DEVELOPMENT OF RAPID AND SIMPLE SPECTROSCOPIC TECHNIQUES BASED ON CHEMOMETRICS DATA ANALYSIS FOR THE DETERMINATION OF GOAT MILK ADULTERATION WITH COW MILK

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
Hülya SAMANCIOĞLU

December 2015
İZMİR

We approve the thesis of **Hülya SAMANCIOĞLU**

Examining Committee Members:

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

Doç. Dr. Hasan ERTAŞ
Department of Chemistry, Ege University

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

29 December 2015

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

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

Prof. Dr. Bilge KARAÇALI
Dean of the Graduate School of
Engineering and Sciences

ACKNOWLEDGEMENTS

I would like thank several people who really accompanied and supported me during my thesis studies.

Firstly, I would like to express my deepest gratitude to my advisor Prof. Dr. Durmuş ÖZDEMİR for his continuous support, patience, motivation, and immense knowledge throughout my study and research. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my master study. In addition, I express my appreciation to Prof. Dr. Figen TOKATLI and Assoc. Prof. Dr. Hasan ERTAŞ for having served on my committee. Their thoughtful questions and comments were valued greatly.

I would also like to thank the members of Chemometrics Research Group for helping me during my graduate years at İYTE.

My sincere thanks and deep appreciation also go to Kemal Volkan ÖZDOKUR, who helped me writing of this thesis. Also I owe special thanks to my best friends Duygu EREN AŞIK and Fulden YILMAZ for their all assistance, motivation and encouragement, through my whole life.

This study would not be possible without the support of The Scientific and Technical Research Council of Turkey (TUBİTAK-Hızlı Destek Project number 114Z076).

Finally, I would like to express my special thanks to my parents and my brother for their unconditional love and support; I would not have been able to complete this thesis without their continuous love and encouragement.

ABSTRACT

DEVELOPMENT OF RAPID AND SIMPLE SPECTROSCOPIC TECHNIQUES BASED ON CHEMOMETRICS DATA ANALYSIS FOR THE DETERMINATION OF GOAT MILK ADULTERATION WITH COW MILK

Milk and milk products are one of the most consumed foods. However, there has been an increase in milk adulteration, especially goat milk. Current methods for detection of milk adulteration are expensive, impractical and are not sufficient to answer quantitatively. Therefore, to meet this demand, a rapid, easy to use and inexpensive method was developed in this study.

The proposed methodology is based on Fourier Transform Infrared Spectroscopy (FTIR) with the combination of multivariate calibration techniques for determination of adulteration of goat milk with cow milk. During the study, 150 raw goat milk samples were collected from different goats in 3 different seasons (June-2014, December 2014, March2015). Then, adulterated samples were prepared with goat milk, cow milk and water. Both adulterated and raw goat milk samples were analyzed with FTIR. Afterwards, 7 different multivariate calibration models (for each season (3), binary combination of seasons(3) and a ternary combination of all the seasons) were generated with synthetically prepared samples by using Genetic Inverse Least Squares (GILS) and Partial Least Squares (PLS) methods. These models were used to predict both adulterated and raw goat milk samples in order to evaluate the success of the models. Standard Error of Prediction (SEP) values for goat milk, cow milk and water are 7.9, 6.3 and 4.9 (w/w %), respectively, indicate satisfactory predictions by GILS. On the other hand, PLS models gave SEP values ranging between 6.4 and 12.9 (w/w %).

ÖZET

KEÇİ SÜTÜNÜN İNEK SÜTÜ İLE TAĞŞİŞİNİN BELİRLENMESİ İÇİN KEMOMETRİK VERİ ANALİZ YAKLAŞIMLARINA DAYALI BASİT VE HIZLI SPEKTROSKOPİK YÖNTEMLER GELİŞTİRİLMESİ

Süt ve süt ürünleri dünyada en çok tüketilen gıdalar arasındadır. Fakat özellikle keçi sütünde olmak üzere, süt tağşişi gündemdeki popüler konulardan biri haline gelmiştir. Hali hazırda var olan süt tağşişini belirlemeye yönelik yöntemlerin pahalı ve zaman alıcı olmasının yanı sıra, miktar belirtmede yetersiz kaldığı görülmektedir. Bu problem ve eksiklerden yola çıkılarak, keçi sütünün inek sütü ile tağşişinin çok değişkenli kalibrasyon yöntemlerine dayalı Fourier Dönüşümlü İnfrared Spektroskopisi (FTIR) ile belirlenmesine yönelik hızlı, pratik ve masraf gerektirmeyen metot geliştirilmiştir.

Çalışma boyunca her sezondan (Haziran-2014, Aralık 2014, Mart-2015) 50 adet olmak üzere toplamda 150 adet saf keçi sütü örneği toplanmıştır. Saf keçi sütü, inek sütü ve su kullanılarak tağşişli örnekler hazırlanmıştır. Hem hazırlanan örneklerin, hem de saf keçi sütlerinin FTIR spektrumları alınmıştır. Daha sonra bu karışım örnekleri kullanılarak, Genetik Ters En Küçük Kareler (GILS) ve Kısmi En Küçük Kareler (PLS) metotları ile 7 farklı (her sezonun (3), sezonların ikili kombinasyonlarının (3) ve sezonların üçlü kombinasyonu (1)) çok değişkenli kalibrasyon modelleri kurulmuştur. Modellerin başarısının ölçülmesi amacı ile hem tağşişli örnekler, hem de saf keçi sütleri bu modeller kullanılarak tahmin ettirilmiştir. GILS ile kurulan nihai üçlü kombinasyon modeli ile yapılan tahmin sonuçlarında keçi sütü, inek sütü ve su için elde edilen Standart Tahmin Hataları (SEP) sırası ile 7.9, 6.3 ve 4.9 (w/w%) olarak bulunmuştur. PLS metodu ile bulunan SEP değerlerinin ise 6.4 ile 12.9 (w/w%) arasında değiştiği gözlemlenmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

LIST OF ABBREVIATIONS

FTIR	Fourier transform infrared spectroscopy
ATR	attenuated total reflectance
GILS	genetic inverse least squares
PLS	partial least squares
ILS	inverse least squares
CLS	classical least squares
SECV	standard error of cross validation
SEPv	standard error of prediction for validation set
SEP_{avr}	average standard error of prediction for second independent validation (raw goat milk samples) set
SEP_J	standard error of prediction for June 2014
SEP_D	standard error of prediction for December 2014
SEP_M	standard error of prediction for March 2015
PC	principle component
PCA	principle component analysis

CHAPTER 1

INTRODUCTION

1.1. Motivation

Milk has a very high nutritional value among the foodstuffs. Goat milk which is similar with breast milk in terms of nutritional value is preferred much, owing to the possibility of cow milk allergy in people, in particular newborn babies. Therefore, the reliability of the goat milk is very vital. However, the financial value of goat milk is much more than cow milk on the market. Accordingly, this causes goat milk adulteration with cow milk. In our country, Ministry of Food, Agriculture and Livestock, does not put any restrictions on adulterated milk and dairy products. This situation may result in an incorrect product labeling, so consumers exposed to consume adulterated product and this situation may lead to undesirable consequences for the people who are allergic to cow milk. Therefore, the development of analytical methods for the determination of a precise adulteration ratio in mixed dairy products both for the protection of consumers and honest producers have great importance.

Already some biochemical analysis methods [e.g. PCR (Polymerase Chain Reaction - PCR)], can be easily detect goat milk and cow milk. However, these techniques do not have the high precision on quantitative analysis. For these reasons, the aim of this project primarily based on the development of rapid and practical molecular spectroscopic technique based on multivariate chemometric data analysis methods for identification of adulteration of goat milk with cow milk, quantitatively.

1.2. Structure and Scope of the Thesis

The first chapter of this thesis study gives the purpose of the research and a brief overview of literature review. Chapter 2 includes detailed information about Fourier Transform Infrared Spectroscopy (FTIR) used in the study. In chapter 3, comprehensive description of multivariate calibration techniques are given. Experimental and

instrumental part is covered chapter 4 whereas in chapter 5 results are discussed. Lastly, in chapter 6, concluding remarks are given.

1.3. Literature Review

In recent years, there have been a lot of studies about determining goat milk adulteration with the cow milk in Europe and many other countries. Although most studies are just for determination of adulteration, recent developments in multivariate calibration methods on food stuffs turned it into a very useful tool which can be used for the quantitative analysis of adulteration. There are different techniques to determine milk adulteration such as diffuse reflectance infrared fourier transform spectroscopy (DRIFTS), polymerase chain reaction (PCR), electronic tongue taste, gas chromatography-mass spectrometry (GC-MS), infrared microspectroscopy, FTIR, high performance liquid chromatograph (HPLC), matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS), capillary electrophoresis. Some of these studies received support from chemometric data analysis as multivariate analysis, partial least squares (PLS), discriminant analysis, principle component analysis (PCA), Soft Independent Modeling of Class Analogy (SIMCA), Statistical Package for the Social Sciences (SPSS). Such techniques compared with chemometric methods give more precise and accurate results, according to these studies.

Classical and traditional techniques are generally response that whether there is an adulteration or not for milk samples. PCR is the most commonly used method for this purpose. Rodrigues et al. 2012 have studied about to investigate the adulteration of goat milk with bovine milk by smallholders in Brazil. They developed and standardized a duplex PCR assay. The detection limit of the method was found 0.5% bovine milk in goat milk. After developing the method, they tested the commercial goat milk in the market they found bovine milk in 41.2% of sample. Then they sold the method to dairy plants to reduce adulteration in goat milk. Another method for milk adulteration with PCR was performed by López-Calleja et al. 2005. They detected goat milk in sheep milk using spesific RNA gene. This goat spesific gene was not found in sheep, cow or buffalo milk. So, they prepared binary mixtures of different milk species and detected goat milk in them with a sensitivity treshold of 0.1%. As a result, they decided that the proposed PCR method is not useful to detect adulteration of milk products. Golinelli et al. 2014 have used PCR method to detect bovine milk in goat cheese. They used RNA

gene of both goat and bovine and PCR method found useful for detecting et least 0.5% bovine milk added goat cheese formulation. Another study was performed by Scano et al. 2014. They investigated the discrimination of caprine milk from goat milk with GC-MS technique and multivariate statistical data analysis method. The discrimination was performed by the polar metabolite profiles of goat and cow milk samples. Valine and glycine metabolites are specific for goat milk, talose and malic acid metabolites are specific for cow milk. With the aid of these metabolites, they developed the method and determined cow milk in goat milk with the error of 5%. The method was demonstrated that GC-MS and multivariate data analysis are useful for discriminate goat and cow milk, so they proposed an easy and new analytical method to discover and protect milk fraud. There is an interesting study by Dias et al. 2009 in literature. They built an electronic tongue with 36 cross-sensibility sensors to detect adulteration. The method showed the 97 % sensibility and 93 % specificity. Furthermore, they set a model with the aid of Principal Component Analysis (PCA), and the model showed that the unknown milk samples were classified correctly with a sensibility and specificity of 87 % and 70 %, respectively. Spectroscopic techniques based on cluster analysis also uses for identification or differentiation. Pappas et al. 2008 have studied on this topic. They used DRIFTS and analyzed data with Statistical Package for the Social Sciences (SPSS) program to discriminate goat and sheep milk from each other. They implified that the proposed method is fast and accurate for identification of defatted milk species with cluster analysis. They obtained satisfactory result that the samples of goat milk can be differentiated from sheep milk samples.

On the other hand, determination of adulteration qualitatively can be mostly inadequate. Consequently, in order to determine amount of the adulteration, some analytical method is needed to combined with multivariate calibration methods. Santos, Pereira-Filho, and Rodriguez-Saona 2013 have studied about on this subject. They used attenuated total reflectance (ATR) mid-infrared (MIR) microspectroscopy and chemometric methods (PLS Regression and Pattern recognition analysis by Soft Independent Modeling of Class Analogy (SIMCA)) to detect and quantify the milk adulteration. The adulteration were done with urea, whey, synthetic urine, hydrogen peroxide and synthetic milk. PLS gave satisfactory results according to low standard error of prediction (SEP) values. Results showed that (MIR) microspectroscopy can provide an alternative methodology to the dairy industry to screen potential fraudulent for economic adulteration of cow milk. Another study with FTIR and PLS performed by

Nicolaou, Xu, and Goodacre 2010. They used 3 different milk species, sheep-cow-goat, and quantified the different levels of adulterated samples for use in routine milk analysis by dairy industry. They prepared 4 type milk combinations that are 3 binary mixture and a ternary mixture samples. Mixtures were analyzed with FTIR, then multivarite statistical analysis was performed on the data. Results proved that the method has excellent potential for use in the dairy industry as a rapid technique of identification and quantification with the low error levels of 3.4% to 4.9%. In addition to these techniques, chromatographic techniques may also be used for quantitative analysis. Rodriguez et al. 2010 performed a study on detection and quantification milk adulteration to prevent possible fraud and to give protection to companies and consumers. In this study, data were obtained from HPLC with diode-array detector of cheese and milk extracts and data was studied through PLS calibration model. Thus, they determined the amounts of each milk species in the samples. In an another recent work, Nicolaou, Xu, and Goodacre 2011 prepared different amounts of samples from goat, sheep and cow milk and used them for determining qualitatively and quantitatively by MALDI-MS and PLS methods. When they looked at the results, they decided that MALDI-MS and PLS were quite good at measuring adulteration in milk samples. Cartoni et al. 1999 demonstrated that capillary zone electrophoresis can be used to determine cow milk in goat milk products by the aid of specific whey proteins. They found the minimum detectable of cow milk is 2% in milk mixtures and 4% in cheese. The developed method was expected to widespread routine analyses. The milk adulteration is not only done with goat, cow or sheep milk, but also can be done with soy milk. Jaiswal et al. 2015 were investigated to detect soy milk in cow-buffalo milk with FTIR-ATR. PCA method was performed on data and study showed that the proposed method is a useful tool to detect soymilk easily.

In order to identify the fraud and adulteration of the goat milk, some biological and biochemical methods have been developed but these methods offer only qualitative information. This also explains the insufficiency of these methods. The proposed methodology which is based on FTIR with the combination of multivariate calibration techniques for determination of adulteration of goat milk with cow milk, reveals the original value of the project. The improvements to pure ILS provided by genetic algorithm have made this quantitative analysis possible and more reliable. Therefore, the results of the project will contribute to the literature.

CHAPTER 2

FOURIER TRANSFORM INFRARED SPECTROSCOPY

2.1. Introduction to Infrared Spectroscopy

Spectroscopy is the science that deals with the interactions between electromagnetic radiations and matter. These interactions can be transitions, emission or adsorption of electromagnetic radiation. Spectroscopic techniques are divided into two:

- Atomic Spectroscopy
- Molecular Spectroscopy

While atomic spectroscopy based on the measurements of free atomic species in the vapor state, molecular spectroscopy based on molecular species in solution, in the vapor or solid state (Ingle and Crouch 1988). Infrared (IR) spectroscopy is one of the molecular spectroscopic technique which is used for identification of the compounds. The technique deals with the molecules which absorb specific frequencies of infrared radiation those are characteristic of their structure. Infrared spectroscopy is a simple technique which has lots of application field and is commonly used in both for inorganic and organic compounds (Setle 1997).

The region of the infrared spectrum ranges between the wavenumber from 12.800 to 10 cm^{-1}, and between the wavelengths from 780 to $1x10^6$ nm. Infrared Spectrum is usually subdivided into three regions because of the requirements that differ from applications and the instruments. The wavelength ranges of these three infrared regions are shown in Table 2.1. (Skoog, Holler, and Crouch 1998)

Table 2.1. The corresponding wavelengths and wavenumbers of the Infrared regions.

Regions	Wavelength Range (nm)	Wavenumber range (cm^{-1})
Near–Infrared (NIR)	780 - 2500	12.800 - 4000
Mid–Infrared (MIR)	2500 - 50.000	4000 - 200
Far–Infrared (FIR)	50.000 - 1×10^6	200 - 10

For qualitative and quantitative analysis, mid–infrared (MIR) is used, in general. Functional groups of the molecules have specific information in MIR region. These groups are named ‘fingerprints’. Both organic and inorganic molecules can be analyzed in the far–infrared (FIR) region according to their metallic bands. In addition, near infrared (NIR) region is usually used for the quantitative analysis of complex samples with the aid of chemometrics methods. Molecules in the sample are analyzed with an infrared spectroscopy with these following steps: light from a light source strikes to the sample and then the radiation of the energy is absorbed by molecules in the sample in order to be excited to the vibrational or rotational states. The molecular interactions and their corresponding infrared regions are shown in the Table 2.2. (Workman 1996).

Table 2.2. The molecular interactions and their corresponding infrared regions.

Regions	Characteristics Measured
Near–Infrared (NIR)	Overtone & combination bands of fundamental molecular vibrations
Mid–Infrared (MIR)	Fundamental molecular vibrations
Far–Infrared (FIR)	Molecular rotations

2.2. Fourier Transform Infrared Spectroscopy (FTIR)

Fourier Transform Infrared (FTIR) systems were developed with the invention of Michelson interferometer by Albert Abraham Michelson. This interferometer is the heart of most FTIR instruments. The interferometer takes a beam of light, then separates it into two beams using a beam splitter, and makes them travel different distances. This is called *optical path difference* (OPD). The Michelson interferometer is composed of four members. The first one is a source of infrared light, the second one is a stationary mirror, the third member is a movable mirror and the last one is a slit. These two mirrors are mounted perpendicular to each other where there is a beamsplitter, which reflects half of the light and transmits half of it, to these mirrors. In consequence, reflected light strikes movable mirror and transmitted light strikes fixed mirror. These two beams recombine after reflecting from mirrors, then goes out of interferometer for the interaction with sample. (Smith 1996). A Michelson interferometer diagram is shown in Figure 2.1.

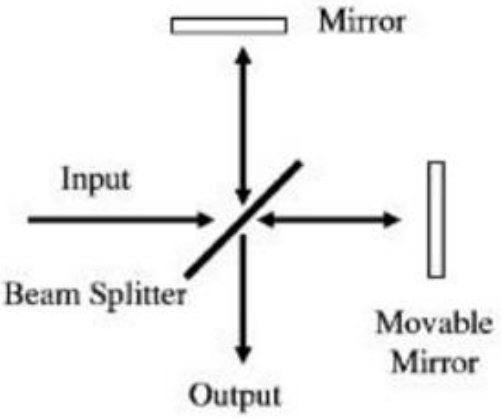

Figure 2.1. A Michelson interferometer diagram.
(Source: Cheng 2005)

When the movable mirror is moved at constant speed, infrared radiation intensity fluctuates as a result of constructive and destructive interferences. Detector measures that fluctuation of light intensity as a result of the optical path difference (OPD) and the plot between light intensity and OPD forms interferogram. When the movable mirror moves back and forth once, it is called as *scan* and a complete interferogram is generated which is basically composed of a number of sinusoidal signals. If an infrared spectrum considered as a mathematical function and an interferogram is sum of sinusoidal waves, a Fourier Transform (FT) procedure is used to calculate the frequency domain spectrum from the interferogram easily. In a sense, FT inverts the interferogram which is a time (distance) domain signal to an infrared spectrum, as shown in Figure 2.2. (Smith 1996) (Skoog, Holler, and Crouch 1998).

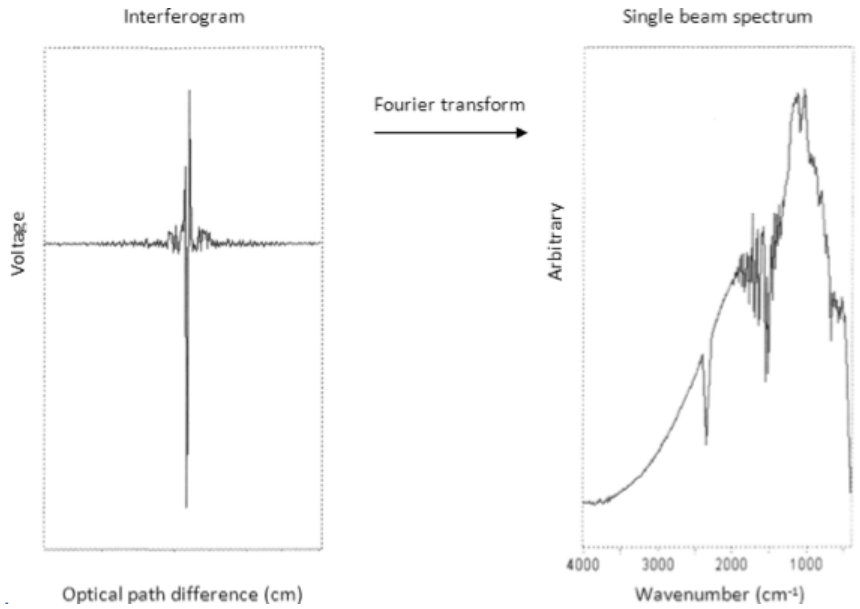

Figure 2.2. An illustration of how an interferogram is Fourier transformed to obtain an infrared spectrum. (Source: Smith 1996)

An FTIR instrument is consisted of an interferometer, detector, beamsplitter and source. Michelson interferometers are generally used in FTIR systems, as mentioned above. Deuterated triglycine sulfate (DTGS) and Mercury Cadmium Telluride (HgCdTe or MCT) are most commonly used detectors in FTIR. DTGS detectors are quite simple inexpensive and robust but they have a major drawbacks such as less sensitivity. Also, MCT detectors are 10 times more sensitive than DTGS. However, MCT detectors are usable with the range of 4000 - 700 cm^{-1}. Between 700 and 400 cm^{-1} MCT generates 5 - 10 times higher noisy compared to DTGS. The other drawback of MCT detectors are that they must be cooled, otherwise noise signals gets even worst.

Potassium Bromide (KBr) is the most commonly used beamsplitter in MIR region. However, a thin germanium can be placed between two pieces of KBr layers, and this reflects light well but is still somewhat transparent. Also Cesium Iodide (CsI) is used as a beamsplitter. However, while KBr is usable for mid-infrared region (4000 - 400 cm^{-1}), CsI can transmit from 4000 to 200 cm^{-1}, that means CsI is 200 cm^{-1} wider than the others (Ingle and Crouch 1988) (Smith 1996).

2.3. Attenuated Total Reflectance (ATR)

The Attenuated Total Reflectance (ATR) accessory is used for the analysis of liquids, solids, semi-solids and thin films with FTIR spectroscopy, which is placed onto the sample compartment of an FTIR instrument. ATR provides no sample preparation and non-destructive measurements of samples are possible (Boyd and Kirkwood 2011). This accessory also requires only a single drop for liquid samples that is applied directly to the crystal. An ATR accessory diagram is shown in Figure 2.3.

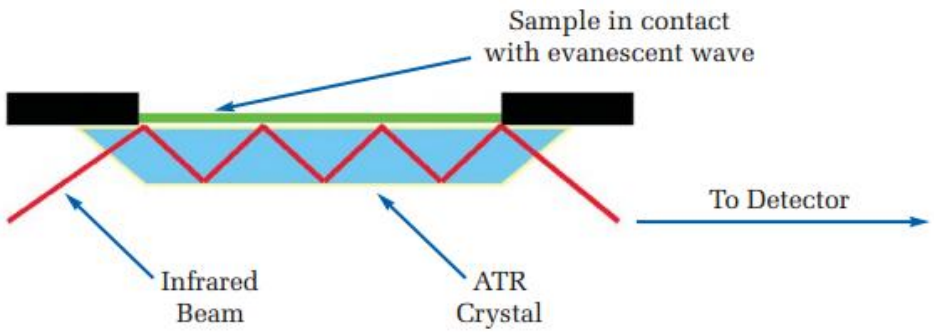

Figure 2.3. A schematic diagram of a multiple reflection Attenuated Total Reflectance accessory. (Source: Perkin Elmer)

There is a crystal which is made of infrared transparent material of high refractive index at the heart of the ATR accessory. A number of different crystal types are available such as Diamond, Zinc Selenide (ZnSe), Germanium and Thallium Iodide/Thallium Bromide, for various applications.

With ATR sampling, IR beam is directed into the crystal. The IR beam reflects from the internal surface of the crystal and creates an evanescent wave which projects orthogonally into the sample in intimate contact with the ATR crystal. Some of the energy of the evanescent wave is absorbed by the sample and the reflected radiation (some now absorbed by the sample) is returned to the detector (Gayete, Guaerdia, and Garrigues 2006). Graphical illustration of a typical single reflection of an ATR setup is shown Figure 2.4.

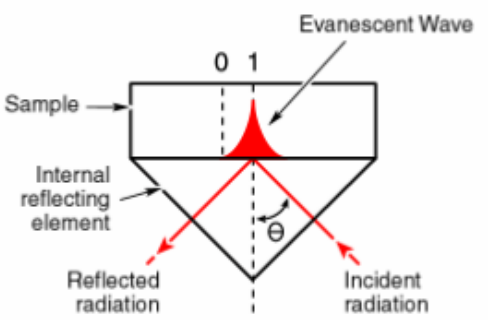

Figure 2.4. Graphical representation of a single reflection ATR.
(Source: PIKE Technologies)

When a sample contacts with the crystal, interaction is occurred with the evanescent wave and absorb infrared radiation. Then, the absorption of IR beam by the sample attenuates the evanescent wave, where the name of *attenuated total reflectance* comes from. It is very important that the surface of crystal must be kept smooth and clean because there is a need to be ensured that the evanescent wave penetrates into the sample. ATR is quite usable for solid and liquid samples. The accessory is designed to do an analysis easily both kinds of materials. So, nearly every sample can be analyzed using an ATR accessory (Smith 1996).

2.4. Advantages of FTIR

The FTIR spectroscopy has three major advantages over the other infrared spectrometers. First one is called multiplex advantage since the interferometer does not split energy into individual frequencies. Each wavelength of light being measured is contained in the resultant interferogram. Since multiple scans in a short period of time is possible these scans are averaged unlike the other infrared systems. The second advantage is that throughput advantage. The FTIR does not have a slit, so the light reaching to detector is not limited to certain frequencies. In a sense, more energy reaches to the sample resulting in higher signals are obtained. This leads to higher signal-to-noise ratio (SNR) and higher SNR means that sensitivity of the FTIR is greater. The third advantage is related to precision where FTIR spectrometers are controlled by the velocity of the movable mirror which saves time to collect data along the mirror stroke by a laser. This reference laser signal is also used within the

instrument, which higher provides accuracy and precision. (Smith 1996) (Grdadolnik 2002)

CHAPTER 3

MULTIVARIATE CALIBRATION METHODS

3.1. Overview

Chemometrics is a science that seriously depends on the use of statistical and mathematical methods to extract chemical information from the chemical data (Wold 1995). In modern analytical chemistry and biochemistry, chemometric approaches have become famous in quantitative and qualitative analysis of samples from spectroscopic data (Miller and Miller 2000). The process of data evaluation is called calibration.

Calibration is a mathematical model which is constructed with the known concentrations of a number of calibration samples and the spectral information of the same samples from a spectrometer. However, prediction is performed with the calibration model by using spectral information of unknown samples. In univariate calibration, concentration of one component is related to only one spectral response which is the maximum absorbing wavelength (Beebe, Pell, and Seasholtz 1998). On the other hand, multivariate calibration methods use all or several of the responses to determine the concentrations of multi-component mixture.

3.2. Univariate Calibration

In univariate calibration the aim is to find a relationship which relates a sample property, such as peak area, ratio of peak areas and spectral intensity at characteristic positions. When there is only one variable to measure (x) and one variable to predict (y), univariate calibration method is used (Naes et al. 2002). This technique is commonly used for quantitative analysis, where the correlation of the concentration of a sample and the instrument response is stated by Lambert Beer's Law. When this correlation is considered as a linear, there are two options:

- Classical Univariate Calibration
- Inverse Univariate Calibration

Classical univariate method uses a single wavelength in a spectrum and concentration is modeled by the absorbance. Classical calibration is representing with the formula of:

$$\mathbf{a} = \mathbf{c} \,.\, s + \mathbf{e} \tag{3.1}$$

a is the vector (*m x 1*) of absorbance at the single wavelength for a number of samples and **c** is the vector (*m x 1*) of corresponding concentrations. The scalar *s* depends **a** and **c**, and can be calculated with to the formula of 3.2:

$$s = (\mathbf{c}' \cdot \mathbf{c})^{-1} \cdot \mathbf{c}' \cdot \mathbf{a} \tag{3.2}$$

c′ is the transpose of the concentration vector. Once *s* is calculated, the concentration of an unknown can be estimated simply:

$$\hat{c} = a / s \tag{3.3}$$

scalar *a* represents the absorbance value for the unknown sample and $\hat{c}$ refers to the concentration of the unknown sample. Residuals (**e**) of the model are calculated with the difference between the observed (**c**) and predicted ($\hat{\mathbf{c}}$) concentrations in the calibration set.

$$\mathbf{e} = \mathbf{c} - \hat{\mathbf{c}} \tag{3.4}$$

Quality of a model can be assessed by examining the magnitude of residuals. For example the property of a component is studied between 0 - 100% and residuals were found to be ranging in between ± 1%, then it can be certainly concluded that this a good model since the maximum variability of the residual is only 1% in an hundreds unit interval (Brereton 2003).

Even if classical univariate calibration is generally used in chemistry, there are two main reasons that this method is not mostly convenient approach. The first reason is that the predicted concentration is obtained from the instrumental response. Second one is relates to error distributions due to instrumental performance. While instruments have

become more reproducible, the sensitivity and reproducibility of instruments has increased, but the quality of volumetric flasks, syringes and so on has not improved much. However, the concentration is usually calculated by weighing, dilutions and so on, they have become largest source of errors. So, Errors are mostly originated from concentration which is larger than instrumental error. Figure 3.1 represents the errors, where (a) is obtained from instrument and (b) is from concentration (Brereton 2003).

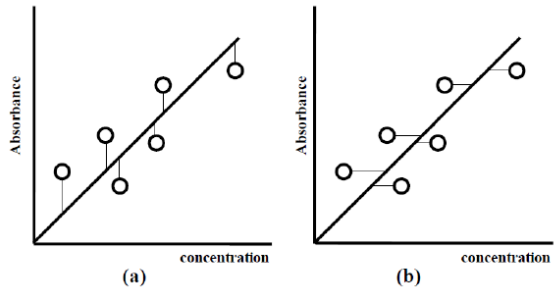

Figure 3.1. Difference between errors in (a) classical and (b) inverse calibration (Source: Brereton 2003).

Inverse calibration method can be expressed by the formula of:

$$\mathbf{c} = \mathbf{a} \cdot b \tag{3.5}$$

b is a scalar coefficient and is approximately inverse of **s** because each model makes assumptions on errors differently. b can be calculated by the following formula:

$$b = (\mathbf{a}' \cdot \mathbf{a})^{-1} \cdot \mathbf{a}' \cdot \mathbf{c} \tag{3.6}$$

and predicted concentration of an unknown sample is calculated as:

$$\hat{c} = \hat{a} \,.\, b \tag{3.7}$$

3.3. Multivariate Calibration

Today, rapid development of electronic and instrument technologies in parallel within the enormously growing computer technology has led the modern instruments produce very large data sets in a short time. (Miller and Miller 2000). Multivariate calibration is the process of relating these multiple responses to properties of a sample. The samples could be a mixture of chemical components, and the aim is to predict the concentrations of the different chemical components in the mixture from measurements (Beebe, Pell, and Seasholtz 1998).

Multivariate calibration method has some advantages over univariate methods;

- It is possible to analyze more than one component synchronically. With univariate analyses this is not possible (Beebe, Pell, and Seasholtz 1998).
- Multivariate calibration also enables fault-detection. Univariate analysis has no fault-detection capabilities (Beebe, Pell, and Seasholtz 1998).
- Multivariate calibration is an enhanced tool for selectivity and reliability, and can also increase analytical capacity and reliability of instruments (Martens and Naes 1989).
- When a peak overlap problem is occurred in a spectrum by virtue of different absorbers, it is not usually possible to predict the concentration of one absorbers whit the use of absorbance at a single wavelength (Naes et al. 2002).
- If the analyte is not stable or homogeneous and interferences contaminate the measurements, multivariate calibration methods are needed (Martens and Naes 1989).

In this study, partial least squares method (PLS) and genetic inverse least squares (GILS) method which is a hybrid method of genetic algorithms and inverse least squares are used. Therefore, before giving information about GILS, firstly classical least squares (CLS) and inverse least squares (ILS) methods have to be explained as an introduction to the multivariate calibration topic.

3.3.1. Classical Least Squares (CLS)

The classical least squares (CLS) method is also known as Beer's Law, where the absorbance is directly as a function of the analyte concentration. In this method, errors come from the instrument responses as classical univariate method. CLS is modeled by the following equation:

$$\mathbf{A} = \mathbf{C}\,\mathbf{K} + \mathbf{E_A} \tag{3.8}$$

where **A** is a *m x n* matrix that contains absorbance spectra of *m* calibration samples at *n* wavelengths of the samples at different wavelengths. **C** is the *m x h* matrix which consists of concentrations of each of the *h* components in the *m* calibration samples. K is the *h x n* matrix of absorptivity-pathlength constants and $\mathbf{E_A}$ is the *m x n* matrix of random errors or residuals that are not fit by the model. **K** matrix is estimated by the least square with the following equation:

$$\mathbf{K} = (\mathbf{C'} \cdot \mathbf{C})^{-1} \cdot \mathbf{C'} \cdot \mathbf{A} \tag{3.9}$$

After **K** matrix is calculated with the equation 3.9, concentration of an unknown sample is predicted from their spectrums by the following equation 3.10

$$\hat{\mathbf{c}} = (\mathbf{K} \cdot \mathbf{K'})^{-1} \cdot \mathbf{K'} \cdot \mathbf{a} \tag{3.10}$$

where **a** is the spectrum of unknown sample and $\hat{\mathbf{c}}$ is the vector of the predicted component concentrations. The residual is the difference between the predicted and reference concentration values are calculated by:

$$\mathbf{e} = \mathbf{c} - \hat{\mathbf{c}} \tag{3.11}$$

CLS is very useful method to set up a calibration model for the whole spectrum, but this method has a main disadvantage that all interfering chemical components must be known and their concentrations included in the model. Otherwise, CLS method will fail (Özdemir and Öztürk 2004). This requirement can be fixed by Inverse Least Square (ILS) method.

3.3.2. Inverse Least Squares (ILS)

The drawback of CLS can be eliminated by inverse least squares (ILS) method. This method is the inverse of Beer's Law. In this method, concentrations are modeled as a function of absorbance. The ILS method for m calibration samples with n wavelengths is given in the following equation:

$$\mathbf{C} = \mathbf{A}\,\mathbf{P} + \mathbf{E_C} \tag{3.12}$$

where **C** and **A** are the same as in CLS (**C** is the concentration matrix and **A** is the absorbance matrix). $\mathbf{E_C}$ is the *m x h* matrix of errors in the concentrations not fit by the model. **P** is the *n x h* matrix of the unknown calibration coefficients relating *h* component concentrations to the spectral intensities. The major advantage of ILS is that equation 3.12 can be reduced for the analysis of one component at a time. The reduced model is:

$$\mathbf{c} = \mathbf{A}\mathbf{p} + \mathbf{e_c} \tag{3.13}$$

where **c** is the *m x 1* vector of concentrations for the analyte that is being analyzed, **p** is *n x 1* vector of calibration coefficients and $\mathbf{e_c}$ is the *m x 1* vector of concentration residuals not fit by the model. During the calibration step, $\hat{\mathbf{p}}$ which is estimated **p** can be calculated with:

$$\hat{\mathbf{p}} = (\mathbf{A}' \cdot \mathbf{A})^{-1}\ \mathbf{A}' \cdot \mathbf{c} \tag{3.14}$$

Once $\hat{\mathbf{p}}$ is calculated, the concentration of the analyte of interest can be predicted with the equation 3.15:

$$\hat{c} = \mathbf{a}' \cdot \hat{\mathbf{p}} \tag{3.15}$$

where $\hat{c}$ is the scalar estimated concentration and $\mathbf{a}'$ is the spectrum of the unknown sample. The major advantage and ability of ILS is that the method can be predict one component at a time without knowing the concentrations of interfering species and this makes ILS one of the most preferably calibration method (Özdemir 2006).

Although ILS has more advantage than CLS, the drawback of the method is seen in Equation 3.14, where **A** matrix which must be inverted has huge dimensions equal to the number of wavelengths in the spectrum. However, this number needs to be equal to or smaller than the number of calibration samples. Furthermore, when more wavelengths are added to the model, it causes overfitting. Due to this effect, predicted concentrations will not be reasonable (Arpakçı 2013).

3.3.3. Partial Least Squares (PLS)

In order to calculate all possible variation in data matrix, principle components analysis (PCA) can be used. PCA allows the data to be represented with loadings and scores where the first column of scores contains the maximum variation in the data. Thus using only a few number of principle components can cover most of the information, effectively reduces dimensions. Hence, PCA can be directly or indirectly used as a classification method for multivariate data. PCA can also be used as a calibration method (principle components regression) since the scores can be regressed with concentrations.

Partial least squares (PLS) modeling is successfully applied to the quantitative analyses as a powerful multivariate statistical technique (Haaland and Thomas 1988). When there is partial knowledge of the data, this method becomes a useful tool.

PLS method has a main advantage over other multivariate calibration methods. It can analyses one chemical component at a time. The method does not require knowing all chemical components in the sample and there is no wavelength selection problem. The major feature of PLS is that it does not takes into account only concentration prediction errors, but also consider errors from spectra since it applies Principle Components Analysis to both concentrations and absorbances. PLS calibration method is shown to be composed of a series of simplified CLS and ILS steps (Brereton 2003).

The PLS model equation is described as:

$$\mathbf{A} = \mathbf{T}\,\mathbf{B} + \mathbf{E}_{\mathbf{A}} \tag{3.16}$$

where **A** is an *m x n* matrix of spectral absorbance, **B** is a *h x n* matrix of loading vectors or loading spectra. **T** is an *m x h* matrix of intensities or scores in the new coordinate

system of the *h* loading vectors for the m sample spectra. $\mathbf{E_A}$ is the *m x n* matrix residuals not fit by the model (Haaland and Thomas 1988) . The loading vectors in **B** are not pure component spectra, however they are linear combinations of the original calibration spectra. This is the difference between PLS and CLS. The number of basis vectors, *h*, to signify original calibration spectra which is determined by an algorithm during the calibration step. The spectral intensities (**T**) can be related to concentrations of the inverse least-squares analysis.

Concentration of the analyte is calculated by:

$$\mathbf{c = T\,v + e_c} \tag{3.17}$$

where **c** is the *m x 1* vector of component concentrations, **v** is the *h x l* vector of coefficients which relate spectral intensities to the component concentration and $\mathbf{e_c}$ is the *m x l* vector of errors in reference values of the component that is being modelled. These matrices are given in Figure 3.2. The product of **T** and **B** approximates to the spectral data and the product of **T** and **v** to the true concentrations.

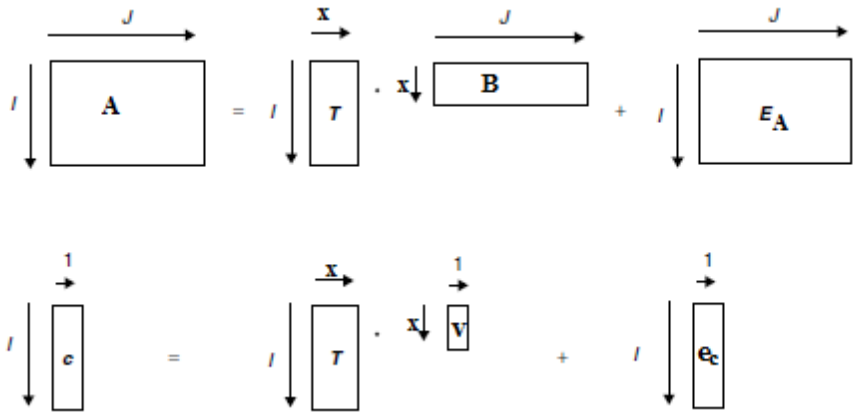

Figure 3.2. Principles of PLS
(Source: Brereton 2003).

The least-squares solution for **v** vector that is similar to the Equation 3.14 in ILS is given in the following equation:

$$\hat{\mathbf{v}}_{\mathbf{h}} = (\mathbf{T}' \mathbf{T})^{-1} \mathbf{T}' \mathbf{c} \qquad (3.18)$$

where $\hat{\mathbf{v}}_{\mathbf{h}}$ is the least-squares estimate of **v**. The **T** and **B** matrices are calculated in a stepwise manner (one vector at a time) till the desired model has been obtained.

For the complex chemical mixtures, there are two PLS algorithms are available. They are called PLS1 and PLS2 methods. PLS1 is used one component at a time to for the model building step. This method is generally used in PLS and it is reported that the predictions obtained with PLS1 are better that those obtained PLS2. So, it is recommended that PLS2 algorithm is more likely successful for qualitative application.

The PLS1 algorithm starts with the calculation of the estimated first weighed loading vector, $\widehat{\mathbf{w}}_{\mathbf{h}}$, by setting *h* to 1. This is done with the method of least squares and is given as:

$$\widehat{\mathbf{w}}_{\mathbf{h}} = \mathbf{A}' \mathbf{c} (\mathbf{c}' \mathbf{c})^{-1} \qquad (3.19)$$

where $\widehat{\mathbf{w}}_{\mathbf{h}}$ is an *n x l* vector representing the first order approximation of the pure component spectra for the component that is being analyzed. Then this calculated weighted loading vector is used to form the score vector $\hat{\mathbf{t}}_{\mathbf{h}}$, with an ILS prediction model. The first estimated $\hat{\mathbf{t}}_{\mathbf{h}}$ vector is estimated by:

$$\hat{\mathbf{t}}_{\mathbf{h}} = \mathbf{A}\, \widehat{\mathbf{w}}_{\mathbf{h}} \qquad (3.20)$$

This score vector can be related to the component concentrations with a linear least-squares regression. The scalar regression coefficient, $\hat{v}_h$, is estimated by:

$$\hat{v}_h = \hat{\mathbf{t}}_{\mathbf{h}}' \mathbf{c} (\hat{\mathbf{t}}_{\mathbf{h}}' \hat{\mathbf{t}}_{\mathbf{h}})^{-1} \qquad (3.21)$$

The residuals of concentration are obtained by using the least-squares estimated regression coefficients. In order to eliminate collinearity problems, the PLS loading

vector, $\hat{\mathbf{b}}_\mathbf{h}$, can be calculated now with a new model for **A**. Once again the method of least squares is used to find estimated b vector by:

$$\hat{\mathbf{b}}_\mathbf{h} = \hat{\mathbf{t}}'_\mathbf{h}\ \mathbf{A}\ (\hat{\mathbf{t}}'_\mathbf{h}\ \hat{\mathbf{t}}_\mathbf{h})^{-1} \tag{3.22}$$

where $\hat{\mathbf{b}}_\mathbf{h}$ is an *n x 1* vector. It is now possible to calculate the first PLS approximation to the calibration spectra by multiplying the score vector ($\hat{\mathbf{t}}_\mathbf{h}$) with transpose of PLS loading vector ($\hat{\mathbf{b}}'_\mathbf{h}$). Final calibration coefficients, $\mathbf{r_f}$, are obtained in the prediction step of PLS1. Once the $\mathbf{r_f}$ is calculated, it is simple to estimate the concentration of a new sample using the average concentration of the analyte and its spectra.

The prediction step in PLS1 is given by the following formula:

$$\mathbf{r_f} = \hat{\mathbf{W}}\ (\hat{\mathbf{B}}\ \hat{\mathbf{W}}'\ )^{-1}\ \hat{\mathbf{v}} \tag{3.23}$$

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

$$\hat{c} = \mathbf{a}'\ \mathbf{r_f} + c_0 \tag{3.24}$$

where $\hat{c}$ is the predicted unknown sample, $\mathbf{a}'$ is the spectrum of that sample and c_0 is the average concentration of calibration samples (Haaland and Thomas 1988).

The determining process of the optimal number of PLS factors are based on an algorithm and the cross-validation is one of the methods for this (Malinowski 1977). For n calibration spectra, the PLS1 algorithm is performed on n-1 spectra and the left out spectrum is used for the model validation. This process is continued until each spectrum is left out once in the calibration set. Afterwards, 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:

$$PRESS = \sum_{i=1}^{m} \left(c_i - \hat{c}_i\right)^2 \tag{3.25}$$

where c_i is the reference (known) concentration of the ith sample and $\hat{c}$ is the predicted concentration of the ith sample for m calibration standard (Öztürk 2003).

3.3.4. Genetic Inverse Least Squares (GILS)

The principle of Genetic inverse least squares (GILS) is a combination of ILS calibration method and Genetic Algorithms (GA). GA is a global search and optimization method based on natural evolution and selection which is developed by Darwin. (Wang, Veltkamp, and Kowalski 1991). Darwin's theory says that populations produce their offspring which will be with the best genetic fitness for the environment. So, the next generation will have evolved with a higher representation. (McClean 1997).

Recent decades, scientists have tried to take advantages of the natural evolutions in order to improve solving large scale optimization problems. Holland has performed the experiments with genetic algorithms in his researches and he is considered the father of the field (Gilbert et al. 1997).

The application of GA has five basic steps. These are initialization, evaluation, breeding and mating, crossover and mutation, and replacing as shown in Figure 3.3. These steps names come from the biological foundation of the algorithm.

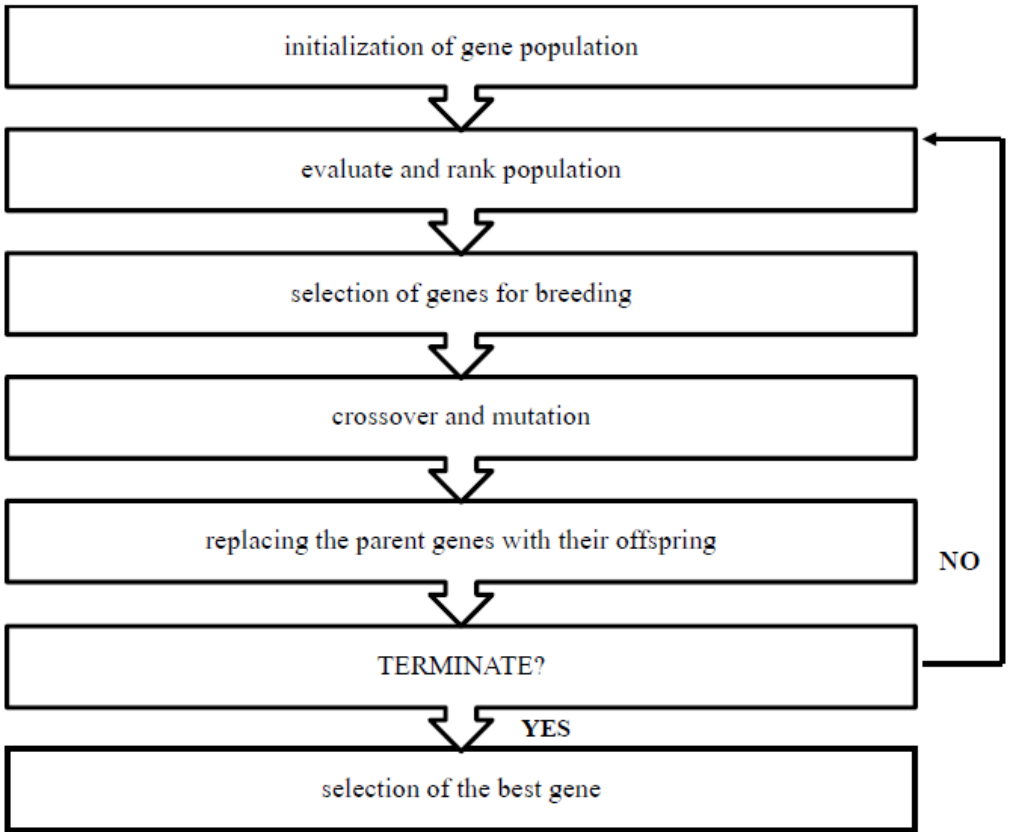

Figure 3.3. Flow chart of genetic algorithm used in GILS.

The term ‘gene’ describes collection of instrumental responses of the data set which are randomly selected and the collection of individual genes in the current generation is referred as ‘population’. A gene is shown in the Formula 3.26:

$$\mathbf{S = [A_{1754}\, A_{926}\, A_{2268}\, A_{596}\, A_{1255}\, A_{3500}]} \qquad (3.26)$$

where **S** is a gene, **A** is the absorbance which is measured at the indicated wavelength in the subscript. The initialization step creates the first generation of genes randomly with a fixed population size. In order to minimize bias and maximize the number of possible recombination, genes are selected randomly. The size of the gene pool is defined by the user. In order to allow for breeding of genes in the population, size must be even number. Population size determines the time that it takes to complete an individual run of the algorithm. The number of wavelength point in a gene is obtained randomly between fixed high limit and low limit. The lower limit was set to two in order to allow

single-point crossover, and the higher limit was set to eliminate overfitting problems and reduce computation time.

Once the initial gene population is produced, the next step is to evaluate and rank the genes using a fitness function, which is calculated by the inverse of the standard error of calibration with cross validation (SEC) (Özdemir 2005). The success of the model is confirmed by this function:

$$\text{Fitness} = \frac{1}{\text{SEC}} \tag{3.27}$$

where SEC is calculated from the ILS model with the formula given in the following equation:

$$SEC = \sqrt{\frac{\sum_{i=1}^{m}\left(c_i - \hat{c}_i\right)^2}{m-2}} \tag{3.28}$$

where c and $\hat{c}_i$ are the reference and predicted concentrations of analyte, respectively, for m samples. (m-2) is called as degrees of freedom. The number 'two' comes from that there are two parameters to be extracted. These parameters are the slope between the reference vs. predicted concentration values and the intercept. The SEC value is used for the determination of the success of each gene.

The third step 'selection of genes for breeding' based on the fundamental principle of natural evolution. Parent genes are selected from the current population to breed of a new generation. The aim is to create the best performing members of the population with the highest fitness value. This gives a chance to those genes to survive in the long run and will be able to pass their information to the next generations. Therefore, better offspring will be generated with the genes which are better suited for the problems. Otherwise, if the genes have low fitness value, breeding chance will be decrease and so most of them will be unable to survive.

There are several methods for parent selection such as roulette wheel method and top down method, tournament selections methods. In GILS Roulette wheel selection method is used where each part of roulette wheel symbolizes a gene. In the

Figure 3.4 illustrated a roulette wheel in which five genes are shown where each area on the wheel correspondes to the fitnness of the genes.

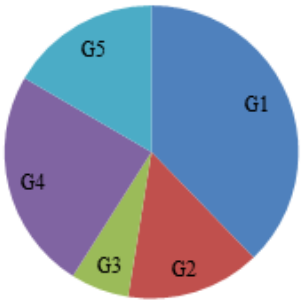

Figure 3.4. An example roulette wheel in which five genes are shown where each area on the wheel correspondes to the fitnness of the genes.

The gene with the highest fitness value has the largest part of roulette wheel and the one with lowest fitness value has the smallest part. A gene with high fitness has more possibilities to be selected than a gene with low fitness, when the wheel is spun. While some of genes will be selected several times, some of them will not be selected or will be thrown out from the gene pool. After genes are selected, they are allowed to mate top-down, whereby the first gene (S1) mates with the second gene (S2), S3 with S4 and so on in the order they are being selected until all the genes mates. There is a possibility of the genes with low fitness can be mate with better performing genes, because of the genes selected by roulette wheel having no ranking. Thus, possibility of recombination is increased.

'Crossover and mutation' step involves breaking of the genes at random points and the offspring genes are generated by cross-coupling. Parent genes (S1 and S2) and their corresponding offspring (NewS1 and NewS2) are illustrated in the following sample:

Parents:

S1= [A_{452} A_{3732} # A_{1237} A_{2890}]

S2 = [A_{923} A_{1457} A_{1743} # A_{832} A_{3022}]

The points where the genes are cut for mating are indicated by # and the place where crossover takes place.

Offspring:

NewS1 = [A_{452} A_{3732} A_{832} A_{3022}]

NewS2 = [A_{923} A_{1457} A_{1743} A_{1237} A_{2890}]

where A_{452} and the others represent the instrument response at the wavenumber written in subscript. S1 is the first and S2 is the second parent genes. Here, the first part of S1 is combined with the second part of S2 to give NewS1, likewise the second part of S1 with the first part of S2 to give NewS2. This process is called single point crossover and it is commonly used in GILS.

After crossover, the parent genes are replaced by their offspring and the offspring are evaluated. The ranking process is based on their fitness values follows the evaluation step. Then the selection for breeding/mating starts all over again. This is repeated until a predefined number of iterations are reached. In the end, the gene with the lowest SEC (highest fitness) is selected for the model building and the model is used to predict the concentration of component analyzed in the validation (test) set. The success of the model in the prediction of the validation set is evaluated using standard error of prediction (SEP) which is calculated with the following formula:

$$SEP = \sqrt{\frac{\sum_{i=1}^{m}\left(c_i - \hat{c}_i\right)^2}{m}} \tag{3.29}$$

where *m* now signifies the number of independent validation samples.

The final step, termination of the algorithm is done by setting predefined iteration number for the number of breeding/mating cycles. The best run which has the lowest SEC for the calibration set and at the same time produced SEP for validation set that is in the same range with SEC is selected for evaluation and further analysis.

CHAPTER 4

EXPERIMENTATION & INSTRUMENTATION

4.1. Sample Preparation

In this study, totally 150 raw goat milk samples were collected in 3 different seasons (Each sample from different goat and 50 samples for each season –June 2014, December 2014 and March 2015), from a goat farm in Çeşme, İzmir. The whole cow milk samples which were daily milk and SEK branded were bought from a supermarket (produced by microfiltration and pasteurized at 72^{o}C which were indicated on the label). Adulterated goat milk samples were prepared with the raw goat milk, whole cow milk and water. Water was added into the adulterated samples, in order to reduce correlation between goat and cow milk as a result of the nature of binary mixtures, directly. Otherwise, in a binary mixture as one of the component increased, the other must be decreased. Another reason for the use of water in the adulterated samples was that the adulteration of goat milk is not only with cow milk, but also the possibility of could be made with the water. Compositions of the ternary mixture samples including different amounts (percentage by weight between 0% –20%) of water and other components of the adulterated samples are given in the Table 4.1, 4.2 and 4.3. Correlation graphs of goat milk-cow milk, goat milk-water and cow milk-water of June-2014, December-2014 and March-2015 are presented in Figure 4.1, 4.2 and 4.3 with the given percentages of from the Table 4.1, 4.2 and 4.3, respectively.

Table 4.1. Percentages by weight of 50 adulterated samples which prepared from raw goat milk, whole cow milk and water in June 2014.

No	Goat Milk (w/w %)	Cow Milk (w/w %)	Water (w/w %)	No	Goat Milk (w/w %)	Cow Milk (w/w %)	Water (w/w %)
1	100.00	0.00	0.00	26	81.53	6.80	11.67
2	94.90	0.00	5.10	27	58.67	36.23	5.10
3	89.83	0.00	10.17	28	76.76	16.86	6.39
4	84.87	0.00	15.13	29	78.71	13.71	7.58
5	79.79	0.00	20.21	30	86.83	7.54	5.63
6	0.00	100.00	0.00	31	49.32	46.25	4.44
7	0.00	94.84	5.16	32	90.81	6.22	2.97
8	0.00	89.70	10.30	33	58.59	31.49	9.92
9	0.00	84.73	15.27	34	65.32	25.63	9.05
10	0.00	79.55	20.45	35	49.82	42.61	7.57
11	93.03	4.41	2.56	36	63.38	36.62	0.00
12	44.88	43.38	11.75	37	40.56	59.44	0.00
13	75.83	23.64	0.53	38	98.73	1.27	0.00
14	84.84	10.54	4.62	39	45.77	54.23	0.00
15	65.28	26.73	7.99	40	75.55	24.45	0.00
16	47.40	45.22	7.38	41	26.68	73.32	0.00
17	72.65	8.08	19.27	42	84.74	15.26	0.00
18	82.38	15.53	2.10	43	67.72	32.28	0.00
19	89.76	6.32	3.92	44	74.74	25.26	0.00
20	87.41	6.14	6.45	45	59.05	40.95	0.00
21	49.16	41.43	9.41	46	24.73	75.27	0.00
22	72.02	23.92	4.06	47	10.23	89.77	0.00
23	72.37	23.78	3.85	48	15.74	84.26	0.00
24	74.73	15.38	9.89	49	30.58	69.42	0.00
25	73.20	18.99	7.80	50	41.39	58.61	0.00

Table 4.2. Percentages by weight of 50 adulterated samples which prepared from raw goat milk, whole cow milk and water in December 2014.

No	Goat Milk (w/w %)	Cow Milk (w/w %)	Water (w/w %)	No	Goat Milk (w/w %)	Cow Milk (w/w %)	Water (w/w %)
1	86.80	3.21	9.98	26	99.69	0.31	0.00
2	81.31	13.42	5.27	27	31.88	68.12	0.00
3	88.86	1.16	9.98	28	21.06	78.94	0.00
4	98.74	1.26	0.00	29	51.66	28.96	19.38
5	85.60	12.45	1.95	30	77.31	5.47	17.22
6	2.06	97.94	0.00	31	95.86	4.14	0.00
7	75.48	14.43	10.09	32	59.60	26.88	13.52
8	73.43	26.57	0.00	33	59.95	20.49	19.55
9	58.35	24.17	17.48	34	70.27	10.38	19.35
10	3.80	96.20	0.00	35	99.35	0.65	0.00
11	69.42	30.58	0.00	36	55.32	44.68	0.00
12	23.42	76.58	0.00	37	87.58	4.63	7.80
13	80.10	14.74	5.16	38	79.52	15.84	4.63
14	37.25	62.75	0.00	39	42.29	57.71	0.00
15	36.54	63.46	0.00	40	88.35	2.46	9.19
16	91.40	5.93	2.68	41	85.55	7.95	6.50
17	64.85	28.34	6.81	42	62.47	27.14	10.39
18	67.51	25.00	7.49	43	78.27	21.73	0.00
19	31.48	68.52	0.00	44	92.91	6.05	1.04
20	77.69	9.96	12.36	45	72.86	16.44	10.69
21	64.00	21.45	14.55	46	69.44	24.10	6.46
22	88.05	7.99	3.96	47	81.74	1.89	16.37
23	56.27	26.52	17.21	48	82.41	10.75	6.84
24	91.29	8.71	0.00	49	30.08	69.92	0.00
25	65.35	34.65	0.00	50	79.35	17.73	2.92

Table 4.3. Percentages by weight of 50 adulterated samples which prepared from raw goat milk, whole cow milk and water in March 2015.

No	Goat Milk (w/w %)	Cow Milk (w/w %)	Water (w/w %)	No	Goat Milk (w/w %)	Cow Milk (w/w %)	Water (w/w %)
1	100.00	0.00	0.00	26	82.74	8.73	8.53
2	100.00	0.00	0.00	27	59.17	30.54	10.29
3	100.00	0.00	0.00	28	60.55	27.58	11.87
4	0.00	100.00	0.00	29	66.70	24.30	9.00
5	0.00	100.00	0.00	30	83.41	3.42	13.16
6	0.00	100.00	0.00	31	82.66	7.21	10.13
7	54.88	45.12	0.00	32	90.14	4.33	5.53
8	87.78	12.22	0.00	33	77.78	14.64	7.58
9	52.67	47.33	0.00	34	54.68	37.73	7.59
10	63.71	36.29	0.00	35	66.76	25.60	7.64
11	60.81	39.19	0.00	36	75.58	9.57	14.84
12	72.22	27.78	0.00	37	91.26	1.87	6.87
13	52.74	47.26	0.00	38	70.81	23.74	5.45
14	77.71	22.29	0.00	39	54.80	37.10	8.11
15	70.56	29.44	0.00	40	63.01	25.79	11.21
16	85.61	14.39	0.00	41	49.71	33.19	17.10
17	80.73	19.27	0.00	42	80.94	8.73	10.34
18	69.22	30.78	0.00	43	83.23	6.30	10.47
19	77.52	22.48	0.00	44	93.18	2.63	4.19
20	84.41	15.59	0.00	45	50.20	33.75	16.05
21	95.16	2.91	1.93	46	69.62	29.49	0.90
22	97.49	0.70	1.80	47	75.72	21.81	2.47
23	67.15	18.45	14.40	48	48.84	36.37	14.79
24	54.88	35.76	9.36	49	79.29	13.45	7.26
25	57.46	35.64	6.90	50	71.37	17.80	10.83

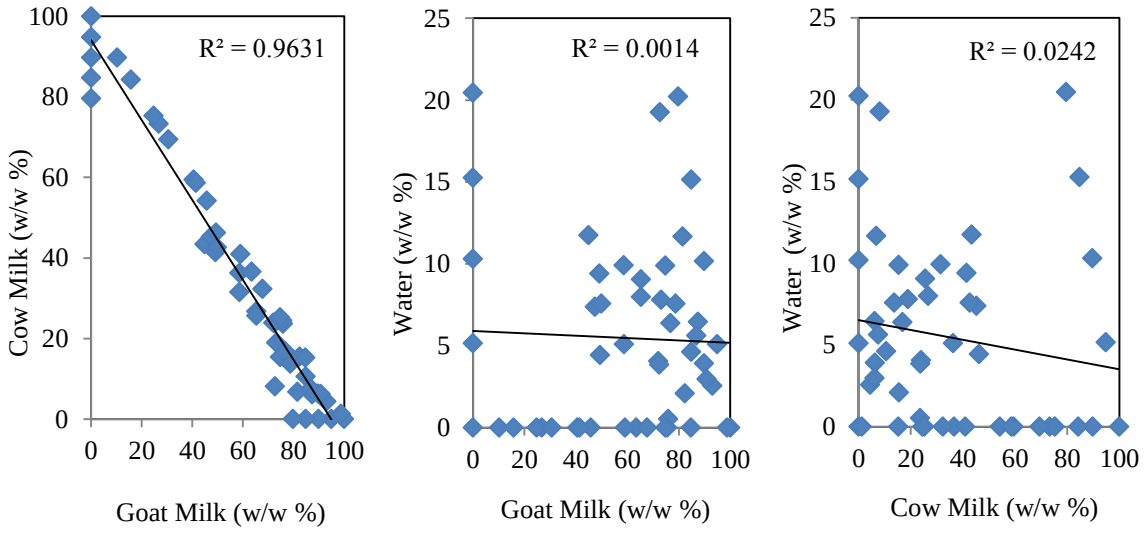

Figure 4.1. Correlation graphs of Goat Milk-Cow Milk, Goat Milk-Water and Cow Milk-Water of June 2014.

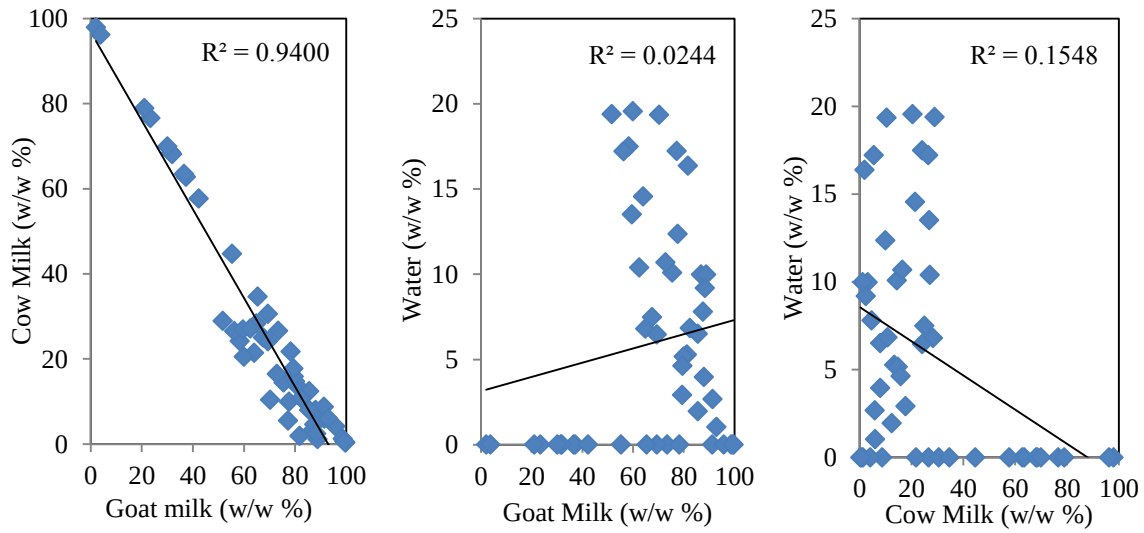

Figure 4.2. Correlation graphs of Goat Milk-Cow Milk, Goat Milk-Water and Cow Milk-Water of December 2014.

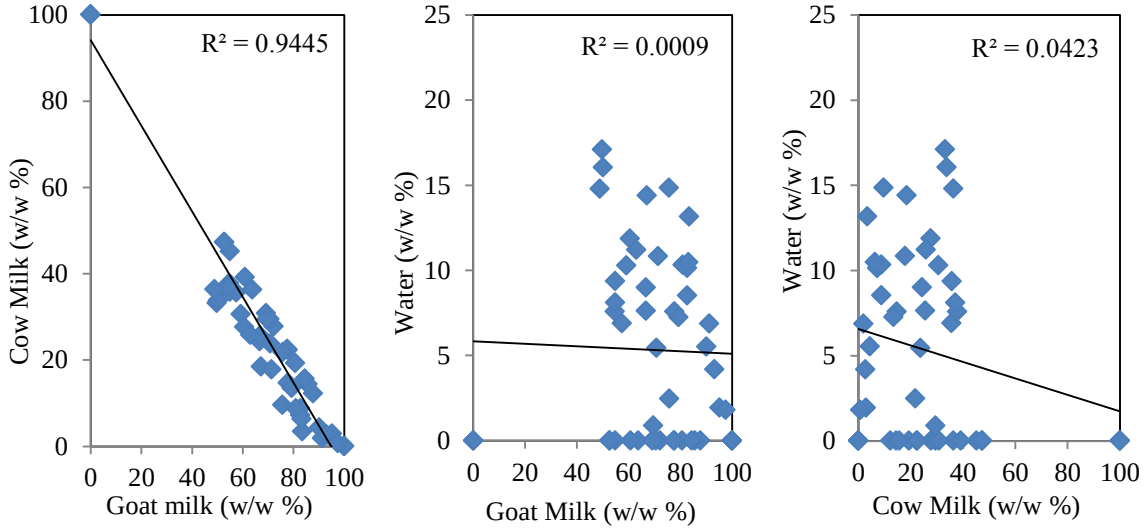

Figure 4.3. Correlation graphs of Goat Milk-Cow Milk, Goat Milk-Water and Cow Milk-Water of March 2015.

As it can be followed from the correlation graphs, correlation coefficients of goat milk-cow milk are quite high. In order to reduce the correlation, water was added to the samples and correlation was reduced between goat milk and water consequently, like cow milk and water.

4.2. Instrumentation and Data Processing

Fourier Transform Infrared spectroscopic analyses were performed with Perkin Elmer Spectrum 100 FTIR spectrometer which equipped with a Ne-He laser, Cr-Ni as a light source, a DTGS (Deuterated Triglycine Sulphate) as a detector, and KBr as a beam splitter. FTIR spectra of both raw and adulterated samples were collected using 3 reflection Diamond ATR (Attenuated Total Reflectance) accessory, between 4000 – 650 cm^{-1} with the spectral resolution of 4 cm^{-1}. Each spectrum was recorded as log (1/R) against to water background and saved in order to generate multivariate calibration models. Because of the 80-90% of goat milk is water in average (Russ et al. 2010), with the background of air the required data may be hidden in the peak of water instead of water background. Calibration, independent validation and second independent validation (raw goat milk samples) sets –which contains all the raw goat samples- were prepared as text files with the aid of Microsoft Excel (MS Office 2010, Microsoft Corporation) program using the adulterated samples given in the Table 4.4 and 4.5 and raw goat samples. The Genetic Inverse Least Squares (GILS) and Partial Least Squares (PLS) multivariate calibration methods were coded in Matlab programming (Matlab R2013a - MathWorks Inc., Natick, MA). Both method were performed with the option of ‘’Leave-One-Out Cross Validation’’ mode.

4.3. Design of Data Sets

Design of the calibration set is the first step of developing the calibration model. The estimated maximum and minimum values must be included in the calibration range. Also, the successes of the models on prediction are tested by analyzing the independent validation sets (validation and second independent validation set). In order to prepare calibration and independent validation sets, 50 synthetically prepared ternary mixture samples along with 10 raw goat milk samples were used. Among these 60 samples, 10 raw goat milk samples and 34 randomly selected ternary and binary mixture samples from the Tables 4.1, 4.2 and 4.3 were used to prepare calibration set and the remaining 16 ternary and binary mixture samples were reserved as independent validation set. After preliminary modeling, one of the calibration samples is found to be outlier in June 2014 and March 2015, one of the validation samples is found to be outlier in December

2014, and therefore they were taken out from the data sets. Table 4.4 shows concentration profiles of calibration sets that were prepared for the 3 different seasons. The calibration set of June contains 11 raw samples (goat or cow milk), 13 binary samples (goat and cow milk, goat milk and water, cow milk and water) and 19 ternary samples (goat milk, cow milk and water). The calibration set of December contains 9 raw samples (goat or cow milk), 16 binary samples (goat and cow milk) and 19 ternary samples (goat milk, cow milk and water). The calibration set of March contains 13 raw samples (goat or cow milk), 10 binary samples (goat and cow milk) and 20 ternary samples (goat milk, cow milk and water). Table 4.5 shows concentration profiles of validation sets that were prepared for the 3 different seasons. The validation set of June contains 10 binary samples (goat and cow milk, goat milk and water, cow milk and water) and 6 ternary samples (goat milk, cow milk and water). The validation set of December contains 4 binary samples (goat and cow milk) and 11 ternary samples (goat milk, cow milk and water). The validation set of March contains 2 raw samples (goat or cow milk), 4 binary samples (goat and cow milk) and 10 ternary samples (goat milk, cow milk and water).

Since the possibility of not predict the other seasons' raw goat milk sample contents by the model of single season, binary and ternary models were prepared which are combined with different seasons' data. Consequently, 3 more scenarios were generated for the combination of two seasons (June 2014 – December 2014, June 2014 – March 2015 and December 2014 – March 2015). Finally last scenario was generated with the combining of 3 seasons' data. Table 4.6 represents the number of samples for all scenarios. 120 raw goat milk samples in the second independent validation set consist of a combination of 40 raw goat milk samples from three different seasons. Each seasons' samples were combined in order to prepare binary and ternary models' calibration and validation sets. Totally 7 different scenario were generated. All scenarios were analyzed with GILS individually. PLS method was only performed to final scenario to compare the success of GILS and PLS one to each other.

Table 4.4. Concentration profiles of Calibration sets for goat milk, cow milk and water samples.

No	Calibration Set - June 2014			Calibration Set - December 2014			Calibration Set - March 2015		
	Goat Milk (w/w%)	Cow Milk (w/w%)	Water (w/w%)	Goat Milk (w/w%)	Cow Milk (w/w%)	Water (w/w%)	Goat Milk (w/w%)	Cow Milk (w/w%)	Water (w/w%)
1	100.00	0.00	0.00	86.80	3.21	9.98	100.00	0.00	0.00
2	94.90	0.00	5.10	81.31	13.42	5.27	100.00	0.00	0.00
3	79.79	0.00	20.21	88.86	1.16	9.98	0.00	100.00	0.00
4	0.00	100.00	0.00	98.74	1.26	0.00	0.00	100.00	0.00
5	0.00	84.73	15.27	85.60	12.45	1.95	54.88	45.12	0.00
6	0.00	79.55	20.45	2.06	97.94	0.00	87.78	12.22	0.00
7	93.03	4.41	2.56	75.48	14.43	10.09	63.71	36.29	0.00
8	44.88	43.38	11.75	73.43	26.57	0.00	60.81	39.19	0.00
9	65.28	26.73	7.99	58.35	24.17	17.48	52.74	47.26	0.00
10	47.40	45.22	7.38	3.80	96.20	0.00	77.71	22.29	0.00
11	72.65	8.08	19.27	69.42	30.58	0.00	85.61	14.39	0.00
12	87.41	6.14	6.45	23.42	76.58	0.00	80.73	19.27	0.00
13	49.16	41.43	9.41	80.10	14.74	5.16	77.52	22.48	0.00
14	72.02	23.92	4.06	37.25	62.75	0.00	84.41	15.59	0.00
15	72.37	23.78	3.85	36.54	63.46	0.00	97.49	0.70	1.80
16	74.73	15.38	9.89	91.40	5.93	2.68	67.15	18.45	14.40
17	73.20	18.99	7.80	64.85	28.34	6.81	57.46	35.64	6.90
18	76.76	16.86	6.39	67.51	25.00	7.49	82.74	8.73	8.53
19	78.71	13.71	7.58	31.48	68.52	0.00	60.55	27.58	11.87
20	86.83	7.54	5.63	77.69	9.96	12.36	66.70	24.30	9.00
21	49.32	46.25	4.44	64.00	21.45	14.55	82.66	7.21	10.13
22	90.81	6.22	2.97	88.05	7.99	3.96	90.14	4.33	5.53
23	58.59	31.49	9.92	56.27	26.52	17.21	54.68	37.73	7.59
26	98.73	1.27	0.00	99.69	0.31	0.00	70.81	23.74	5.45
27	45.77	54.23	0.00	31.88	68.12	0.00	63.01	25.79	11.21
28	75.55	24.45	0.00	21.06	78.94	0.00	49.71	33.19	17.10
29	67.72	32.28	0.00	51.66	28.96	19.38	83.23	6.30	10.47
30	74.74	25.26	0.00	77.31	5.47	17.22	93.18	2.63	4.19
31	59.05	40.95	0.00	95.86	4.14	0.00	69.62	29.49	0.90
32	24.73	75.27	0.00	59.60	26.88	13.52	75.72	21.81	2.47
33	10.23	89.77	0.00	59.95	20.49	19.55	79.29	13.45	7.26
34	41.39	58.61	0.00	70.27	10.38	19.35	71.37	17.80	10.83
35	100.00	0.00	0.00	99.35	0.65	0.00	100.00	0.00	0.00
36	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00
37	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00
38	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00
39	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00
40	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00
41	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00
42	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00
43	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00
44				100.00	0.00	0.00			

Table 4.5. Concentration profiles of Validation sets for Goat milk, Cow milk and Water samples.

No	Validation Set - June 2014			Validation Set - December 2014			Validation Set - March 2015		
	Goat Milk (w/w%)	Cow Milk (w/w%)	Water (w/w%)	Goat Milk (w/w%)	Cow Milk (w/w%)	Water (w/w%)	Goat Milk (w/w%)	Cow Milk (w/w%)	Water (w/w%)
1	89.83	0.00	10.17	55.32	44.68	0.00	100.00	0.00	0.00
2	84.87	0.00	15.13	87.58	4.63	7.80	0.00	100.00	0.00
3	0.00	94.84	5.16	79.52	15.84	4.63	52.67	47.33	0.00
4	0.00	89.70	10.30	42.29	57.71	0.00	72.22	27.78	0.00
5	75.83	23.64	0.53	88.35	2.46	9.19	70.56	29.44	0.00
6	84.84	10.54	4.62	85.55	7.95	6.50	69.22	30.78	0.00
7	82.38	15.53	2.10	62.47	27.14	10.39	95.16	2.91	1.93
8	89.76	6.32	3.92	78.27	21.73	0.00	54.88	35.76	9.36
9	81.53	6.80	11.67	92.91	6.05	1.04	59.17	30.54	10.29
10	58.67	36.23	5.10	72.86	16.44	10.69	83.41	3.42	13.16
11	63.38	36.62	0.00	69.44	24.10	6.46	77.78	14.64	7.58
12	40.56	59.44	0.00	81.74	1.89	16.37	75.58	9.57	14.84
13	26.68	73.32	0.00	82.41	10.75	6.84	54.80	37.10	8.11
14	84.74	15.26	0.00	30.08	69.92	0.00	80.94	8.73	10.34
15	15.74	84.26	0.00	79.35	17.73	2.92	50.20	33.75	16.05
16	30.58	69.42	0.00				48.84	36.37	14.79

Table 4.6. The number of samples that used for the 7 different scenarios of the models.

	Scenario	Calibration Set	Independent Validation Set	Second Independent Validation (raw goat milk samples) Set
Single Model	June 2014	43	16	120 (40+40+40)
	December 2014	44	15	120 (40+40+40)
	March 2014	43	16	120 (40+40+40)
Binary Models	June 2014 – December 2014	87 (43+44)	31 (16+15)	120 (40+40+40)
	June 2014 – March 2015	86 (43+43)	32 (16+16)	120 (40+40+40)
	December 2014 – March 2015	87 (44+43)	31 (15+16)	120 (40+40+40)
Ternary Model	June 2014 – December 2014 – March 2015	130 (43+44+43)	47 (16+15+16)	120 (40+40+40)

CHAPTER 5

RESULTS AND DISCUSSION

In this thesis study, a FTIR spectroscopic method based on multivariate calibration data analysis for the determination of goat milk adulteration with cow milk was developed. Totally, 150 raw goat milk samples were obtained from a goat farm in the 3 different sampling periods. Raw goat milk samples were adulterated with whole cow milk and water. Both raw samples and adulterated samples were analyzed with FTIR and were saved their spectra. Multivariate calibration techniques (GILS and PLS) were performed on data of the spectra. After analyzing data with chemometric methods, the results were evaluated and compared.

5.1. FTIR-ATR Results

Figure 5.1, 5.2 and 5.3 represent FTIR spectra of 50 raw goat milk samples for each, obtained from a goat farm in June 2014, December 2014 and March 2015, respectively. In addition, the spectral differences between goat milk and cow milk are compared in Figure 5.4. As depicted in the figure almost similar FTIR spectrum was observed for goat and cow milk.

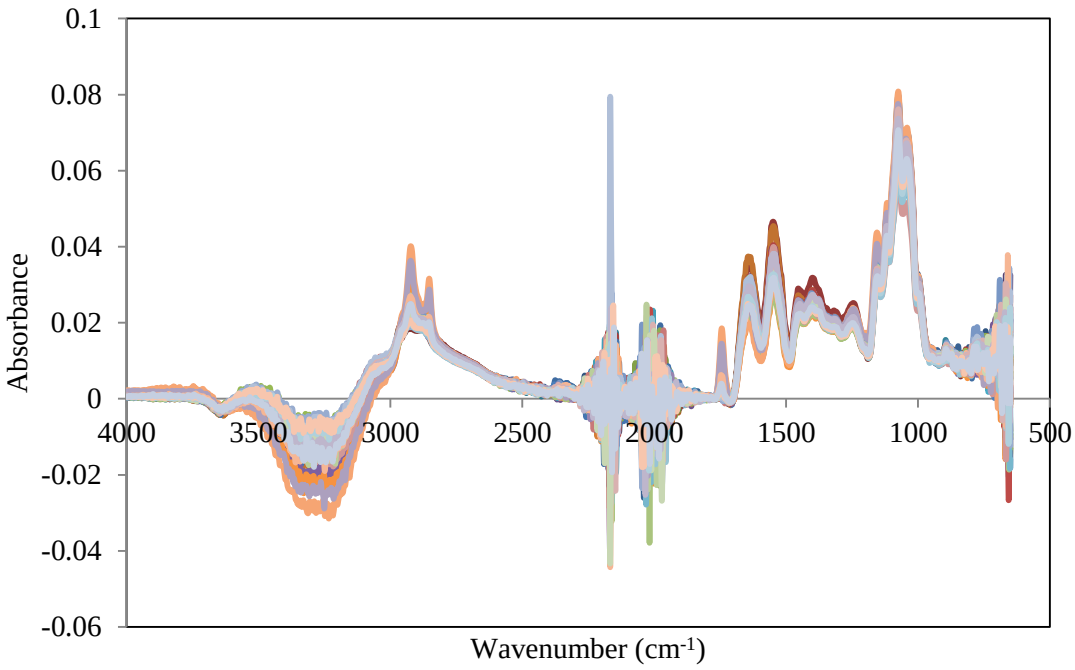

Figure 5.1. FTIR spectra of milk samples collected from 50 different goats in June 2014.

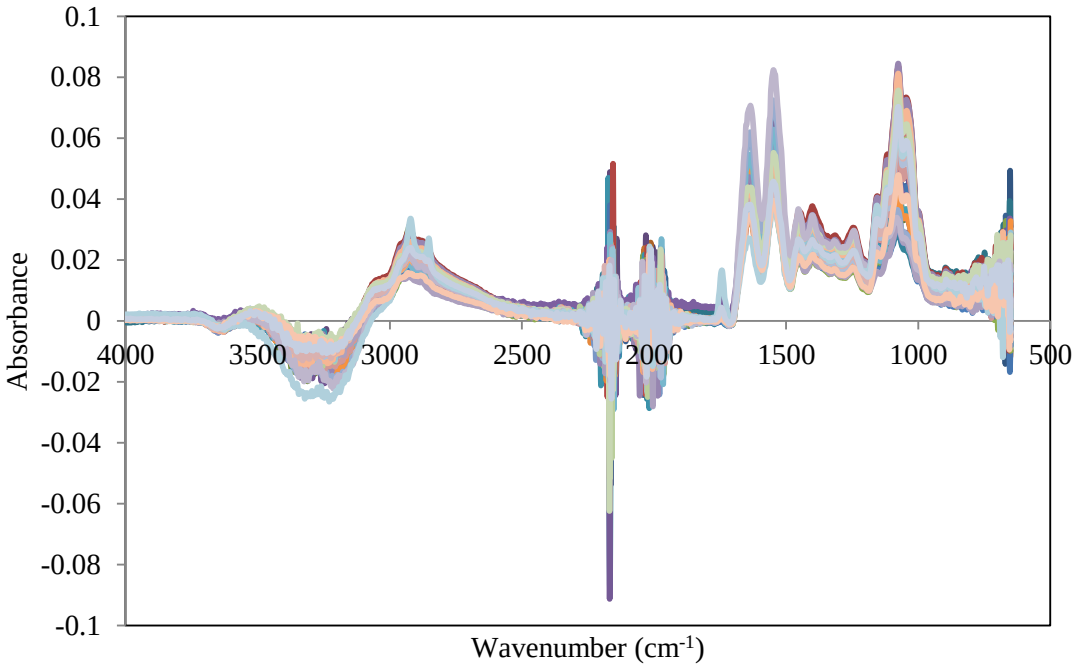

Figure 5.2. FTIR spectra of milk samples collected from 50 different goats in December 2014.

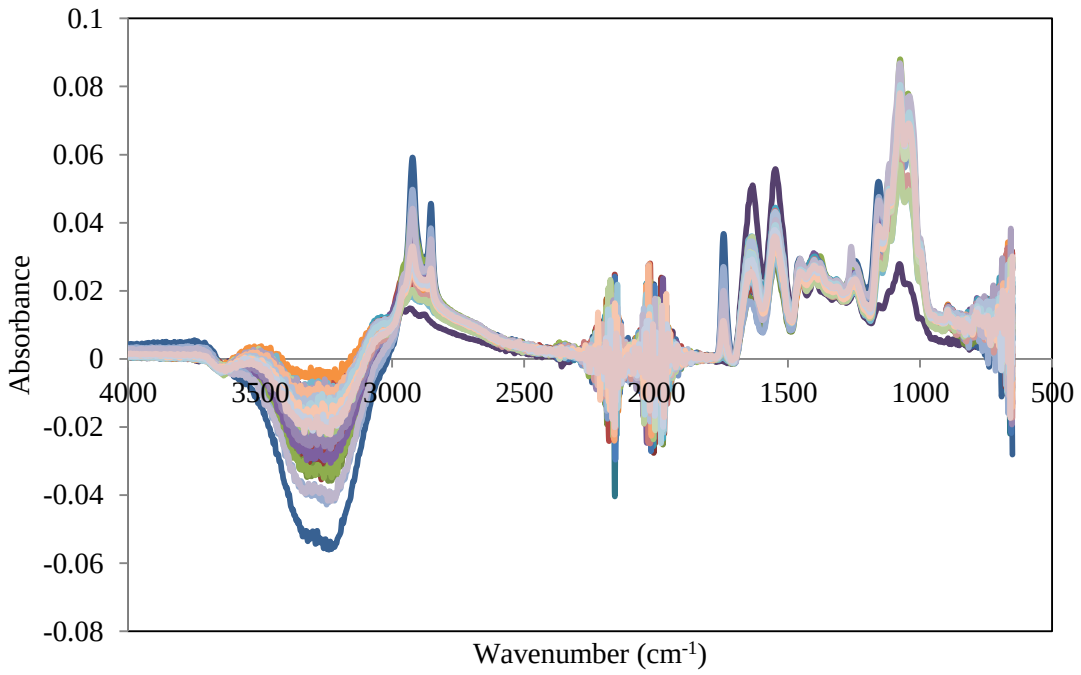

Figure 5.3. FTIR spectra of milk samples collected from 50 different goats in March 2015.

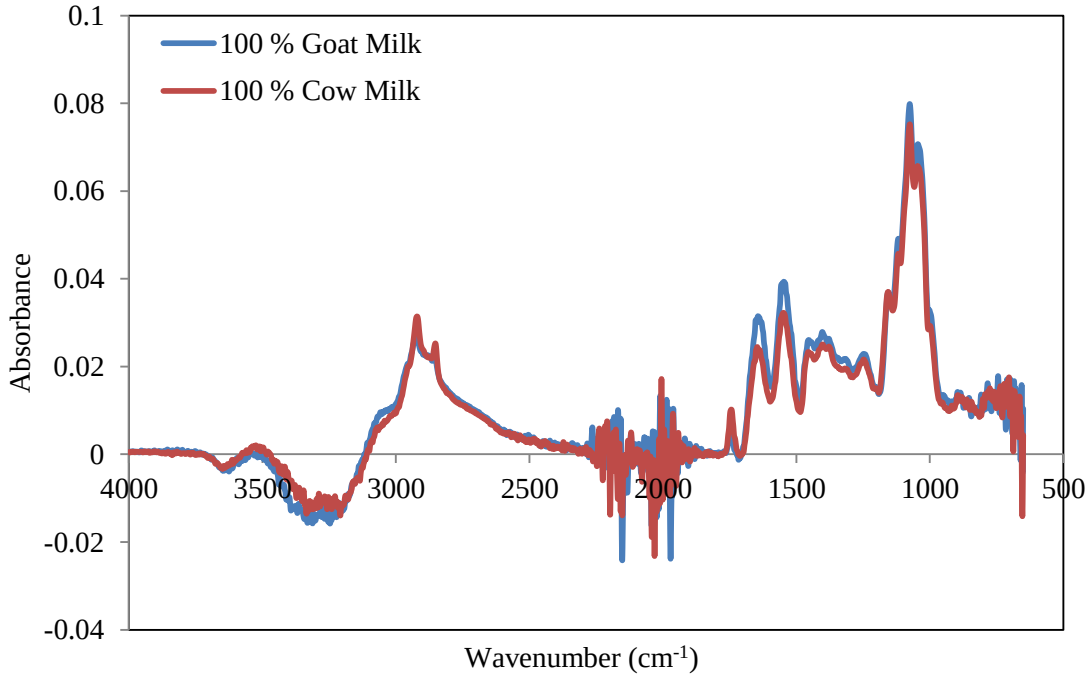

Figure 5.4. FTIR spectra of milk and raw cow milk samples.

FTIR spectra of 50 ternary mixture samples are given in Figure 5.5, 5.6 and 5.7 for June 2014, December 2014 and March 2015, respectively, which were prepared according to the given ratios of Table 4.1, 4.2 and 4.3.

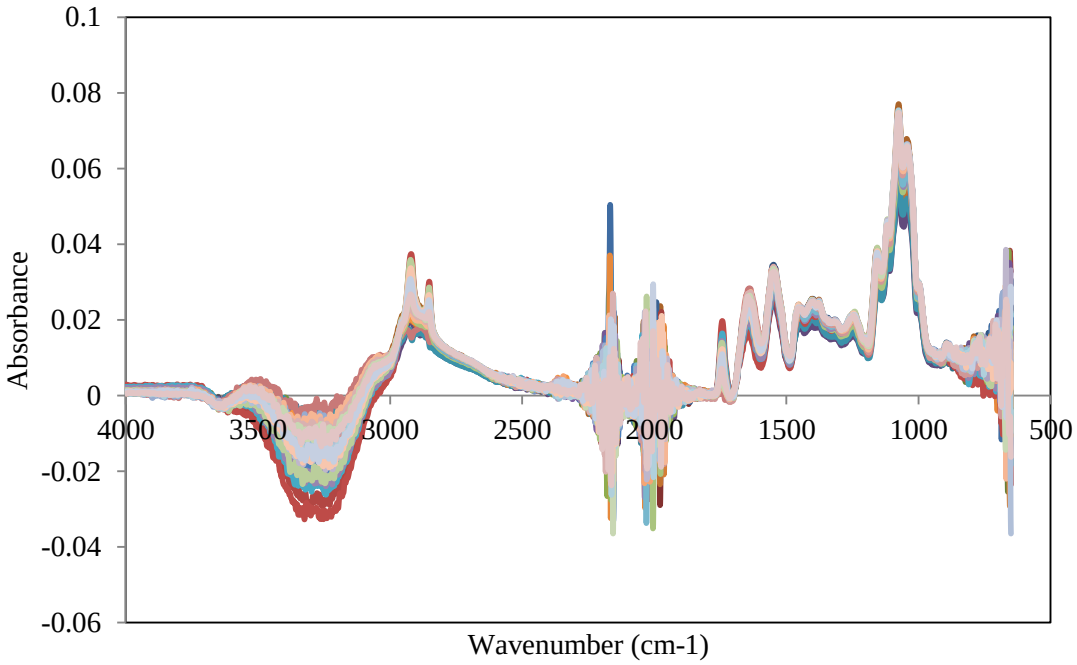

Figure 5.5. FTIR spectra of 50 ternary mixture samples in June 2014.

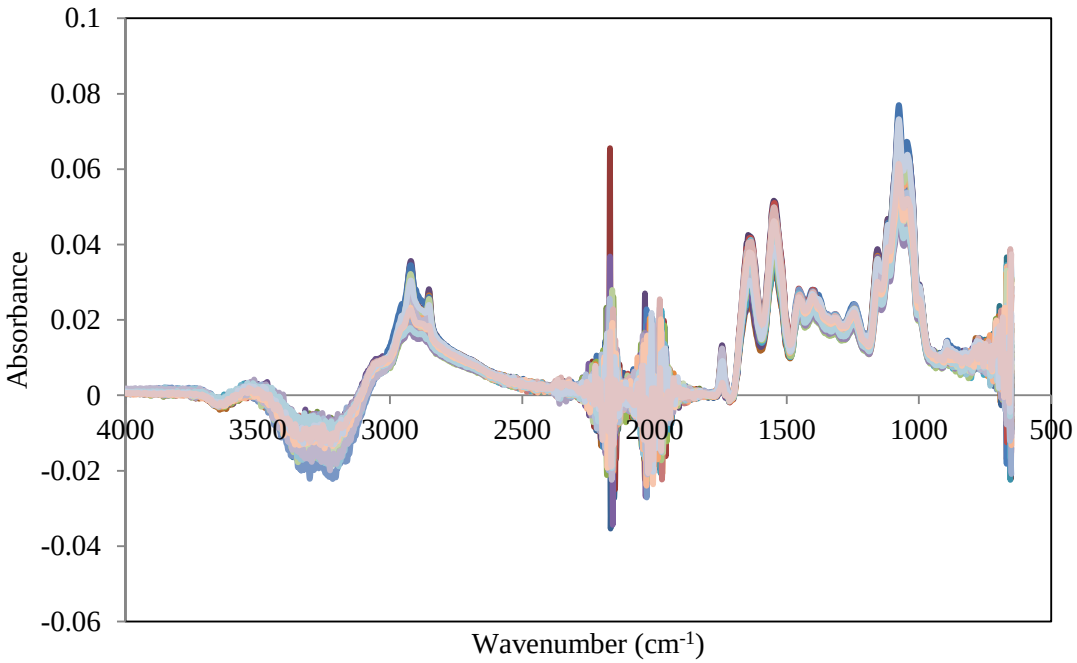

Figure 5.6. FTIR spectra of 50 ternary mixture samples in December 2014.

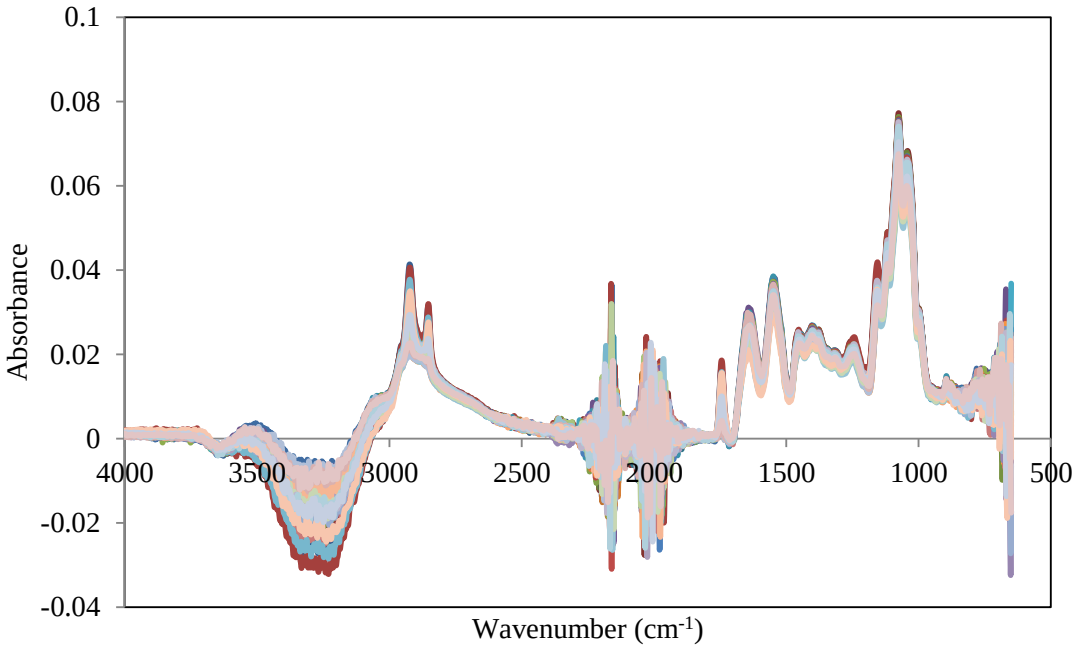

Figure 5.7. FTIR spectra of 50 ternary mixture samples in March 2015.

5.2. Multivariate Calibration Results

7 different multivariate calibration models (given in Table 4.6) were generated with synthetically prepared samples by using Genetic Inverse Least Squares (GILS) and Partial Least Squares (PLS) and both adulterated samples and raw goat milk samples were predicted with these models.

5.2.1. GILS Results for 3 Different Single Season Models

The content of the synthetically prepared adulterated milk samples are given in Table 4.1, Table 4.2 and Table 4.3 that were obtained in June 2014, December 2014 and March 2015, respectively. The samples were separated into calibration and validation sets as given in Table 4.4 and Table 4.5 which are available in the previous chapter. GILS models were set up for each season, for goat milk, cow milk and water, separately with the data obtained from FTIR. Correlation graphs, for June 2014 data set, were plotted for predicted percentages of goat milk, cow milk and water by GILS model versus actual values are given in Figure 5.8, 5.10 and 5.12, respectively. Further, the developed GILS models were applied to 120 raw goat milk samples and the predicted results for goat, cow milk and water percentages are given in Figure 5.9, 5.11 and 5.13, respectively.

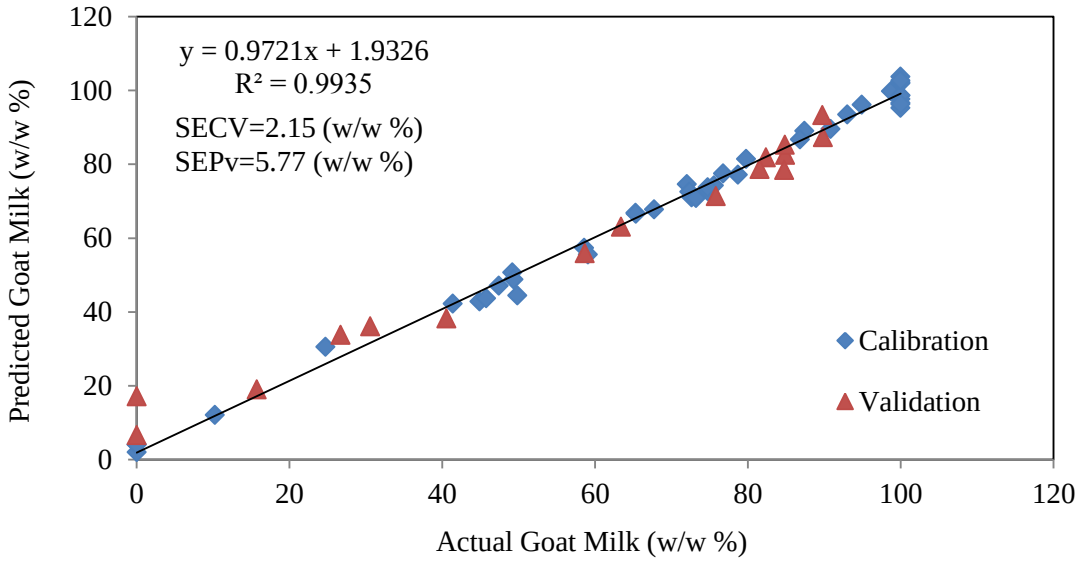

Figure 5.8. Actual versus predicted plot of goat milk content obtained from GILS model by using June 2014 data set.

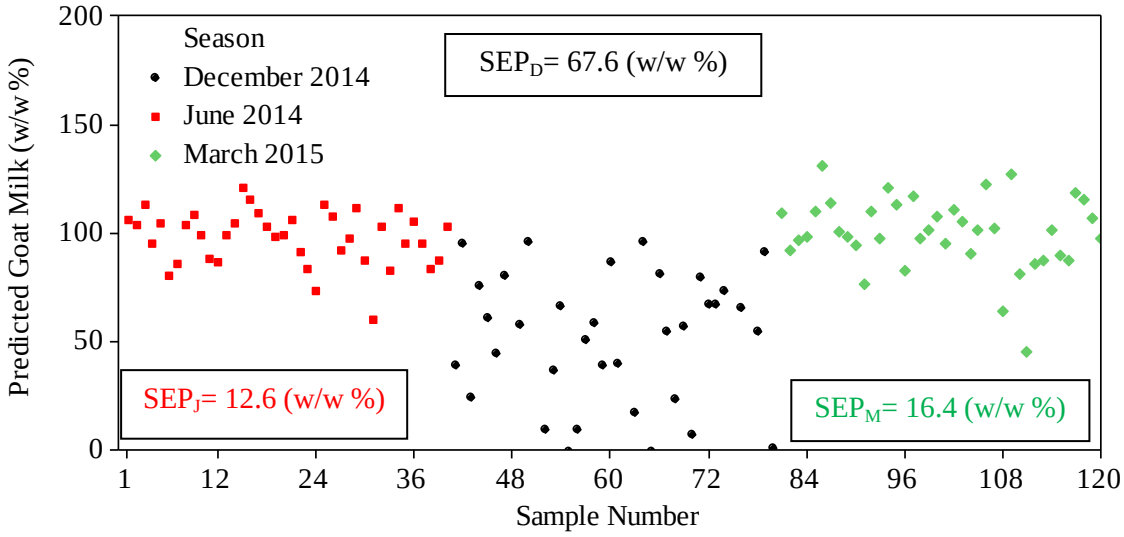

Figure 5.9. Predicted goat milk content of raw goat milk samples, estimated by GILS model by using the data set from June 2014 sampling period.

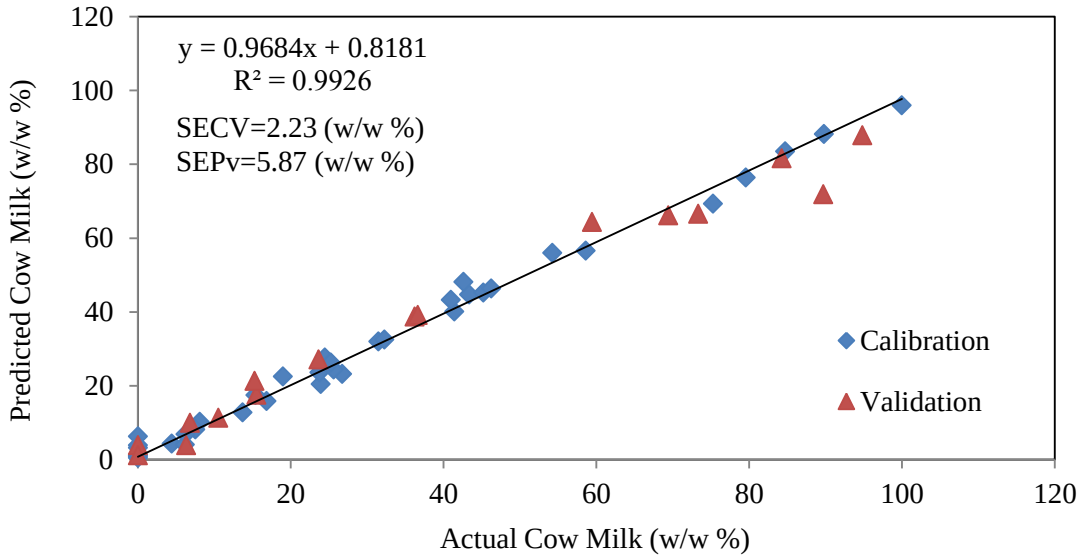

Figure 5.10. Actual versus predicted plot of cow milk content obtained from GILS model by using June 2014 data set.

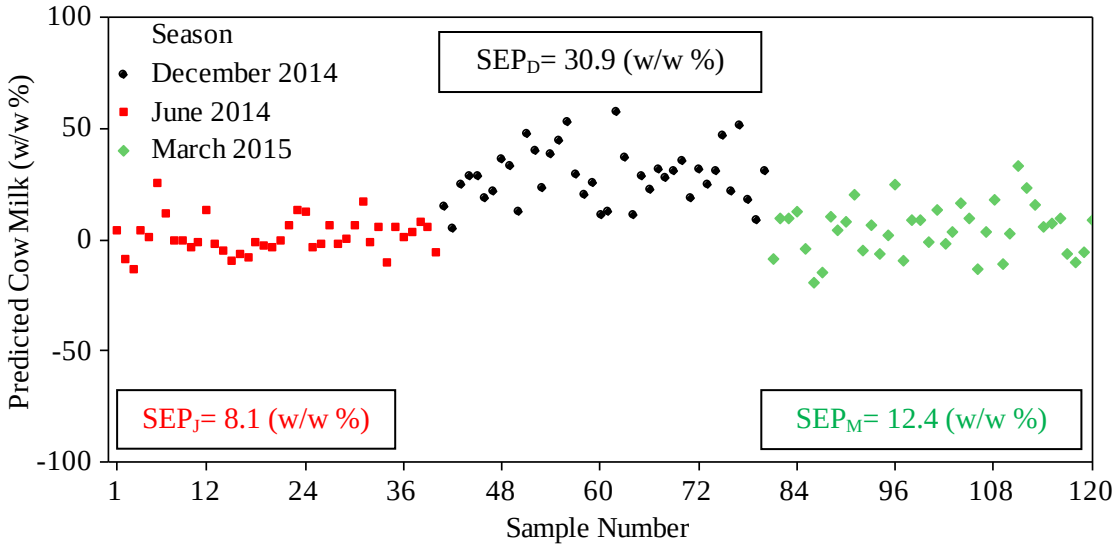

Figure 5.11. Predicted cow milk content of raw goat milk samples, estimated by GILS model by using the data set from June 2014 sampling period.

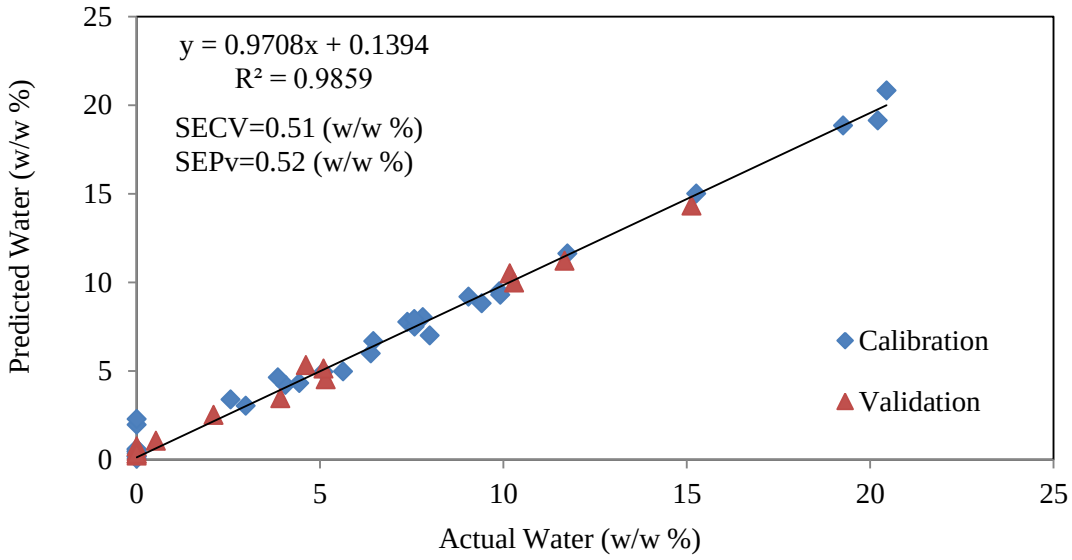

Figure 5.12. Actual versus predicted plot of water content obtained from GILS model by using June 2014 data set.

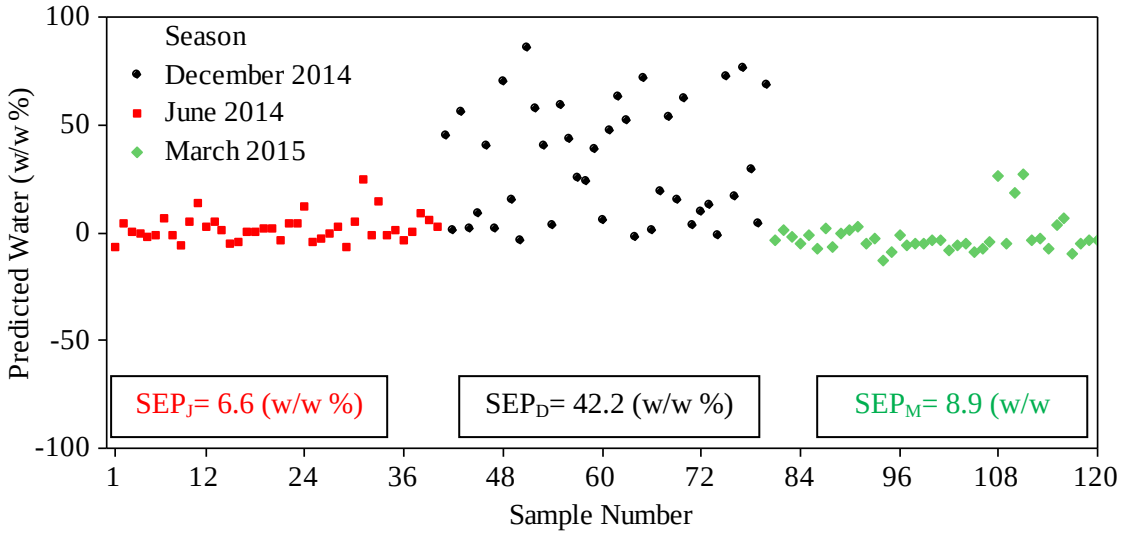

Figure 5.13. Predicted water content of raw goat milk samples, estimated by GILS model by using the data set from June 2014 sampling period.

The results of GILS model was discussed for the elucidation of the method performance. As it can be followed from the GILS model graphs of June 2014, the correlation coefficients were fluctuated between 0.986 and 0.994. In addition, the graphs also shows standard error of cross-validation (SECV) values for the calibration

set and standard error of prediction (SEPv) values for the validation set of the model which were varied between 0.51 and 2.23 (w/w %) and 0.52 and 5.87 (w/w %) respectively.

The water content of mixed samples varied in the dynamic range between 0 – 20% (w/w) however, this range is between 0 – 100% for both goat and cow milk. Therefore, the individual examination of water, apart from goat and cow milk, should be performed during the SECV and SEPv evaluation.

On the other hand, the results from the models of goat and cow milk can be compared with each other. Accordingly, it is seen that SECV and SEPv values of calibration and independent validation sets were almost found same for goat and cow milk models. It can be concluded that the predicted values of the independent validation sets by the models seem to be successful for each component of mixed samples. However, the key point of the method success is prediction of the second independent validation (raw goat milk samples) set. As it can be followed from the Figure 5.9, 5.11 and 5.13, the satisfactory results for milk cannot be obtained, for any season and for December 2014 in particular. Another important point is that, although predicted goat milk values of December 2014 are below 100 % by the model of goat milk, predicted cow milk values are over 0% by the model of cow milk.

This situation can be explained by the use of adulterated milk samples obtained in only June 2014; however the predicted goat milk samples were obtained in any season. The seasonal variation in the milk content had resulted low accurate results.

In particular, December 2014 is the beginning of the first period of breeding goats, the water content in milk composition varies highly according to other seasons. On the other hand, the variation corresponded to the FTIR system can affect the results as the water model results showed similar fluctuations (Figure 5.6).

Correlation graphs of predicted goat milk, cow milk and water percentages obtained from GILS model results versus actual values are given in Figure 5.14, 5.16 and 5.18, respectively, for December 2014. In the sequel, by use of that GILS models, goat milk, cow milk and water percentages were predicted in 120 milk samples and the obtained results are given in Figure 5.15, 5.17 and 5.19, respectively.

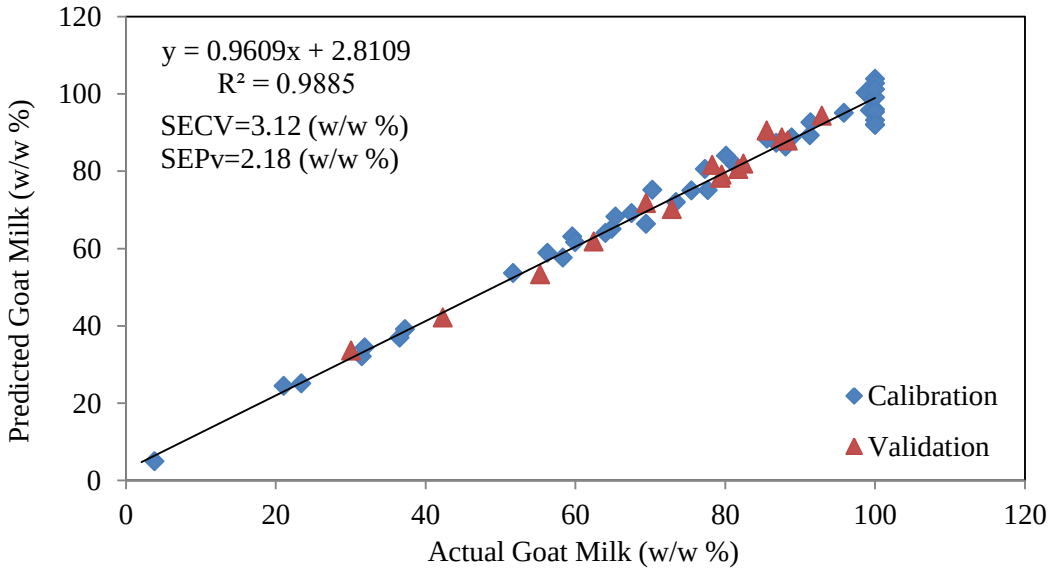

Figure 5.14. Actual versus predicted plot of goat milk content obtained from GILS model by using December 2014 data set.

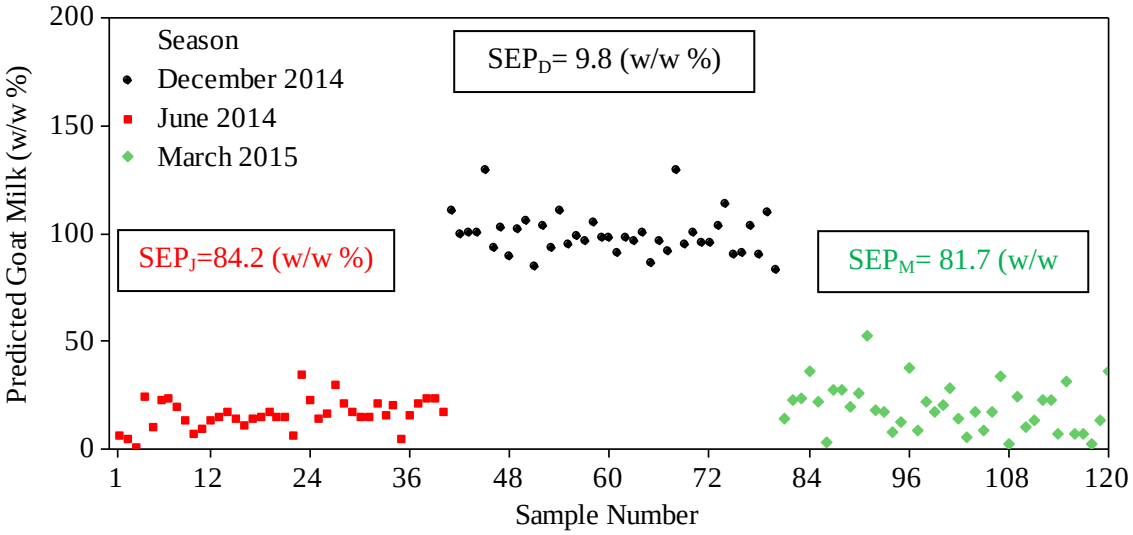

Figure 5.15. Predicted goat milk content of raw goat milk samples, estimated by GILS model by using the data set from December 2014 sampling period.

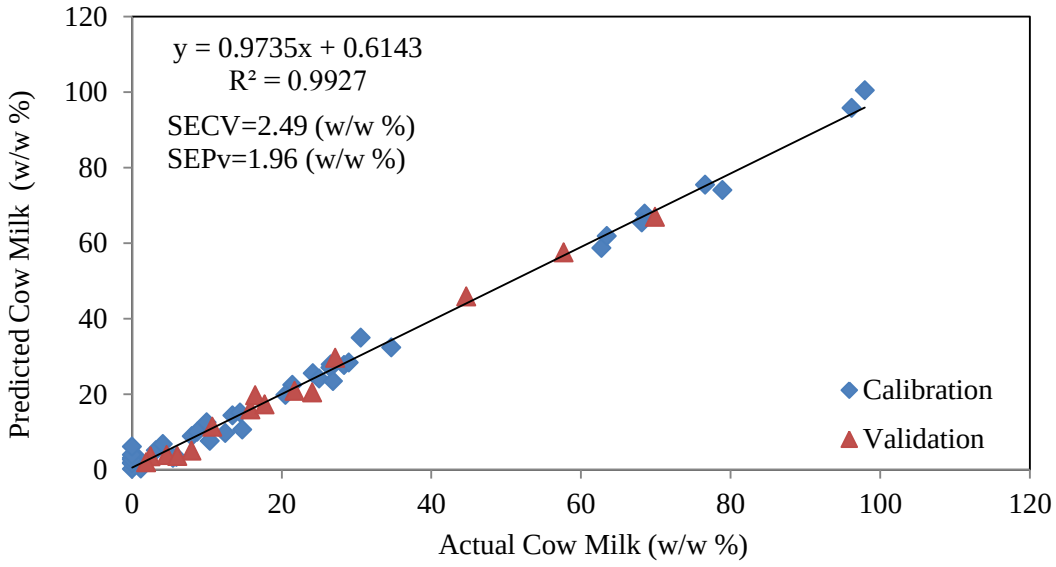

Figure 5.16. Actual versus predicted plot of cow milk content obtained from GILS model by using December 2014 data set.

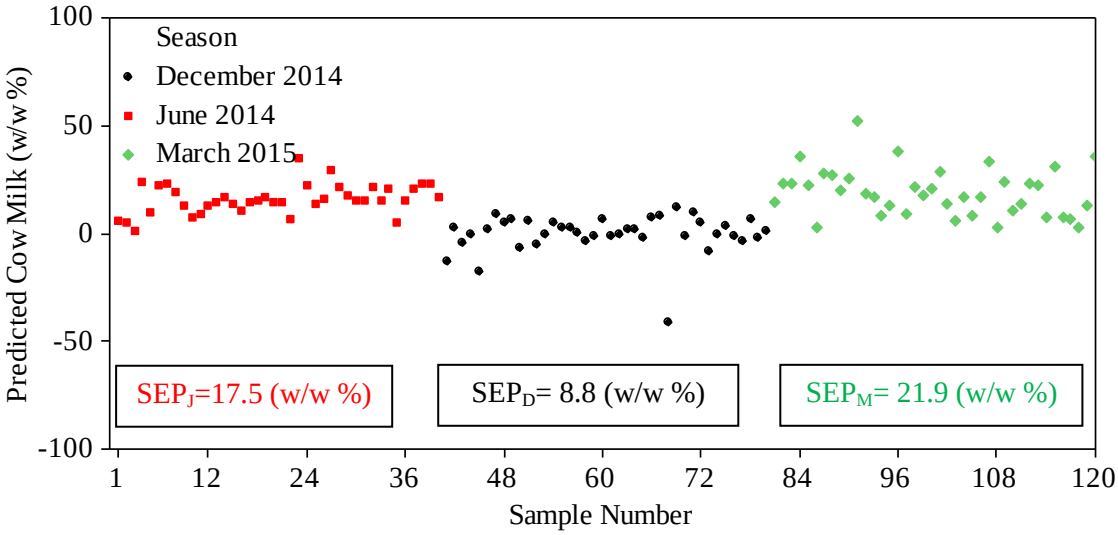

Figure 5.17. Predicted cow milk content of raw goat milk samples, estimated by GILS model by using the data set from December 2014 sampling period.

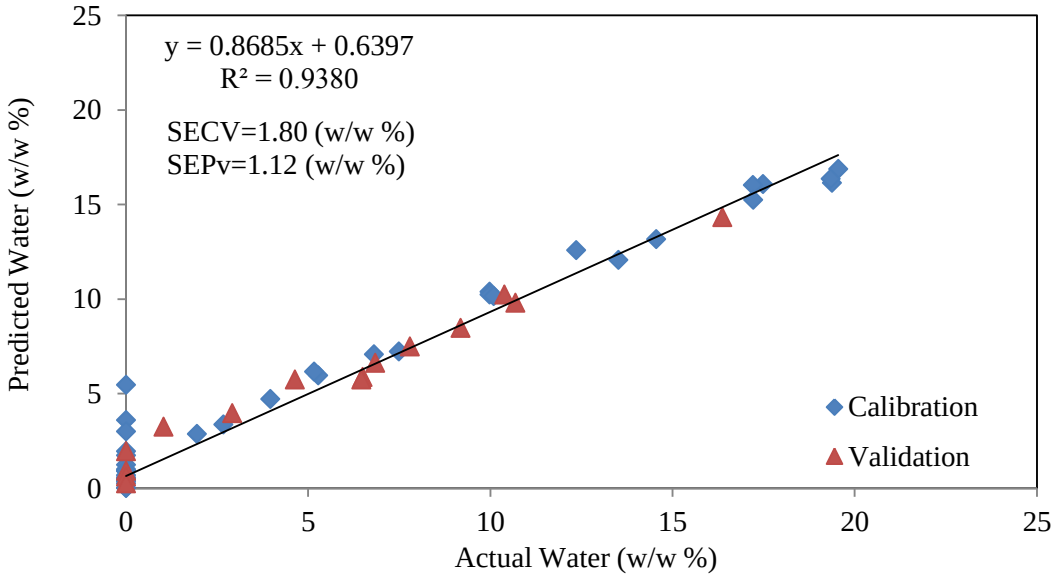

Figure 5.18. Actual versus predicted plot of water content obtained from GILS model by using December 2014 data set.

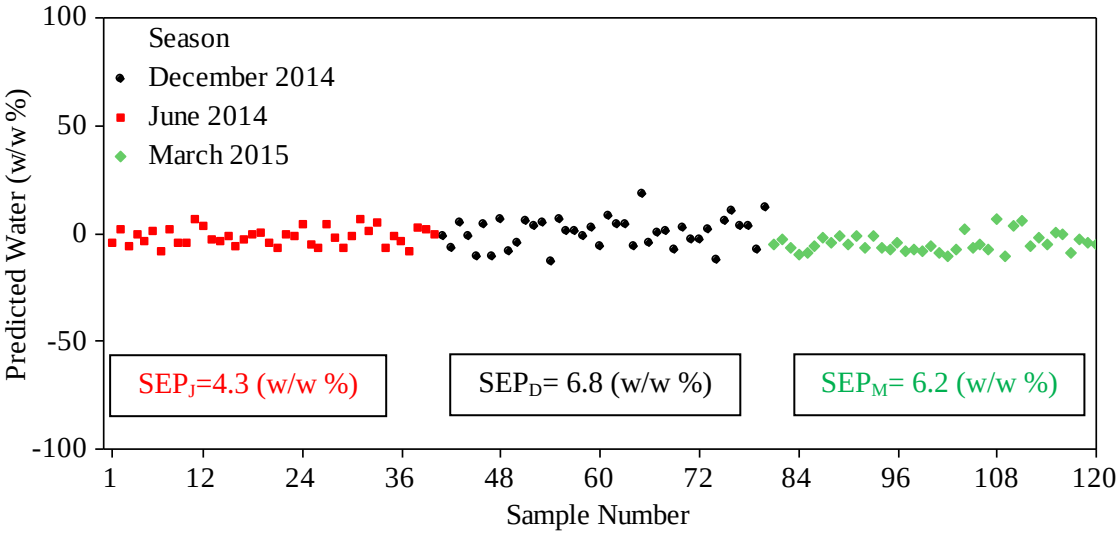

Figure 5.19. Predicted water content of raw goat milk samples, estimated by GILS model by using the data set from December 2014 sampling period.

It can be followed from the figure 5.14, 5.16 and 5.18 correlation coefficients of GILS models were found to be, 0.9885, 0.9927 and 0.9380 for goat milk, cow milk and water, respectively. On the other hand, SECV were fluctuated between 1.80 and 3.12 (w/w %), and found to be acceptable SEPv was calculated for independent validation

ranged from 1.12 and 2.18 (w/w %). As the indicated parameters shown in the graphs the models were found to be successful for validation sets. However, as indicated in the figure 5.15, 5.17 and 5.19 unsatisfactory results for raw goat milk samples were obtained by the model and these results can be attributed to seasonal variations as explained in June 2014 results. As a conclusion, December 2014 model were unable to predict other seasons' raw goat milk samples.

Correlation graphs of predicted goat milk, cow milk and water contents obtained from GILS model results versus actual values are given in Figure 5.20, 5.22 and 5.24, respectively, for March 2015 sampling period. Further, by use of that GILS models, goat milk, cow milk and water percentages were predicted in 120 milk samples and the obtained results are given in Figure 5.21, 5.23 and 5.25, respectively.

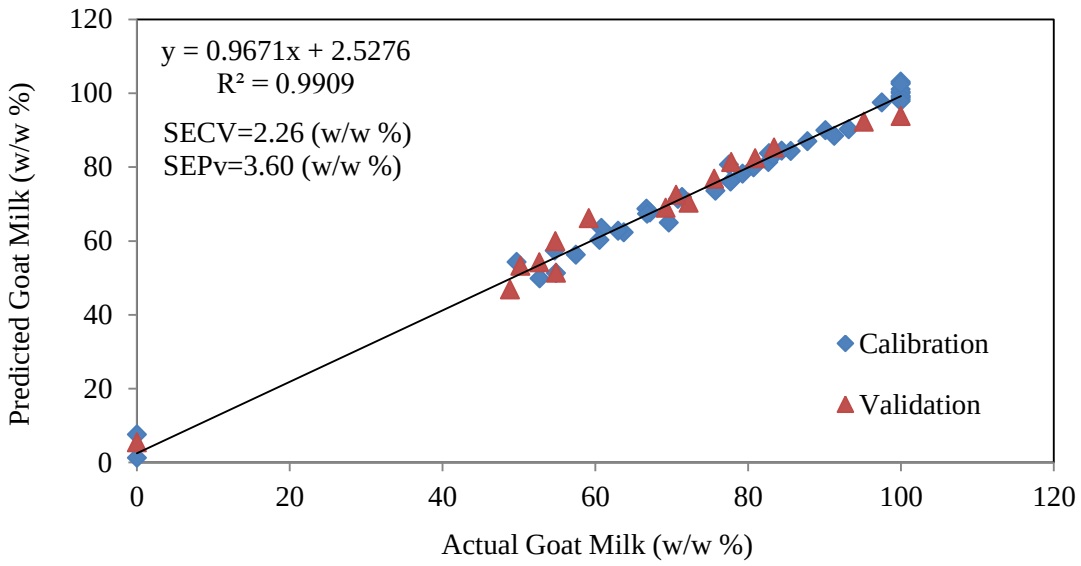

Figure 5.20. Actual versus predicted plot of goat milk content obtained from GILS model by using March 2015 data set.

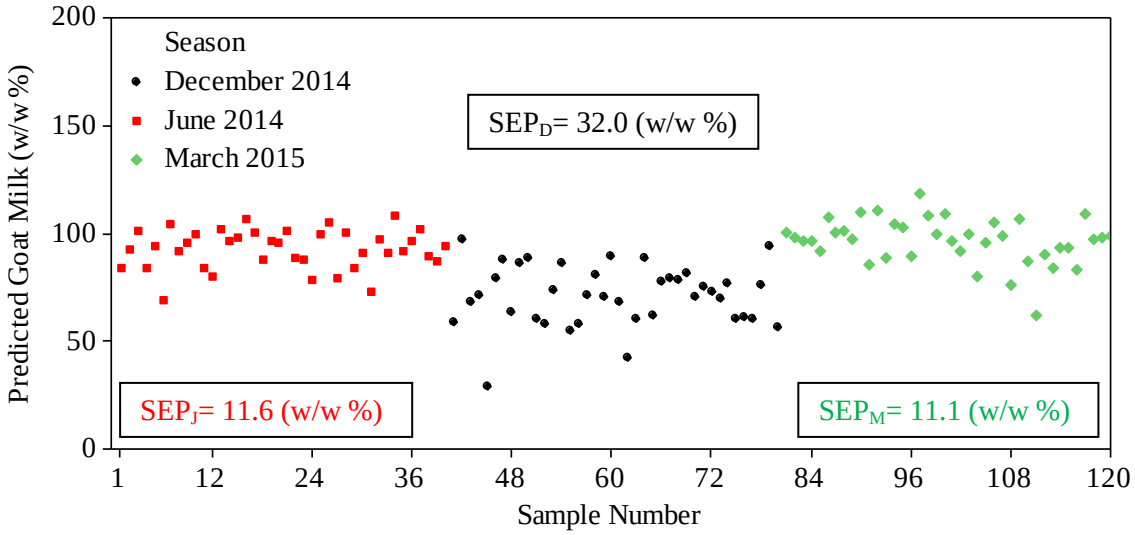

Figure 5.21. Predicted goat milk content of raw goat milk samples, estimated by GILS model by using the data set from March 2015 sampling period.

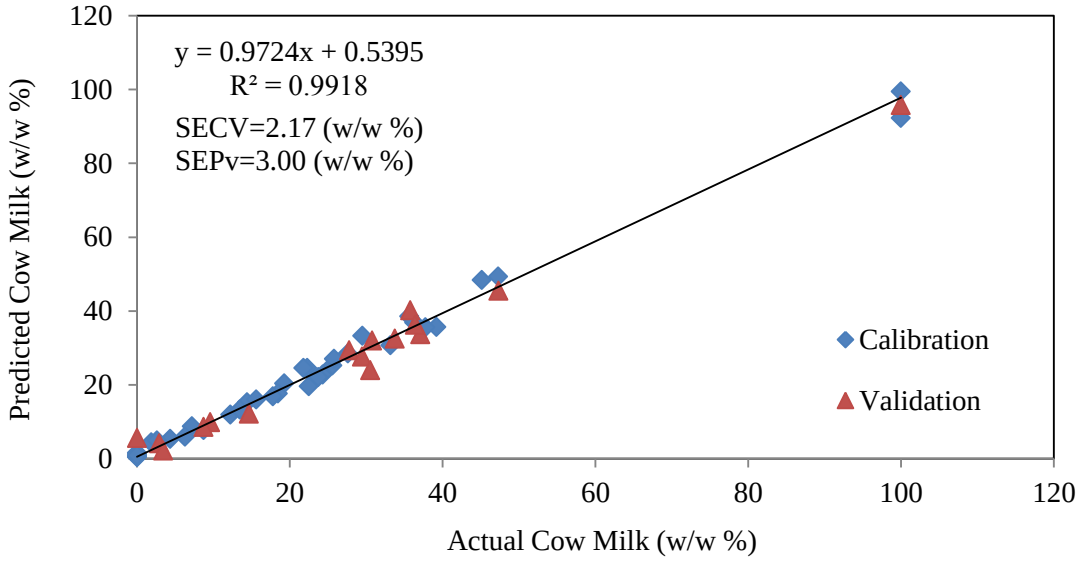

Figure 5.22. Actual versus predicted plot of cow milk content obtained from GILS model by using March 2015 data set.

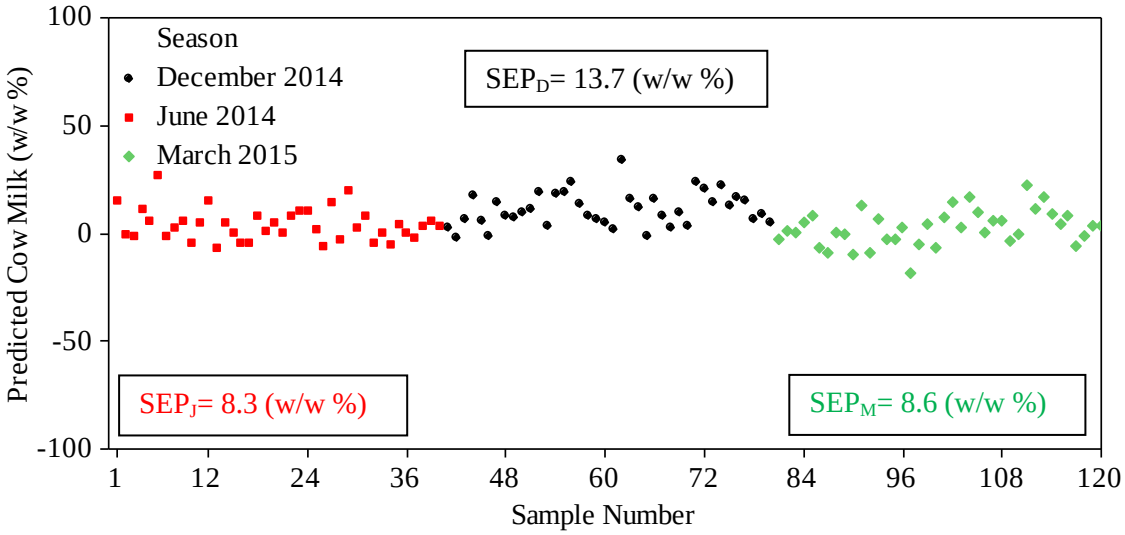

Figure 5.23. Predicted cow milk content of raw goat milk samples, estimated by GILS model by using the data set from March 2015 sampling period.

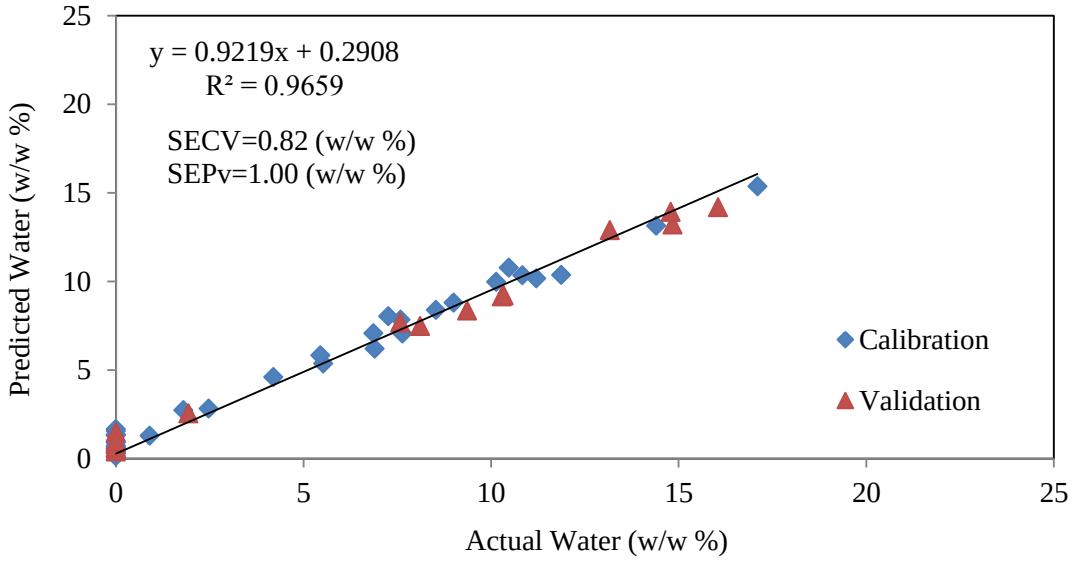

Figure 5.24. Actual versus predicted plot of water content obtained from GILS model by using March 2015 data set.

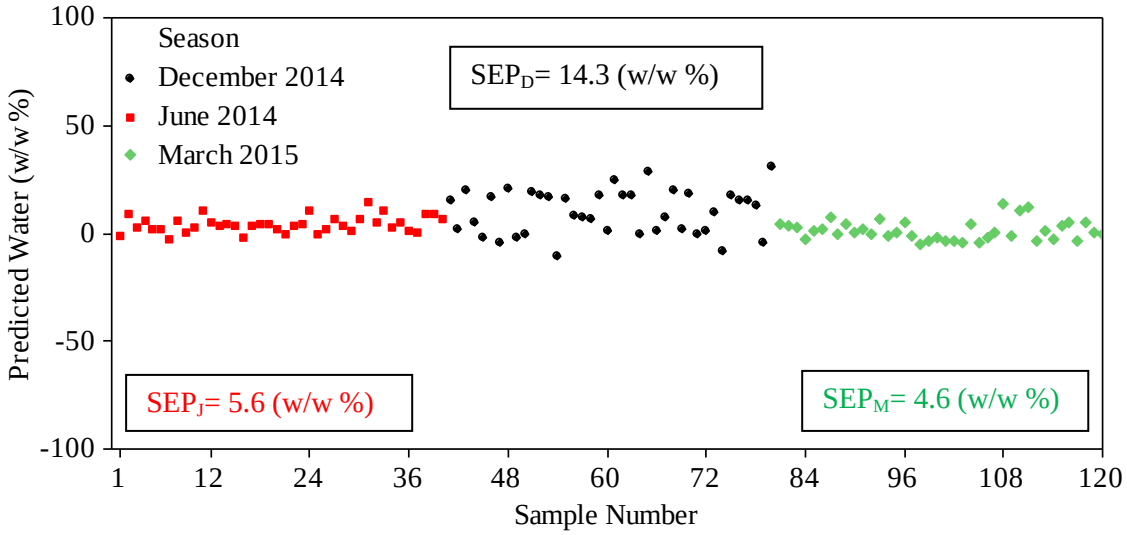

Figure 5.25. Predicted water content of raw goat milk samples, estimated by GILS model by using the data set from March 2015 sampling period.

As depicted in the figure 5.20, 5.22 and 5.24 correlation coefficients of GILS models were found to be, 0.9909, 0.9918 and 0.9659, respectively. On the other hand, SECV were varied between 0.82 and 2.26 (w/w %), and found to be acceptable. SEPv was calculated for independent validation sets ranged from 1.00 and 3.60 (w/w %). As the calculated parameters shown in the graphs, the models were found to be successful for validation sets. However, as indicated in the figure 5.21, 5.23 and 5.25, the seasonal variations caused unsatisfactory results for raw goat milk samples. As a conclusion, March 2015 model cannot be able to predict other seasons' raw goat milk samples.

As an overall assessment for all single GILS models, each season's models can not predict the other seasons' raw goat milk samples properly. Different contents of water in the samples may be the result of this situation. In order to improve the prediction of raw goat milk samples, seasons should be combined, as mentioned in the next chapter.

5.2.2. GILS Results for Binary Season Models

As the single models cannot be able to predict the raw goat milk samples properly, the binary mixtures of different seasons (June 2014-December 2014, June 2014-March 2015and December 2014-March 2015) were tested to improve the success of the model. These binary mixtures were generated from the single seasons' data.

Initially, the combination of June 2014-December 2014 season samples were used to set the model. Correlation graphs of predicted goat milk, cow milk and water contents obtained from GILS model results versus actual values are given in Figure 5.26, 5.28 and 5.30, respectively. Further, by use of that GILS models, goat milk, cow milk and water percentages were predicted in 120 milk samples and the obtained results are given in Figure 5.27, 5.29 and 5.31, respectively.

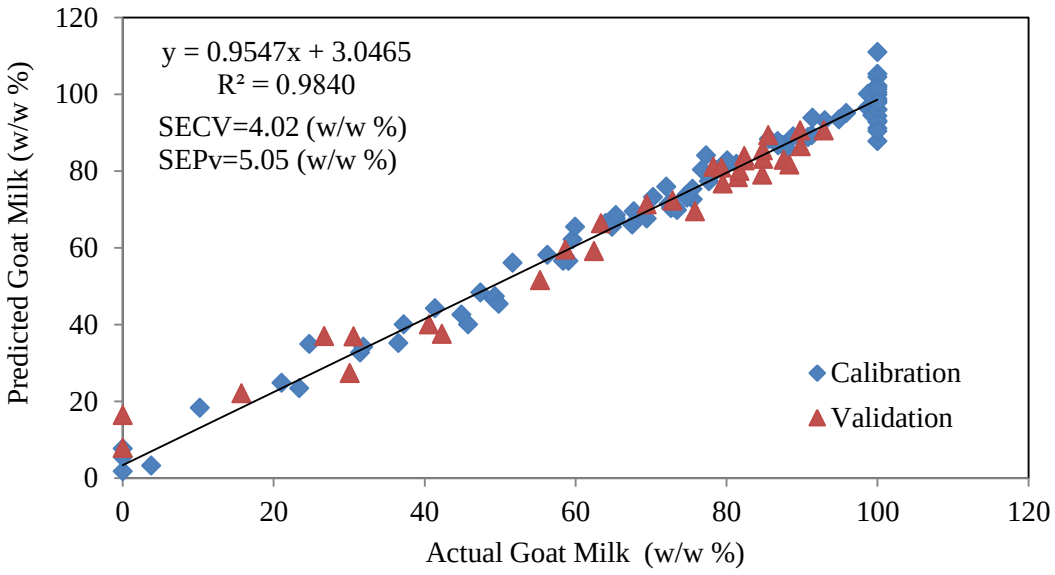

Figure 5.26. Actual versus predicted plot of goat milk content obtained from GILS model by using June 2014 and December 2014 combined data sets.

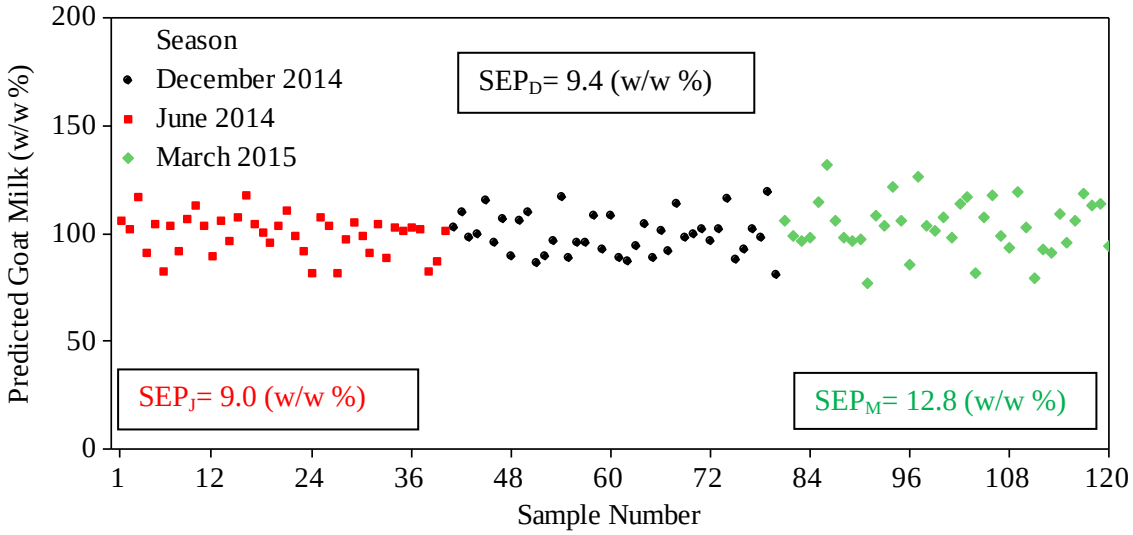

Figure 5.27. Predicted goat milk content of raw goat milk samples, estimated by GILS model by using the combined data set from June 2014 and December 2014 sampling periods.

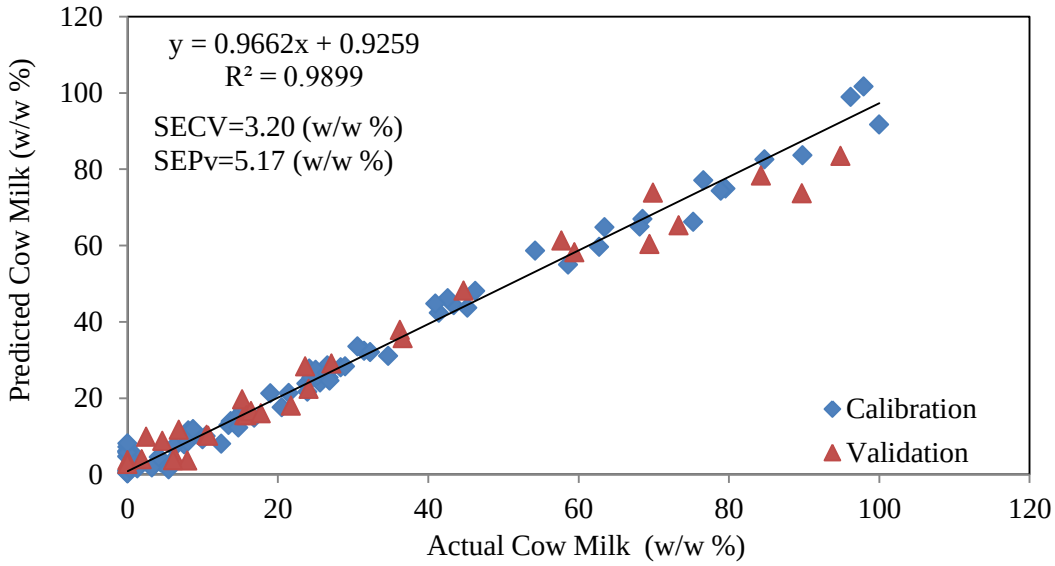

Figure 5.28. Actual versus predicted plot of cow milk content obtained from GILS model by using June 2014 and December 2014 combined data sets.

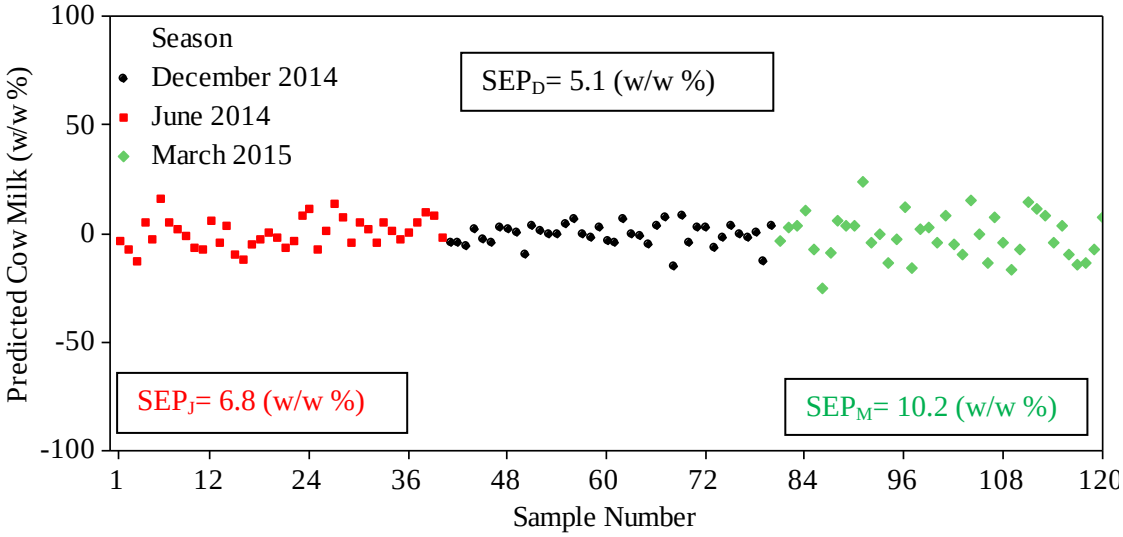

Figure 5.29. Predicted cow milk content of raw goat milk samples, estimated by GILS model by using the combined data set from June 2014 and December 2014 sampling periods.

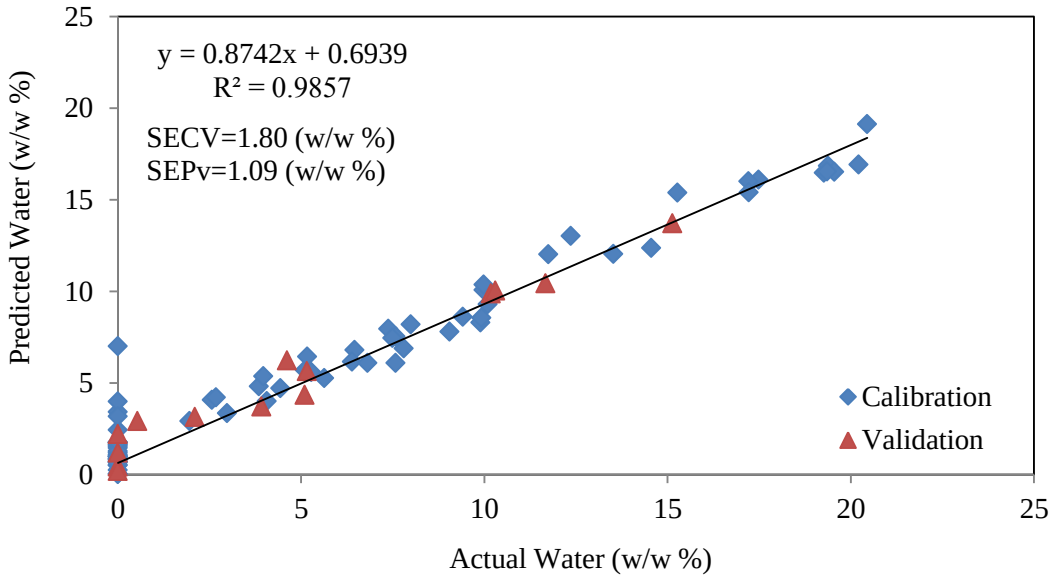

Figure 5.30. Actual versus predicted plot of water content obtained from GILS model by using June 2014 and December 2014 combined data sets.

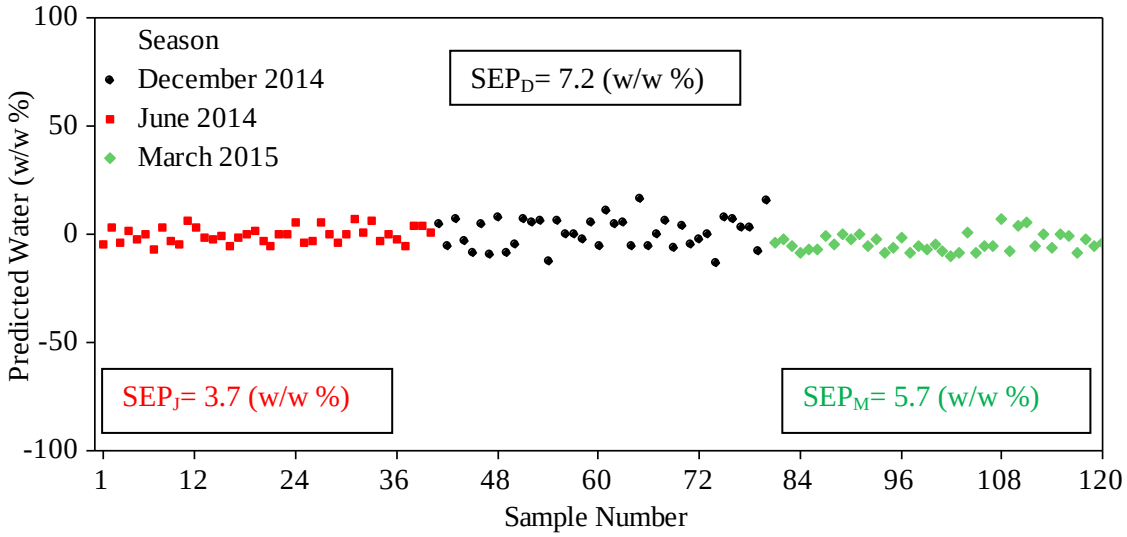

Figure 5.31. Predicted water content of raw goat milk samples, estimated by GILS model by using the combined data set from June 2014 and December 2014 sampling periods.

As it can be followed from the binary GILS model graphs of June 2014-December 2014, the correlation coefficients were found as 0.9840, 0.9899 and 0.9857. In addition, SECV and SEPv values of the model varied between 1.80 and 4.02 (w/w %) and 1.09 and 5.17 (w/w %) respectively.

The combination of June 2014-December 2014 data sets improved the prediction in June 2014 and December 2014 raw goat milk samples. Almost similar composition of June 2014 and March 2015 samples has resulted close SEP values (SEP$_J$ and SEP$_M$). Therefore, predictions of March 2015 raw goat milk samples were found to be acceptable.

Secondly, June 2014 and March 2015 data were combined for the binary combination and utilized to set the model. Correlation graphs of predicted goat milk, cow milk and water contents obtained from GILS model results versus actual values are given in Figure 5.32, 5.34 and 5.36, respectively. Sequentially, by use of that GILS models, goat milk, cow milk and water percentages were predicted in 120 milk samples and the obtained results are given in Figure 5.33, 5.35 and 5.37, respectively.

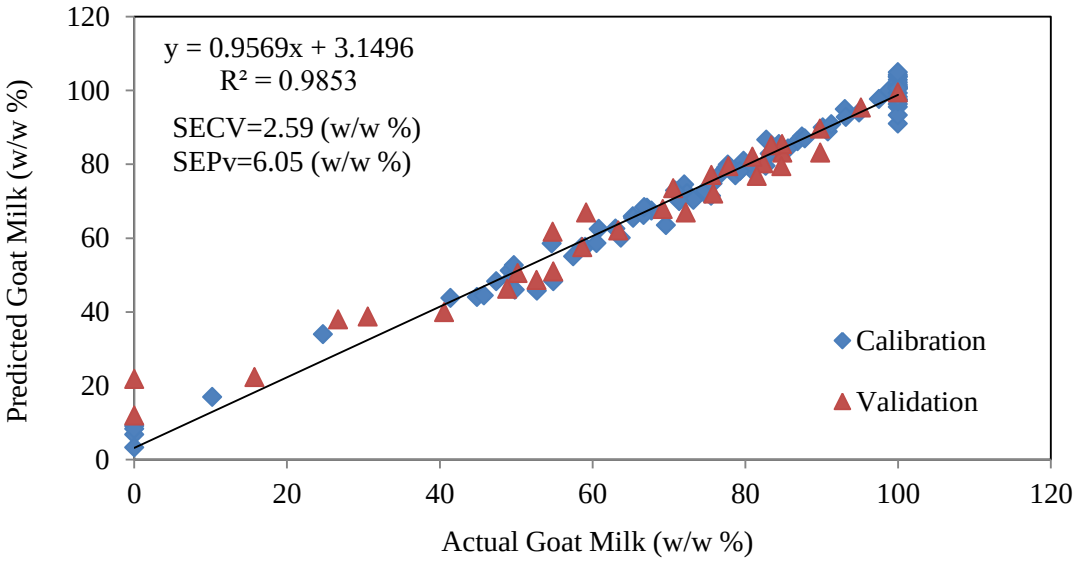

Figure 5.32. Actual versus predicted plot of goat milk content obtained from GILS model by using June 2014 and March 2015 combined data sets.

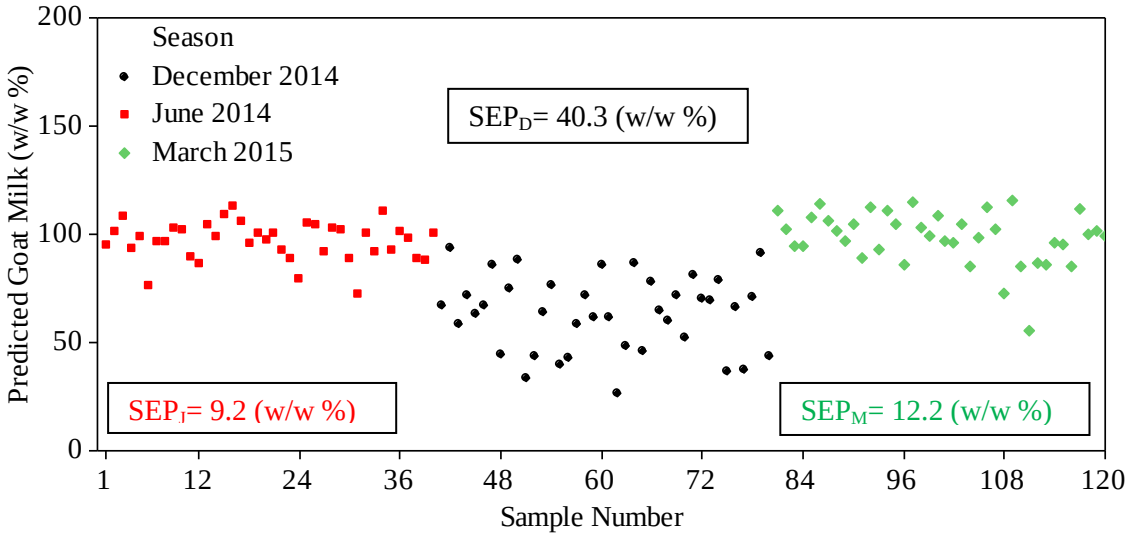

Figure 5.33. Predicted goat milk content of raw goat milk samples, estimated by GILS model by using the combined data set from June 2014 and March 2015 sampling periods.

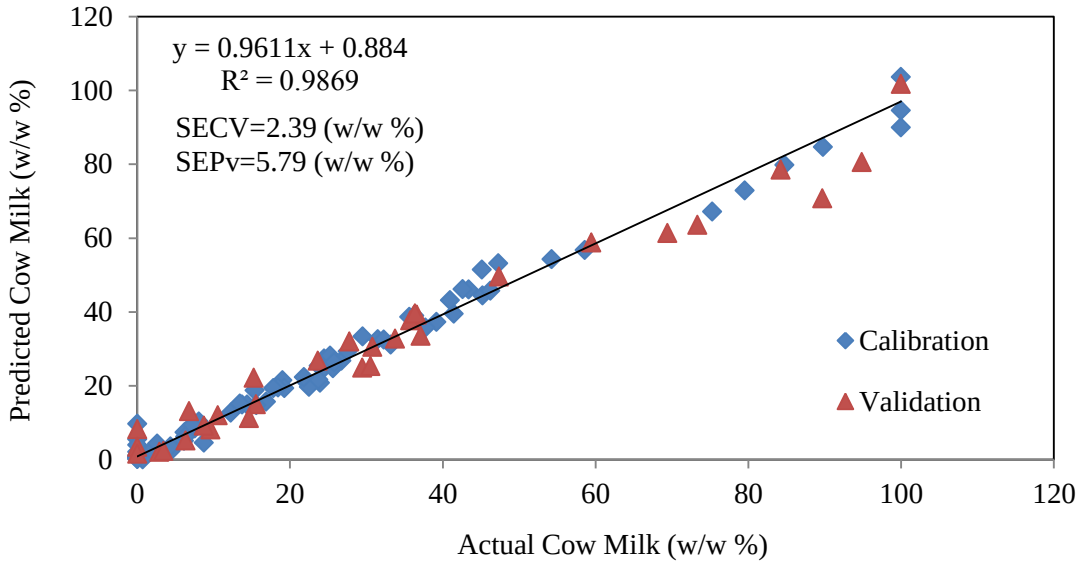

Figure 5.34. Actual versus predicted plot of cow milk content obtained from GILS model by using June 2014 and March 2015 combined data sets.

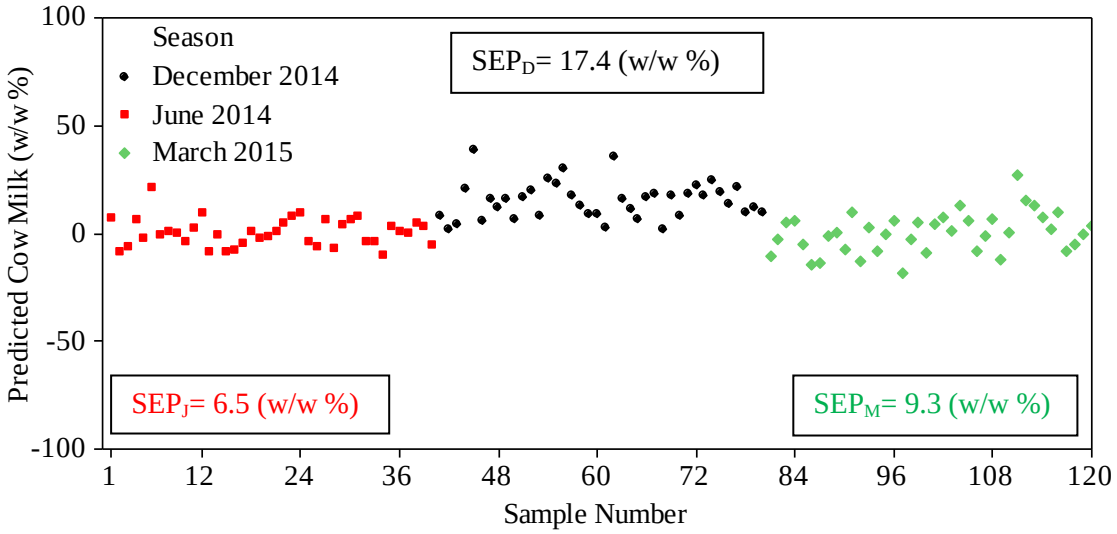

Figure 5.35. Predicted cow milk content of raw goat milk samples, estimated by GILS model by using the combined data set from June 2014 and March 2015 sampling periods.

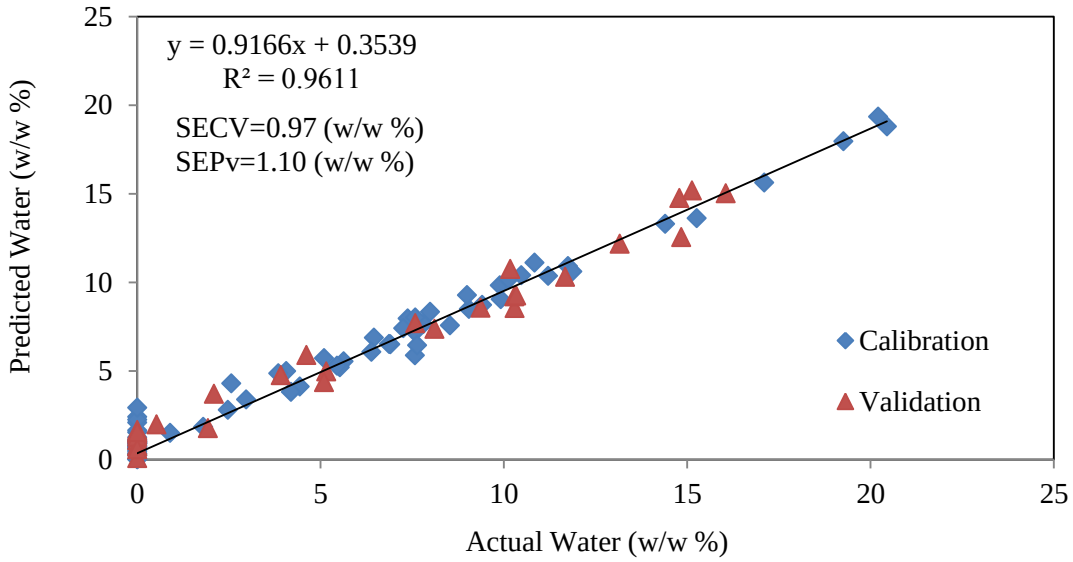

Figure 5.36. Actual versus predicted plot of water content obtained from GILS model by using June 2014 and March 2015 combined data sets.

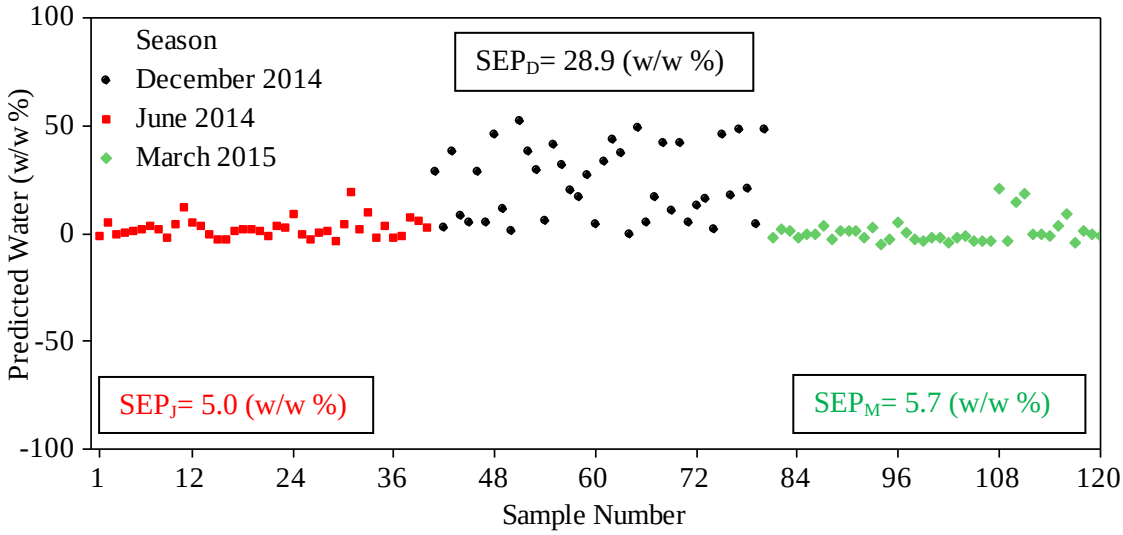

Figure 5.37. Predicted water content of raw goat milk samples, estimated by GILS model by using the combined data set from June 2014 and March 2015 sampling periods.

As it can be followed from the binary GILS model graphs of June 2014-March 2015, the correlation coefficients were found as 0.9853, 0.9869 and 0.9611. In addition, SECV and SEPv values of the model varied between 0.97 and 2.59 (w/w %) and 1.10 and 6.05 (w/w %) respectively.

The combination of June 2014-March 2015 data sets improved the prediction in June 2014 and March 2015 raw goat milk samples. However, similar improvement was not observed for December 2014 raw goat milk samples probably due to the lack of December 2014 data in the model.

Finally, December 2014 and March 2015 data were combined for the binary combination and utilized to set the model. Correlation graphs of predicted goat milk, cow milk and water contents obtained from GILS model results versus actual values are given in Figure 5.38, 5.40 and 5.42, respectively. Sequentially, by use of that GILS models, goat milk, cow milk and water percentages were predicted in 120 milk samples and the obtained results are given in Figure 5.39, 5.41 and 5.43, respectively.

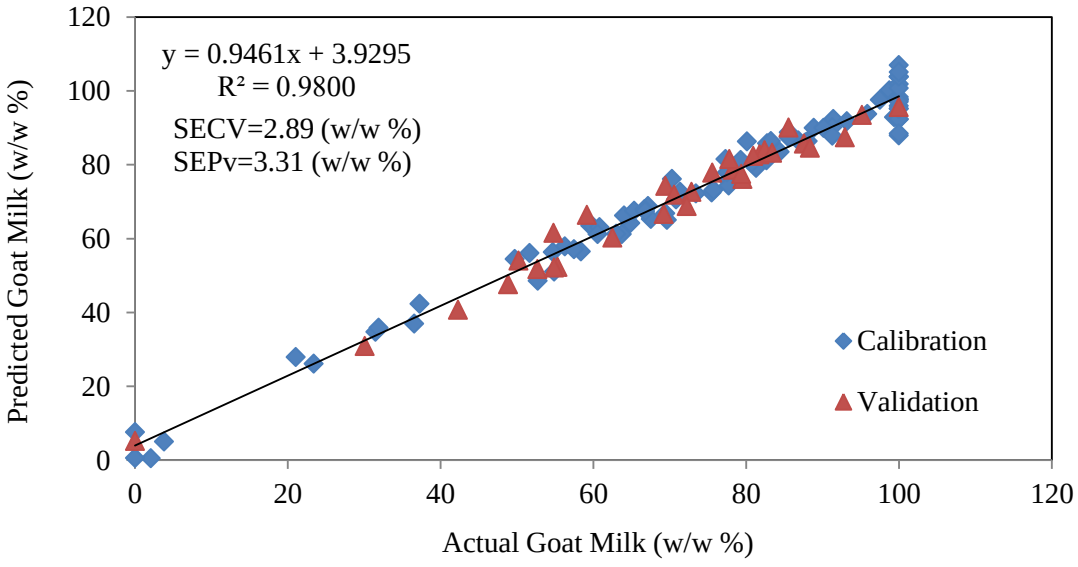

Figure 5.38. Actual versus predicted plot of goat milk content obtained from GILS model by using December 2014 and March 2015 combined data sets.

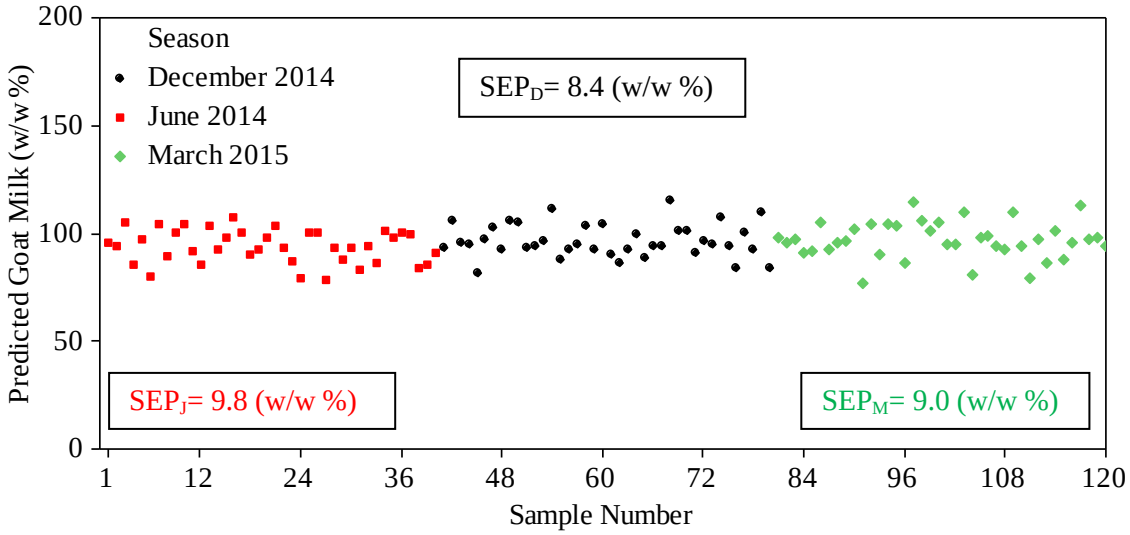

Figure 5.39. Predicted goat milk content of raw goat milk samples, estimated by GILS model by using the combined data set from December 2014 and March 2015 sampling periods.

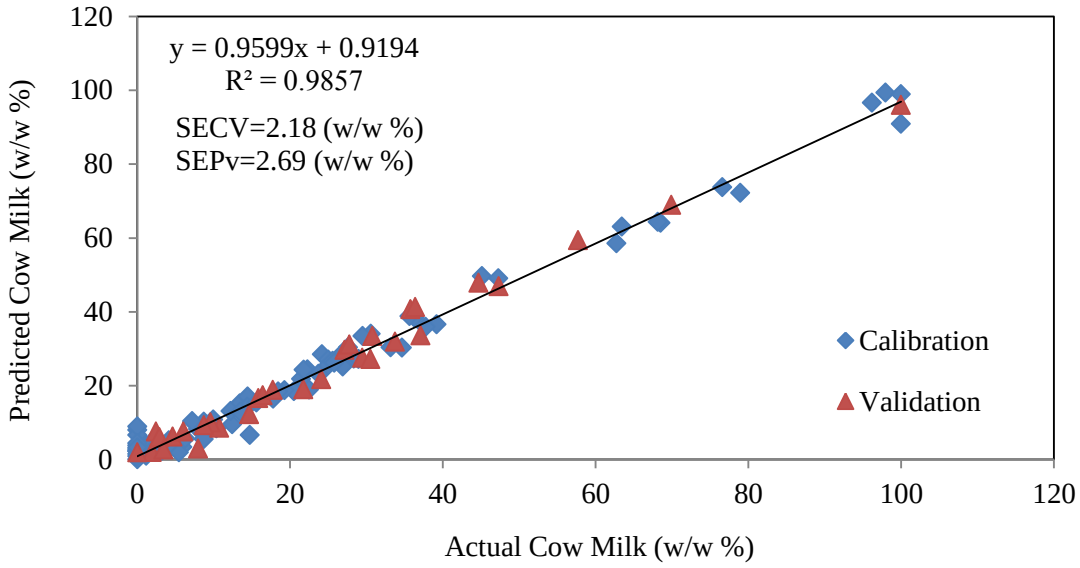

Figure 5.40. Actual versus predicted plot of cow milk content obtained from GILS model by using December 2014 and March 2015 combined data sets.

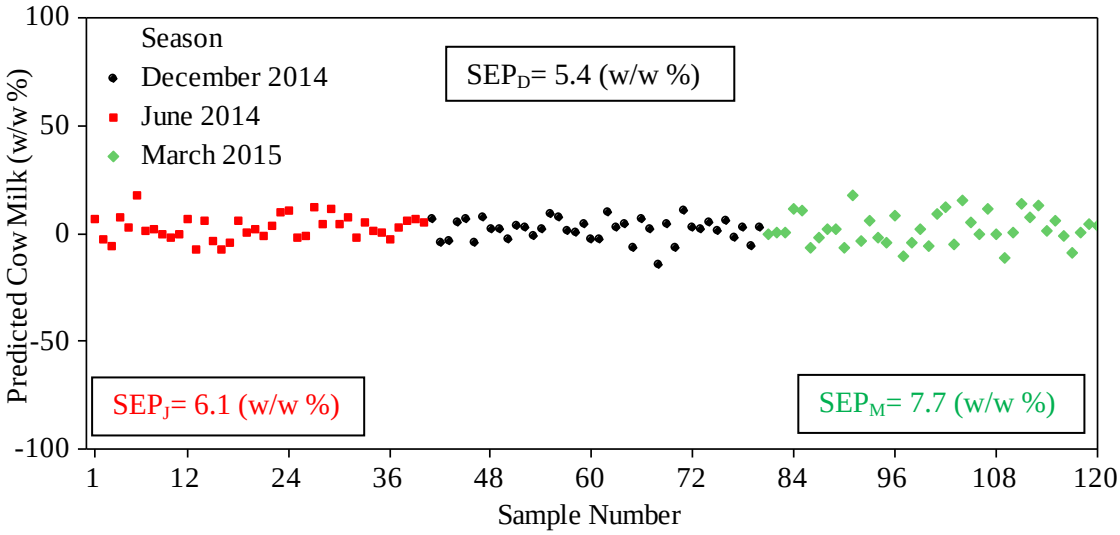

Figure 5.41. Predicted cow milk content of raw goat milk samples, estimated by GILS model by using the combined data set from December 2014 and March 2015 sampling periods.

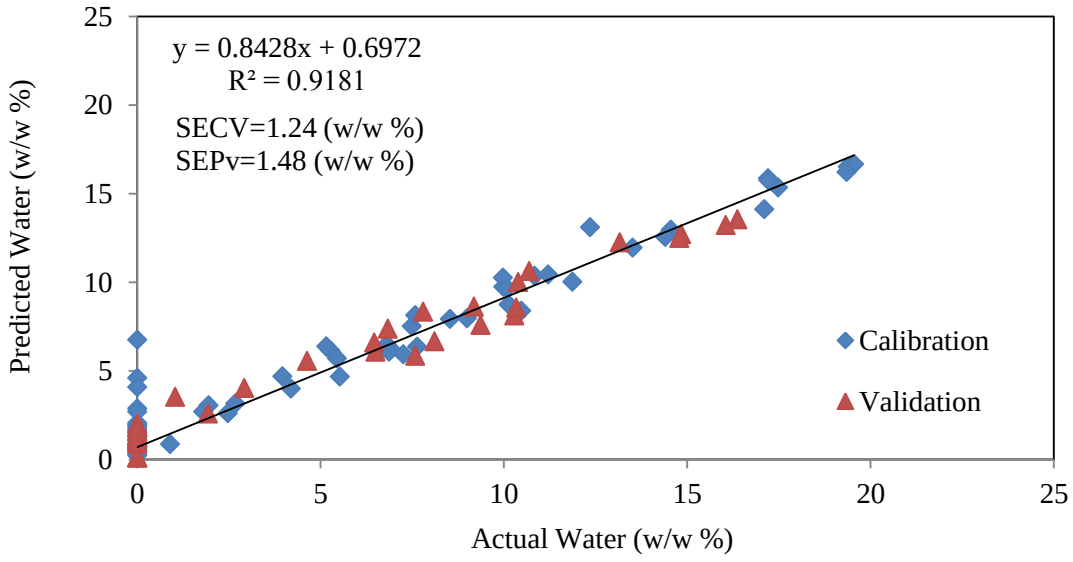

Figure 5.42. Actual versus predicted plot of water content obtained from GILS model by using December 2014 and March 2015 combined data sets.

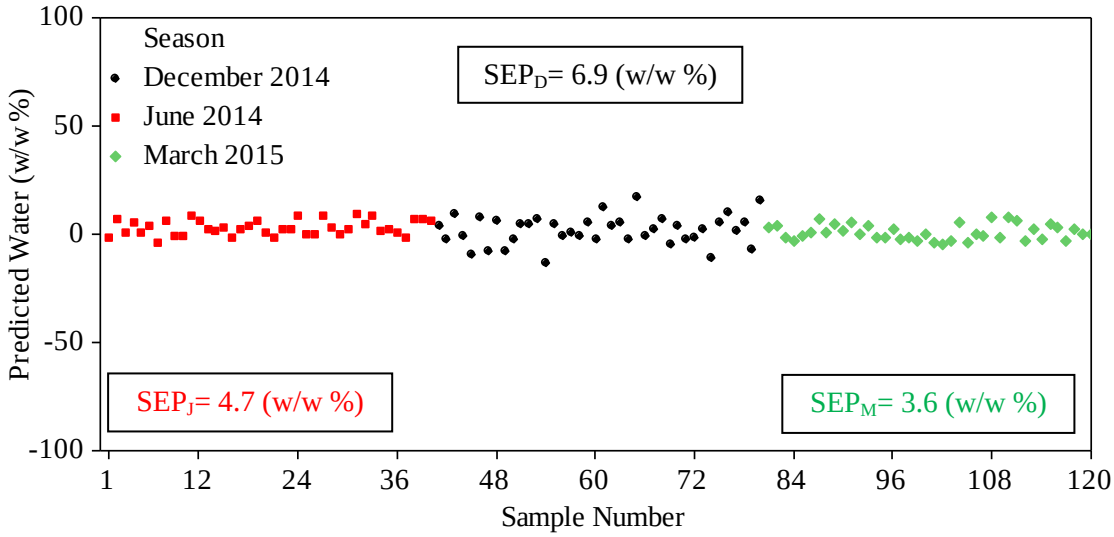

Figure 5.43. Predicted water content of raw goat milk samples, estimated by GILS model by using the combined data set from December 2014 and March 2015 sampling periods.

As it can be followed from the binary GILS model graphs of June 2014-March 2015, the correlation coefficients were found as 0.9800, 0.9857 and 0.9181. In addition, SECV and SEPv values of the model varied between 1.24 and 2.89 (w/w %) and 1.48 and 3.31 (w/w %) respectively.

The combination of December 2014-March 2015 data sets improved the prediction in June 2014, December 2014 and March 2015 raw goat milk samples. This improvement can be attributed to similar compositions of June 2014 and March 2015 samples, although June 2014 data was not included in the model.

As a result, binary models showed more successful results than their individual models. This improvement in prediction can be attributed to partial elimination of seasonal variation by combination of them with each other.

5.2.3. GILS Results for Ternary Model

Single and binary combination results showed that success of the model was affected by the seasonal variation. Accordingly, the ternary combination of three seasons data were used to set final model (June 2014-December 2014-March 2015) for eliminating seasonal variations.

Correlation graphs of predicted goat milk, cow milk and water contents obtained from GILS model results versus actual values are given in Figure 5.44, 5.46 and 5.48, respectively. Further, by use of that GILS models, goat milk, cow milk and water percentages were predicted in 120 milk samples and the obtained results are given in Figure 5.45, 5.47 and 5.49, respectively.

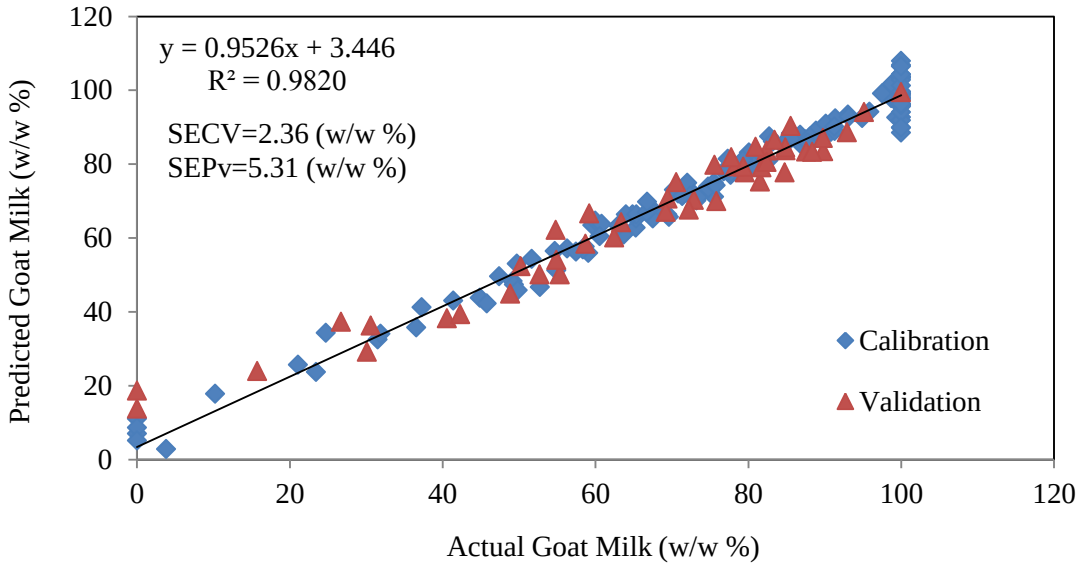

Figure 5.44. Actual versus predicted plot of goat milk content obtained from GILS model by using June 2014, December 2014 and March 2015 combined data sets.

.

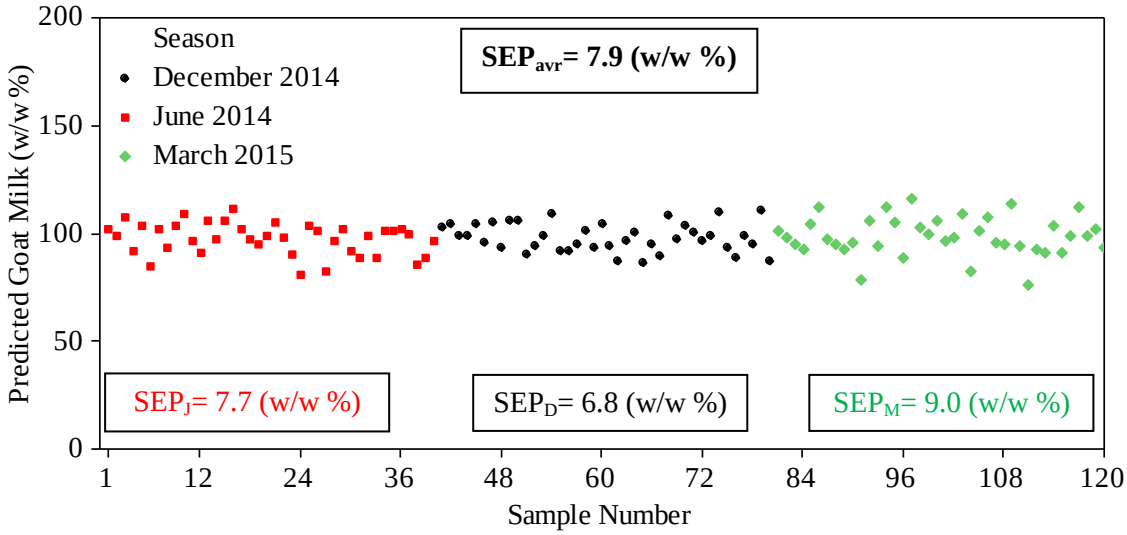

Figure 5.45. Predicted goat milk content of raw goat milk samples, estimated by GILS model by using the combined data set from June 2014, December 2014 and March 2015 sampling periods.

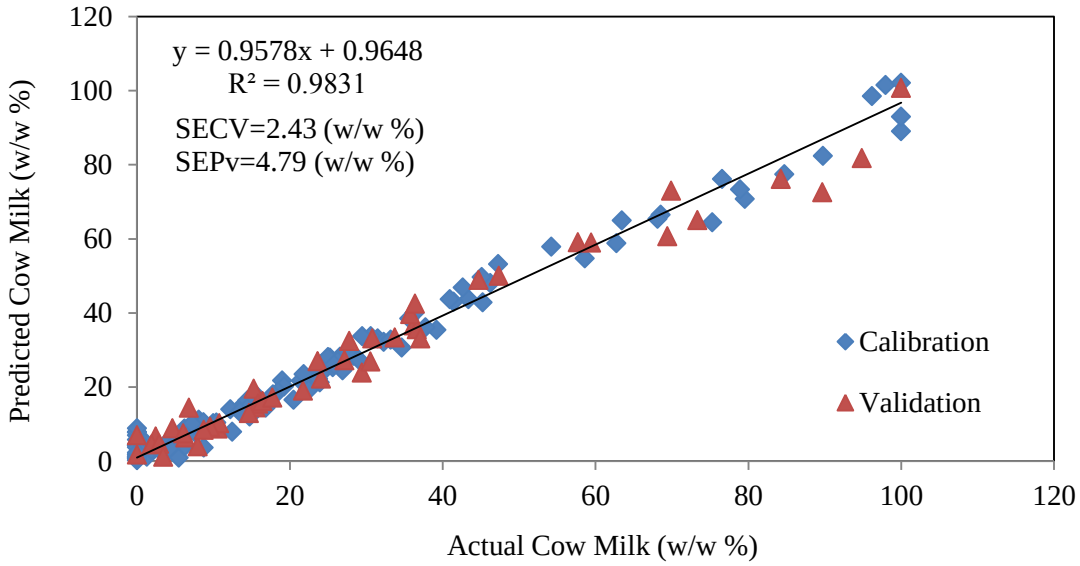

Figure 5.46. Actual versus predicted plot of cow milk content obtained from GILS model by using June 2014, December 2014 and March 2015 combined data sets.

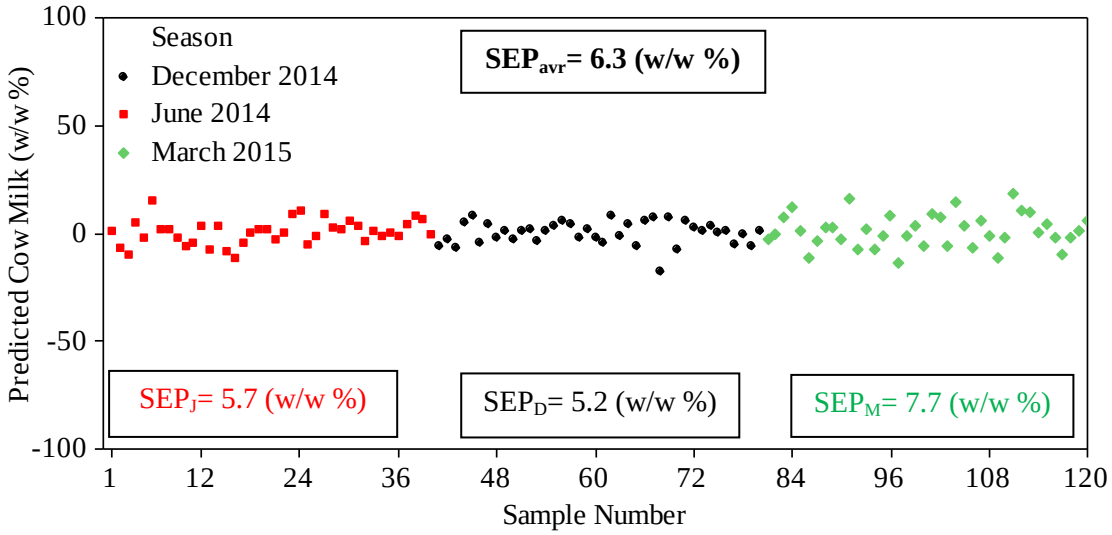

Figure 5.47. Predicted cow milk content of raw goat milk samples, estimated by GILS model by using the combined data set from June 2014, December 2014 and March 2015 sampling periods.

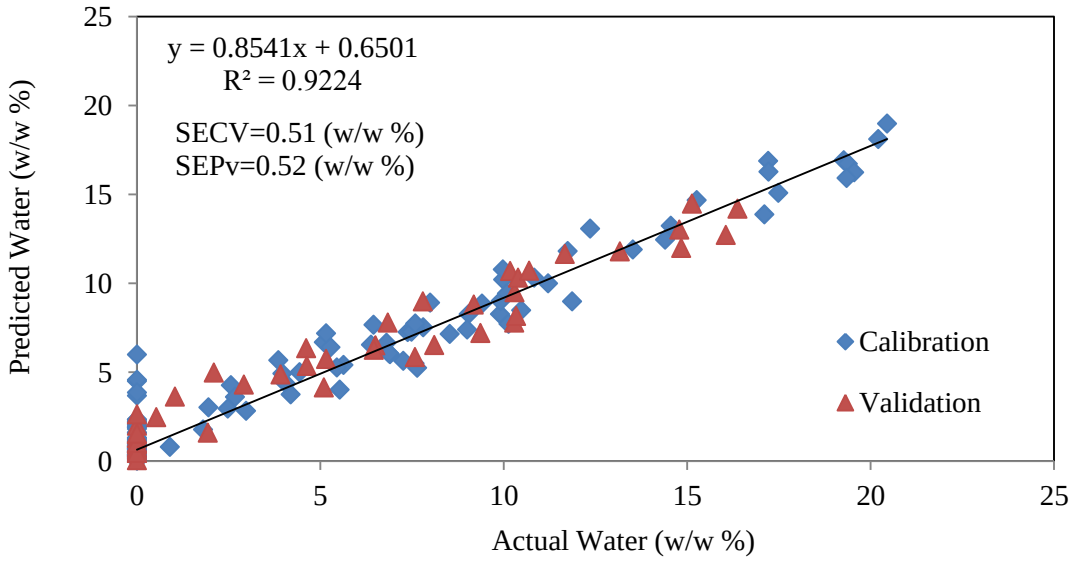

Figure 5.48. Actual versus predicted plot of water content obtained from GILS model by using June 2014, December 2014 and March 2015 combined data sets.

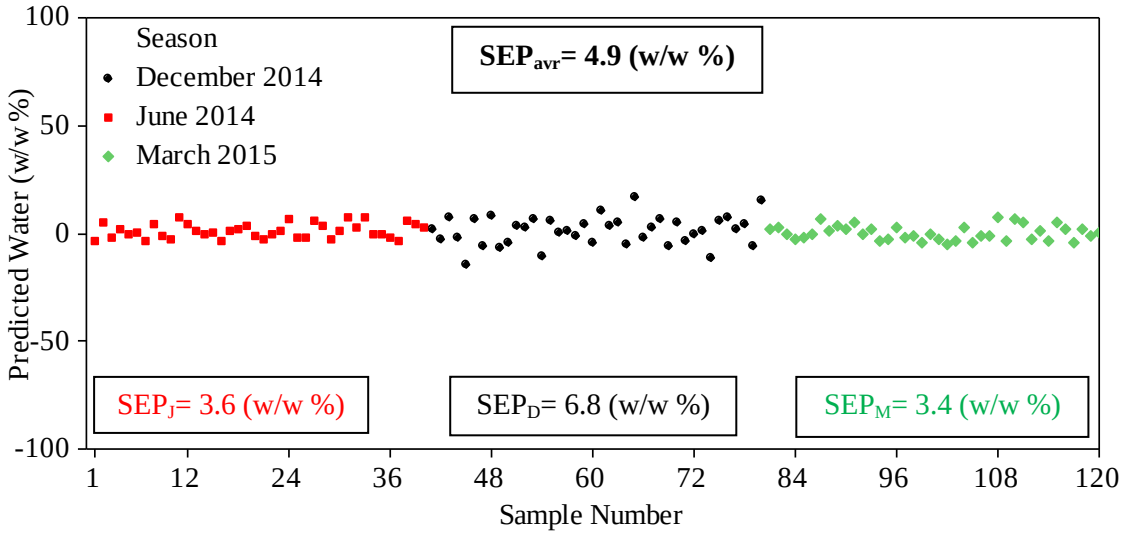

Figure 5.49. Predicted water content of raw goat milk samples, estimated by GILS model by using the combined data set from June 2014, December 2014 and March 2015 sampling periods.

As it can be seen from the binary GILS model graphs of June 2014-December 2014-March 2015, the correlation coefficients were found as 0.9820, 0.9831 and 0.9224. In particular, the regression coefficients for goat and cow milk were found higher than 0.9800 in a wide dynamic range (0-100%). In addition, SECV and SEPv values of the model varied between 0.51 and 2.43 (w/w %) and 0.52 and 5.31 (w/w %) respectively.

SEP values of raw goat milk samples showed that the success of the prediction was improved. The main success of the final model is to predict December 2014 raw goat milk samples accurately that has the most seasonal variations. However, higher SEP value of December 2014 raw goat milk samples (SEP$_D$= 6.8%) was found nearly doubled according to SEP$_J$ and SEP$_M$ (3.6 and 3.4%, respectively). This can be attributed to higher water content of December 2014 samples and/or the use of water as a blank sample in FTIR measurement.

As an overall assessment for all GILS models, correlation coefficients of single models are higher than the binary and ternary models. However, this situation is not showed that the predictions are better in single models. Moreover, standard error of prediction values for different seasons were found in a satisfactory range. In addition to these, low correlation coefficient values in binary and ternary models are a result of an increasing of variability.

5.2.4. PLS Results for Ternary Model

In order to compare the results obtained from the GILS model for ternary mixture samples (final model), PLS model was tested which is another multivariate calibration method. Before PLS analysis, noisy region (between 650 – 881 cm^{-1} and 1820 – 2391 cm^{-1}) was taken out. Thus, better results were observed, instead of full spectrum. Correlation graphs of predicted goat milk, cow milk and water contents obtained from PLS model results versus actual values are given in Figure 5.50, 5.52 and 5.54, respectively. Further, by use of that PLS models, goat milk, cow milk and water percentages were predicted in 120 milk samples and the obtained results are given in Figure 5.51, 5.53 and 5.55, respectively. It is important to emphasize that PLS model was tested just for ternary mixture samples, not for the single and binary mixtures due to the quite insufficient results when compared with GILS results.

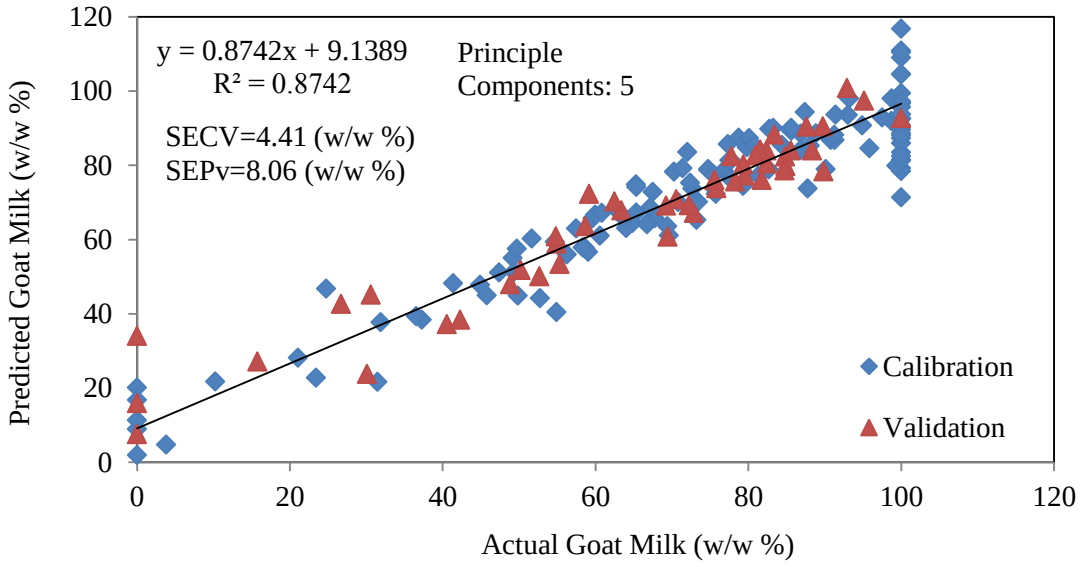

Figure 5.50. Actual versus predicted plot of goat milk content obtained from PLS model by using June 2014, December 2014 and March 2015 combined data sets.

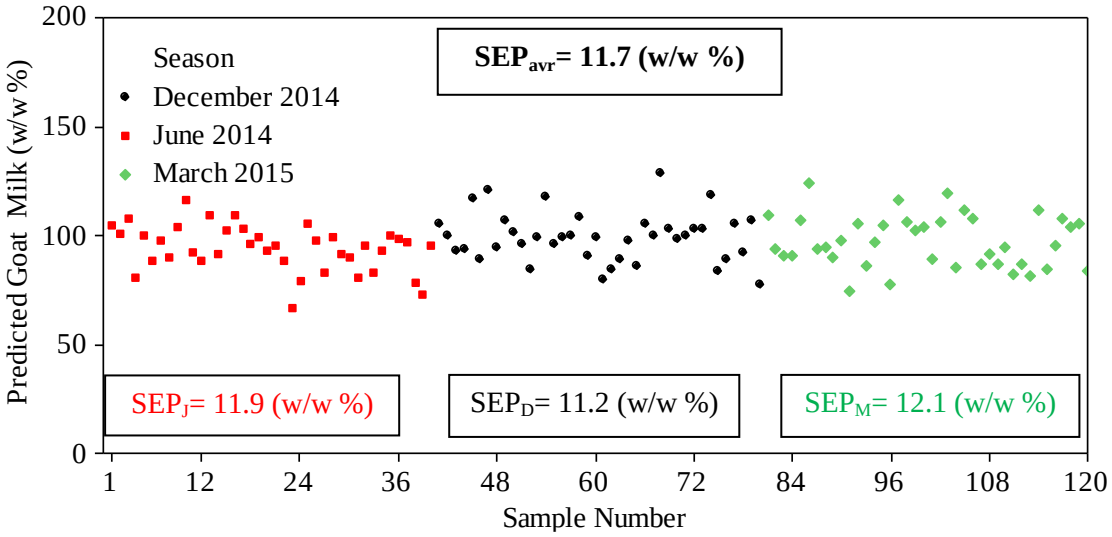

Figure 5.51. Predicted goat milk content of raw goat milk samples, estimated by PLS model by using the combined data set from June 2014, December 2014 and March 2015 sampling periods.

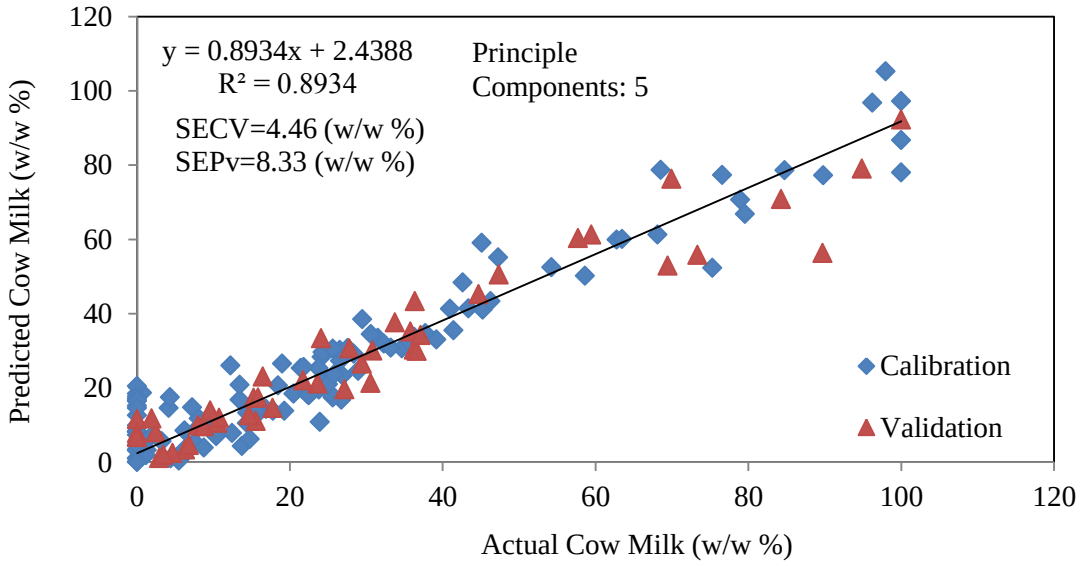

Figure 5.52. Actual versus predicted plot of cow milk content obtained from PLS model by using June 2014, December 2014 and March 2015 combined data sets.

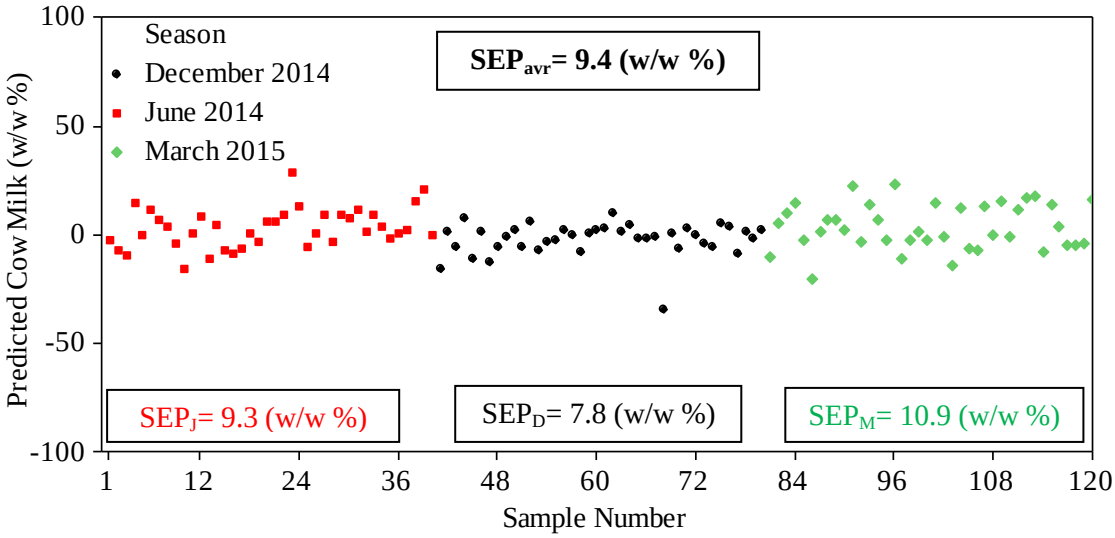

Figure 5.53. Predicted cow milk content of raw goat milk samples, estimated by PLS model by using the combined data set from June 2014, December 2014 and March 2015 sampling periods.

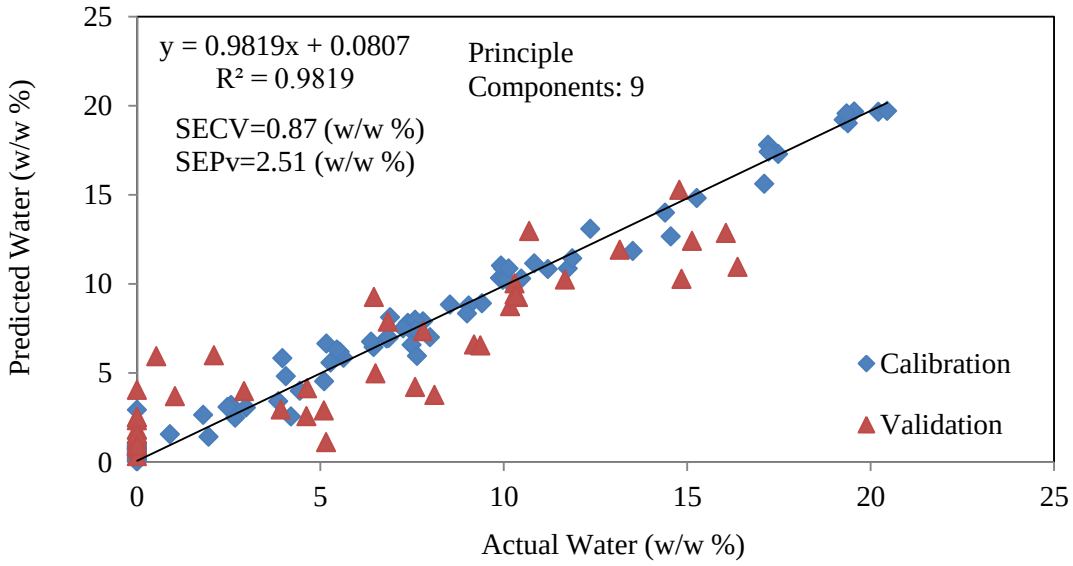

Figure 5.54. Actual versus predicted plot of water content obtained from PLS model by using June 2014, December 2014 and March 2015 combined data sets.

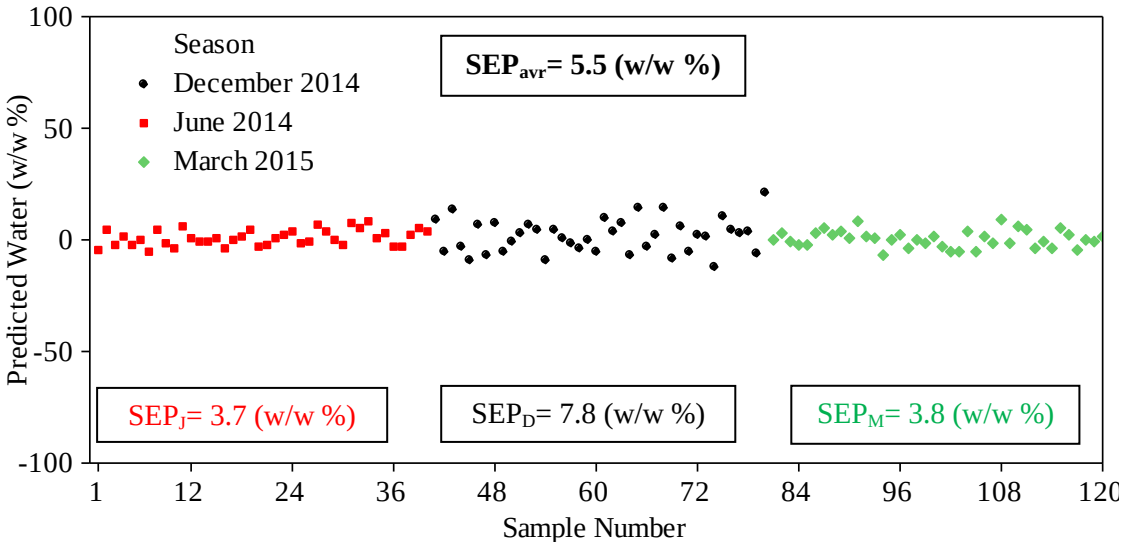

Figure 5.55. Predicted water content of raw goat milk samples, estimated by PLS model by using the combined data set from June 2014, December 2014 and March 2015 sampling periods.

From the ternary PLS model graphs of June 2014-December 2014-March 2015 it is apparent that the correlation coefficients for goat milk, cow milk and water were found as 0.8742, 0.8934 and 0.9819, respectively. In addition, SECV and SEPv values of the model varied between 0.87 and 4.46 (w/w %) and 2.51 and 8.33 (w/w %) respectively. The results obtained with this model were found nearly doubled form GILS results according to SECV and SEP_V values. This can be explained by the genetic algorithm of GILS method chooses the data in the best compatible wavenumber while modeling from the parameters of spectral data, unlike PLS. In order to allow easier comparison, correlation coefficient values obtained as a result of 8 different scenarios from the models and their SECV and SEP (SEP_V, SEP_J, SEP_D, SEP_M, SEP_{avr}) values are given in Table 5.1. Predicted goat milk, cow milk and water contents with the 8 different models are given in Table A.1, A.2 and A.3, respectively, in Appendix A.

5.3. Principle Component Results

Results of principle components analysis (PCA) for raw goat milk samples from each season were illustrated in Figure 5.56, 5.57 and 5.58. In addition to single season PCA analysis for each season a complete PCA was performed for the three seasons samples in order to observe possible clustering among seasons, especially for the

samples from December 2014 season. For the plots belonging to individual seasons, only the first two principle components were used while in the figure 5.59 containing all seasons the first three principle components were used.

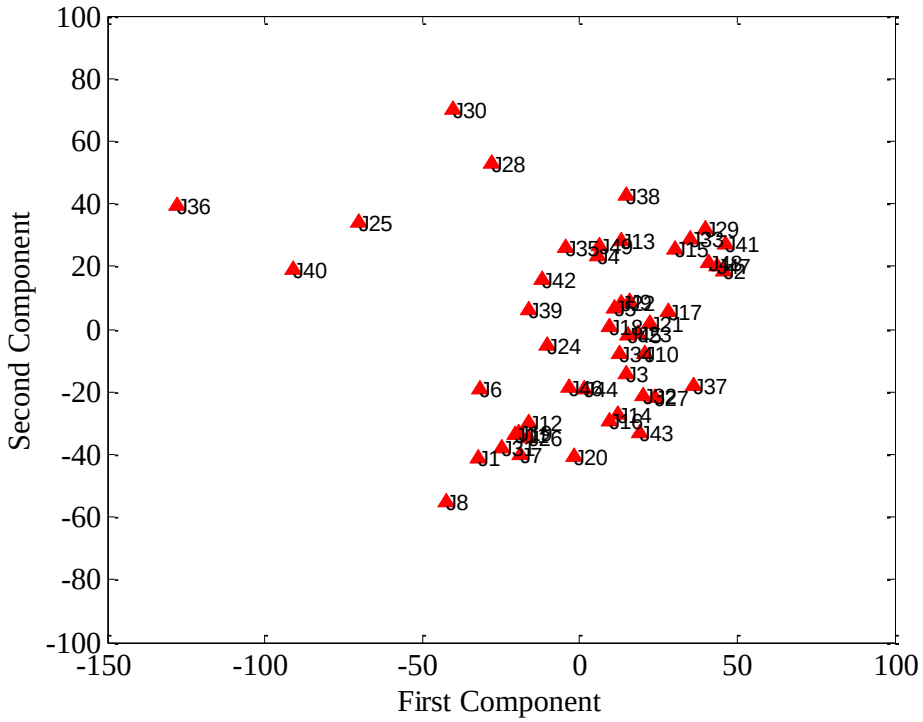

Figure 5.56. The scores plot of the first component (PC1) versus the second component (PC2) for 50 raw goat milk samples from June 2014 season using FTIR spectra.

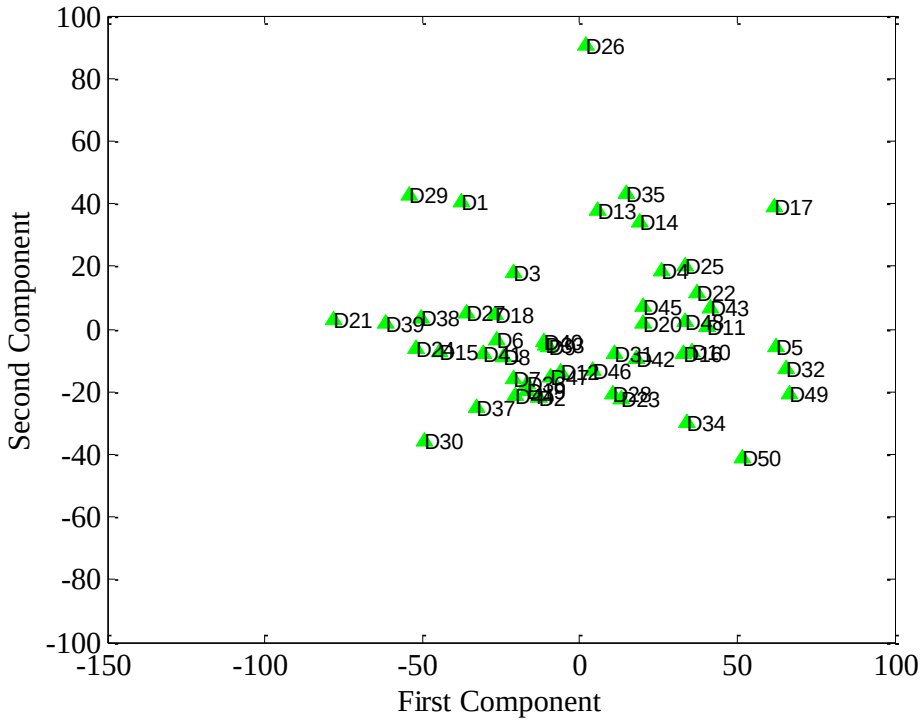

Figure 5.57. The scores plot of the first component (PC1) versus the second component (PC2) for 50 raw goat milk samples from December 2014 season using FTIR spectra.

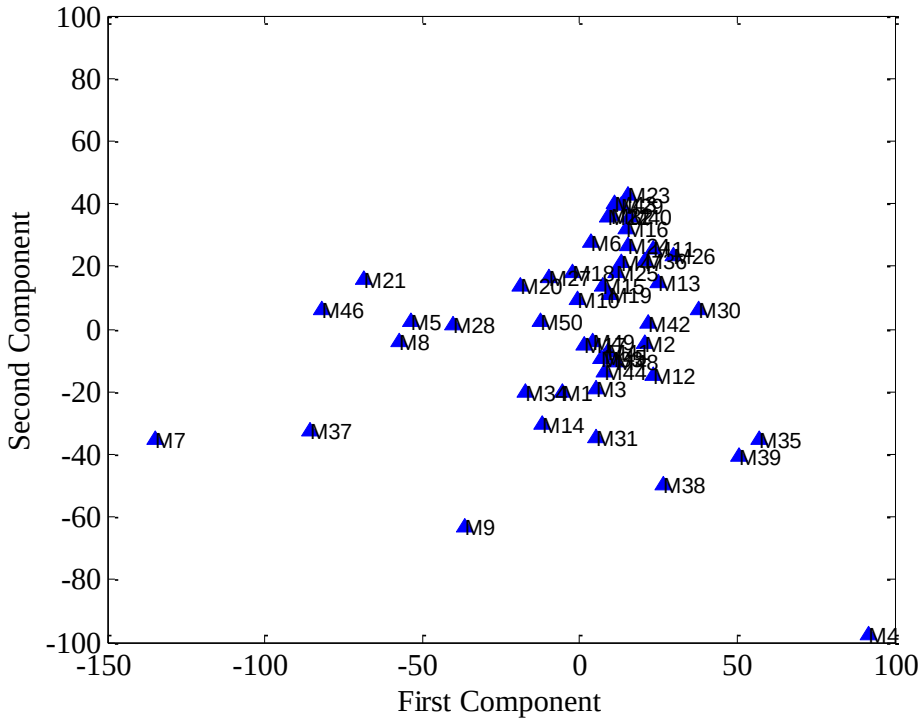

Figure 5.58. The scores plot of the first component (PC1) versus the second component (PC2) for 50 raw goat milk samples from March 2015 season using FTIR spectra.

For the scores plot of June 2014 season, the first principle component explains 11.6% of the total variance and the second component explains 9.7% of variance. For December 2014 season, 13.3% of overall variability was explained by the first PC and 8.8% was explained by the second PC. For March 2015 season, the first PC covers 13.8% of the total variance and the second PC covers 10.5% of the total variance.

In order to make these seasons comparable the axis limits have been set to same values that is -150 to 100 for the first PC and -100 to 100 for the second PC. From the individual plots it is possible to say that each season formed a cluster however there are few samples that appear to be slightly away from the cluster center compared to the most of the group members. When the spectra of these distant samples were examined, the difference appears to be in between 3500-4000 cm^{-1} wavenumber range which corresponds to the water content. In the March 2015 season, however, the 4^{th} sample is relatively further due to the difference in the range corresponding to fingerprint region.

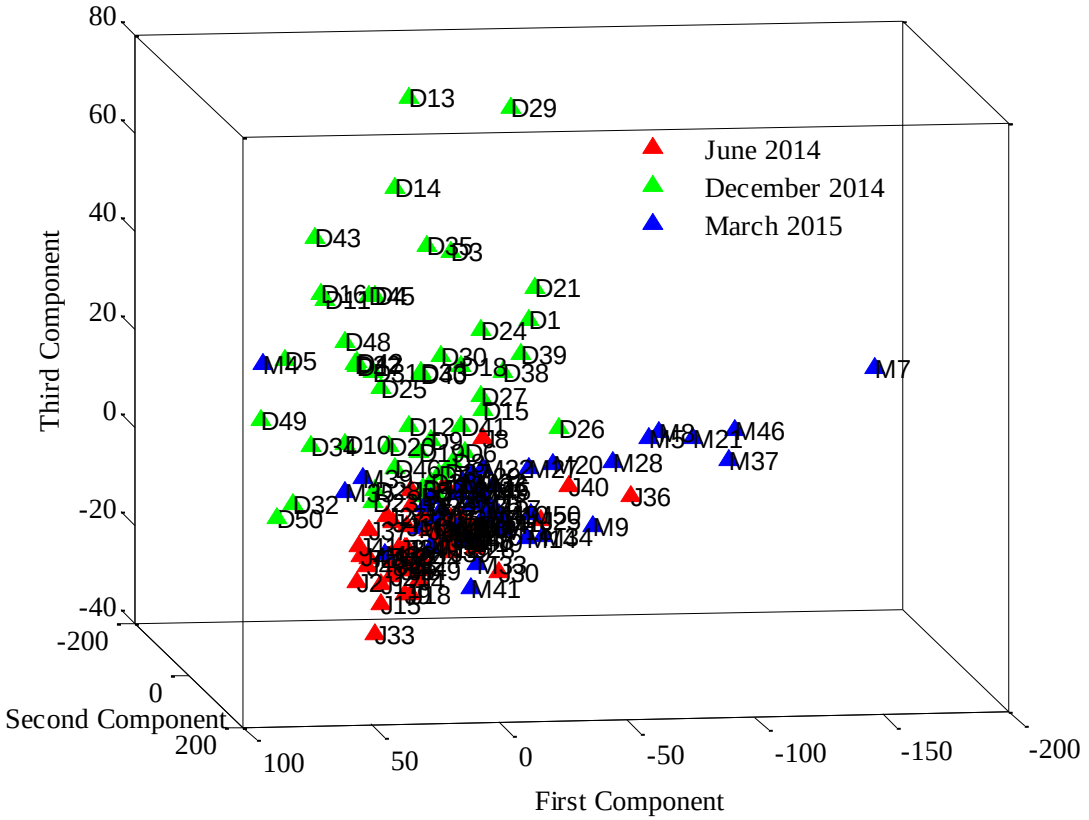

Figure 5.59. The 3 dimensional scores plot of the PC1, PC2 versus PC3 for 150 raw goat milk samples from June 2014, December 2014 and March 2015 using FTIR spectra.

In the Figure 5.59 all the seasons are combined and their scores were plotted using the first 3 PCs, where the first, second and third PCs contain 8.6%, 7.3%, 4.8% (cumulative 20.7%) of the total variance respectively, in order to see whether there is a clustering among seasons. While the two seasons, June 2014 and March 2015 seasons, are very close which indicates a similar content, December 2014 season appears separated from the other seasons in the Z axis that is the third principle component. This situation was also mentioned in the single season model results and binary season model results. As a result, the PCA score plots support the unusual predictions for December 2014 samples of single and binary season model results.

Table 5.1. The statistical parameters obtained from all GILS and PLS models.

Scenario		R^2	SECV	SEPv	SEP_J	SEP_D	SEP_M	SEP_{avr}
	Unit		(w/w %)	(w/w %)	(w/w %)	(w/w %)	(w/w %)	(w/w %)
June 2014 (GILS)	Goat	0.9940	2.15	5.77	12.60	67.60	16.40	40.82
	Cow	0.9930	2.23	5.87	8.10	30.90	12.40	19.77
	Water	0.9860	0.51	0.52	6.60	42.20	8.90	25.16
December 2014 (GILS)	Goat	0.9890	3.12	2.18	84.20	9.80	81.70	67.97
	Cow	0.9890	2.49	1.96	17.50	8.80	21.90	16.97
	Water	0.9380	1.80	1.12	4.30	6.80	6.20	5.87
March 2015 (GILS)	Goat	0.9910	2.26	3.60	11.60	32.00	11.10	20.69
	Cow	0.9920	2.17	3.00	8.30	13.70	8.60	10.49
	Water	0.9660	0.82	1.00	5.60	14.30	4.60	9.28
June 2014-December 2014 (GILS)	Goat	0.9840	4.02	5.05	9.00	9.40	12.80	10.54
	Cow	0.9900	3.20	5.17	6.80	5.10	10.20	7.67
	Water	0.9860	1.80	1.09	3.70	7.20	5.70	5.70
June 2014-March 2015 (GILS)	Goat	0.9860	2.59	6.05	9.20	40.30	12.20	24.90
	Cow	0.9870	2.39	5.79	6.50	17.40	9.30	11.98
	Water	0.9610	0.97	1.10	5.00	28.90	5.70	17.24
December 2014-March 2015 (GILS)	Goat	0.9800	2.89	3.31	9.80	8.40	9.00	9.07
	Cow	0.9860	2.18	2.69	6.10	5.40	7.70	6.46
	Water	0.9180	1.24	1.48	4.70	6.90	3.60	5.23
June 2014-December 2014-March 2015 (GILS)	Goat	0.9820	2.36	5.31	7.70	6.80	9.00	7.90
	Cow	0.9830	2.43	4.79	5.70	5.20	7.70	6.30
	Water	0.9220	0.51	0.52	3.60	6.80	3.40	4.90
June 2014-December 2014-March 2015 (PLS)	Goat	0.8742	4.41	8.06	11.90	11.20	12.10	11.70
	Cow	0.8934	4.46	8.33	9.30	7.80	10.90	9.40
	Water	0.9819	0.87	2.51	3.70	7.80	3.80	5.50

CHAPTER 6

CONCLUSION

In this study, a new method which is based on Fourier Transform Infrared Spectroscopy (FTIR) coupled with chemometric multivariate calibration was developed for the determination of goat milk adulteration with cow milk. Two different multivariate calibration methods namely Genetic Inverse Least Squares (GILS) and Partial Least Squares (PLS) were used for the development of multivariate calibration models.

Quite successful results were obtained both for goat milk and adulterants (cow milk and water) by using GILS, although there were significant variabilities among the different season milk samples due to the seasonal changes. On the other hand, GILS models for single season data sets were unable to predict raw goat milk samples from the seasons which were not used in the calibration sets. In order to improve the prediction of raw goat milk samples, binary and ternary combinations of the seasons were also used. The results of these binary and ternary season combinations demonstrated that the raw goat milk samples from the seasons included in the binary and ternary combinations in the calibration sets were successfully predicted. Multivariate calibration modeling with PLS was only performed with the ternary combination of the seasons, because of the unsatisfactory results for the raw goat milk samples coming from the seasons that were not taken into account. GILS models gave more satisfactory results than PLS models. This is because, while genetic algorithm in GILS selects data in the most compatible wavenumbers with parameters, there is no such option in the PLS algorithm.

In summary, within the scope of this thesis, a rapid and simple molecular spectroscopy based analytical method was developed for the determination of goat milk adulteration with cow milk by the use of chemometric multivariate calibration methods. This study could be considered as a milestone for the other type of dairy product adulteration studies such as goat cheese and yoghurt where the determination of possible adulterants is more crucial issue.

REFERENCES

Arpakçı, T. 2013. "Multivariate Statistical Optimization of Enzyme Immobilization Onto Solid Matrix Using Central Composite Design."M.Sc., İzmir Institute of Technology

Beebe, K. R., R. J. Pell, and M. B. Seasholtz. 1998. *Chemometrics: A Practical Guide*: A Wiley-Interscience: John Wiley & Sons, Inc.

Boyd, S., and J. Kirkwood. 2011. "Quantitative Analysis using ATR-FTIR Spectroscopy." https://www.agilent.com/cs/library/applications/si-1374.pdf.

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

Cartoni, Giampaolo, Franco Coccioli, Renata Jasionowska, and Maurizio Masci. 1999. "Determination of cows' milk in goats' milk and cheese by capillary electrophoresis of the whey protein fractions." *Journal of Chromatography A* 846 (1–2):135-141. doi: http://dx.doi.org/10.1016/S0021-9673(98)01032-2.

Cheng, C. 2005. "Asymmetrical interleaver structure based on the modified Michelson interferometer." *Journal of optical Engineering* 44 (11). doi: doi:10.1117/1.2126586.

Dias, L. A., A. M. Peres, A. C. A. Veloso, F. S. Reis, M. Vilas-Boas, and A. A. S. C. Machado. 2009. "An electronic tongue taste evaluation: Identification of goat milk adulteration with bovine milk." *Sensors and Actuators B: Chemical* 136 (1):209-217. doi: 10.1016/j.snb.2008.09.025.

Gayete, J.F, M Guaerdia, and S Garrigues. 2006. "Attenuated total reflectance infrared determination of sodium nitrilotriacetate in alkaline liquid detergents." *Talanta* 70 (4):870-875. doi: 10.1016/j.talanta.2006.02.007.

Gilbert, R. J., R. Goodacre, A. N. Woward, and D.B. Kell. 1997. *Analytical Chemistry*. Vol. 69.

Golinelli, L. P., A. C. Carvalho, R. S. Casaes, C. S. C. Lopes, R. Deliza, V. M. F. Paschoalin, and J. T. Silva. 2014. "Sensory analysis and species-specific PCR detect bovine milk adulteration of frescal (fresh) goat cheese." *Journal of Dairy Science* 97 (11):6693-6699. doi: http://dx.doi.org/10.3168/jds.2014-7990.

Grdadolnik, J. 2002. "ATR-FTIR Spectroscopy: Its Advantages and Limitations." *Acta Chim. Slov.* 49:631 - 642.

Haaland, D. M., and E. V. Thomas. 1988. "Partial Least-Squares Methods for Spectral Analyses. 1. Relation to Other Quantitative Calibration Methods and the Extraction of Qualitative Information." *Analytical Chemistry,* 60 (11):1193-1202.

Ingle, J. D., and S. R. Crouch. 1988. *Spectrochemical Analysis*: Prentice-Hall, Inc.

Jaiswal, P., S. N. Jha, A. Borah, A. Gautam, M. K. Grewal, and G. Jindal. 2015. "Detection and quantification of soymilk in cow-buffalo milk using Attenuated Total Reflectance Fourier Transform Infrared spectroscopy (ATR-FTIR)." *Food Chem* 168:41-7. doi: 10.1016/j.foodchem.2014.07.010.

López-Calleja, I., I. González, V. Fajardo, I. Martín, P. E. Hernández, T. García, and R. Martín. 2005. "Application of Polymerase Chain Reaction to Detect Adulteration of Sheep's Milk with Goats' Milk." *Journal of Dairy Science* 88 (9):3115-3120. doi: http://dx.doi.org/10.3168/jds.S0022-0302(05)72993-3.

Malinowski, E. R. 1977. *Theory of Error in Factor Analysis*. Vol. 4, *Anal. Chem.*

Martens, H., and T. Naes. 1989. *Multivariate Calibration*: John Wiley & Sons Ltd.

McClean, P. 1997. *Population and Evolutionary Genetics*.

Miller, J. N., and J. C. Miller. 2000. *Statistics and Chemometrics for Analytical Chemistry*. 4th Edition ed: Pearson Education Limited.

Naes, T., T. Isaksson, T. Fearn, and T. Davies. 2002. *Multivariate Calibration and Classification*: NIR Publications.

Nicolaou, N., Y. Xu, and R. Goodacre. 2010. "Fourier transform infrared spectroscopy and multivariate analysis for the detection and quantification of different milk species." *J Dairy Sci* 93 (12):5651-60. doi: 10.3168/jds.2010-3619.

Nicolaou, N., Y. Xu, and R. Goodacre. 2011. "MALDI-MS and multivariate analysis for the detection and quantification of different milk species." *Anal Bioanal Chem* 399 (10):3491-502. doi: 10.1007/s00216-011-4728-6.

Özdemir, D., and Betül Öztürk. 2004. "Genetic Multivariate Calibration Methods for Near Infrared (NIR) Spectroscopic Determination of Complex Mixtures." *Turk J Chem* 28:497-514.

Özdemir, Durmuş. 2005. "Determination of Octane Number of Gasoline Using Near Infrared Spectroscopy and Genetic Multivariate Calibration Methods." *Petroleum Science and Technology* 23 (9-10):1139-1152. doi: 10.1081/lft-200035547.

Özdemir, Durmuş. 2006. "Genetic multivariate calibration for near infrared spectroscopic determination of protein, moisture, dry mass, hardness and other residues of wheat." *International Journal of Food Science and Technology* 41 (s2):12-21. doi: 10.1111/j.1365-2621.2006.01243.x.

Öztürk, Betül. 2003. "Monitoring the Esterification Reactions Of Carboxylic Acids With Alcohols Using Near–Infrared Spectroscopy And Multivariate Calibration Methods."M.Sc., Izmir Institute of Technology.

Pappas, C. S., P. A. Tarantilis, E. Moschopoulou, G. Moatsou, I. Kandarakis, and M. G. Polissiou. 2008. "Identification and differentiation of goat and sheep milk based on diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) using cluster analysis." *Food Chemistry* 106 (3):1271-1277. doi: 10.1016/j.foodchem.2007.07.034.

Perkin Elmer. "FT-IR Spectroscopy - Attenuated Total Reflectance (ATR)." http://www.utsc.utoronto.ca/~traceslab/ATR_FTIR.pdf.

PIKE Technologies. "ATR Theory and Applications." http://www.piketech.com/files/pdfs/ATRAN611.pdf.

Rodrigues, N. P., P. E. Givisiez, R. C. Queiroga, P. S. Azevedo, W. A. Gebreyes, and C. J. Oliveira. 2012. "Milk adulteration: Detection of bovine milk in bulk goat milk produced by smallholders in northeastern Brazil by a duplex PCR assay." *J Dairy Sci* 95 (5):2749-52. doi: 10.3168/jds.2011-5235.

Rodriguez, N., M. C. Ortiz, L. Sarabia, and E. Gredilla. 2010. "Analysis of protein chromatographic profiles joint to partial least squares to detect adulterations in milk mixtures and cheeses." *Talanta* 81 (1-2):255-64. doi: 10.1016/j.talanta.2009.11.067.

Russ, A., M. Barnett, W. McNabb, R. Anderson, G. Reynolds, and N. Roy. 2010. "Post-weaning effects of milk and milk components on the intestinal mucosa in inflammation." *Mutat Res* 690 (1-2):64-70. doi: 10.1016/j.mrfmmm.2009.12.006.

Santos, P. M., E. R. Pereira-Filho, and L. E. Rodriguez-Saona. 2013. "Rapid detection and quantification of milk adulteration using infrared microspectroscopy and chemometrics analysis." *Food Chem* 138 (1):19-24. doi: 10.1016/j.foodchem.2012.10.024.

Scano, P., A. Murgia, F. M. Pirisi, and P. Caboni. 2014. "A gas chromatography-mass spectrometry-based metabolomic approach for the characterization of goat milk compared with cow milk." *J Dairy Sci* 97 (10):6057-66. doi: 10.3168/jds.2014-8247.

Setle, F.A. 1997. *Handbook of Instrumental Techniques for Analytical Chemistry*. 1st Edition ed: Prentice Hall PTR.

Skoog, D. A., F. J. Holler, and S. R. Crouch. 1998. *Principles of Instrumental Analysis*: Brooks/Cole, Cengage Learning.

Smith, B. C. 1996. *Fundamentals of Fourier Transform Infrared Spectroscopy*: CRC Press LLC.

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

Wold, S. 1995. "Chemometrics; what do we mean with it, and what do we want from it? ." *Chemometrics and Intelligent Laboratory Systems* 30:109-115.

Workman, J. J. 1996. "Interpretive Spectroscopy for Near-Infrared." *Applied Spectroscopy Reviews* 31 (3):251.

APPENDIX A

SUPPORTING INFORMATION

Table A.1. Predicted goat milk contents obtained from all GILS and PLS models.

Sample Number	Actual Goat Milk (w/w %)	Predicted Goat Milk (w/w %)							
		Single Models of GILS			Binary Models of GILS			Ternary Model of GILS	Ternary Model of PLS
		June 2014	December 2014	March 2015	June 2014-December 2014	June 2014-March 2015	December 2014-March 2015	June 2014-December 2014-March 2015	June 2014-December 2014-March 2015
1	100.00	106.00	6.27	84.25	106.18	94.95	95.84	102.10	104.50
2	100.00	103.17	5.03	92.24	102.30	101.65	94.05	99.05	100.77
3	100.00	113.22	0.98	101.47	116.78	108.48	105.15	107.47	107.76
4	100.00	95.00	24.25	83.59	91.15	93.66	85.63	92.00	80.84
5	100.00	104.48	10.15	94.11	104.46	99.40	97.20	103.40	100.15
6	100.00	80.17	22.73	68.97	82.67	76.85	80.01	84.74	88.08
7	100.00	85.53	23.35	104.05	103.37	96.76	104.25	102.05	97.85
8	100.00	103.74	19.42	91.70	91.96	97.03	89.29	93.04	89.86
9	100.00	108.21	13.19	95.38	106.29	102.89	100.57	103.60	103.98
10	100.00	98.80	7.41	99.90	112.86	102.67	104.30	108.82	116.48
11	100.00	88.03	9.09	84.06	103.72	89.81	91.43	96.59	92.44
12	100.00	86.33	13.13	80.25	89.71	87.00	85.20	90.79	88.28
13	100.00	98.50	14.62	101.66	106.07	105.06	103.75	105.66	109.15
14	100.00	104.20	16.89	96.11	96.72	99.59	92.78	96.99	91.49
15	100.00	120.47	13.77	98.33	107.35	109.43	98.19	105.92	102.69
16	100.00	115.52	10.71	107.04	117.67	113.18	107.06	111.57	109.24
17	100.00	108.64	14.46	100.18	104.37	106.29	100.76	101.96	102.71
18	100.00	102.98	15.06	87.73	100.25	96.00	90.17	97.30	95.96
19	100.00	97.81	17.23	96.36	96.04	100.69	92.91	94.67	99.08
20	100.00	99.18	14.84	95.74	103.42	97.32	98.16	98.72	93.22
21	100.00	105.87	14.66	100.98	110.37	100.88	103.86	105.19	95.56
22	100.00	90.83	6.43	88.72	98.76	93.26	93.21	97.88	88.66
23	100.00	83.51	34.61	87.98	91.54	88.89	86.94	90.49	66.91
24	100.00	73.13	22.63	78.55	81.57	79.93	78.88	80.95	79.17
25	100.00	113.00	13.85	99.39	107.69	105.86	100.12	103.82	105.33
26	100.00	107.58	16.39	104.72	103.29	104.63	100.42	101.31	97.69
27	100.00	92.19	29.64	78.85	81.26	91.82	78.52	82.61	82.87
28	100.00	97.51	21.46	100.13	97.25	103.44	93.55	96.13	98.97
29	100.00	111.13	17.47	84.24	105.13	102.11	87.64	102.13	91.80
30	100.00	86.81	15.05	90.77	99.21	88.84	93.27	92.08	89.83

(cont. on next page

Table A.1 (cont.)

Sample Number	Actual Goat Milk (w/w %)	Predicted Goat Milk (w/w %)							
		Single Models of GILS			Binary Models of GILS			Ternary Model of GILS	Ternary Model of PLS
		June 2014	December 2014	March 2015	June 2014-December 2014	June 2014-March 2015	December 2014-March 2015	June 2014-December 2014-March 2015	June 2014-December 2014-March 2015
31	100.00	59.53	15.15	72.74	91.36	72.60	83.19	88.72	80.53
32	100.00	102.69	21.55	97.13	104.27	101.02	94.27	98.51	95.03
33	100.00	82.41	15.35	91.15	88.88	91.81	86.44	88.26	82.80
34	100.00	111.00	20.77	107.88	102.48	110.88	101.56	100.92	93.10
35	100.00	94.92	4.79	91.40	100.94	92.95	97.82	100.82	99.97
36	100.00	104.88	15.50	96.68	102.43	101.65	100.34	101.84	98.24
37	100.00	95.06	21.11	101.62	101.66	98.53	99.98	99.80	97.16
38	100.00	82.92	23.32	89.66	82.38	89.35	83.80	85.68	78.56
39	100.00	87.08	23.52	87.28	87.05	88.28	85.75	88.30	72.97
40	100.00	102.73	17.08	94.24	100.79	101.17	90.84	96.17	95.27
41	100.00	38.94	110.73	58.86	102.58	67.45	93.57	102.86	105.63
42	100.00	95.04	99.96	97.34	109.83	93.64	106.27	104.38	99.96
43	100.00	24.06	100.26	68.38	98.12	58.92	95.51	99.22	93.01
44	100.00	75.83	100.78	71.68	99.81	72.20	94.81	98.46	93.93
45	100.00	60.97	129.25	28.63	115.33	62.92	81.35	104.06	116.82
46	100.00	44.67	93.01	79.04	95.98	67.02	96.97	95.36	89.24
47	100.00	79.91	102.73	87.59	107.05	85.72	102.64	105.27	120.94
48	100.00	-5.98	89.22	63.66	89.72	44.67	92.76	93.14	94.49
49	100.00	57.68	101.81	86.45	106.22	75.08	105.99	106.07	107.07
50	100.00	95.88	105.58	88.96	110.15	88.57	105.34	105.90	101.20
51	100.00	-33.57	84.70	60.24	86.20	33.46	93.11	90.02	96.37
52	100.00	9.13	103.48	58.04	89.26	43.47	93.94	94.08	84.16
53	100.00	36.78	93.09	73.76	96.61	64.24	96.46	99.11	99.45
54	100.00	66.19	110.38	86.58	116.89	76.92	111.42	109.25	118.05
55	100.00	-0.94	94.83	55.17	88.66	40.15	87.78	91.41	96.49
56	100.00	9.42	99.17	58.00	95.75	43.02	92.70	91.96	99.32
57	100.00	50.40	96.09	71.76	95.93	58.94	94.52	95.14	100.25
58	100.00	58.61	104.95	80.63	107.91	72.24	103.32	101.46	108.39
59	100.00	39.19	97.85	70.57	92.58	61.97	92.23	93.39	90.74
60	100.00	86.16	98.04	89.23	108.52	85.74	104.14	104.14	99.12
61	100.00	39.56	90.62	68.05	88.25	62.05	90.32	93.85	80.20
62	100.00	-13.04	98.42	42.45	87.45	26.66	86.22	87.41	84.81
63	100.00	16.86	96.56	60.59	94.43	48.59	92.76	96.46	88.99
64	100.00	95.71	100.22	88.44	104.49	86.87	99.31	100.21	97.96
65	100.00	-0.53	86.02	61.66	88.98	46.27	88.53	86.45	85.72

(cont. on next page)

Table A.1 (cont.)

Sample Number	Actual Goat Milk (w/w %)	Predicted Goat Milk (w/w %)							
		Single Models of GILS			Binary Models of GILS			Ternary Model of GILS	Ternary Model of PLS
		June 2014	December 2014	March 2015	June 2014-December 2014	June 2014-March 2015	December 2014-March 2015	June 2014-December 2014-March 2015	June 2014-December 2014-March 2015
66	100.00	80.68	96.59	77.90	101.54	78.43	93.98	94.90	105.25
67	100.00	54.19	91.69	78.99	91.91	64.83	94.03	89.13	99.74
68	100.00	23.66	129.09	78.51	113.48	60.18	114.91	108.46	128.94
69	100.00	56.86	95.15	81.91	98.06	71.76	101.36	97.32	103.10
70	100.00	7.02	100.67	70.24	99.59	52.14	101.20	103.30	98.10
71	100.00	79.21	95.93	75.03	101.62	81.51	90.95	100.08	100.22
72	100.00	66.60	95.46	72.57	96.85	69.93	96.32	96.76	103.00
73	100.00	66.90	103.31	69.86	102.27	69.40	94.70	98.76	103.02
74	100.00	73.23	113.68	77.20	115.71	78.66	107.33	109.53	118.94
75	100.00	-13.54	89.98	60.18	87.57	36.53	94.08	93.51	84.08
76	100.00	65.11	90.73	61.14	92.79	66.51	84.23	88.89	88.95
77	100.00	-23.29	103.64	60.16	102.12	37.26	100.26	98.94	105.26
78	100.00	54.29	90.34	75.72	97.99	71.04	92.68	95.24	92.32
79	100.00	91.04	110.03	94.00	119.22	91.73	109.91	110.40	107.26
80	100.00	0.76	83.13	56.59	80.79	43.83	83.71	87.17	77.22
81	100.00	109.29	14.19	100.65	105.85	110.68	97.88	100.82	109.00
82	100.00	92.19	22.76	98.11	98.70	102.11	95.33	98.23	93.80
83	100.00	96.44	23.52	96.53	96.46	94.29	97.45	94.69	90.58
84	100.00	98.37	35.74	96.64	98.23	94.46	91.16	92.24	90.57
85	100.00	109.61	22.13	91.83	114.60	108.18	92.07	104.32	106.63
86	100.00	130.51	3.11	107.07	131.64	114.42	105.29	111.91	123.68
87	100.00	113.83	27.57	100.50	105.82	106.02	92.90	97.21	94.18
88	100.00	100.35	27.43	100.80	98.01	101.67	96.06	94.90	94.67
89	100.00	98.11	19.88	97.33	96.76	96.69	96.25	92.79	90.19
90	100.00	94.09	25.87	109.85	97.55	104.46	102.02	95.48	97.67
91	100.00	76.17	52.17	85.14	76.79	88.95	77.25	78.19	74.77
92	100.00	110.04	18.27	110.44	108.44	112.23	104.47	106.22	105.41
93	100.00	97.05	16.90	88.34	103.78	93.14	89.97	94.39	85.81
94	100.00	120.29	7.99	104.60	121.28	110.73	104.50	112.01	96.99
95	100.00	112.75	12.61	102.75	105.98	104.55	103.50	105.47	104.43
96	100.00	82.16	37.75	89.67	85.27	85.99	85.88	88.85	77.84
97	100.00	116.61	8.69	118.31	126.44	114.72	114.72	115.76	116.64
98	100.00	97.33	21.79	108.34	103.53	103.47	105.98	102.63	106.29
99	100.00	101.29	17.62	99.47	101.51	99.14	101.54	99.42	102.10
100	100.00	107.10	20.76	109.23	107.33	108.59	105.36	105.59	103.71

(cont. on next page)

Table A.1 (cont.)

Sample Number	Actual Goat Milk (w/w %)	Predicted Goat Milk (w/w %)							
		Single Models of GILS			Binary Models of GILS			Ternary Model of GILS	Ternary Model of PLS
		June 2014	December 2014	March 2015	June 2014-December 2014	June 2014-March 2015	December 2014-March 2015	June 2014-December 2014-March 2015	June 2014-December 2014-March 2015
101	100.00	94.70	28.37	96.45	97.74	97.14	95.03	96.78	89.06
102	100.00	110.84	13.89	91.60	113.58	96.13	94.60	97.93	106.13
103	100.00	104.92	5.50	99.44	116.88	105.04	109.89	109.38	119.48
104	100.00	90.11	17.00	79.82	81.32	84.85	80.53	82.58	85.05
105	100.00	101.30	8.52	95.89	107.39	98.49	98.13	101.54	111.64
106	100.00	121.96	17.05	104.84	117.93	112.64	98.80	107.11	108.13
107	100.00	102.04	33.64	98.95	98.81	102.55	93.86	95.33	87.10
108	100.00	63.63	2.58	75.87	93.24	72.87	92.49	95.23	91.82
109	100.00	126.55	23.98	106.37	119.50	115.72	110.06	113.67	86.53
110	100.00	81.28	10.41	86.99	102.98	84.96	94.14	93.74	94.56
111	100.00	44.83	13.70	62.33	79.00	55.20	78.90	76.20	82.18
112	100.00	85.31	22.97	90.27	92.46	86.58	96.91	92.37	87.01
113	100.00	87.34	22.70	83.53	90.61	86.31	86.19	91.31	81.24
114	100.00	101.54	7.23	93.10	109.21	95.92	100.99	103.22	111.69
115	100.00	89.67	31.02	93.35	95.64	95.27	88.16	90.92	84.30
116	100.00	86.83	7.32	83.24	105.92	84.83	95.98	98.83	95.07
117	100.00	118.32	6.80	108.89	118.59	111.70	113.06	112.29	107.98
118	100.00	114.80	2.73	97.18	112.98	100.21	97.11	98.57	103.90
119	100.00	106.99	13.30	97.94	113.48	101.25	97.82	102.03	105.69
120	100.00	96.96	36.08	98.85	93.95	99.01	94.29	92.97	84.04

Table A.2. Predicted cow milk contents obtained from all GILS and PLS models.

Sample Number	Actual Cow Milk (w/w %)	Predicted Cow Milk (w/w %)							
		Single Models of GILS			Binary Models of GILS			Triple Model of GILS	Triple Model of PLS
		June 2014	December 2014	March 2015	June 2014-December 2014	June 2014-March 2015	December 2014-March 2015	June 2014-December 2014-March 2015	June 2014-December 2014-March 2015
1	0.00	4.36	6.27	15.66	-3.33	7.71	7.01	0.79	-3.01
2	0.00	-8.91	5.03	-0.09	-7.06	-8.26	-3.05	-6.43	-7.23
3	0.00	-13.45	0.98	-1.10	-12.75	-5.71	-5.86	-10.07	-9.41
4	0.00	4.23	24.25	11.76	4.90	6.80	7.29	4.80	14.23
5	0.00	1.40	10.15	5.86	-2.99	-1.86	2.99	-2.23	-0.45
6	0.00	25.76	22.73	26.81	15.85	21.31	17.94	15.45	11.30

(cont. on next page)

Table A.2 (cont.)

Sample Number	Actual Cow Milk (w/w %)	Predicted Cow Milk (w/w %)							
		Single Models of GILS			Binary Models of GILS			Triple Model of GILS	Triple Model of PLS
		June 2014	December 2014	March 2015	June 2014-December 2014	June 2014-March 2015	December 2014-March 2015	June 2014-December 2014-March 2015	June 2014-December 2014-March 2015
7	0.00	11.69	23.35	-1.44	5.18	-0.42	1.13	1.93	6.40
8	0.00	-0.44	19.42	2.52	2.26	1.24	1.97	1.88	3.56
9	0.00	-0.04	13.19	5.80	-1.36	0.43	-0.70	-2.32	-3.98
10	0.00	-3.39	7.41	-4.30	-6.54	-3.63	-2.35	-5.67	-15.76
11	0.00	-1.31	9.09	4.78	-7.62	2.75	-0.56	-4.67	0.64
12	0.00	13.20	13.13	15.16	5.54	9.82	6.92	3.15	8.00
13	0.00	-2.13	14.62	-6.41	-4.62	-8.53	-7.54	-7.40	-10.90
14	0.00	-4.74	16.89	4.77	3.75	-0.58	5.76	3.84	4.36
15	0.00	-9.24	13.77	0.20	-9.94	-8.07	-3.50	-8.04	-7.04
16	0.00	-6.41	10.71	-4.31	-12.10	-7.84	-7.47	-11.28	-8.82
17	0.00	-8.34	14.46	-3.98	-5.37	-4.31	-4.11	-4.51	-6.60
18	0.00	-1.09	15.06	8.20	-2.45	1.17	5.78	0.41	0.15
19	0.00	-2.72	17.23	1.06	0.29	-1.97	0.23	1.67	-3.83
20	0.00	-3.09	14.84	4.73	-2.31	-1.06	2.11	1.58	5.58
21	0.00	-0.47	14.66	0.31	-6.25	0.87	-1.47	-2.54	6.09
22	0.00	6.33	6.43	8.21	-3.30	5.01	3.83	0.78	9.27
23	0.00	13.46	34.61	10.20	7.95	8.34	10.01	8.73	28.64
24	0.00	12.90	22.63	10.66	11.32	9.42	10.95	10.44	12.88
25	0.00	-3.30	13.85	1.98	-7.28	-3.82	-2.11	-5.30	-5.54
26	0.00	-2.25	16.39	-5.57	1.16	-6.04	-0.84	-0.95	0.66
27	0.00	6.71	29.64	14.48	13.85	6.34	12.24	9.20	8.65
28	0.00	-1.80	21.46	-3.06	7.39	-6.51	4.14	2.45	-3.22
29	0.00	0.71	17.47	19.95	-4.00	4.42	11.01	1.86	8.64
30	0.00	6.43	15.05	2.61	5.21	6.95	4.36	5.59	7.43
31	0.00	17.40	15.15	8.34	2.10	8.01	7.39	3.36	11.23
32	0.00	-0.85	21.55	-4.03	-4.36	-3.54	-1.82	-3.39	1.27
33	0.00	5.72	15.35	0.51	5.10	-3.92	4.87	1.16	8.82
34	0.00	-10.06	20.77	-5.40	1.47	-10.05	0.81	-1.04	3.84
35	0.00	5.62	4.79	4.10	-2.79	3.85	0.20	0.02	-2.02
36	0.00	1.04	15.50	0.39	0.05	0.90	-2.88	-0.95	0.21
37	0.00	3.80	21.11	-1.75	5.16	0.55	2.52	4.49	1.76
38	0.00	8.30	23.32	3.64	9.70	5.05	6.17	8.49	14.85
39	0.00	5.99	23.52	5.61	8.20	3.15	6.92	6.81	20.29
40	0.00	-6.01	17.08	3.38	-1.94	-4.89	4.78	-0.64	-0.71
41	0.00	15.15	-12.71	3.07	-4.16	8.56	6.73	-5.91	-15.93

(cont. on next page)

Table A.2 (cont.)

Sample Number	Actual Cow Milk (w/w %)	Predicted Cow Milk (w/w %)							
		Single Models of GILS			Binary Models of GILS			Triple Model of GILS	Triple Model of PLS
		June 2014	December 2014	March 2015	June 2014-December 2014	June 2014-March 2015	December 2014-March 2015	June 2014-December 2014-March 2015	June 2014-December 2014-March 2015
42	0.00	5.27	2.47	-2.29	-4.22	2.06	-4.59	-2.76	1.25
43	0.00	24.51	-4.12	6.48	-5.97	4.46	-3.23	-6.43	-5.65
44	0.00	28.53	-0.66	17.63	1.59	20.86	5.20	5.02	7.11
45	0.00	28.81	-17.26	6.20	-2.91	39.03	7.03	8.31	-11.02
46	0.00	18.68	1.74	-1.43	-3.90	5.60	-4.03	-4.48	1.40
47	0.00	21.88	9.38	14.50	3.03	15.89	7.32	4.70	-12.46
48	0.00	36.40	4.77	7.85	2.08	12.43	1.81	-1.88	-5.49
49	0.00	33.38	6.44	7.41	0.69	15.74	2.29	1.35	-1.01
50	0.00	12.29	-6.59	9.73	-9.74	6.96	-2.93	-2.44	1.98
51	0.00	47.24	5.90	11.29	3.74	16.63	3.27	1.50	-5.57
52	0.00	39.83	-4.77	18.84	1.34	19.78	2.96	2.14	5.99
53	0.00	23.22	-0.69	3.57	-0.27	8.57	-1.07	-3.69	-7.48
54	0.00	38.57	5.17	18.16	-0.31	25.74	1.65	1.03	-3.67
55	0.00	44.43	2.61	19.54	4.05	22.79	8.69	3.37	-2.90
56	0.00	52.54	2.54	24.13	6.47	30.24	7.44	6.27	2.29
57	0.00	29.10	0.76	13.95	-0.56	17.40	1.36	4.03	-0.70
58	0.00	20.15	-3.70	8.15	-2.03	13.04	0.35	-1.80	-8.01
59	0.00	25.34	-1.07	6.36	2.87	9.25	4.17	2.18	0.23
60	0.00	11.18	7.00	5.25	-3.22	9.16	-2.55	-1.64	1.67
61	0.00	12.43	-1.52	1.79	-4.11	2.39	-2.54	-4.03	2.44
62	0.00	57.10	-0.34	34.08	6.95	35.71	10.12	8.29	9.67
63	0.00	36.58	2.03	15.93	-0.09	15.81	2.39	-1.47	1.10
64	0.00	10.88	1.61	12.17	-0.85	11.58	4.61	4.24	4.27
65	0.00	28.76	-2.26	-0.98	-4.71	6.39	-6.79	-6.24	-1.59
66	0.00	22.48	7.24	15.83	3.72	17.21	6.72	5.89	-1.64
67	0.00	31.25	7.85	8.09	7.38	18.68	1.61	7.25	-1.15
68	0.00	27.48	-41.51	2.50	-14.79	1.61	-14.29	-17.35	-34.78
69	0.00	30.84	12.19	9.96	8.54	17.68	4.52	7.37	0.47
70	0.00	35.01	-1.09	3.62	-4.44	8.34	-7.04	-7.55	-6.37
71	0.00	18.49	9.61	23.66	2.35	18.11	10.33	5.55	2.73
72	0.00	31.85	5.43	20.60	3.08	22.74	2.80	3.05	-0.47
73	0.00	24.51	-8.44	14.72	-6.69	17.37	1.69	0.82	-4.04
74	0.00	31.06	-0.26	22.48	-2.12	24.77	5.27	3.15	-5.84
75	0.00	46.50	3.39	12.95	3.56	19.50	1.54	0.01	5.21
76	0.00	21.71	-1.11	16.60	-0.13	13.97	5.64	1.51	3.63

(cont. on next page)

Table A.2 (cont.)

Sample Number	Actual Cow Milk (w/w %)	Predicted Cow Milk (w/w %)							
		Single Models of GILS			Binary Models of GILS			Triple Model of GILS	Triple Model of PLS
		June 2014	December 2014	March 2015	June 2014-December 2014	June 2014-March 2015	December 2014-March 2015	June 2014-December 2014-March 2015	June 2014-December 2014-March 2015
77	0.00	51.04	-3.52	15.63	-1.73	21.85	-1.82	-5.25	-9.20
78	0.00	18.06	7.03	6.51	0.57	9.46	2.60	-0.06	1.21
79	0.00	8.41	-1.60	8.91	-12.83	12.05	-5.64	-6.09	-1.94
80	0.00	30.55	1.06	5.44	3.53	9.77	2.63	0.87	2.05
81	0.00	-8.72	14.19	-3.09	-3.61	-10.76	-0.51	-2.59	-10.45
82	0.00	9.31	22.76	1.05	2.59	-3.06	0.28	-0.35	5.41
83	0.00	9.26	23.52	0.40	3.54	4.78	0.47	7.34	9.85
84	0.00	12.47	35.74	5.27	10.15	5.90	11.47	12.34	14.62
85	0.00	-4.49	22.13	7.97	-7.12	-4.89	10.95	1.06	-2.78
86	0.00	-19.66	3.11	-6.67	-25.34	-14.26	-6.73	-11.48	-20.68
87	0.00	-14.71	27.57	-9.25	-9.20	-14.03	-1.59	-3.70	1.12
88	0.00	10.58	27.43	0.26	5.63	-0.81	1.98	2.56	6.77
89	0.00	3.86	19.88	-0.77	3.82	0.50	2.10	2.60	6.71
90	0.00	8.29	25.87	-9.96	3.76	-7.58	-6.34	-2.41	2.33
91	0.00	20.22	52.17	13.04	23.57	9.75	17.83	16.03	21.93
92	0.00	-4.99	18.27	-8.95	-4.50	-13.01	-3.76	-7.34	-3.77
96	0.00	24.62	37.75	2.64	12.28	5.65	8.07	8.06	22.98
97	0.00	-9.20	8.69	-18.23	-15.90	-18.46	-10.71	-13.92	-11.50
98	0.00	8.71	21.79	-5.36	1.90	-2.64	-4.56	-1.52	-2.80
99	0.00	8.82	17.62	4.64	2.39	4.74	2.30	3.72	1.13
100	0.00	-0.87	20.76	-6.72	-4.55	-8.72	-5.74	-5.66	-2.34
101	0.00	13.53	28.37	7.32	7.85	4.36	9.16	8.80	14.54
102	0.00	-1.76	13.89	14.53	-5.25	7.70	12.53	7.69	-1.42
103	0.00	3.70	5.50	2.57	-9.84	1.43	-5.33	-5.60	-14.20
104	0.00	16.44	17.00	17.04	14.94	12.59	15.07	14.64	12.30
105	0.00	9.15	8.52	9.50	-0.30	5.64	4.91	3.22	-6.24
106	0.00	-13.54	17.05	0.19	-13.45	-7.98	-0.58	-6.41	-7.64
107	0.00	3.65	33.64	5.61	7.59	-1.33	11.49	5.60	13.20
108	0.00	18.13	2.58	6.20	-4.47	6.44	-0.55	-1.45	-0.49
109	0.00	-11.13	23.98	-3.14	-16.90	-12.51	-11.58	-11.10	15.33
110	0.00	2.31	10.41	-0.28	-7.78	0.50	0.24	-1.74	-1.49
111	0.00	33.30	13.70	22.62	14.21	27.21	13.94	18.77	11.08
112	0.00	22.92	22.97	11.04	11.07	15.22	7.80	10.23	16.86
113	0.00	15.88	22.70	16.83	8.55	13.22	13.31	10.10	17.56
114	0.00	5.80	7.23	8.90	-4.45	7.29	1.09	0.16	-7.84

(cont. on next page)

Table A.2 (cont.)

Sample Number	Actual Cow Milk (w/w %)	Predicted Cow Milk (w/w %)							
		Single Models of GILS			Binary Models of GILS			Triple Model of GILS	Triple Model of PLS
		June 2014	December 2014	March 2015	June 2014-December 2014	June 2014-March 2015	December 2014-March 2015	June 2014-December 2014-March 2015	June 2014-December 2014-March 2015
115	0.00	7.42	31.02	4.42	3.85	1.96	5.59	3.99	13.37
116	0.00	9.34	7.32	8.14	-9.43	10.13	-1.12	-2.00	3.48
117	0.00	-6.83	6.80	-5.62	-14.10	-7.85	-9.35	-9.58	-5.17
118	0.00	-10.51	2.73	-1.28	-13.50	-4.93	0.47	-1.72	-5.35
119	0.00	-5.61	13.30	3.56	-7.35	-0.09	4.46	1.04	-4.03
120	0.00	8.41	36.08	3.57	7.02	3.40	3.36	6.17	16.30

Table A.3. Predicted water contents obtained from all GILS and PLS models.

Sample Number	Actual Water (w/w %)	Predicted Water (w/w %)							
		Single Models of GILS			Binary Models of GILS			Ternary Model of GILS	Ternary Model of PLS
		June 2014	December 2014	March 2015	June 2014-December 2014	June 2014-March 2015	December 2014-March 2015	June 2014-December 2014-March 2015	June 2014-December 2014-March 2015
1	0.00	-6.58	-4.62	-0.82	-4.48	-1.33	-1.32	-3.48	-3.89
2	0.00	4.65	2.15	8.98	3.50	5.10	6.80	5.40	4.95
3	0.00	0.02	-5.80	2.93	-3.67	-0.57	0.48	-2.33	-1.77
4	0.00	-0.55	-0.14	5.98	1.46	0.75	5.69	2.35	2.18
5	0.00	-2.02	-3.33	2.23	-2.18	1.34	0.74	-0.24	-2.09
6	0.00	-1.18	1.48	1.89	-0.18	2.20	3.55	0.57	0.75
7	0.00	6.76	-8.47	-2.63	-6.69	3.18	-3.74	-3.27	-5.30
8	0.00	-0.86	1.89	6.21	3.43	1.98	6.63	4.52	4.67
9	0.00	-5.50	-4.68	0.23	-3.25	-2.16	-0.95	-1.28	-1.07
10	0.00	5.47	-4.59	2.85	-4.72	4.20	-0.82	-2.98	-3.36
11	0.00	14.08	6.95	10.24	6.20	12.19	8.43	7.80	6.16
12	0.00	2.50	3.17	4.84	3.46	5.10	6.25	4.44	0.86
13	0.00	5.00	-3.14	3.51	-1.64	3.74	2.40	1.05	0.00
14	0.00	0.88	-3.84	4.27	-2.04	-0.57	1.78	-0.44	-0.15
15	0.00	-5.21	-1.37	3.32	-0.85	-2.45	3.17	0.20	1.05
16	0.00	-4.15	-6.09	-1.72	-5.46	-2.98	-1.24	-3.30	-3.72
17	0.00	0.57	-2.97	3.89	-1.28	0.82	2.72	0.86	0.37

(cont. on next page)

Table A.3 (cont.)

Sample Number	Actual Water (w/w %)	Predicted Water (w/w %)							
		Single Models of GILS			Binary Models of GILS			Ternary Model of GILS	Ternary Model of PLS
		June 2014	December 2014	March 2015	June 2014-December 2014	June 2014-March 2015	December 2014-March 2015	June 2014-December 2014-March 2015	June 2014-December 2014-March 2015
18	0.00	0.49	-0.41	4.00	-0.02	1.98	3.86	1.70	1.59
19	0.00	1.65	0.07	4.59	1.75	1.80	6.44	3.56	4.76
20	0.00	1.81	-4.13	1.91	-3.08	0.88	0.53	-0.89	-2.75
21	0.00	-3.77	-6.60	-0.47	-5.28	-1.13	-1.86	-2.37	-2.18
22	0.00	4.14	-0.01	3.42	0.14	3.56	1.99	-0.19	1.16
23	0.00	4.64	-1.48	4.32	0.01	2.43	2.58	1.17	2.81
24	0.00	12.27	4.47	10.62	5.52	9.22	8.82	6.34	4.18
25	0.00	-4.13	-4.84	-0.19	-4.03	-0.54	0.24	-2.16	-0.96
26	0.00	-2.61	-6.64	1.94	-3.44	-2.61	0.31	-2.20	-0.01
27	0.00	-0.04	4.07	6.60	5.19	0.76	8.55	5.65	7.01
28	0.00	2.92	-2.04	3.81	0.28	1.52	3.26	3.16	4.47
29	0.00	-6.67	-6.85	0.96	-4.28	-3.73	-0.23	-2.94	0.01
30	0.00	5.22	-1.50	6.30	-0.02	4.60	2.68	1.51	-2.08
31	0.00	24.57	6.53	14.28	6.70	19.18	9.52	7.72	7.63
32	0.00	-1.22	0.85	5.26	0.90	1.58	4.99	2.70	5.48
33	0.00	14.55	4.78	10.28	5.91	9.74	8.85	7.82	8.99
34	0.00	-1.06	-6.88	2.89	-3.36	-2.18	1.58	-0.77	0.84
35	0.00	1.23	-1.41	5.14	-0.24	3.81	2.19	-0.15	3.31
36	0.00	-3.74	-3.64	1.49	-2.72	-2.24	0.72	-1.92	-2.72
37	0.00	0.41	-8.00	0.46	-5.23	-1.07	-1.94	-3.61	-2.62
38	0.00	9.32	2.89	9.06	3.80	7.29	7.04	5.84	2.33
39	0.00	6.04	2.24	8.63	4.08	5.57	6.84	4.39	5.90
40	0.00	2.97	-0.55	6.46	0.98	2.54	5.89	2.84	3.96
41	0.00	44.84	-1.21	15.65	4.35	28.81	4.04	2.24	9.72
42	0.00	1.05	-6.39	1.77	-5.26	2.50	-2.20	-2.91	-4.63
43	0.00	56.08	4.81	20.35	7.15	37.67	9.04	7.83	14.07
44	0.00	1.93	-1.18	4.85	-3.17	7.85	-0.43	-1.95	-2.38
45	0.00	8.75	-10.66	-2.04	-8.78	5.40	-9.10	-14.46	-8.63
46	0.00	40.31	4.58	16.97	4.84	28.24	7.63	6.46	7.17
47	0.00	1.96	-10.37	-4.16	-9.45	5.11	-7.48	-6.11	-6.64
48	0.00	70.43	6.65	20.79	7.87	45.73	6.61	8.16	8.10
49	0.00	15.01	-8.18	-2.15	-8.28	11.09	-7.56	-6.40	-5.32
50	0.00	-3.60	-4.17	-0.01	-5.00	1.46	-2.06	-4.11	-0.22
51	0.00	85.95	5.76	19.45	6.88	51.99	4.76	3.87	3.65
52	0.00	57.67	3.89	17.55	5.35	37.85	4.52	3.07	6.89

(cont. on next page)

Table A.3 (cont.)

Sample Number	Actual Water (w/w %)	Predicted Water (w/w %)							
		Single Models of GILS			Binary Models of GILS			Ternary Model of GILS	Ternary Model of PLS
		June 2014	December 2014	March 2015	June 2014-December 2014	June 2014-March 2015	December 2014-March 2015	June 2014-December 2014-March 2015	June 2014-December 2014-March 2015
53	0.00	40.63	4.71	16.59	6.41	29.78	6.90	6.63	5.18
54	0.00	3.20	-13.18	-10.50	-12.74	5.83	-12.95	-10.24	-9.11
55	0.00	59.15	6.38	16.41	6.09	40.98	4.80	5.80	5.02
56	0.00	43.51	1.06	8.18	0.04	31.56	-0.88	0.54	1.22
57	0.00	25.20	1.53	7.69	0.09	20.37	0.91	1.34	-1.06
58	0.00	24.09	-1.47	6.43	-2.61	16.78	-0.46	-1.13	-3.31
59	0.00	38.97	2.90	17.97	5.33	27.32	5.45	4.64	0.54
60	0.00	5.52	-5.60	1.04	-5.12	4.49	-2.04	-4.12	-5.23
61	0.00	47.61	8.51	24.49	10.93	33.52	12.13	10.54	9.92
62	0.00	63.22	4.65	17.28	4.93	43.16	3.70	3.73	4.54
63	0.00	52.26	4.24	17.66	5.36	37.21	5.84	5.34	8.06
64	0.00	-2.34	-6.18	-0.55	-5.35	-0.54	-2.66	-4.84	-6.17
65	0.00	71.37	18.24	28.90	16.19	49.13	16.86	16.97	14.82
66	0.00	0.92	-4.11	1.39	-5.19	5.07	-1.13	-2.28	-2.46
67	0.00	19.28	0.14	7.52	0.08	16.76	2.43	2.71	2.92
68	0.00	53.47	1.53	20.20	6.14	42.10	6.74	6.77	14.61
69	0.00	15.61	-7.22	1.73	-6.51	10.37	-4.88	-5.84	-7.99
70	0.00	62.19	2.67	18.41	3.56	41.87	4.28	5.16	6.77
71	0.00	3.79	-3.02	-0.61	-4.82	5.32	-2.69	-3.14	-5.29
72	0.00	10.02	-2.60	1.07	-2.29	12.96	-1.87	-0.28	2.29
73	0.00	12.80	1.76	9.42	0.09	15.78	2.64	1.39	2.16
74	0.00	-1.20	-12.09	-8.46	-13.41	2.08	-11.27	-11.11	-12.15
75	0.00	72.62	5.80	17.93	7.46	46.20	5.27	5.78	10.67
76	0.00	16.86	10.54	15.53	7.27	17.81	9.97	7.64	4.65
77	0.00	76.78	3.32	15.50	3.31	48.61	1.42	2.35	3.29
78	0.00	29.53	3.16	12.77	3.35	20.79	5.45	4.39	4.45
79	0.00	4.16	-7.49	-4.31	-7.58	4.17	-7.10	-5.90	-6.00
80	0.00	68.26	12.14	31.02	15.62	48.29	15.83	15.54	21.39
81	0.00	-3.89	-4.81	4.58	-3.64	-2.12	3.47	1.68	0.14
82	0.00	1.37	-2.81	3.54	-2.53	2.20	4.18	2.73	3.06
83	0.00	-1.59	-6.85	2.53	-5.25	0.85	-1.20	-0.66	-0.70
84	0.00	-4.97	-9.58	-2.66	-8.85	-1.82	-3.36	-2.42	-1.68
85	0.00	-1.01	-9.25	1.44	-7.33	-0.40	-0.93	-1.62	-1.64
86	0.00	-7.69	-5.76	2.12	-6.76	-0.72	0.58	-0.01	3.39
87	0.00	2.05	-1.74	7.52	-0.51	3.20	6.89	6.40	5.82

(cont. on next page)

Table A.3 (cont.)

Sample Number	Actual Water (w/w %)	Predicted Water (w/w %)							
		Single Models of GILS			Binary Models of GILS			Ternary Model of GILS	Ternary Model of PLS
		June 2014	December 2014	March 2015	June 2014-December 2014	June 2014-March 2015	December 2014-March 2015	June 2014-December 2014-March 2015	June 2014-December 2014-March 2015
88	0.00	-6.82	-4.65	-0.48	-4.74	-2.41	0.89	0.91	2.97
89	0.00	-0.66	-1.49	4.51	-0.33	0.91	4.58	3.51	4.45
90	0.00	0.91	-4.91	0.35	-2.17	1.22	1.94	2.33	0.84
91	0.00	2.92	-1.28	2.13	-0.28	1.39	5.43	5.15	8.68
92	0.00	-5.34	-6.71	-0.57	-5.33	-2.21	0.15	-0.30	1.64
93	0.00	-2.68	-0.85	6.30	-2.62	3.08	4.18	1.92	1.40
94	0.00	-12.92	-6.48	-0.79	-8.68	-4.72	-1.20	-3.80	-6.32
95	0.00	-9.09	-7.41	0.22	-6.11	-3.06	-1.38	-3.04	0.22
96	0.00	-1.40	-4.58	4.90	-1.80	5.19	2.41	3.05	2.43
97	0.00	-6.22	-8.24	-1.23	-8.67	0.27	-2.33	-2.22	-3.09
98	0.00	-4.85	-7.12	-5.01	-5.19	-2.56	-1.66	-1.47	0.54
99	0.00	-5.27	-8.33	-3.67	-7.09	-3.35	-2.91	-4.06	-0.90
100	0.00	-3.55	-6.22	-1.75	-4.78	-1.81	0.02	-0.61	2.17
101	0.00	-3.64	-8.78	-3.42	-8.11	-1.87	-3.62	-3.12	-2.68
102	0.00	-8.49	-10.61	-3.26	-10.53	-4.06	-5.03	-5.25	-5.08
103	0.00	-6.14	-7.81	-4.52	-8.40	-2.10	-3.43	-3.35	-4.63
104	0.00	-5.15	2.33	4.01	1.09	-0.89	5.67	3.04	3.90
105	0.00	-8.73	-7.02	-4.23	-8.21	-3.38	-3.56	-4.47	-5.06
106	0.00	-7.22	-5.20	-1.81	-5.74	-3.85	0.17	-0.78	2.13
107	0.00	-4.26	-7.80	0.22	-5.65	-3.18	-0.62	-1.07	-0.94
108	0.00	26.47	6.55	13.46	6.70	20.71	7.46	7.45	9.36
109	0.00	-5.03	-10.79	-0.83	-8.01	-3.89	-1.51	-3.54	-1.19
110	0.00	18.37	3.14	10.91	4.17	14.50	7.57	6.56	6.79
111	0.00	27.36	5.81	12.00	5.15	18.68	6.17	5.46	4.88
112	0.00	-3.89	-6.09	-3.28	-5.39	-0.56	-2.75	-3.02	-3.52
113	0.00	-2.85	-2.08	0.80	-0.18	-0.27	2.72	1.00	-0.03
114	0.00	-7.13	-5.13	-2.92	-6.64	-1.07	-2.47	-3.32	-3.72
115	0.00	3.58	0.16	3.68	0.18	3.73	4.92	5.25	5.88
116	0.00	6.51	-0.42	4.88	-1.08	8.76	3.31	2.25	2.86
117	0.00	-9.82	-8.65	-3.57	-8.29	-4.41	-3.51	-4.69	-3.82
118	0.00	-5.42	-3.10	4.76	-2.67	0.99	2.23	1.87	0.53
119	0.00	-3.52	-4.26	0.17	-5.27	-0.50	0.12	-0.98	-0.36
120	0.00	-3.71	-5.19	-0.77	-3.98	-1.27	0.35	0.76	1.68