**1.76**	**1.58**	**1.75**	**1.98**	**1.64**	**1.78**	**1.74**	**1.54**	**1.74**
MG	0.13	0.13	0.13	0.13	0.13	0.13	0.14	0.13	0.12	0.13
MGO	0.26	0.25	0.26	0.25	0.25	0.26	0.29	0.25	0.21	0.27
SA*	**2.28**	**2.27**	**2.39**	**2.33**	**2.34**	**2.32**	**2.29**	**2.28**	**2.42**	**2.12**
EA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OA*	**65.74**	**66.63**	**65.95**	**65.52**	**65.34**	**66.68**	**67.15**	**65.77**	**67.19**	**67.78**
VA	3.12	2.19	2.40	2.52	2.69	2.38	2.46	2.71	2.42	2.10
LO*	**9.32**	**10.22**	**10.45**	**10.57**	**9.27**	**9.25**	**7.57**	**9.92**	**7.83**	**8.20**
TLO +TLN	0.08	0.07	0.07	0.05	0.04	0.08	0.06	0.06	0.06	0.05
LN*	**1.03**	**1.03**	**1.03**	**1.05**	**1.08**	**1.12**	**1.13**	**1.14**	**1.15**	**1.20**
AA	0.43	0.42	0.40	0.44	0.44	0.45	0.46	0.06	0.51	0.44
GA	0.31	0.31	0.30	0.31	0.29	0.32	0.32	0.32	0.34	0.33
BA	0.12	0.12	0.12	0.12	0.12	0.13	0.13	0.13	0.15	0.13
LG	0.08	0.07	0.07	0.07	0.07	0.08	0.09	0.08	0.11	0.09
SQ	0.95	1.06	0.86	0.99	0.89	0.86	1.06	1.15	1.08	1.19
TFA	0.08	0.07	0.07	0.05	0.04	0.08	0.06	0.06	0.06	0.05
SFA*	**18.53**	**17.55**	**17.95**	**17.97**	**19.04**	**18.18**	**19.24**	**17.68**	**19.25**	**18.30**
MUFA	71.00	71.14	70.49	70.35	70.55	71.28	72.00	70.79	71.70	72.22
PUFA*	**10.35**	**11.25**	**11.48**	**11.62**	**10.35**	**10.37**	**8.70**	**11.06**	**8.98**	**9.40**

Table 3.3. : TAG components of olive oil samples

FATTY ACIDS	ZAE-B	ZAE-KP	MLG	KY1	KY2	KY3	KY4	KY5	KY 6	AKH1	AKH2	YTG	İLY	DNG	KRK	DK	ZO 2	KH	BY	BLC1	BLC2	BLC3
LLL	0.55	0.35	0.17	0.30	0.31	0.53	0.13	0.49	0.24	0.29	0.60	0.67	0.31	0.18	0.27	0.25	0.58	0.38	0.37	0.44	0.59	0.29
LOLn+POLL	0.52	0.46	0.32	0.65	0.37	0.44	0.26	0.38	0.54	0.64	0.15	0.37	0.09	0.47	0.55	0.44	0.62	0.84	0.71	0.83	0.62	0.39
PLLn	0.24	0.14	0.08	0.13	0.10	0.14	0.04	0.31	0.26	0.13	0.08	0.23	0.18	0.18	0.24	0.30	0.17	0.22	0.32	0.11	0.15	0.08
OLL	4.31	3.32	2.13	3.75	3.42	2.66	1.92	2.98	2.29	3.50	0.87	1.40	0.97	1.93	1.82	2.66	4.11	3.85	1.96	3.74	4.57	2.36
OLnO	1.50	1.42	1.51	1.40	1.51	1.77	1.55	1.64	2.11	1.43	1.87	1.61	1.15	1.98	2.08	1.96	1.53	1.75	2.20	1.60	1.71	1.44
PLL	1.64	1.03	0.67	1.31	1.26	0.89	0.54	1.30	0.87	1.11	0.74	0.44	0.25	0.64	0.59	0.87	1.69	0.99	0.72	1.15	1.20	0.57
POLn	1.02	0.79	0.74	0.75	0.85	0.82	0.59	1.11	1.05	0.66	0.14	0.88	0.49	1.15	1.10	0.97	0.72	0.82	1.23	0.89	0.84	0.62
LOO+PLnP	13.50	13.13	12.50	13.58	14.76	11.88	11.22	12.92	10.60	15.00	8.69	9.39	8.04	8.67	10.05	13.51	13.87	15.89	10.36	14.31	16.83	13.54
PoOO	2.32	2.75	1.42	3.65	1.21	2.47	2.06	2.18	2.78	2.26	1.86	1.68	1.55	2.95	1.83	1.75	1.78	0.73	1.88	2.44	2.36	1.13
PLO + SLL	9.51	7.97	6.82	8.54	8.72	7.21	6.25	7.27	7.34	8.84	4.26	4.63	3.81	6.70	6.20	7.45	9.73	8.46	6.58	8.33	8.63	6.58
PoOP	0.79	0.71	0.77	1.10	0.57	0.89	1.13	0.68	0.70	0.71	0.91	0.66	0.59	0.92	1.02	0.94	0.83	0.45	1.01	0.80	0.81	0.45
PLP	1.58	1.21	1.06	1.46	1.11	1.24	0.89	1.04	1.01	1.04	0.95	0.91	0.84	1.22	1.15	1.09	1.23	1.31	1.16	1.33	1.20	1.09
OOO	28.13	33.18	35.80	31.36	31.61	33.58	36.38	33.25	33.51	31.32	39.81	41.79	43.70	33.72	34.77	34.24	27.61	30.13	34.02	30.19	31.66	37.54
SLO + POO	23.30	23.34	25.17	22.54	24.43	24.64	26.12	25.00	26.14	22.07	25.61	24.68	25.61	27.23	27.24	24.77	23.62	23.37	26.70	23.03	21.55	23.84
POP	4.81	3.97	3.74	4.31	4.00	4.21	4.07	3.78	4.60	4.46	3.88	3.11	3.58	5.43	5.12	3.68	4.36	4.82	5.10	4.36	3.15	3.65
PPP	0.71	0.67	0.43	0.51	0.46	0.55	0.40	0.60	0.55	0.42	1.00	0.77	0.92	0.78	0.62	0.26	0.79	0.99	0.64	0.72	0.37	0.67
SOO	4.01	4.32	5.18	3.26	3.76	3.90	3.76	0.88	4.20	3.84	5.77	4.90	5.47	4.38	4.18	3.64	4.48	3.84	4.04	3.52	2.75	4.14
POS	1.47	1.26	1.49	1.05	1.09	1.10	1.00	0.35	1.22	1.48	1.63	1.02	1.45	1.51	1.16	1.11	1.66	1.18	0.99	1.37	0.71	1.03
ECN42	1.31	0.95	0.57	1.08	0.78	1.11	0.43	1.18	1.04	1.06	0.83	1.27	0.58	0.83	1.06	0.99	1.37	1.44	1.40	1.38	1.36	0.76
ECN44	8.47	6.56	5.05	7.21	7.04	6.14	4.60	7.03	6.32	6.70	3.62	4.33	2.86	5.70	5.59	6.46	8.05	7.41	6.11	7.38	8.32	4.99
ECN46	27.70	25.77	22.57	28.33	26.37	23.69	21.55	24.09	22.43	27.85	16.67	17.27	14.83	20.46	20.25	24.74	27.44	26.84	20.99	27.21	29.83	22.79
ECN 48	36.95	61.16	65.14	58.72	60.50	62.98	66.97	62.63	64.80	58.27	70.30	70.35	73.81	67.16	67.75	62.95	56.38	59.31	66.46	58.30	56.73	65.70
ECN 50	5.48	5.58	6.67	4.31	4.85	5.00	4.76	1.23	5.42	5.32	7.40	5.92	6.92	5.89	5.34	4.75	6.14	5.02	5.03	4.89	3.46	5.17
PLO/OOO	0.34	0.24	0.19	0.27	0.28	0.21	0.17	0.22	0.22	0.28	0.11	0.11	0.09	0.20	0.18	0.22	0.35	0.28	0.19	0.28	0.27	0.18
LLL/ECN42	0.42	0.37	0.30	0.28	0.40	0.48	0.30	0.42	0.23	0.27	0.72	0.53	0.53	0.22	0.25	0.25	0.42	0.26	0.26	0.32	0.43	0.38
PLL/OLL	0.38	0.31	0.31	0.35	0.37	0.33	0.28	0.44	0.38	0.32	0.85	0.31	0.26	0.33	0.32	0.33	0.41	0.26	0.37	0.31	0.26	0.24
ECN48/ECN46	2.06	2.37	2.89	2.07	2.29	2.66	3.11	2.60	2.89	2.09	4.22	4.07	4.98	3.28	3.35	2.54	2.05	2.21	3.17	2.14	1.90	2.88
LOO/PLO	8.23	12.75	18.66	10.37	11.71	13.35	20.78	9.94	12.18	13.51	11.74	21.34	32.16	13.55	17.03	15.53	8.21	16.05	14.39	12.44	14.03	13.05
OOO/POO	1.21	1.42	1.42	1.39	1.29	1.36	1.39	1.33	1.28	1.42	1.55	1.69	1.71	1.24	1.28	1.38	1.17	1.29	1.27	1.31	1.47	1.57

(cont. on next page)

Table 3.3. (cont.)

FATTY ACIDS	MCD1	MCD2	MCD3	MCD4	MCD5	MCD6	MCD7	MCD8	MCD9	MCD10	SR1	SR 2	SR 3	KRK	BLN1	BLN2	BLN3	BLN4	BLN5	BLN5	KD	BR 1
LLL	0.30	0.12	0.16	0.26	0.46	0.33	0.28	0.57	0.51	0.18	0.33	0.28	0.22	0.13	0.13	0.41	0.10	0.57	0.27	0.16	0.28	0.48
LOLn +POLL	0.47	0.21	0.55	0.44	0.55	0.54	0.36	0.59	0.53	0.34	0.27	0.63	0.34	0.45	0.32	0.49	0.38	0.38	0.68	0.46	0.85	0.34
PLLn	0.10	0.05	0.09	0.11	0.15	0.22	0.10	0.18	0.16	0.08	0.09	0.60	0.08	0.17	0.15	0.14	0.04	0.10	0.16	0.07	0.31	0.15
OLL	**2.10**	**2.30**	**1.93**	**2.96**	**4.28**	**3.08**	**2.95**	**3.82**	**3.69**	**2.43**	**2.38**	**1.84**	**2.59**	**1.47**	**1.63**	**2.11**	**1.60**	**3.44**	**2.03**	**2.38**	**2.35**	**3.66**
OLnO	2.16	1.36	1.64	1.78	1.51	1.85	1.43	1.73	1.70	1.42	1.20	1.52	1.61	1.21	1.45	1.99	1.46	1.36	1.68	1.45	2.02	1.60
PLL	0.79	0.63	0.67	0.99	1.11	1.00	0.97	1.01	1.00	0.56	0.84	0.80	0.89	0.41	0.57	0.82	0.21	1.32	0.99	0.73	1.17	1.14
POLn	1.03	0.66	0.77	0.88	0.78	1.03	0.76	0.83	0.74	0.69	0.56	0.73	0.94	0.46	0.75	1.04	0.74	0.80	1.25	0.73	1.06	0.73
LOO +PLnP	10.84	12.31	11.30	13.23	16.78	12.92	14.44	14.03	16.40	10.74	13.62	11.55	12.51	11.35	11.23	11.44	12.15	15.24	10.83	13.84	11.83	16.39
PoOO	**2.02**	**1.22**	**1.54**	**1.61**	**1.41**	**1.62**	**1.41**	**3.54**	**1.38**	**2.66**	**1.18**	**1.57**	**1.63**	**1.11**	**1.66**	**1.66**	**1.43**	**0.93**	**1.92**	**1.11**	**2.07**	**1.39**
PLO + SLL	6.92	6.06	6.45	7.46	8.24	7.56	7.73	7.98	7.94	6.19	7.56	6.42	7.41	5.69	6.24	7.38	4.47	8.91	6.86	7.09	7.38	8.01
PoOP	1.26	0.51	0.84	0.82	0.63	0.85	0.79	0.79	0.78	0.89	0.73	0.81	0.76	0.45	0.75	0.84	0.50	0.76	1.28	0.65	1.02	0.77
PLP	1.45	0.96	1.24	1.10	1.00	1.19	1.14	1.11	1.06	1.14	1.17	1.34	1.18	0.90	1.24	1.34	1.29	1.22	1.96	1.05	1.29	1.21
OOO	**32.74**	**38.70**	**35.86**	**33.59**	**32.16**	**32.99**	**34.32**	**32.45**	**33.50**	**38.24**	**34.97**	**35.35**	**34.00**	**39.97**	**36.17**	**32.86**	**45.77**	**31.92**	**31.14**	**36.43**	**32.37**	**34.25**
SLO + POO	25.69	23.85	25.56	24.96	22.26	24.85	23.77	22.63	22.65	23.69	25.67	25.14	25.35	25.00	26.14	25.87	19.91	24.25	24.39	24.36	25.04	21.81
POP	**4.86**	**3.15**	**4.26**	**4.02**	**3.41**	**3.99**	**3.29**	**3.72**	**3.26**	**4.10**	**3.90**	**4.67**	**4.09**	**3.73**	**4.37**	**4.71**	**2.83**	**3.89**	**5.97**	**3.66**	**4.72**	**2.78**
PPP	1.00	0.42	0.58	0.38	0.39	0.40	0.38	0.45	0.30	0.83	0.25	0.99	0.65	0.64	1.25	0.57	1.42	0.72	1.43	0.70	0.69	0.47
SOO	**3.89**	**4.75**	**4.62**	**3.78**	**3.47**	**3.76**	**4.09**	**3.16**	**2.80**	**3.91**	**3.88**	**4.26**	**4.08**	**4.88**	**4.84**	**4.34**	**4.91**	**4.19**	**3.76**	**4.02**	**3.67**	**3.59**
POS	1.48	1.34	1.42	1.19	1.01	1.00	1.09	0.89	1.46	1.07	0.94	1.43	0.91	1.16	2.32	1.41	0.82	0.00	2.57	1.11	1.15	0.70
ECN42	0.87	0.38	0.80	0.81	1.16	1.09	0.74	1.34	1.20	0.60	0.69	1.51	0.64	0.75	0.60	1.04	0.52	1.05	1.11	0.69	1.44	0.33
ECN44	6.08	4.95	5.01	6.61	7.68	6.96	6.11	7.39	7.13	5.10	4.98	4.89	6.03	3.55	4.40	5.96	4.01	6.92	5.95	5.29	6.60	7.13
ECN46	22.49	21.06	21.37	24.22	28.06	24.14	25.51	27.45	27.56	21.62	24.26	21.69	23.49	19.50	21.12	22.66	19.84	27.06	22.85	23.74	23.59	27.77
ECN 48	64.29	66.12	66.26	62.95	58.22	62.23	61.76	59.25	59.71	66.86	64.79	66.15	64.09	69.34	67.93	64.01	69.93	60.78	62.93	65.15	62.82	59.31
ECN 50	5.37	6.09	6.04	4.97	4.48	4.76	5.18	4.05	4.26	4.98	4.82	5.69	4.99	6.04	7.16	5.75	5.73	4.19	6.33	5.13	4.82	4.29
PLO/OOO	0.21	0.16	0.18	0.22	0.26	0.23	0.23	0.25	0.24	0.16	0.22	0.18	0.22	0.14	0.17	0.22	0.10	0.28	0.22	0.19	0.23	0.23
LLL/ECN42	0.34	0.32	0.20	0.32	0.40	0.30	0.38	0.43	0.43	0.30	0.48	0.19	0.34	0.17	0.22	0.39	0.19	0.54	0.24	0.23	0.19	1.45
PLL/OLL	0.38	0.27	0.35	0.33	0.26	0.32	0.33	0.26	0.27	0.23	0.35	0.43	0.34	0.28	0.35	0.39	0.13	0.38	0.49	0.31	0.50	0.31
ECN48/ECN46	2.86	3.14	3.10	2.60	2.07	2.58	2.42	2.16	2.17	3.09	2.67	3.05	2.73	3.56	3.22	2.82	3.52	2.25	2.75	2.74	2.66	2.14
LOO/PLO	13.72	19.54	16.87	13.36	15.12	12.92	14.89	13.89	16.40	19.18	16.21	14.44	14.06	27.68	19.70	13.95	57.86	11.55	10.94	18.96	10.11	14.38
OOO/POO	1.27	1.62	1.40	1.35	1.44	1.33	1.44	1.43	1.48	1.61	1.36	1.41	1.34	1.60	1.38	1.27	2.30	1.32	1.28	1.50	1.29	1.57

(cont. on next page)

Table 3.3. (cont.)

FATTY ACIDS	BR 2	KK	GMI	GM 2	KYH	DMB1	DMB2	PZ	BLK	DR 1	DR 2	TND	KST	GRC	MND	MNV	AKY	AH	EDN	ED	MDN	TRI	TR 2	TR3
LLL	0.13	0.22	0.35	0.30	0.40	0.22	0.27	0.26	0.31	0.36	0.49	0.33	0.17	0.35	0.36	0.20	0.13	0.48	0.13	0.10	0.36	0.15	0.13	0.26
LOLn +POLL	0.41	0.25	0.56	0.31	0.53	0.68	0.43	0.49	0.44	0.42	0.67	0.55	0.21	0.37	0.49	0.47	0.44	0.52	0.24	0.26	0.28	0.32	0.46	0.47
PLLn	0.08	0.06	0.15	0.09	0.29	0.19	0.16	0.17	0.17	0.08	0.41	0.24	0.05	0.09	0.13	0.08	0.08	0.19	0.06	0.08	0.07	0.09	0.06	0.28
OLL	**1.51**	**2.03**	**1.32**	**2.01**	**2.61**	**2.67**	**2.22**	**2.32**	**3.12**	**2.75**	**2.71**	**3.10**	**1.01**	**2.59**	**2.57**	**2.70**	**2.18**	**2.09**	**1.79**	**1.35**	**2.35**	**1.80**	**2.02**	**2.27**
OLnO	1.39	1.38	1.66	1.41	1.74	1.63	1.75	1.54	1.78	1.55	2.15	2.00	1.27	1.49	1.48	1.44	1.30	2.04	1.30	1.22	1.20	1.29	1.60	1.57
PLL	0.39	0.56	0.46	0.58	0.89	0.74	0.83	0.75	1.15	1.16	1.20	1.17	0.32	1.00	0.98	1.01	0.80	0.85	0.50	0.36	0.57	0.51	0.71	0.64
POLn	0.55	0.57	0.91	0.60	0.80	0.76	0.87	0.72	0.92	0.88	1.17	0.95	0.56	0.63	0.65	0.70	0.54	0.97	0.54	0.52	0.56	0.62	0.72	0.84
LOO +PLnP	11.31	12.90	9.81	11.91	11.79	13.86	10.59	11.97	13.03	11.63	12.65	12.67	9.11	11.76	12.63	13.51	12.02	11.13	10.72	10.22	11.06	11.88	12.27	12.39
PoOO	**1.37**	**1.86**	**1.84**	**1.71**	**1.49**	**1.28**	**1.64**	**1.56**	**1.41**	**2.06**	**1.89**	**1.69**	**1.85**	**2.46**	**1.94**	**1.48**	**1.51**	**1.82**	**1.43**	**1.54**	**1.48**	**1.95**	**2.14**	**1.82**
PLO + SLL	5.29	6.38	5.20	6.15	7.14	6.49	7.11	7.07	7.93	7.45	8.03	8.03	4.27	7.83	7.39	8.04	7.07	7.68	5.52	4.97	5.19	6.02	6.82	6.86
PoOP	0.63	0.79	0.73	0.92	0.81	0.75	0.89	0.80	0.97	0.90	1.13	0.82	0.66	1.04	1.10	0.75	0.87	0.75	0.66	0.71	0.67	0.91	1.10	0.74
PLP	0.97	0.96	0.84	0.96	1.23	1.16	1.40	1.10	1.43	1.49	1.52	1.16	0.80	1.44	1.21	1.13	1.06	1.45	0.93	0.77	0.82	0.82	1.13	1.05
OOO	**42.12**	**37.71**	**41.19**	**37.56**	**33.86**	**36.76**	**33.85**	**34.07**	**32.00**	**32.32**	**31.91**	**32.20**	**42.75**	**32.12**	**32.97**	**33.12**	**33.77**	**31.99**	**39.40**	**40.50**	**40.90**	**37.04**	**35.78**	**36.20**
SLO + POO	24.45	24.33	25.50	25.40	25.33	22.15	26.33	26.04	24.52	24.61	24.33	24.80	25.05	24.79	24.80	25.68	25.67	25.78	24.31	25.11	23.34	24.77	25.31	25.13
POP	**3.25**	**3.26**	**3.26**	**4.17**	**4.72**	**3.72**	**5.52**	**4.89**	**4.87**	**4.63**	**4.38**	**4.56**	**3.29**	**4.21**	**4.06**	**4.18**	**4.49**	**5.02**	**3.42**	**3.64**	**2.68**	**3.92**	**3.86**	**3.05**
PPP	0.72	0.29	0.69	0.13	0.62	0.89	0.80	0.56	0.78	0.83	0.26	0.66	0.60	0.56	0.59	0.17	1.17	0.67	0.46	0.73	0.39	0.55	0.50	0.80
SOO	**4.39**	**4.50**	**4.55**	**4.40**	**4.35**	**4.27**	**3.45**	**4.47**	**3.78**	**4.63**	**3.15**	**3.96**	**5.51**	**4.96**	**4.35**	**4.09**	**4.84**	**4.11**	**5.88**	**5.62**	**5.62**	**5.55**	**4.21**	**4.70**
POS	0.74	1.15	0.99	1.10	1.41	1.12	1.64	1.24	1.40	1.67	1.17	1.12	1.30	1.41	1.56	1.25	1.44	1.41	1.37	0.62	1.36	0.59	1.20	1.05
ECN42	0.32	0.46	1.06	0.70	1.22	1.09	0.86	0.92	0.92	0.86	1.57	1.12	0.43	0.81	0.98	0.75	0.65	1.19	0.43	0.44	0.71	0.56	0.65	1.01
ECN44	3.84	4.54	4.35	4.60	6.04	5.80	5.67	5.33	6.97	6.34	7.23	7.22	3.16	5.71	5.68	5.85	4.82	5.95	4.13	3.45	4.68	4.22	5.05	5.32
ECN46	19.57	22.89	18.42	21.65	22.46	23.54	21.63	22.50	24.77	23.53	25.22	24.37	16.69	24.53	24.27	24.91	22.53	22.83	19.26	18.21	19.22	21.58	23.46	22.86
ECN 48	70.54	65.59	70.64	67.26	64.53	63.52	66.50	65.56	62.17	62.39	60.88	62.22	71.69	61.68	62.42	63.15	65.10	63.46	67.59	69.98	67.31	66.28	65.45	65.18
ECN 50	5.13	5.65	5.54	5.50	5.76	5.39	5.09	5.71	5.18	6.30	4.32	5.08	6.81	6.37	5.91	5.34	6.28	5.52	7.25	6.24	6.98	6.14	5.41	5.75
PLO/OOO	0.13	0.17	0.13	0.16	0.21	0.18	0.21	0.21	0.25	0.23	0.25	0.25	0.10	0.24	0.22	0.24	0.21	0.24	0.14	0.12	0.13	0.16	0.19	0.19
LLL/ECN42	0.41	0.48	0.33	0.43	0.33	0.20	0.31	0.28	0.34	0.42	0.31	0.29	0.40	0.43	0.37	0.27	0.20	0.40	0.30	0.23	0.51	0.27	0.20	0.26
PLL/OLL	0.26	0.28	0.35	0.29	0.34	0.28	0.37	0.32	0.37	0.42	0.44	0.38	0.32	0.39	0.38	0.37	0.37	0.41	0.28	0.27	0.24	0.28	0.35	0.28
ECN48/ECN46	3.60	2.87	3.83	3.11	2.87	2.70	3.07	2.91	2.51	2.65	2.41	2.55	4.30	2.51	2.57	2.54	2.89	2.78	3.51	3.84	3.50	3.07	2.79	2.85
LOO/PLO	29.00	23.04	21.33	20.53	13.25	18.73	12.76	15.96	11.33	10.03	10.54	10.83	28.47	11.76	12.89	13.38	15.03	13.09	21.44	28.39	19.40	23.29	17.28	19.36
OOO/POO	1.72	1.55	1.62	1.48	1.34	1.66	1.29	1.31	1.31	1.31	1.31	1.30	1.71	1.30	1.33	1.29	1.32	1.24	1.62	1.61	1.75	1.50	1.41	1.44

(cont. on next page)

Table 3.3. (cont.)

FATTY ACIDS	TR 4	TR 5	TR 6	İZN	TCR	DRA	ÇKR	BYL	GM	MF
LLL	0.19	0.40	0.12	0.12	0.08	0.09	0.17	0.24	0.32	0.11
LOLn +POLL	0.35	0.15	0.39	0.24	0.30	0.36	0.20	0.23	0.27	0.21
PLLn	0.08	0.27	0.05	0.07	0.06	0.12	0.04	0.06	0.06	0.04
OLL	2.63	1.69	1.65	1.44	1.21	1.32	1.88	2.13	2.43	1.39
OLnO	1.49	1.61	1.32	1.30	1.42	1.14	1.37	1.23	1.23	1.13
PLL	0.79	0.55	0.46	0.34	0.45	0.33	0.50	0.52	0.69	0.33
POLn	0.60	0.84	0.50	0.55	0.45	0.46	0.60	0.47	0.60	0.45
LOO +PLnP	13.84	11.50	11.60	10.78	10.17	10.77	11.34	12.00	11.77	10.38
PoOO	1.66	1.90	1.43	1.15	1.32	1.34	0.97	1.31	1.48	1.14
PLO + SLL	7.26	5.31	5.63	4.57	4.51	4.71	4.76	5.51	5.40	4.72
PoOP	0.80	0.84	0.46	0.61	0.47	0.55	0.41	0.57	0.60	0.57
PLP	0.87	0.70	0.84	0.72	0.84	0.78	0.73	0.86	0.71	0.68
OOO	35.91	41.78	41.41	42.73	41.94	43.00	43.18	41.47	39.90	42.40
SLO + POO	24.48	24.01	23.97	24.74	23.90	24.32	22.75	23.44	23.97	24.32
POP	3.33	2.81	2.92	3.07	3.41	3.04	2.74	2.57	2.82	3.05
PPP	0.38	0.40	0.54	0.82	1.19	0.72	0.73	0.49	0.52	0.56
SOO	4.31	4.04	5.35	5.86	6.00	5.71	6.39	5.16	5.40	5.98
POS	1.04	1.20	1.34	1.30	1.46	1.22	1.25	1.10	1.28	1.42
ECN42	0.62	0.82	0.56	0.43	0.44	0.57	0.41	0.53	0.65	0.36
ECN44	5.51	4.69	3.93	3.63	3.53	3.25	4.35	4.35	4.95	3.30
ECN46	24.43	20.25	19.96	17.83	17.31	18.15	18.21	20.25	19.96	17.49
ECN 48	64.10	69.00	68.84	71.36	70.44	71.08	69.40	67.97	67.21	70.33
ECN 50	5.35	5.24	6.69	7.16	7.46	6.93	7.64	6.26	6.68	7.40
PLO/OOO	0.20	0.13	0.14	0.11	0.11	0.11	0.11	0.13	0.14	0.11
LLL/ECN42	0.31	0.49	0.21	0.28	0.18	0.16	0.41	0.45	0.49	0.31
PLL/OLL	0.30	0.33	0.28	0.24	0.37	0.25	0.27	0.24	0.28	0.24
ECN48/ECN46	2.62	3.41	3.45	4.00	4.07	3.92	3.81	3.36	3.37	4.02
LOO/PLO	17.52	20.91	25.22	31.71	22.60	32.64	22.68	23.08	17.06	0.02
OOO/POO	1.47	1.74	1.73	1.73	1.75	1.77	1.90	1.77	1.66	1.74

3.2.2 Methods

Examples of natural oil varieties of Gemlik, were filtered by dehydrated sodium sulfate of which placed over Whatman - No. 42 filter paper. Oil samples were used on tri-achyl glycerol (TAG) analyses. Besides, with fatty acid esterification process, fatty acids were converted into methyl esters and FAME analyses were applied. To be esterification by oil samples, it was used cold methylation method (IUPAC Method 2.301) which certificated from International Olive Oil Council (Anonymous, 1979). Implementation of the method is as follows:

- approximate 200 mg (few drops) oil samples were weighed in vial
- It was added over 10 ml hexane that be chromatographic purity, then shaked by hand.
- And then, it was added 500 μl from 2M methanolic KOH solution, and since clear solution, it was shaked with hand for twice.
- It was injected from upper side clear phase to chromatography after dividing phase of glycerol.

3.2.2.1 Analyses of Fatty Acids Components

Fatty acids that belong to fatty samples converted to the methyl esters were analyzed by gas chromatography GC HP 6890 model.

Flame ionization detector (FID) and the capillary column (DB -23, bonded 50% cyanopropyl, 30 m X 0.25 mm ID x 0.250 micro M, J & W Scientific, Folsom, CA, USA) were used. The device operating parameters are shown below Table 3.2.

Table 3.4 : Parameters of GC instrument

The detector temperature	250ºC
Injector temperature	250ºC
Injection	Split model 1/100
Carrier gases : Helium flow rate Hydrogen flow rate Air flow rate	0.5 ml/min 30ml/min 300ml/min
Make up gases: Nitrogen flow rate	24 ml/min
Oven temperature program	Programmed between 170 – 210 ^{0}C On analyses, between 170 ^{0}C' - 210 ^{0}C Oven Program is to be implementedincrementally20^{o}C/min and analysis will be completed with samples will be waited 10 minutes at 210^{0}C

Fatty acids used in the analysis were myristic acid (C14:0), palmitic acid (C16:0), palmitoleic acid (C16:1), margaric acid (C17:0), margoleic acid (C17:1), stearic acid (C18:0), elaidic acid (C18:1 *trans*), oleic acid (C18:1), linoelaidic acid (C18:2 *trans*), linoleic acid (C18:2), *trans* linolenic acid (C18:3 *trans*), linolenic acid (C18:3), arachidic acid (C20:0), gadoleic acid (C20:1), behenic acid (C22:0), lignoseric acid (C24:0). Each sample was analyzed at least two times. 16 main fatty acids in olive oil samples were determined by retention time of each one according to the reference of standard fatty acids. The area of the each peak which belonged to these fatty acids was integrated by using Chem-station software. The integrated area of each fatty acid was converted to the % concentration by dividing the calculated area of each acid to total area content of all related fatty acids existed in olive oil. These fatty acids are shown in Figure3.5.

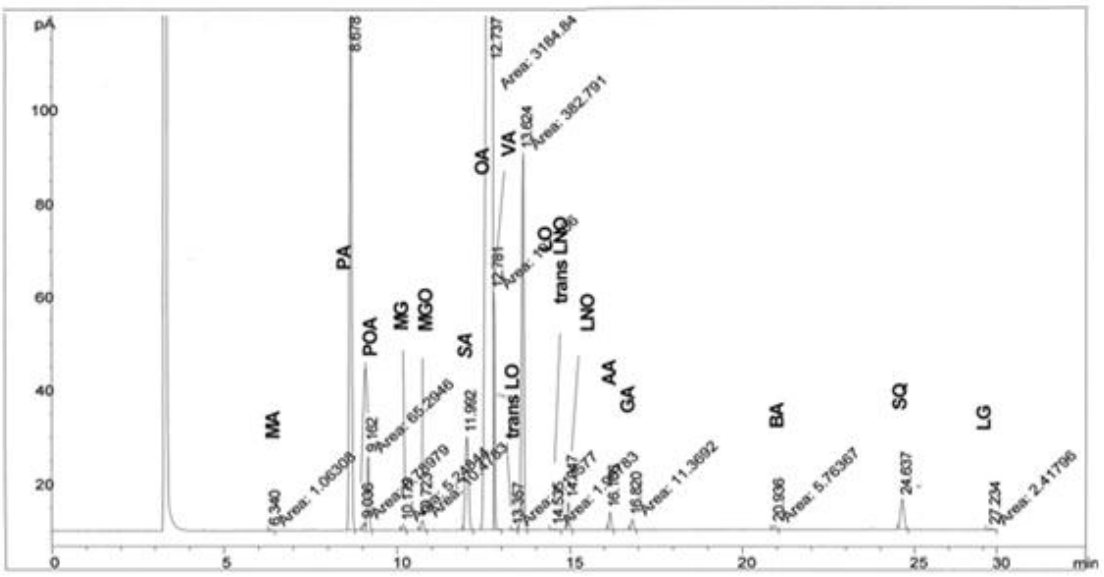


Figure 3.6. Fatty acids in Gas chromatogram

In this study, as a result of analyzing the fatty acids; palmitic acid, palmitoleic acid, stearic acid, oleic acid, linoleic acid and linolenic acid were preferred due to amount of mass percentages were more. The chromatographic analysis results for fatty acid methyl esters were used as the reference method. In addition, spectroscopic measurements of fatty acid methyl esters are made simultaneously chromatographic measurements, spectra are recorded.

3.2.2.2 Analysis of Triacylglycerol Components

Analyses of triacyl glycerol components of natural olive oil samples (triglyceride (TAG)) were determined by HPLC Agilent 1200 device, according to international standards(IUPAC 2324) that is recommended by EC 2568-91 directive of the European Union and based on the principle HPLC that is adopted by IOOC. Refractive index detector as a detector (RID), as column Superspher 100 RP-18 column (244 x 4 mm ID x 4 m) were used.TAG was analyzed at 350^{o}C temperature, up to 200 bar maximum pressure and 1.2 mL / min mobile phase flow rate. Mobile phase is Acetone + 63.6% 36.4% acetonitrile and injection volume is 500 μl. Tri acyl glycerol components are shown in Figure 3.6

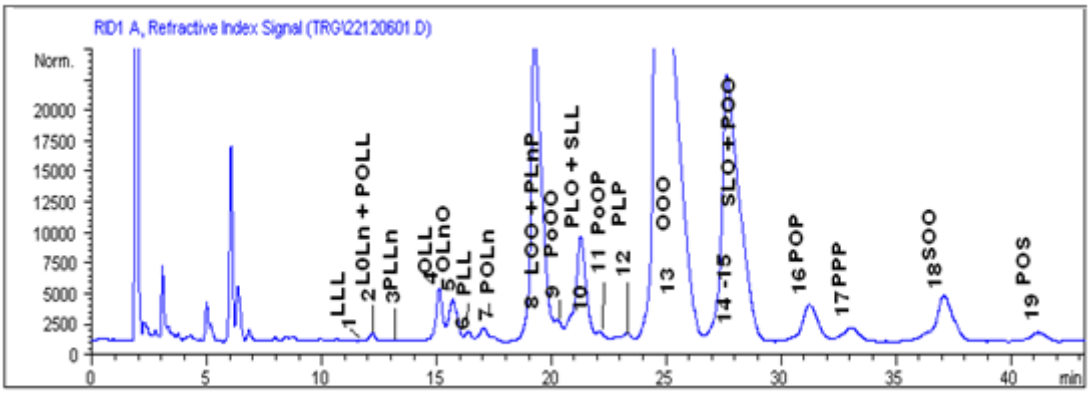


Figure 3.7. Tri acyl glycerol components in HPLC chromatogram

In this study PoOO, OLL, OOO, POP and SOO which were the main components of olive oil triacyl glycerol, of the chromatographic and spectroscopic measurements were evaluated.

3.3 Spectroscopic Analysis

In parallel with the chromatographic analysis, simultaneously two different molecular spectroscopic methods have been used for spectroscopic analysis. These spectroscopic methods are Fourier transform infrared spectroscopy and near infrared spectroscopy.

Fourier transform infrared spectroscopic analyses were performed by Perkin Elmer Spectrum 100 model equipped with diamond-ZnSe ATR accesory. The working range was set 600-4000 cm^{-1}wavenumber with 4 cm^{-1} resolution by averaging 64 scan numbers. Absorbance spectrums were collected at room temperature. Background spectrum was obtained empty and dry ATR cell. Before and after each sample analyses background was collected to reduce the contaminations that were come from the ATR crystal. ATR crystal was cleaned with pure ethanol and allowed to dry.

Near-infrared spectroscopic analyses were performed at room temperature with FTS-3000 NIR spectrometer (Bio-Rad, Excalibur, Cambridge, MA) and the working range was set 4500 -10000 cm^{-1}wavenumber with 8 cm^{-1} resolution by averaging 64 scan numbers. Sample spectra were collected in absorbance method. Before and after each sample analyses background was collected to reduce the contaminations that might

come from sample cuvette. Quartz cuvette was cleaned with pure acetone and allowed to dry.

3.4 Data Analysis

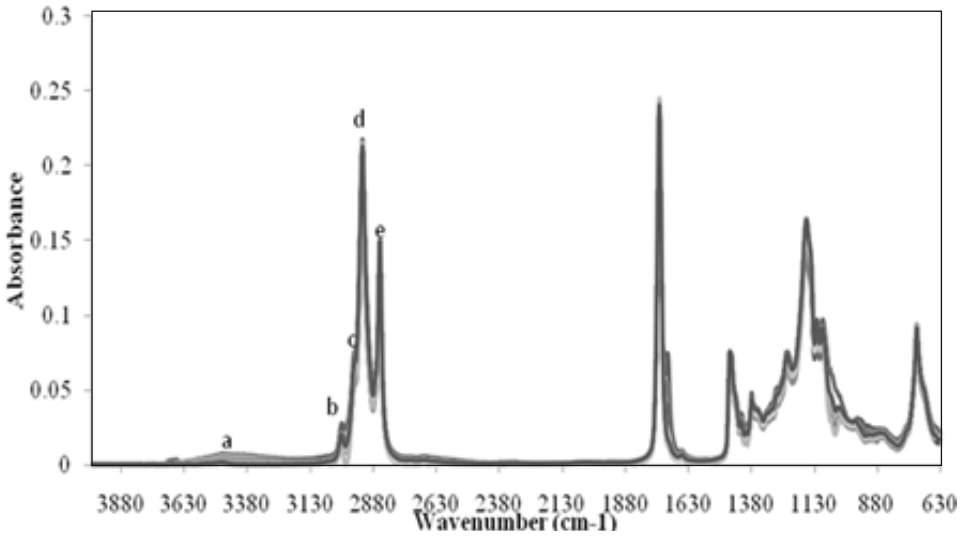
The collected spectra were transferred in ASCII file format and were combined with Microsoft Excel program. Then, for multivariate analyses data files (calibration and validation) were prepared as text files. Genetic algorithm based calibration method was written in MATLAB programming language Version 7.0 (MathWorks Inc., Natick, MA).Partial Least Square was also applied in Minitab programme and data files were collected ASCII file format.

CHAPTER 4

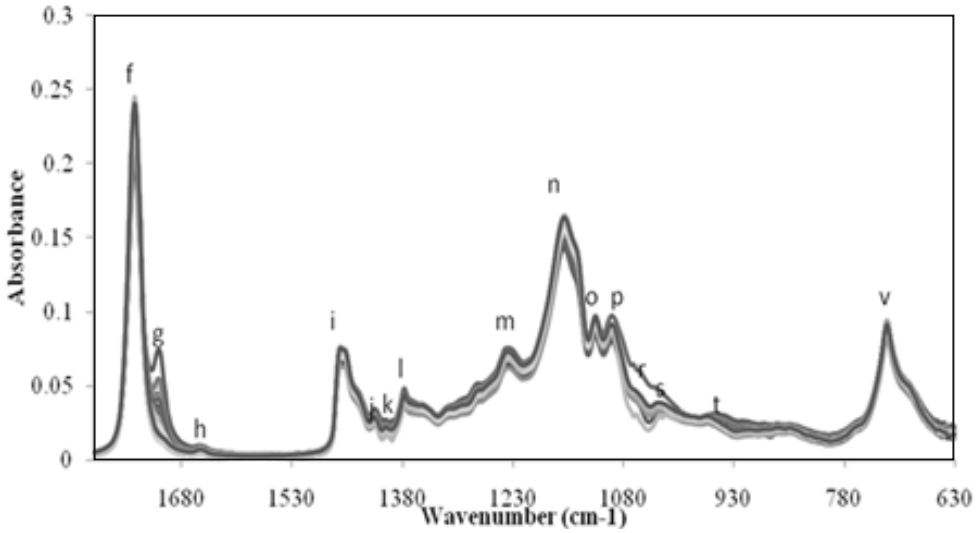
RESULT AND DISCUSSION

4.1. GILS Results

Spectrum 100 (Perkin Elmer, Waltham, MA, USA) Fourier Transform Infrared (FTIR) spectrometer system coupled to attenuated total reflectance (ATR) accessory was used to measure the olive oil samples. Both blank and sample spectra were collected in absorbance method. FTIR was used to develop of multivariate calibration methods for the determination of olive oil compositions. This calibration technique was used to determine for both the FAME and the TAG compositions of olive oil samples. The collected spectra are shown in Figure 5.1.



(a)



(b)

Figure 4.1. FTIR spectra of olive oil samples measured in the range of a)4000-600 cm^{-1}.b) 1800-600 cm^{-1} using ATR accessory attached diamond ZnSe crystal.

Figure 4.1 shows the spectrum of olive oil samples which are scanned between 4000 and 600 cm^{-1}. Maximum absorbance wave number that obtained from spectrum is 1746cm^{-1}. The band seen in 3470 cm^{-1} is the overtone of the glyceride ester carbonyl absorption. At 3009 cm^{-1} band are corresponded to CH stretching of = CH. 2960 cm^{-1} value band is shown symmetric and asymmetric vibration of aliphaticCH_3 groups. Band at 2925 cm^{-1} is corresponded asymmetric stretching of aliphatic CH_2 groups and the band at 2854 cm^{-1} is also corresponded symmetric stretching of aliphatic CH_2groups.Stretching of ester carbonyl functional group of triglycerides (C=O) are seen at 1745 cm^{-1}. The band seen as a shoulder in 1710 cm^{-1} is the acid group of free fatty acids. The peak illustrated in 1655 cm^{-1} is C=C stretching vibration of olefins. It is shown that the band at1460 $cm^{-1}$is bending vibration of CH_2 and CH_3 aliphatic groups and rocking vibration of CH bonds of cis-disubstituted olefins is seen at 1418 cm^{-1}.The peak between 1241 and 1033 cm^{-1} is stretching vibration of C-O ester and the peak can be seen at 950 cm^{-1} is bending vibration of out of plane of trans-disubstituted olefinic groups. Finally the band at 723 cm^{-1} is methylene rocking vibration and out-of –plane bending vibration of cis-disubstituted olefins. (Guillén and Cabo 1997, Vlachos, et al. 2006)

Near infrared spectra of the olive oil samples were collected with FTS-3000 NIR spectrometer (Bio-Rad, Excalibur, Cambridge, MA) and spectrums were collected

between 4500 and 10000 cm^{-1}at room temperature. The sample spectra were collected in absorbance method. Before and after each sample analyses background was collected to reduce the contaminations. Pure ethanol was used to clean the instrument. NIR was also used to develop of spectroscopic multivariate calibration methods for the determination of olive oil compositions. This spectroscopic technique was used for both FAME and TAG compositions of olive oil samples. The collected spectra are shown in Figure 4.2.

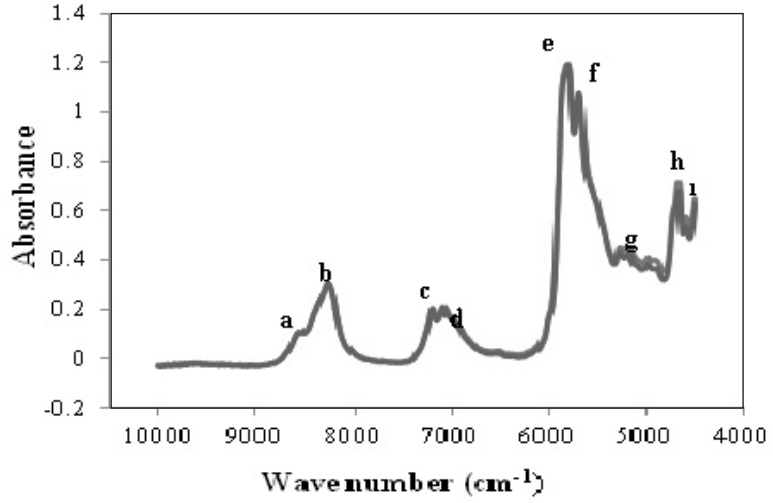


Figure 4.2. NIR spectra of olive oil samples measured in the range of 10000– 4500^{cm-1}.

Figure 4.2 shows the spectrum of olive oil samples which are scanned between 4500 and 10000 cm^{-1}. As can be seen band at 4590 cm^{-1}(ı) belongs to-CH=CH- and corresponds asymmetric stretching of C-H, and asymmetric stretching of C=C. The band at 4656 cm^{-1} (h) also belongs to-CH=CH- group and shows stretching =C-H and C=C. Maximum absorbance wave number that obtained from spectrum is 5794 cm^{-1} (e) .This maxima shows C-H stretching first overtone of CH_3 groups, The shoulder peak at 5675 cm^{-1} (f) C-H stretching first overtone of -CH_2 groups . At 7170 cm^{-1} (c) weak band corresponds 2C-H stretching of CH_3 group and at 7074 cm^{-}1(d) band shows 2C-H stretching of-CH_2 group. At 8560 cm^{-1} (a) weak and broad band second overtone of -CH_3groups, peak at 8240 cm^{-1} (b) corresponds second overtone of- C-H of CH_2 group. (Ozaki, et al. 2004).

After spectroscopic analysis, data are collected and calibration models are separately modeled both FAME and TAG compositions. First results will be given for FAME components.

4.1.1 Results of FAME Compositions

In order to construct FTIR multivariate calibration models for fatty acid methyl ester compositions, firstly interested fatty acid was chosen. Percentage values of olive oil samples for linolenic acid (LN) that is reference oil acid content are given in the following tables. The sets of measurement consist of two parts: calibration and validation sets. 53 olive oil samples were included in calibration set. On the other hand 26 samples were also assigned as validation set (Table 4.1 and Table 4.2). Likewise calibration and validation sets were created for NIR measurements. 53 olive oil samples were included in calibration and 26 samples were also assigned as validation set. For each set of samples absorbance matrices included spectral answers were created. Calibration models are constructed separately for each component by GILS method and tested. GILS is a wavelength selection based method, so when working with GILS, the program set to 30 genes, 40 iteration numbers, and 100 runs. These parameters are used for each calibration models for all TAG and FAME compositions.

Table 4.1. Percentage content by mass of Linolenic Acid (LN) in oil samples of calibration set

sample no	LN acid (w/w%)	sample no	LNacid(w/w%)	sample no	LN acid (w/w%)
1	0.64	19	0.89	37	0.64
2	0.87	20	0.79	38	0.78
3	0.69	21	0.62	39	0.72
4	0.92	22	0.68	40	0.56
5	0.68	23	0.72	41	0.75
6	0.8	24	0.75	42	1.14
7	0.71	25	1.03	43	0.99
8	0.79	26	0.57	44	0.62
9	0.74	27	1.05	45	1.03
10	0.75	28	1.03	46	0.61
11	0.62	29	0.99	47	0.76
12	1.02	30	0.95	48	1.13
13	0.78	31	0.71	49	1.08
14	0.68	32	1.02	50	1.12
15	1.02	33	0.77	51	0.56
16	0.8	34	0.74	52	1.2
17	0.92	35	0.66	53	1.15
18	0.8	36	0.9		

Table 4.2. Percentage content by mass of Linolenic acid (LN) in oil samples of validation set

sample no	LN acid (w/w%)	sample no	LN acid (w/w%)
1	0.96	14	0.77
2	0.88	15	0.67
3	0.75	16	0.85
4	0.78	17	0.94
5	0.82	18	0.66
6	0.66	19	0.83
7	0.66	20	0.69
8	0.64	21	0.98
9	0.67	22	0.67
10	0.81	23	0.78
11	0.86	24	0.97
12	0.74	25	0.93
13	0.69	26	0.94

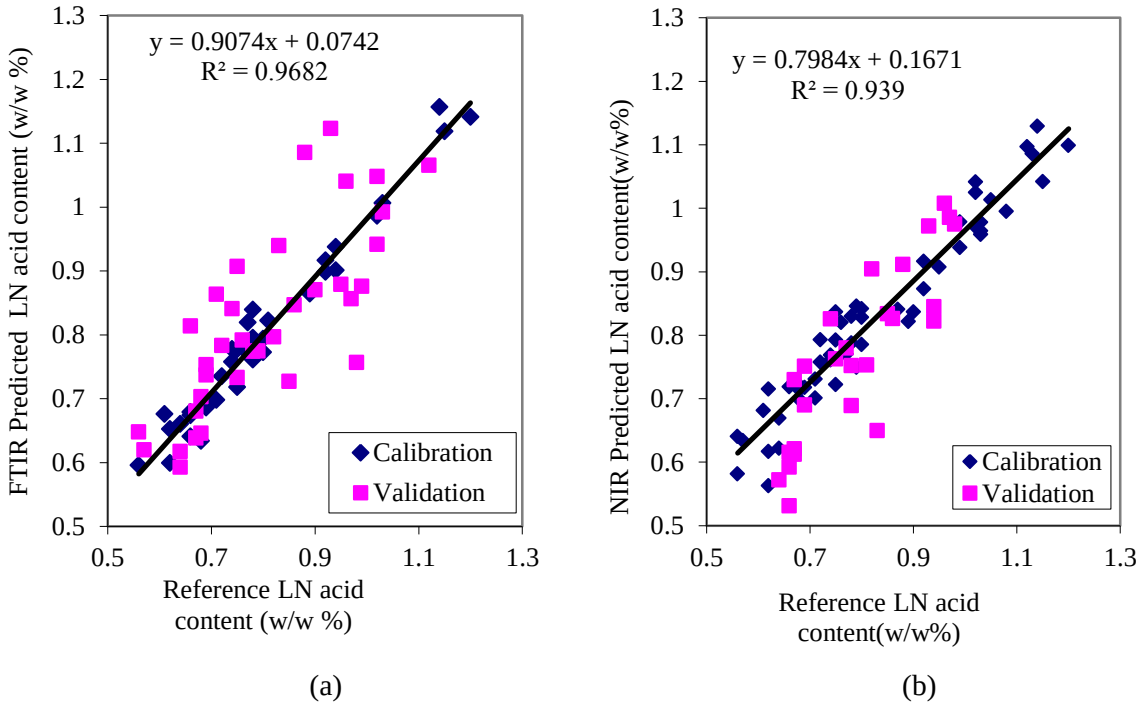


Figure 4.3. (a) Reference Linolenic acid (LN) content vs. predicted values based on FTIR-ATR spectra using GILS method(b) Reference Linolenic acid (LN) content vs. predicted values based on NIR spectra using GILS method

Reference linolenic acid contents versus predicted values based on FTIR and NIR spectra using GILS method are shown in Figure 4.3. Calibration models based on FTIR spectra for linolenic acid content determination gave standard error of calibration (SEC) and standard error of prediction (SEP) values as 0.0486%(w/w) and 0.1147%(w/w) for calibration and prediction sets. On the other hand calibration models based on NIR spectra for linolenic acid content determination, the SEC and SEP values were 0.0789%(w/w) and 0.1398%(w/w) for calibration and prediction sets, respectively. The R^2value of regression lines was 0.9682 for FTIR and that for linolenic acid content was 0.939 for NIR spectra. When compare the results of two device, it can be said that calibration model of FTIR showed a more uniform distribution than calibration model of NIR looking at R^2values.

Percentage values of olive oil samples for linoleic acid (LO) that is reference oil acid content are given in the Table 4.3 and Table 4.4.

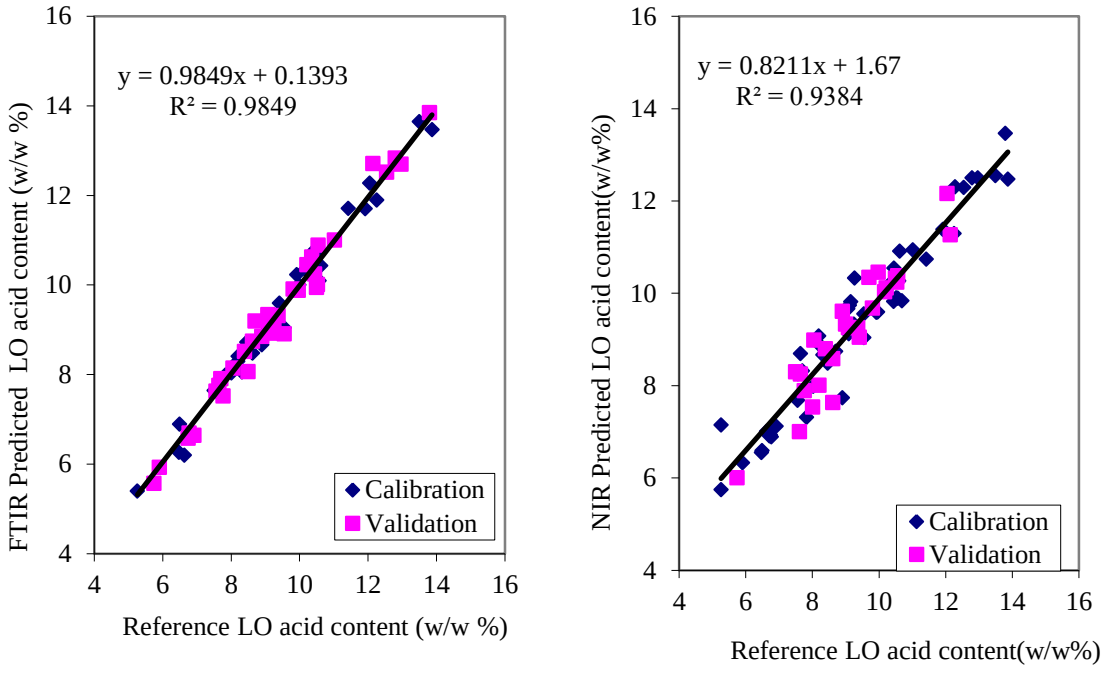
Table 4.3. Percentage content by mass of Linoleic acid (LO) in oil samples of calibration set

sample no	LO acid (w/w%)	sample no	LO acid (w/w%)	sample no	LO acid (w/w%)
1	8.46	19	9.25	37	7.7
2	8.03	20	9.15	38	8.9
3	9.55	21	12.29	39	5.9
4	6.71	22	8.49	40	12.8
5	9.96	23	7.57	41	12.97
6	7.83	24	6.74	42	11.92
7	9.1	25	10.44	43	6.63
8	8.7	26	9.07	44	10.52
9	10.45	27	9.27	45	12.26
10	12.55	28	7.56	46	8.19
11	9.32	29	11.02	47	11.42
12	8.19	30	9.92	48	13.5
13	6.91	31	10.36	49	13.8
14	6.49	32	8.32	50	9.12
15	10.62	33	10.6	51	7.64
16	6.47	34	10.57	52	13.87
17	6.77	35	5.26	53	5.26
18	9.55	36	10.69		

Table 4.4. Percentage content by mass of Linoleic acid (LO) in oil samples of validation set

sample no	LO acid (w/w%)	sample no	LO acid (w/w%)
1	8.62	14	8.01
2	9.09	15	8.62
3	8.39	16	9.98
4	12.05	17	8.05
5	9.37	18	7.50
6	10.17	19	8.20
7	9.81	20	7.76
8	9.00	21	10.57
9	10.22	22	10.54
10	7.62	23	12.14
11	9.41	24	9.70
12	10.50	25	5.74
13	8.90	26	7.64

53 olive oil samples were included in calibration and 26 samples were also assigned as validation set. For each set of samples absorbance matrices included spectral answers were created. Calibration models are constructed separately for each component by GILS method and tested.



(a) (b)

Figure 4.4. (a) Reference Linoleic acid (LO) content vs. predicted values based on FTIR-ATR spectra using GILS method (b) Reference Linoleic acid (LO) content vs. predicted values based on NIR spectra using GILS method

Looking at the Figure 4.4 in the SEC, SEP, and R^2 values, Calibration models based on FTIR spectra for linoleic acid content determination gave standard error of calibration (SEC) and standard error of prediction (SEP) values as 0.2525%(w/w) and 0.2699 % (w/w) for calibration and prediction sets. On the other hand calibration models based on NIR spectra for linoleic acid content determination, the SEC and SEP values were 0.9280 %(w/w) and 1.5198%(w/w) for calibration and prediction sets, respectively.FTIR spectrum of the calibration model can be seen in a better model was established. When SEC and SEP values are examined in the first model, it is seen that the agreement between these values are better than the NIR based model. On the other hand, the R^2 of calibration lines of NIR model were now lower than FTIR based model. In conclusion, linearity is provided between reference and predicted values in calibration model based on FTIR spectrum. By using GILS method and spectroscopic methods, calibration models are constructed and necessary and useful information can be easily obtained from the spectrum.

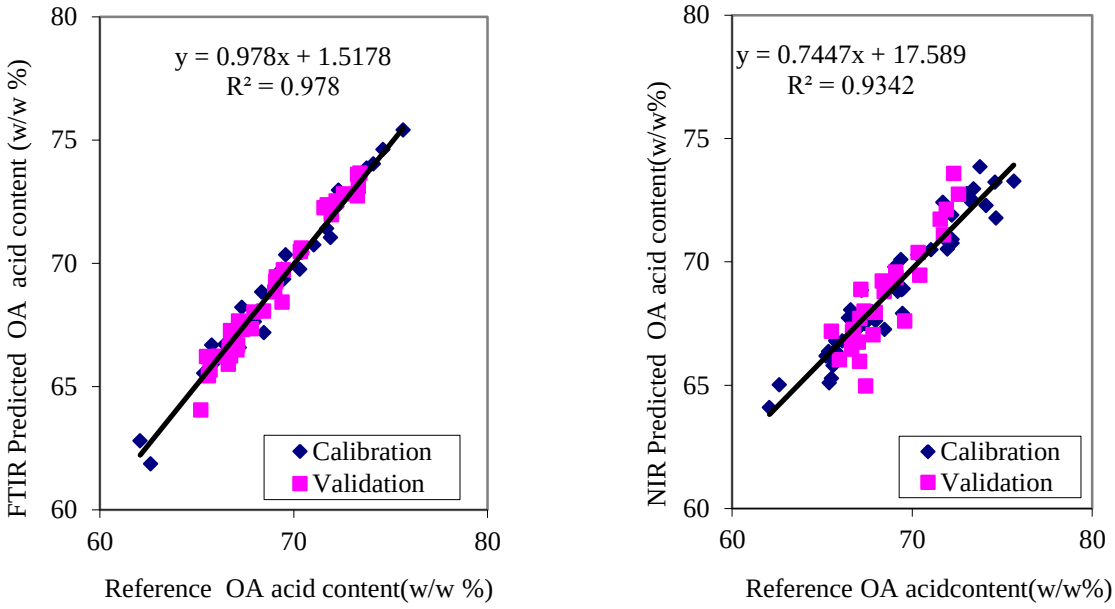
Oleic acid (OA) is the most abundant poly unsaturated fatty acid in olive oil. Percentage values of olive oil samples for oleic acid (OA) that is reference oil acid content are given in the Table 4.5 and Table 4.6 and calibration models are constructed.

Table 4.5. Percentage content by mass of Oleic acid (OA) in oil samples of calibration set

sample no	OA(w/w%)	sample no	OA(w/w%)	sample no	OA(w/w%)
1	65.39	19	69.38	37	66.94
2	65.63	20	65.52	38	73.29
3	65.21	21	66.12	39	68.47
4	72.20	22	66.45	40	66.58
5	71.70	23	71.95	41	70.42
6	67.97	24	67.15	42	72.22
7	67.78	25	73.76	43	66.46
8	69.49	26	69.41	44	74.61
9	73.31	27	66.78	45	65.74
10	67.14	28	65.64	46	72.22
11	71.04	29	65.34	47	69.47
12	73.03	30	65.77	48	62.07
13	65.84	31	65.36	49	67.19
14	65.69	32	67.16	50	74.11
15	73.33	33	66.75	51	62.61
16	67.36	34	69.03	52	75.65
17	69.24	35	71.94	53	74.66
18	70.32	36	73.41		

Table 4.6. Percentage content by mass of Oleic acid (OA) in oil samples of validation set

sample no	OA(w/w%)	sample no	OA(w/w%)
1	72.31	14	71.93
2	68.45	15	67.15
3	67.01	16	72.58
4	67.41	17	71.56
5	69.08	18	69.11
6	65.95	19	68.34
7	66.74	20	67.82
8	67.33	21	69.59
9	70.42	22	65.51
10	66.68	23	67.15
11	67.08	24	71.73
12	66.63	25	66.73
13	70.34	26	67.95



(a) (b)

Figure 4.5. (a) Reference OA acid content vs. predicted values based on FTIR-ATR spectra using GILS method (b) Reference OA acid content vs. predicted values based on NIR spectra using GILS method

Reference extractives and OA content versus predicted values based on NIR and FTIR-ATR spectra using GILS method are shown in Figure 4.4.Looking at the figures created and modeled for oleic acid, calibration models for OA content determination gave standard error of calibration (SEC) and standard error of prediction (SEP) values as 0.4942% (w/w) and 0.4486% (w/w) for FTIR results. Besides SEC and SEP values of NIR spectrum results were 1.6843% (w/w) and 2.3737 % (w/w). When these SEC and SEP values for NIR are examined, it is seen that the values are higher than the FTIR values. The R^2 value of regression lines for FTIR-ATR based model was 0.978 and that for NIR based model was 0.9342. When compare two methods, FTIR based model was more successful than NIR based calibration model.

Table 4.7. Percentage content by mass of Palmitic Acid (PA) in oil samples of calibration set

sample no	PA(w/w%)	sample no	PA(w/w%)	sample no	PA(w/w%)
1	13.82	19	11.94	37	12.13
2	14.98	20	15.37	38	15.37
3	12.77	21	12.32	39	15.59
4	12.95	22	15.05	40	15.66
5	14.33	23	12.26	41	11.77
6	14.14	24	14.48	42	11.99
7	14.86	25	14.86	43	11.89
8	13.43	26	13.92	44	14.44
9	14.62	27	14.19	45	11.86
10	13.19	28	12.85	46	13.26
11	14.79	29	14.68	47	15.53
12	12.83	30	14.87	48	16.11
13	14.39	31	15.24	49	12.64
14	13.79	32	15.46	50	15.92
15	13.39	33	12.43	51	10.41
16	14.13	34	13.05	52	12.46
17	12.03	35	11.54	53	15.92
18	13.48	36	14.31		

Table 4.8. Percentage content by mass of Palmitic Acid (PA) at oil samples of validation set

sample no	PA(w/w%)	sample no	PA(w/w%)
1	13.17	14	13.86
2	14.45	15	14.82
3	12.54	16	12.47
4	13.35	17	14.24
5	13.79	18	13.25
6	14.34	19	14.57
7	13.82	20	12.87
8	12.71	21	14.37
9	14.94	22	13.54
10	12.87	23	13.59
11	13.62	24	14.51
12	12.25	25	12.11
13	14.86	26	14.36

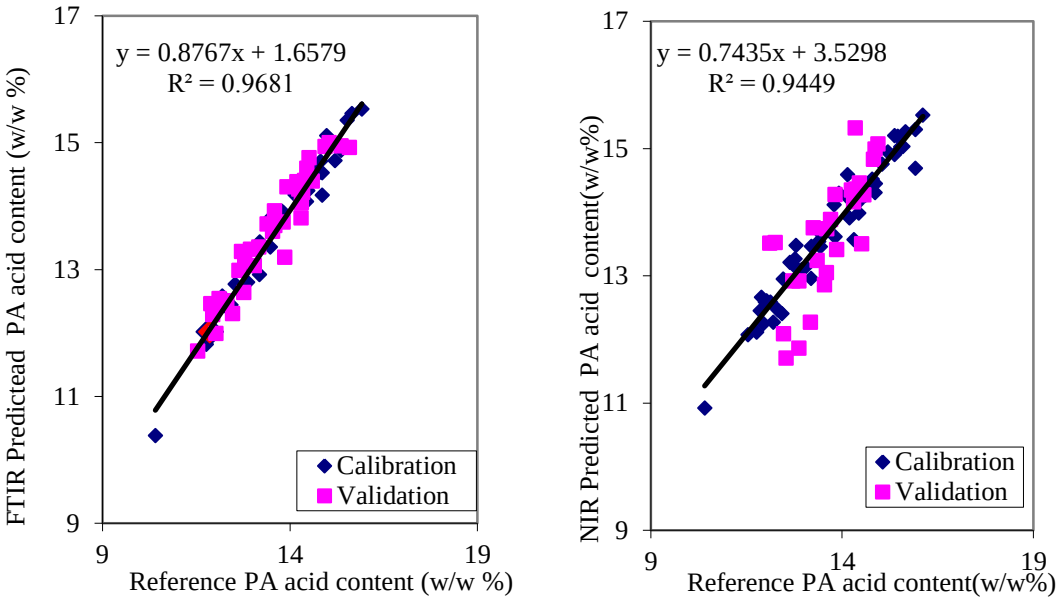


Figure 4.6. (a) Reference Palmitic acid (PA) content vs. predicted values based on FTIR-ATR spectra using GILS method (b) Reference Palmitic acid (PA) content vs. predicted values based on NIR spectra using GILS method

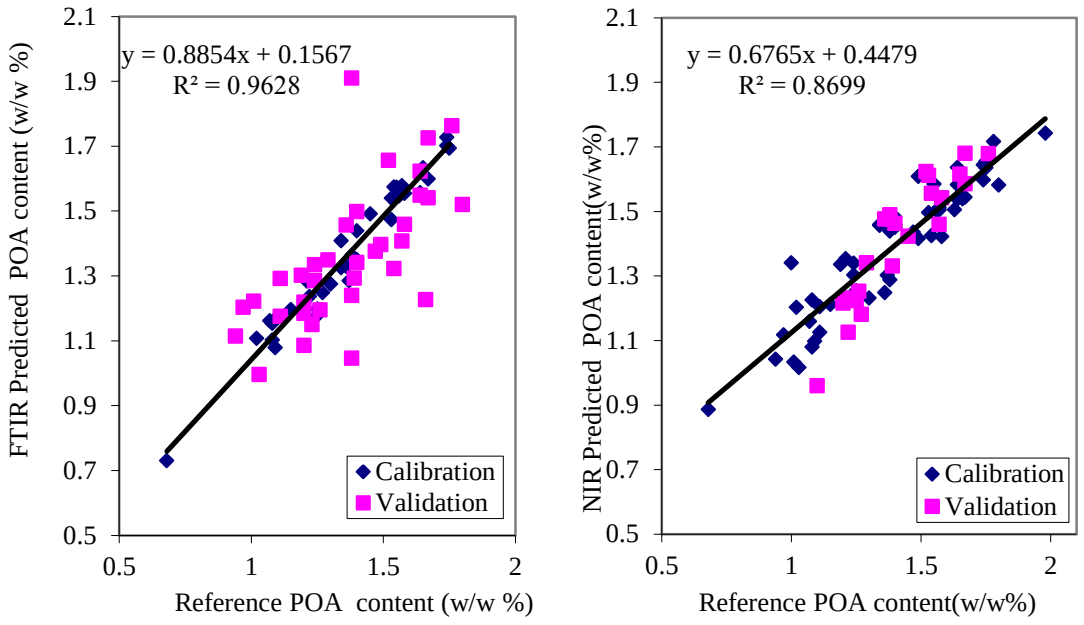
Table 4.7 and Table 4.8 were percentage content by mass of PA values for calibration and validation sets. Reference PA content versus predicted values based on NIR and FTIR spectra using GILS method are shown in Figure 4.6. When compare calibration models based on two instruments, it is seen that better calibration model is obtained from the results of FTIR-ATR. Reference values and predicted values showed good linearity in this method. Calibration models based on FTIR spectra for palmitic acid content determination gave standard error of calibration (SEC) and standard error of prediction (SEP) values as 0.4604%(w/w) and 0.4287%(w/w) for calibration and prediction sets.SEC and SEP values of FTIR based model are close to each other. But the plot of Figure 4.6.b, SEP value (1.3725%) is two times the value of SEC (0.7028%). The R^2 value of regression lines for FTIR based model was 0.9681 and R^2value for NIR based model 0.9449.

Table 4.9. Percentage content by mass of Palmitoleic Acid (POA) in oil samples of calibration set

sample no	**POA(w/w%)**	**sample no**	**POA(w/w%)**	**sample no**	**POA(w/w%)**
1	1.42	19	1.37	37	1.63
2	1.25	20	1.38	38	1.67
3	1.09	21	1.64	39	1.38
4	1.57	22	1.49	40	1.08
5	1.56	23	1.01	41	1.36
6	1.57	24	1.64	42	1.74
7	1.03	25	1.15	43	1.07
8	1.49	26	1.47	44	1.58
9	1.08	27	1.11	45	0.97
10	1.11	28	1.55	46	0.94
11	1.43	29	1.21	47	1.02
12	1.47	30	1.24	48	1.78
13	1.54	31	1.34	49	1.66
14	1.45	32	1.75	50	0.68
15	1.33	33	1.34	51	1.84
16	1.75	34	1.53	52	1.98
17	1.24	35	1.74	53	1.02
18	1.64	36	1.19		

Table 4.10. Percentage content by mass of Palmitoleic Acid (POA) in oil samples of validation set

sample no	POA(w/w%)	sample no	POA(w/w%)
1	1.36	14	1.22
2	1.65	15	1.27
3	1.29	16	1.17
4	1.25	17	1.24
5	1.23	18	1.53
6	1.67	19	1.67
7	1.45	20	1.26
8	1.52	21	1.76
9	1.26	22	1.38
10	1.48	23	1.29
11	1.25	24	1.57
12	1.58	25	1.54
13	1.39	26	1.38



(a) (b)

Figure 4.7. (a) Reference Palmitoleic acid (POA) content vs. predicted values based on FTIR-ATR spectra using GILS method (b) Reference Palmitoleic acid(POA) content vs. predicted values based on NIR spectra using GILS method

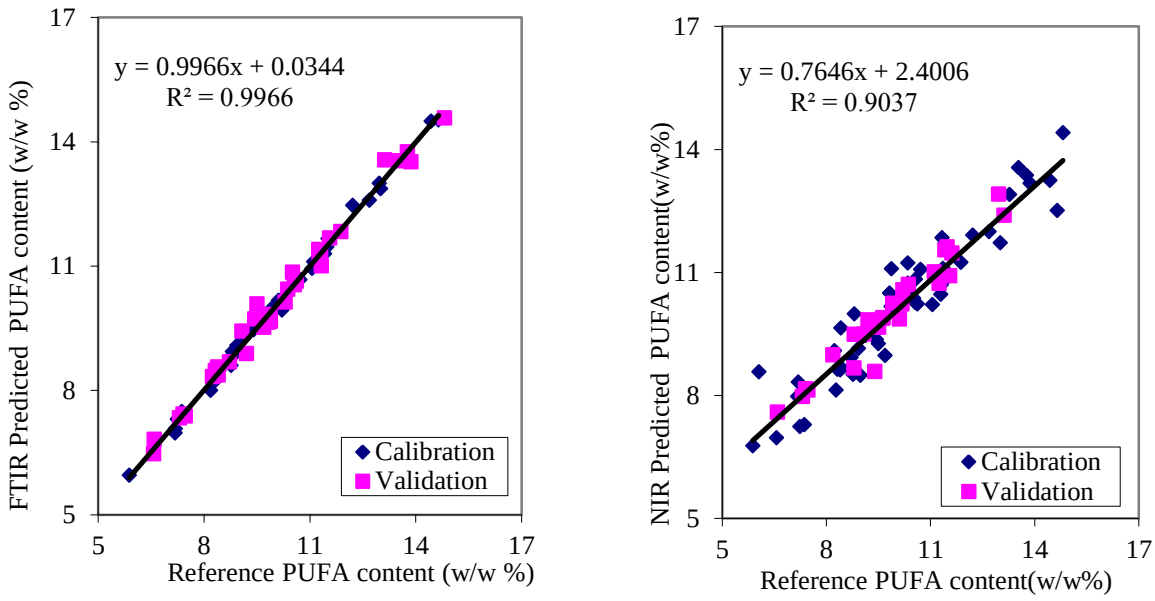
Table 4.9 and 4.10 and Figure 4.7 belong to palmitoleic acid which is the type of fatty acid. 53 olive oil samples were included in calibration and 26 samples were also assigned as validation set. When Figure 4.7 examined, calibration model based on FT-IR measurements gave the better results with calibration model based on FT-NIR data. Calibration models based on FT-NIR spectra for palmitoleic acid content determination gave standard error of calibration (SEC) and standard error of prediction (SEP) values as 0.158%(w/w) and 0.189%(w/w) for calibration and prediction sets. SEC (0.0817%w/w) and SEP (0.174%w/w) values of FTIR based models are narrower. Also predicted values and reference values have good linearity. When compared with the FT-NIR and FTIR-ATR results, SEC and SEP values became higher and thus regression became smaller. One possible explanation of this improvement could be attributed to increased number of calibration and prediction samples. The R^2 value of regression lines for FTIR was 0.9628 and that for NIR was 0.8699.

Table 4.11. Percentage content by mass of Polyunsaturated fatty acid(PUFA) in oil samples of calibration set

sample no	PUFA(w/w%)	sample no	PUFA(w/w%)	sample no	PUFA(w/w%)
1	8.32	19	8.72	37	11.52
2	10.34	20	7.37	38	11.26
3	7.18	21	8.77	39	14.82
4	13.54	22	11.35	40	8.81
5	10.62	23	9.14	41	11.31
6	10.57	24	8.93	42	13.01
7	9.44	25	8.98	43	11.36
8	11.33	26	9.69	44	9.82
9	8.28	27	10.72	45	11.06
10	10.56	28	6.57	46	10.35
11	10.51	29	11.88	47	7.19
12	8.75	30	13.77	48	9.88
13	13.28	31	11.35	49	8.42
14	10.35	32	8.24	50	5.88
15	9.53	33	12.69	51	14.44
16	7.24	34	8.39	52	14.65
17	8.43	35	13.87	53	6.06
18	9.87	36	12.22		

Table 4.12. Percentage content by mass of Polyunsaturated fatty acid(PUFA) in oil samples of validation set

sample no	PUFA(w/w%)	sample no	PUFA(w/w%)
1	11.56	14	10.18
2	11.62	15	9.64
3	9.91	16	10.37
4	10.11	17	11.25
5	11.11	18	10.21
6	11.42	19	7.47
7	12.97	20	11.48
8	9.51	21	9.08
9	8.81	22	7.31
10	8.19	23	7.41
11	8.82	24	6.59
12	9.43	25	9.62
13	9.28	26	13.12



(a) (b)

Figure 4. 8. (a) Reference Polyunsaturated fatty acid (PUFA) content vs. predicted values based on FTIR-ATR spectra using GILS method (b) Reference Polyunsaturated fatty acid(PUFA) content vs. predicted values based on NIR spectra using GILS method

Table 4.11 and Table 4.12 were percentage content by mass of PUFA values for calibration and validation sets. Reference PUFA content versus predicted values based on NIR spectra using GILS method are shown in Figure 4.8.When Figure 4.8 was examined, linearity of both models is observed that more than 90%.

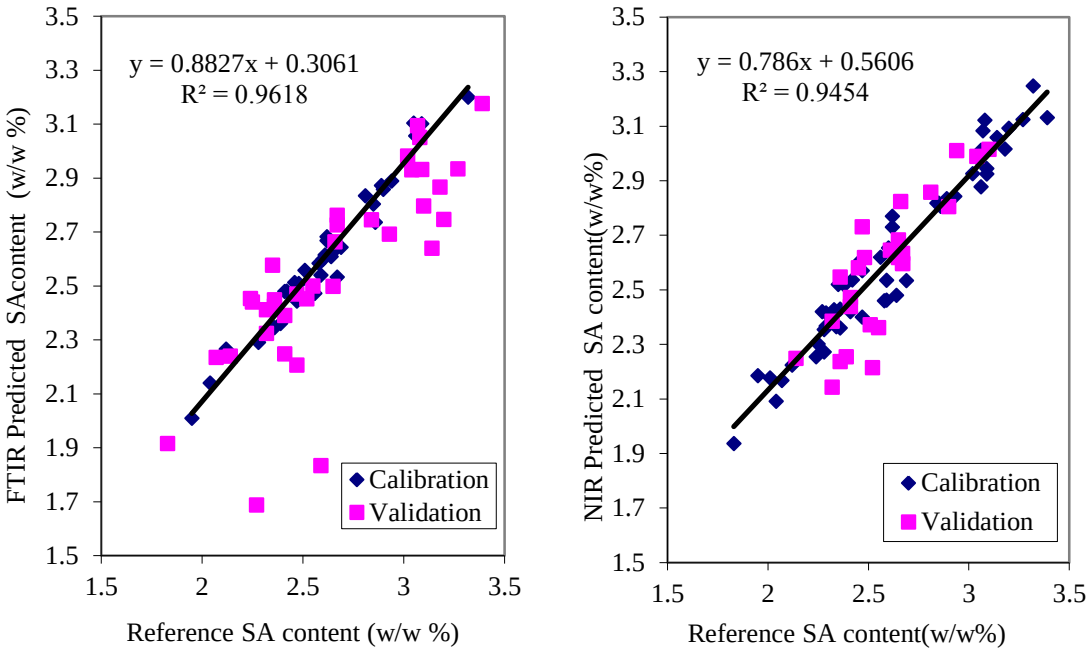
When compare calibration models based on two instruments, it is seen that better calibration model is obtained from the results of FTIR-ATR. Reference values and predicted values showed good linearity in this method.

Table 4.13. Percentage content by mass of Stearic acid (SA) in oil samples of calibration set

sample no	SA(w/w%)	sample no	SA(w/w%)	sample no	SA(w/w%)
1	2.59	19	2.93	37	2.12
2	2.85	20	2.62	38	2.86
3	2.36	21	2.24	39	2.69
4	3.08	22	3.32	40	3.05
5	2.66	23	2.46	41	3.27
6	2.41	24	2.47	42	2.07
7	2.27	25	2.47	43	2.62
8	2.56	26	2.29	44	2.42
9	3.02	27	2.38	45	2.25
10	2.28	28	2.89	46	2.01
11	2.84	29	3.14	47	3.09
12	2.34	30	2.04	48	3.39
13	2.33	31	2.69	49	3.09
14	3.07	32	2.59	50	3.06
15	2.36	33	3.06	51	1.83
16	2.58	34	2.29	52	3.18
17	2.64	35	3.26	53	1.95
18	2.28	36	2.35		

Table 4.14. Percentage content by mass of Stearic acid (SA) in oil samples of validation set

sample no	SA(w/w%)	sample no	SA(w/w%)
1	2.67	14	2.65
2	2.65	15	3.04
3	2.96	16	2.66
4	2.61	17	2.94
5	2.81	18	3.17
6	2.39	19	2.47
7	2.36	20	2.67
8	2.67	21	2.32
9	2.45	22	2.41
10	2.52	23	2.36
11	2.51	24	2.41
12	2.55	25	2.14
13	2.48	26	2.32

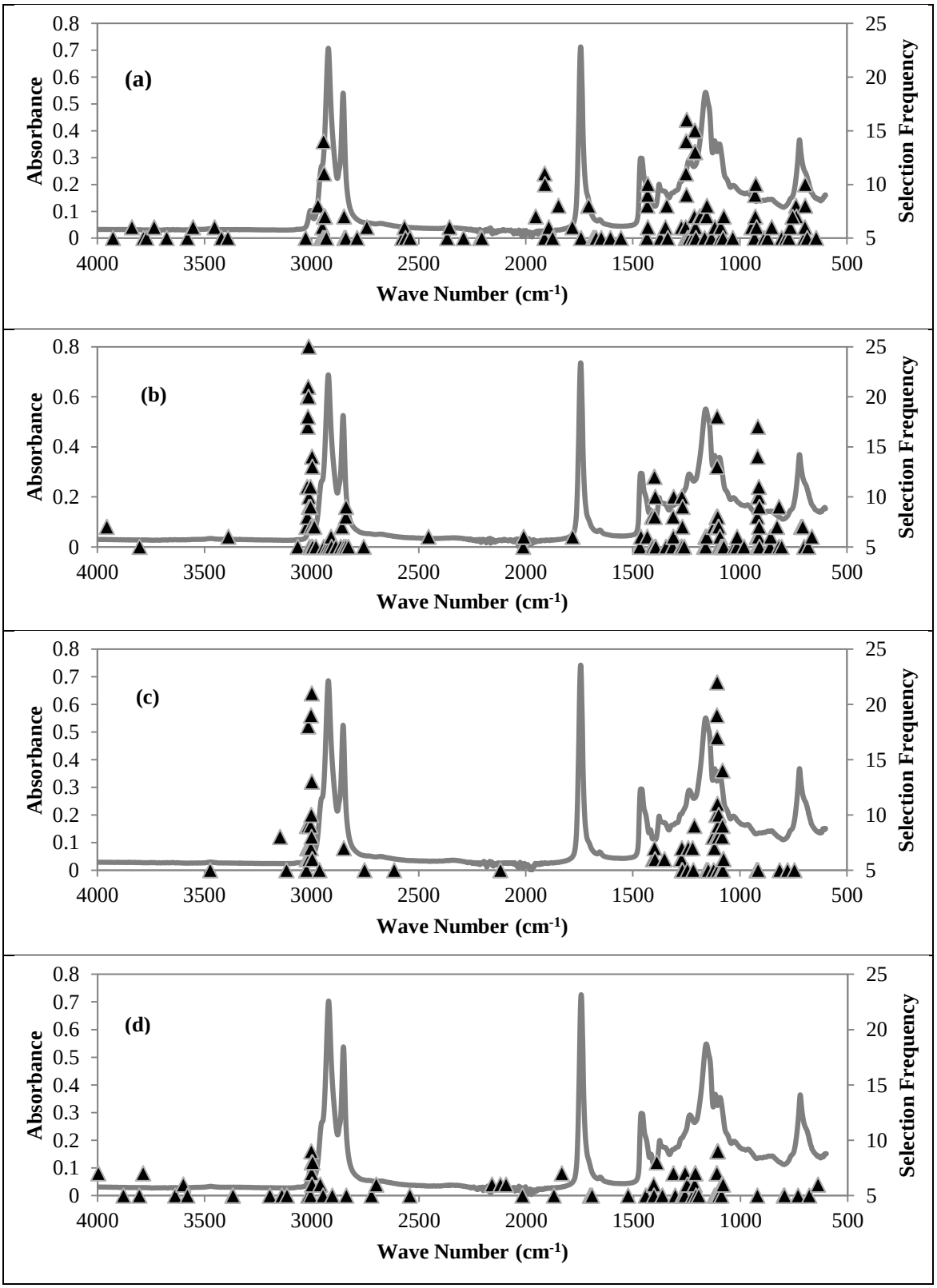


(a) (b)

Figure 4.9. (a) Reference Stearic acid (SA) content vs. predicted values based on FTIR-ATR spectra using GILS method (b) Reference Stearic acid (SA) content vs. predicted values based on NIR spectra using GILS method

Reference extractives and SA content versus predicted values based on NIR and FTIR-ATR spectra using GILS method are shown in Figure 4.9.Looking at the figures created and modeled for stearic acid, Calibration models for SA content determination gave standard error of calibration (SEC) and standard error of prediction(SEP) values as 0.1811% (w/w) and 0.2824 % (w/w) for NIR results. Besides SEC and SEP values of FTIR spectrum results were %0.0696 (w/w) and %0.2512 (w/w). When these SEC and SEP values are examined, it is seen that the values are smaller than the FT-NIR values. The R^2 value of regression lines for FTIR-ATR based model was 0.9618 and that for NIR based model was 0.9454. When compare two methods, FTIR based model was more successful than FT-NIR based calibration model.

Because GILS is a wavelength selection based method, it is interesting to observe the distribution of selected wavelengths in multiple runs over the entire full spectral region. Figure 4.10 illustrates the frequency distribution of selected wavelengths in 100 runs with 40 genes and 30 iterations for FAME of olive oil samples.


(a)
(b)
(c)
(d)
Absorbance
Selection Frequency
Wave Number (cm-1)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
25
20
15
10
5
4000
3500
3000
2500
2000
1500
1000
500

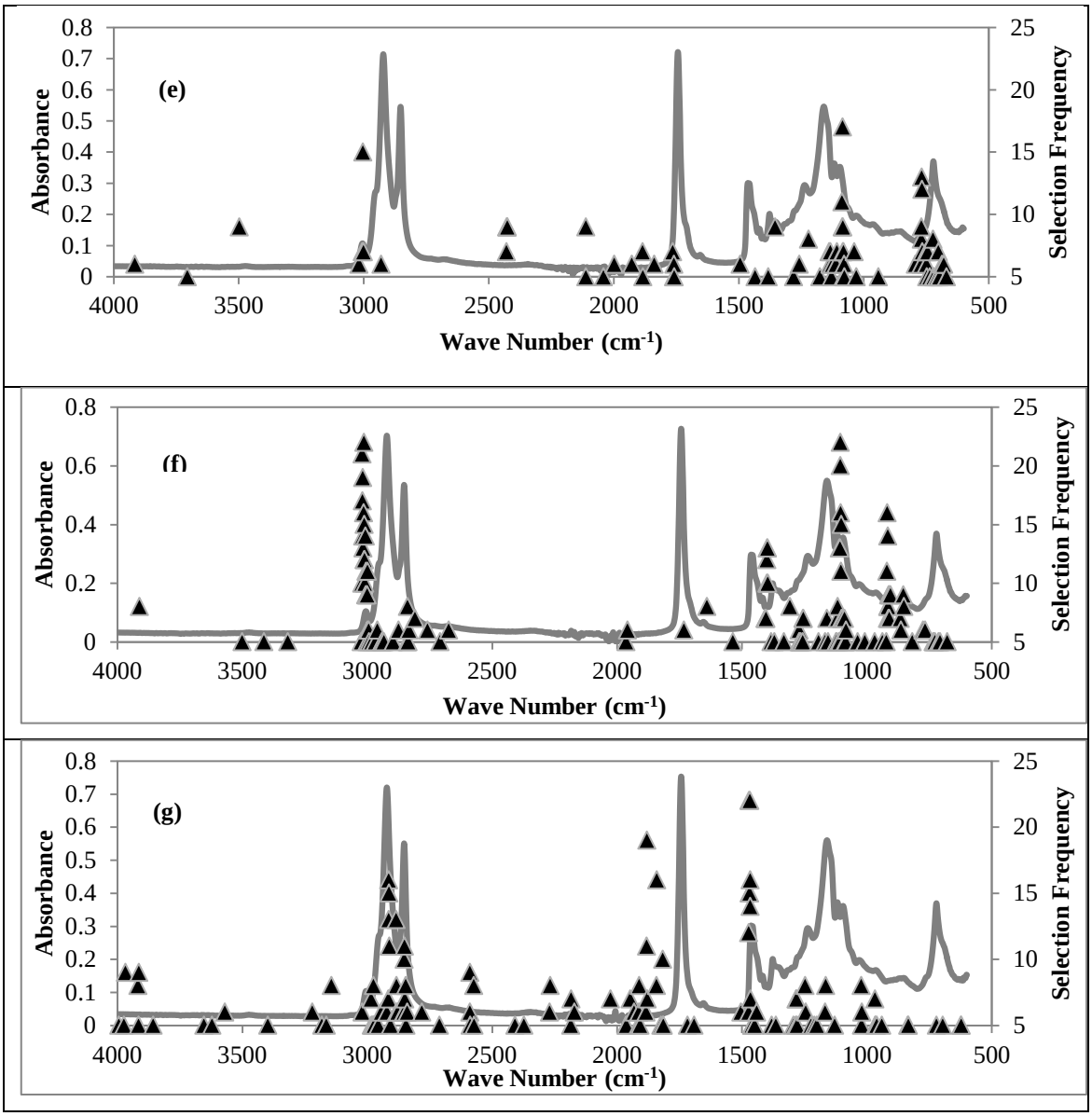


Figure 4.10. Frequency distribution of GILS selected FTIR wavelengths for FAME (a)Linolenic acid (LN) (b)Linoleic acid (LO) (c) Oleic acid(OA) (d)Palmitic acid(PA) (e) Palmitoleic acid(POA) (f)Polyunsaturated fatty acid (PUFA) (g) Stearic acid(SA) contents of olive oil samples

As can be seen from Figure 4.10 there are a number of regions where selection frequencies are very high compared to the rest of the spectrum. The wavelength region around 1240 and 3000 cm^{-1} for LN acid, around 900 and 3000 cm^{-1} for LO acid, around 1100 and 3000 cm^{-1} for OA acid, around 1105 and 3000 cm^{-1} for PA, around 1085 and 760 cm^{-1} for POA acid, around 1100 and 3020 cm^{-1} for PUFA indicates a strong

tendency for GILS method to select while for SA, around 1470 and 1470 cm^{-1} is the most frequently selected region.

4.2. Results of Tri-achyl Glycerol (TAG) Compositions

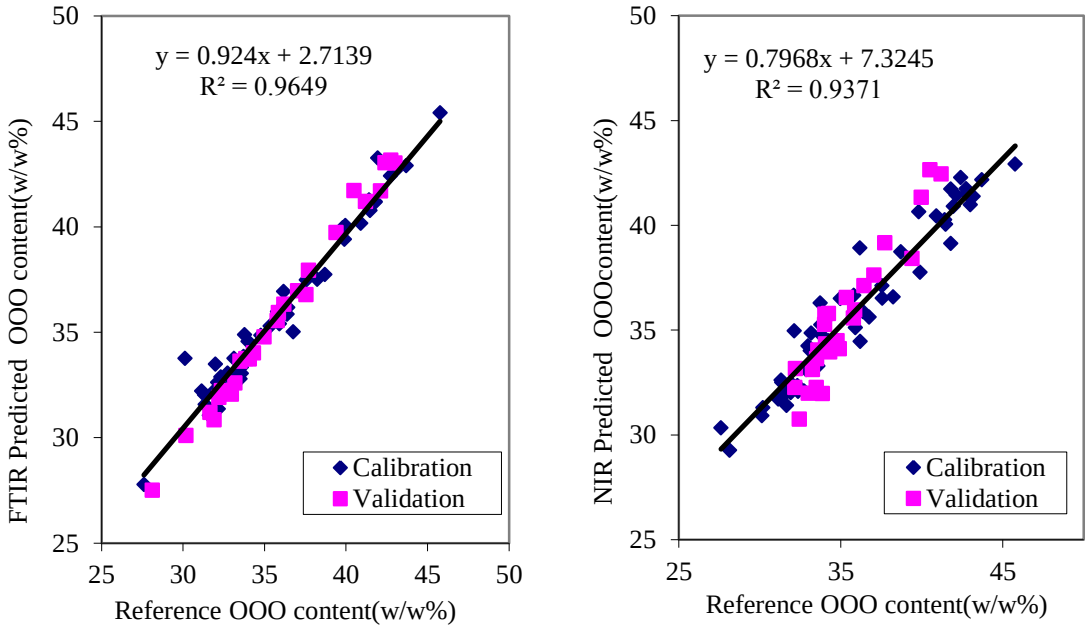
Triachylglycerol (TAG) compositions are important components in olive oil. Results from HPLC chromatography were used to reference in model. Also measurements were taken from FTIR and NIR instruments. Finally these methods were evaluated in success of prediction. First, results of oleic-oleic-oleic (OOO) component are given.

Table 4.15. Percentage content by mass of triolein (OOO) in oil samples of calibration set

sample no	OOO(w/w%)	sample no	OOO(w/w%)	sample no	OOO(%w/w)
1	34.02	19	41.79	37	41.47
2	32.99	20	33.72	38	32.12
3	34.25	21	27.61	39	33.12
4	31.36	22	39.96	40	34.03
5	32.45	23	32.86	41	41.94
6	42.73	24	36.38	42	33.58
7	33.85	25	40.90	43	45.77
8	37.56	26	36.43	44	41.41
9	39.97	27	31.91	45	32.32
10	35.83	28	41.78	46	31.32
11	33.25	29	34.24	47	31.14
12	35.35	30	35.91	48	43.78
13	34.77	31	33.77	49	36.17
14	33.59	32	32.74	50	36.76
15	31.61	33	38.79	51	31.99
16	39.81	34	32.05	52	30.13
17	33.51	35	43.18		
18	32.16	36	38.24		

Table 4.16. Percentage content by mass of triolein (OOO) in oil samples of validation set

sample no	OOO(w/w%)	sample no	OOO(w/w%)
1	31.66	14	33.18
2	32.37	15	37.04
3	32.26	16	42.12
4	39.48	17	35.78
5	36.24	18	32.97
6	41.19	19	37.54
7	43.97	20	31.92
8	30.19	21	42.48
9	34.97	22	33.51
10	28.13	23	34.32
11	34.07	24	40.53
12	33.86	25	42.75
13	37.71	26	35.86



(a) (b)

Figure 4.11. (a) Reference triolein (OOO) content vs. predicted values based on FTIR-ATR spectra using GILS method (b) Reference triolein (OOO) content vs. predicted values based on NIR spectra using GILS method

Reference triolein contents versus predicted values based on FTIR and NIR spectra using GILS method are shown in Figure 4.11. Calibration models based on

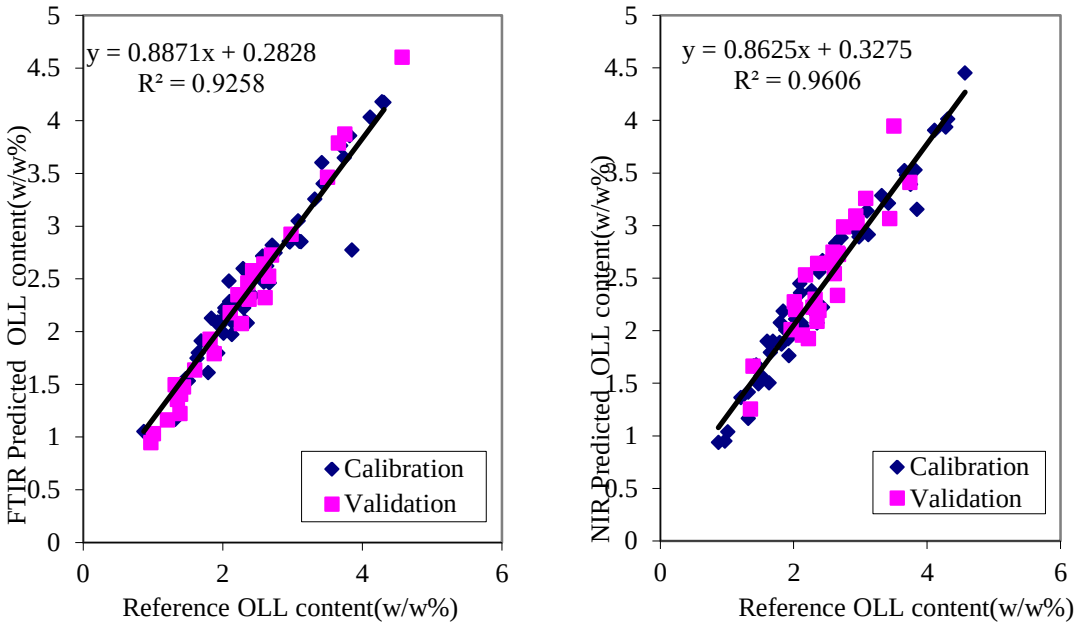
FTIR spectra for triolein content determination gave standard error of calibration (SEC) and standard error of prediction (SEP) values as 1.14%(w/w) and 1.40 %(w/w) for calibration and prediction sets. The other hand calibration models based on NIR spectra for triolein content determination, the SEC and SEP values were 2.11%(w/w) and 3.23%(w/w) for calibration and prediction sets, respectively. When these SEC and SEP values of FTIR based calibration model are examined, it is seen that these values are compatible with each other, which illustrates a good prediction. The R^2 value of regression lines was 0.9649 for FTIR and that for triolein content was 0.9371 for NIR spectra. When compare results of two devices, the calibration model of FTIR showed uniform distribution when looking at R^2 values.

Table 4.17. Percentage content by mass of 1,2-dilinoleyl-3-oleylglycerol (OLL) in oil samples of calibration set

sample no	**OLL(w/w%)**	**sample no**	**OLL(w/w%)**	**sample no**	**OLL(w/w%)**
1	1.32	19	3.08	37	1.79
2	2.71	20	2.96	38	1.84
3	3.74	21	1.80	39	3.32
4	2.01	22	2.63	40	2.10
5	3.44	23	2.66	41	4.28
6	3.42	24	1.93	42	2.18
7	2.75	25	4.31	43	3.10
8	1.63	26	2.35	44	2.67
9	2.59	27	2.32	45	1.69
10	2.95	28	2.13	46	2.35
11	3.82	29	2.03	47	0.87
12	2.03	30	4.11	48	2.02
13	3.69	31	2.13	49	3.12
14	1.92	32	2.57	50	2.29
15	1.96	33	2.38	51	2.09
16	1.47	34	1.65	52	3.85
17	1.51	35	2.43		
18	2.30	36	1.93		

Table 4.18. Percentage content by mass of OLL (1,2-dilinoleyl-3-oleylglycerol) at oil samples of validation set

sample no	OLL(w/w%)	sample no	OLL(w/w%)
1	1.40	14	1.01
2	4.57	15	1.44
3	1.35	16	3.66
4	2.70	17	2.66
5	1.60	18	2.98
6	2.61	19	1.21
7	3.50	20	1.32
8	2.22	21	2.59
9	2.43	22	2.38
10	0.97	23	2.36
11	2.27	24	1.88
12	2.11	25	1.39
13	3.75	26	1.82



(a) (b)

Figure 4.12. (a) Reference 1,2-dilinoleyl-3-oleylglycerol (OLL) content vs. predicted values based on FTIR-ATR spectra using GILS method (b) Reference 1,2-dilinoleyl-3-oleylglycerol (OLL) content vs. predicted values based on NIR spectra using GILS method

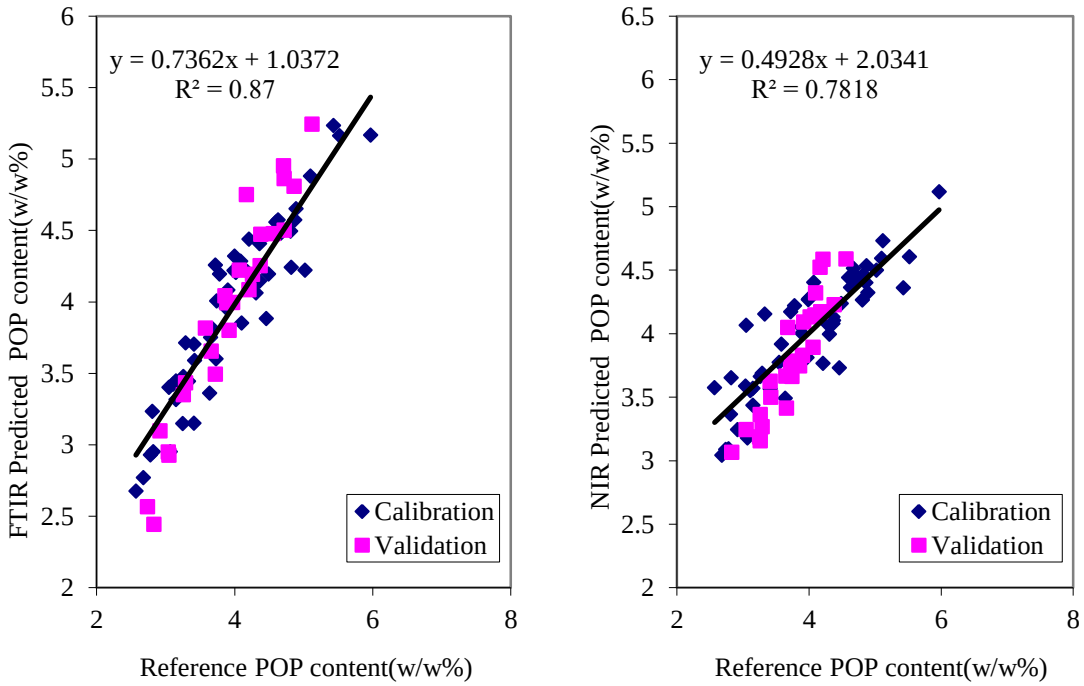
Table 4.17 illustrates that percentage content of OLL in calibration, Table 4.18 also illustrates that percentage content of OLL in validation sets. 52 olive oil samples were used in calibration, 26 olive oil samples were also used in validation model. Calibration models based on FTIR spectra for OLL content determination gave SEC and SEP values as 0.2688%(w/w) and 0.2598%(w/w) for calibration and prediction sets. The other hand calibration models based on NIR spectra for OLL content determination, the SEC and SEP values were 0.3504%(w/w) and 0.5966%(w/w) for calibration and prediction sets, respectively. When compared with FTIR and NIR results, SEC and SEP values became higher and thus regression became smaller, but NIR based model gave a compatible agreement with reference and predicted values. The models based on FTIR and NIR spectrum results, NIR results gave good linearity with reference values. The R^2valueof regression line is 0.9606 for NIR, so NIR based model is better than the model based on FTIR results.

Table 4.19. Percentage content by mass of 1,2-dipalmitoyl-3-oleylglycerol(POP) in oil samples of calibration set

sample no	POP(w/w%)	sample no	POP(%w/w)	sample no	POP(%w/w)
1	3.25	19	3.26	37	3.99
2	4.36	20	3.33	38	3.11
3	4.36	21	3.26	39	4.81
4	3.65	22	2.82	40	4.31
5	3.88	23	2.68	41	5.52
6	3.73	24	3.68	42	3.29
7	4.09	25	4.02	43	3.05
8	4.06	26	3.41	44	2.57
9	4.21	27	3.97	45	4.14
10	4.63	28	5.43	46	3.78
11	4.63	29	3.74	47	3.15
12	4.53	30	2.81	48	4.46
13	3.42	31	4.89	49	3.72
14	4.18	32	4.49	50	4.82
15	3.41	33	3.15	51	5.97
16	3.07	34	3.64	52	5.02
17	5.12	35	4.67		
18	4.87	36	2.78		

Table 4.20. Percentage content by mass of 1,2-dipalmitoyl-3-oleylglycerol(POP) in oil samples of validation set

sample no	POP(%w/w)	sample no	POP(%w/w)
1	2.74	14	4.72
2	4.07	15	3.92
3	4.72	16	3.29
4	3.05	17	2.83
5	2.92	18	3.66
6	4.17	19	3.26
7	3.97	20	3.04
8	4.56	21	4.37
9	3.89	22	3.58
10	4.21	23	3.86
11	4.86	24	4.71
12	4.26	25	3.72
13	4.38	26	5.12



(a) (b)

Figure 4.13. a) Reference 1,2-dipalmitoyl-3-oleylglycerol (POP) content vs. predicted values based on FTIR-ATR spectra using GILS method (b) Reference 1,2-dipalmitoyl-3-oleylglycerol (POP) content vs. predicted values based on FT-NIR spectra using GILS method

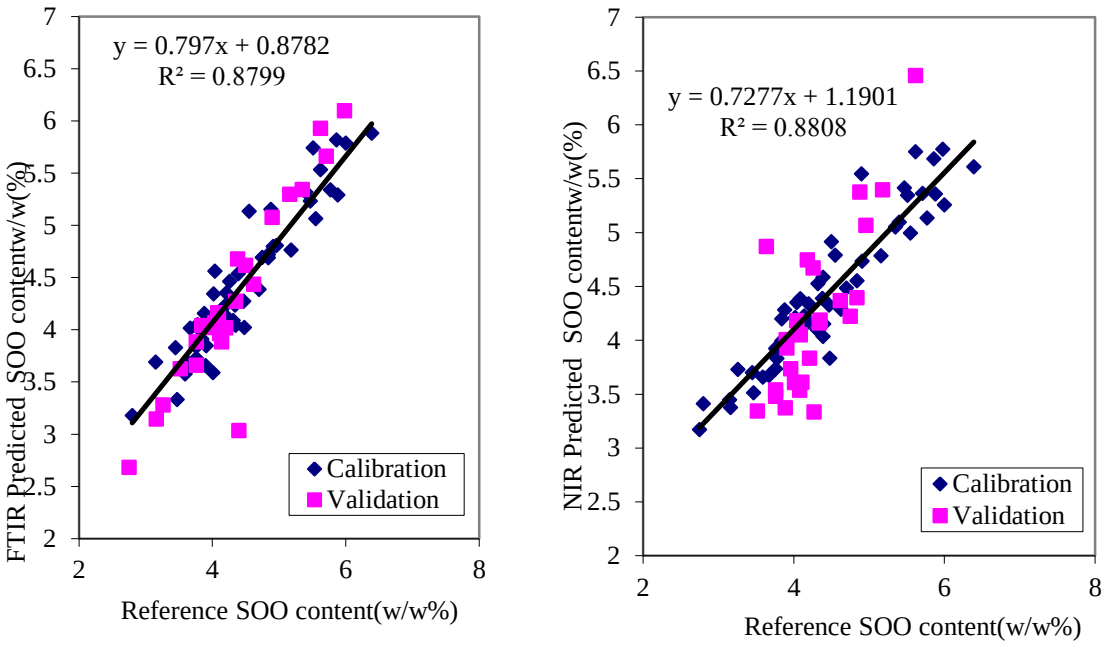
Reference POP (1,2-dipalmitoyl-3-oleylglycerol) contents versus predicted values based on FTIR and NIR spectra using GILS method are shown in Figure 4.13. Calibration models based on FTIR spectra for POP (1,2-dipalmitoyl-3-oleylglycerol)content determination gave standard error of calibration (SEC) and standard error of prediction (SEP) values as 0.39%(w/w) and 0.45 %(w/w) for calibration and prediction sets. The other hand calibration models based on NIR spectra for content determination, the SEC and SEP values were 0.60%(w/w) and 0.56%(w/w) for calibration and prediction sets, respectively. The R^2 value of regression lines was 0.87 for FTIR and that for POP (1,2-dipalmitoyl-3-oleylglycerol) content was 0.7818 for NIR spectra. It can be said that the calibration model of FTIR instrument has a good distribution than the model of FT-NIR.R^2 values of FTIR showed this result.

Table 4.21. Percentage content by mass of 2,3-dioleyl-1-stearoylglycerol (SOO) in oil samples of calibration set

sample no	**SOO(w/w%)**	**sample no**	**SOO(w/w%)**	**sample no**	**SOO(w/w%)**
1	5.46	19	6.25	37	5.47
2	4.18	20	5.51	38	4.02
3	4.19	21	3.59	39	3.45
4	3.91	22	3.78	40	5.18
5	4.84	23	3.76	41	3.88
6	3.84	24	4.96	42	2.83
7	4.91	25	4.39	43	5.55
8	3.89	26	4.84	44	4.55
9	5.86	27	4.34	45	5.77
10	4.75	28	4.32	46	4.01
11	4.63	29	3.95	47	5.88
12	3.76	30	5.62	48	6.39
13	3.64	31	4.73	49	3.15
14	3.67	32	3.78	50	4.48
15	4.21	33	4.27	51	4.04
16	4.47	34	4.88	52	2.88
17	4.31	35	3.47		
18	4.26	36	4.35		

Table 4.22. Percentage content by mass of 2,3-dioleyl-1-stearoylglycerol (SOO) in oil samples of validation set

sample no	SOO(w/w%)	sample no	SOO(w/w%)
1	3.16	14	2.75
2	4.35	15	4.23
3	4.94	16	4.48
4	4.04	17	4.09
5	3.52	18	4.38
6	3.76	19	3.96
7	4.11	20	4.08
8	4.09	21	5.62
9	3.26	22	4.14
10	5.71	23	3.76
11	5.35	24	5.98
12	4.50	25	4.62
13	3.84	26	5.16



(a) (b)

Figure 4.14. (a) Reference 2,3-dioleyl-1-stearoylglycerol (SOO) content vs. predicted values based on FTIR-ATR spectra using GILS method (b) Reference 2,3-dioleyl-1-stearoylglycerol (SOO) content vs. predicted values based on FT-NIR spectra using GILS method

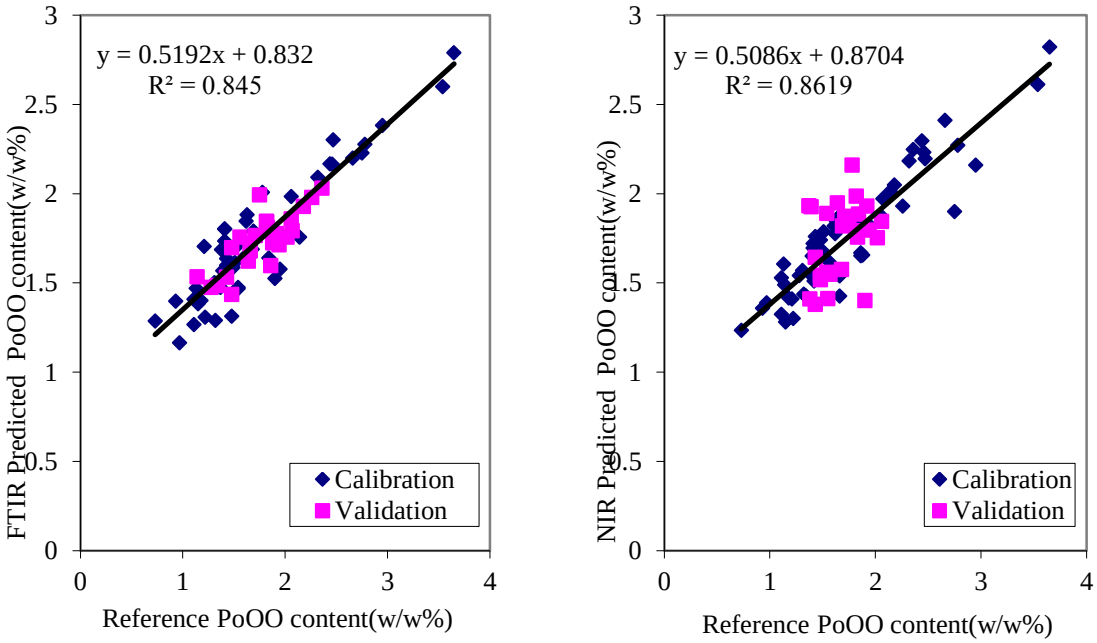
The Table 4.21 and 4.22 and figure4.14 belongs to triachyl glycerol SOO (2,3-dioleyl-1-stearoylglycerol) type triachyl glycerol. 52 olive oil samples were included in calibration and 26 samples were also assigned as validation set. When Figure 4.14 examined, calibration model based on FT-NIR measurements gave approximately the same results with calibration model based on FTIR-ATR data. When compared with the NIR and FTIR-ATR results, SEC and SEP values became higher and thus regression became smaller. Lastly the calibration model based on FT-NIR gave the better result than the model based on NIR.

Table 4.23. Percentage content by mass of 1-palmitioleoyl,2,3-dioleoyl (PoOO) in oil samples of calibration set

sample no	**PoOO(w/w%)**	**sample no**	**PoOO(w/w%)**	**sample no**	**PoOO(w/w%)**
1	1.66	19	1.43	37	2.44
2	1.37	20	1.41	38	1.21
3	1.43	21	1.49	39	1.63
4	1.42	22	1.15	40	2.14
5	1.61	23	1.11	41	2.75
6	1.57	24	1.95	42	1.13
7	1.55	25	1.85	43	1.41
8	1.22	26	1.93	44	0.93
9	1.39	27	1.69	45	1.78
10	1.68	28	1.54	46	2.66
11	1.84	29	0.97	47	2.46
12	1.54	30	1.48	48	0.73
13	1.32	31	2.06	49	2.95
14	1.31	32	1.51	50	2.78
15	2.47	33	1.62	51	3.54
16	2.32	34	1.18	52	3.65
17	1.86	35	1.38		
18	1.11	36	1.41		

Table 4.24. Percentage content by mass of 1-palmitioleoyl,2,3-dioleoyl (PoOO) in oil samples of validation set

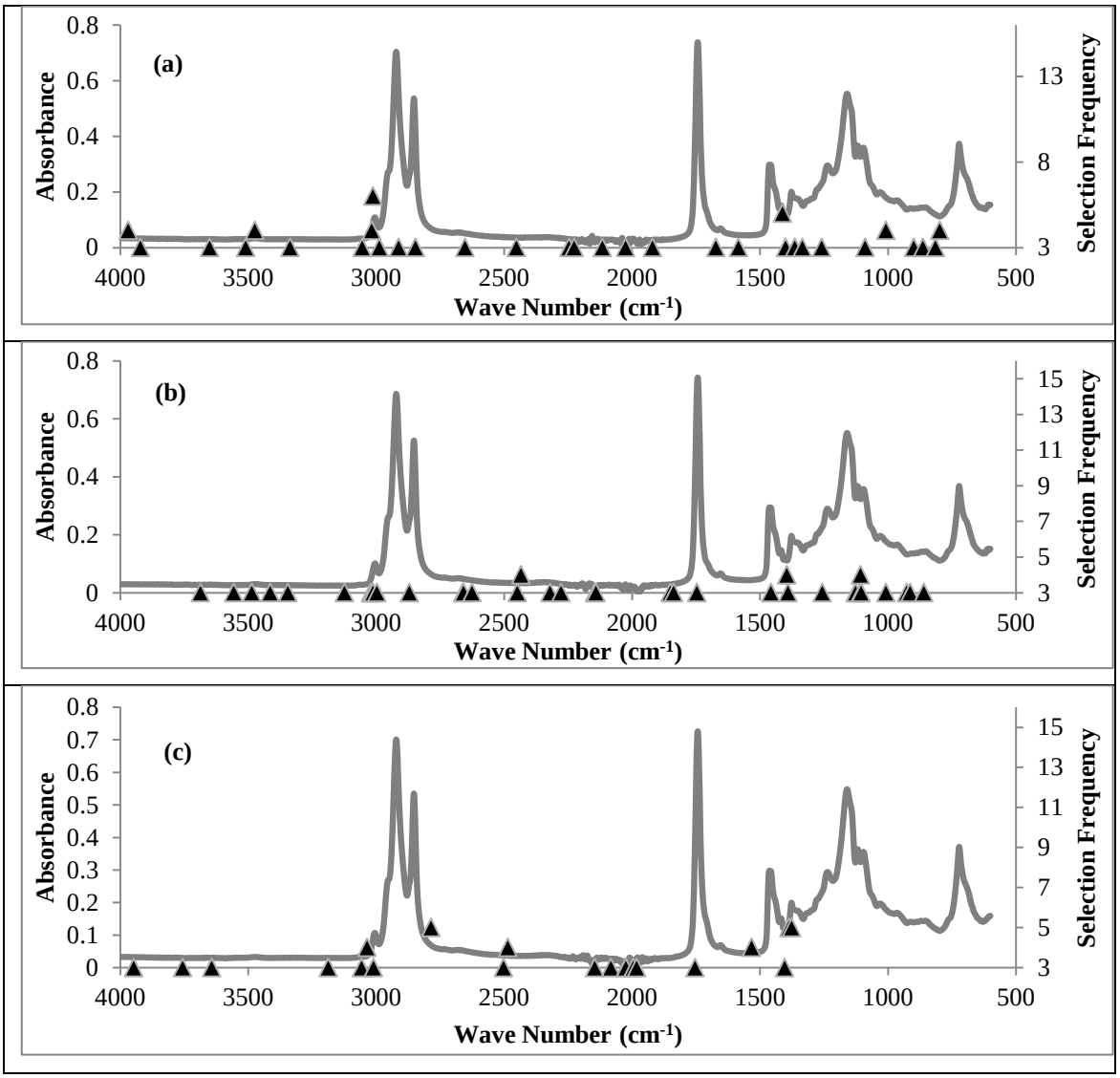
sample no	PoOO(w/w%)	sample no	PoOO(w/w%)
1	2.18	14	2.02
2	2.06	15	1.56
3	1.75	16	1.83
4	1.66	17	1.48
5	1.82	18	1.82
6	1.43	19	1.71
7	1.86	20	1.66
8	2.36	21	2.26
9	1.88	22	1.28
10	1.92	23	1.94
11	2.07	24	1.14
12	1.48	25	1.89
13	1.64	26	1.34



(a) (b)

Figure 4.15. (a) Reference 1-palmitioleoyl,2,3-dioleoyl (PoOO) content vs. predicted values based on FTIR-ATR spectra using GILS method (b) Reference 1-palmitioleoyl,2,3-dioleoyl (PoOO) content vs. predicted values based on FT-NIR spectra using GILS method

Table 4.23 and Table 4.24 are percentage content of calibration and validation sets of PoOO values. Looking at Figure 4.15, the linearity of both models is observed that more than 85%.The results are very comparable but NIR results seem to have somewhat better R^2values for calibration models. Reference and predicted values have good linearity and so it can be said the model of NIR is better than the model of FTIR. In conclusion, predicted values based on both two instrument using GILS method shows the success of calibration model. Figure 4.16 illustrates the frequency distribution of selected wavelengths in 100 runs with 40 genes and 30 iterations for TAG of olive oil samples.



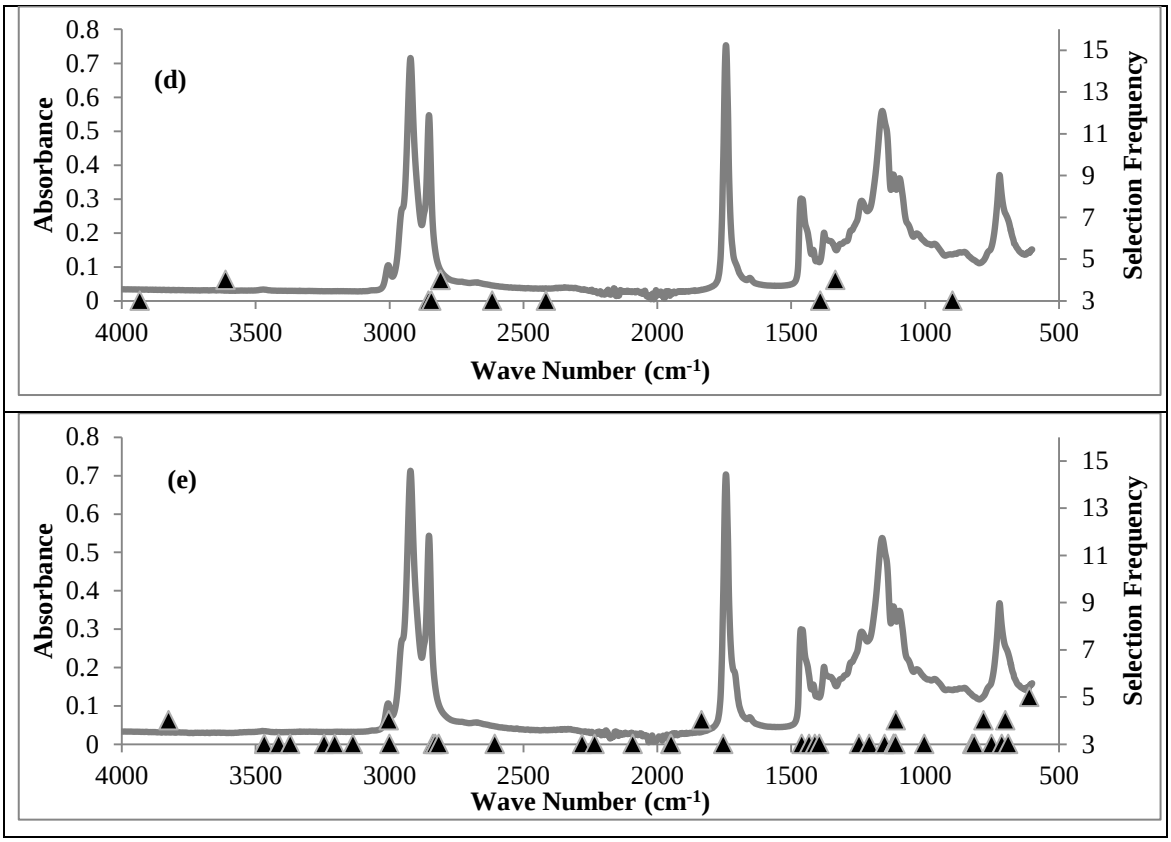


Figure 4.16. Frequency distribution of GILS selected FTIR wavelengths for TAG (a)OLL acid (b)OOO acid (c) PoOO (d)SOO (e) POP contents of olive oil samples

As can be seen from Figure 4.16 there are a number of regions where selection frequencies are very high compared to the rest of the spectrum. The wavelength region around 1500 and 3000 cm^{-1} for OLL acid, around 950 and 2400 cm^{-1} for OOO acid, around 2000 and 3000 cm^{-1} for PoOO acid, around 1300 and 2850 cm^{-1} for SOO, around 3300 and 1300 cm^{-1} for POP acid indicates a strong tendency for GILS method to select which is the most frequently selected region.

4.3. PLS Results

PLS is a kind of factor based methods that combines ILS approach using factor data. PLS assumes that the error can be derived from both absorbance readings and from the measurement of component concentration. Like ILS, PLS does not require the user to prepare a calibration standard where all of the interfering species are known.

Also, it tries to find the factors which have the greatest relevance for prediction. The PLS calibration models for each components were first calculated using cross–validation. And all data was mean–centered not scaled. There are two main advantages of cross–validation methods. The first is estimation of the performance of the model. Since the predicted samples are not same as the samples used build the model. The second benefit of cross–validation is better outlier detection since each sample is left out of the models during the cross–validation process. On the other hand, it is a very time–consuming process. Mean centering translates the collection of data to the origin of multivariate space where analysis will be performed. It also removes the need for an intercept from the regression model. Since fewer terms in the regression model may need to be estimated and estimated analyte concentrations may be more precise following mean centering of the data. More of the information content of a data set can usually be described with a simpler model if the data is mean centered. The major effect of mean–centering is removing the broad sloping background from the data collection. Also PLS–1 algorithm (one component at a time) was used here.

In this study, PLS was performed for both sets of FAME and TAG using the data from the FTIR and NIR spectra. The aim is to provide the most successful model for the prediction of concentrations. From Figure 4.17 to Figure 20 illustrates PLS results for the set of FAME compositions. According to between Figure 4.17. to Figure 4.20 , the models for FAME components have low prediction ability of concentrations with low regression coefficients. The models are also suffering from high collinearity and cannot be trusted for producing a high calibration quality and for making prediction of concentrations in the validation samples.

Generally, as can be seen Figure 4.17.,Figure 4.18, Figure 4.19 and Figure 4.20 reference versus predicted values are not compatible with each other. When overall calibration performance models examine, it is possible to state that PLS is not best calibration method; since all FAME components have worst regression coefficient when comparing GILS method.

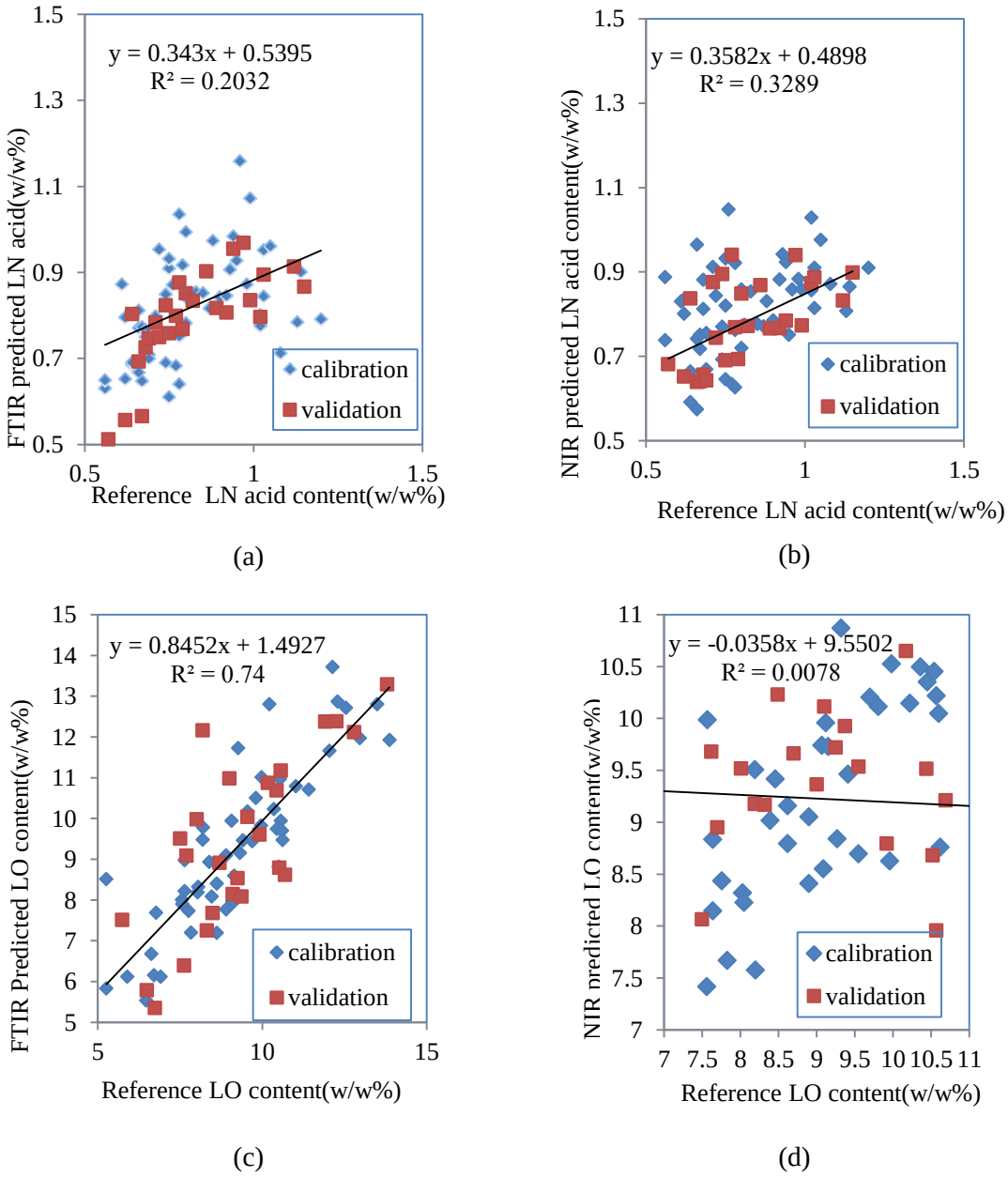


Figure 4.17. (a) Reference Linolenic acid (LN) content vs. predicted values based on FTIR-ATR spectra using PLS method (b) Reference Linolenic acid (LN) content vs. predicted values based on NIR spectra using PLS method (c) Reference Linoleic acid (LO) content vs. predicted values based on FTIR-ATR spectra using PLS method (d) Reference Linoleic acid (LO) content vs. predicted values based on NIR spectra using PLS method

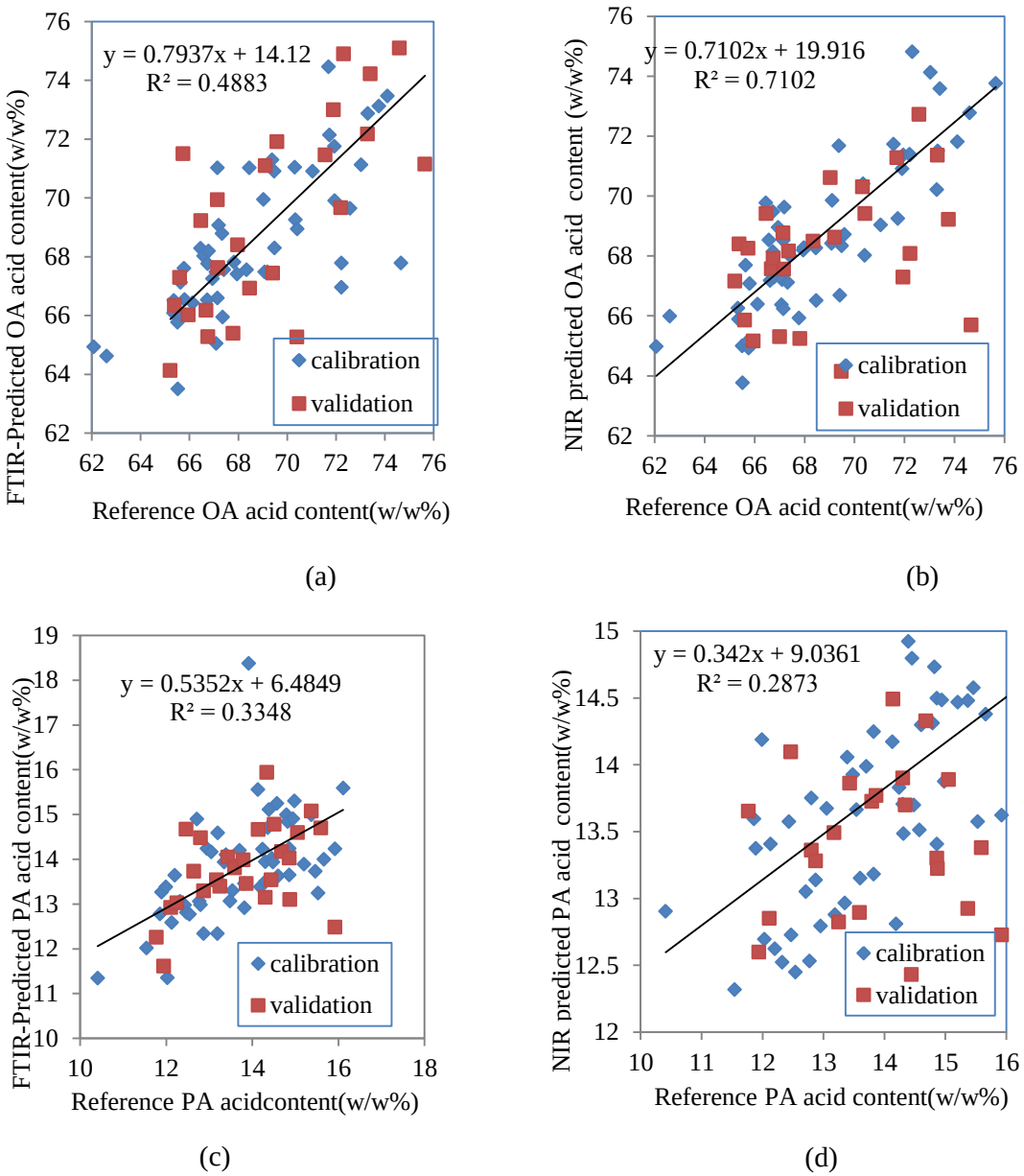


Figure 4.18. (a) Reference Oleic acid (OA) content vs. predicted values based on FTIR-ATR spectra using PLS method (b) Reference Oleic acid (OA) content vs. predicted values based on NIR spectra using PLS method (c) Reference Palmitic acid (PA) content vs. predicted values based on FTIR-ATR spectra using PLS method (d) Reference Palmitic acid (PA) content vs. predicted values based on NIR spectra using PLS method

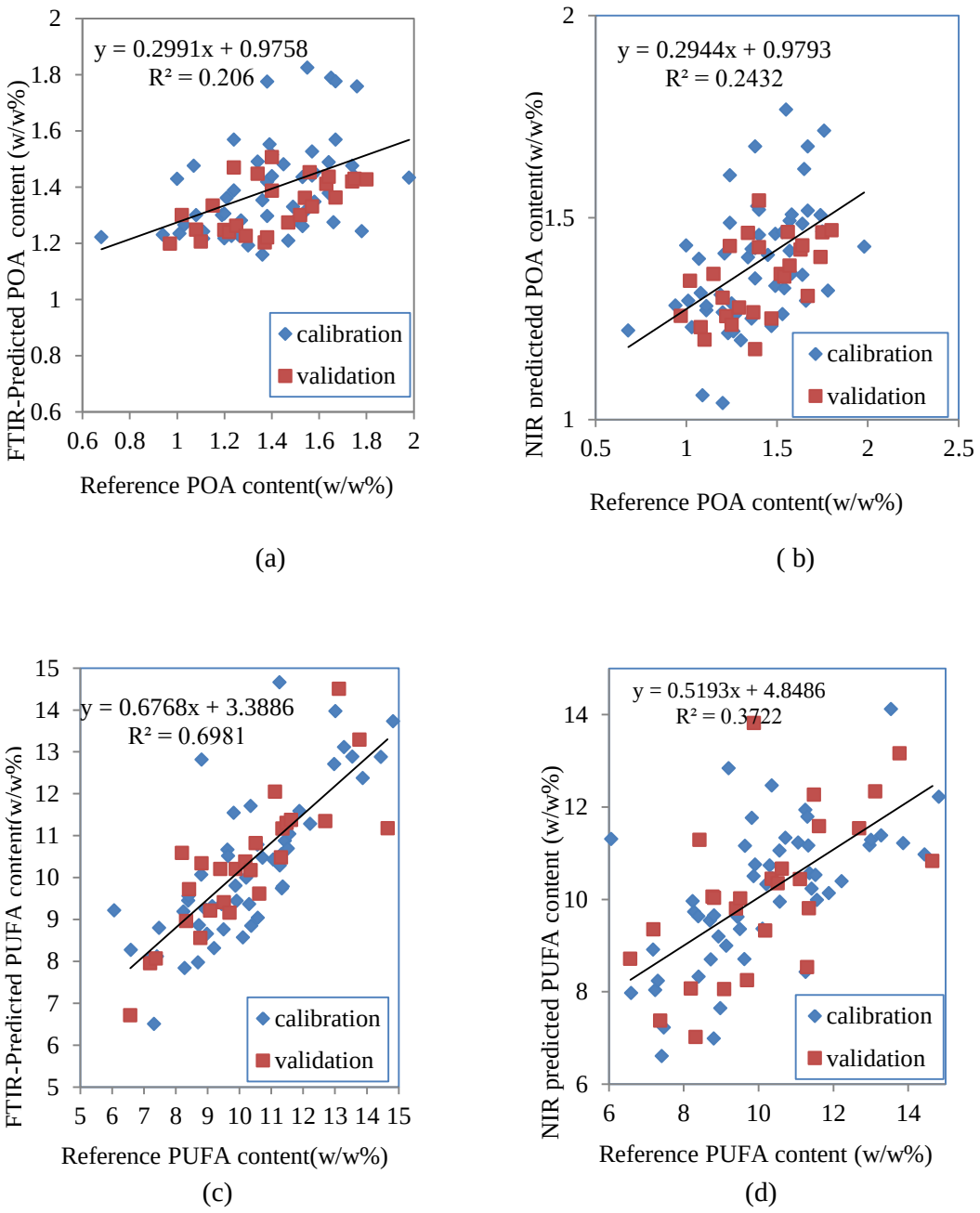
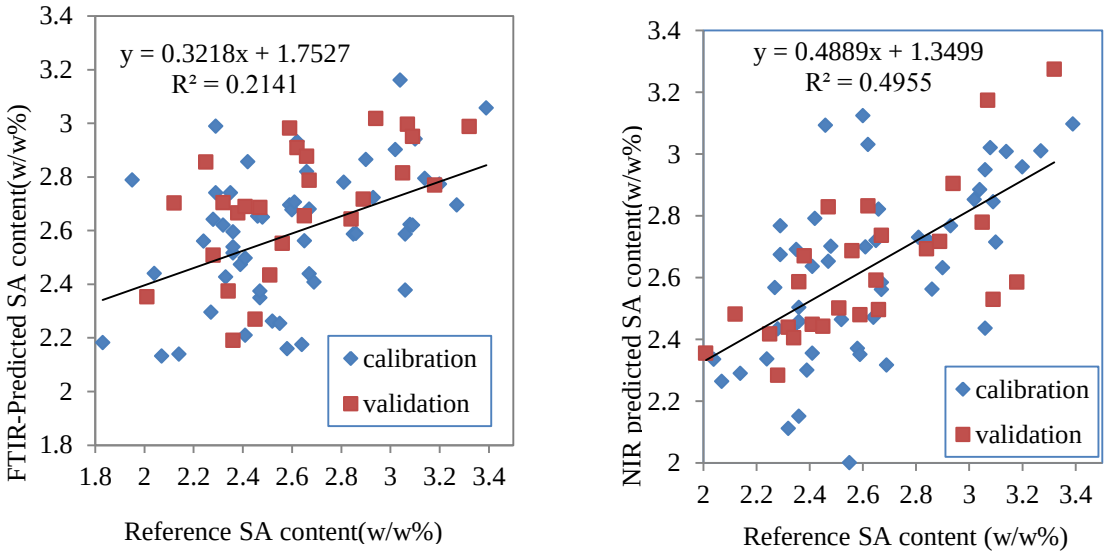


Figure 4.19. (a) Reference Palmitoleic acid (POA) content vs. predicted values based on FTIR-ATR spectra using PLS method (b) Reference Palmitoleic acid (POA) content vs. predicted values based on NIR spectra using PLS method (d) Reference Polyunsaturated fatty acid (PUFA)content vs. predicted values based on FTIR-ATR spectra using PLS method (d) Reference Polyunsaturated fatty acid (PUFA) content vs. predicted values based on NIR spectra using PLS method



(a) (b)

Figure 4.20. t(a) Reference Stearic acid (SA) content vs. predicted values based on FTIR-ATR spectra using PLS method (b) Reference Stearic acid (SA) content vs. predicted values based on NIR spectra using PLS method

According to Figure 4.21. to Figure 4.22 , the models for TAG components have low prediction ability of concentrations with low regression coefficients. And also the models are suffering from high collinearity and cannot be trusted for producing a high calibration quality and for making prediction of concentrations in the validation samples.

Generally, as can be seen Figure 4.21., and Figure 4.22 reference versus predicted values are not compatible with each other. When overall calibration performance models examine, it is possible to state that PLS is not best calibration method; since all FAME components have worst regression coefficient when comparing GILS method

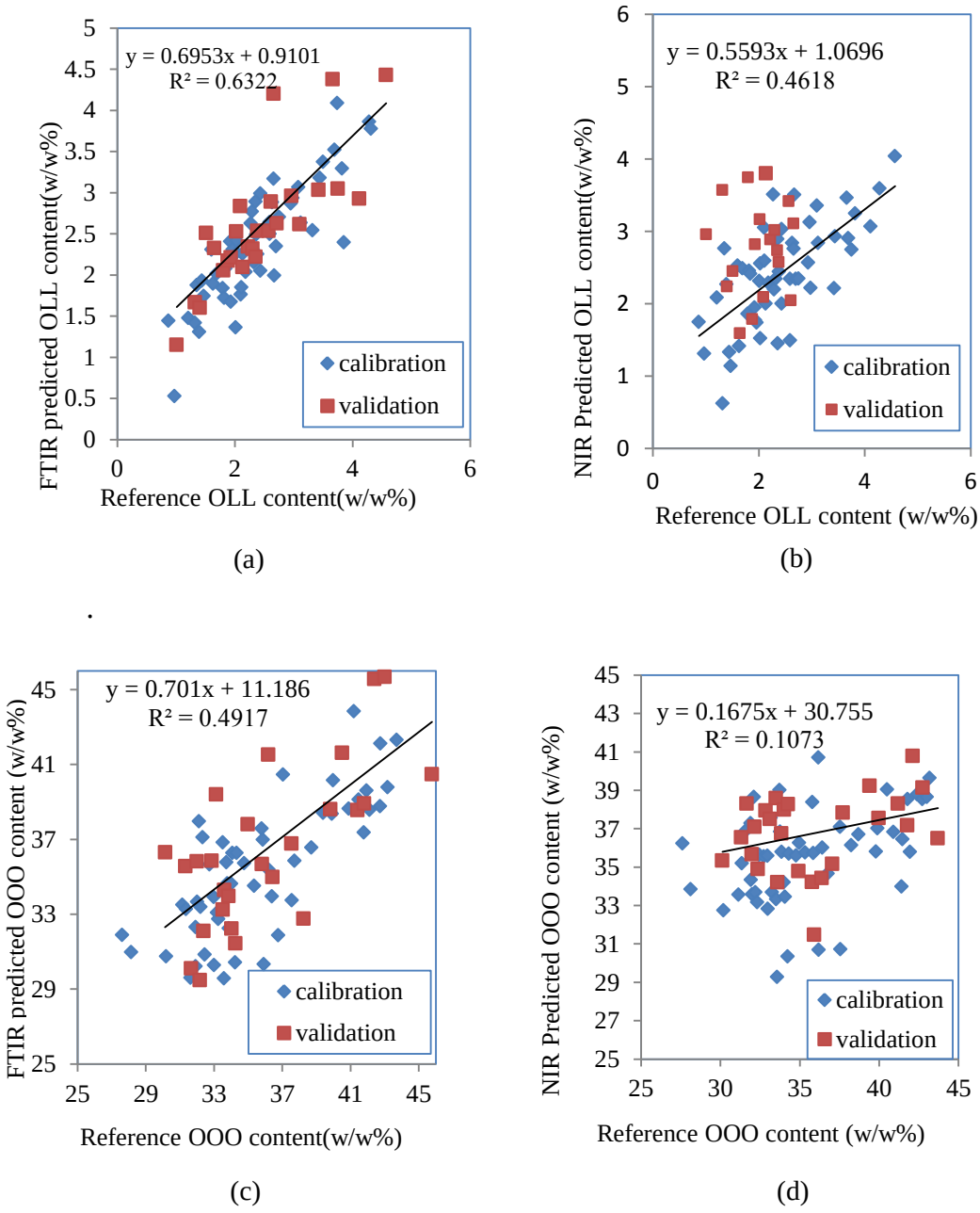
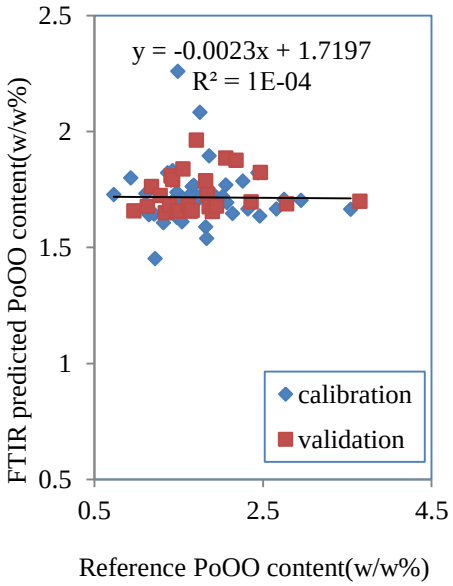
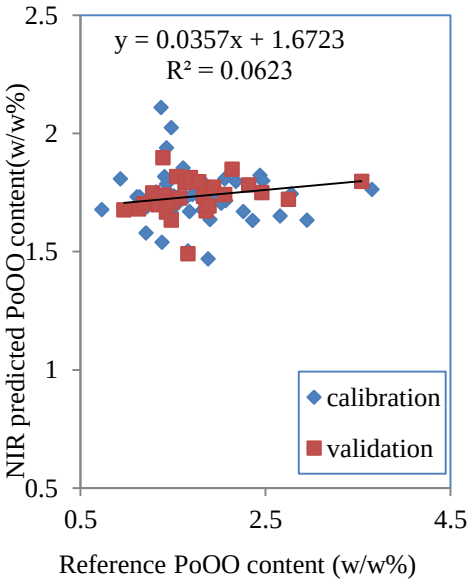


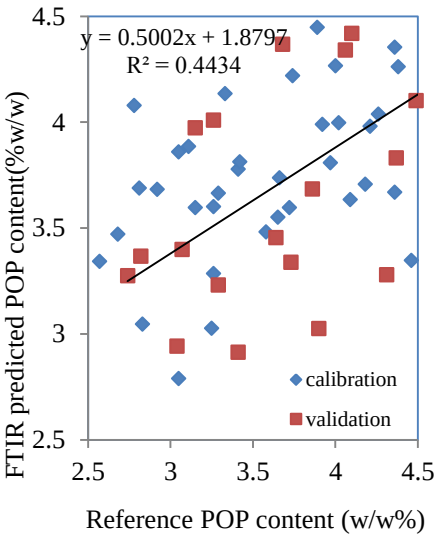
Figure 4.21. (a) Reference 1,2-dilinoleyl-3-oleylglycerol (OLL) content vs. predicted values based on FTIR-ATR spectra using PLS method (b) Reference 1,2-dilinoleyl-3-oleylglycerol (OLL) content vs. predicted values based on NIR spectra using PLS method (c) Reference triolein (OOO) content vs.predicted values based on FTIR-ATR spectra using PLS method (d) Reference triolein (OOO) content vs. predicted values based on NIR spectra using PLS method


y = -0.0023x + 1.7197
R² = 1E-04
FTIR predicted PoOO content(w/w%)
Reference PoOO content(w/w%)
calibration
validation
2.5
2
1.5
1
0.5
0.5
2.5
4.5

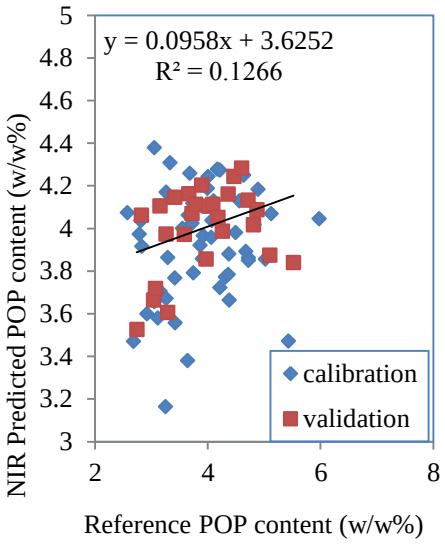
(a)


y = 0.0357x + 1.6723
R² = 0.0623
NIR predicted PoOO content(w/w%)
Reference PoOO content (w/w%)
calibration
validation
2.5
2
1.5
1
0.5
0.5
2.5
4.5

(b)


y = 0.5002x + 1.8797
R² = 0.4434
FTIR predicted POP content(%w/w)
Reference POP content (w/w%)
calibration
validation
4.5
4
3.5
3
2.5
2.5
3
3.5
4
4.5

(c)


y = 0.0958x + 3.6252
R² = 0.1266
NIR Predicted POP content (w/w%)
Reference POP content (w/w%)
calibration
validation
5
4.8
4.6
4.4
4.2
4
3.8
3.6
3.4
3.2
3
2
4
6
8

(d)

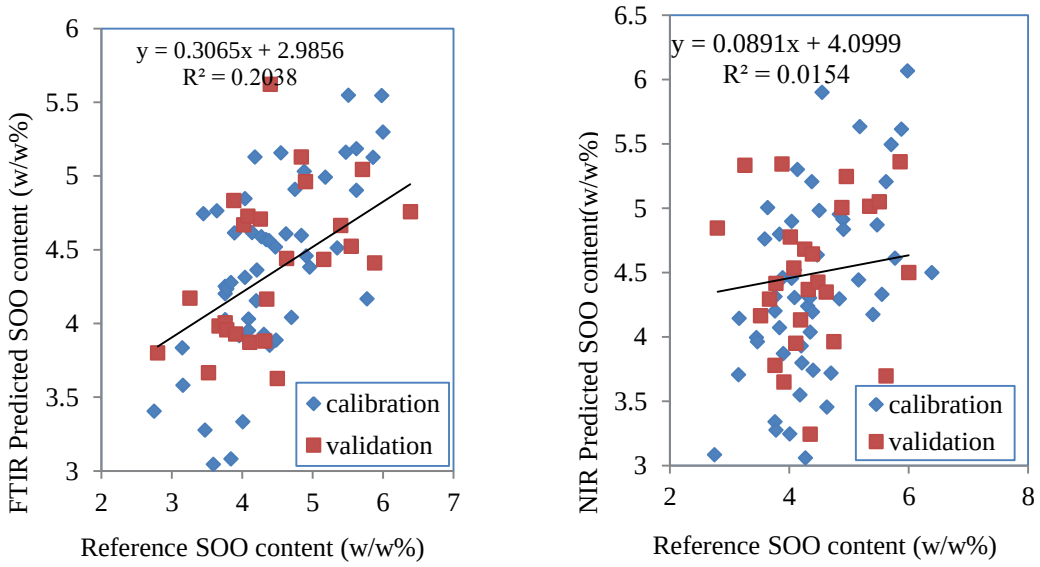


Figure 4.22. (a) Reference 1-palmitioleoyl,2,3-dioleoyl (PoOO) content vs. predicted values based on FTIR-ATR spectra using PLS method (b) Reference 1-palmitioleoyl,2,3-dioleoyl (PoOO) content vs. predicted values based on FT-NIR spectra using PLS method (c) Reference 1,2-dipalmitoyl-3-oleylglycerol (POP) content vs. predicted values based on FTIR-ATR spectra using PLS method (d) Reference 1,2- dipalmitoyl-3-oleylglycerol (POP) content vs. predicted values based on FT-NIRspectra using PLS method(e) Reference 2,3-dioleyl-1-stearoylglycerol (SOO) content vs. predicted values based on FTIR-ATR spectra using PLSmethod (f) Reference 2,3- dioleyl-1-stearoylglycerol (SOO) content vs. predicted values based on FT-NIR spectra using PLS method

CHAPTER 5

CONCLUSION

In this thesis, it is aimed to develop molecular spectroscopic multivariate calibration models for the determination some of fatty acid methyl esters and triachyl glycerol compositions of Gemlik type olive oils provided by Olive oil Research Institute. The data obtained from High Performance Liquid Spectroscopy (HPLC) was chosen reference method for tri-achyl glycerol compositions and also the data obtained from Gas Chromatography (GC) was chosen reference method for fatty acid methyl esters. Analyzes were done by Olive Oil Research Institute. Samples were also measured by spectroscopic techniques. These spectroscopic techniques are Fourier Transform Infra red spectroscopy (FTIR) and Near Infra Red spectroscopy (NIR). Data were obtained by both two instruments for TAG and FAME compositions. GILS (Genetic inverse least square) chemometric method which is multivariate calibration model based on genetic algorithm was used to construct calibration models for both compositions. Reliability of the calibration models was determined by SEC and SEP values as well as with the R^2 values from the reference vs. predicted content plots.

In conclusion, the models based on FTIR instrument data for both FAME compositions and TAG compositions gave successful results by using GILS method. From the results, it is seen that successful calibration models can be constructed by using the methods mentioned to provide fast and non-destructive determination of FAME and TAG compositions. This might give rise to improvements in the olive oil industry in economical manner. In addition, by wavelength selection feature of GILS method, the wavelengths which carry information of FAME and TAG could be determined in order to develop case specific analysis models. NIR and FTIR combined with multivariate calibration models could be more advantageous compare to chromatographic methods because of their simplicity and speed.

REFERENCES

B.H. Jennings, C.C. Akoh, "Enzymatic modification of triacylglycerols of high seicosapentaenoic and docosahexaenoic acids content to produce structured lipids," *Journal of the American Oil Chemists'* Society, vol. 7, no. 10, pp. 1133-1137, **1999**.

Beebe K.R., Pell. R.J., and Seasholtz M.B. **1998**. *Chemometrics, a practical guide.*Wiley-Interscience: John Wiley & Sons, Inc.

Bendini, A.; Cerretani, L.; Di Virgilio, F.; Belloni, P.; Bonoli-Carbognin, M.; Lercker, G. 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*. **2007**, *30*, 424-437.

Bewig, K.M., A.D. Clarke, C. Roberts, and N. Unklesbay, Discriminant Analysis of Vegetable Oils by Near-Infrared Reflectance Spectroscopy, *J. Am. Oil Chem. Soc. 71*:195–200(**1994**).

Brereton R.G. **2000.**Introduction to multivariate calibration in analytical chemistry.*The Analyst* 125:2125-2154.

Christy, A. A., Kasemsumran, S., Du, Y., Ozaki, Y. **2004** The detection and Quantification of Adulteration in Olive Oil by Near-Infrared Spectroscopy and Chemometrics. *Analytical Sciences 20:935-940.*

Cinquanta, L.; Esti, M.; La Notte, E. Evoulation of phenolic compounds in virgin oliive oil during storage. *Journal of American Oil Chemists' Society*. **1997**, *74*, 1259-1264.

Cong, P. and Li, T. **1994.** Numeric genetic algorithm part I. theory, algorithm and simulated experiments.*AnalyticaChimicaActa*293:191-203.

Çelik, D. **2013.** Development of Chromatographic and Molecular Spectroscopic Multivariate Chemometric Models for the Geographical Classification of Olive Oils. *İzmir Institute of Technology of Ms Thesis.*

Dıraman H., Dibeklioglu H., Characterization of Turkish Virgin Olive Oils Produced from Early Harvest Olives, J Am Oil Chem Soc., 86, 663–674, (**2009**).

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

Firestone, D., **2005**.Olive Oil.Bailey's Industrial Oil and Fat Products(edited by F.Shahidi), Sixth edition,John Wiley Inc.303-331.

Fontain, E. **1992**.The problem of atom-to-atom mapping.An application of geneticalgorithms.*AnalyticaChimicaActa*265:227-232.

Geladi, P.; Kowalski, B. R.; "Partial Least Squares Regression"; Anal. Chim. Acta; 185; **1986**; p. 1.

Gilbert R. J., Goodacre, R., Woward, A. N., and Kell, D.B. **1997**. Genetic programming: a novel method for the quantitative analysis of pyrolysis mass spectral data. *Analytical Chemistry 69*:4381-4389.

Guillén, M.D., Cabo, N. **1997**.Characterization of Edible Oils and Lard by Fourier Transform Infrared Spectroscopy. Relationships between Composition and Frequency of Concrete Bands in the Fingerprint Region.*Journal of the American Oil Chemists' Society* 74(10):1281-1286.

Güler, M., Cesur, R., Sarı, N. Zeytinde Bakım İşlemleri.38 s. Adana, (**2010**).

Haaland, D. M.; Melgaard, D. K.; "New augmented classical least squares methods for improved quantitative spectral analysis"; Vibrational Spectroscopy; Vol 29; **2002**; p. 171.

Hartnett, M. K.; Diamond, D.; Anal. Chem.; Vol 69; 1997; 1909.

Harwood, J.L., Yaqoop, P., **2002**. Nutritional and health aspects of olive oil.Eur.J.Lipid Sci.Technol.104 ,685 – 697.

Hibbert, D.B. **1993**. Genetic algorithms in chemistry.*Chemometrics and Intelligent Laboratory Systems* 19:277-293.

Howard, M., Campbell,B., **2008**. An introduction to Near Infrared spectroscopy and associated chemometrics.

http://www.boomer.org (accessed May 15,**2013**).

http://www.unctad.org/infocomm/anglais/olive/market.html# (accessed May 12, **2011**).

http://www.oliveoillife.com/en/market/html/76.html (accessed May 22, **2011**).

Karaman, İ. **2008**. Prediction of extractives and lignin contents of Anatolian black pine (pinus nigra arnold. Varpallasiana) and Turkish pine (pinus brutia ten.) Trees using infrared spectroscopy and multivariate calibration. *İzmir Institute of Technology of Ms Thesis.*

Kiritsakis, A. K. Flavour components of olive oil- A review. *Journal of the American Oil Chemists' Society*. **1998**, *75*(6), 673-681.

Lai, Y.W., E.K. Kemsley, and R.H. Wilson, Quantitative Analysis of Potential Adulterants of Extra Virgin Olive Oil Using Infrared Spectroscopy, *Food Chem. 53*:95–98 (**1995**).

Lucasius, C.B. and Kateman, G. **1991**. Genetic algorithms for large-scale optimization in chemometrics: an application. *Trends in Analytical Chemistry* 10:254-261.

Lucasius, C.B., Beckers, M.L.M., and Kateman, G. **1994**. Genetic algorithms in wavelength selection: a comparative study. *AnalyticaChimicaActa*286:135-153.

Lucasius and 19 Kateman **1992**, Li, T., Lucasius, C.B., and Kateman, G. 1992.Optimization of calibration data with the dynamic genetic algorithm. *AnalyticaChimicaActa*268:123-134.

Malinowski, E. R.; "Theory of error in Factor Analysis"; Anal. Chem.; Vol 4; 49;**1977**; p. 606.

Mark, H.; Anal. Chem.; Vol 58; **1986**; p.2816.

Ozdemir, D., Mosley, R.M., and Williams, R.R. **1998**a. Hybrid calibration models analternative to calibration transfer. *Applied Spectroscopy* 52:599-603(5).

Ozdemir, D., Mosley, R.M., and Williams, R.R. **1998**b. Effect of wavelength drift onsingle- and multi-instrument calibration.

Ozdemir, D. and Williams, R.R. **1999**. Multi-instrument calibration with geneticregression in UV-visible spectroscopy.*Applied Spectroscopy* 53:210-217(8).

Öztürk, B. **2003**.Monitoring the esterification reactions of carboxylic acids with alcohols using near infrared spectroscopy and multivariate calibration methods.*Izmir Institute of Technology thesis of M.Sc.*

Paradkar, R.P. and Williams, R.R. **1997**. Correcting fluctuating baselines and spectral overlap with genetic regression.*Applied Spectroscopy* 51:92-100(9).

Skoog, D. A.; Holler, F. J.; Nieman, T. A. Principles of instrumenatl analysis. *Saunders College Publishing, Harcourt Brace College Publishers*. **1998**,357-383.

Tsimidou, M., and K.X. Karakostas, Geographical Classification of Greek Virgin Olive Oil by Non-Parametric Multivariate Evaluation of Fatty Acids Composition. *J. Sci. Food Agric. 62*:253–257 (**1993**).

UNCTAD based on data from the Report on the proceedings of the 86th session of the International Oil Council June **2002.**

Wang, Y., Veltkamp, D. J., and Kowalski, B. R. **1991**.Multivariate instrument standardization.*Analytical Chemistry* 63:2750-2756.

Wienke, D., Lucasius, C. B., Ehrlich, M., and Kateman, G. **1993**.Multicriteria target vector optimization of analytical procedures using a genetic algorithm: part II. polyoptimization of the phothometric calibration graph of dry glucose sensors for quantitative clinical analysis. *AnalyticaChimicaActa*271:253-268.

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

Wold, H.; Multivariate Analysis; John Wiley & Sons Publications NY; **1966**; p. 391.