DETERMINATION OF THE GEOGRAPHICAL ORIGIN OF MILK BASED ON SPECTROSCOPIC AND CHEMOMETRIC TECHNIQUES

SHIMA BEHKAMI

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

DETERMINATION OF THE GEOGRAPHICAL ORIGIN OF MILK BASED ON SPECTROSCOPIC AND CHEMOMETRIC TECHNIQUES

SHIMA BEHKAMI

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: SHIMA BEHKAMI

Matric No: SHC100069

Name of Degree: Doctor of philosophy (Ph.D.)

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

DETERMINATION OF THE GEOGRAPHICAL ORIGIN OF MILK BASED ON SPECTROSCOPIC AND CHEMOMETRIC TECHNIQUES

Field of Study: FOOD CHEMISTY AND CHEMOMETRICS

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

ABSTRACT

The Purpose of this research is to develop methods to determine the geographical origin of raw and factory cow milk samples in Peninsular Malaysia and some selected regions of the world. To achieve this, the concentration of mineral and trace elements, isotopic ratio and nutritional values have been studied in this dissertation. We have used various analytical methods have been used for analysis of milk samples such as Inductively Coupled Plasma Mass Spectrometry (ICP-MS), Isotopic Ratio Mass Spectrometry (IRMS), ultrasonic based scanner (Milkoscan), UV/Vis spectrometry and Near Infrared Spectroscopy (NIR). Besides milk, samples of food, pellet, water and plant consumed by cattle as well as rain, soil and cattle hair were also collected from each farm and analyzed in order to observe if there are any possible correlation between these samples and milk. Cattle hair in particular was analyzed with the same procedure as milk samples in order to confirm whether it can be used in place of milk for the determination of geographical origin. The analytical data obtained from these methods were then analyzed using various chemometric methods such as principal component analysis (PCA), hierarchal cluster analysis (HCA), discriminant analysis (DA), Factorial analysis (FA) and artificial neural network (ANN). From the elemental analyses of ICP-MS it is observed that as far as Malaysian raw cow milk samples are concerned, although the separation of samples based on the state where the farms are located is observed, a clearer mapping is ascertained between northern and southern sampling regions. Northern raw cow milk samples are richer in Mo, Al, Mn, Na, Fe, Mg, Zn, Ca, K, Ba, Se and Cu compared to the south where they are mainly loaded with Ni. Malaysian factory milk samples are clustered away from factory milks of some selected countries and the separation is based on the loadings of Ca, Fe, Na, Ba, Zn, Mg, Mn and K. This indicates that Malaysian factory milk samples are richer in these metals compared to factory milks from the selected countries. Multi elemental information on

the pellet used in different farms lead to two clusters of pellets which could be explained by the fact that there are only two major factories producing cattle pellets in Peninsular Malaysia. Multi element isotopic analyses of IRMS shows separation of samples in the northern and southern part of Malaysia. The discriminating factors for the southern milk samples are $\delta^{18}O$ and $\delta^{13}C$ while that of the northern samples is $\delta^{15}N$. Data from milk samples analyzed by NIR and UV/Vis were used to train ANNs and the trained ANNs are able to predict the origin of the milk samples very well. Based on the different analytical and chemometric techniques used in analyzing our samples it can be concluded that even though the chemical and physical variables investigated are different in each method, distinct separation between samples from different geographical origins can be clearly observed.

ABSTRAK

Tujuan penyelidikan ini adalah untuk menentukan asal usul sampel susu mentah dan susu kilang di Semenanjung Malaysia dan beberapa kawasan di dunia. Untuk itu, kepekatan galian dan unsur surih, nisbah isotop dan nilai nutrisi akan dikenalpasti di dalam tesis ini. Kami telah menggunakan pelbagai metod analisis untuk menganalisis sampel susu kami. Antaranya adalah *Inductively Coupled Plasma Mass Spectrometry (ICP-MS), Isotopic Ratio Mass Spectrometry (IRMS), ultrasonic based scanner (Milkoscan), UV/Vis spectrometry* dan *Near Infrared Spectroscopy (NIR).* Di samping susu, sampel makanan, pelet air dan rumput yang dimakan oleh lembu, termasuk air yang diminum lembu serta bulu lembu telah juga dikumpulkan dari setiap ladang. Sampel ini telah juga dianalisis bagi menentukan samada terdapat korelasi antara sampel tersebut dan sampel susu. Bulu lembu telah dianalisis dengan prosidur yang sama seperti penganalisaan susu bagi memastikan samada bulu lembu boleh digunakan sebagai pengganti untuk susu bagi menentukan asal-usul geografi. Data yang diperoleh dari metod analisis kimia ini kemudiannya dianalisis menggunakan pelbagai metod kimometrik seperti analisis komponen utama (PCA), analisis kluster berperingkat (HCA), analisis diskriminan (DA), analisi factorial, dan rangkaian neural tiruan (ANN). Dari analisis unsur yang diperoleh dari ICP-MS, didapati bahawa bagi susu mentah dari Malaysia, walaupun terdapat pemisahan berdasarkan negeri tempat ladang ternak berada, pemetaan yang lebih jelas dapat dilihat bagi kedudukan ladang di utara dan selatan semenanjung. Susu dari kawasan utara kaya dengan Mo, Al, Mn, Na, Fe, Mg, Zn, Ca, K, Ba, Se and Cu berbanding dengan selatan yang lebih kaya dengan Ni. Susu kilang Malaysia berkumpul jauh dari kumpulan susu kilang dari negara lain. Pemisahan ini adalah berdasarkan pemberatan pada Ca, Fe, Na, Ba, Zn, Mg, Mn and K. Ini menunjukkan bahawa sampel susu kilang di Malaysia adalah lebih kaya dengan logam tersebut berbanding negara lain. Informasi banyak unsur untuk pelet yang digunakan di

ladang yang berbeza menunjukkan dua kumpulan dan ini adalah kerana terdapat hanya dua kilang yang mengeluarkan pelet di Semenanjung Malaysia. Dari analisis IRMS, pemisahan sampel berdasarkan kawasan utara dan selatan Malaysia dapat diperhatikan. Faktor pemisah bagi sampel susu selatan adalah $\delta^{18}O$ dan $\delta^{13}C$ manakala sampel utara pula terpisah kerana $\delta^{15}N$. Data dari sampel susu yang dianalisis menggunakan NIR dan UV/Vis kemudiannya digunakan untuk melatih ANN dan ANN yang dilatih ini berjaya meramalkan asal usul geografi sampel susu dengan baik. Berdasarkan teknik analisis dan kimometrik yang berbeza, kami membuat keputusan bahawa meskipun pembolehubah kimia dan fizik yang dikaji adalah berbeza, kami masih dapat memerhatikan pemisahan yang jelas antara sampel dari utara dan selatan.

Dedicated to:

GOD, who gifted me those who were the keys to my success,

My advisor who taught me how to achieve my dreams by giving me his time and knowledge,

My grandparents, who gifted me wonderful parents,

My mom, who left for heaven and never got to see me following the path she drew successfully,

My dad, who supported me academically and financially,

My brother, who made me laugh even when I didn't want to smile,

My daughter, who spent wonderful moments of her childhood in lab with me,

My husband, who shared his experiences with me.

Love you all

ACKNOWLEDGEMENT

Each and every single page of this thesis is full of unforgettable sweet and bitter experiences that my advisor and I have shared. Thankyou God, for loving me unconditionally. Your blessings have allowed me to reach to this stage.

To my advisor Prof. Sharifuddin M. Zain, please accept my heartfelt thanks for accepting me as your PhD student; it's difficult for me to find the words to truly express my gratitude to you. I will never forget that whenever I stumbled, you didn't catch me, but taught me how to fly — a wiser decision that could only have been made by you who cared a lot.The answer to all my needs was "try to be patient" which was vital in my quest to become a successful academician. Your smiley face was the only courage to move on!

To my mom, Shahin Soltanian, Mom, you were the first one to teach me chemistry; remember you always talked to me about how you analysed milk and ice cream as a quality control manager. You encouraged me to be a food chemist, but you left me without seeing me achieving it. Thanks for all the times I forgot to thank you.

To my dad, Prof. Aziz Behkami, thanks for being such a wonderful father. Dad, by establishing what came to be your extensive academic background, you engineered the success of our family. Thanks for understanding the things that I said, and even those that I never planned on telling you. Thanks for unlimited financial support.

To my brother, Dr. Nima Behkami, with whom I shared lovely memories of my childhood. Your shoulder has been a place for me to lean on. The financial and emotional supports is unforgettable. To my lovely daughter Rima, for all the sacrifices you have made. Thanks for always praying for me with your clean heart. Thanks for always coming out on top in your studies, ranking first without my help. To my

husband, Dr. Mehrdad Gholami, for his guidance and expertise which has been the light to my journey.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

%A	Absorbance
%R	Reflectance
%T	Transmittance
°C	Degree centigrade
°K	Degree kelvin
µg kg^{-1}	Microgram per kilogram
µg/day	Microgram per day
µgL^{-1}	Microgram per liter
µm	Micrometer
‰	Part per thousands
ANN	Artificial neural network
ANOVA	Analysis of variance
C_3	Reductive pentose phosphate cycle
C_4	Dicarboxylic acid cycle
CA	Canonical analysis
CAM	Crassulacean acid metabolism
CF-IRMS	Continuous flow isotopic ratio mass spectrometer
CV	Coefficient of variation
DA	Discriminant analysis
DF	Degree of freedom
DA	Discriminant analysis
DI-IRMS	Dual-inlet isotopic ratio mass spectrometer
EA-IRMS	Elemental analyser isotopic ratio mass spectrometer
FA	Factorial analysis

FTIR	Fourier transform infrared spectroscopy
FTNIR	Fourier transform near infrared spectroscopy
GC-FID	Gas chromatography-flame ionization detector
GC-MS	Gas chromatography mass spectrometry
HCA	Hierarchical cluster analysis
HR-ICP-MS	High resolution inductively coupled plasma mass spectrometer
HTC-IRMS	High temperature conversion isotopic ratio mass spectrometer
HTCR-IRMS	High-temperature carbon reduction-Isotopic ratio mass spectrometer
HTP-IRMS	High temperature pyrolysis
ICP-AES	Inductively coupled plasma atomic emission spectroscopy
ICP-MS	Inductively coupled plasma mass spectrometer
ICP-OES	Inductively coupled plasma optical emission spectrometry
IR	Infrared spectroscopy
IRMS	Isotopic ratio mass spectrometer
KNN	K nearest neighbor
LDA	Linear discriminant analysis
LOD	Limit of detection
LOQ	Limit of quantification
LS-SVM	Least Squares Support Vector Machines
Max	Maximum
mbar	Milibar
Mean Abs Dev	Mean absolute deviation
mg kg^{-1}	microgram per kilogram
mg	milligram
Min	Minimum

MIR	Magnetic resonance imaging
MLP-ANN	Multilayer perceptron - artificial neural network
N	North
N	Number
NAA	Neutron activation analysis
ND	Not detected
NIR	Near infrared spectroscopy
NIST	National institute of standard and technology
nm	nano meter
NMR	Nuclear magnetic resonance
Nov	November
NW	North west
ORS	Octopole reaction system
p	Probability level
PCA	Principal component analysis
PLS	Partial least square
PLS-DA	Partial least square- discriminant analysis
ppb	Parts per billion
RMSE	Root-mean-square error
SD	Standard deviation
Sep	September
SNF	Solid nonfat
SRM	Standard reference material
UHT	Ultra-high temperature processing milk
UV/Vis	Uv-Violate/visible
UVE-SPA-LS-SVM	Variable elimination- successive projections algorithm -least

	square support vector machine
VPDB	Vienna pee dee belemnite
VSMOW	Vienna standard mean ocean water
W	Watt
WHO	World Health Organization
XRF	X-ray fluorescence
δ	Isotopic ratio

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Background of study

Today, quality and safety of food are of special concern as they are directly related to human health. Consequently, reliable methods are needed to ascertain the quality and safety of food (Bunaciu *et al.,* 2016). One of the most important sources of nutrients for humans are milk and dairy products since they provide micronutrients (enzymes, elements and vitamins) as well as macronutrients (lipids, carbohydrates and proteins) needed for the body (Bunaciu *et al.*, 2016; Domingo *et al.*, 2014; Hilding-Ohlsson *et al.*, 2012; Souza *et al.,* 2011). Furthermore, as mammalian milk is the first and the only source of nutrients for all mammals providing all the nutrients and energy needed for a proper growth during the early years, it is important to study the fingerprint nutritional value and chemical composition of milk (Pereira, 2014). This study will most notably lead to the geographical origin of the product.

1.2 Geographical origin of milk

Clustering food samples based on their chemical composition is of utmost importance for the determination of geographical origin and authenticity (Palacios-Morillo *et al*., 2013). In recent years, after the advent of food borne diseases there has been an increasing demand in obtaining information on the authenticity, geographical origin, agricultural practices safety and quality of food all over the world (de la Guardia & Illueca, 2013).

The two most considered topics in adulteration of foodstuff for food producers and consumers are food authenticity and its geographical origin hence information on food composition and geographical origin of the food is important especially in the case of animal products (Zhao *et al.,* 2013). Adulteration of food products could be carried out in various ways such as changing the whole or part of the product with other cheap

material or even by adding flavor and aroma to improve the cheap product so that the poor quality is not easily recognized (Hrbek *et al*., 2014; van Leeuwen *et al*., 2014). As milk is one of the seven important and most frequently consumed foods that are exposed to adulteration (Moore *et al*., 2012)the need for verification of the geographical origin of milk and dairy products is in demand (Baroni *et al*., 2015) as these could be indicators of their authenticity and quality (Cheajesadagul *et al*., 2013).

Milk adulteration could be due to addition of melamine, salt, sugar, whey or mixing of different types of milk (Kartheek *et al*., 2011). Another aspect of milk adulteration is in the form of labeling (Molkentin & Giesemann, 2007). Milk adulteration could also be related to methods used in the processing technology and also geographical origin (Hrbek *et al*., 2014). In other words, geographical origin of milk is related to the nature of the food consumed by the cow, which depends on the processing conditions, technology and quality of the raw materials; and the characteristics of the raw materials depends on the environmental conditions of the geographical location (Cubero-Leon *et al*., 2014; Monfreda, 2012).

Consequently determination of the geographical origin of milk is undeniably linked to its authenticity (de la Fuente & Juarez, 2005). The adulteration, geographical origin and authenticity of milk and dairy products can be ascertain by several analytical techniques such as (UV/Vis), (NIR), fluorescence spectroscopy, (ICP-MS), (IRMS), (NMR) and a few other method (Kamal & Karoui, 2015). In this work, IRMS, ICP-MS, ultrasonic scanner (Milkoscan), UV/Vis and NIR spectroscopy are used to determine the elemental concentration, isotopic ratio and nutritional values of milk.

1.2.1 Multi elemental information

Throughout the world special attention is given to heavy metals as they can cause environmental pollution. Presence of heavy metals in food can also lead to health risk

problems (Saei-Dehkordi & Fallah, 2011). If animals graze on a contaminated land, contamination will be present in their meat and milk (Licata *et al*., 2012). Moreover, contamination of food with heavy metals and other pollutants have increased due to the use of new technologies in food production (Shahbazi *et al*., 2016). Dairy produce such as milk is one of the most important natural foodstuff that consist of various trace and minor elements. Although the amount of metals in dairy products is comparatively low, their concentration could vary depending on many factors. For example, the elemental composition of food products can be linked to the geochemical and geological condition of the milking farms (Baroni *et al*., 2015). Some of these factors are environmental such as geographical location of the farm, natural food source and season and others are related to physiology of the animal, species and lactation state (Caroli *et al*., 2009; O'Brien *et al*., 1999; Rodriguez *et al*., 2001). Moreover, the concentration of elements in milk could signal that there is pollution in the environment. There are studies that show that elemental composition of food reflects frequently the environmental and geographical conditions of the location where the foodstuff are produced (Baroni *et al*., 2011; Di Paola-Naranjo *et al*., 2011; Pellerano *et al*., 2012; Podio *et al*., 2013).

Metals in human body are categorized as essential and non-essential elements (Fraga, 2005). The essential elements are important for normal physiological functions of the human body so deficiency or absence of these essentials can lead to some disorders (Hsieh *et al*., 2011; Komatsu *et al*., 2012; Schroeder *et al*.,1966). Among the trace elements there are essential ones for metabolism and growth and these are zinc, chromium, manganese and copper (Stawarz *et al*., 2007). Although trace elements are necessary for biological functions of the body they could be toxic in high doses, where they go beyond the body's needs (Fraga, 2005). Among the trace elements Zn, Mn, Fe and Cu are needed in sufficient amounts to control physiological functions although

consuming them in excessive amounts can cause toxic effects (Nagpal *et al*., 2012; Pereira, 2014; Zamberlin *et al*., 2012).

Toxicity of heavy metals is of high concern as well although they have several industrial uses, hence, it is necessary to control their level. Measurement of trace and minor elements is a must in controlling the quality of milk during production and manufacturing (Khan *et al*., 2014). Most of the unsafe metals and metalloids enter the food chain from agricultural and industrial environment. The presence of metal pollutants such as Pb and Cd which enter the bodies of animals and humans via food chain could prove toxic (Zheng *et al*., 2007). Knowing that food is the most abundant source of toxic elements such as Hg, Ni, Pb and Cd, high levels of these metals can cause carcinogenic or neurological disorders (Organization, 1996). In addition, metals such as nickel, cobalt, cadmium and chromium might enter the feed and environment of cow farms, passing into the milk and milk products, thus causing various problems (Schuhmacher *et al*., 1991).

Food cycle could act as a source of Cd and Pb if the soil that plants grow in is polluted and significant amounts of these pollutants could enter into the greens consumed by animals, especially cattles (Alonso *et al*., 2002; Miranda, Lopez-Alonso *et al*., 2005). Accumulation of Pb and Cd cause toxic effects in cattle, as well as human's upon consumption of milk and meat polluted by these elements (Cai *et al*., 2009; Gonzalez-Weller *et al*., 2006; Vromman *et al*., 2008). Pb and Cd poisoning is more common in ruminant animals in the farms as they are more sensitive to these metals (Swarup *et al*., 2005).

Since the last couple of years, milk pollution is one of the most important issues frequently discussed globally. Among metals and metalloids Fe, Cu, Cd, Pb and Zn are mainly linked to toxicity and pollution issues (Malhat *et al*., 2012). Metals can easily

bio-accumulate in milk and milk is consumed in high amounts by children and infants. Consequently, determination of metal residues in milk are of high concern (Tripathi *et al*., 1999). In this research, the content of metals on milk was determined by Inductively Coupled Plasma Mass Spectrometry (ICP-MS).

1.2.2 Multi isotopic information

Multi analysis of stable isotopes of C, N, O and H is important value in determination of geographical origin of food. Isotopic ratios are used to analyze the stable isotopes in the tissue and produce of animals which are representative of the feeding regime of the cattle and might vary depending on the produce or tissue. Information on the isotopic ratios can point to their geographical origin.

It is important to mention that multi elemental stable isotope ratios of C, N, O and H in animal products such as meat and milk as well as tissues such as hair and blood are correlated with the water and food consumed by the animal. Stable isotope ratios might change depending on the tissue or the lipid content. As the isotopic ratios of diet are related to the origin and environment of cattle farms, multi isotopic ratio analysis is one of the possible spectroscopic methods used for determining the geographical origin and authenticity of milk and dairy products, discriminating the milk of cow with different feeding regimes and production system in various regions (Bontempo *et al*., 2011; Bontempo *et al*., 2012; Donghui Luo *et al*., 2015). Multi elemental isotopic ratio is commonly used to differentiate between milk of different areas that have used different production system and diet as the water and food consumed by animals are related to the environmental factors (Bontempo *et al*., 2012; Chesson *et al*., 2010).

An example could be the isotopic ratios of C and N in living beings are related to those in the environment. Isotopic ratios of C and N in milk can fluctuate depending on the food and water consumed by the animal. Isotopic ratio of oxygen for example

explains the amount of oxygen in milk water which varies with geographical factors such as longitude, latitude and distance from sea or seasonal differences. In the summer for example, animals eat more fresh grass enriched in oxygen-18 due to the more efficient evaporation of water in plants, therefore the water of milk has higher oxygen-18 content (Chung *et al*., 2014; Molkentin, 2013; Sacco *et al*., 2009).

Among the isotopic ratios studied in food products, isotopes of nitrogen and carbon are important in living beings as they are both related to the isotopes found in the environment. In milk, $\delta^{15}N$ is related to the use of fertilizers and soil type whereas $\delta^{13}C$ is related to the diet of animals and most importantly to the pathway of plants either C_4 or C_3 (Scampicchio *et al*., 2012). $\delta^{18}O$ in milk is related usually to the ground water and describes the water taken by the animals and can be connected to geographical factors such as latitude, altitude, distance from sea as well as seasonal effects (Chung *et al*., 2014; Molkentin, 2013; Sacco *et al*., 2009). To ascertain these ratios, Isotopic Ratio Mass Spectroscopy (IRMS) was employed.

1.2.3 Nutritional information

The nutritional importance of milk in human diet relates to its possible role in preventing several chronic disorders such as obesity, diabetics, cancer and cardiovascular diseases (Pereira, 2014). Besides minerals and trace elements there are a few other factors that influence the quality of milk such as lactose, fat and lipids (Rebechi et al., 2016; Ruth, Bremer *et al*., 2009), water, solids and proteins. In this work, the nutritional information was determined by ultrasonic scanner and UV/Vis and NIR spectroscopy. These nutritional factors are also important in ascertaining the quality and geographical origin of milk.

1.3 Objectives of this study

The aims of this research is to measure metal content, isotopic ratio and nutritional information of raw and factory cow milk samples obtained directly from farms or from supermarkets in Malaysia and from selected countries using analytical methods such as ICP-MS, IRMS, ultrasonic scanner (Milkoscan) and UV/Vis and NIR spectrometry. The physico-chemical characteristics of milk are important parameters for the authentication and adulteration aspects of milk and dairy products (Gabbi *et al*., 2013).These characteristics can also be used to identify the geographical origin of the produce. The obtained physico-chemical data obtained from the various analytical methods are then analyzed using various chemometric tools to determine the geographical origin of the milk samples.

The detailed objectives of this research are as follows:

a) To determine the physico-chemical characteristics of the milk via different analytical methods. The analytical methods include:
   i) ICP-MS: To determine the elemental composition of samples.
   ii) IRMS: To determine the isotopic ratio information of samples.
   iii) UV/Vis, NIR: To determine the nutritional values of milk samples by studying their absorbance at different wavelengths
   iv) Ultrasonic scanner: To determine fat, solids-nonfat (SNF), density, protein, lactose, salts and water contents in milk samples as well as their freezing point and temperature.
b) To determine physico-chemical characteristics of cow hair as a representative of animal tissue (instead of milk) using ICP-MS and IRMS
c) To ascertain the geographical origin of cow milk (and cow-hair) based on various chemometric analyses of the physico-chemical data.

1.4 Thesis outline

This thesis contains five chapters.

Chapter 1. General introduction to analytical information and methods that have been used in this research in order to determine geographical origin of milk.

Chapter 2. Literature review of the various work carried out on milk and dairy products using various analytical methods.

Chapter 3. Discussion of the research methodology, instruments used in the chemical and physical analyses of milk samples. Also includes discussion on statistical and chemometric methods used for analyzing the data obtained from the chemical and physical analyses.

Chapter 4. Results and discussion contain subsections on ICP-MS, IRMS, ultrasonic, UV-Vis and NIR analyses as well as various chemometric data analyses carried out on the physical and chemical analytical results.

Chapter 5. Conclusion

CHAPTER 2: LITERATURE REVIEW

2.1 Quality of milk

Milk is a fluid that is secreted from female mammary glands to meet primary nutritional needs of their infants (Nunez-Sanchez *et al*., 2016; Trienekens & Zuurbier, 2008). Milk produced from the mammary gland has been the food for mammals since the time of their birth, which ensures proper growth and nutrition. Milk provides all the nutrients needed for a proper growth and is especially important in bone mass formation. The nutritional importance of milk in human diet and its possible role in preventing several chronic disorders such as obesity, diabetics, cancer and cardiovascular diseases is confirmed by several epidemiological studies (Pereira, 2014).

Milk is a mixture of several substances such lactose, lipids, proteins, amino acids, urea, creatinine, etc. As milk is highly nutritious, it is purchased and consumed by a large population nowadays. The portion of nutrients in milk is determined according to human needs and these are the things that will affect its market price.

Parameters that influence the color, flavor and composition of milk are the cow's breed, diet, lactation state, farming system, physical environment and season. In other words milk and dairy products are nutrient rich food can be used in occasions where there is limited access to animal source products or foods are low in fat.

Although consumers can judge the quality of milk based on color, smell, taste and some other factors but these are not accurate, hence accurate analytical methods must be used for the determination of the quality of milk. For this reason, consumers are always concerned with the origin of the food they consume. They may also be concerned with information on the farming methods and the animals' diet. This is because grass feeding or pasture grazing are factors that can be linked to the quality of milk and dairy products (Elgersma *et al*., 2006). In fact, food quality and safety have

always been important issues as food is directly related to human health. Thus, there is need to evaluate the quality of food and among the concerns of food quality is its geographical origin.

2.2 Geographical origin of food

Geographical origin of foodstuff can give an indication of the authenticity and quality of the product. The geographical origin of food is in turn related to the raw material quality. High quality raw materials guarantee the quality of the food that humans consume. Information on raw food materials are important for determination of the quality of the final product (Monfreda, 2012; Rutkowska *et al*., 2015). Geographical origin of milk is highly desired for the determination of the authenticity of dairy products (Brescia *et al*., 2005). Problems related to food authentication are mostly related to mislabeling and adding or removing of a foreign substance (Drivelos & Georgiou, 2012). By ascertaining the geographical origin of food, problems of authenticity can partly be solved.

The most frequent methods used for the determination of geographical origin of food produce are mass spectroscopic, spectroscopic and separation techniques. Mass spectroscopic techniques measure the mass to charge ratio of ions. In general, mass spectroscopy measures the composition of a sample by measuring the masses of the samples detected in a mass spectrum. After the sample is ionized, ions of different masses are separated and by measuring the intensities of ion flux the relative abundance is recorded. Among the mass spectroscopic methods used in food analysis are the isotopic ratio mass spectrometer (IRMS), inductively coupled plasma mass spectrometer (ICP-MS) and GC-MS (Gas chromatography mass spectrometer).

Spectroscopic methods use electromagnetic radiation to stimulate the sample and different electromagnetic response such as absorption, reflection, emission verses

stimulated wave length is observed in the form of spectrum. Most used spectroscopic techniques for determination of geographical origin of food are NMR, IR, and Atomic spectroscopy (Luykx & Van Ruth, 2008).

Separation technique is one of the other methods used to ascertain food adulteration. It works based on the affinity of a substrate to differentiate different materials. Usually the substrate is called the stationary phase, and the mixture of materials are allowed to pass through the stationary phase and are separated from each other because of the different affinity to the stationary phase. Among the various separation techniques used for this purpose are gas chromatography and high performance liquid chromatography (Diaz *et al*.,2005; Macatelli *et al*., 2009; Rutkowska *et al*., 2015).

In this research we have used ICP-MS, IRMS, ultrasonic scanner, UV/Vis and NIR spectrometry in ascertaining the factors that allow for the determination of geographical origin of milk, cattle food and cattle hair.

2.3 Instruments

Various spectroscopic techniques have been used to determine the micro and macro elements in milk samples. These are ICP-MS, IRMS and UV/Vis and NIR spectrometers.

2.3.1 Mass spectroscopic techniques

Mass spectrometric techniques works by measurement of the mass to charge ratio of ions (Aebersold & Mann, 2003) and are applied to samples to elucidate their composition which is manifested by a mass spectrum. Upon ionization, the ions of different masses are separated the abundance of the ions are measured by the ion flux

intensities (Luykx & Van Ruth, 2008). Amongst the instrumental techniques that uses spectroscopic methods are ICP-MS and IRMS.

2.3.1.1 Inductively coupled plasma mass spectrometry (ICP-MS)

ICP-MS was first invented in 1980's (Houk *et al*., 1980) and since then it has served as a versatile detection technique for mineral and trace elements in sample matrices (Gray & Date, 1983; Houk, 1986; Profrock & Prange, 2012). The first ICP-MS was introduced in 1983 by Perkin Elmer.

In milk, the amount of minerals is not constant and varies due to different aspects such as environmental condition, the animal's nutrition, stage of lactation and animal breed. The values reported for minerals are mostly due to the above reasons and some could be from contamination due to milking and processing equipment or analytical errors (Caroli *et al*., 2009; Kalač & Samková, 2010). The chemical form of trace elements and macronutrients present in milk is of importance since it might influence intestinal utilization and absorption (Cashman, 2002).

Trace elements such as Zn, Mn, Co and Cr are of utmost importance for normal growth and metabolism (Joint *et al*., 2007; Stawarz *et al*., 2007). Metals are important in the physiological functions of animals and humans as they are essential cofactors of enzymes and thus any deficiencies may cause problems. The amount of metals in pure milk and dairy products is usually small but the content varies according to packing and manufacturing processes.

There are metals such as Cd, Pb, Ni, Co and Cr, of which high levels of concentration might contaminate the environment and thus transfer of these metals to milk and milk products often causes serious problems (Schuhmacher *et al*., 1991). Toxic metals and metalloids frequently enter the food chain from agricultural and

industrial sources. Metal pollutants such as Cd and Pb which have toxic effects on human and animals and usually enter the body of living beings via the food chain, mostly via polluted plants. These elements can be transferred from contaminated soil to grass and plant causing accumulation of toxic metal in grazing animals (Alonso *et al*., 2002; Miranda *et al*., 2005; Pilarczyk *et al*., 2013) and also in humans who consume meat and milk contaminated with these toxins (Gonzalez-Weller *et al*., 2006; Vromman *et al*., 2008).

One of the essential trace elements for human health is Se (Rayman, 2000, 2008) which is recommended to be taken at 55 µg/day by adults but there are a documentations that reports that intake up to 200 µg/day will protect the human body from diseases such as cancers which is caused by free radicals (Alzate *et al*., 2010; Clark *et al*., 1996).

Toxicity of heavy metals on humans and animals is of concern because these metals have extensive industrial uses. The effect of some elements are cumulative, hence it is important to control their level in consumed foods. Consequently, it is important to measure the concentrations of essential and trace elements in assessing the quality of milk during production and manufacturing (Khan, *et al*., 2014).

Trace elements are of importance due to their relationship with the environment and health of humans, plants and animals. Elements in environmental matrices have trace (10 mg kg^{-1}) or ultra trace (1µg kg^{-1}) concentration levels. Elements such as Fe, Cu, Mn and Zn are essential micronutrients if no more than few milligrams of them are consumed each day. However they might be harmful if consumed more than the recommended doze. Toxicity of elements varies, hence heavy metals intake from food chain is of main importance in assessing the health risk of humans who consume them. Environmental pollutions have made the heavy metal concentration an important issue.

By ingestion and inhalation, heavy metals enter a human body, however, the ingestion pathway depends on the food consumed. For all of the essential elements there is a range of intake which is adequate for the body and beyond this range, toxic effects and deficiency can be noticed. WHO has provided information on dietary intake of various elements (Patra *et al*., 2010).

Milk can carry various xenobiotic substances which cause risk factors for dairy products and more importantly for human health. Determining the residual concentration of elements in milk could be of importance in its hygienic status as well as ascertaining the degree of pollution of the milk from the environment it was produced in (Gonzalez-Montana *et al*., 2012; Pilarczyk *et al*., 2013). One of the main problems with milk is its ability to bio-accumulate metals (Tripathi *et al*., 1999).

The variation in the multi elemental content of milk samples from different regions could also be explained by seasonal variation, soil type, altitude, latitude, distance from sea, rain fall and the effects of industrial manufacturing processes (Oliveira *et al*., 2015). Different analytical methods have been used across the world to determine essential and trace elements in various food products (Reinholds *et al*., 2015). Among the analytical methods used is inductively coupled plasma mass spectrometry (ICP-MS) (Barbosa *et al*., 2014; Batista *et al*., 2012; Borges *et al*., 2015; Chevallier *et al*., 2015; Kruzlicova *et al*., 2013; Miller-Ihli & Baker, 2001; Millour *et al*., 2011; Potortì *et al*., 2013; Ródenas *et al*., 2009) inductively coupled plasma optical emission spectrometry (ICP-OES) (Bressy *et al*., 2013; Fallah *et al*.,2011; Kira & Maihara, 2007; Larrea-Marin *et al*., 2010; Luis *et al*., 2015).

For most cases, ICP-MS is used for analysis of elements in food (Chudzinska & Baralkiewicz, 2011; Giannenas *et al*., 2009; Nardi *et al*., 2009; Tuzen *et al*., 2007). ICP-MS is believed to be the best for elemental analysis (Yilmazcan *et al*., 2014) due to its

ability to carry multi elemental and isotopic analysis, its high sensitivity and ability to carry out bulk sample analysis (Thomas, 2013).

Among the most recent studies carried out on various kind of milk and dairy products using ICP-MS was on the elemental composition of donkey's milk where researchers have analyzed donkey milk for its nontoxic and toxic elements and they noted that although it is rich in Se, Fe, Zn and Co, seasonal variation in the concentration of elements were noticed as well. The researchers also found that by using elemental analysis of milk, it is indeed possible to determine the geographical origin of milk (Potortì *et al*., 2013).

In another work the multi elemental composition of milk and yogurt from South Korea was studied. The researchers used heating block for digestion of their samples. It was found that the concentration of mineral and trace elements were similar for the different kinds of milk and yogurt under that study. Se was found to be high in milk and Mn and Co were high in fruity yogurts (Khan, Jeong, *et al*., 2014).

In a research done on butter samples from black sea region of Turkey ICP-MS was used to determine its elemental content where they dissolved the ash from butter samples to detect the elements and then analyzed the samples. However the concentration as they of the elements analyzed was not related to the geographical origin of the butter (Dervisoglu *et al*., 2014).

Table 2.1 lists down other works that had been carried out on foodstuff using various instruments including the ICP-MS.

2.3.1.2 Isotopic ratio mass spectrometry (IRMS)

Isotopic ratio analysis is a new technology firstly introduced by European wine industry to confirm the authenticity of the geographical origin of their products in order

to prevent adulteration of wine (Almeida & Vasconcelos, 2001; R. Crittenden *et al*., 2007) and since then it has been used for determining the authenticity of food especially when traditional methods are not accurate enough (Guo *et al*., 2010).

Quality control of food products is dependent on analysis of chemical composition of food stuff but for this, significant and fast methods are needed. Recently there had been studies done on isotopes done on milk and dairy products (Bontempo *et al*., 2012; Chung *et al*., 2014; Donghui Luo *et al*., 2015).

IRMS is one of the ways for tracing food as it analyzes the isotopic ratio of food products in order to verify the geographical origin. The stable isotope ratios of $^{2}H/^{1}H$, $^{13}C/^{12}C$, $^{15}N/^{14}N$, $^{34}S/^{32}S$ and $^{18}O/^{16}O$ represent the characteristics of the substance analyzed (Diomande *et al*., 2015) and are mostly used for evaluation of food (Luykx & Van Ruth, 2008) as well as milk (Crittenden *et al*., 2007; Gonzalvez *et al*., 2009; Longobardi *et al*., 2015).

In general, the isotopic ratio of elements in plant materials gives information on the composition of isotopes. The isotopic ratio of carbon is related to the carbon fixation pathway namely the so called C3, C4 and CAM pathways (Guler *et al*., 2014; Wang *et al*., 2015). Depending on the pathways a plant uses, the $^{12}C/^{13}C$ ratios will be different. C4 plants will have isotopic ratio of carbon varying from -10 ‰ to -20 ‰ and C3 plants from -22 ‰ to -33 ‰ while for CAM the range is much larger. Unlike carbon, nitrogen isotopic composition depends on the soil composition and nutrients (Elflein & Raezke, 2008; Kropf *et al*., 2010). $\delta^{16}O$ gives the oxygen isotopic composition and it is related to the extent of precipitation and ground water (Craig, 1961; Gray & Thompson, 1976; Suzuki *et al*., 2008).

There has been many research done on geographical origin of foods using IRMS such as a study carried out on rice in Japan where IRMS was used to differentiate samples based on the geographical origin using isotopic composition and the researchers reported that $\delta^{13}C$, $\delta^{15}N$ and $\delta^{18}O$ was reliable in discriminating geographical origin of rice (Suzuki *et al*., 2008). Isotopic ratio information had also used to differentiate Italian, Spanish and French cheese according to variability factors such as geographical origin, seasonal and climate condition and the diet of the animal (Camin *et al*., 2004).

Generally, EA-IRMS has been used for determining the geographical origin as well as detecting fraud and mislabeling of products in many fields such as archaeology, geochemistry, food science, biochemistry, petroleum science and others (Ambrose & Deniro, 1986; Hayes *et al*., 1990; Rozanski *et al*., 1992). IRMS is considered a very efficient method in various aspects of research especially authenticity of food products as it is able to discriminate products based on their origin (Longobardi *et al*., 2011).

Traceability and authenticity of food products is important for consumers and producers as each of the food products are linked to specific geographical regions which may have faded over time due to globalization of food industry (Drivelos & Georgiou, 2012). Verifying food authenticity is important for safety and quality control of food products. Consumers may want to seek information on the production region of the product especially when there is an outbreak of diseases (Gonzalvez, Armenta, & de la Guardia, 2009). Camin et al in (2008) proved that carbon and nitrogen isotopic ratios are reliable parameters for the determination of geographical origin of milk (Camin *et al*., 2008). Matteo et al (2012) confirmed this by analyzing milk as well as forage samples from different farms (Scampicchio *et al*., 2012). On a different note, to avoid mislabeling, a robust method for verifying authenticity of milk needs to be developed.

Elements and isotopic ratio in milk of bio elements is the best way to do this (Camin *et al*., 2012; Gonzalvez *et al*., 2009).

Kelly (2003), Bateman (2009) and Grundy et al. (2012) mentioned that isotopic ratio analysis of food samples can also provide information on added flavors, water, sugar and other additives to food products besides diet of the animal and also the variation in organic and conventional farming.

The food and drinks consumed by humans contain hydrogen, oxygen, carbon and nitrogen isotopes and these are indicators of geographical origin and climate of that particular origin (Manning & Soon, 2014). To this effect, there had been some new research who that use IRMS as a tool to differentiate the geographical origin of food products. There had been recent studies carried out isotopic ratio of milk and dairy products using IRMS such as a study done on pure milk of some selected namely Australia, New Zealand, France, Germany, U.S.A and China to determine their geographical origins. The researchers extracted protein from milk samples and analyzed them for carbon and nitrogen content. They also used raw milk samples in order to analyze for hydrogen and oxygen isotopic ratios and reported that four isotopic ratios of carbon, oxygen, nitrogen and hydrogen are able to discriminate the geographical origin of the milk samples (Donghui Luo *et al*., 2015).

A study carried out on cattle hair for stable isotopic ratio in order to predict the geographical origin of beef from different regions of China showed that there were considerable differences between carbon, nitrogen and hydrogen isotopic values of hair from the different regions (Liu *et al*., 2013). Moreover, isotopic ratio analysis combined with NMR had been used in order to differentiate conventional milk samples from organic milk. This was also done by protein isolation or extraction of fat from milk samples and they found that the combination of methods in separating these samples

was much more effective than only considering one single method (Erich *et al*., 2015). Cheese from Italy was had also been used in the analysis of isotopic ratios of hydrogen, carbon, nitrogen and sulfur to confirm the authenticity of the dairy product. Casein from cheese samples was extracted from the samples and analyzed, confirming IRMS to be a suitable method for determining authenticity of cheese (Camin *et al*., 2012).

Another study was carried out on forage, milk and cheese from two different sites in Italy where they were analyzed for carbon, oxygen, nitrogen and hydrogen isotopic ratios. The researchers found that isotopic information gave good separation of the products from different types of pasture at two different regions (Bontempo *et al*., 2012). Isotopic ratio of carbon and nitrogen was also used to determine the authenticity of organic milk samples from Korea where the researchers observed that $\delta^{13}C$ value was higher in organic milk compared to conventional milk but for $\delta^{15}N$, the organic milk had a lower value (Chung *et al*., 2014).

Isotopic ratio information besides being used for determination of geographical origin can also be used to differentiate milk samples prepared with different heating programs based on a study on raw, UHT and pasteurized milk in Italy. They researchers found that heat changes the composition of milk, which in turn changes the isotopic ratio of carbon and nitrogen. They observed that UHT and pasteurized milk cannot be separated by $\delta^{15}N$ and $\delta^{13}C$ information and discrimination is only seen between raw and processed milk (Scampicchio *et al*., 2012). Description of more work done on milk using IRMS can be found in Table 2.1.

2.3.2 UV/Vis and NIR spectroscopic techniques

Spectroscopy is the study of interaction between light and matter. Spectroscopic techniques are widely used in the analysis of foodstuff and discussion of some of the methods used are presented in this section. In this work, UV/Vis and NIR spectroscopic

measurements were carried out. There had been several studies on determining the geographical origin of food, particularly milk and dairy products for the different regions of the electromagnetic spectrum. However, information on geographical origin of milk (freeze dried and plasma) based on UV/Vis and NIR region is reported for the first time in this research, as far as we know. Nevertheless, UV spectroscopy have been used to study milk for other objectives. A study done in Switzerland determined the content of protein and casein in milk which used UV spectroscopy to do this. They observed determination of casein and fat is possible by the use of fourth derivative spectra obtained from UV spectrophotometer (Luthi-Peng & Puhan, 1999). There was another work that measured fat content in milk using UV spectrophotometric method where the authors measured the absorption of the sample at 208 nm in a UV/Vis spectrophotometer. They claimed that the information obtained correlated with the results of milkotester, a device that measures fat by IR spectroscopy (Forcato *et al*., 2005). There was also a study done on Emmental cheese from selected countries which used combined infrared and fluorescence spectroscopic methods to determine its geographical origin. The authors mentioned that NIR and MIR results were comparable and the classification of samples was good.The differentiation between emmental cheese made from thermalized and raw milk using fluorescence spectroscopy was also obtained (Forcato *et al*., 2005). Another research done on cow milk to determine its geographical origin and authentication of cow feeding using visible and near infrared spectroscopy had also been carried out. In the research near infrared spectroscopy was able to separate pasture and non-pasture milk although it did not trace the milk's origin when the feeding effect was isolated from the altitude effect (Coppa *et al*., 2012).

2.3.3 Ultrasonic scanner

One of the most powerful and widely used methods in food science sensory descriptive analysis which explains sophisticated sensory information on food products is the ultrasonic scanner (Varela & Ares, 2012).

Dairy products sensory analysis explains human reaction to nature of the product by evaluating five senses of sound, smell, tough, sight, taste (Bodyfelt *et al*., 2008). Sound, aroma, flavor, aftertaste, texture and appearance which are characteristics of food products are qualitative factors which differentiate samples from one another. There are various research that had been done on adulteration of milk and dairy products based on their sensory information (Kamal & Karoui, 2015).

Sensory properties of dairy products could be explained due to its aroma, flavor, texture, appearance and color (Brighenti *et al*., 2008). Authors have also studied the sensory characteristics of cream cheese, that contains different levels of fat and found that there were significant differences in cheese fat content (Brighenti *et al*., 2008).

2.4 Methods used for data analysis

Chemometrics has become a standard tool in analytical and food chemistry. The huge amount of data obtained from various analytical instrument prompts the use of chemometrics in analyzing these data. There are various pattern recognition tools in chemometrics which differ in the ways of clustering the samples (Massart *et al*., 1997). Various supervised and unsupervised pattern recognition methods have been used in food analysis. The methods used are grouped into two: first group are the ones that look at discrimination of clusters such as PCA, HCA, CA, DA, LDA and KNN. The second group are methods that deal with modeling of class analogy which will not be discussed in this research (Berrueta *et al*., 2007). The first group are methods that fall under unsupervised pattern recognition. Besides these methods, data on this thesis were also

analyzed using artificial neural networks (ANN) which can be grouped under artificial intelligence methodology.

Chemometric tools are often used to discriminate various geographical sites on the bases of trace and major elements, isotopic ratios and nutritional information. Some of the applications of chemometry in food science are studies done on milk (Bassbasi *et al*., 2014; Bontempo *et al*., 2012; Erich *et al*., 2015; Souza *et al*., 2011) cheese and ice cream (Monakhova *et al*., 2013). Description of some of the works on food products which employed chemometrcis can be found in Table 2.1.

2.5 Various works on food analysis based on the tools used for this thesis

This part describes in detail the works that had been carried out on foodstuff based on the tools used in this thesis. The table includes the references and a short description of the work done. Of course the reviews are not exhaustive but they do indicate the usefulness of the methods used in this work in determining the geographical origin of milk of Malaysia and some selected countries.

Table 2.1: Reviews on food analysis; (Ataro *et al*., 2008); (Bilandzic *et al*., 2015)

Publication information	Summary of the work
Quantification of trace elements in raw cow milk by ICP-MS.	24 freeze dried raw cow milk samples collected from dairy farms close to mines in Gauteng and North West Provinces of South Africa were digested using supra pure and microwave digestion system which consumed small amount of sample and reagent and then analyzed for elemental concentration of V, Cr, Mn, Sr, Cd and Pb using ICP-MS. The precision of the method was checked by NIST SRM 8435. In conclusion they found that the elemental concentration of these elements followed the SRM of milk.
Determination of Macro- and Microelements in Cow, Goat and Human	45 cow and 14 goat milk samples were collected from farms in Croatia including

Milk using ICP-OES.	5 human breast milk samples and digested using analytical grade acid and base in high pressure laboratory microwave oven in order to look for concentrations of Ca, Na, Zn, Mg, Sr, Cr, Mo, Fe, K, Mn, Zn, Se and Cu. These elements concentrations are compared for the three kind of cow, goat and human milk It was concluded that the highest concentration of Ca, Na, Zn and Sr was found in cow milk while in Mn, K, Mg, Fe, Cr, Mo and Li was more in goat milk and Cu was the highest in human milk. The content of Mo was the same for cow and goat milk.

Table 2.1 continued: (Khanmohammadi *et al*., 2014); (Khan, Choi, *et al*., 2014); (Kim *et al*., 2015); (Zhao *et al*., 2013)

Publication information	Summary of the work
Classification of persimmon fruit origin by near infrared spectrometry and least squares-support vector machines	FTNIR had been used to determine the authenticity of persimmon fruits from different regions of Spain and it was observed that the best separation was seen by applying LS-SVM classification method.
Determination of mineral elements in milk products by inductively coupled plasma- optical emission spectrometry.	6 elements of Ca, Mg, K, Na , P and Fe were studied in 207 milk products purchased from supermarkets in South Korea which were powdered and analyzed with ICP-OES after the samples were digested using wet digestion and microwave-assisted combustion with analytical grade nitric acid and hydrogen peroxide. The accuracy of the method was checked by milk SRM (NIST-1549). They found that elemental concentrations were very close to one another in different milk products. Furthermore, the most abundant elements were Ca, K, P, Na, Mg and Fe. The concentration levels were within the recommended dietary allowance.
Applicability of stable C and N isotope	C & N isotopic ratios in three commercial

analysis in inferring the geographical origin and authentication of commercial fish (Mackerel, Yellow Croaker and Pollock).	fish from different countries were measured using IRMS analysis in order to determine the traceability and authenticity of this product. They concluded that both C & N can be used to trace the traceability and authenticity of commercial fish from different regions.
Multi-element composition of wheat grain and provenance soil and their potentialities as fingerprints of geographical origin.	Concentration of 22 elements in wheat grain and soil in China were analyzed by high resolution inductively coupled plasma mass spectrometry (HR-ICP-MS) and x-ray fluorescence (XRF) and the relation between wheat and soil was studied. The elements associated with soil were used to differentiate wheat by applying PCA and LDA. As a result there was significant correlation between elements in soil and wheat.

Table 2.1, continued: Reviews on food analysis; (Andueza *et al*., 2013); (Orun *et al*., 2011); (Rey-Crespo *et al*., 2013); (Wu *et al*., 2009)

Publication information	Summary of the work
Using Visible or near infrared spectroscopy (NIRS) on cheese to authenticate cow feeding regimes.	Two spectroscopic methods UV-VIS and NIR were used to differentiate between cheese samples from two cow diets. The reflectance spectra of fresh and freeze dried samples were recorded in both ranges and there was no difference observed when fresh and freeze dried sample were used furthermore, NIR is able to cluster cheese samples from different regimes.
Breast milk lead and cadmium levels from suburban areas of Ankara.	144 breast milk samples were collected from mothers in suburban area who were not occupationally exposed to toxic metals. Dried milk samples were weighed and digested by 65% nitric acid, (concentrated grade) in microwave oven and then analyzed by ICP-MS. They noted that Pb concentration was more than the limit in breast milk reported by the World Health Organization (WHO).
Essential trace and toxic element concentrations in organic and conventional milk in NW Spain.	50 samples of milk from organic and conventional farms in NW Spain were collected and the concentration of trace

	and toxic elements were determined after digestion with acid using ICP-MS. They found that the concentration of essential trace element in organic milk were lower in comparison to conventional milk
Exploring Near and Mid infrared Spectroscopy to Predict Trace Iron and Zinc Contents in Powdered Milk.	Seven brand of powdered milk samples produced in China were analyzed using NIR and MIR spectroscopy. They concluded that that MIR spectroscopy combined with UVE-SPA-LS-SVM could be used as a rapid method to determine the concentration of trace elements such as zinc and iron in powdered milk

Table 2.1 continued: Reviews on food analysis; (Rutkowska *et al*., 2015); (Bassbasi *et al*., 2014); (Caredda *et al*., 2016); (X. Liu *et al*., 2013)

Publication information	Summary of the work
Differentiation of geographical origin of cream products in Poland according to their fatty acid profile.	63 cream samples were collected from dairy plants in three different regions of Poland in order to be used to identify the region of origin using gas chromatography and by applying chemometric methods. From 44 fatty acid identified two groups were distinguished and it was concluded that cream origin could be identified by gas chromatography together with chemometric analysis of milk fat.
Prediction of the geographical origin of butters by partial least square discriminant analysis (PLS-DA) applied to infrared spectroscopy (FTIR) data.	54 butter samples were analyzed using FTIR together with chemometric methods to separate butters from different locations of Morocco. It was shown than FTIR together with chemometrics is a powerful technique used for this discrimination. PCA showed that the important absorption bands were in the range of 3000–600 cm^{-1} consisting of bands associated with protein and lipid compound absorptions. PLS was 100% able to assign the origin and quality of butters.
Prediction of fatty acid content in sheep	250 sheep milk samples were collected

milk by Mid-Infrared spectrometry with a selection of wavelengths by Genetic Algorithms.	from farms in northern, central and southern locations of Sardinia in order to build a prediction model for measuring fatty acid composition in sheep milk using mid infrared spectrometry to evaluate milk quality control in order to be able to set a fair price for this product. It was mentioned that the models had good predictive ability.
Stable isotope analysis of cattle tail hair: A potential tool for verifying the geographical origin of beef.	Stable isotopic information of cattle tail hair was used to classify beef from different regions of China. Variance analysis and LDA was used to differentiate hair samples using isotopic ratios of C, N and H. As noted the stable isotopic information of cattle tail hair can be used in establishing beef origin.

Table 2.1, continued: Reviews on food analysis (Ehtesham *et al*., 2015); (Osorio *et al*., 2015); (Coppa *et al*., 2015); (Erich *et al*., 2015); (H. Zhao *et al*., 2011)

Publication information	Summary of the work
Influence of feed and water on the stable isotopic composition of dairy milk.	18 dairy farms in New Zealand were studied to find possible correlation between isotopic composition of δ^{2}H, δ^{18}O, δ^{13}C and δ^{15}N in milk fatty acid, water and feed consumed by animals and it was mentioned that a significant relationship was observed.
Multi-Element (C, N, H, O) Stable Isotope Ratio Analysis for Determining the Geographical Origin of Pure Milk from Different Regions.	60 samples of milk, cheese and plant were collected from three farms in different locations of Cyprus where major and trace elements were studied using ICP-AES. It was noted that milk and cheese showed significant differences in the elemental composition.
Potential of milk fatty acid composition to predict diet composition and authenticate feeding systems and altitude origin of European bulk milk.	Milk samples were taken from 10 different European countries to study the potential of milk fatty acid to predict cow diet composition.
Combined chemometric analysis of 1H NMR, 13C NMR and stable isotope data to differentiate organic and conventional milk.	85 pasteurized and (UHT) milk samples were collected from organic and conventional farms in south Germany plus samples from Australia and Swiss. Data

	obtained from chemometric analysis of 1H NMR-, 13C NMR-spectroscopy, stable-isotope data (IRMS) and α-linolenic acid content (gas chromatography) were used to map organic and conventional milk samples. When mixing the data from various methods more efficient differentiation was observed than when single method data was used.
Determining the Geographic Origin of Wheat Using Multielement Analysis and Multivariate Statistics.	Wheat from four major areas of China was analyzed for 15 elements. Analysis of variance and linear discriminant analysis was used for analysis and as it was concluded multi elemental analysis is a promising result for determination of origin of wheat.

Table 2.1, continued: Reviews on food analysis; (Capuano *et al*., 2014); (Capici *et al*., 2015); (Huck-Pezzei *et al*., 2014); (Bilandzic *et al*., 2015)

Publication information	Summary of the work
Verification of fresh grass feeding, pasture grazing and organic farming by FTIR spectroscopy analysis of bovine milk.	116 milk samples were collected from Netherland and analyzed by FTIR. Classification method of PLS-DA was used to dibranchiate fresh grass feeding from pasture grazing and organic farming. They came to a conclusion that spectra from FTIR analysis gives information on cow's food which could be used for authentication of food.
Determination of Cheese Authenticity by Carbon and Nitrogen Isotope Analysis: Stelvio cheese as a Case Study.	In this research IRMS is used to trace milk samples used for production of Stelvio cheese in Alpine and they found that isotopic fractionation does not happen during the manufacturing period of Stelvio cheese regardless of the kind of milk used. They mentioned that IRMS was able to trace milk samples during cheese manufacturing and significantly detect addition of milk or reconstituted powdered milk with different isotopic values.
Alps food authentication, typicality and intrinsic quality by near infrared	Cheese, meat and apple samples from the Alp's region were collected and then by

spectroscopy.	the use of NIR the origin was defined. After applying multivariate analysis they confirmed that NIR is a nondestructive fast and cheap method for determining the origin of foods in the Alps region as well as for other regions.
Differences in macro- and microelement contents in milk and yoghurt.	The study was done on different samples of milk and yogurt to evaluate the concentration of macro elements and essential elements in samples collected from Croatian market using ICP-OES. They mentioned that concentration of Ca, K, Na and Cu were higher in milk than in yogurt and so elemental variation compared to other countries confirms the good process production of yogurt.

Table 2.1, continued**:** Reviews on food analysis; (Yang *et al*., 2015); (Matteo Scampicchio *et al*., 2015); (Valenti *et al*., 2013)

Publication information	Summary of the work
Discrimination and characterization of different intensities of goaty flavor in goat milk by means of an electronic nose.	Goaty flavor in goat milk samples was measured with electronic nose based on metal oxide sensors. The found that there was a good correlation between electron nose evaluation and sensory evaluation and that electron nose together with LDA and PCA as well as some other methods differentiated among goat milk samples with goaty flavor milk samples.
Multi-method Approach to Trace the Geographical Origin of Alpine Milk: a Case Study of Tyrol Region.	Raw milk samples were collected from north and south Tyrol and the UHT samples from European Union that were discriminated according to their geographical origin heat treatment and season of production using different techniques of IRMS, MID-NIR,NIR and GC-FID. They came to a conclusion that each of these methods separately gave limited amount of separation while combined together gave more obvious separation of their geographical origin.
Infrared spectroscopic methods for the	Bulk milk samples from France were

discrimination of cows' milk according to the feeding system, cow breed and altitude of the dairy farm.	differentiated according to their feeding system using MIR and NIR. MIR method differentiated milk from hey and pasture fed and those from maize silage and pasture feed. MIR did not show discrimination between milk from hey and maze silage feed. Similar results was obtained with NIR but with lower efficiency. Two methods did not show discrimination of breed. There were significant differences for discrimination of samples using these two infrared methods based on feed and breed clustering while there were no significant differences between the two methods used for discriminating upland and low land samples.

Table 2.1 continued**:** Reviews on food analysis; (Y. Zhao *et al*., 2014); (Codina-Torrella *et al*., 2015); (Karabagias *et al*., 2013); (Casale *et al*., 2015)

Publication information	Summary of the work
Recent developments in application of stable isotope analysis on agro-product authenticity and traceability.	Several agro samples from animal and plant source were analyzed using IRMS for authenticity plus factors such as altitude, latitude, evaporation and climate condition. They noted that IRMS is a reliable method for determining the geographical origin of agro products
Characterization and comparison of tiger nuts (Cyperus esculentus L.) from different geographical origin Physico-chemical characteristics and protein fractionation.	Four types of tiger nuts from Spain, Burkina Faso and Niger were compared according to their physico chemical characteristics and protein characterization. Nuts from Spain were the largest. More over the Spanish tubers were the richest in protein and fat and African tubers had similar values. There was little amount of tannin and phytate detected in all of the samples. Results showed that physico-chemical characteristics of tiger nuts were dependent on their geographical origin although protein was similar in different origins.

Classification of Western Greek virgin olive oils according to geographical origin based on chromatographic, spectroscopic, conventional and chemometric analyses.	47 olive oil samples were collected from western Greek island and analyzed for 14 physico-chemical parameters. The conclusion was that when physico-chemical characteristics are combined with chemometrics there was indeed possibility to differentiate olive oils.
NIR spectroscopy as a tool for discriminating between lichens exposed to air pollution.	NIR was used to investigate the fingerprint of lichens. PCA was applied to the spactra and it was observed that the separation of samples due to their species was much less than the level of environmental pollutants. Then LDA was used to differentiate samples based on their exposure to pollutants. Overall, NIR was able to separate the lichens from one another based on their exposure to different environmental pollutant.

Table 2.1, continued: Reviews on food analysis; (Longobardi *et al*., 2015b); (D. Luo *et al*., 2015); (Y. Wu *et al*., 2015); (Liu *et al*., 2013)

Publication information	Summary of the work
Discrimination of geographical origin of lentils (Lens culinaris Medik.) using isotope ratio mass spectrometry combined with chemometrics.	The geographical origin of lentils was determined using IRMS in combination with chemometrics. Samples were collected from two origins in Italy and Canada for isotopic ratios of C, N, O, H and S. All isotopic parameters were significantly different except N. Principal component analysis was applied for grouping the samples. Besides PLS and KNN algorithm were used for analysis and both models showed prediction abilities.
The application of stable isotope ratio analysis to determine the geographical origin of wheat.	35 wheat samples from different regions of Australia, China, America and Canada were collected and analyzed for isotopic ratios of C& N. The results showed that these milk can be separated from one another by isotopic ratios of C and N. The overall conclusion was that isotopic information can be used for determination of geographical origin of wheat and cereal grains.
Geographical origin of cereal grains based on element analyser-stable isotope ratio	C&N isotopic ratios of cereal grains from different origins were determined using

mass spectrometry.	EA-IRMS. The results pointed out that C isotopic ratio of rice, soya bean, millet, corn and wheat were significantly varied within the regions while isotopic ratio of N was not. In conclusion five kinds of cereals were separated thus comparison of C isotopic ratio is useful for determination of origin of cereal grains.
NIR Spectroscopy and Imaging Techniques for Evaluation of Fish Quality—A Review.	NIR and imaging techniques have been studied with regard to microbiological, sensory and chemical composition for determining the quality and authentication of fish and fishery products. It was observed that NIR provides more information on chemical composition compared to traditional methods. NIR is capable of differentiating different species of fish as well as fish feeding regime.

Table 2.1, continued**:** Reviews on food analysis; (Guo *et al*., 2013); (Muniz-Valencia *et al*., 2014); (Anjos *et al*., 2015); (dos Santos *et al*., 2009)

Publication information	Summary of the work
Chemometric Classification of Apple Juices According to Variety and Geographical Origin Based on Polyphenolic Profiles.	Apple juice samples were collected in China and were used to classify apple varieties and geographical origin on the basis of polyphenol composition. Chemometric methods, and it was observed that polyphenols can serve as indicator of apple variety and geographical origin.
Characterization of Mexican coffee according to mineral contents by means of multilayer perceptrons artificial neural networks	The concentration of Cu, Fe, Ca, Mn, Mg, K and Zn in roasted Mexican coffee analyzed by ICP-OES had been used to determine the origin of the samples. Initially, PCA was used for observing the natural distribution of samples followed by LDA and ANN. It was observed that ANN gave higher prediction ability than LDA.
Neural networks applied to discriminate botanical origin of honeys	Physical and chemical properties of 49 honey samples from different classes were used as input variables to build an ANN model capable of predicting the botanical origin of honey samples. As a conclusion electrical conductivity and

	colorimetric are reliable in predicting the origin of honey.
Chemical composition of bovine milk from Minas Gerais State, Brazil.	Composition of bovine milk samples from 32 dairy farms in Minas Gerais State were collected and the analyses were done on lyoph-ilized samples using ICP-MS and neutron activation analysis (NAA) and they concluded that there were significant differences in Ba, K, Na and Fat. Furthermore, the analytical methods used were able to detect differences in the quality of milk from different farm mainly for Na and K. They mentioned that Na/K ratio can be used to test the health of cattle

CHAPTER 3: METHODOLOGY

In this research different spectroscopic methods, namely ICP-MS, IRMS, ultrasound scanner and UV/Vis and NIR spectrometry have been used for analyzing milk and other related samples. They will be discussed in their individual sections. Gathering information on the geographical location of each farm was the first step in this research.

3.1 Geographical location of sampling

Table 3.1 shows the geographical information (longitude and latitude) of each farm where the samples have been taken during three different sampling times following two monsoon wind seasons, southwest (May-Sep) and Northeast (Nov-March). These information were used in the illustration of the maps using JMP®, Version <*Pro 12.2.0* > SAS Institute Inc., Cary, NC, U.S.A, 2015.

Table 3.1: Sampling sites

Origin	Sampling1		Sampling2		Sampling3	
	Latitude	Longitude	Latitude	Longitude	Latitude	Longitude
Perlis	-	-	6.501	100.312	6.654	100.274
Kedah	-	-	-	-	5.492	100.575
Pinang	-	-	5.415	100.195	-	-
Terengganu	5.321	103.027	5.268	103.094	4.992	103.067
Perak	4.631	101.105	4.683	101.173	4.799	101.091
Pahang	-	-	-	-	3.973	101.796
Kuala Selangor	3.364	101.313	3.364	101.313	-	-
Melaka	-	-	2.387	102.229	2.385	102.202
Johor	2.040	103.309	1.915	102.949	1.814	102.972

3.2 Sampling

Raw and processed cow milk samples have been collected from various farms and supermarkets both in Malaysia and some selected regions of the world. Farm samples were collected early in the morning. 1st sampling (S1) followed the northeast monsoon while 2nd (S2) and 3rd sampling (S3) was carried out during the southwest monsoon season. As shown in Figure 3.1 (b), 72 raw cow milk samples have been directly collected from farms in the northern region (Perlis, Kedah and Terengganu, Ipoh and Pinang) and southern region (Johor, Melaka, Selangor, and Pahang) of Peninsular Malaysia.

Upon travel to Iran we have collected seven raw cow milk samples directly from farms in the north and south of Iran. In addition to raw cow milks some factory milk samples from Malaysia and selected countries have been collected. The collected samples which were bought in supermarkets in Malaysia consist of Malaysian (5), New Zealand (1) and Australian (1) factory milks. Factory milk samples from other selected regions of the world were collected from supermarkets upon travel to those areas. The selected countries are Turkey (3), Iran (3), Azerbaijan (1), Belgium (1), Canada (1) and U.S.A. (2). Figure 3.1 (a) shows the sampling regions for either factory or raw cow milk from the selected countries.

Besides milk, various samples such as pellet, Napier grass, water, rain, hair, mixed food and soil were also taken from farms in Perlis, Perak, Terengganu, Johor, Melaka, Penang, Kuala Selangor and Kedah. However, farms in Pahang did not provide any of the said samples besides milk.

Tail hair samples from seven farms were analyzed in triplicate for the ICP-MS analysis. The origins are namely Perlis, Pinang, Terengganu, Perak, Kuala Selangor,

Johor and Melaka as shown in Figure 3.2. For the UV analysis, the sampling regions were the same as that of the ICP-MS analysis.

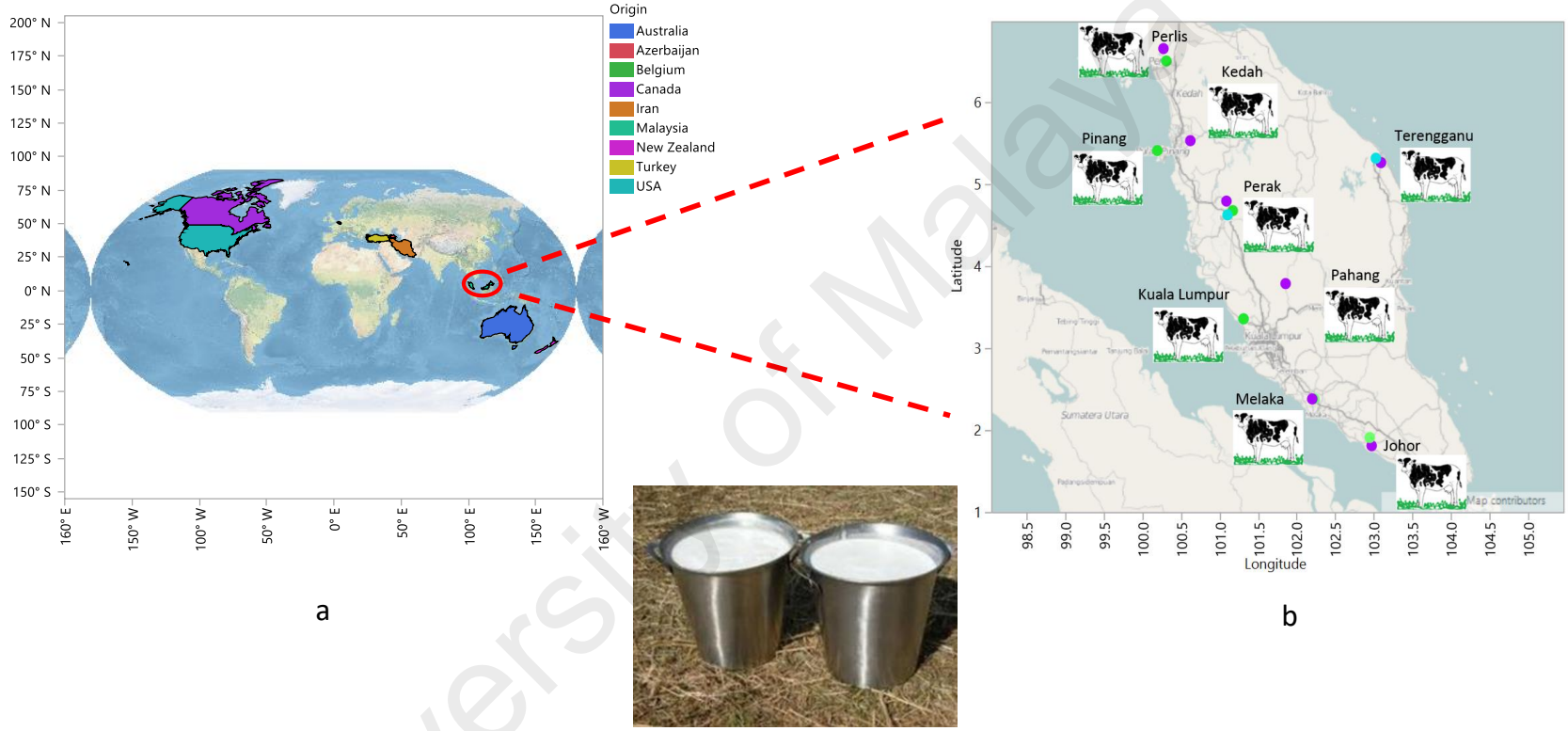

Figure 3.1: Milk sampling regions of the world and milk sampling sites in Malaysia.

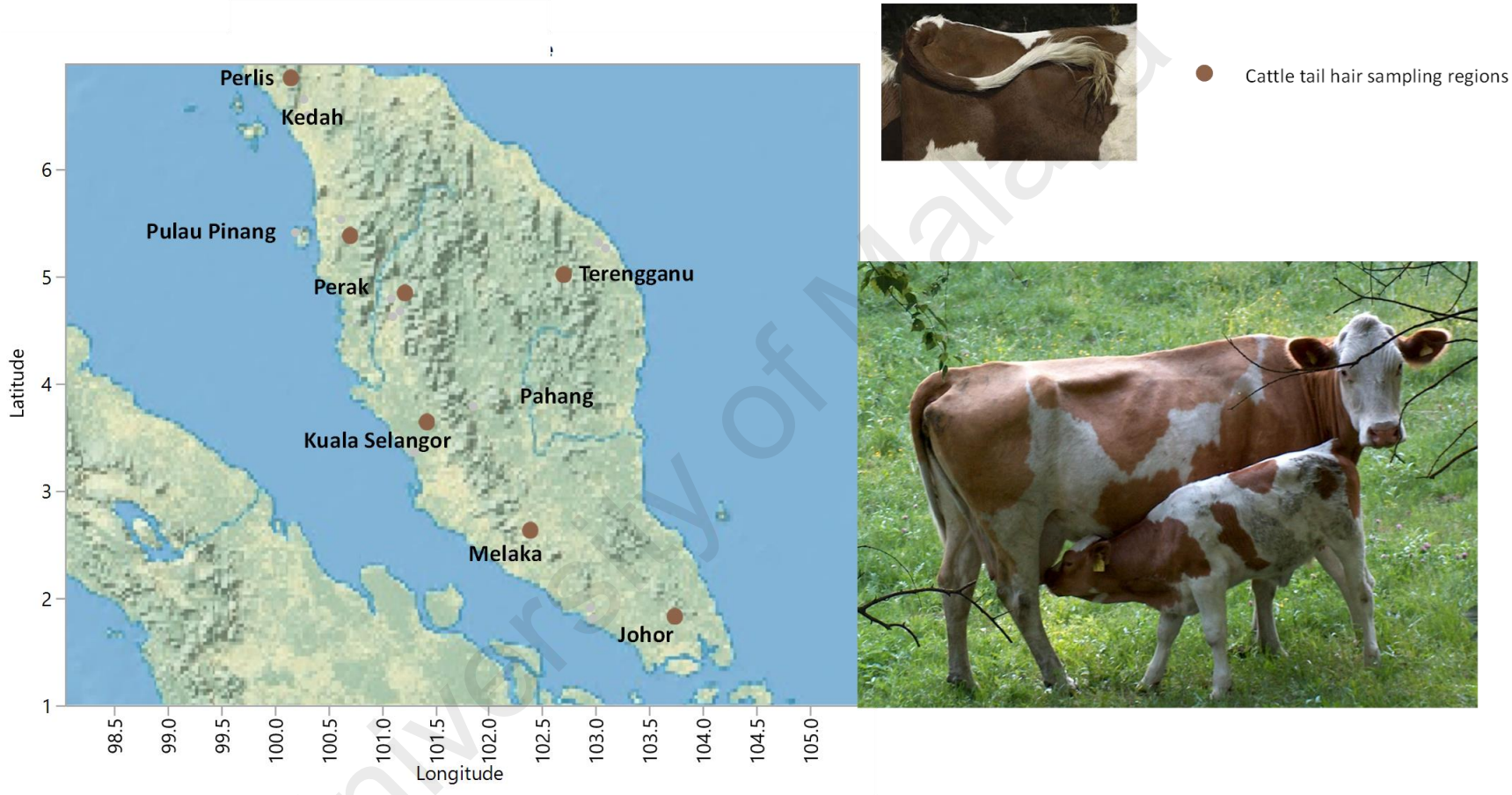

Figure 3.2: Tail hair sampling regions.

3.3 Analysis of samples by ICP-MS

3.3.1 Apparatus and Instruments

For preparing the samples to be analyzed by ICP-MS, two apparatus were used. PURELAB® UHQ II system (ELGA®, UK) with a resistivity of more than 18 MΩ cm was utilized to produce deionized water for the purpose of washing and rinsing of bottles as well as for the dilution of the digested samples. Microwave-assisted digestion (MAD) device, CEM Mars-Xpress (CEM Corporation, Matthews, NC, USA) was used for the digestion of freeze dried milk samples.

After digesting the samples in the microwave digestion system and diluting them with de-ionized water they were analyzed with ICP-MS, which is a technique for analyzing mineral and trace elements. ICP-MS 7500ce Agilent with an octopole reaction system (ORS) was used in our study. An octopole reaction system consists of an octopole ion guide inside a pressurized reaction cell and can be operated in both collision and reaction mode for the removal of polyatomic spectral interferences that can reduce the polyatomic ion effects by chemical reaction. The gas used for the whole experiment was Ar gas (99.999% pure). The instrument was tuned every day before the experiment with 1 ppb Agilent tuning solution consisting of Li, Co, Y, Ce and Tl in 2% HNO3 and 0.5% HCl. The setting of the instrument is reported in Table 3.2.

The main body of the ICP-MS instrument consists of the sample introduction, ICP torch, interface and cons, lens, quadrupole mass analyzer and ion detector (Thomas, 2001) as shown in Figure 3.3.

The liquid samples are pumped into the sample introduction which consists of the nebulizer and spray chamber where the samples were converted to small droplets which pass to the ICP torch consisting of plasma.

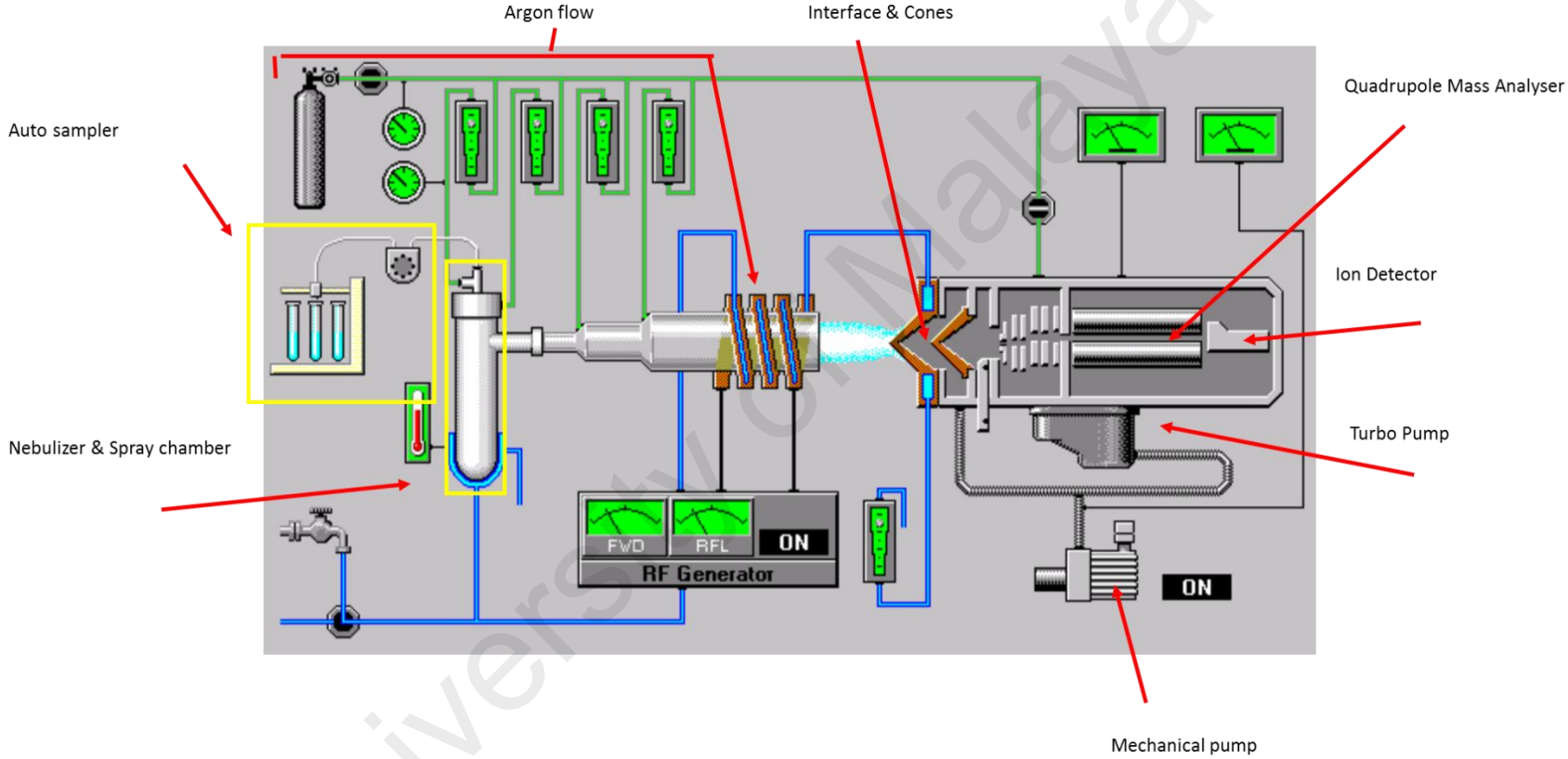

Figure 3.3: Main body of the ICP-MS instrument.

As it passes through the plasma torch, the sample dries, vaporizes, atomizes and ionizes. During this time liquid samples have changed to gas. At temperature of 7,500–10,000 K, the samples arrive at the analytical zone of plasma where they are excited and ionized. The detection and movement of these positively charged ions makes ICP-MS an ultra-trace detector. Because of the plasma's high temperature it is better to introduce the samples in their liquid form to ICP-MS (Profrock & Prange, 2012).

Table 3.2: Instrumental parameters and operating condition for ICP-MS 7500ce

Plasma Power	1500 W
Sample depth	6-8 min
Reflected power	<5 W
Plasma gas flow	15 $L.min^{-1}$
Carrier gas flow	0.8-1 $L.min^{-1}$
Collision gas flow	4.5-5 $L.min^{-1}$
S/C temp	2°C
Sampler and skimmer cons	Ni

The ICP-MS instrument is capable of measuring nearly most of the periodic table elements with detection limits as low as parts per trillion (Elmer, 2001; Profrock & Prange, 2012) as shown in Figure 3.4. However there are elements that either do not have naturally abundant isotopes or are not measurable by ICP-MS. The number of relative abundance of natural isotope for each element explains the isotopic finger print. Different number of neutrons in the nucleus of an isotope cause the naturally occurring isotopes in nature to have different atomic mass but the same atomic number. ICP-MS can detect more than one hundred isotopes with its high sensitivity (Liu *et al.*, 2014). The gases used for ICP-MS are mostly noble gases such as Ar.

Figure 3.4: Periodic table representing the elements measured by ICP-MS.

Quadrupole ICP-MS has served as a method used for multi elemental analysis, however, due to the presence of naturally formed interferences in argon plasma or matrix effects of the sample or the environment, precise detection has not been that easy. Table 3.3 lists the most common polyatomic interferences and the isotopes or modes of gas that could be used to eliminate these condition. The most important group of interferences in ICP-MS are spectroscopic ones which are caused when ions are produced from sample, plasma or a mixture of them which have mass to charge ratio equal to the analyte of interest (May & Wiedmeyer, 1998).

Table 3.3: Polyatomic interferences.

Interference	Isotope	Gas mode
ArAr	^{80}Se	^{2}H
$^{38}Ar\ ^{16}O\ ^{1}H$	^{55}Mn	He
$^{36}Ar\ ^{16}O$	^{52}Cr	He
$^{40}Ar\ ^{16}O$	^{56}Fe	He
$^{40}Ar\ ^{16}O\ ^{1}H$	^{57}Fe	He
$^{38}Ar\ ^{1}H$	^{39}K	He
^{40}Ar	^{40}Ca	He
$^{40}Ar\ ^{12}C$	^{52}Cr	He
$^{12}C\ ^{15}N$, $^{12}C\ ^{14}NH$	^{27}Al	He
$^{35}Cl\ ^{16}O$	^{51}V	He
$^{40}Ar\ ^{35}Cl$	^{75}As	He
$^{35}Cl\ ^{16}O\ ^{1}H$	^{52}Cr	He

The advantages of ICP-MS as multi elemental detective are its wide detection range, low matrix effects, low detection limit, high spectral resolution, high sensitivity and wide dynamic range. It is also able to obtain isotopic ratio information for elements having multiple isotopes when accurate quantitative results are needed (Wang *et al*., 2010). ICP-MS has been applied to various food samples in order to determine the elemental (Lachas *et al*., 2000; Vanhoe, 1993) composition of the product and furthermore its geographical origin (Ariyama & Yasui, 2006).

3.3.2 Sample preparation

Before sampling, all containers used for storing and analyzing samples were soaked in 10% nitric acid solution overnight, and then rinsed with ultrapure water with resistivity of 18Ω, before they were dried at room temperature.

Then, the milk samples were directly poured into the polypropylene bottles, labeled properly and transferred to the lab in an icebox. Samples delivered to the lab were kept in a -20 °C freezer. For preparing milk samples for ICP-MS analysis, 5 ml of liquid milk samples were poured in polyethylene bottles and then placed in the freezer for 24 hours. Christ Freeze Dryer model ALPHA 1-2 LO plus was used under a pressure of 0.089 mbar and a temperature of -49°C for freeze-drying the samples. The dried milk samples were then crushed using mortar agate until homogenized.

Food pellet samples were crushed by mortar agate and then sieved and the powder was kept in oven at 70°C for 1 hour until dried. As Napier grass is the most common grass eaten by the cows, the grass samples were also taken. The Napier grass samples were washed with deionized water and left to dry at room temperature on a piece of paper. They were then cut by a plastic cutter and left in the oven to dry at a temperature of 75°C for 8 hours. Then they were powdered with a mortar agate and sieved. The powder was taken and stored for digestion and analysis. Finally, after preparing all the samples, we attempted to produce mixture of food from all the food types consumed by the cattle in each single farm. To do this, we weighed 0.1mg of each food type given to the cows and placed them in a container. They were then mixed several times by shaking for a few days until they are homogenized.

It should be mentioned that besides all the samples that have been processed there were some additional foods which differed from farm to farm such as soya waste, palm waste, rice bran, wheat bran, maize, fish, fruit, bread and sago waste. So in the

mix food prepared for each farm we have added 0.1mg of whatever extra food type that was used by that farm.

Hair samples which were also collected from the tail of each cattle. They were washed with acetone and rinsed three times with de-ionized water in order to remove any possible contamination. In the case of soil samples, they were left at room temperature on a piece of paper to dry and they were then sieved and the powder was stored in dark glass containers until analysis. Cattle drinking water samples were collected in polypropylene bottles and kept until analysis by ICP-MS.

3.3.2.1 Reagents used for microwave digestion

60% ultrapure nitric acid and 31% ultrapure hydrogen peroxide (Merck Germany) were used for digestion of the samples in this research. We used Standard Reference Material 1849a (Infant/Adult Nutritional formula obtained from NIST U.S department), BCR-14R sludge, 1547 peach leave, 1573a tomato leave, 1570a Spinach leave and 1515 apple leave to check the validation of the methods used in this work.

3.3.2.2 Microwave digestion

For the digestion of the dried milk and all the other solid samples mentioned previously, a microwave oven with eight reaction vessels was used with a three step power/temp program. First, the power was set to 300 (W) for 2 minutes, then, the power was raised to 600 (W) for 10 minutes, and finally, the power was set to 300 (W) for 5 minutes. 0.1g of freeze-dried milk sample from each farm or factory was weighed and placed in Teflon bomb vessels for digesting powdered milk samples, and then under hood, 2.0 mL of ultra-pure 60% nitric acid and 1.0 mL of ultra-pure 31% hydrogen peroxide were added and left for 10 minutes to stabilize the reaction. After digestion, vessels were set aside under the hood to cool down and afterwards the content of each vessel was transferred into 50.0 mL polypropylene volumetric flasks. Then, the

volumetric flask solution was diluted with ultra-pure water to the marked level. Diluted samples were stored in polyethylene ICP vials until analysis is done using ICP-MS. After investigating the results we got from ICP-MS and calculating the recoveries, we believe that our method, which involves a smaller amount of sample together with very small volumes of acid and base, is as effective as other conventional digestion methods and contributes towards a greener environment.

3.3.2.3 Cleaning procedure for microwave vessels

The Teflon bombs were washed by the same program as digestion of milk samples except that the nitric acid used for washing the vessels was analytical grade Merck. After the microwave washing program was completed, the vessels were rinsed with deionized water three times.

3.3.3 Sample analysis

After digestion, the milk and other solid samples were filtered using 0.45µm filters and analyzed. Drinking water samples were directly analyzed without any sample preparation step.

The concentration of 24 elements in raw cow milk, soil, mixed food, grass, water, pellet, rain and hair samples were measured by ICP-MS 7500ce analyzer which was equipped with an octopole reaction system. Calibration curve was used for determining the concentration of elements in the digested samples. Each sample was divided into three and analyzed separately. Subsequently, the blank concentration was measured and subtracted from sample concentrations.

3.3.4 Standards

Agilent multi element environmental calibration standard consisting of 1000 mg L^{-1} (Ca, Fe, K, Mg, Na) and 10 mg L^{-1} (Ag, Al, As, Ba, Be, Cd, Co, Cr, Cu, Mn, Zn,

Mo, Ni, Pb, Sb, Se, Th, Tl, U, V) in 5% HNO3 was used to prepare different concentration of elements which are needed for plotting their calibration curve. All the standards were diluted using ultra-pure water.

3.3.4.1 Calibration curves for ICP-MS analysis

Agilent multi element standard with a concentration of 10 mg L^{-1} for Ag, Al, As, Ba, Be, Cd, Co, Cr, Cu, Mn, Mo, Ni, Pb, Sb, Se, Th, Tl, U ,V and Zn and 1000 mg L^{-1} for Ca, Fe, K, Mg and Na was diluted and used for calibration to carry out the quantitative analysis of the samples.

To plot the calibration curve, for Ca, Fe, K, Mg and Na five different calibration standards, of 50, 100, 300, 1000 and 5000 ppb, were used and for the rest of the elements concentrations of 0.5, 1, 3, 10 and 50 ppb were used. The correlation coefficient in the calibration curves varied from 0.998 to 1.000 depending on the element that shows good linearity in the range of the concentrations made. The calibration plots for each single element are reported in appendix B.

To overcome polyatomic interferences, different isotopes of the required elements were checked to avoid these interferences and finally the one having the least interference from the digested matrix was chosen. The mode of gas was changed in some cases to reduce interference.

3.3.5 Quality assurance

To confirm the analytical characteristics of the methods, the limit of detection (LOD) and limit of quantification (LOQ) were calculated and expressed with the unit of μgL^{-1}. These amounts are the minimum detectable values with signal to noise ratio of 3/1 and 10/1 respectively (Oliveira *et al*., 2015).

The quality of the readings were assured by running several blanks after each batch of samples. The limit of quantification for the digested samples was calculated as ten times the standard deviation of signal of blank over the slope of the calibration curve whereas the limit of detection is calculated as three times the standard deviation of the signal of the blank over the slope of the calibration curve.

The accuracy of the method was controlled by using the powdered infant milk formula Standard Reference Material (NIST SRM 1849a) which was digested and analyzed using the same exact method used for the other milk samples. Recoveries, coefficients of variation percent (CV %), limits of detection (LOD) and limits of quantitation (LOQ) are listed in Table 3.4.

We have also used other SRMs based on their similarity to our samples. These include silage SRM (BCR-14R), tomato leaves SRM (1573a), peach leaves SRM (1547), spinach leaves SRM (1570a) and surface water SRM (X). They were analyzed in the same way as other samples and reported respectively in Appendix A to check the accuracy of the methods used for analysis.

Milk SRM1849 is used to evaluate reproducibility, reputability and accuracy in each run. The closeness of our result to the certified value and the recoveries ranging from (95-104) % indicates the high accuracy of this measurement and as reported in Table 3.4 coefficient of variation (CV) values are good representative of repeatability and reproducibility. It is clear that this value is between 0.01 and 2.29 verifying the quality assurance.

Table 3.4: Quality control parameters for SRM (1849a).

Elements	Certified value (mg kg^{-1})	Obtained value (mg kg^{-1})	Recovery (%)	CV (%)	LOD (µgL^{-1})	LOQ (µgL^{-1})
Na	4265±83	4379±17.95	102%	0.41	0.346	1.15
Mg	1648±36	1646.44±12.93	100	0.78	0.157	0.52
K	9220±110	9510±91.29	104.2	0.71	3.408	11.36
Ca	5253±51	5290±71.94	100.7	1.36	20.620	68.73
Cr	1.072±0.032	1.06±0.01	98.67	1.44	0.122	0.41
Mn	49.59±0.97	51.4±0.29	103	0.57	0.064	0.21
Fe	175.6±2.9	181.35±4.15	103	2.29	0.227	0.76
Cu	19.78±0.26	20.17±0.003	101.1	0.01	0.134	0.45
Zn	151.0±5.6	143.48±1.95	95.02	0.5	0.038	0.13
Se	0.812±0.029	0.8±0.01	98	1.78	0.003	0.01
Mo	1.707±0.40	1.74±0.01	102	0.85	0.016	0.05

CV= coefficient of variation, LOD= limit of detection, LOQ= limit of quantitation.

3.4 Analysis of samples by IRMS

IRMS is one of the mass spectroscopic techniques used to measure the abundance of isotopes in a given sample. IRMS analysis is generally carried out in four steps:

- Combustion or thermal conversion of the samples in an elemental analysis
- Conversion of the sample gas into ions
- Isolation of the ion gas as well as separation and detection of ions in the mass spectrometer
- Collection analysis of data

These steps are implicitly discussed in the upcoming sections.

3.4.1 Apparatus and Instrument

For preparing the samples for IRMS analysis a freeze dryer (CHRIST, Germany) was used to dry liquid milk samples to form milk powder. The freeze dryer was first warmed up until the temperature of ice condenser falls down to around -49ºC and the pressure to around 0.089 mbar. The freeze milk samples were then placed in the chamber. All of the milk samples were left in the freeze dryer overnight to complete the freeze drying process. After freeze drying the milk samples they were ready to be analyzed by IRMS analyzer.

There are two kinds of IRMS instruments. They are the dual IRMS (DI-IRMS) and continuous flow IRMS (CF-IRMS). Continuous flow IRMS uses elemental analyzers (EA-IRMS). In this research stable isotope analyses were performed at the Malaysian Nuclear Agency using a SCRON GEO 20-20 (UK) combined with SECRON elemental analyzer.

3.4.2 Sample preparation for IRMS

After the milk samples were dried using freeze dryer, they were analyzed for isotopic ratios of C, N and O. The samples were crushed by mortar until homogenized and the analyses were directly carried out on freeze dried milk samples. In the case of hair the analysis was carried out directly on hair. Depending on the isotope wanted, the samples were prepared for analyzes either via combustion or pyrolysis.

For combustion, approximately 1.5 mg of freeze dried milk sample was weighed and placed in tin capsules for both C and N analysis using the two-step method. For pyrolysis, 0.5 mg of the freeze dried milk sample was weighed and placed in silver capsules for O analysis.

3.4.3 Standards

The variation in the values of stable isotope ratio are determined as part per thousand (‰) deviation from internationally accepted standard Mean Ocean Water (VSMOW) for oxygen, Vienna Pee Dee Belemnite (VPDB) for carbon atmospheric and nitrogen (AIR) for nitrogen. The relative deviation values were obtained using:

$$\delta\ ‰ = (R_{sample} - R_{standard} / R_{standard}) \times 1000 \quad (3\text{-}1)$$

In the above equation, R stands for the ratio of heavy to light isotope, hence, the bigger δ is, the more is the sample enriched in heavy isotope. Rsample represents the isotopic ratio of C, N, H and O in the sample and Rstandard represents the isotopic ratio of the reference material. Each of the samples was analyzed three times together with the standards to ensure the accuracy of the readings. In this work, R001 was used for C and N isotopes while NBS18 was used for O and NBS22 for hydrogen analysis. IRMS calibration was carried out using the δ-scale instead of m/z scale.

3.4.4 Sample analysis

Samples can be analyzed using EA-IRMS either in solid or liquid form using tin (C and N analysis) or silver (H and O analysis) capsules as the liquids have limited Viscosity to be analyzed directly using the liquid inlet. Key parts of EA-IRMS used to analyze carbon and nitrogen isotopic ratios in a variety of samples is shown in Figure 3.5.

For hydrogen and oxygen, analysis was carried out by decomposing the samples in high temperature and this could be either by using the high temperature conversion IRMS (HTC-IRMS), high temperature pyrolysis IRMS (HTP-IRMS) or high temperature carbon reduction IRMS (HTCR-IRMS). The schematics of the EA-IRMS used for high temperature pyrolysis is shown in Figure 3.6 Further explanation on how

the combustion and pyrolysis of samples before their analyses were carried out are explained in the next subsection.

3.4.4.1 Combustion (C and N analysis)

This technique involves the coupling of a preparation system employing the Dumas principle with a stable isotope mass spectrometer that analyzes samples using a two-step method. The analyzer has a two reactor system, a combustor and a reductor, which can also be combined in a tube. 1.5 mg of powdered freeze dried milk or tail hair samples containing carbon and nitrogen were weighed and loaded into tin capsules and sealed tightly. The samples are placed in an auto sampler and introduced to the elemental analyzer by dropping them one by one into a furnace at 1000^{o}C. The tin capsules burned exothermically, in an atmosphere of oxygen and the temperature rises to about 1800^{o}C, oxidizing the sample.

Complete combustion was ensured by passing the combustion products through a reduction tube of chromium oxide at 1000^{o}C, using a helium carrier gas. The products were then passed through a second furnace containing copper at 600^{o}C where excess oxygen is absorbed and nitrogen oxides are reduced to elemental nitrogen. Water was removed by a trap containing a hydrous magnesium perchlorate and carbon dioxide, by a trap containing Carbosorb™. The gas stream passes into a gas chromatograph where components of interest are separated and then bled into a mass spectrometer where the isotope species are ionised then separated in a magnetic field. The isotopic species were detected separately and from their ratios, the level of ^{15}N or ^{13}C were calculated. Calibration of the system is made using known standards allowing both total nitrogen and ^{15}N content (or total carbon and ^{13}C content) to be obtained from each sample.

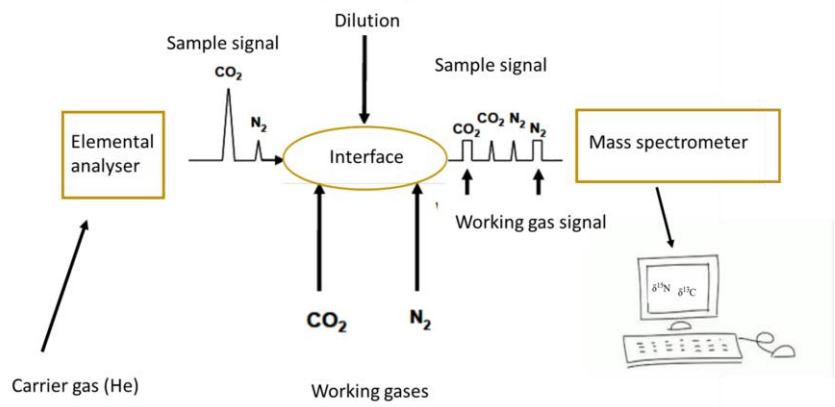

Figure 3.5: Schematic diagram of EA-IRMS used for C and N analysis.

3.4.4.2 Pyrolysis (O and H analysis):

Analysis of oxygen is carried out using pyrolysis. The samples are weighed and sealed in silver capsules. After 30 seconds the samples are dropped one by one from the auto sampler into the reduction tube where the temperature is around 1300°C. At this temperature, organic and inorganic compounds are converted to gases of CO, N_2 and H_2. The separation of the evolved gases is achieved via an isothermal packed column GC. The samples are then ready to be analyzed by the IRMS for their oxygen isotopes.

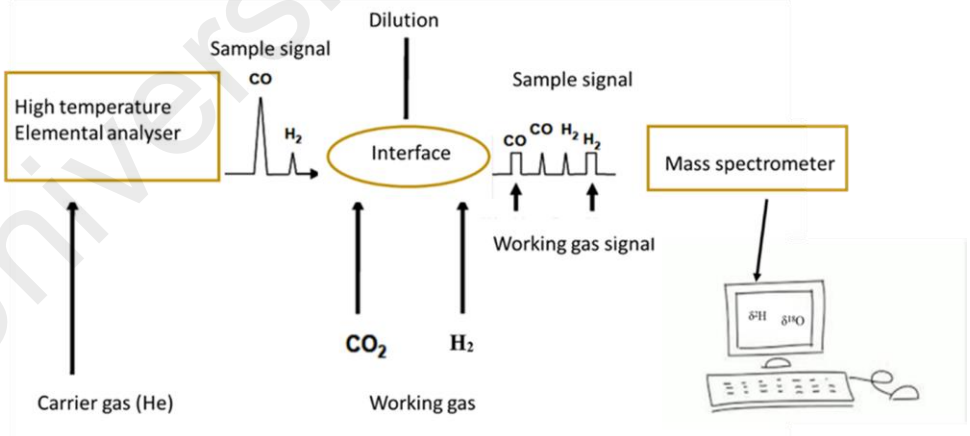

Figure 3.6: Schematic diagram of EA-IRMS used for O and H.

3.5 Analysis of samples by (UV/Vis and NIR)

3.5.1 Instrument

As milk is a colloidal liquid mix of proteins, lactose, lipid, amino acids and etc (Kamizake *et al*., 2003) determining its composition is important in the dairy industry in

order to establish its value both for quality control and consumers information. The colloidal nature of milk allows it to absorb and scatters light. Fats and proteins scatter light and the degree of scattering depends on the size of the scattering particles (Raty & Peiponen, 1999). Consequently, molecular spectroscopic methods have a wide application in food analysis because food is composed of fat, protein, carbohydrates and water.

In this thesis we have used the UV/Vis and NIR region of the electromagnetic spectrum to study the absorption of electromagnetic radiation by milk as is shown in Figure 3.7.

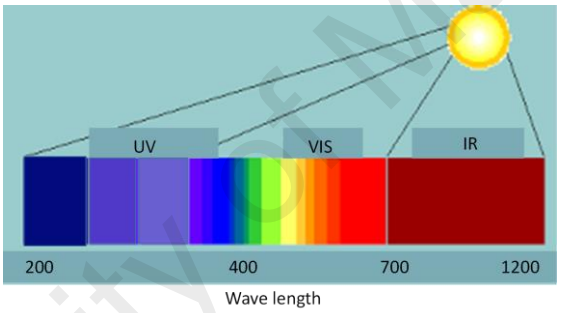

Figure 3.7: Wavelength range of electromagnetic spectrum.

A simple schematic of a typical spectrophotometer is shown in Figure 3.8 that consists of the light (UV/Vis/NIR) source, spectrometer, sample compartment and the detector.

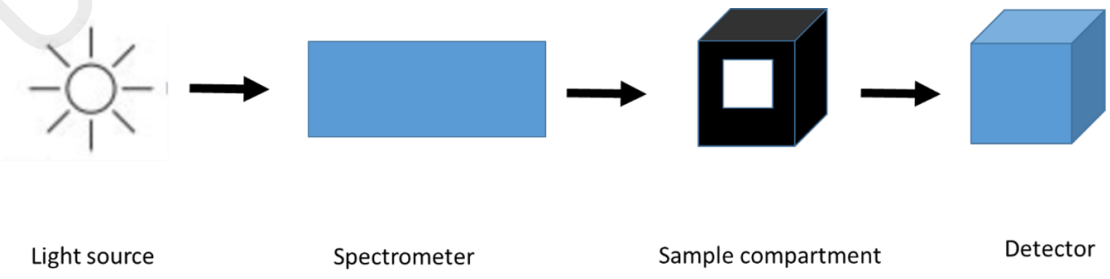

Figure 3.8: Schematic diagram of a spectrometer.

3.5.2 Sampling

The sampling regions involved are the same as had been discussed in the previous sections.

3.5.3 Sampling preparation

For analyzing the milk samples in solid phase the freeze drying system described in previous sections was used. In order to analyze milk samples in the liquid form with UV a centrifuge was used to separate the plasma.

3.5.4 Sample analysis

In this research UV-NIR 3600 Shimadzu spectrometer was used. For solid samples, the spectra were taken as reflectance measurements as shown in Figure 3.8. Initially, the intensity of light (I_0) is measured with barium sulphate as blank as is shown in Figure 3.9 (a). Then, as shown in Figure 3.9 (b) the intensity of the milk sample (I_R) was measured. The reflectance is then calculated by;

$$R = \frac{I_R}{I_0} \tag{3-2}$$

and the sum of reflected, transmitted and absorbed light is presented as;

$$\%A + \%T + \%R = 100\% \tag{3-3}$$

but since the transmittance is zero in solid samples, so the equation changes to:

$$A\% = 100\% - \%R \tag{3-4}$$

The absorbance is then calculated as:

$$A = \log_{10} \frac{1}{R} \tag{3-5}$$

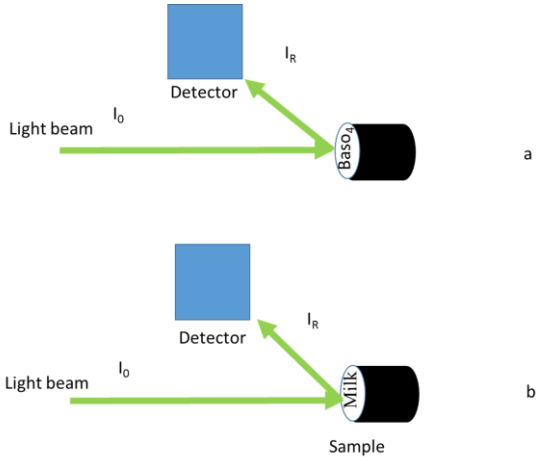

Figure 3.9: UV/Vis/NIR spectrometric measurements; a) blank, b) milk.

As noted the variation is observed in two regions of UV/Vis and NIR region which will be further analyzed in the result and discussion section. The difference between the spectra of the milk samples from various regions are linked to different factors such as the size of the particles in the milk as well as the contents of fat, lactose, protein and casein. In the infrared region the absorbance properties are determined by presence of functional groups such as –NH, -OH and –CH- that are responsible for the vibrational peaks. Moreover, components of fat, casein and lactose in milk have specific bands respectively as (1720, 1730, 1780, 2270, 2310, 2340) nm, (1450, 1680, 1720, 1730, 1780, 1820, 1980, 2100, 2310, 2340, 2790) nm and (1450, 1820, 2100, 2340) nm (Laporte & P. Paquin, 1999).

3.6 Analysis of samples by Ultrasonic scanner

3.6.1 Instrument

The milk analyzer used in this research is the Master Eco Milkotester which uses ultrasonic principles to ascertain the fat, lactose and various other contents of milk. Basically, this device contains piezoelectric transducers at two ends of an aluminum tube. One is used as a transmitter, and the other as a receiver as seen in Figure 3.10. Milk sample was pumped into the ultrasonic chamber, filling the tube. An ultrasonic pulse is sent through the milk sample. The time taken for the ultrasound wave to

traverse the distance between the transmitter and receiver was then determined. From the variation of sound speed at different frequencies, the concertation of different the components of the milk can be measured (De Luca *et al*., 2014).

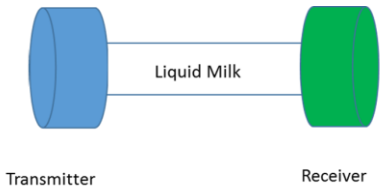

Figure 3.10: Schematic diagram of milkotester.

The advantages of ultrasonic techniques are that they are cheap, energy saving and simple and may be implemented online, hence, food quality can be evaluated in a safe, hygienic and economic way (Awad *et al*., 2012).

3.6.2 Sampling

Milk samples used are the same samples obtained from the sampling regions discussed previously.

3.6.3 Sample preparation

There was no sample preparation needed except that milk samples were shaken several times to bring the samples to a homogeneous state.

3.6.4 Sample analysis

Samples were poured into the analyzer's container were sucked into the ultrasonic chamber and heated to a uniform temperature. Ultrasonic pulses were transmitted through the sample in the ultrasonic chamber, after which absorption of the ultrasound were processed by the microprocessor and results were displayed.

3.7 Data analysis and data pretreatment

Statistical data analyses were carried out on the data obtained from the various physical and chemical analyses. This section discusses the data analysis methodology used to analyze the data.

Descriptive statistics and chemometric tools have been used to analyze the data obtained from this research. Prior to analysis, data pre-treatment was carried out. It has been noted that data pre-treatment is essential, especially standardization and normalization as without normalization of the data, the elements with higher concentration will have more influence compared to the low level concentration elements (Li et *al*., 2014).

Therefore, in this work before applying chemometric methods, data pre-treatment by either mean centering or z-transforming the data was done. Then, multivariate methods of Principal Component Analysis (PCA), Hierarchal Cluster Analysis (HCA), Canonical Analysis (CA), Discriminant Analysis (DA), factorial analysis (FA) and artificial neural network (ANN) were used to analyze the data.

3.7.1 Principal component analysis (PCA)

PCA is one of the unsupervised linear multivariate data reduction methods which is used for variable reduction and is based on the correlations between the variables (Brereton, 2003). PCA maps samples into new uncorrelated latent variables named principal components (PCs) (Granato et al., 2010). From the PC's two or three dimensional score plots can be drawn with loadings presented as vectors (Zhang *et al*., 2013).

The scores of a principal component is expressed with;

$$C_{ij} = b_{i1}X_{1j} + b_{i2}X_{2j} + \ldots\ldots + b_{im}X_{mj} \quad (3\text{-}6)$$

where C is score of the component, b is the loading of the component, X is the measured value of the original variable, i is the component number, j is the sample number and m is the total number of original variables.

PCA has been applied for determining the authenticity and geographical origin of milk and dairy products (Brescia *et al*., 2005; Guerreiro *et al*., 2013).

3.7.2 Hierarchical cluster analysis (HCA)

Hierarchical cluster analysis is carried out to determine the clusters of samples. It is also named as agglomerative clustering. HCA starts with each point (row) as its own cluster, then, at each step of the clustering process, it calculates the distance between each cluster, and groups the two clusters that are closer to each other. This grouping continues until all the points are in one final cluster. There are five rules for determining the distance between clusters, they are, the centroid, complete, average, Ward and single. Each rule can generate different sequence of clusters. However the one used for this research is Ward. In the Ward methodology, the distance between two clusters is the ANOVA sum of square between the two clusters added up over all variables. At each generation, the within cluster sum squares is minimized over all partitions which are obtained by merging the two previous clusters. The methodology is strongly biased towards producing clusters with approximately the same number of observations and is very sensitive to outliers. Distance for the Ward method is calculated from the equation below:

$$D_{KL} = \frac{\|\overline{X_K} - \overline{X_L}\|^2}{\frac{1}{N_K} + \frac{1}{N_L}} \quad (3\text{-}7)$$

Where,

$\overline{X_K}$: Mean vector for cluster C_K

N_K: Number of observations in C_K

$\|X\|$: Square root of the sum of squares of elements of X (The Euclidean length of vector X)

$\bar{X}$: Sample mean vector

The Ward's method joins clusters with a small number of observations Hierarchical cluster analysis displaces the data into two-dimensional space where Euclidean distances is measured between samples and presented into a matrix having similarity index of 0 to 1. HCA is mostly for classification purposes. HCA has been applied in so many cases such as in ascertaining the authenticity of milk (Brescia et al., 2005).

3.7.3 Discriminant analysis (DA)

DA is an unsupervised classification technique which uses a set of independent variables to find the linear combination of the variables that are responsible for clustering. This is done by building a discriminant function (DF) for each individual group (Härdle & Simar, 2012) which is calculated by the equation below:

$$f(G_i) = K_i + \sum_{j=1}^{n} w_{ij} P_{ij} \quad (3\text{-}8)$$

Where,

i: Number of groups

K_i: Constant inherent to each group

n: Number of parameters used to classify a set of data in to a given group

W_j: Weight coefficient assigned by DF analysis

P_j: Given parameter

DA takes into account the mutual relation of the distance of samples in order to reduce the similarity distance in a low dimension space (Gnanadesikan, 2011). In this research DA was applied to some parts of data to investigate the differences in the means of the variable and consequently using that variable for assigning group memberships.

Backward stepwise mode is applied for analyzing the data using DA which is done by selecting p values that are smaller than 0.05 and then eliminating variables that have p values larger than 0.05 while building the model.

There are four methods for conducting and fitting the discriminant method namely linear, quadratic, regularized and wide linear. The one used in this research was the linear method. In this method it was assumed that the within group covariance matrices are equal and the covariate means for the groups were assumed different. This was done by the stepwise variable selection which begins by selecting variables that have the smallest p values between all covariates and uses the backward stepwise methods which removes the least significant covariant from covariant entered but not locked. As validation is used prob>F values are based on training set. All the methods estimate the distance from observation to each group multivariate mean using Mahalanobis distance and observations are grouped to the closest group. Fitting method used in this research is linear which assumes equal covariance matrices with in a group.

3.7.4 Factorial Analysis (FA)

Factorial analysis is an explorative analysis similar to cluster analysis that describes observed variables in terms of unobserved latent factors which are represented as linear combination of the observed variables. Factorial analysis reduces the number

of variables by reducing the dimensions of the observations and helps in providing meaningful explanation of the observed variables via the latent factors. The key concept of factorial analysis is that the observed variables have similar pattern of response because they are all in relation with a latent variable (the latent factor) (Sall *et al*., 2012). The scores of a latent factor is given by;

$$A_{ij} = b_{i1}X_{1j} + b_{i2}X_{2j} + \ldots + b_{im}X_{mj} + e_1 \quad (3\text{-}9)$$

where A is the score of the latent factor component, b is the loading of the component, X is the measured value of the original variable, i is the component number, j is the sample number and m is the total number of original variables.

3.7.5 Artificial neural network (ANN)

Artificial neural networks are network models made of nodes consisting of layers which are connected to one another by weights. Artificial Neural Networks (ANNs) are improved machine learning modeling techniques used when other statistical methods are unable to explain the complete phenomena.

ANNs are able to learn and improve by experience and this is done by the help of specific algorithms. ANN simulates models of human brain and connections of neurons in the brain (Fausett, 1994; McCulloch & Pitts, 1943). Each layer has nodes and the weights of these nodes express the importance of each node (Anjos *et al*., 2015). MLP-ANN is the most common feed forward multilayer network consisting of neurons placed in layers and is usually trained by the backward propagation algorithm (Olszewski *et al*., 2008; Siripatrawan & Harte, 2007). The data are fed forward from the input layer, through the hidden layer and out towards the output layer without feedback. The backpropagation algorithm works on the correction of errors obtained from the

prediction of the network. These error corrections begins from the output layer, passes through hidden layer and finally to the input layer.

The input layer is mapped to the output layer via the hidden layer. The nodes in the hidden layer maps the input and output using transfer or activation functions. In this work, the software used for ANN modeling is the JMP Pro12 software which offers three type of activation function – TanH (sigmoid), linear and Gaussian. As an example, the hyperbolic tangent (TanH or sigmoid) function can be determined by equation (3-9).

$$\frac{e^{2x}-1}{e^{2x}+1} \tag{3-9}$$

The errors caused during the training time are continuously minimized to let the model learn the characteristics of the input data by improving its performance (Ripley, 2007). Since MLP models are supervised methods, they require dependent variables and in this work these are geographical region (N/S) (Cancilla *et al*., 2016). Designing an ANN model architecture is the most important aspect in obtaining a good prediction. The number of hidden layers and the size and training cycles are determined by randomly inputting different weight values and checking the accuracy of the ANN's prediction. The optimal number of epoch in ANN model is obtained when the error on the training set reaches minimum. To determine the best number of hidden nodes it is better to start with the simplest architecture having 1 hidden layer and adding one node at a time until the network produces the lowest error of prediction. Moreover, the root mean square error is used to determine the quality of an ANN architecture (Barile *et al*., 2006; Basheer & Hajmeer, 2000).

The successfulness of the ANN model can be understood in various ways such as the ability of recognizing the percentage of samples in the training set which are

correctly classified, the ability of predicting the percentage of samples in the test set which are correctly classified and the ability of classifying the percentage of samples in both the training and testing sets (Berrueta *et al*., 2007). In food analysis, cross validation is commonly used where the prediction ability of the model is obtained by constructing a model where some parts of the data are used for training or learning and another part of the data are used for testing/validation (da Rocha *et al*., 2015).

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Data analysis of ICP-MS results

4.1.1 Raw milk concentration

Determining the geographical origin of milk is indeed possible by various spectroscopic methods, one of which is ICP-MS. This method provides information on the elemental composition of milk. To ascertain this,70 raw cow milk samples from farms in north and south of Malaysia is collected, digested and analyzed by ICP-MS for their multielemental concentration of Na, Mg, K, Fe, Ca, Zn, Ag, Al, As, Ba, Be, Cd, Co, Cr, Cu, Mn, Mo, Ni, Pb, Sb, Se, Th, Tl, U and V. Between these elements 13 were detected, namely Na, Mg, K, Ca, Mn, Fe, Ni, Cu, Zn, Se, Mo, Ba and Al. Among these elements there are those that are not detected in some of the regions.

In Table 4.1 the concentrations of 13 elements as mean ± SD are reported for nine states in Peninsular Malaysia. It is observed that some of the obtained concentration values are below the detection limit. Concentration of elements below detection limit was replaced by zero in order that statistical analysis could be carried out.

In this table the concentrations of the detected elements of Na, Mg, K, Ca, Mn, Fe, Ni, Cu, Zn, Se, Mo, Ba and Al are respectively in the range of (686.02-1664.29), (137.50-280.14), (194.51-3607.22), (1490.74-2696.30), (ND-0.09), (ND-1.22), (ND-0.02), (ND-0.15), (4.47-8.48), (ND-0.06), (ND-0.08), (ND-0.31) and (ND-5.30).

4.1.2 Radar plot analysis

The concentration of elements in milk samples varies from region to region due to some facts such as the types of food taken by the animal and the environment the animal lives in as well as the lactation period.

Table 4.1: Mean ± SD of raw cow milk samples from different regions of Peninsular Malaysia.

	Kuala Selangor	Johor	Melaka	Pahang	Perak	Terengganu	Pinang	Kedah	Perlis
Na	686.02 ± 51.79	843.95± 66.04	914.26 ± 2.90	796.48 ±24.42	741.78 ± 35.76	1664.29 ± 87.30	1184.91 ± 43.55	914.26 ± 2.90	695.77 ± 12.39
Mg	137.50 ± 5.03	148.81 ± 21.77	213.08 ± 1.37	160.68 ± 4.37	242.89 ± 3.01	280.14 ± 11.32	225.94 ± 5.81	213.08 ± 1.37	219.76 ± 10.37
K	2883.37 ± 73.24	2507.84 ± 81.02	2998.96 ± 9.93	1941.51 ± 42.62	3607.22 ± 15.86	3407.77 ± 91.81	2540.93 ± 56.09	2998.96 ± 9.93	3195.57 ± 66.80
Ca	2183.59 ± 66.64	1966.80 ± 50.88	2594.46 ± 17.52	1490.74 ± 50.53	2696.30 ± 39.04	2600.26 ± 39.48	1963.21 ± 35.87	2594.46 ± 17.52	2417.46 ± 60.43
Mn	0.03 ± 0.003	0.07 ± 0.003	0.03 ± 0.003	0.05 ± 0.001	ND	0.09 ± 0.01	0.07 ± 0.008	0.03 ± 0.003	ND
Fe	0.57 ± 0.06	0.36 ± 0.02	0.79 ± 0.05	0.51 ± 0.007	0.72 ± 0.12	1.03 ± 0.01	1.22 ± 0.04	0.79 ± 0.05	ND
Ni	0.02 ± 0.002	0.02 ± 0.001	ND	0.02 ± 0.002	ND	0.03 ± 0.002	0.01 ± 0.001	ND	ND
Cu	0.14 ± 0.04	0.15 ± 0.04	ND	ND	ND	0.15 ± 0.02	ND	ND	ND
Zn	8.31 ± 0.98	6.07 ± 0.49	6.72 ± 0.12	4.47 ± 0.03	6.84 ± 0.15	8.48 ± 1.16	5.55 ± 0.13	6.72 ± 0.12	5.78 ± 0.21
Se	0.03 ± 0.01	0.03 ± 0.007	ND	0.02 ± 0.003	ND	0.06 ± 0.008	0.02 ± 0.003	0.05 ± 0.004	ND
Mo	0.05 ± 0.01	0.05 ± 0.009	0.06 ± 0.004	0.07 ± 0.004	0.03 ± 0.003	0.08 ± 0.003	0.08 ± 0.006	ND	0.02 ± 0.003
Ba	0.06 ± 0.02	0.08 ± 0.01	ND	0.05 ± 0.002	0.31 ± 0.03	0.16 ± 0.06	0.15 ± 0.01	ND	0.19 ± 0.006
Al	0.15 ± 0.09	0.45 ± 0.25	ND	0.74 ± 0.06	1.83 ± 0.01	0.70 ± 0.1	5.30 ± 0.20	ND	0.28 ± 0.02

ND: Not detected

Radar plot has been used as a simple but powerful statistical tool for differentiating milk samples based on their elemental composition. Since there are only 6 elements that are detected in all of the milk samples from the 24 analyzed, we have used these 6 elements for the radar plots to see if these graphs will lead to some form of clustering. Before analyzing, data normalization was carried out for all the data. Z-transform was applied to all of the concentrations using the equation (4-1);

$$Z\ score = C - mean/SD \quad (4\text{-}1)$$

In this equation C stands for the concentration of each element, while mean is the average concentration of an element in an individual region and SD is the overall standard deviation obtained from the total concentration of an individual element. After applying z-transform to the data, since there are different ranges obtained, in order to be able to better compare the samples, data for all the sampling regions are set to a -2.5 to 2.5 range. Elemental pattern distribution of raw cow milk samples from individual regions in the north and south of Peninsular Malaysia gave relatively different patterns.

The radar plot shown in Figure 4.1 are plotted for samples from different sampling sites in the southern sampling regions of Peninsular Malaysia namely Johor, Melaka, Kuala Selangor and Pahang. It is observed that Kuala Selangor by the highest in terms of Zn, but rather poorly in Mg and Na, while the concentrations are similar for Ca, Mo and K. For samples from Johor, the highest scores are attributed to Mo and Zn while the others are similar. For Pahang the scores are highest in Mo and lowest in Ca and K but more or less similar for the other elements. Melaka scores higher for Na but low in Zn and the other elements are approximately the same.

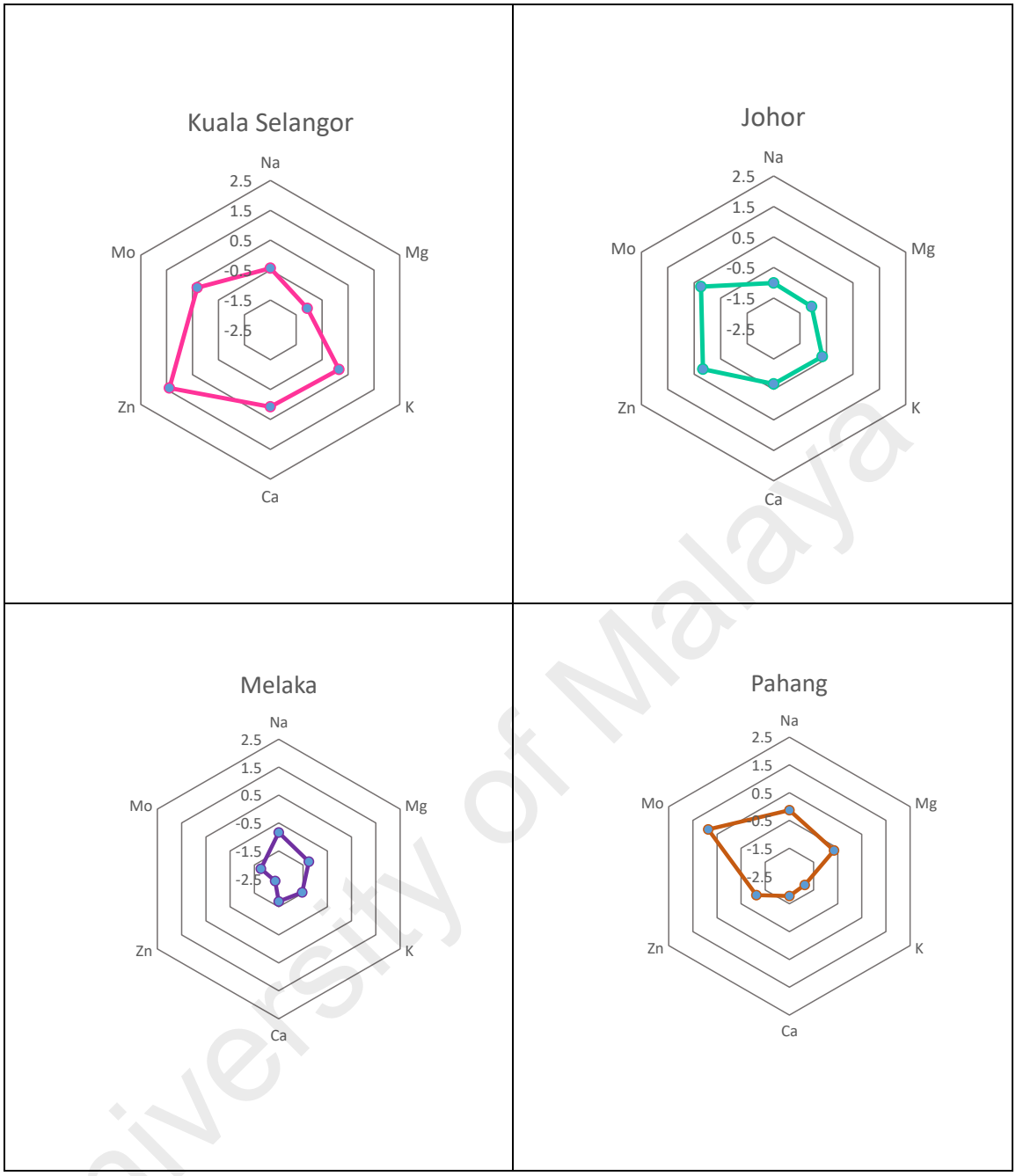

Figure 4.1: Radar plot of raw cow milk samples from southern regions of Peninsular Malaysia.

Figure 4.2 shows the radar plot of raw cow milk samples from various northern regions of Peninsular Malaysia such as Perak, Pulau Pinang, Perlis, Kedah, and Terengganu. From this, it is observed that milk samples from Kedah has the highest score for Ca but fairly similar scores for K, Zn, Mo, Na and Mg. Milk from Perlis is high in K, Ca and Mg but fairly similar in Zn, Mo and Na. Pulau Pinang has milk scoring highest in Mo, Na and Mg but is rather low in Ca, Zn and K. Terengganu scores

highest in Na but similar in Mg, Mo, Zn, Ca and K. In Perak Ca, K, Mg and Zn are high, but low in Mo and Na.

Obviously, the radar plot for each of the sampling regions is different from one another due to differences in their elemental distribution. To have a clearer picture, if possible, of the elemental concentration distribution pattern, the raw cow milk samples from the northern and southern sampling regions were also grouped in two radar plots.

Fig 4.3 (a) shows a radar plot displaying elemental concentrations of milk from combined sampling sites in the northern region. It is clear that there is a meaningful variation in Ca, Mg, K and Zn levels. In the case of Mo, there is no big variation in concentration due to the overlapping of the points on the plot. In the case of Na, there is almost no variation except that Kedah can be easily differentiated from other northern regions by virtue of its high Na concentration.

It is obvious in Fig. 4.3 (b) that there are variations between the concentrations of the elements in Johor, Melaka, Kuala Selangor and Pahang as well. Variations are apparent for K, Ca and Mo, but less apparent for Na, Zn and Mg (overlapping in some cases).

It is rather difficult to differentiate the two regions just based on the radar plots. The shape of the plots do not show obvious patterns of discrimination between the north and the south. A note can be made on the concentration of Ca based on the plot that could serve as a reason for discrimination, but it would not be a strong enough reason. Consequently, multivariate pattern recognition methods, PCA and DA were applied.

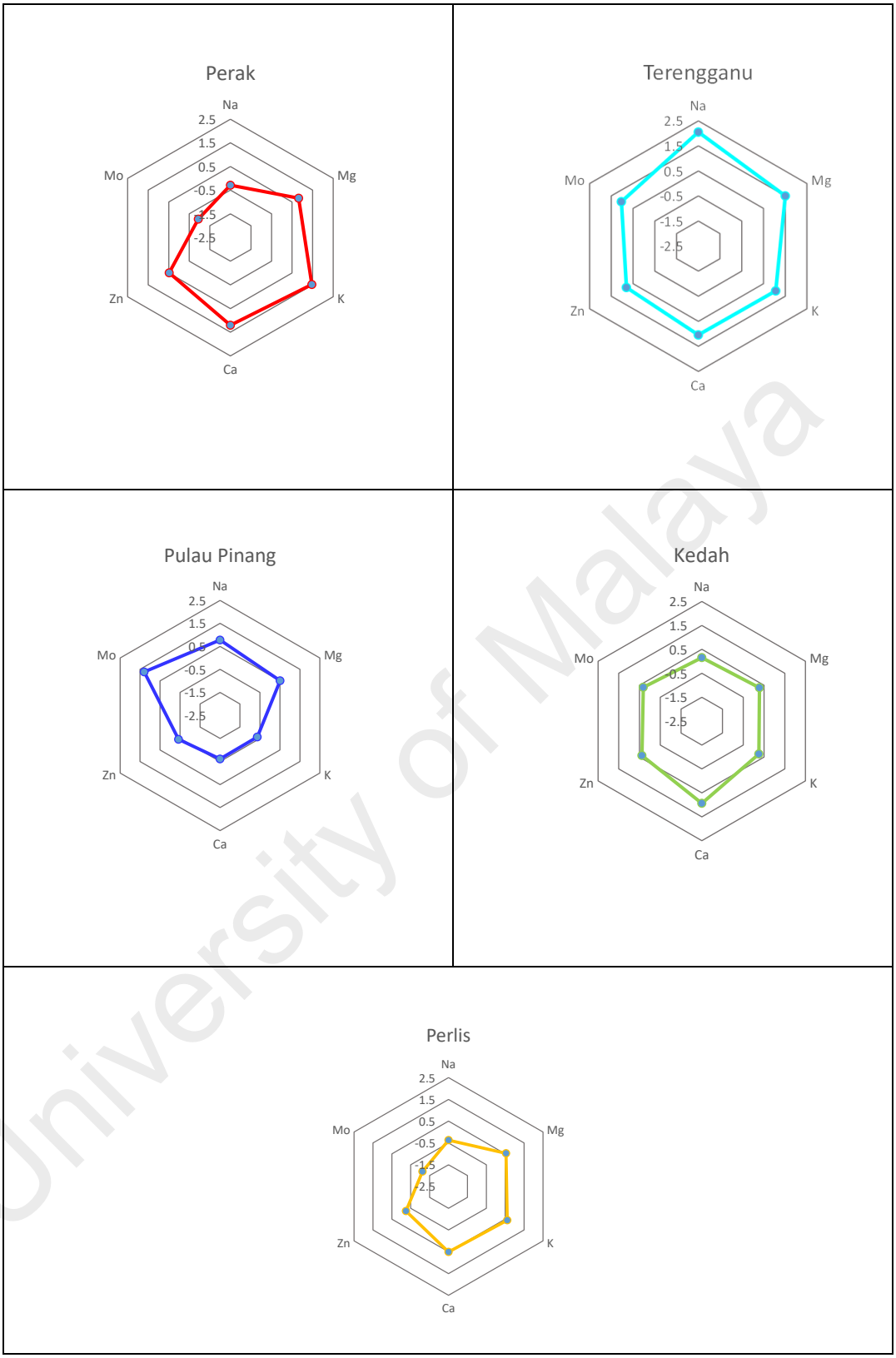

Figure 4.2 Radar plot of raw cow milk samples from northern regions of Peninsular Malaysia.

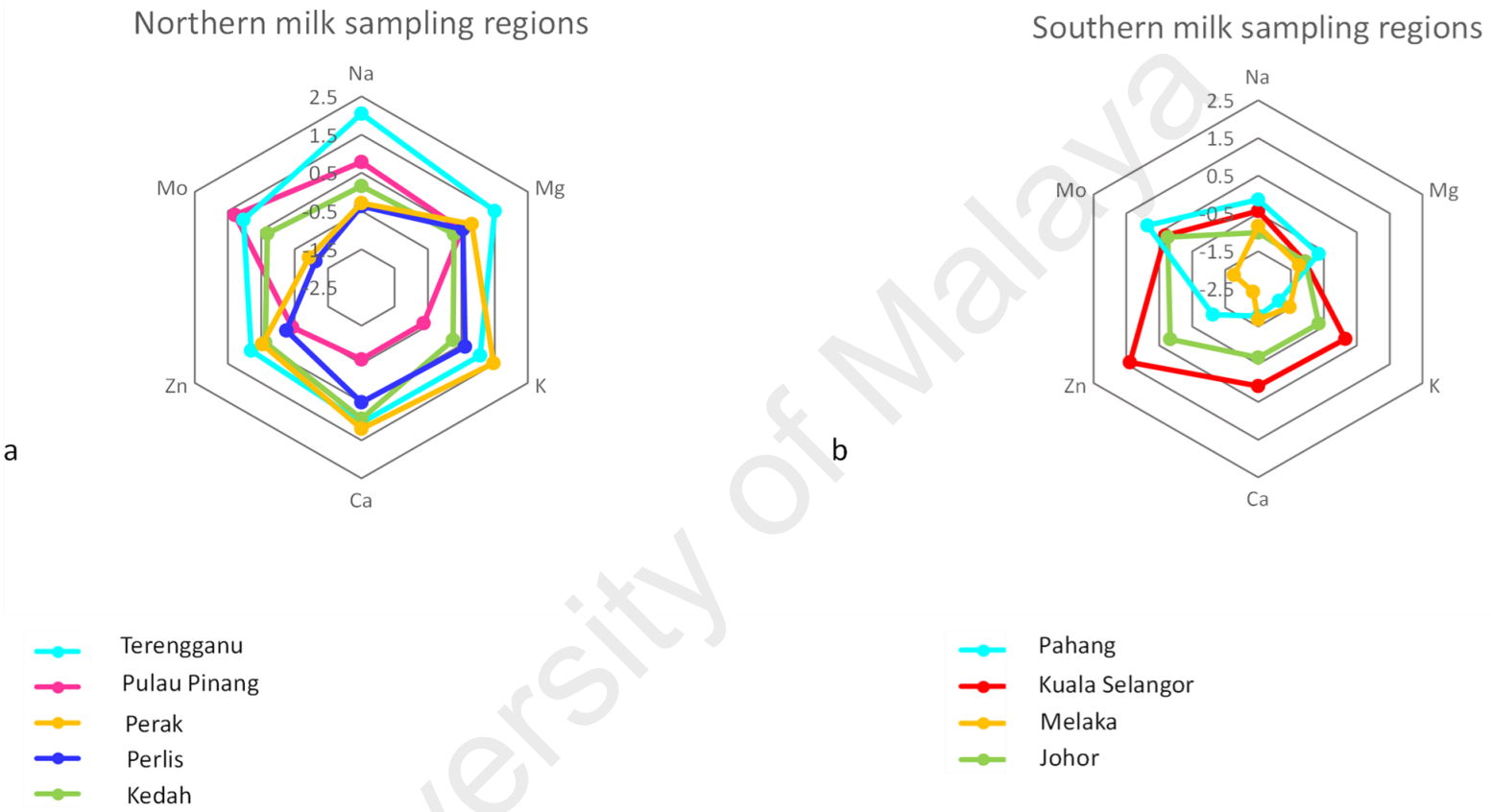

Figure 4.3: Radar plots of raw cow milk samples a) from northern region b) southern region.

4.1.3 Principal component analysis (PCA)

70 raw cow milk samples from various regions in Peninsular Malaysia were analyzed by PCA with a matrix of thirteen analytical parameters in order to cluster the samples according to their elemental distribution. It is found that 2 principal components (PCs) with eigenvalues larger than one explained a total variance of 67.5 %.

In Figure 4.4 there is a clear separation between the northern and southern clusters and the loading plot shows the variables causing the separation. It is observed that PC1 is responsible for the separation of samples into two groups of north and south. Samples from the south have negative PC1 scores while those from the north have positive PC1 scores. In the case of PC2, northern and southern samples share both negative and positive PC1. Ba, K, Ca, Mg, Zn, Fe, Na, Mo, Ni, Mn, Se, Cu and Al are the elements which dominate samples from the northern region, discriminating them from southern samples. PC1 explains 39.3% of the total variance with PC2 and PC3 each contribute 28.2% and 16.31% of total variance respectively. For all the other PC's the contribution is below 5%. The loadings of the PC's are listed in Table 4.2 where it is noted that PC1 is associated with Mg, K, Ca, Na, Se, Zn, Ba and Mo, PC2 is loaded with Mn, Mo, Ni, Se, Ba, Na, Fe, K, Ca and Al and PC3 by Cu, Ni, Zn, Ba and Al.

From the loading plot of PC2 vs PC1 in Figure 4.4 it is observed that Mn, Ni, K, Ca, Mg, Cu, Zn, Se, Mo, Ba, Na, Fe and Al loaded PC1positively while K, Ca, Se and Ba negatively loaded PC2. Therefore, it can be concluded that based on PCA, a separation based on northern and southern milk samples can be observed.

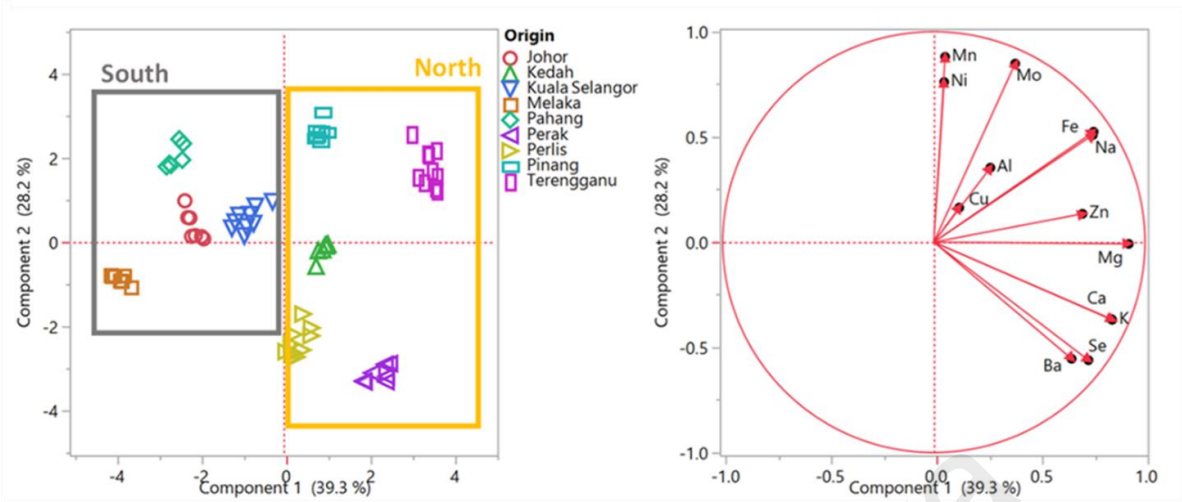

Figure 4.4: Scores and loading plots of milk samples.

4.1.4 Discriminant analysis (DA)

Discriminant analysis as one of the multivariate chemometric methods has been used to cluster the raw cow milk samples collected in this research. The calculations were carried out using backward selection based on p values by eliminating variables that have p values larger than 0.05. Among the elements, Ca, Mg, Na, Zn, Mo and K were selected for canonical analysis since they were the only elements that had been detected in all of the sampling regions. Canonical1 explains 62.77% and Canonical2 explains 28.35% of the matrix variation. The closeness of some of the milk samples in the southern region suggests that milk samples from these regions do not depend on important variables noted in PC1 and PC2. Milk samples distribution according to their geographical origin is shown in Figure 4.5.

From the plot, separation of northern and southern milk samples is noted by virtue of Canonical1. The unique separation of milk samples in Peninsular Malaysia shows that elemental fingerprint can be considered for clustering milk samples based on their geographical origin.

Table 4.2: Loading matrix of the PC's.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13
Na	0.75658	0.52419	-0.07772	-0.32104	0.16228	0.10692	0.02461	-0.03343	0.00446	-0.00570	-0.03453	0.00430	-0.05430
Mg	0.92292	-0.00842	-0.22096	-0.23809	0.07251	-0.04014	-0.04103	-0.00479	0.06807	0.15960	-0.04093	-0.03934	0.02713
K	0.84436	-0.36934	0.23586	0.07251	-0.14465	0.19518	-0.05639	-0.04659	-0.10888	-0.07787	-0.08407	-0.00084	0.02004
Ca	0.84069	-0.36541	0.21429	0.02345	-0.00451	-0.24087	0.17501	0.03466	-0.14614	0.03809	0.00924	0.02059	-0.00726
Mn	0.05361	0.88022	-0.05442	0.08305	0.40605	-0.15553	0.04652	-0.08978	0.01991	-0.09795	-0.05126	0.00759	0.02276
Fe	0.75233	0.50851	-0.32517	0.10880	-0.12078	0.05304	0.10662	0.11240	0.00112	-0.08133	0.06259	-0.07664	0.00580
Ni	0.04887	0.76057	0.54204	-0.27967	0.01486	0.16695	0.09710	0.03193	-0.02356	0.03983	0.05694	0.04984	0.02648
Cu	0.12040	0.16381	0.76125	0.52918	0.28625	0.07563	-0.02154	0.06388	0.01285	0.07618	-0.00624	-0.02775	-0.01236
Zn	0.70398	0.13455	0.50970	0.13066	-0.39748	-0.08645	0.03718	-0.12767	0.15884	-0.02666	0.00282	0.01802	-0.00521
Se	0.73216	-0.56089	0.11835	-0.10878	0.21202	-0.04160	-0.11617	0.21695	0.08890	-0.07975	0.01771	0.04082	0.00376
Mo	0.38509	0.84791	-0.03270	0.07083	-0.15140	-0.10867	-0.28519	-0.00770	-0.09268	0.01828	0.03857	0.01069	-0.00362
Ba	0.65247	-0.55447	-0.32539	0.14907	0.27894	0.08797	-0.02212	-0.21003	-0.00192	0.00262	0.09134	0.01547	0.00325
Al	0.26792	0.35376	-0.77540	0.40487	-0.10819	0.07764	0.06721	0.08197	0.02361	0.05398	-0.03227	0.06655	0.00008

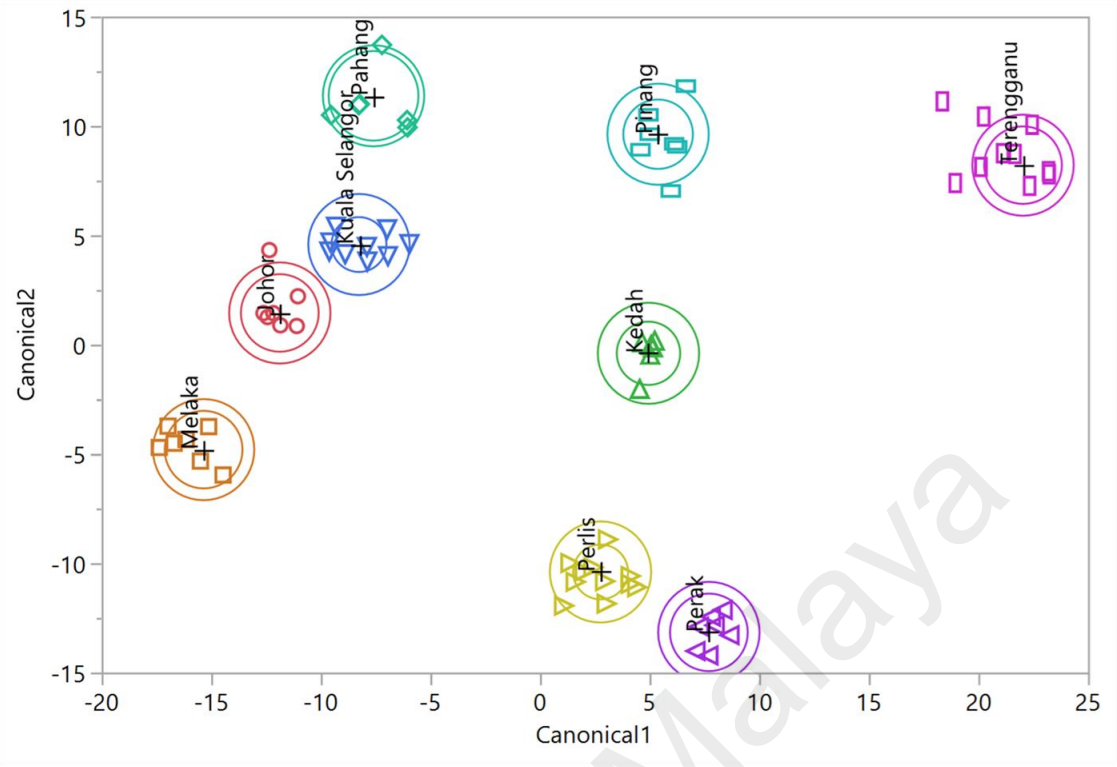

Figure 4.5: Canonical plot of milk samples from different regions.

In order to ascertain the reliability of this model, cross validation has been used to compute the grouping probability for the raw cow milk samples and the information is presented in Table 4.3. It is observed that from the information in the Table, 100% correct classification of all milk samples from Peninsular Malaysia is obtained. Moreover, off diagonal counts are zero, indicating correct classification of samples. Furthermore, the -2Loglikelihood value is small indicating that the model is good.

Discriminant analysis was also applied to the same data set to determine the loadings responsible for the separation of samples in the north and south as shown in Fig 4.6. The two groups of milk samples from the north and south are separated by Canonical1.

Table 4.3: Score summaries.

Source	Count	Number Misclassified	Percent Misclassified	Entropy RSquare	-2LogLikelihood
Training	46	0	0.00000	1.00000	1.34e-7
Validation	9	0	0.00000	1.00000	
Test	15	0	0.00000	1.00000	

Training

Actual	Predicted								
Origin	**Johor**	**Kedah**	**Kuala Selangor**	**Melaka**	**Pahang**	**Perak**	**Perlis**	**Pinang**	**Terengganu**
Johor	4	0	0	0	0	0	0	0	0
Kedah	0	6	0	0	0	0	0	0	0
Kuala Selangor	0	0	8	0	0	0	0	0	0
Melaka	0	0	0	4	0	0	0	0	0
Pahang	0	0	0	0	3	0	0	0	0
Perak	0	0	0	0	0	4	0	0	0
Perlis	0	0	0	0	0	0	8	0	0
Pinang	0	0	0	0	0	0	0	5	0
Terengganu	0	0	0	0	0	0	0	0	4

Validation

Actual	Predicted								
Origin	**Johor**	**Kedah**	**Kuala Selangor**	**Melaka**	**Pahang**	**Perak**	**Perlis**	**Pinang**	**Terengganu**
Johor	0	0	0	0	0	0	0	0	0
Kedah	0	0	0	0	0	0	0	0	0
Kuala Selangor	0	0	0	0	0	0	0	0	0
Melaka	0	0	0	1	0	0	0	0	0
Pahang	0	0	0	0	2	0	0	0	0
Perak	0	0	0	0	0	1	0	0	0
Perlis	0	0	0	0	0	0	2	0	0
Pinang	0	0	0	0	0	0	0	1	0
Terengganu	0	0	0	0	0	0	0	0	2

Testing

Actual	Predicted								
Origin	**Johor**	**Kedah**	**Kuala Selangor**	**Melaka**	**Pahang**	**Perak**	**Perlis**	**Pinang**	**Terengganu**
Johor	3	0	0	0	0	0	0	0	0
Kedah	0	0	0	0	0	0	0	0	0
Kuala Selangor	0	0	1	0	0	0	0	0	0
Melaka	0	0	0	2	0	0	0	0	0
Pahang	0	0	0	0	1	0	0	0	0
Perak	0	0	0	0	0	2	0	0	0
Perlis	0	0	0	0	0	0	1	0	0
Pinang	0	0	0	0	0	0	0	1	0
Terengganu	0	0	0	0	0	0	0	0	4

It is noted that covariates are loaded only on Canonical1. Six variables are used to cluster the samples into two groups of northern and southern sampling regions and this separation is based on Canonical1 being positive for the northern sampling regions and negative for southern sampling regions. Ca, K, Mo and Mg are positively associated with Canonical1. The order of highest degree of association for milk samples from the north is Ca>Mo>Mg>K, which are located at the right hand side of the plot. Zn and Na are negatively associated with Canonical1. Zn has the highest degree of association for samples from the southern sampling regions and is located at the left hand side of the graph. Na on the other hand has the lowest degree of association. The summary report defines that there is 100% classification with no misclassification observed. From this, 65% of the data was automatically selected for training, 20% for testing and 15% for validation. Furthermore, the entropy source for all three sources is 0.999 and, -2Loglikelihood value being very small confirms the good fit of the model obtained.

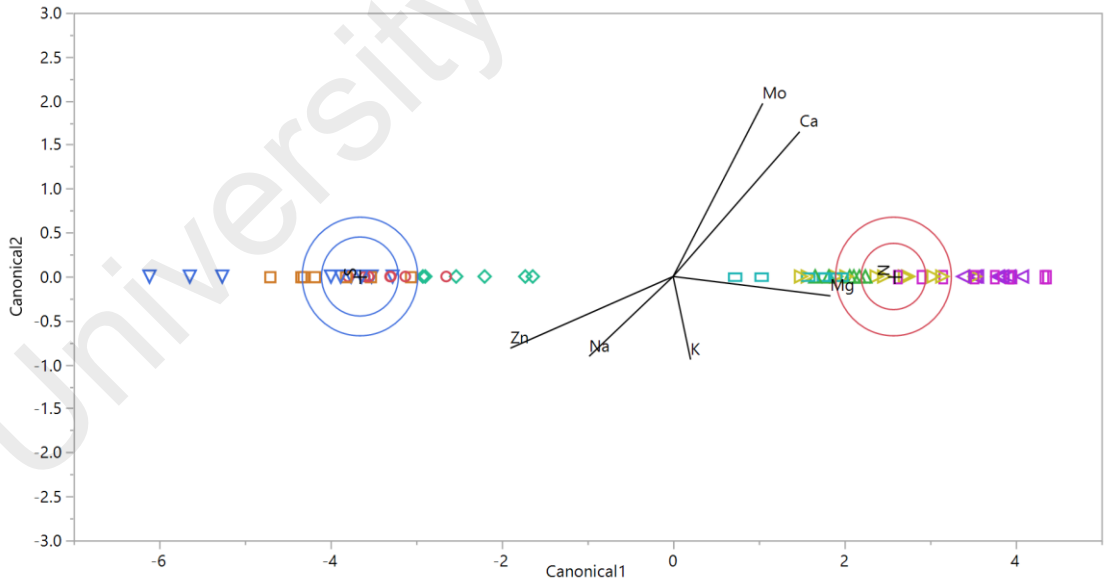

Figure 4.6: Canonical plot of milk samples based on northern and southern sampling regions.

4.1.5 Mean concentration of essential and trace elements in raw cow milk of Malaysia

52 raw cow milk samples which were digested by microwave digestion system were analyzed for 24 mineral and trace elements by ICP-MS. In Table 4.4 milk samples from northern and southern sampling regions are compared based on their elemental composition and it is observed that there are no significant difference in the concentrations of Na, Mg, K, Ca, Mn, Zn, Mo and Cu between northern and southern Malaysia at the level of 95% confidence interval. However, concentration of Fe in milk samples from southern Malaysia is significantly higher ($p<0.05$) compared to the north of Malaysia. Levels of Ba and Se are found to be significantly higher in the north compared to the south and concentrations of Mn and Mo is the same in the north and south of Malaysia. This seem to confirm that the elemental composition of raw cow milk varies depending on geographical origin. The concentration of elements in milk of Malaysia and Iran has been compared and it is observed that the mean concentration of Zn, Fe and Cu are significantly higher ($p<0.05$) in raw cow milks of Malaysia and Mo is significantly higher in Iran milk samples. Additionally, the concentration of Na, Mg, K, Ca, Ba, Se and Mn is more or less the same ($p>0.05$) in raw cow milk samples of Malaysia and Iran.

4.1.6 Box plot of factory milk samples based on the country of origin

Elemental composition of factory milk samples have been investigated in this work using the box and whisker plot for each individual element based on the milk's country of origin namely U.S.A, Canada, Belgium, Turkey, Malaysia, New Zealand, Iran, Australia and Azerbaijan.

Table 4.4: Comparison of the average concentration range of various essential and trace elements of raw cow milk samples from northern and southern parts of Peninsular Malaysia and Iran ($mgKg^{-1}$).

Elements	Mean ± SD for farms in Malaysia		Mean ± SD for farms in Iran	
	North N=63	South N=93	North N=12	South N=9
Na	510.6±16.7	479.6±14.9	474.8±2.0	386.3±1.4
Mg	115.7±16.0	116.5±26.2	106.4 ±0.8	96.970±1.1
K	1550.8±187.7	1596.0±371.9	1503.9±10.9	1473.5±10.9
Ca	1237.9±152.3	1219.7±281.1	1082.0±14.7	1037.7±7.7
Mn	0.080±0.003	0.080±0.003	0.079±0.004	0.079±0.004
Zn	5.93±0.69	6.48±1.30	3.53±0.02	3.09±0.01
Se	0.098±0.007	0.080±0.007	0.080±0.003	0.098±0.002
Mo	0.050±0.010	0.050±0.010	0.098±0.001	0.098±0.001
Fe	2.39±0.09	4.58±0.06	1.44±0.05	1.33±0.03
Cu	1.29±0.03	1.30±0.04	0.83±0.06	0.93±0.09
Ba	0.079±0.003	0.053±0.003	0.079±0.007	0.056±0.007

SD: standard deviation; ND: not detected

Figure 4.7 illustrates the box plot of Na in different countries. It is observed that milk samples from Belgium and Azerbaijan have the lowest concentration of Na and have approximately the same median. The upper quartile of samples from Azerbaijan shows that the Na content is a bit higher than Belgium. This is also apparent for the lower quartile. Samples from Belgium are distributed normally but in the case of samples from Azerbaijan the distribution is skewed to the right (where the mean is greater than median). Samples from New Zealand have the highest concentration of Na compared to samples from other countries and is normally distributed. The box plots of milk samples from Canada, Iran, Malaysia, Turkey and U.S.A overlap with each other. The mean of these samples are similar but the spread of data is different. The concentration of samples from Canada and U.S.A are normally distributed but for Iran, Turkey and Malaysia the distribution of data is not normal and in all three, the data is skewed to the left. The box and whisker plots for all the other detected elements can be found in Appendix C.

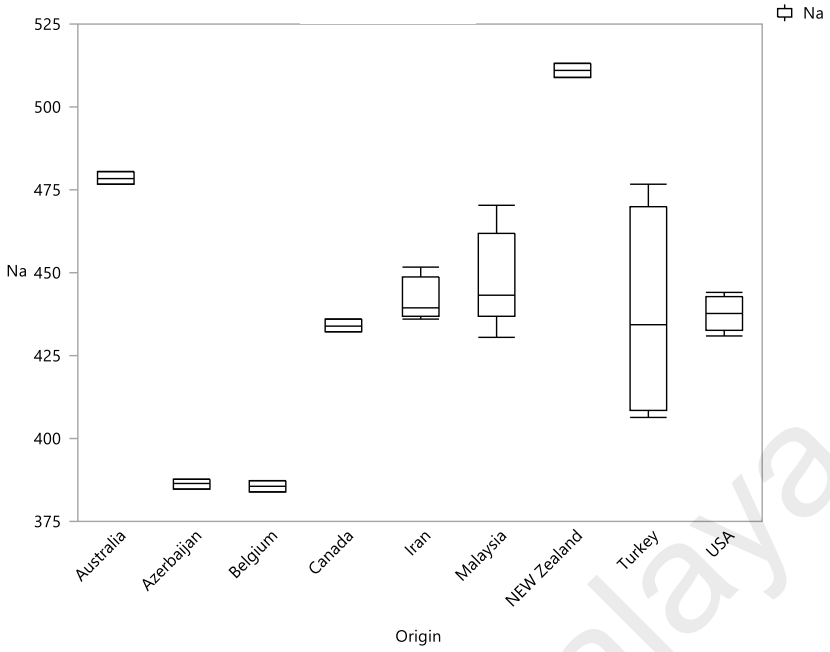

Figure 4.7: Box plot of Na in factory milk samples of some selected countries.

4.1.7 Mean concentrations of essential and trace elements in factory milks of Malaysia and other selected regions

18 factory cow milk samples consisting of Malaysia (5), New Zealand (1), Australia (1), Turkey (3), Iran (3), Azerbaijan (1), Belgium (1), Canada (1) and U.S.A (2) were analyzed by ICP-MS and were compared with one another based on their elemental composition as shown in Table 4.5. Between 24 elements, only eleven elements of Na, Mg, K, Ca, Mn, Zn, Mo, Se, Ba, Fe and Cu are detected in all the factory milk samples. The concentration of these elements varied respectively in the following range from 385.6 to 511.0, 93.1 to 123.7, 1473.5 to 1690.5, 1028.4 to 1207.8, ND to 0.084, 3.09 to 5.05, 0.027 to 0.099, 0.060 to 0.120, 0.046 to 0.110, 1.30 to 5.30 and 0.73 to 1.21 mg kg^{-1}.

Comparing Malaysian and Australasian milk samples, the concentrations of concentrations of Cu, Ca, Mn, Zn, Fe and Ba are similar and the concentration of Na, Mg, K, Mo, and Se are significantly different ($p<0.05$). In Europe and Malaysia, concentrations of Cu, Zn and Ca are similar while concentrations of Na, Mg, K, Mn, Fe,

Ba, Se and Mo in Malaysian milk samples are significantly different ($p<0.05$) compared to Europe. In comparing Malaysian milk samples with that of the Middle East, it is observed that they have similar Cu concentration but are significantly different ($p<0.05$) in the concentrations of Na, Mg, K, Ca, Mn, Zn, Mo, Fe, Ba and Se. In turn, the Malaysian milk samples are similar in concentrations of Zn, Se, Ca, Ba and Cu compared with America but significantly different ($p<0.05$) in the concentrations of Na, Mg, K, Mo, Fe and Mn. The variation in the concentration of the elements is also shown by cell plots in Figure 4.8 and 4.9. The information from these plots will be used when applying the two way analysis which is a combination of HCA and the heat map. Figure 4.8 presents the distribution of elements based on individual sampling regions of Malaysia, Australia, New Zealand, Turkey, Azerbaijan, Iran, Canada, U.S.A and Belgium. There were eleven elements that were detected in all of the sampling regions and are presented in the figure where concentration of each element is shown in the legend. For each region, different columns of elements manifest variation in colors. From the color variation, it can be seen if that particular element has the highest or lowest or fairly high or fairly low concentration. Moreover, the more intense is the red color of the cell, whereas the more intense is the blue color the region has less concentration of that particular element. For example, the Na concentration in New Zealand is more intense in red compared to other regions meaning that the concentration of this element is the highest in this particular region. In Azerbaijan and Belgium, on the other hand, Na is lowest. In Figure 4.9 the same concept is followed but this time, the concentration of elements are compared in different continents of Asia, Australasia, Europe, America and Middle East. Na is observed to be lowest in Europe. Although it is observed that Na is low in Azerbaijan, which is one of the regions located in Middle East, this is not necessarily true for all other areas in that region. This is illustrated in Australasia as well where Na is highest in New Zealand but is not necessarily true for

all other areas in Australasia. The distribution of all the other elements could be explained by the same fact using these cell plots.

Table 4.5: Average elemental concentrations of various factory milks in mg kg^{-1} (mean ± standard deviation) in dry weight.

Element	New Zealand N=3	Belgium N=3	Iran N=9	Australia N=3	Malaysia N=15	Azerbaijan N=3	Turkey N=9	USA N=6	Canada N=3
Na	511.0±2.2	385.6±1.8	441.8±6.1	478.6±2.0	450.2±13.9	386.3±1.4	430.6±33.2	437.7±5.3	434.1±1.8
Mg	99.2±0.8	105.8±0.8	105.6±3.1	99.4±0.9	123.7±7.6	97.0±1.1	108.6±2.1	96.8±0.9	93.1±1.5
K	1573.4±11.4	1493.2±11.4	1598.3±15.0	1598.7±11.8	1690.5±44.5	1473.5±11.0	1501.6±12.5	1583.9±12.1	1624.6±11.8
Ca	1178.0±16.0	1162.3±19.0	1048.0±32.4	1133.3±15.5	1160.4±45.7	1028.4±8.1	1105.2±22.2	1207.8±15.1	1103.3±18.1
Mn	0.080±0.005	ND	0.079±0.008	0.081±0.005	0.080±0.009	0.080±0.005	0.084±0.005	ND	ND
Zn	4.41±0.02	4.76±0.02	3.25±0.17	4.10±0.02	5.05±3.55	3.09±0.02	3.68±0.17	4.13±0.16	3.51±0.02
Mo	0.090±0.002	0.091±0.002	0.083±0.001	0.090±0.002	0.027±0.004	0.099±0.002	0.090±0.006	0.090±0.005	0.090±0.002
Se	0.060±0.005	0.066±0.005	0.074±0.010	0.060±0.005	0.071±0.005	0.090±0.007	0.100±0.010	0.065±0.005	0.120±0.001
Ba	0.110±0.005	0.046±0.005	0.0510±0.010	0.071±0.004	0.090±0.011	0.057±0.004	0.072±0.005	0.067±0.010	0.095±0.005
Fe	5.30±0.06	2.76±0.03	1.59±0.11	3.75±0.06	3.65±0.86	1.33±0.03	1.30±0.07	2.60±0.07	3.24±0.03
Cu	0.85±0.03	1.21±0.03	0.77±0.06	0.73±0.03	0.87±0.01	0.93±0.09	0.87±0.11	0.81±0.06	0.93±0.03

ND= not detected, N= number of samples triplicated

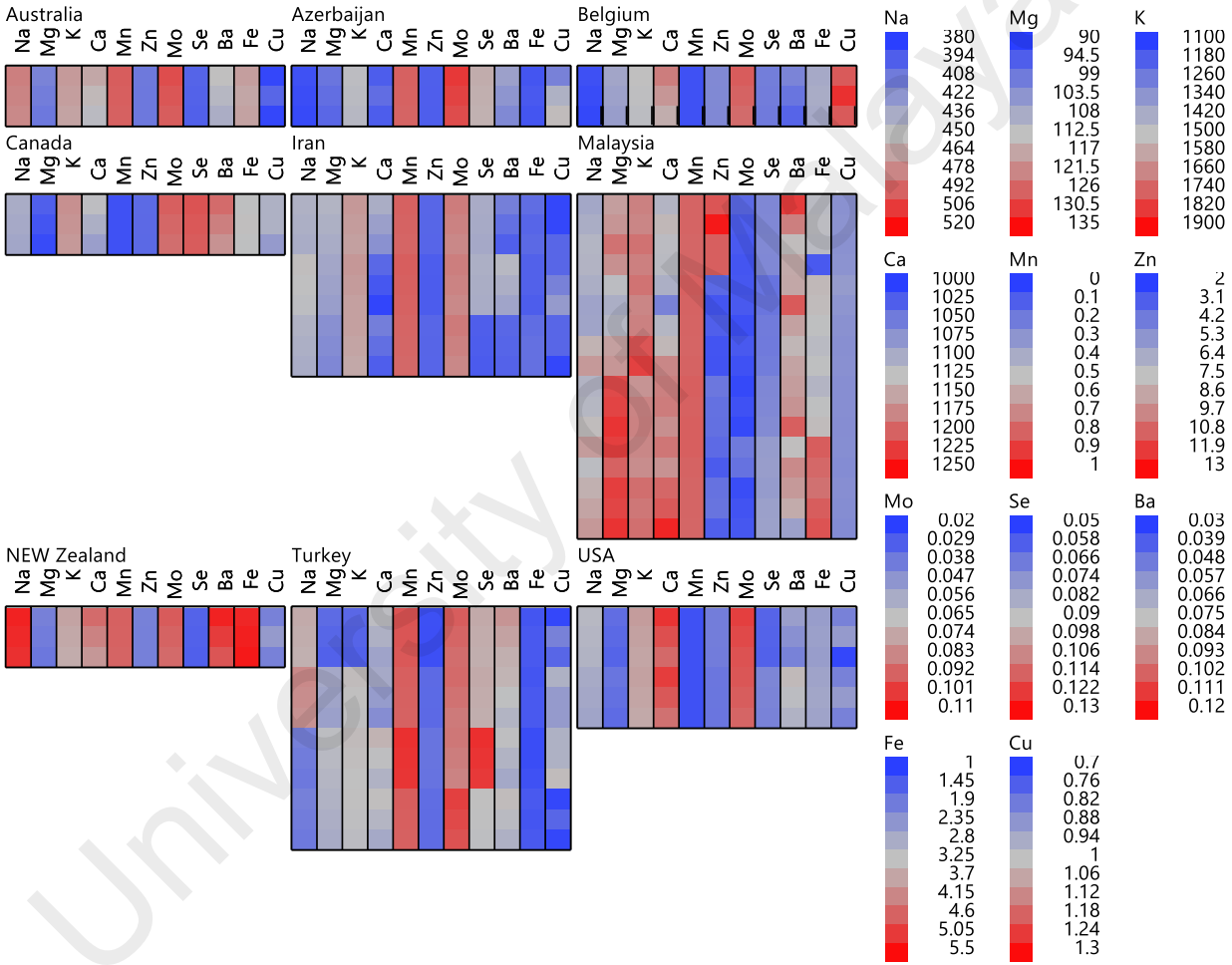

Figure 4.8: Cell plot of factory milk samples based on some selected countries.

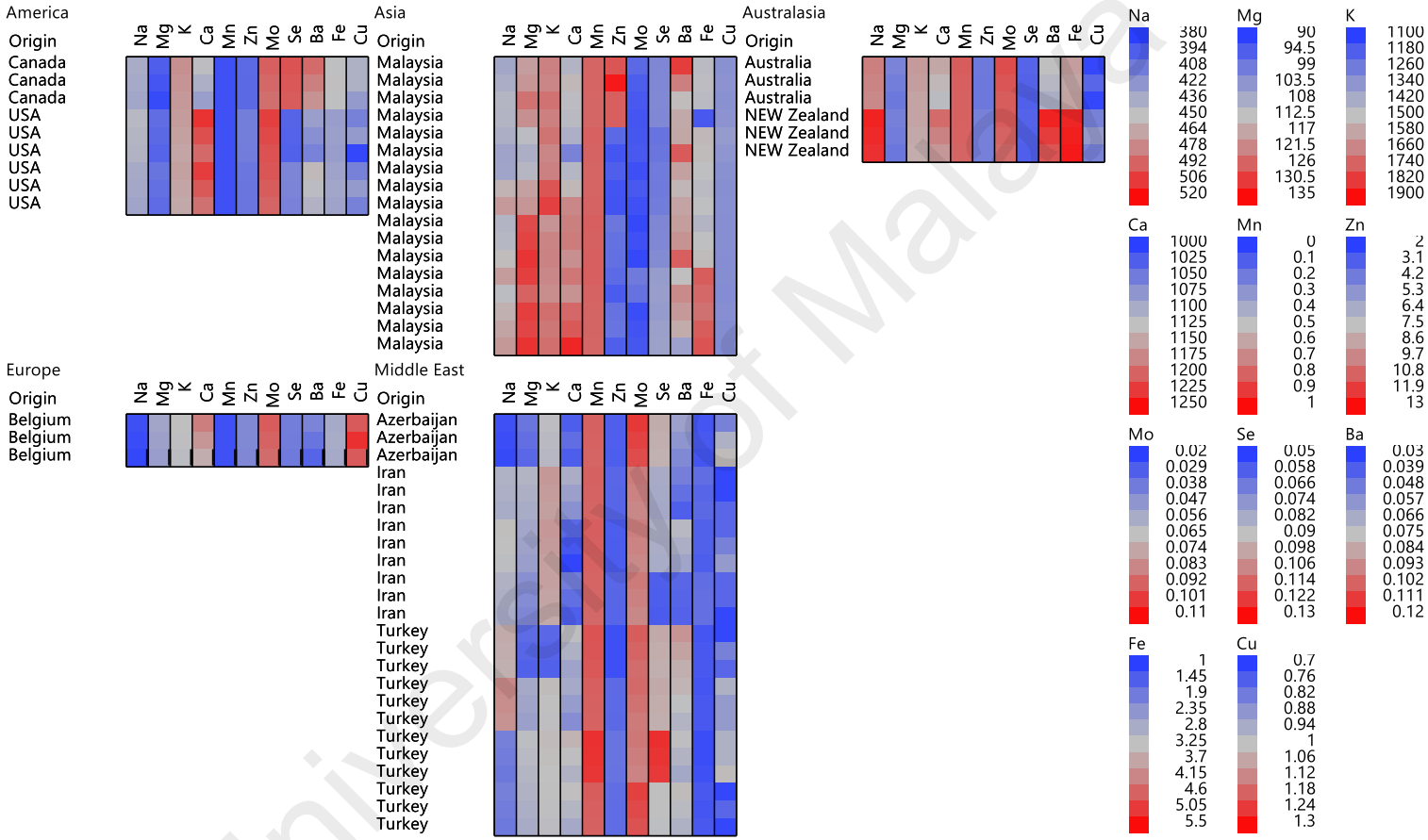

Figure 4.9: Cell plot of factory milk samples based on the continents.

4.1.8 Box plot of factory milk samples of various continents

The elemental composition of eleven detected elements have been compared based on the continent of origin using the box and whisker plots. Figure 4.10 represents the multi box plots for Na. Europe is seen to have the lowest concentration of Na whereas Australasia has the highest. The distribution of data in these two regions is by to be normal. It is observed that the boxplots of milk from the Middle East, Asia and America overlap each other. From the box plots of America and Asia it is understood that the median is close to lower range of the data and that these two boxplots are right skewed. In the Middle East the data is more spread out and the median is close to the higher range of the data and is left skewed. The box plots for rest of the elements are presented in appendix D.

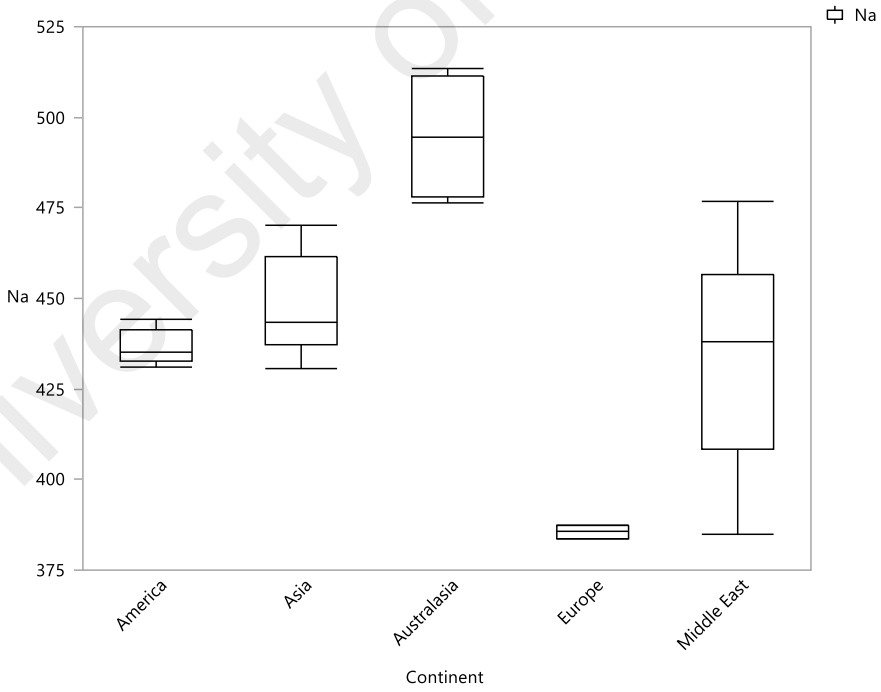

Figure 4.10: Box plot of Na in factory milk samples based of continents.

4.1.9 PCA of factory cow milks

To show the pattern distribution of metals between factory cow milk samples of Malaysia and other selected regions of the world, PCA was applied to a matrix of eleven analytical parameters for 54 factory milk samples. There are four PC's with eigenvalues over one that are extracted which explained 53.1 % of the total variance.

4.1.10 Biplots of factory cow milk samples

In Fig. 4.11, the biplots for all factory milk samples studied in this work is presented. Milk samples are clustered into two groups of clusters, 1 and 2. PC1 is responsible for the separation of samples in cluster 1 from cluster 2. Milk samples that have positive PC1 and are located in the right hand side are loaded by K, Ba, Na, Fe, Ca, Mg, Mn and Zn whereas samples having negative PC1 scores located at left hand side of the graph are loaded by Mo, Cu and Se. Overall, these could be considered as discriminating factors for separation of the two clusters from one another.

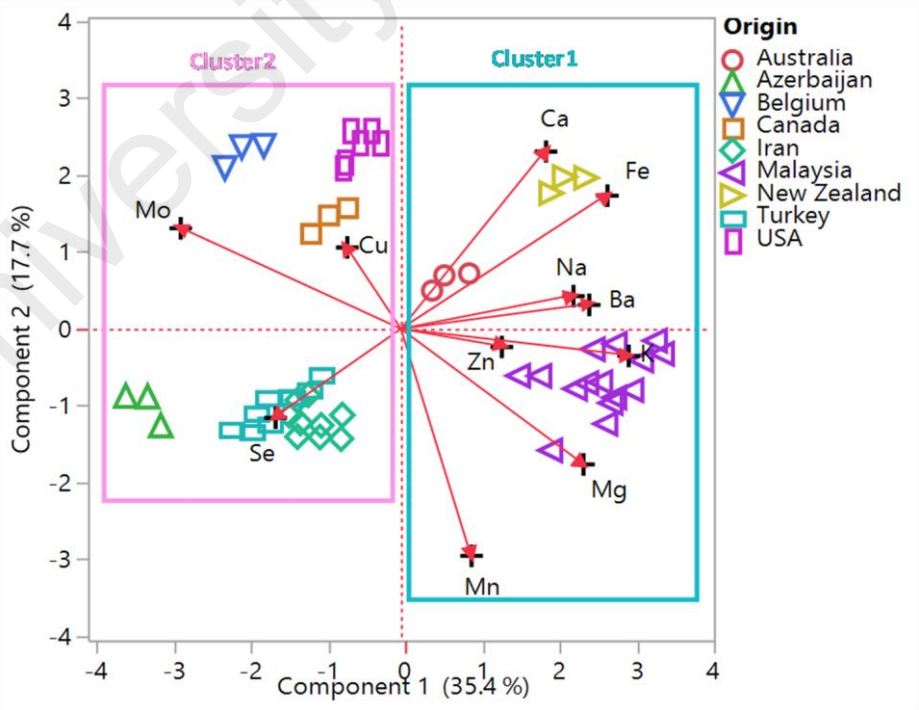

Figure 4.11: Biplots of all factory milks studied in this work.

In order to investigate the discriminating factors in the separation of Malaysian milk samples from other selected regions of the world biplots in Fig. 4.12 (a, b, c and d) are considered.

Malaysian milk samples are separated from Australasia samples by PC1 scores as presented in Fig. 4.12 (a). Milk samples from Malaysia are loaded by Cu, Zn, Se, K and Mg and have positive PC1 scores. Australian milk samples are loaded by Fe, Mn, Mo, Na, Ca and Ba and have negative PC1 scores. Milk samples of Australasia consisting of Australia and New Zealand can be separated by virtue of the sign of PC2 scores. Samples from New Zealand are located at the upper left hand side of the quadrant and have positive PC2 scores while the ones from Australia are located at the lower left hand side of the quadrant and have negative PC1 scores. Samples from New Zealand are generally loaded by Ba, Ca, Na and Fe while samples from Australia are loaded by Mn and Mo.

In Fig. 4.12 (b), biplot of milk samples from Malaysia and Europe is shown. In this plot, PC1 is responsible for the separation of the two observed clusters being negative for samples from Europe and positive for samples from Malaysia. European milk samples are loaded by Mo and Cu whereas loadings on Zn, Mn, K, Na, Ba, Mg, Se, Ca and Fe dominate the Malaysian samples.

Fig. 4.12 (c) represents the separation of Malaysian and Middle Eastern milk samples which is due by PC1. Milk samples from Malaysia are located at the right hand side of the plot and have positive PC1 scores and are mainly loaded by Fe, K, Mg, Ba, Na, Zn, Ca and Cu. Middle Eastern samples have negative PC1 scores and are loaded by Mn, Se and Mo. In the case of samples from Middle East, three clusters of Iran, Turkey and Azerbaijan are observed although some of the samples from Iran and Turkey have overlap which could be due to the closeness of lands having the same environmental

condition. Biplot illustrated in Fig.4.12 (d) presents the separation of Malaysian milk samples from the American region which is possible by PC1 scores. Milk samples from Malaysia are loaded by Zn, Fe, Mn, Mg, Na, Ba, Cu and K. Milk from the America region which consist of U.S.A and Canada are separated based on PC2. Samples from U.S.A located in the lower left of the quadrant have negative PC1 and PC2 scores whereas samples of Canada have negative PC1 and positive PC2 scores.

Generally, it seems that there are differences in the elemental composition of milk samples in the tropical and Australasian regions compared to other selected regions of the world. Malaysian milk samples are clustered in to a separate group located at the lower right hand of the plot and are separated by PC2 scores where Zn, K and Mg represent the separation. The discriminating factors for the clustering of tropical and Australasian regions from other selected regions of the world are the concentration of Mn, Mg, Na, Ca, Ba, Fe, K and Zn. Overall, cow milks from tropical and Australian regions are mainly loaded by Mn, Zn, Ba, Fe, K, Mg, Ca and Na. Cu, Se and Mo seem to dominate the loadings for other regions studied in this research. Mg, K, Ca, Fe and Zn are the discriminating factor responsible for the separation of Malaysian and Australasian milk samples from other parts of the world.

4.1.11 3D PCA of factory milk samples

To confirm the sufficiency of only considering the first two principal components, a 3D plot of varimax rotated PC1 vs. PC2 vs. PC3 is also shown in Figure 4.13. The clustering of samples in the 3D plot bears a similar cluster structure to the 2D plot of PC1 vs. PC2, hence, discussion of the clustering based on the 2D plot is deemed sufficient.

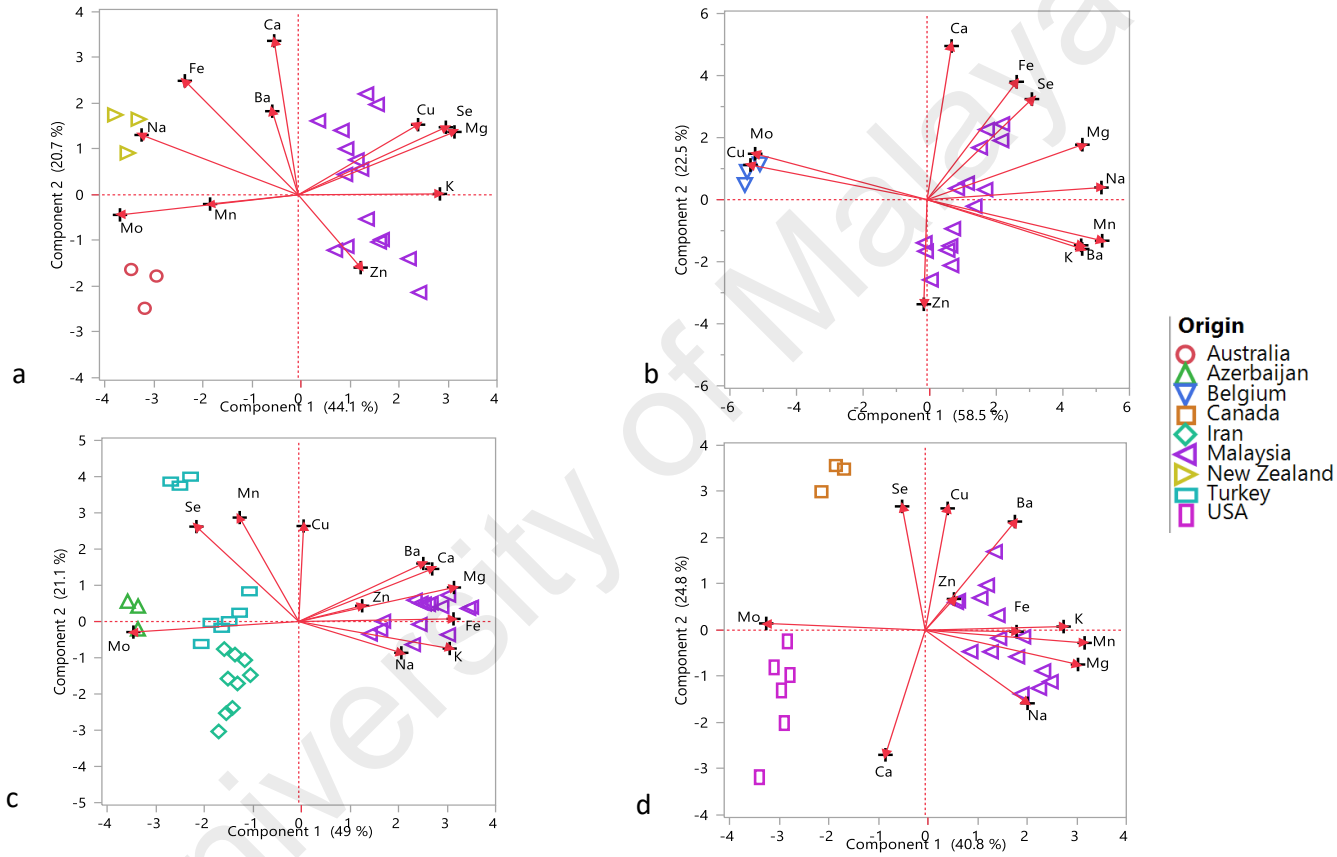

Figure 4.12: (a) Factory milks from Malaysia and Australasia. (b) Factory milks from Malaysia and Europe. (c) Factory milks from Malaysia and the Middle East. (d) Factory milks from Malaysia and America.

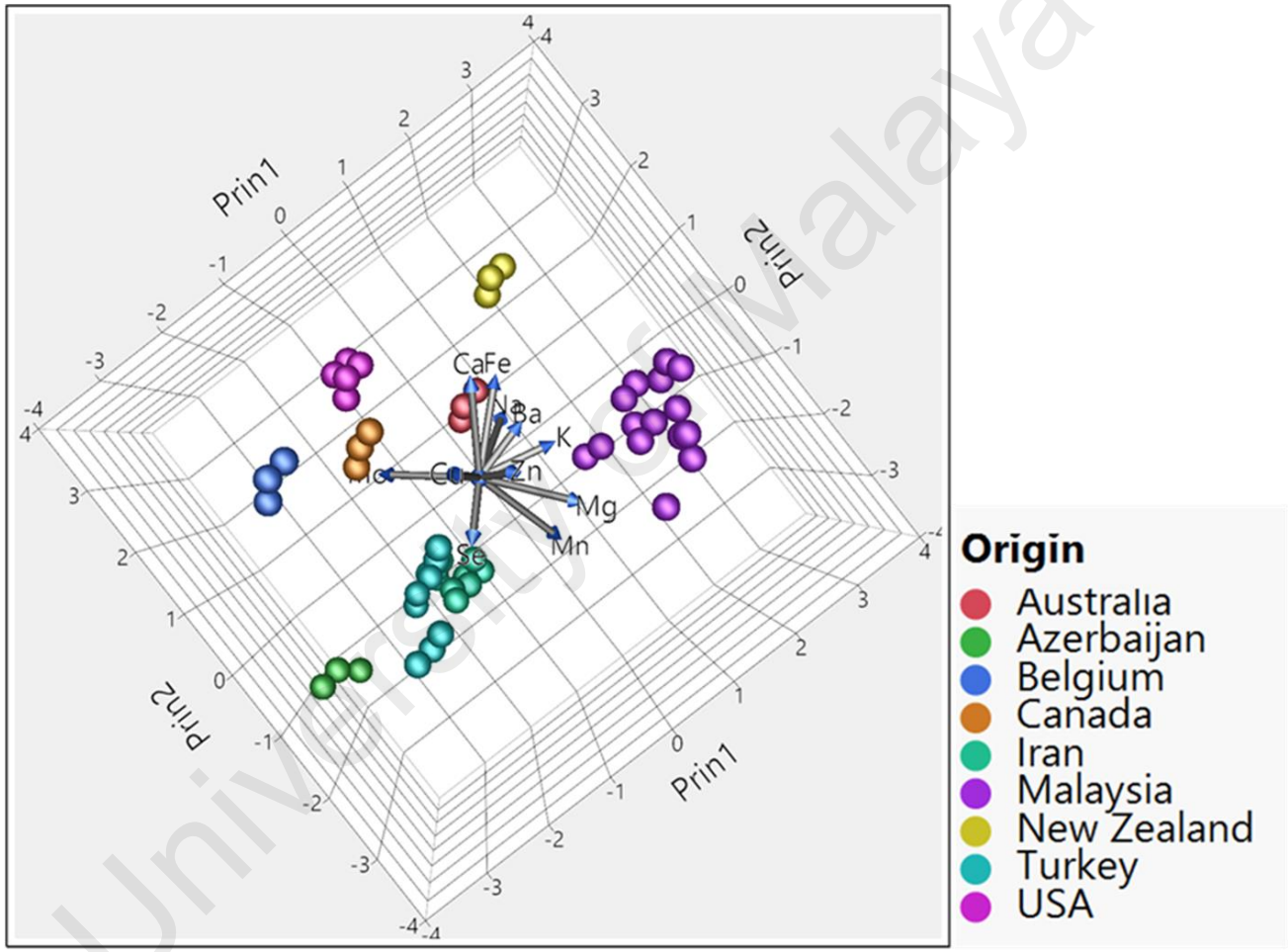

Figure 4.13: 3D varimax rotation plot of PC1 vs PC2 vs PC3.

4.1.12 Hierarchical cluster analysis (HCA) for factory milks

HCA results points out a clear separation of tropical and Australasian factory milk samples from the selected regions of the world. In Figure 4.14, Ward's algorithm was applied and a dendogram was obtained based on the same descriptors used for PCA. It is noted that there are two main clusters. Cluster 1 consists of samples from Malaysia and Australasia and cluster 2 includes samples from Europe, America, Middle East and Europe. In addition to this it is observed that samples from the Middle East consisting of Turkey, Iran and Azerbaijan and samples from America including U.S.A and Canada and Australasian samples which include Australia and New Zealand are all grouped next to their sub groups.

A two way analysis (mixture of HCA and heat plots) was also applied to the data. Figure 4.15 shows the elemental concentration of milk samples from some selected countries. There are two cluster analysis in the graph one of which clusters the samples based on their country of origin and the other is used for clustering the samples based on the detected elements. As an example Belgium that has the bluest cell is expected to have the lowest concentration of Na among all the other selected countries. Based on the combined plot it is easier to understand the variation and grouping of each sampling location. Referring to what has been explained previously, the elemental concentration shown in the cells as the red color gets more intense indicating that the sample has a high concentration of the particular element in the region and vice versa, as the cell is more intense in blue it is clear that the sample has its lowest concentration of that element in the region. Figure 4.16 shows the elemental distribution of milk samples based on the continent of origin, which can be explained in a similar way to Figure 4.15.

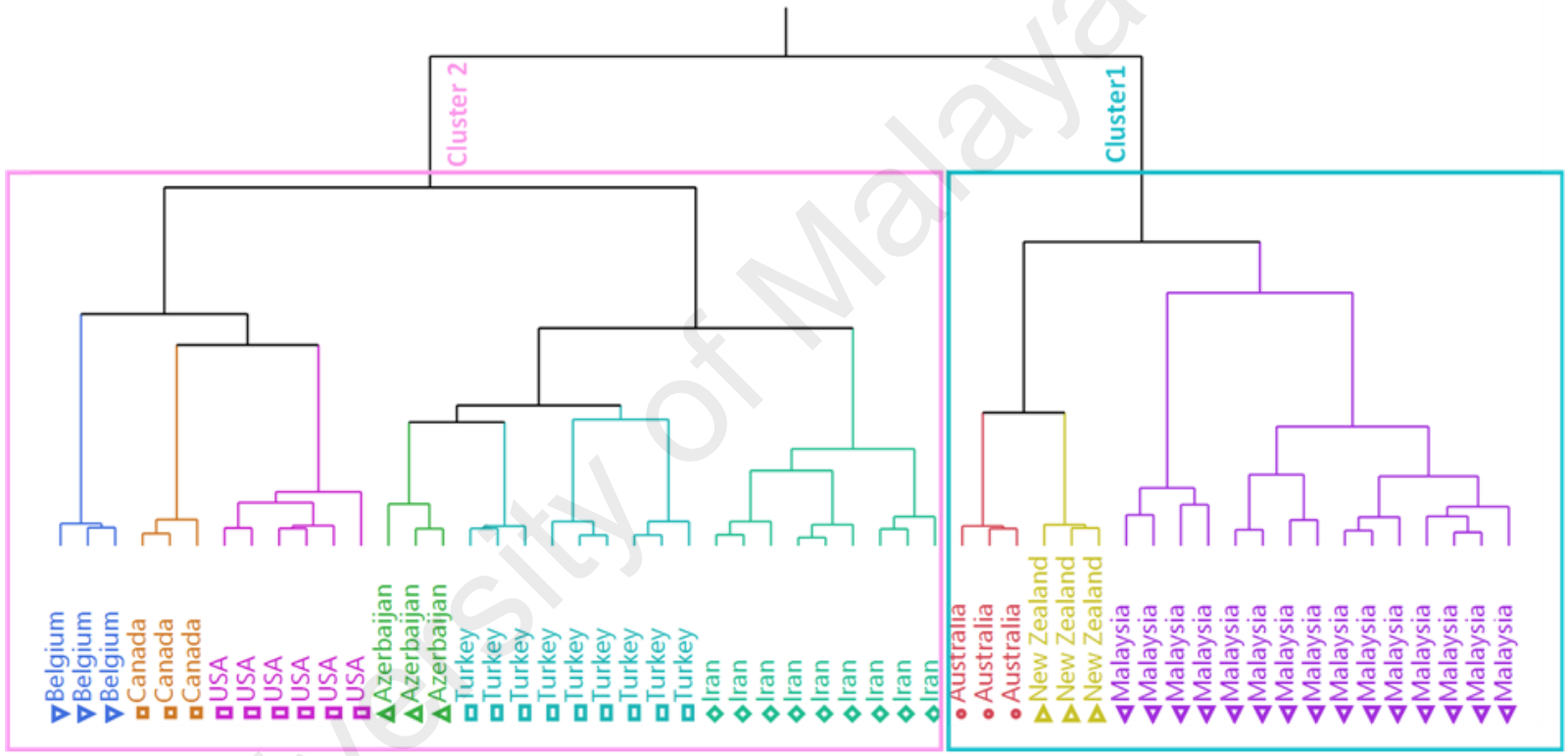

Figure 4.14: Hierarchical graph for all factory milk samples.

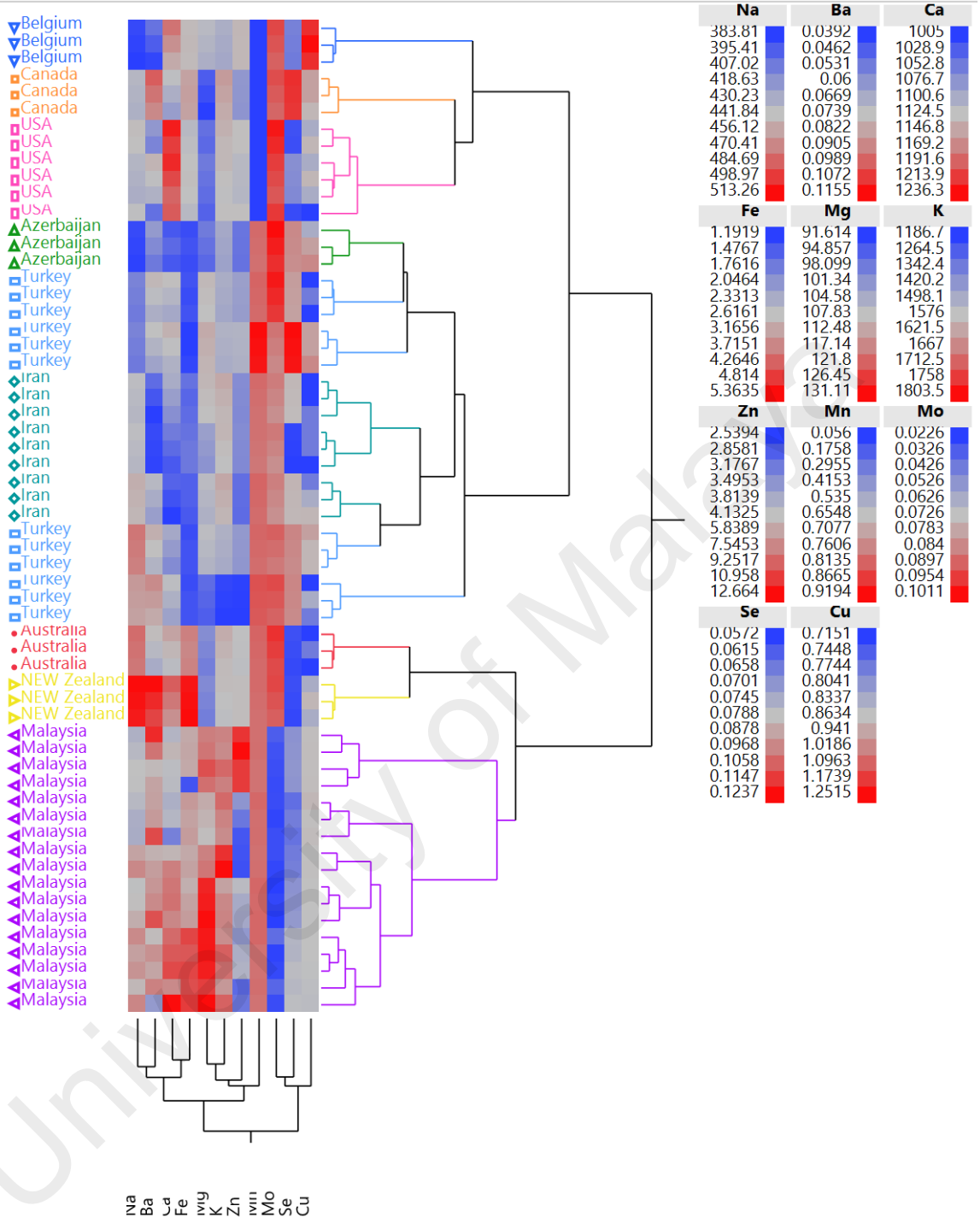

Figure 4.15 : HCA and heat plots of elemental distribution in milk samples elements based on the countries of origin.

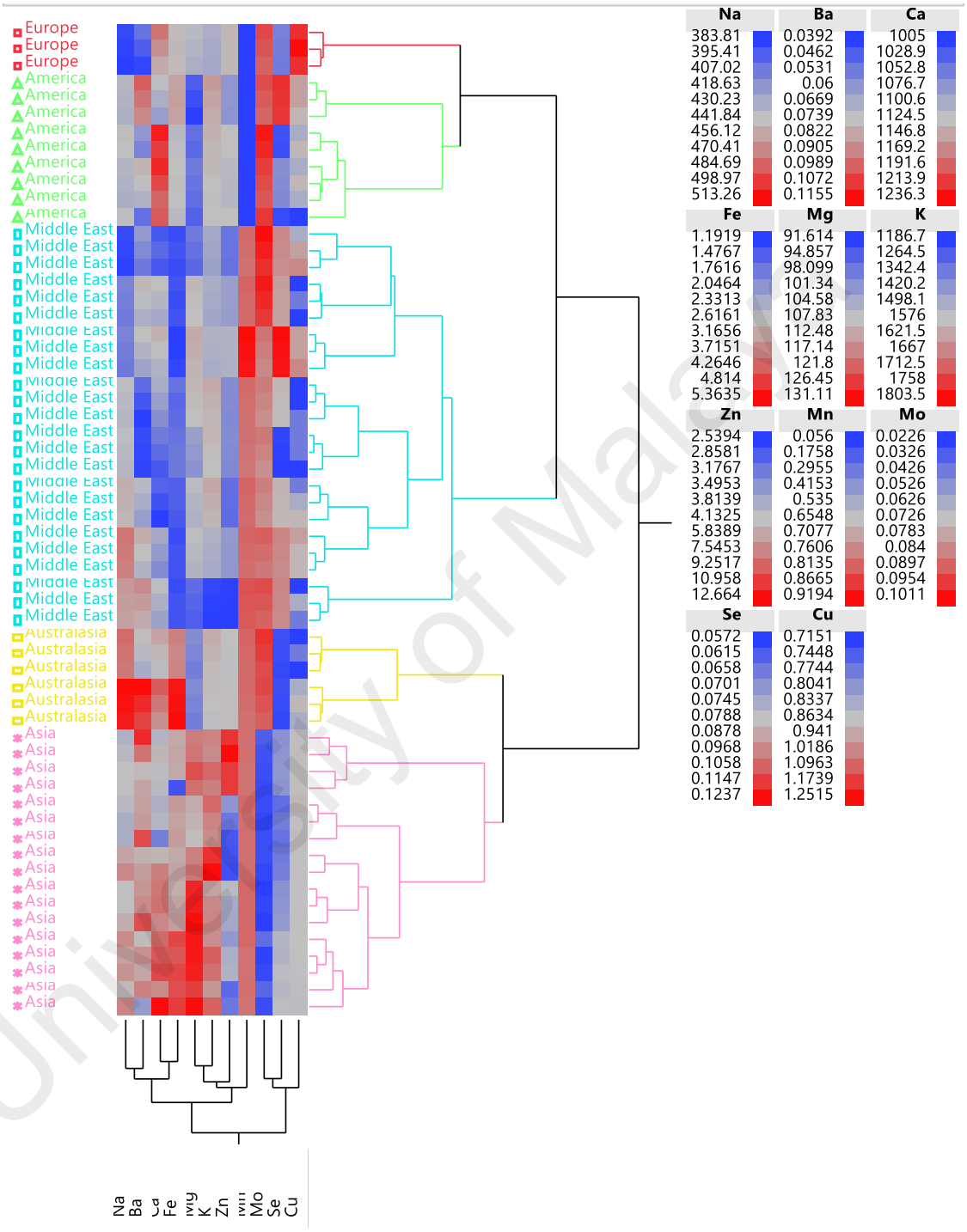

Figure 4.16 : HCA and heat plots of elemental distribution of milk samples from some selected continents of origin.

4.1.13 Constellation plots

Constellation plots are another way of presenting the hierarchical structure of the samples. In these plots, countries are specified at the end point of the branches. Each group joins a new point with lines that define the members of that cluster. As observed from Figure 4.17, two main clusters, of 1 and 2 are obtained. Malaysian and Australasian samples are grouped in cluster 1 while cluster 2 is representative of samples from other regions of the world.

Similar to the clustering shown in Fig 4.11, where milk samples from Malaysia and Australasia cluster together and are separated from milk samples of other selected regions of the world, the constellation plot in Fig. 4.17 shows this differentiation in a clearer manner. The constellation plots of Fig. 4.18 (a, b) and Fig. 4.19 (a, b) are also much clearer compared to the principal component biplots in showing that Malaysian milk samples are grouped separately from the other regions. The constellation plots, which are also based on mineral and trace elements content in the milk samples give a very clear picture and are very effective in clustering the milk samples in this research. The plots indicate the separation of the sub clusters clearly. The grouping observed in these constellation plots is the same as that observed in the biplots except that constellation plots show the clustering in a clearer manner. Clearly, constellation plots are an effective way in showing the separation of milk samples as the separation of sub clusters is also seen clearly.

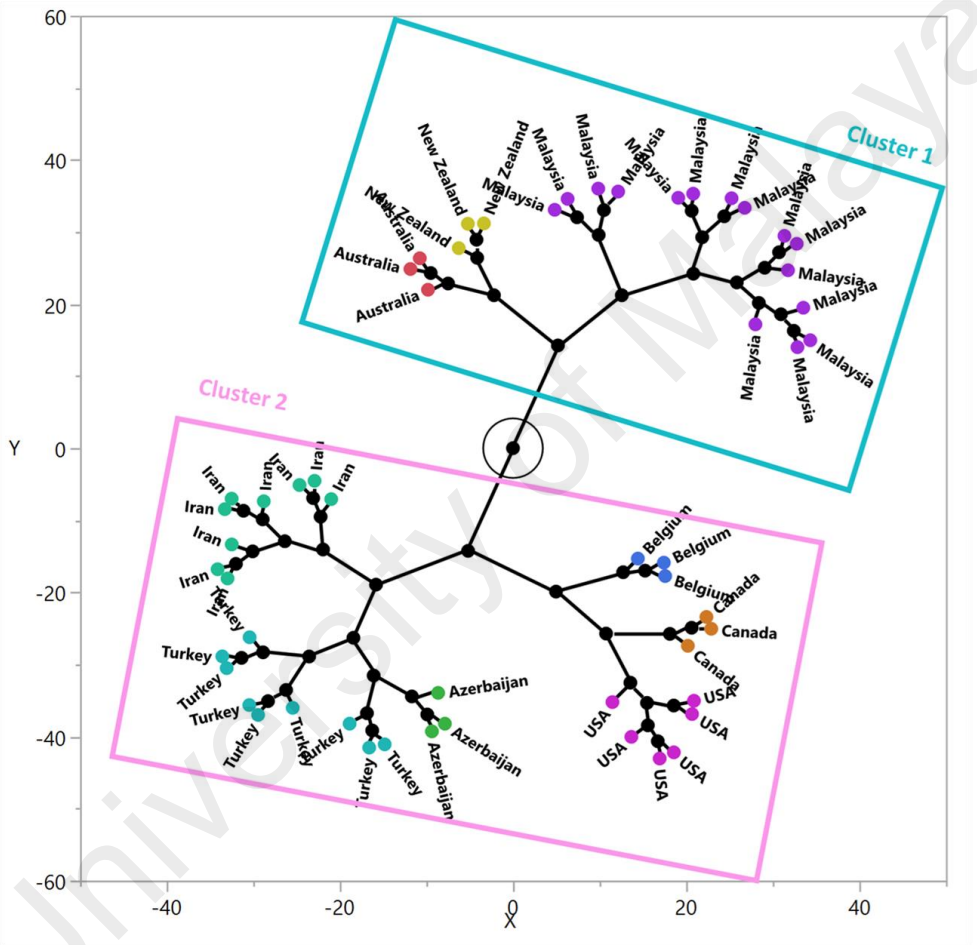

Figure 4.17: Constellation plot used for separation of Malaysian and Australasian milk samples from other selected regions of the world.

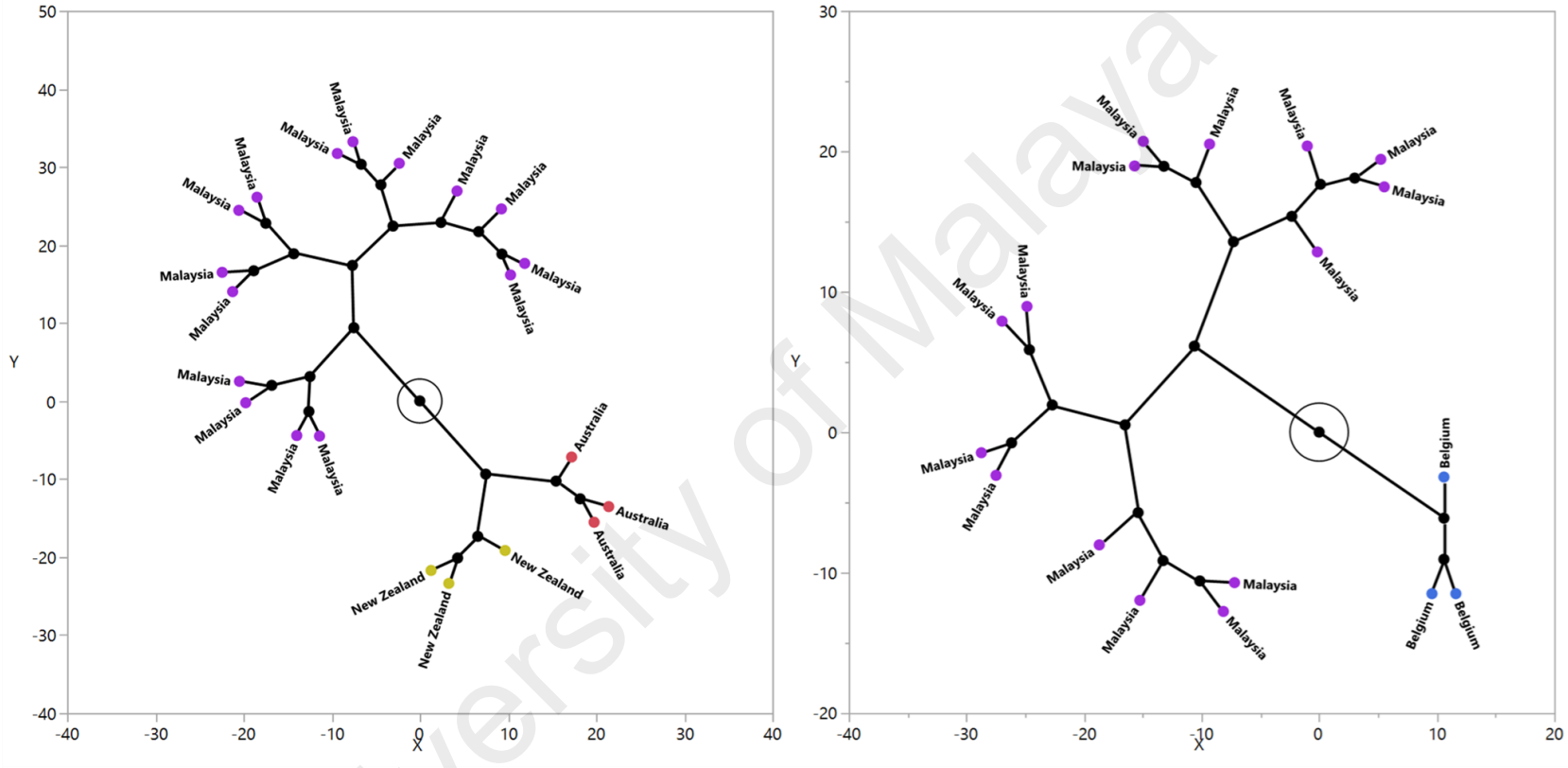

Figure 4.18: Constellation plots of (a) Malaysia and Australasia (b) Malaysia and Europe (Belgium).

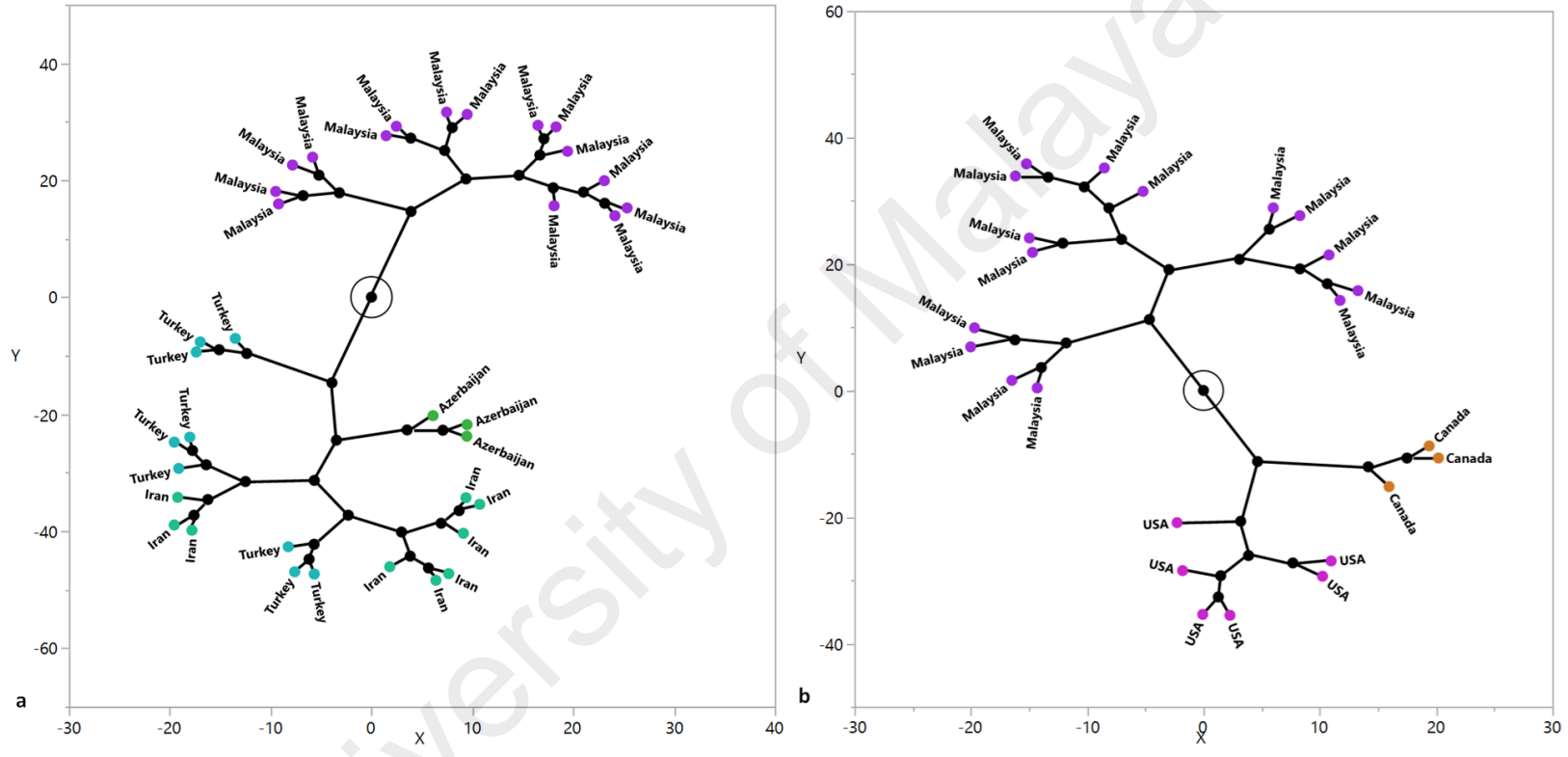

Figure 4.19: Constellation plots of (a) Malaysia and Middle Eastern countries of (Turkey, Iran and Azerbaijan) (b) Malaysia and America (Canada and U.S.A).

4.1.14 Toxic elements in all types of milk

Toxic trace elements such as U, As, Ag, Cd, Pb and Hg are not detected in any of the raw cow milk samples under this study.

4.1.15 Comparison of concentration of elements in this research with the literature

Concentrations of milk samples of this research are compared to samples from Italy (Benincasa *et al*., 2008), China (Jiang *et al*., 2015), U.SA. (Nóbrega *et al*., 1997), Turkey (Simsek *et al*., 2000) Spain (Sola-Larrañaga & Navarro-Blasco, 2009), Brazil (Soares *et al*., 2010). The concentration of the elements listed in the Table 4.6 follows the order of K>Ca>Na>Mg both in our data and in the literature whereas trace elements follow the order of Zn>Fe>Cu>Mo>Se>Mn>Ba.

From our results it can be observed that the concentration of macronutrients of Na, K and Ca and trace elements Fe and Zn are higher in Malaysian milk samples compared to samples from Spain, Brazil, Italy, China, Turkey, U.S.A and Iran. Brazil has the highest concentration of Cu in comparison to Malaysia>Iran>Turkey. Malaysia has the highest Mo concentration compared to U.S.A and Italy. Se of Malaysia and Iran are fairly the same but still higher compared to other regions. The only regions that are nearly similar in concentration of Mn and Mg are Malaysia and Brazil and these are higher compared to other regions.

In this work, we have also looked at the correlation of milk samples with food and environmental samples taken from the farms where the milk were obtained to see if there are similarities between these samples.

Table 4.6: Literature data for raw cow milk samples from other countries compared to Malaysia and Iran.

Elements	Spain	Brazil	Italy	China	Turkey	U.S.A	Malaysia	Iran
Na	372±40 [a]	NR	NR	475[a]	NR	NR	495.1±15.8[b]	430.5±1.7[b]
Mg	91.8±9.4 [a]	105.0±15.0 [a]	NR	87[a]	NR	NR	116.1±21.1[b]	101.7±0.9[b]
K	1344±65[a]	NR	1190.0[b]	1490[a]	NR	NR	1573.4±279.7[b]	1488.7±10.9[b]
Ca	970 ± 98[a]	888±131[a]	1220.0[b]	890[a]	NR	NR	1228.8±216.7[b]	1059.9±11.2[b]
Mn	0.021±0.009 [a]	0.081±0.058 [a]	0.032[b]	0.040[a]	NR	0.023 ± 0.001[a]	0.080±0.003[b]	0.079±0.004[b]
Zn	4.63±0.85 [a]	4.59±1.37 [a]	3.81[b]	3.02[a]	3.77[b]	4.85±0.41[a]	6.20±0.99[b]	3.31±0.02[b]
Se	0.009±0.008 [a]	NR	NR	NR	NR	NR	NR	0.089±0.025[b]
Mo	NR	NR	0.029[b]	NR	NR	NR	NR	0.098±0.001[b]
Ba	NR	NR	NR	NR	NR	NR	NR	0.067±0.007[b]
Fe	NR	1.05±0.89[a]	NR	NR	1.01[b]	NR	0.056 ± 0.005[a]	0.089±0.007[b]
Cu	NR	1.730±0.890[a]	NR	NR	0.390[b]	NR	0.037 ± 0.001[a]	0.050±0.010[b]

[a] mg L^{-1}, [b]mg kg^{-1} NR: not reported

4.1.16 Correlation between elements in milk with other samples

Table 4.7 presents the Pearson correlation of each single element in milk samples with that of hair, pellet, Napier grass, soil, drinking water and mixed food samples. From this table it can be inferred that as the correlation coefficient gets closer to one there is a strong correlation between the elements in milk and any one of the samples. High positive and negative correlation between elements in milk with elements in hair, pellet, Napier grass, soil, drinking water, and mix food samples suggests that sources of elements imparted to milk from animal are correlated.

It is noticed that Na in milk does not have a high correlation with Na in any of the samples digested. In relation to this, as we were informed by the farmers that cows are always treated with salt rock so the presence of Na in milk is mostly related to licking the rock salt as an everyday habit. Mg in milk has a correlation coefficient of 0.6791 with soil, 0.6585 with mix food and 0.5456 with pellet from here it is noticed that Mg in milk could have come from soil, mix food or pellet. K in milk does not have high correlation with K in any of the food and environmental samples. Ca in milk has a high correlation of 0.8171 with mixed food and around 0.5716 with pellet. Here, it can be said that K could be transferred to milk from either the mix food or pellet. On the other hand Cu has a high positive correlation with water and mix food. A low correlation is observed between Mn in milk and the other samples but there is a correlation of 0.4808 seen with mixed food.

Zn in milk has a correlation of 0.5791 with Zn in pellet. Se in milk is correlated with that of pellet with a correlation coefficient of 0.4357. Mo in milk is correlated with Mo in hair with correlation coefficient of 0.6238. Ba in milk is correlated with Ba in hair with a correlation coefficient of 0.7148. Fe in milk is correlated with iron in pellet with a correlation coefficient of 0.5254.

Table 4.7: Correlation of elements various matrices.

	Na Milk	**Na Pellet**	**Na grass**	**Na soil**	**Na water**	**Na Hair**	**Na Mix**
Na Milk	1.0000	-0.3898	-0.3329	0.1709	-0.1112	-0.2601	0.2116
Na Pellet	-0.3898	1.0000	0.7352	-0.2996	-0.1914	0.1697	0.4643
Na grass	-0.3329	0.7352	1.0000	0.3606	0.1793	-0.3393	0.4810
Na soil	0.1709	-0.2996	0.3606	1.0000	0.6865	-0.9288	0.1038
Na water	-0.1112	-0.1914	0.1793	0.6865	1.0000	-0.7477	0.0957
Na Hair	-0.2601	0.1697	-0.3393	-0.9288	-0.7477	1.0000	-0.2109
Na Mix	0.2116	0.4643	0.4810	0.1038	0.0957	-0.2109	1.0000
	Mg Milk	**Mg Pellet**	**Mg grass**	**Mg Soil**	**Mg water**	**Mg Hair**	**Mg Mix**
Mg Milk	1.0000	0.5456	0.1440	0.6791	0.2150	0.4013	0.6585
Mg Pellet	0.5456	1.0000	0.1862	0.6293	0.5037	0.6491	0.4206
Mg grass	0.1440	0.1862	1.0000	0.0179	0.0691	0.0852	0.3578
Mg Soil	0.6791	0.6293	0.0179	1.0000	0.7590	0.8985	0.3296
Mg water	0.2150	0.5037	0.0691	0.7590	1.0000	0.9104	-0.0914
Mg Hair	0.4013	0.6491	0.0852	0.8985	0.9104	1.0000	0.0004
Mg Mix	0.6585	0.4206	0.3578	0.3296	-0.0914	0.0004	1.0000
	K Milk	**K Pellet**	**K Grass**	**K Soil**	**K Water**	**K Hair**	**K Mix**
K Milk	1.0000	-0.2821	0.4369	0.4326	-0.2679	0.1978	-0.2265
K Pellet	-0.2821	1.0000	0.4243	0.3510	0.5740	-0.7816	0.7106
K Grass	0.4369	0.4243	1.0000	0.7837	-0.1840	-0.7403	0.4634
K Soil	0.4326	0.3510	0.7837	1.0000	-0.0266	-0.5326	0.2583
K Water	-0.2679	0.5740	-0.1840	-0.0266	1.0000	-0.0009	0.5440
K Hair	0.1978	-0.7816	-0.7403	-0.5326	-0.0009	1.0000	-0.5916
K Mix	-0.2265	0.7106	0.4634	0.2583	0.5440	-0.5916	1.0000
	Ca Milk	**Ca Pellet**	**Ca Grass**	**Ca Soil**	**Ca Water**	**Ca Hair**	**Ca Mix**
Ca Milk	1.0000	0.5716	-0.0102	0.2037	-0.7109	0.0128	0.8171
Ca Pellet	0.5716	1.0000	0.4716	0.3384	-0.1296	0.1801	0.8905
Ca Grass	-0.0102	0.4716	1.0000	0.2917	0.4934	0.5766	0.3330
Ca Soil	0.2037	0.3384	0.2917	1.0000	-0.0319	0.5292	0.3176
Ca Water	-0.7109	-0.1296	0.4934	-0.0319	1.0000	0.5283	-0.3183
Ca Hair	0.0128	0.1801	0.5766	0.5292	0.5283	1.0000	0.2676
Ca Mix	0.8171	0.8905	0.3330	0.3176	-0.3183	0.2676	1.0000
	Mo Milk	**Mo Pellet**	**Mo Grass**	**Mo Soil**	**Mo Water**	**Mo Hair**	**Mo Mix**
Mo Milk	1.0000	-0.0215	0.1836	0.1276	0.4676	0.6238	0.2772
Mo Pellet	-0.0215	1.0000	0.1620	0.4994	-0.1516	-0.3124	0.0920
Mo Grass	0.1836	0.1620	1.0000	0.2709	-0.4007	-0.2105	0.5622
Mo Soil	0.1276	0.4994	0.2709	1.0000	0.3495	0.2492	0.6956
Mo Water	0.4676	-0.1516	-0.4007	0.3495	1.0000	0.8978	0.3580
Mo Hair	0.6238	-0.3124	-0.2105	0.2492	0.8978	1.0000	0.4554
Mo Mix	0.2772	0.0920	0.5622	0.6956	0.3580	0.4554	1.0000

Table 4.7 continued: Correlation of elements various matrices.

	Ba Milk	**Ba Pellet**	**Ba Grass**	**Ba Soil**	**Ba Water**	**Ba Hair**	**Ba Mix**
Ba Milk	1.0000	0.1825	-0.0499	-0.2458	0.7148	0.3567	0.4662
Ba Pellet	0.1825	1.0000	0.3424	0.5183	0.4369	0.7149	0.6079
Ba Grass	-0.0499	0.3424	1.0000	-0.1352	0.2968	-0.1944	0.7518
Ba Soil	-0.2458	0.5183	-0.1352	1.0000	0.1713	0.3012	0.1485
Ba Water	0.7148	0.4369	0.2968	0.1713	1.0000	0.1608	0.8028
Ba Hair	0.3567	0.7149	-0.1944	0.3012	0.1608	1.0000	0.1289
Ba Mix	0.4662	0.6079	0.7518	0.1485	0.8028	0.1289	1.0000
	Fe Milk	**Fe Pellet**	**Fe Grass**	**Fe Soil**	**Fe Water**	**Fe Hair**	**Fe Mix**
Fe Milk	1.0000	0.5254	0.0330	-0.2813	0.0006	-0.3877	-0.4200
Fe Pellet	0.5254	1.0000	-0.1507	-0.3535	0.5313	0.5225	0.4053
Fe Grass	0.0330	-0.1507	1.0000	0.4695	-0.0262	-0.0282	0.0317
Fe Soil	-0.2813	-0.3535	0.4695	1.0000	-0.3307	0.0344	0.3527
Fe Water	0.0006	0.5313	-0.0262	-0.3307	1.0000	0.5114	0.5024
Fe Hair	-0.3877	0.5225	-0.0282	0.0344	0.5114	1.0000	0.9112
Fe Mix	-0.4200	0.4053	0.0317	0.3527	0.5024	0.9112	1.0000
	Cu Milk	**Cu Pellet**	**Cu Grass**	**Cu Soil**	**Cu Water**	**Cu Hair**	**Cu Mix**
Cu Milk	1.0000	0.4451	-0.0730	-0.4336	0.7776	-0.1082	0.5383
Cu Pellet	0.4451	1.0000	-0.2806	-0.0192	0.4717	0.2059	0.7661
Cu Grass	-0.0730	-0.2806	1.0000	0.2536	0.1535	0.7797	-0.1351
Cu Soil	-0.4336	-0.0192	0.2536	1.0000	-0.2442	0.5500	0.1219
Cu Water	0.7776	0.4717	0.1535	-0.2442	1.0000	0.1621	0.2652
Cu Hair	-0.1082	0.2059	0.7797	0.5500	0.1621	1.0000	0.2341
Cu Mix	0.5383	0.7661	-0.1351	0.1219	0.2652	0.2341	1.0000
	Zn Milk	**Zn Pellet**	**Zn Grass**	**Zn Soil**	**Zn Water**	**Zn Hair**	**Zn Mix**
Zn Milk	1.0000	0.5791	-0.0920	-0.3554	0.2123	-0.7511	-0.0767
Zn Pellet	0.5791	1.0000	-0.3365	-0.4199	-0.0652	-0.3193	0.1389
Zn Grass	-0.0920	-0.3365	1.0000	-0.1483	-0.2135	0.2181	0.4567
Zn Soil	-0.3554	-0.4199	-0.1483	1.0000	-0.0464	0.0599	0.1501
Zn Water	0.2123	-0.0652	-0.2135	-0.0464	1.0000	-0.2993	-0.3617
Zn Hair	-0.7511	-0.3193	0.2181	0.0599	-0.2993	1.0000	0.2126
Zn Mix	-0.0767	0.1389	0.4567	0.1501	-0.3617	0.2126	1.0000
	Mn Milk	**Mn Pellet**	**Mn Grass**	**Mn Soil**	**Mn Water**	**Mn Hair**	**Mn Mix**
Mn Milk	1.0000	0.1900	-0.0120	-0.0663	0.3586	0.1198	0.4808
Mn Pellet	0.1900	1.0000	0.2995	-0.4341	-0.1713	0.4170	0.7636
Mn Grass	-0.0120	0.2995	1.0000	0.0774	0.1005	0.2855	0.3786
Mn Soil	-0.0663	-0.4341	0.0774	1.0000	0.2206	-0.8345	-0.2133
Mn Water	0.3586	-0.1713	0.1005	0.2206	1.0000	0.1920	0.3153
Mn Hair	0.1198	0.4170	0.2855	-0.8345	0.1920	1.0000	0.4792
Mn Mix	0.4808	0.7636	0.3786	-0.2133	0.3153	0.4792	1.0000
	Se Milk	**Se Pellet**	**Se Grass**	**Se Soil**	**Se Water**	**Se Hair**	**Se Mix**
Se Milk	1.0000	0.4357	0.3634	-0.1386	0.3358	-0.1060	0.2609
Se Pellet	0.4357	1.0000	-0.3287	0.4712	0.2463	-0.2363	0.0500
Se Grass	0.3634	-0.3287	1.0000	-0.5742	0.2410	0.4067	0.5837
Se Soil	-0.1386	0.4712	-0.5742	1.0000	0.1451	0.0634	-0.1045
Se Water	0.3358	0.2463	0.2410	0.1451	1.0000	-0.1067	0.1404
Se Hair	-0.1060	-0.2363	0.4067	0.0634	-0.1067	1.0000	0.4017
Se Mix	0.2609	0.0500	0.5837	-0.1045	0.1404	0.4017	1.0000

4.2 Data analysis of IRMS results

In this section the results obtained from isotopic ratio mass spectrometer will be discussed.

4.2.1 Isotopic ratio information for milk

The isotopic ratio of cow milk samples is shown in Figure 4.20.

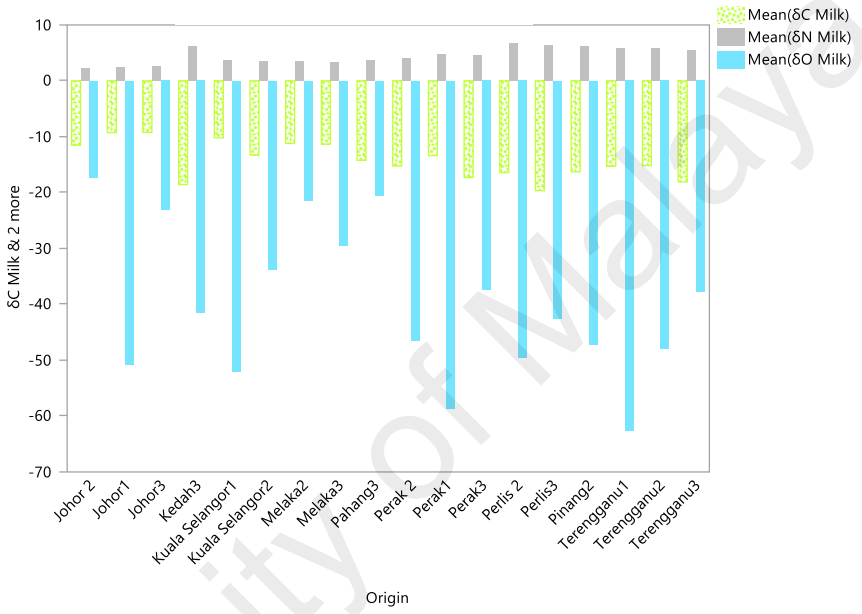

Figure 4.20: Isotopic ratio of cow milk samples.

For a clearer view we have looked into the variation of each isotopic ratio of C, N and O in the northern and southern regions of Peninsular Malaysia using cell plots as shown in Figure 4.21-4.23. It is noticed that the isotopic ratio of carbon in Malaysian milk is related to the abundance of C4 plants near the equator. The value of $\delta^{13}C$ in milk decreases in samples collected as we move from the equator towards the poles. The $\delta^{15}N$ values in cow's milk are higher in the northern region compared to the south and this is most probably due to the fact that cows in the north are mostly grass fed and organic fertilisers are used more frequently compared to chemical fertilisers. The isotopic ratio of oxygen is seen to be dependent on the latitude and the amount of

precipitation of each region and the samples from the south are seen to be richer in heavy oxygen isotope.

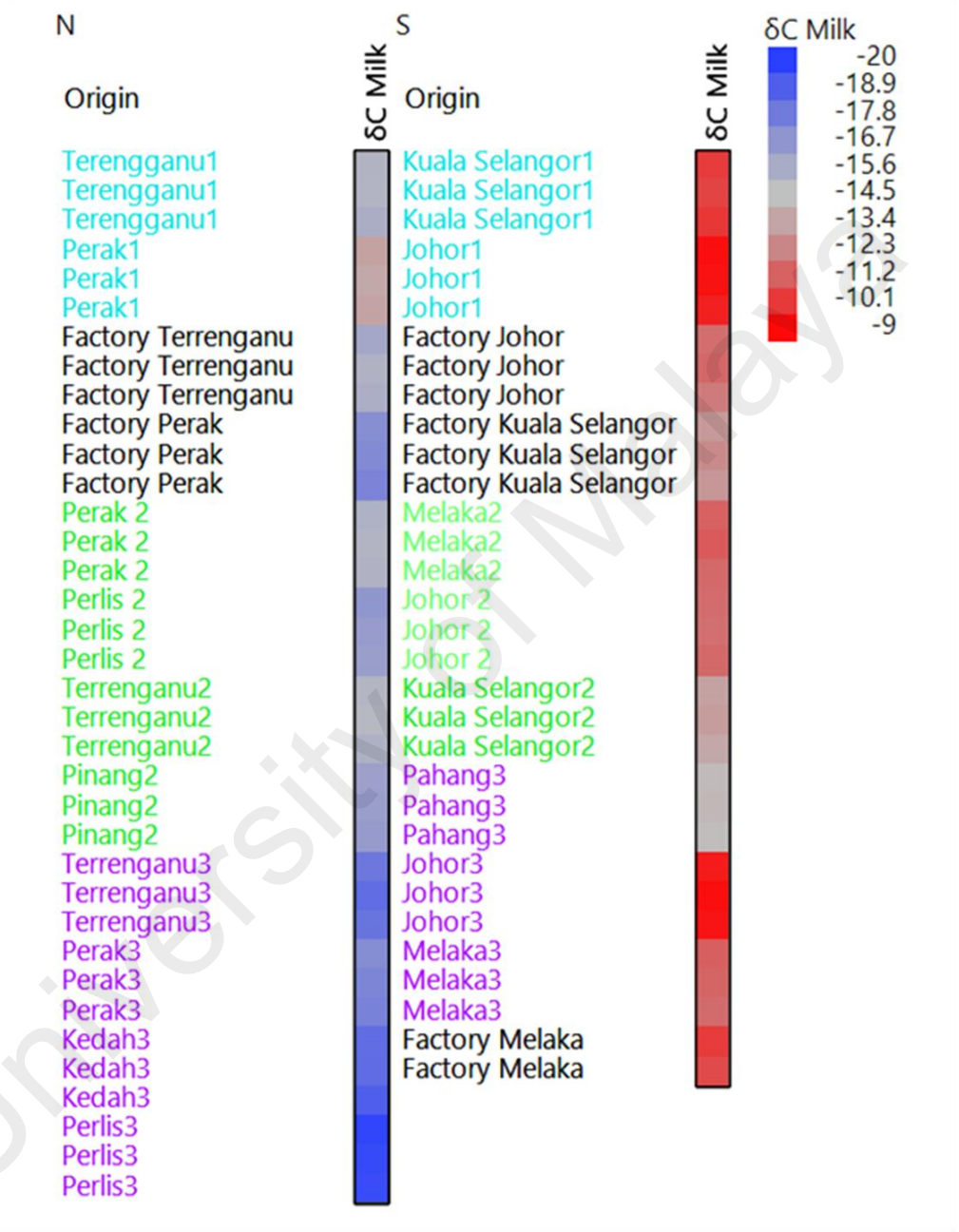

Figure 4.21: Cell plots of isotopic ratios of C.

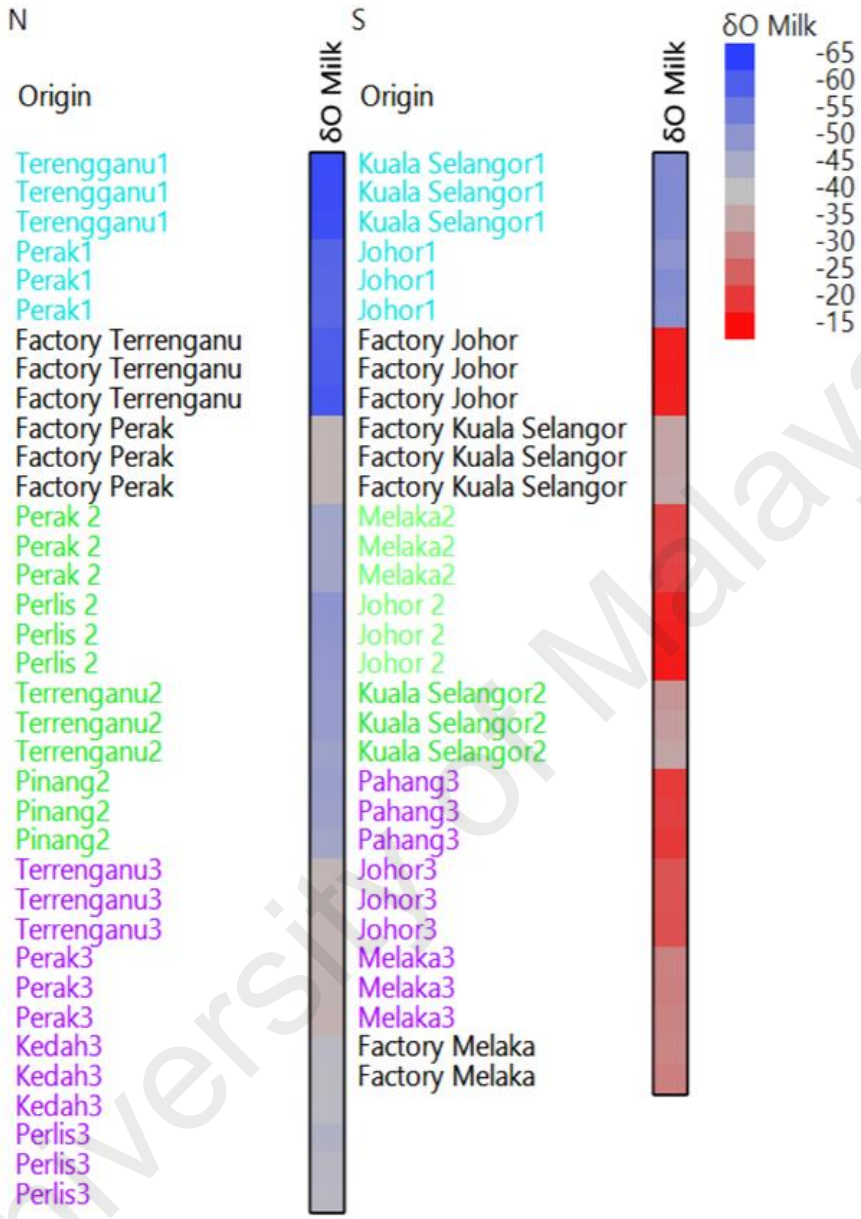

Figure 4.22: Cell plots of isotopic ratios of O.

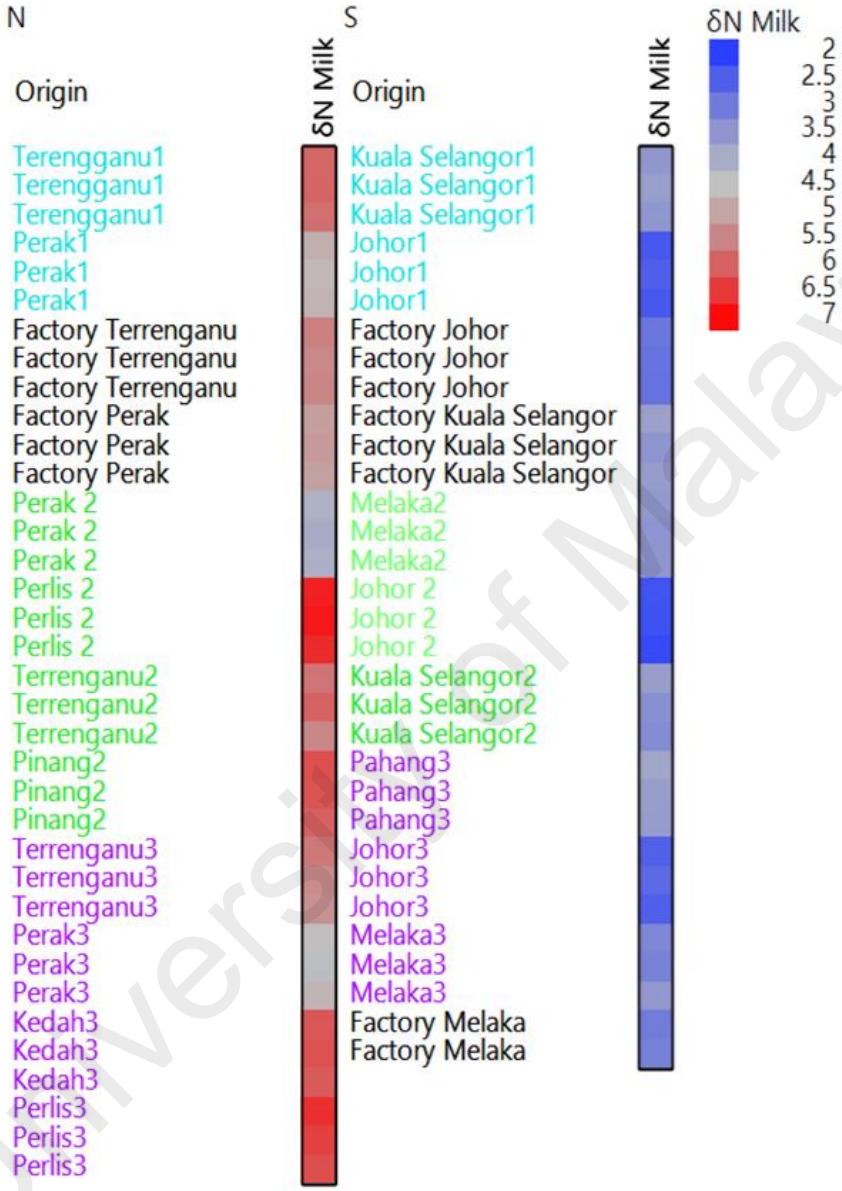

Figure 4.23: Cell plots of isotopic ratios of N.

The mean isotopic ratios of carbon, nitrogen, oxygen and hydrogen for raw cow and factory milk samples collected in sampling S1 and S3 are reported in Table 4.8 and isotopic ratios for milk and hair collected during S2 are reported in Table 4.9. It is noted that the mean value of $\delta^{13}C$ in southern milk samples is in the range of -14.23 to -9.22 which is less negative (higher) in milk samples from south compared to north. Northern milk samples have $\delta^{13}C$ values in the range of -18.10 to -13.40.

In the case of nitrogen the mean isotopic ratio in southern sampling regions is in the range of 2.44 to 3.73 which is lower compared to northern region which is in the range of 4.55 to 6.42. Mean value of $\delta^{18}O$ varied a lot depending to the amount of precipitation but mostly it is richer for samples from the southern region.

Table 4.8: Milk samples Mean±SD for isotopic ratios of C, N and O.

Origin	$\delta^{13}C$ (Mean ±SD)	$\delta^{15}N$ (Mean ±SD)	$\delta^{18}O$ (Mean ±SD)
Kuala Selangor1	-10.21± 0.16	3.59 ± 0.09	-52.03 ± 0.11
Johor1	-9.26 ± 0.23	2.44 ± 0.06	-50.81 ± 0.87
Terrenganu1	-15.31 ± 0.18	5.9 ± 0.08	-62.69 ± 0.36
Perak1	-13.40 ± 0.13	4.75 ± 0.11	-58.59 ± 0.37
Johor3	-9.22 ± 0.14	2.59 ± 0.09	-23.12 ± 0.22
Melaka3	-11.31 ± 0.19	3.31 ± 0.22	-29.47 ± 0.30
Pahang3	-14.23 ± 0.13	3.73 ± 0.11	-20.48 ± 0.47
Terrenganu3	-18.10 ± 0.22	5.47 ± 0.19	-37.80 ± 0.16
Perak3	-17.31 ± 0.24	4.55 ± 0.15	-37.47 ± 0.21
Kedah3	-18.57 ± 0.30	6.15 ± 0.06	-41.52 ± 0.19
Perlis3	-19.67 ± 0.13	6.42 ± 0.19	-42.56 ± 0.82
Factory Johor	-11.64 ± 0.30	2.87 ± 0.04	-17.08 ± 0.19
Factory Melaka	-10.36 ± 0.19	3.14 ± 0.14	-29.26 ± 0.80
Factory Kuala Selangor	-12.77 ± 0.28	3.63 ± 0.12	-35.18 ± 0.42
Factory Perak	-17.2 ± 0.21	5.12 ± 0.06	-37.86 ± 0.10
Factory Terengganu	-15.53 ± 0.27	5.51 ± 0.07	-60.42 ± 0.52

4.2.2 Box plots of elemental isotopic ratio for all milk samples

Box and whisker plot for all the milk samples are plotted in Figure 4.24.

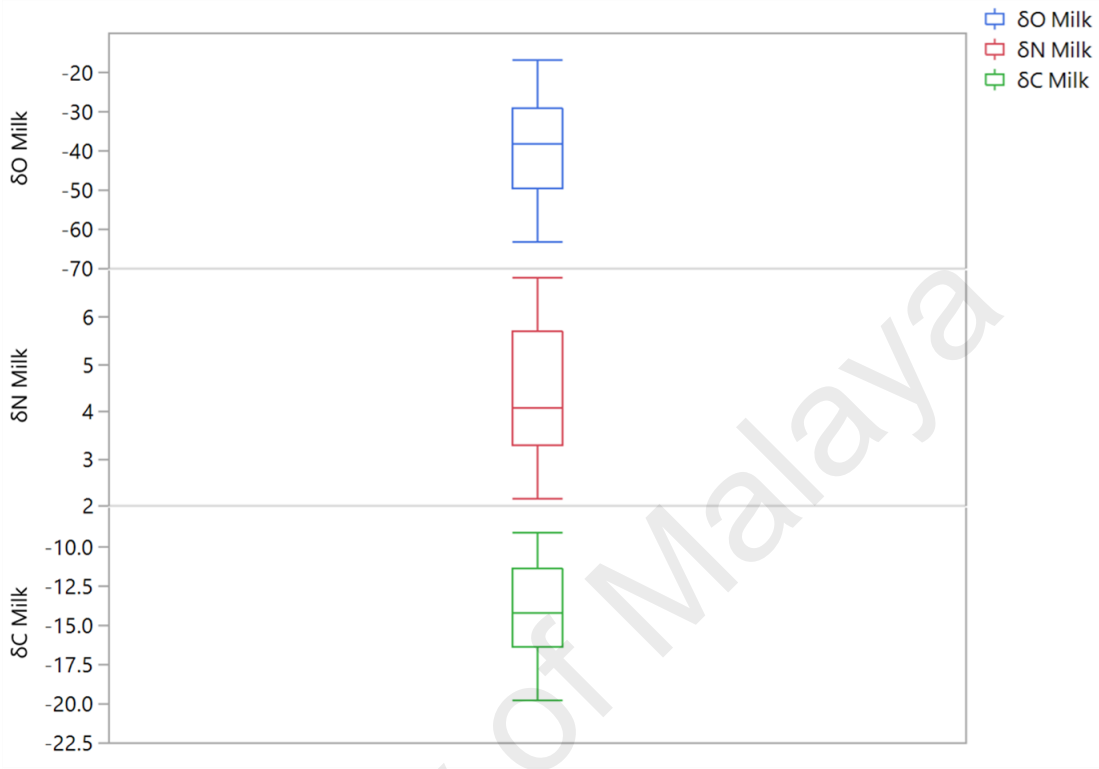

Figure 4.24: Box plot of isotopic ratio of C, N and O for milk samples from different regions.

Table 4.9: Hair samples Mean±SD for isotopic ratios of C and N.

Origin	$\delta^{13}C_{Milk}$ (Mean ±SD)	$\delta^{15}N_{Milk}$ (Mean ±SD)	$\delta^{13}C_{Hair}$ (Mean ±SD)	$\delta^{15}N_{Hair}$ (Mean ±SD)
Kuala Selangor2	-13.28 ± 0.21	3.46 ± 0.18	-13.26 ± 0.25	11.21 ± 0.09
Melaka2	-11.21 ± 0.21	3.53 ± 0.06	-12.42 ± 0.43	10.36 ± 0.25
Johor2	-11.52 ± 0.10	2.25 ± 0.07	-11.33 ± 0.12	9.19 ± 0.15
Perak2	-15.25 ± 0.08	4.06 ± 0.07	-16.12 ± 0.11	11.34 ± 0.07
Terrengganu2	-15.16 ± 0.20	5.74 ± 0.26	-17.34 ± 0.33	12.58 ± 0.09
Perlis2	-16.43 ± 0.20	6.76 ± 0.10	-21.44 ± 0.42	14.09 ± 0.04
Pinang2	-16.29 ± 0.13	6.13 ± 0.11	-18.37 ± 0.33	13.70 ± 0.05

The data for isotopic ratio of C, N and O of the box and whisker plot is reported in Table 4.10.

The data of carbon isotopic ratio as observed from the whisker plots show that the spread of the data is from -19.81 to -9.09. It is also observed that half of the samples have values greater than -13.5 with midpoint of -11.5 and the other half are less than -13.5 with a midpoint of -16. -16 is the median for all samples having values less than real median which is close to the lower end of whole spectrum of data.

In the case of nitrogen isotopic ratio, $\delta^{15}N$ varies between 2.17 to 6.86. The range of data is from 2 to 7.0. From the box and whisker plot it is noted that half of the data are less than 4 in value with a midpoint of 3.5 and the other half that have values more than 4, 5.8 is their midpoint. Overall, the true median is closer to the bottom of the data.

From the values of oxygen isotopic ratio, box and whisker plot shows variation of -16.67 to -63 where half of the samples have values less than -35. The midpoint of the data for samples having $\delta^{18}O$ values less/more than -30.8 is -50/-30. The median for samples having values less/more than the overall median is -50/-30. The median is closer to the upper section of the overall data spectrum.

Table 4.10: Milk boxplot information for C, N and O isotopic ratios.

Isotopic ratio	Median	Mean	Min	Max	1st quartile	3rd quartile
$\delta^{13}C$	-14.23	-14.05	-19.81	-9.09	-16.39	-11.4
$\delta^{15}N$	4.07	4.39	2.17	6.86	3.30	5.71
$\delta^{18}O$	-39.12	-37.89	-63	-16.67	-49.37	-29.15

4.2.3 Correlation plots between C, N and O isotopic ratios in milk samples

Original isotopic ratio correlation plot have been carried out to see if milk samples can be separated based on their geographical origin data by the using of their isotopic ratio information.

Due to the relatively small size of Peninsular Malaysia, we were not able to exactly separate the samples based on sampling collection regions but a separation based on northern and southern regions can be observed.

Figure 4.25, shows the correlation of $\delta^{13}C$ vs $\delta^{15}N$. It is noted that samples separation based on individual states is not very distinct but there is a clear separation based on the northern and southern sampling regions. As observed, samples from south have less negative $\delta^{13}C$ values compared to the north which might be explained by the different feeding regimes, low levels of precipitation in the south and the latitude effect. Moreover, samples from south have lower $\delta^{15}N$ compared to the northern regions which could be due to higher use of organic fertilisers, more pasture regime as well higher amount of legumes in soil.

Plotting $\delta^{15}N$ vs $\delta^{18}O$ for milk samples, two clusters are observed for northern and southern regions of Peninsular Malaysia as presented in Figure 4.26. Samples from the south have lower values in comparison to the north which could be related to the amount of fertilizers used. $\delta^{18}O$ values in milks of north and south do not show a clear separation although it should be noted that $\delta^{18}O$ in milk samples from the south are more negative compared to the northern milk samples.

In Figure 4.27, $\delta^{13}C$ correlation plot of milk vs $\delta^{18}O$ is shown. From this plot it is observed that two clusters of north and south are obtained. Samples from the south have less negative values for $\delta^{13}C$ than northern milk samples. This is related to the

feeding regime of cows and correlated with the amount of C4 as well as lower precipitation in the south and the latitude effect. However, in the case of $\delta^{18}O$ although a clear separation is not observed, the southern group show less negative $\delta^{18}O$ compared to the north.

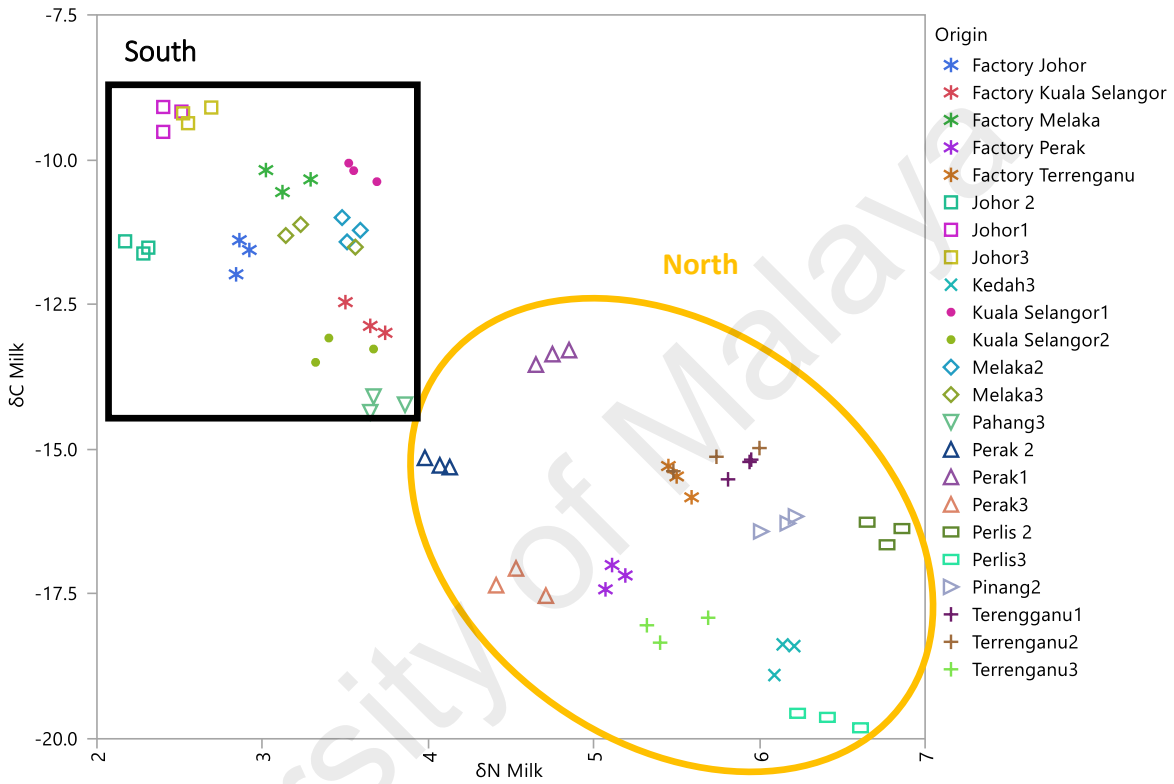

Figure 4.25: Correlation of $\delta^{13}C_{Milk}$ vs $\delta^{15}N_{Milk}$.

In general, as there are still overlaps of groups and using only $\delta^{18}O$ information does not clearly show separation. Hence, the data are further analyzed using standard chemometric methods in order to observe if better separation between the data for the north and the south can be obtained. Multivariate methods might give better separation, nevertheless, simple correlation plots of isotopic ratios do render natural separation between northern and southern milk samples in this study.

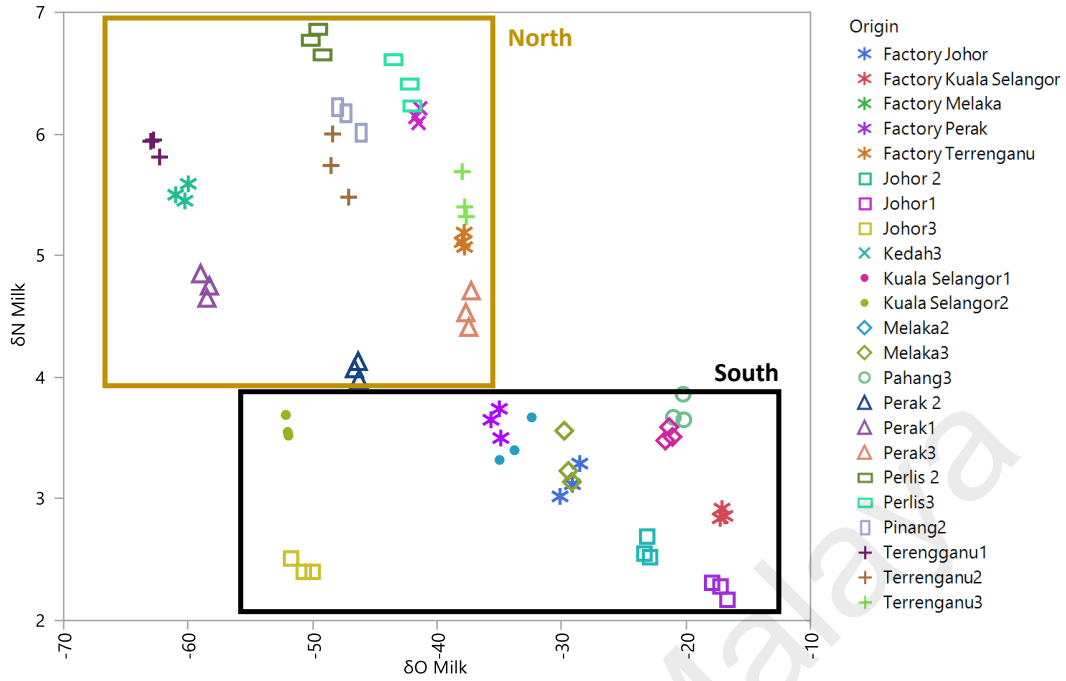

Figure 4.26: Correlation plot of $\delta^{15}N_{Milk}$ vs $\delta^{18}O_{Milk}$.

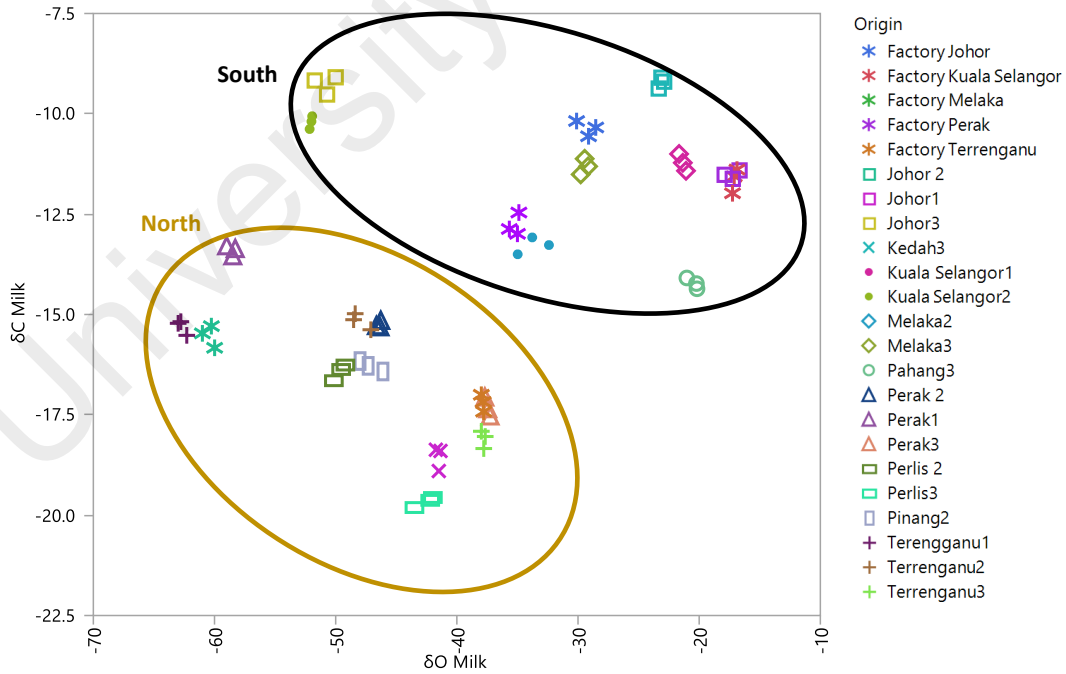

Figure 4.27: Correlation plot of $\delta^{13}C_{Milk}$ vs $\delta^{18}O_{Milk}$.

The next section presents the results of the correlation of $\delta^{15}N$ and $\delta^{13}C$ in milk with that of hair and latitude. This is to see if hair can be used as a representative of animal tissue in place of milk and of animal products for the determination of geographical origin.

4.2.4 Correlation between isotopic ratio of milk and hair

The correlation of isotopic ratio of C and N in cow milk with that of hair has been obtained using S1, S2 and S3. From the correlation in Table 4.11 it can be observed that there is a positive correlation between $\delta^{15}N$ in milk and hair with latitude which could be explained by the latitude relation with nitrogen. It is believed that at higher latitudes, the ecosystem is more sensitive to increases in the availability of nitrogen. Excess nitrogen availability might be due to various reasons, among others, anthropogenic factors and climate warming that accelerates the release of N from soil (Grace, 2014).

As there is high correlation between milk and hair and other researchers have also observed similar correlations, it is thought that hair could be used, and would be a better choice instead of milk in ascertaining elemental isotopic ratios of dairy products (Yanagi et al., 2012). This is because hair is resilient to external factors such as humidity, and temperature change, is easier to collect and offers permanence to isotopic records (Yanagi *et al*., 2012). Furthermore, hair is a keratinized tissue grown regularly, thus it can be used to record time (Schwertl *et al*., 2005; Zazzo *et al*., 2007). Our results shows similarity to work done on cattle meat and hair in Japan and China where there were also high correlations between animal tissue (hair) and animal product (meat) for carbon and nitrogen isotopic ratios (Guo *et al*., 2010; Yanagi *et al*., 2012). $\delta^{13}C$ and $\delta^{15}N$ can vary in milk and hair depending on the food consumed by the cattle

(Molkentin & Giesemann, 2010). The values of $\delta^{13}C$ and $\delta^{15}N$ are found to be higher in hair compared to milk in this work.

The values of $\delta^{13}C$ and $\delta^{15}N$ in milk samples are -9.09 ‰ to -18.9 ‰ and 2.17‰ to 6.86 ‰ respectively while those for cattle hair are from -7.42 ‰ to -23.58 ‰ for carbon and 8.13 ‰ to 15.13 ‰ for nitrogen. From the results it is observed that there are high correlations between isotopic ratios of carbon and nitrogen in milk and hair and this is also observed by other researchers as well. Based on these results and work of others, it can be concluded that animal tissue, in this case hair, could be used instead of animal product, in this case milk, for ascertaining isotopic ratio of the dairy products (Yanagi *et al*., 2012).

In addition, our observations are also in agreement with the work done on meat and hair of cattle in Japan and China as they also observed high correlations between isotopic ratios of carbon and nitrogen in animal tissue (hair) and animal product (meat) (Guo *et al*., 2010; Yanagi *et al*., 2012).

Table 4.11: Correlation table of isotopic ratios of C, N in milk with that of hair and latitude.

	δC Milk	**δN Milk**	**δC Hair**	**δN Hair**	**Latitude**
δC Milk	1.0000	-0.8486	0.9795	-0.9632	-0.8986
δN Milk	-0.8486	1.0000	-0.9069	0.9402	0.9646
δC Hair	0.9795	-0.9069	1.0000	-0.9852	-0.9271
δN Hair	-0.9632	0.9402	-0.9852	1.0000	0.9347
Latitude	-0.8986	0.9646	-0.9271	0.9347	1.0000

4.2.5 Simple correlation plots of carbon and nitrogen isotopic ratios in hair

The plot of $\delta^{13}C$ tail hair against $\delta^{15}N$ is illustrated in Figure 4.28. From the figure, it is also observed that there are two clusters of north and south. The northern cluster consists of Ipoh, Terengganu, Pulau Pinang and Perlis, which exhibit more positive $\delta^{15}N$ compared to the southern samples of Melaka, Johor and Kuala Selangor.

On the other hand, samples from the northern cluster exhibit more negative $\delta^{13}C$ compared to samples of the southern region of Malaysia. What was clear was that the characteristics of the carbon isotopic ratio vs the nitrogen isotopic ratio in hair are very similar to that of milk. Both are able to aggregate the samples based on geographical regions, although hair seemed to manifest a clearer clustering. Since hair samples are resilient to environmental changes, it would certainly be more appropriate to use hair in building the database for the geographical traceability of milk.

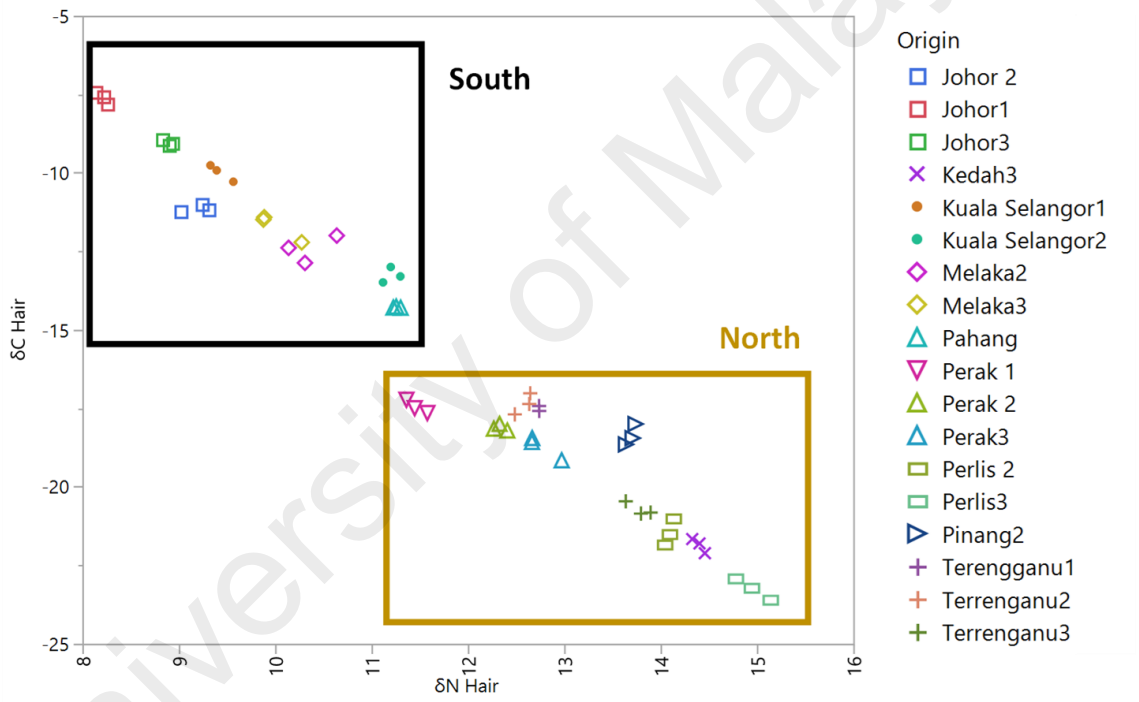

Figure 4.28: Correlation of $\delta^{13}C_{Hair}$ vs $\delta^{15}N_{Hair}$.

4.2.6 Correlation plot of isotopic ratio of milk with precipitation

The correlation of isotopic ratios of H and O of milk with that of precipitation has been obtained based on S1, S2 and S3. From the correlation in Table 4.12 it can be observed that δO_{Milk} and δH_{Milk} are negatively correlated with latitude and δO_{Rain} and δH_{Rain}. This is most probably due to the depletion of the heavier isotopes of hydrogen and oxygen in precipitation in tandem with the increase in latitude and precipitation. This effect is referred to as the latitude and amount effect (Jeon *et al.*, 2015).

Hydrogen and oxygen in milk reflects that of ground water and planst which in turn is related to the amount of these two isotopes in precipitation. Notwithstanding this, oxygen and hydrogen isotopic ratio of milk also depend on geographical factors such as the food taken by the animal (the percentage of fresh and dry grass), water consumed and the isotopic ratio of H and O of precipitation (Longobardi *et al*., 2015a).

Higher values of δ^2H and $\delta^{18}O$ indicate that cows are mostly grass fed and the amount being high due to evaporation of water (Magdas & Puscas, 2011; Renou *et al*., 2004). On the other hand, cattles which are fed with hay, maize and siladge have lower $\delta^{18}O$ level compared to grazing cows. Season is another factor (Henton, MCorriston, Martin, & Oches, 2014) since it had been observed that in the summer, cattles consume more fresh grass, hence the isotopic ratio of oxygen is higher (Bontempo *et al*., 2012). Plants in hot climate have higher $\delta^{18}O$ (less negative) in comparison to cool areas. In other words, higher values of $\delta^{18}O$ reflect warmer climate and high rainfall. $\delta^{18}O$ is also affected by latitude (Crittenden *et al*., 2007) where as the latitude increases the $\delta^{18}O$ and δ^2H decreases.

Table 4.12: Correlation Table between (δO and δH) milk with that of rain and latitude.

	δO Milk	**δH Milk**	**δO Rain**	**δH Rain**	**Latitude**
δO Milk	1.0000	0.5366	-0.6184	-0.8320	-0.9330
δH Milk	0.5366	1.0000	-0.8970	-0.6966	-0.5032
δO Rain	-0.6184	-0.8970	1.0000	0.7902	0.5912
δH Rain	-0.8320	-0.6966	0.7902	1.0000	0.8474
Latitude	-0.9330	-0.5032	0.5912	0.8474	1.0000

The high correlation between milk and precipitation in this work have also been observed by other researchers where similar correlations of δ^2H between precipitation and milk have been attributed to the hydrogen of the fatty acid in milk and bulk milk powder (Ehtesham *et al*., 2013). Based on this correlation, it is thought that precipitation

isotopic information of δ^2H and $\delta^{18}O$ could also be used for building the data base of milk geographical origin traceability.

4.2.7 Isotopic ratio of milk samples

The values of $\delta^{18}O$ in milk samples are from -50.15‰ to -17‰. More specifically this range varies from -34.98‰ to -17‰ for southern milk samples and from -49.98‰ to -37.26‰ for northern samples and they are significantly different ($p<0.05$). The isotopic ratio of hydrogen ranges from 37.11‰ to 83.98‰, however, the values range from 55.64‰ to 79.4‰ for the southern samples and from 55.64‰ to 79.4‰ for the northern samples and they are also significantly different ($p<0.05$). These variations are shown in a clearer manner with a box and whisker plot in Figure 4.29.

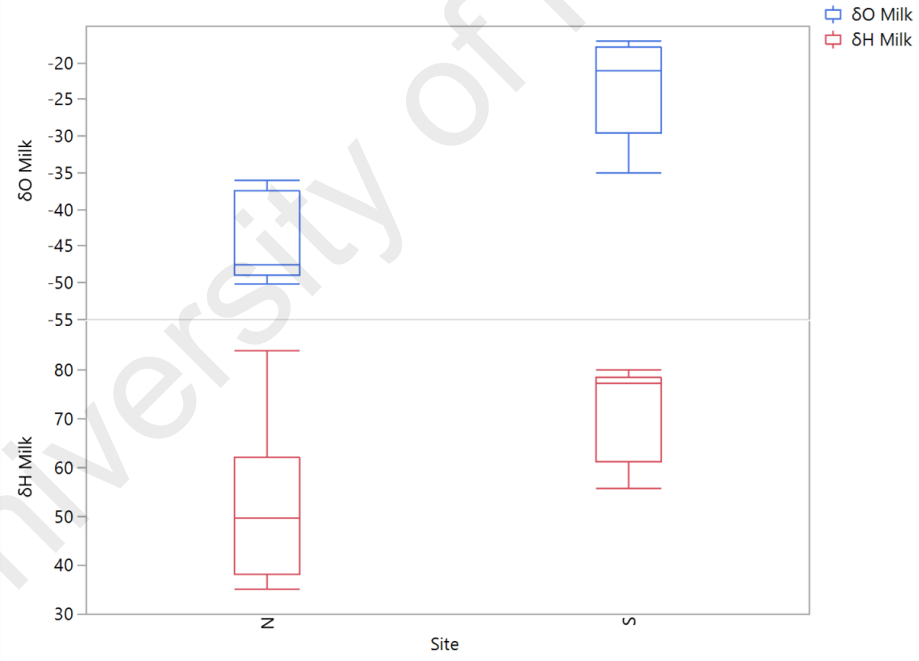

Figure 4.29: Boxplot of δH_{Milk} and δO_{Milk} in southern and northern sampling sites of Peninsular Malaysia.

From the correlations observed, it can be thought that the values of $\delta^{18}O$ and δ^2H in milk are related to the isotopic ratio of ground water and the amount of fresh grass consumed by the cattle. As plants are watered by precipitation and animal drinking

water is mostly precipitation water, it is important to investigate the isotopic ratio of precipitation. Since cattle consume much water, the isotopic ratio of O and H reflects that of drinking water (Magdas *et al*., 2016) and that of precipitation as the source of ground water is precipitation. Since oxygen and hydrogen of plants and water are transferred to the tissue (Rees et al., 2016), it could be assumed that the latitude effect and amount of precipitation could be observed in the milk samples. In grass the content of O and H isotopes is related to water evaporation. Therefore, when cows are fed exclusively by fresh grass O and H values may be seen to affect the isotopic ratios of hydrogen and oxygen in milk (Magdas & Puscas, 2011).

Due to reasons mentioned previously, the depleted values of $\delta^{18}O$ and $\delta^{2}H$ in milk samples from the northern region of the country indicates that the cattle in these part of the country are mostly fed with hay, maize and siladge. The higher values of $\delta^{18}O$ and $\delta^{2}H$ in the milk samples from the southern region of the country indicate that cows mostly graze in these regions (Magdas *et al*.; Renou *et al*., 2004). The values of $\delta^{18}O$ are also observed to be less negative when rain is abundant.

4.2.8 Isotopic ratio in precipitation samples

The box and whisker plots of δH_{rain} and δO_{rain} in the southern and northern sampling sites of Peninsular Malaysia are shown in Figure 4.30. The range of isotopic ratio of $\delta^{18}O$ in rain is -1.12‰ to -7.97‰. The southern part of Peninsular Malaysia has ratio ranging from 3.93‰ to -7.97‰ while in the northern region they are in the range of -1.12‰ to -6.63‰ and they are significantly different ($p<0.05$). $\delta^{2}H$ in rain ranges from -3.66‰ to -53.91‰. The range for southern sampling regions varies between -38.92‰ to -53.91‰. Samples from the south are located in the lower right hand side of the plot and have more negative (lower) values than the northern samples.

4.2.9 Factorial analysis

Factorial analysis as a multivariate method has been used to cluster milk samples based on their geographical origin. The database of milk samples constitutes the learning and training set. The two dimensional plot obtained from the analysis shown in Figure 4.31 explains a total variance of 89.5%. The descriptors used in the analysis are $\delta^{18}O$, $\delta^{2}H$, $\delta^{13}C$ and $\delta^{15}N$ in milk and $\delta^{18}O$, $\delta^{2}H$ in precipitation. The plot shows the highest degree of separation for the samples based on their geographical origin which is illustrated by factor 1 vs factor 2 plot obtained from the analysis. Factor 1 explains 51.9% of the total variance and factor 2, 37.6% of the total variance. From the statistical information provided by this analysis it can be concluded that the most important parameters responsible for factor 1 are δN_{milk} and δH_{rain} which are positively related to factor 1 and δC_{milk} and δO_{milk} which are negatively correlated with factor 1. The most important parameters influencing factor 2 are the δO_{rain} and δH_{rain} which are positively correlated and δH_{milk} which is negatively correlated.

4.2.10 PCA Biplot

In Figure 4.32, principal component analysis (PCA) has been applied to 69 raw cow milk samples with a matrix of 3 analytical parameters. 2 PCs were extracted, explaining 97% of the total variance of 97% that reveals clustering of raw cow milk in Peninsular Malaysia. Separation of samples based on their geographical origin is observed as two main clusters of north and south. Southern samples are placed in right hand side of the PCA plot and have positive PC1 as well as positive and negative PC2 scores. In this plot, the loadings of the isotopic ratio of C and N are in the right hand side of the quadrant indicating that C and O are the dominating factors for southern regions. On the other hand, samples that have negative PC1 scores and negative and positive PC2 scores are placed at the left hand side of quadrant have nitrogen as their sole loading factor which acts as the discriminating factor for this region.

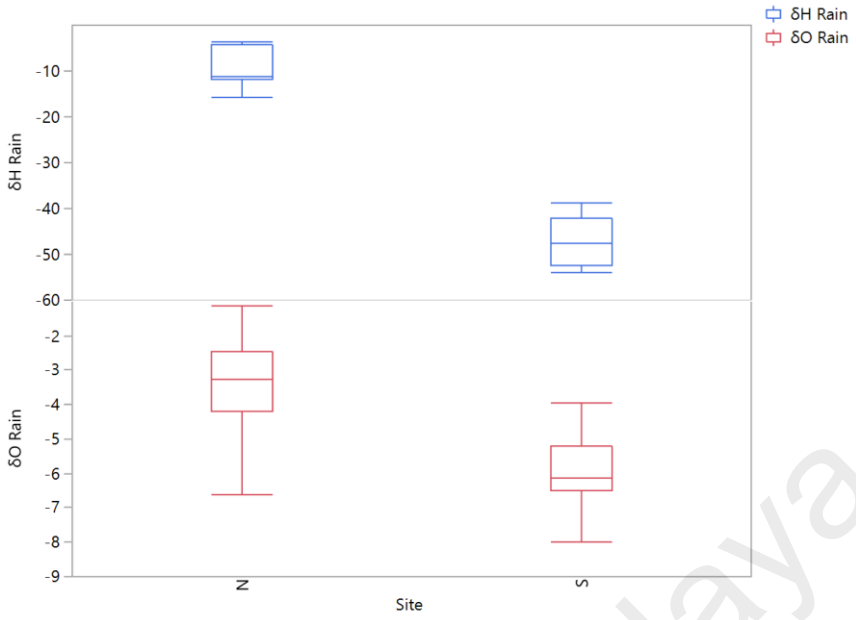

Figure 4.30: Boxplot of δH_{rain} and δO_{rain} in the southern and northern sampling sites of Peninsular Malaysia**.**

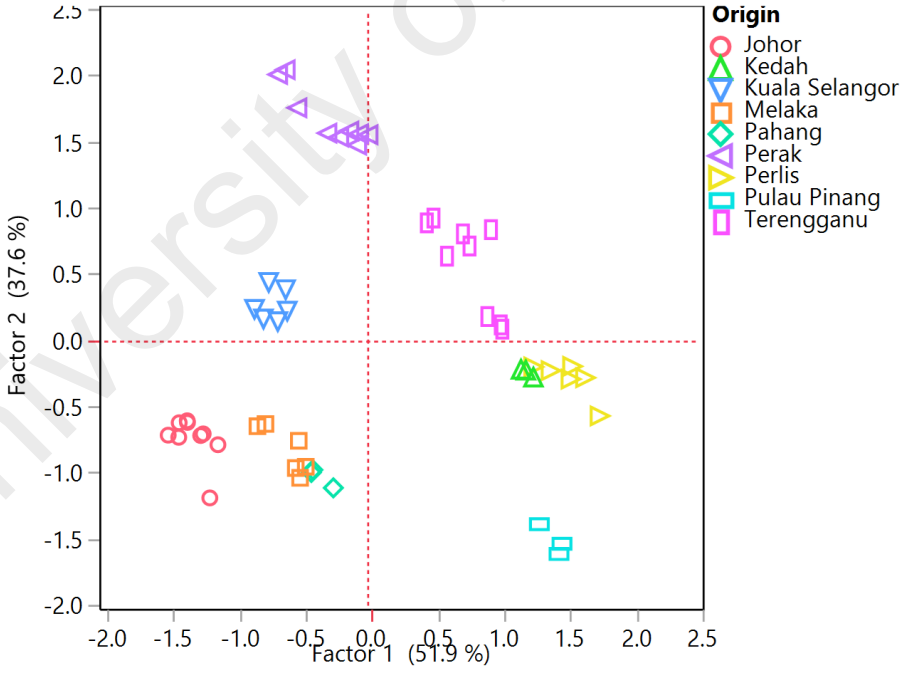

Figure 4.31: Factor plot of geographical origin of milk based on $\delta^{18}O$, $\delta^{2}H$, $\delta^{13}C$ and $\delta^{15}N$ in milk and $\delta^{18}O$, $\delta^{2}H$ in precipitation descriptors.

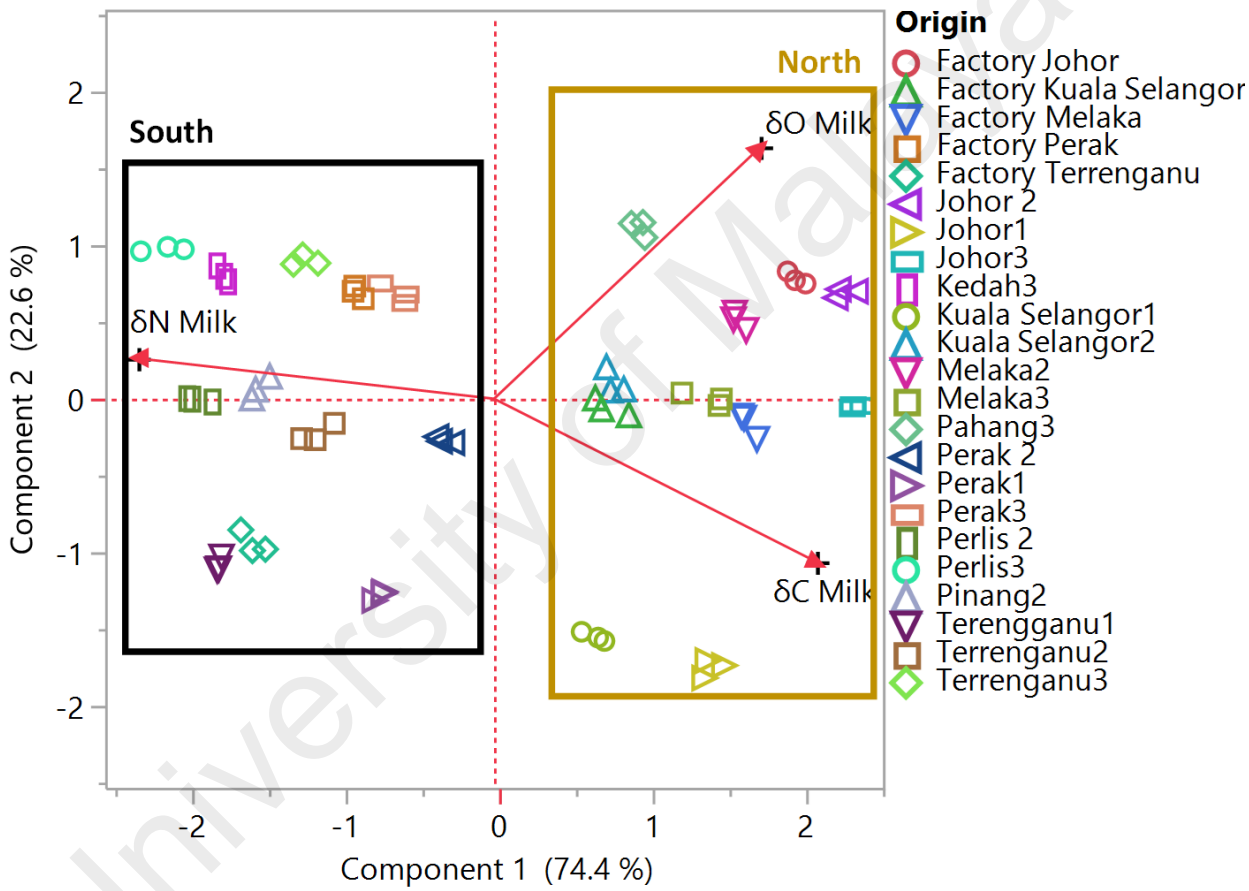

Figure 4.32: PCA biplot of milk samples using isotopic ratios of C, N and O.

4.2.11 Hierarchical cluster analysis

In Figure 4.33, the hierarchical graph for 23 samples is plotted. Two clusters of north and south are also observed from this plot.

In this analysis samples are colored based on their collection time. Light blue color (S1) shows samples from the northeast monsoon wind season which is from Nov-March where rainfall is high. Samples in purple (S3) and green (S2) were collected during southern monsoon wind season from May-Sep. Samples in the southern cluster were collected from farms in Melaka, Kuala Selangor and Johor together with samples from factories of these regions.

S1 samples are seen to cluster in one branch as their isotopic ratios for oxygen are more negative compared to other samples which is probably due to the high rainfall during this sampling time. Samples in green (S2) and purple (S3) can be separated as the sampling time follows the monsoon wind season of May-Sep. Samples from Terengganu, Perlis, Perak, Kedah and Pulau Pinang, on the other hand are grouped in the northern cluster. From the plot it is observed that S1 milk samples are neatly separated from milks in other sampling period most probably due to the higher amount of rain during this season. Moreover, isotopic ratio of oxygen is more negative compared to S2 and S3. To confirm the results obtained from the analysis a few factory milk samples have also been analyzed to see if they are well separated based on their individual regions and clustered the same as their raw cow milk samples. Looking to the clustering of S2 and S3 it is noted that these two sampling seasons are mixed in branches, but this is expected to be because the samples are taken within the same season. Their isotopic ratios for oxygen are also in the same range with only some small variations.

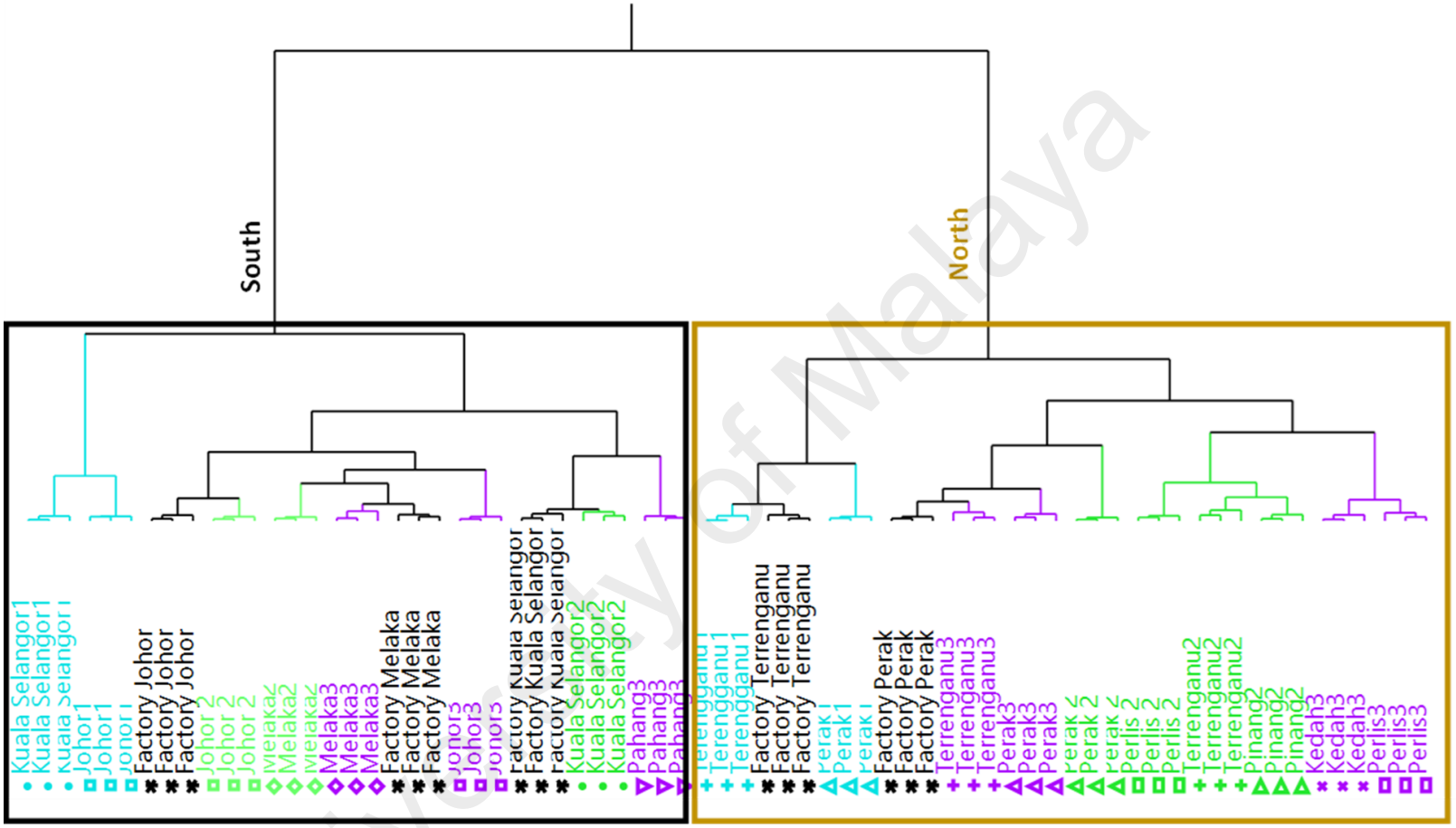

Figure 4.33: Hierarchical cluster analysis of milk samples using isotopic ratio data of C, N and O.

From the observed results, it can be concluded that isotopic ratio could be considered as one of the good discriminating factors differentiating the two sampling season of a tropical area such as Malaysia.

Figure 4.34 presents the cluster plot that resulted from a combination of HCA and heat plots. A clear separation of north and south is observed. Samples from South are marked as S and the ones from the north as N. Each legend shows the variation of each isotopic ratio and how they are distributed in each sampling region. Each cell of the heat map represents one isotopic ratio. The more intense is the cell in red, the higher it the isotopic ratio of the particular element in the region. The more intense the cell gets in blue, the less is the isotopic ratio of the particular element in the region.

4.2.12 Boxplot of samples from the northern and southern regions of Peninsular Malaysia

Boxplot of the mean values of $\delta^{13}C$, $\delta^{15}N$ and $\delta^{18}O$ for northern and southern milk samples have been plotted in Figure 4.35. Due to the fact that other methods gave a separation of samples based on northern and southern sampling regions of the country there is an attempt to ascertain the differences between the samples from these two sampling regions based on means.

After plotting the box and whisker plot, data of descriptive statistics namely median, 1st quartile, 3rd quartile, min and max of the milk samples are reported in Table 4.13. From the $\delta^{13}C$ values reported in the Table, it is noted that there are considerable differences between northern and southern isotopic ratio values for carbon and the calculated p-value is significant ($p<0.05$). This shows that ascertaining the geographical origin of milk using $\delta^{13}C$ data is significant.

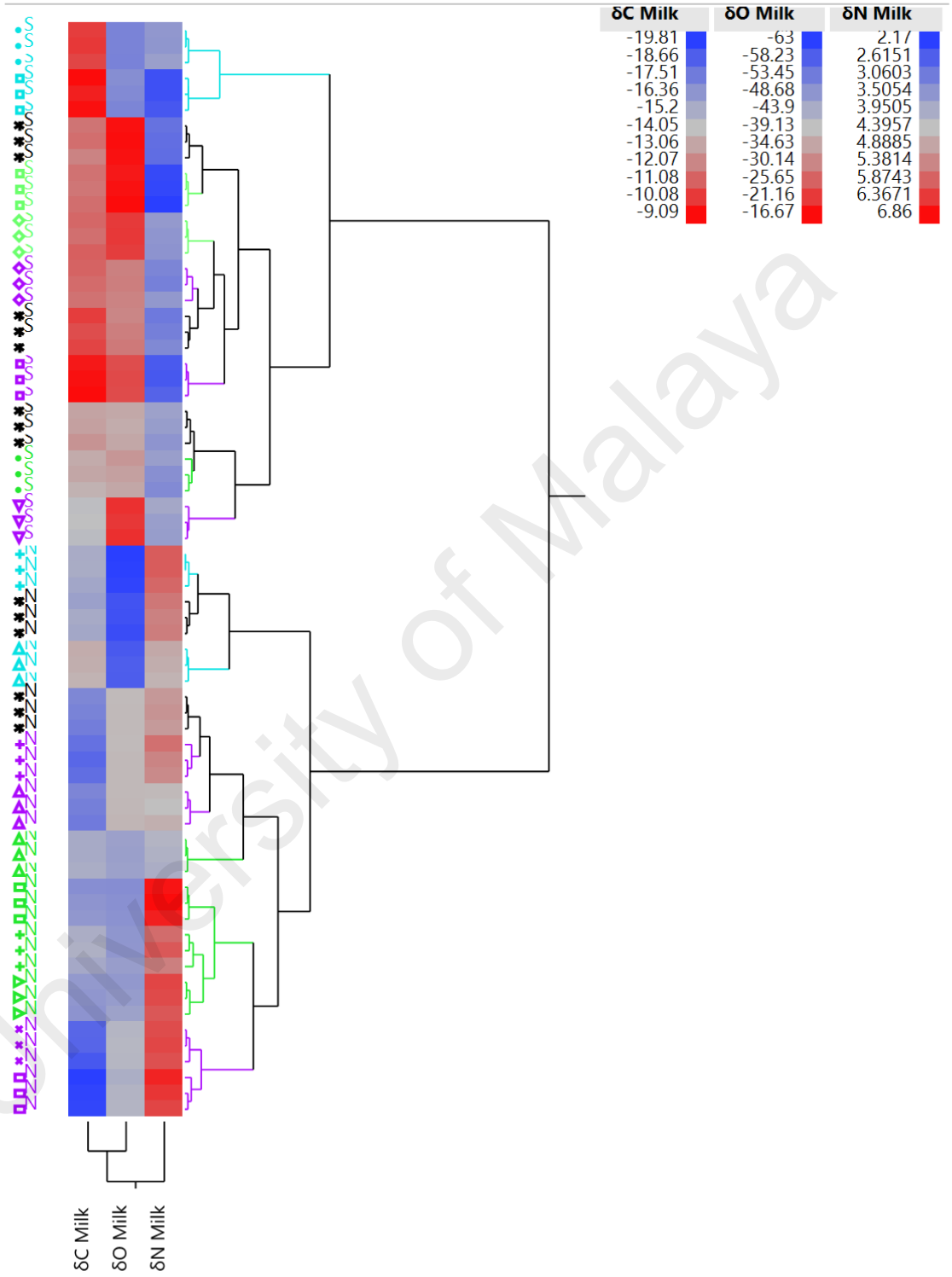

Figure 4.34: Two way analysis of isotopic ratio analysis.

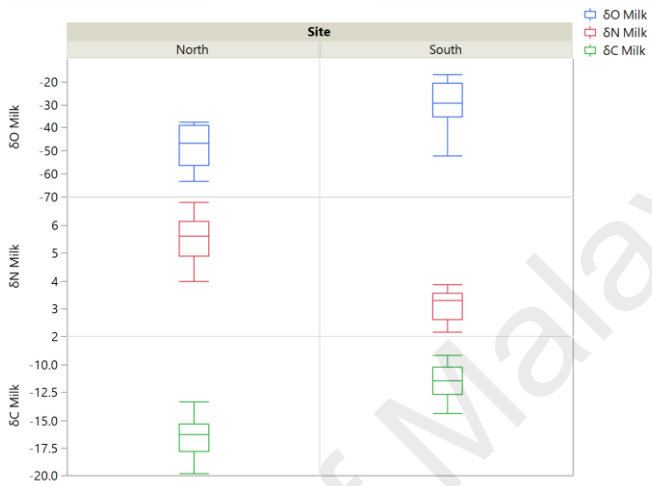

Figure 4.35: Box plot of milk samples in northern and southern regions of Peninsular Malaysia using isotopic ratio data of C, N and O.

Table 4.13: 1st quartile, 3rd quartile, min and max of elemental isotopic ratios of milk samples.

Isotopic ratio (‰)	Geographical origin									
	Northern states of Malaysia (n=36)						Southern states of Malaysia (n=32)			
	Median	Mean	Min	Max	1st quartile	3rd quartile	Median	Mean	Min	Max
$\delta^{13}C$	-16.32	-16.51	-19.81	13.29	-17.81	-15.28	-11.39	-11.36	-14.36	-9.09
$\delta^{15}N$	5.64	5.55	3.98	6.86	4.90	6.16	3.29	3.14	2.17	3.86
$\delta^{18}O$	-46.49	-47.51	-63	-37.26	-56.25	-38.82	-29.13	-29.98	-52.15	-16.67

4.3 Data analysis of Milko tester results

Beside milk samples from different origins, milk from different breed of cows namely Sahiwal, Thai Friesian, Australian Friesian, Friesian Sahiwal, Lamborghini Australia and Local Indian were analyzed by the milko analyzer.

4.3.1 Physical and nutritional information of raw cow milk samples

Raw cow milk samples from different regions have been analyzed for their physical properties of freezing point, solid nonfat particles and density as well as nutritional information of fat, protein, lipid, lactose and salt. These information have been presented in pie charts in Figures 4.36 and 4.37. From Figure 4.36 it is observed that the order of percentage of fat in raw cow milk from different regions follows the order of Pahang > Perak = Kuala Selangor > Pulau Pinang = Kedah > Terengganu > Perlis = Melaka > Johor. In other words Pahang milk has the highest percentage of fat while the lowest percentage of fat is found in Johor milk and this percentage for other regions follows the order mentioned before. Percentage of salt decreases in the order of Perlis

>Terengganu>Perak>Pahang=PulauPinang=Kedah>KualaSelangor>Melaka>Johor,

which means that Perlis milk has the highest percentage of salt and Johor has the lowest and for all the other states it varies as mentioned before. Moreover, the percentage of lactose in milk from different regions under this research follows Kuala Selangor >Pahang>Pinang>Johor=Melaka=Kedah>Perlis=Terengganu=Perak.

Milk of Kuala Selangor has the highest percentage of lactose whereas Perlis, Terengganu and Perak being equal in amount have the lowest percentage of lactose and all the other states follow the order mentioned. The percentage of protein on the other hand, follows Kuala Selangor>Pahang > Johor = Pulau Pinang = Kedah > Perak = Melaka > Terengganu = Perlis.

In other words, the highest percentage of protein belongs to the milks collected from Kuala Selangor while the lowest percentage is for Terengganu and Perlis which have approximately equal percentage of protein.

Figure 4.36: Pie chart for nutritional information of raw cow milk samples from different regions.

Figure 4.37 presents the measured physical parameters in this research. The percentage of Solid nonfat particles (SNF) follows KualaSelangor > Pahang > Pulau Pinang > Perlis = Terengganu = Perak = Melaka > Kedah > Johor. Kuala Selangor has

the highest percentage of solid nonfat particles whereas Johor has the lowest of this. Moreover, the size of nonfat particles in milk varies in milk from different regions as mentioned before.

Freezing point is one of the other physical parameters measured for milk samples in this study and it follows the order of Perak> Terengganu> Johor> Perlis> Kedah> Pahang> Melaka> Pulau Pinang> Kuala Selangor.

Density is another physical parameter measured in raw cow milk samples from the different regions. Milk from Kuala Selangor is denser compared to Melaka and Johor which have the same density. The overall density of milk samples follows the order below; Kuala Selangor > Pahang > Pulau Pinang > Perak > Terengganu > Perlis > Kedah > Melaka = Johor.

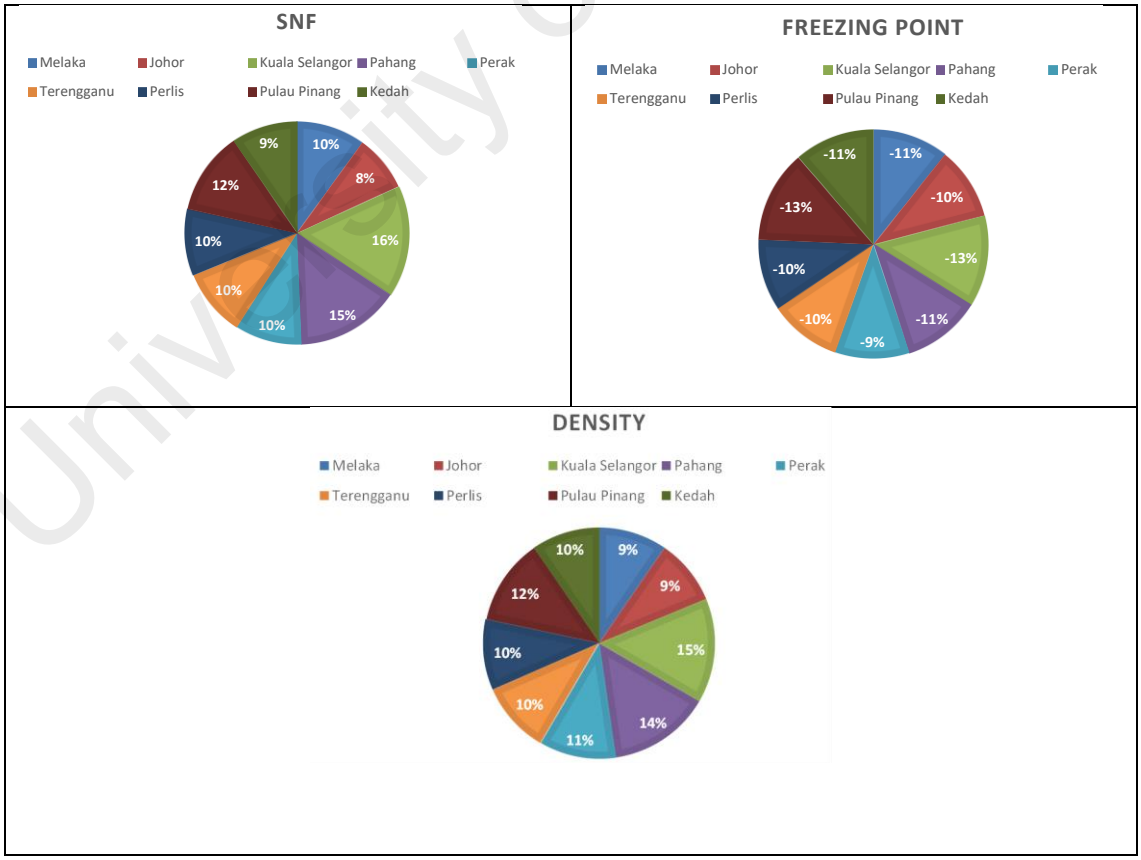

Figure 4.37: Pie chart of physical properties of raw cow milk.

4.3.2 Physical and nutritional information of different breeds of cow

Fat, protein, lactose and salt are the factors measured in each of the milk samples. As observed in Figure 4.38, Australian Friesian, Thai Friesian and Friesian Sahiwal are the breeds that are loaded with protein, lactose, salt and solid nonfat particles. These breeds are located at the right hand side of the graph and are separated by other breeds based on positive PC1 scores. On the other hand, local Indian breed placed at the lower left hand side of quadrant and has negative PC1 and PC2 are loaded by fat. Jersey Friesian and Lamborghini Australian are located at the upper left hand side of the quadrant with negative PC1 and positive PC2. They are loaded by freezing point.

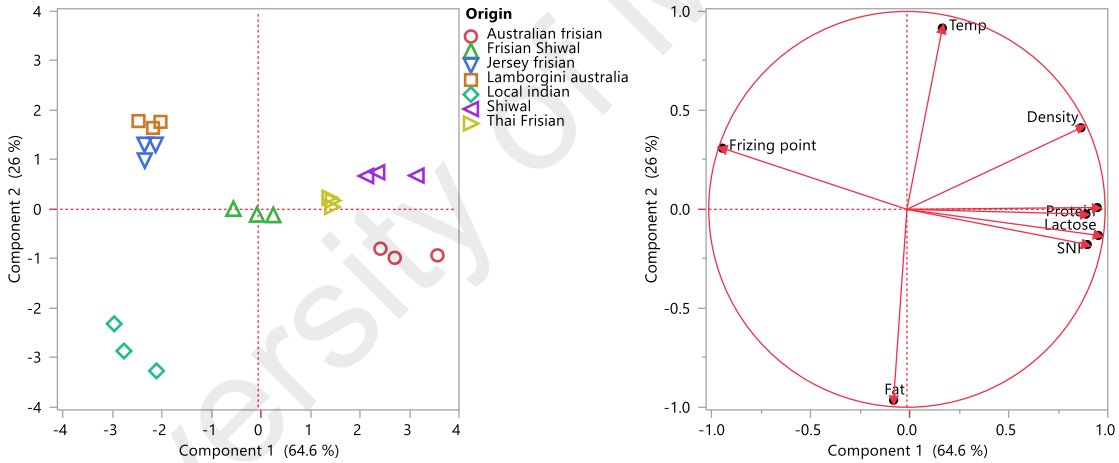

Figure 4.38: PCA scores and loading plots of milk samples from different breeds.

4.4 Data analysis of spectroscopic results

Spectroscopic data in this research are grouped into three types based on the different instruments used. Freeze dried milk samples were analyzed using a Shimadzu UV-3600 (Japan) and a single photodiode Arcoptix FTIR-NIR (Switzerland) while plasma samples of milk were analyzed using a Shimadzu UV-2600 (Japan).

4.4.1 The chemistry of milk (milk constituents)

Milk major composition consists of water, fat, protein, carbohydrate (lactose) (Lei *et al*., 2010) and small amount of minerals, vitamins and enzymes (Huang *et al*., 2016). It has been mentioned that decrease in the content of milk fat could be related to imbalanced forage in the animal diet (Hawke & Taylor, 1983). However, the protein content of milk could changes during lactation stage and when stomach cell account changes whereas lactose is the most stable component present in milk (Tsenkova *et al*., 1999). Moreover, milk fat is a mixture of fatty acid esters called triglycerides being composed of an alcohol named glycerol and various fatty acids. A molecule of fatty acid has a hydrocarbon chain and a carboxyl group (RCOOH). Milk consists of hundred type of proteins which are made up of amino acids. Most important of them being casein family that contain phosphorous and the serum (whey) proteins which do not contain phosphorous but do contain large amount of sulfur containing amino acids. Each amino acid has a carboxylic acid group on one end of the molecule and an amin group on the other end. Milk contains carbohydrates which are composed of carbon, hydrogen and oxygen. The most important carbohydrate in milk is lactose. Lactose is a disaccharide of glucose and galactose. Functional groups for carbohydrates are hydroxyl and carbonyl group.

In NIR spectra mostly three essential bonds of C-H (fats, oil, hydrocarbons), O-H (water, alcohol) and N-H (protein, amines, amides) were observed (Borin *et al*., 2006). Some of the assignment of bands in the UV/NIR region are reported in Table 4.14

Table 4.14: Assignment of bands in UV/NIR region **:**(Bernuy *et al.,* 2008; Forcato *et al*., 2005; Raty & Peiponen, 1999)

Wavelength(nm)	Probable band assignment	component in milk
200-215	Isolated double bonds	fat

220-240	naturally accruing conjugated dienes	fat
290		protein

Table 4.14 continued: (Feng *et al*., 2013); (Osborne & Fearn, 1986; Šašic & Ozaki, 2001; Workman Jr, 1996); (Osborne & Fearn, 1986; Workman Jr, 1996)

Wavelength (nm)	Probable band assignment	components in milk
748,757	C-H and O-H stretching	Protein, fat, lactose and water
928	3rd overtone of C-H stretching	fat
1042	C-H stretching and C-H deformation of fat	fat

Table 4.14 continued: (Osborne & Fearn, 1986; Šašic & Ozaki, 2001)

Wavelength (nm)	Probable band assignment	components in milk
906	3rd overtone of C-H stretching	protein
968	2nd overtone of O-H stretching of water	water
1020	N-H stretching, amide 1	protein
1030	2nd overtone of N-H stretching	protein

Table 4.14 continued: (Feng *et al*., 2013)

Wavelength (nm)	Probable band assignment	components in milk
872	C-H stretching and C-H bending vibration	Fat, lactose
908	3rd overtone of C-H stretching	Fat, lactose
920	Combination and overtone of O-H stretching	water
927	3rd overtone of C-H stretching	fat
937	3rd overtone of C-H stretching	Protein, fat
950	2nd overtone of O-H stretching	Protein, lactose and water
960	2nd overtone of O-H stretching	Lactose, water
966	2nd overtone of O-H stretching	Protein, water
990	2nd overtone of O-H stretching	fat
1017	C-H stretching	Fat, lactose
1046,1054	C-H stretching and deformation	Protein, lactose
1064	C-H stretching and deformation	Fat, lactose

Table 4.14 continued: (Šašic & Ozaki, 2001); (Brandã et al., 2010); (Berzaghi *et al*., 2005); (Laporte & Paquin, 1999) ; (Núñez-Sánchez *et al*., 2016); (Westad *et al*., 2008)

Wavelength (nm)	Probable band assignment	Components in milk
840	C-H stretching and C-H bending vibration	fat
880-890	3rd overtone of C-H stretching	fat
996	2nd overtone of O-H stretching	water
1018	C-H stretching and deformation of fat	fat
1208	2nd overtone CH_2 stretching	fat
1220	2nd overtone CH_2 stretching	fat
1235	Amide 2	protein
1290	Water protein interaction	water
1400	Combination mode, Water interaction with fats	water
1410	Combination mode Weak water interaction with protein	water
1450	O-H 1st over tone, water combination mode $\nu_1+\nu_3$	Water band of fat, protein and lactose
1520	N-H stretch 1st overtone	Water band of protein
1720,1722	1st overtone of CH_2 antisymmetric stretching	Fat, protein
1726,1730,1760,1780	1st overtone from CH stretching vibration of methyl(-CH_3), methylene (-CH_2) and ethenyl (-CH=CH–)	Protein, fat
1754	1st overtone of CH_2 symmetric stretching	fat
1950	Combination of water and -CH=CH–	Water band of protein
2056	Combination of amide I modes and amide A ,NH stretching vibration of various type	protein
2060	-NH related to protein	protein
2100	-CH=CH–	Lactose, protein
2160	-NH related to protein, combination of amide **II** modes and amide B	protein
2174	N-H 2nd overtone, C-H stretch/C=O stretch combination, C=O stretch/N-H amid combination	protein
2302	Combination of CH_2 asymmetric stretching and bending vibration	fat
2310, 2348	2nd overtone, Combination of C-H stretching and bending vibrations of methyl and methylene functional group	Fat, protein
2340	Combination of CH_2 symmetric stretching and bending vibration	fat, protein, lactose

2314,2368	Combination of CH_2 stretching and bending modes of \protein side chain	Protein, fat
2468	C-N-C stretching 1st overtone	protein

4.4.2 Milk spectra obtained from Shimadzu UV-3600

The total spectra of 60 solid milk samples is plotted in the range of 200 to 2500 nm Figure 4.39. It can be visually observed that there seems to be a separation of peaks between the UV/Vis and NIR region. Based on this separation, data analysis are carried out.

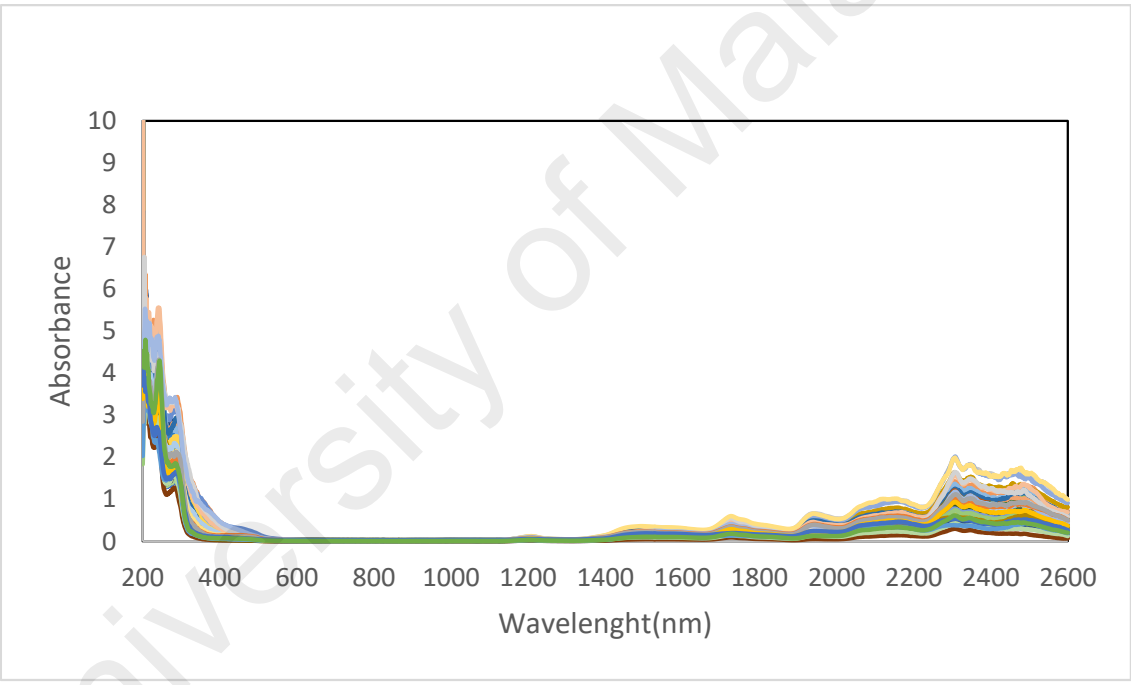

Figure 4.39: UV/Vis/NIR spectra of all milk samples.

Based on the observations obtained from the results of ICP-MS and IRMS discussed in the earlier sections of this chapter, the milk samples are clustered into two groups of north and south. Two sample student t-test was applied to compare the two means obtained in the spectral range of 200 nm to 2500 nm. It is observed that there is no significant difference between the means of the northern and southern sampling regions. Henceforth, the whole spectrum was divided into three spectroscopic range of UV, Vis and NIR regions as indicated in Figures 4.40 – 4.42. The mean test comparison

of two means tests were applied to each region individually and it is observed that the p value was significantly different for UV and Vis regions for the northern and southern milk samples. In the case of NIR there is a separation in the spectrum into two ranges, one of which is 800 nm-1500 nm and the other 1500 nm to 2500 nm. The two sample student t-test comparison of two means tests were applied separately to each one of the regions and it is noted that there is significant difference only in the second half of NIR region.

4.4.2.1 Spectra of the northern and southern samples in the UV/Vis range

Figure 4.40 shows the average spectra of milk samples in the UV range. From the spectra it is noted that there is a significant difference ($p<0.05$) in the average spectra of the northern and southern milk samples which makes it worthwhile to analyse the data further with methods that will be able to separately cluster these two sampling groups.

Figure 4.41 presents the average spectra of samples from the north and south in the visible range. It is also noted that there is a significant difference in the average spectrum of milk sample from the north and south. Visually, this can be clearly observed in the first half of the range between 400 to 565 nm.

4.4.2.2 Spectra of milk from northern and southern sampling regions in NIR region

The average spectra of the northern and southern milk samples illustrated in Figure 4.42, seems to have a separation in this region as well although the more obvious separation is observed in the range of 2000 to 2500 nm.

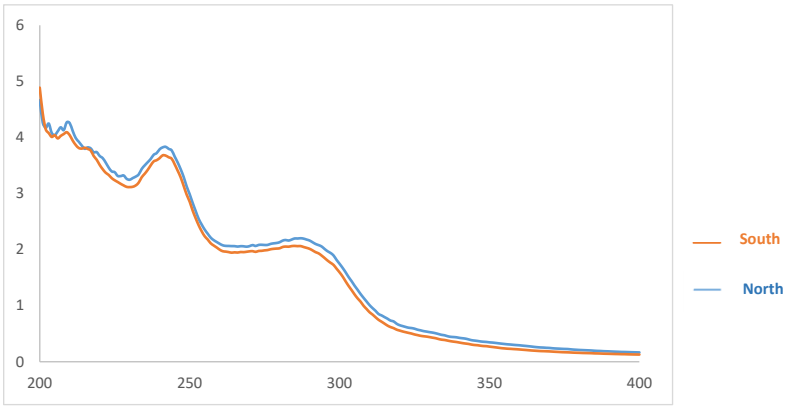

Figure 4.40: Average spectra of milk samples from the north and south of Peninsular Malaysia in the UV range.

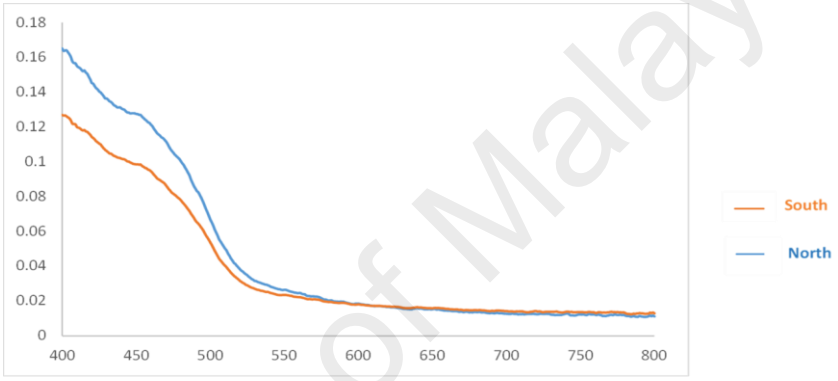

Figure 4.41: Average spectra of milk samples from the north and south of Peninsular Malaysia in the visible range.

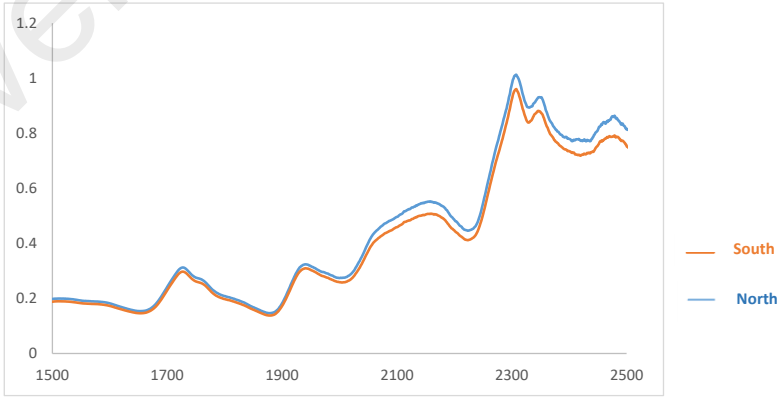

Figure 4.42: Average spectra of milk samples from the north and south of Peninsular Malaysia in the NIR range.

Based on the visual separation observed on the spectra, the data is seen naturally to be separated into northern and southern regions. In order to see whether the conclusion from the visual inspection is accurate, PCA was carried out.

4.4.2.3 PCA

A matrix of 2400 columns and 60 rows was used for the chemometric analysis. PCA was applied to the n-dimensional spectral data set in order to obtain linear combination of original variables called PCs which are uncorrelated. As shown in Figure 4.43 the largest variability is expressed by PC1 and each of the other PC's residual variances.

At first attempt PCA was applied to all of the samples in the wavelength range of 200 nm to 2500 nm. However, it is observed from PCA that the northern and southern samples are not linearly separable. Hence, it is thus reasonable to break the spectral range into UV, Vis and NIR range.

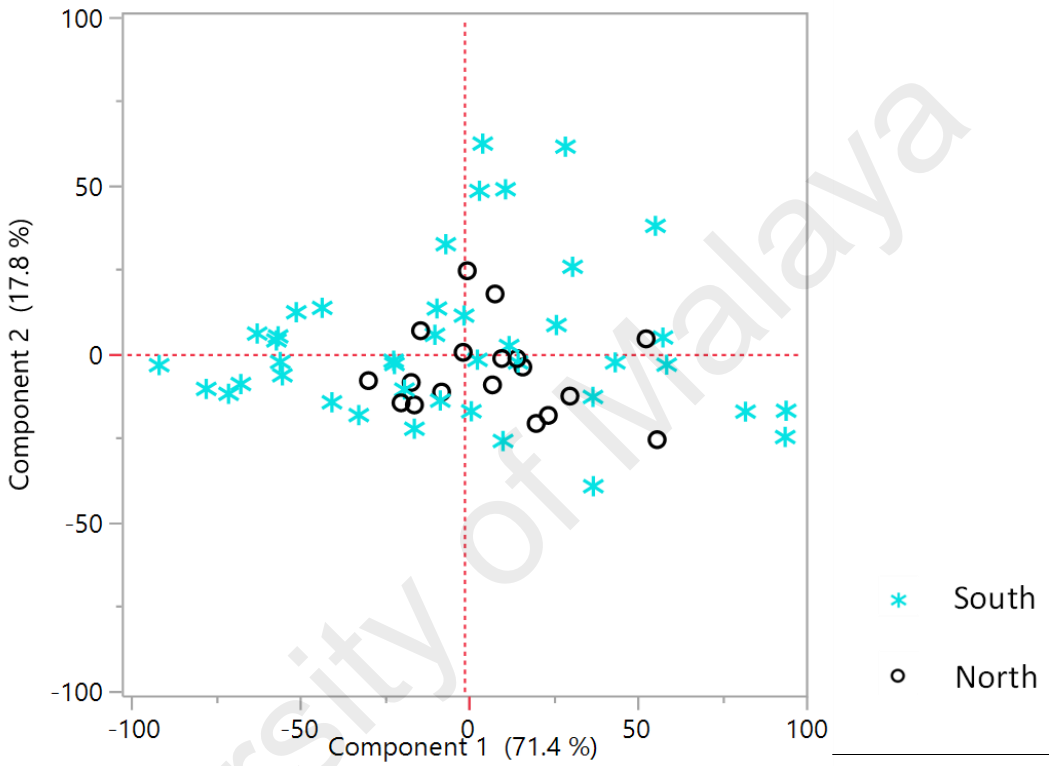

Figure 4.43: PCA of all milk samples analyzed by Shimadzu UV-3600.

Based on the PCA of the whole spectral range, 14 PC's explained 89.7% of total variance. The loading of each variable in the extracted PC are shown for UV, Vis and NIR regions are shown in Figures 4.44 to 4.46. Figure 4.44 shows the PC loading vs. wavelength plot in the UV range of 200-400 nm. As can be seen, almost all wavelengths of the whole range loads PC1 and PC3 positively. PC1 explains 71.9% of the total variance and PC3 explains 6.29 % of the total variance. The other principal components have lower or negative loadings within the whole range compared to PC1 and PC3.

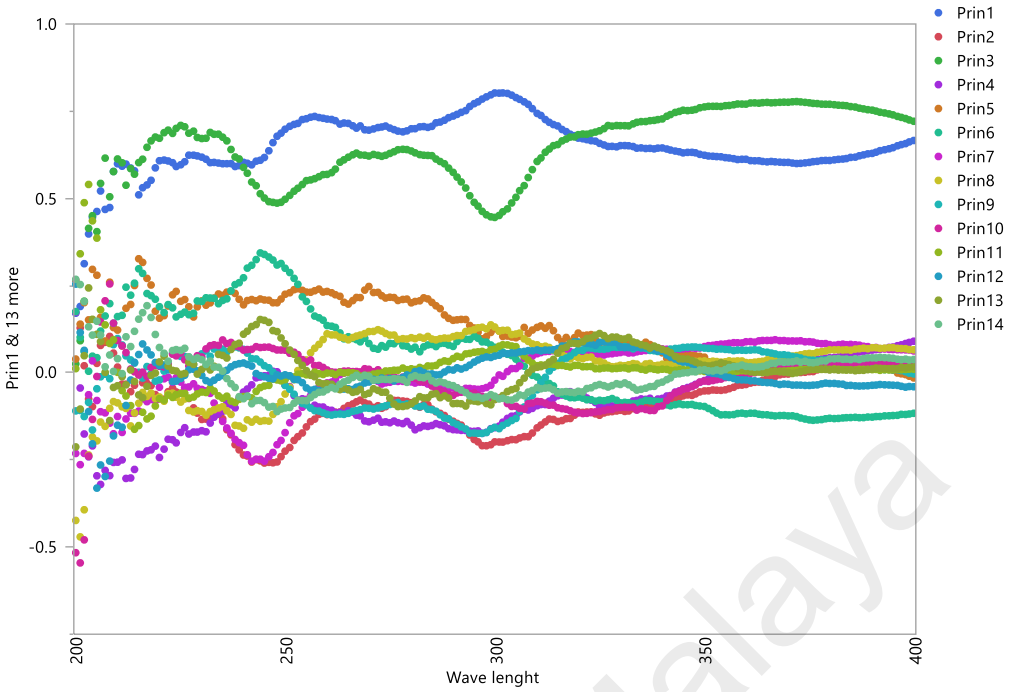

Figure 4.44: Loading plot of PCs vs wavelength in the wavelength range of UV.

Figure 4.45 shows that PC1 is loaded most within the 400-634 nm range while PC2 is loaded by wavelength absorptions in the range of 664-800 nm.

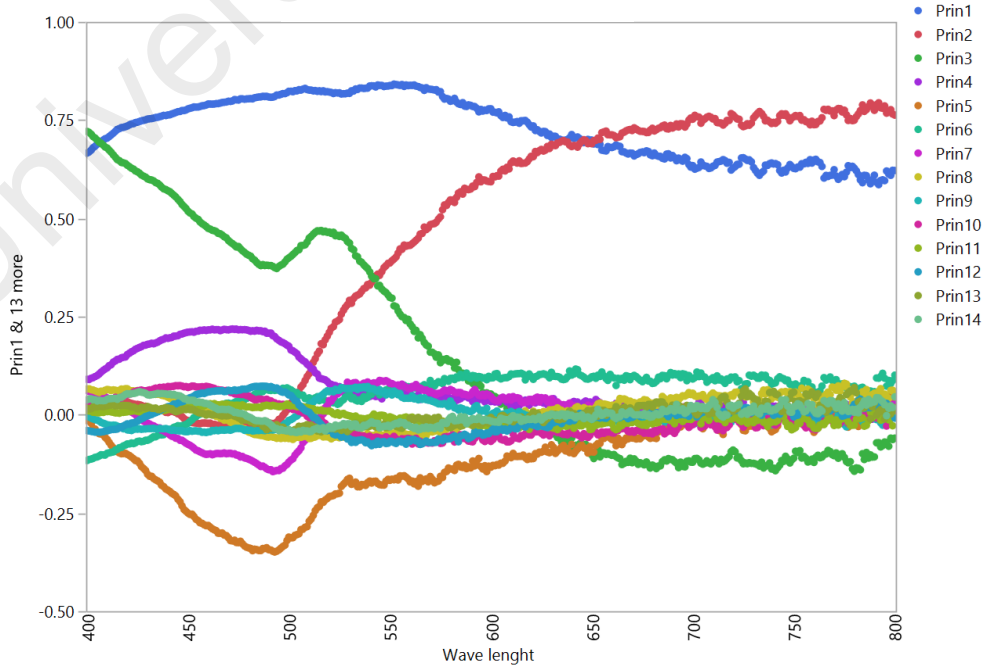

Figure 4.45: Loading plot of PCs vs wavelength in the range of Vis.

From Figure 4.46 it is noted that PC1 seems to be the only component which is loaded by the whole spectral range from 1550 nm to 2500 nm.

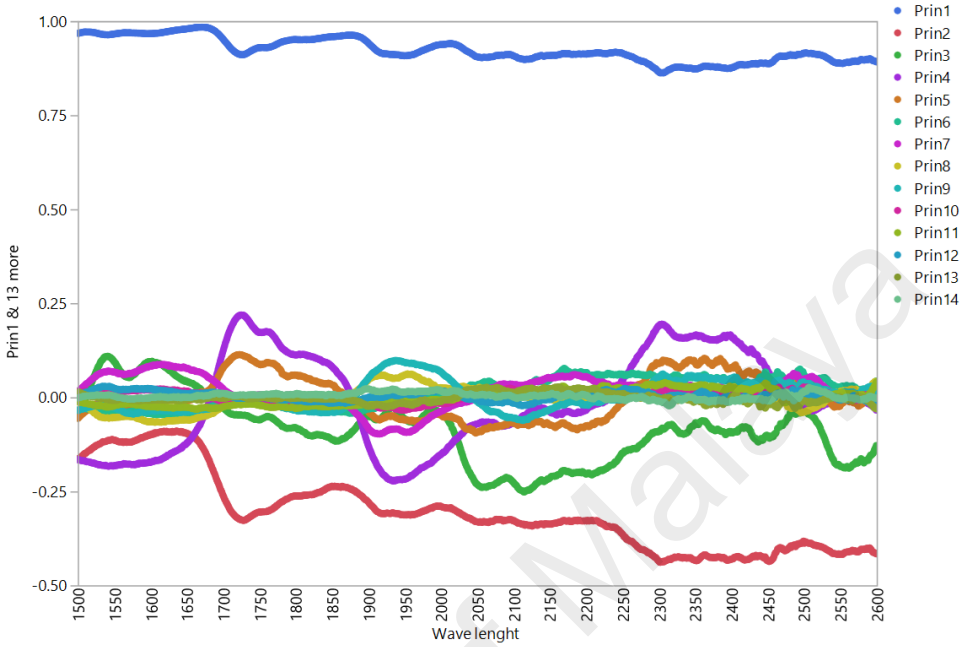

Figure 4.46: Loading plot of PCs vs wavelength in the NIR range.

However, in this study, for a clearer view, the spectral range has been split into three regions. Figures 4.40 to 4.42 respectively shows the average absorption spectra of the northern and southern milk samples in the ranges of UV (200-400) nm, visible (400-800) nm and NIR (1500-2600) nm. Visual inspection of the 800-1500 nm range does not indicate significant variations. Figure 4.40 presents the UV range, that has sharp peaks at wavelengths of 201, 202,209, 210, 218, 226, 242, 289, 290, 404 and 451 nm which indicate the presence of protein, whereas in the NIR region of this study, wavelengths of 1209, 1497, 1724, 1942, 2069, 2156, 2308, 2352 and 2483 nm are dominant and showed the presence of fat, protein and lactose.

From the PCA loadings depicted in Figures 4.44 to 4.46, PC1 which represents 71.9% variation in the broad spectral range can be said to be loaded with all the

important fat, lactose and protein bands of milk from 200 nm to 2500 nm. PC1 dominates the loading over the spectral range of 1550 nm to 2500 nm. These bands are the overtones of vibrational activities mainly due to the protein and fat content of milk.

4.4.2.4 Artificial neural network (ANN) applied to the spectral data of the Shimadzu UV-3600

In constructing a classification model for raw cow milk samples from the various regions of Peninsular Malaysia, MLP-ANN was used. In order to determine the appropriate inputs, PCA was applied as a variable reduction method where PC scores were used as input variables for building the models (Palacios-Morillo *et al*., 2014). As had been mentioned before, 14 PCs with eigenvalues >1 accounting for 89.7% of the total variance were retained. The classification performance of the method was determined by statistical information such as generalized R-square, Entropy R-square, RMSE, Mean Abs Dev, misclassification rate and -Log likelihood.

The ANN models were built and trained until the best performing model is obtained. A sigmoid (TanH) transfer function is assigned to each node. In this work, to built the best performing ANN model, the boosting method available in the JMP software was utilized. Boosting involves building a large additive ANN model by fitting a sequence of smaller models. Each of the smaller models is fit on the scaled residuals of the previous model. The models are combined to form the larger final model. The process uses validation to assess how many component models to fit, not exceeding the specified number of models. The boosting method employs the various statistical tests mentioned above to determine the best performing ANN model. The final network will consist of an input layer with 14 PCs, one hidden layer with number of nodes built using the boosting method and a single output node in the output layer representing the clustering of regions (N or S). ANN model was trained using selected parameters from

the data matrix (PCs) and validation was done using the independent validation data set in order to predict the origin of the milk.

To predict a model using the spectral information of the milk samples, several ANN models were examined starting from simplest one. The leaning rate of 0.1 (min) and 1 (max) were compared. It is observed that by using the max learning rate of 1, high R-Square values and higher classification rates were obtained. Different number of input variables were also compared and it was noted that by using three PCs as input variables, it was enough to build a good model. By using 2 PCs, the prediction rate is 100% and the R-square value explained 95% of training and 92% of validation data with a learning rate of 1. When the learning rate was reduced to 0.1, a minimum number of 6 PCs are needed in order that the prediction rate be sustained at 100%. The summary of results for the different architectures of ANN is reported in Table 4.15. Optimizing the neural network models was done by increasing the number of input variables in order to lower the prediction error, the number of input variables and the number of k-folds in the training stage as it influences the learning rate. The errors of the model were estimated as it was criteria in optimization of the neural model. The number of neurons in input layer was increased until the architecture does not lead to a better performance of the model. Input parameters were chosen based on variable reduction method of PCA. Learning rate is the last factor considered for constructing a model. As learning rate have values between 0 and 1 we have compared our models on the bases of min and max learning rates. Training of neural network was done by increasing the number of neurons in the hidden layer until a fit was obtained. For validation purpose, K-fold cross validation was used (Cancilla *et al*., 2016) by dividing the data to K- dataset (in our case K=5) for building k-models with different verification dataset. Therefore, each model is examined with its verification dataset which leads to estimation of each individual sample.

In general, ANN models give satisfactory results in the discrimination of milk samples by their sampling sites. A minimum of 2 hidden nodes gave us the minimum number of nodes where further increase in the number of nodes does not result in improved performance. From these investigations, it is observed that the best performing model (7 hidden nodes) was obtained by setting the JMP Boosting to 2 base models, 10 boosting models and learning rate of 1 as shown in. The networks are defined as [A, B, C, D] where A is the number of input nodes, B is the number of hidden layer, C is the number of hidden nodes and D is the number of output node. Table 4.15 lists the network models tested based on the definition and the statistics of the training and validation of the models. The statistics are listed in parenthesis ( ) representing the generalized R-Square, entropy R-Square, RMSE, Mean Abs Dev and – Log Likelihood respectively.

The architectures of the model ANNs are shown in Figure 4.47 (a, b) before and after reducing the number of input neurons. In general, 20 network models were examined to determine the best architecture. The prediction performance were compared to one another and investigated by misclassification rate and other statistical information as reported in Table 4.15.

4.4.3 Milk spectra obtained from the Arcoptix FT-NIR

The spectrum of each sample was taken in reflectance mode using a single photodiode Arcoptix FTIR-NIR (Switzerland). The single-photodiode was placed in six different locations within the same sample directly upon which the spectrum was taken. Each spectrum was averages from 6 scans with a spectral range of 900-2600nm.

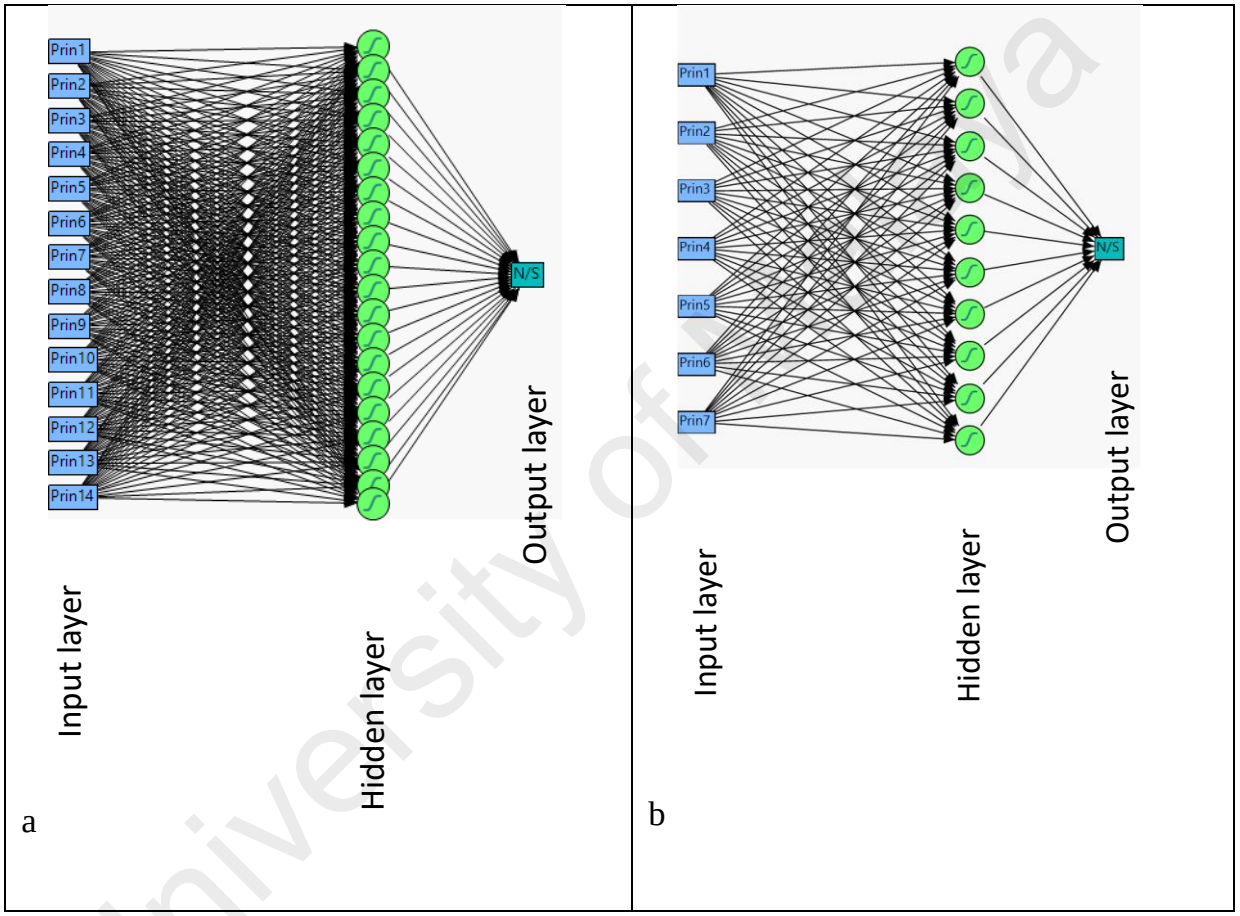

Figure 4.47: The best ANN models using UV/Vis/NIR data a) before reducing the inputs; b) after reducing the inputs.

Table 4.15: ANN models built to predict the best architecture for based on northern and southern sampling regions in Peninsular Malaysia using the Shimadzu UV-3600 data milk samples.

Network	Training	Validation	Learning rate
[14,1,20,1]	(0.9999983, 0.9999967, 3.3823e-6, 1.9685e-6, 9.4487e-5)	(0.9999975, 0.9999949, 4.5153e-6, 3.2171e-6, 0.0000386)	1
[14,1,20,1]	(0.9955997, 0.9914991, 0.0134126, 0.0050388, 0.2463111)	(0.9983436, 0.996661, 0.0065938, 0.0021033, 0.0255039)	0.1
[13,1,14,1]	(0.9999992, 0.9999984, 1.8032e-6, 9.4331e-7, 4.5279e-5)	(0.999999, 0.999998, 2.4678e-6, 1.2582e-6, 0.0000151)	1
[13,1,20,1]	(0.9705571, 0.9447129, 0.0522902, 0.031899, 1.6019196)	(0.9345301, 0.8777671, 0.1124161, 0.0701508, 0.9336358)	0.1
[12,1,10,1]	(0.9999872, 0.9999751, 2.5918e-5, 1.5052e-5, 0.0007225)	(0.999977, 0.9999535, 0.0000426, 2.9627e-5, 0.0003555)	1
[12,1,20,1]	(0.9865628, 0.9743095, 0.0253619, 0.015174, 0.7443727)	(0.9834145, 0.9671905, 0.029332, 0.0204388, 0.2506049)	0.1
[11,1,12,1]	(0.9999939, 0.9999879, 1.554e-5, 7.5435e-6, 0.0003621)	(0.9999848, 0.9999719, 3.0744e-5, 0.0000158, 0.0001897)	1
[11,1,20,1]	(0.998901, 0.9978687, 0.0025594, 0.0012832, 0.0617528)	(0.9988085, 0.9975968, 0.0029748, 0.0015252, 0.018356)	0.1
[10,1,20,1]	(0.9999918, 0.9999842, 0.0000183, 9.5547e-6, 0.0004586)	(0.9999918, 0.9999835, 0.0000193, 1.0524e-5, 0.0001263)	1
[10,1,20,1]	(0.9894901, 0.9798387, 0.0320543, 0.0115997, 0.5841661)	(0.9938748, 0.9877228, 0.0129439, 0.0077294, 0.0937752)	0.1
[9,1,16,1]	(0.9999973, 0.9999948, 8.4387e-6, 3.1415e-6, 0.0001508)	(0.9999946, 0.9999891, 1.1987e-5, 6.9491e-6, 8.339e-5)	1
[9,1,20,1]	(0.9950348, 0.990414, 0.0117582, 0.0057153, 0.2777496)	(0.9933455, 0.9866708, 0.0131469, 0.008396, 0.1018104)	0.1
[8,1,20,1]	(0.989089, 0.97865, 0.0244463, 0.0129469, 0.6364906)	(0.9779288, 0.9601051, 0.0380721, 0.0216657, 0.2692119)	1
[8,1,20,1]	(0.9976232, 0.9953008, 0.0056643, 0.0029024, 0.1400939)	(0.9960397, 0.9927085, 0.0068551, 0.0040765, 0.049203)	0.1
[7,1.10,1]	(0.9999999, 0.9999997, 6.2692e-7, 1.6802e-7, 8.0649e-6)	(0.9999997, 0.9999993, 8.0957e-7, 4.1451e-7, 4.9741e-6)	1
[7,1,20,1]	(0.9944176, 0.9892301, 0.015121, 0.0063817, 0.3120547)	(0.9967735, 0.993509, 0.0089393, 0.0040911, 0.0495797)	0.1

Table 4.15 continued: ANN models built to predict the best architecture for based on northern and southern sampling regions in Peninsular Malaysia using the Shimadzu UV-3600 data milk samples.

Network	Training	Validation	Learning rate
[6,1,16,1]	(0.9999941, 0.9999883, 1.7826e-5, 7.2433e-6, 0.0003477)	(0.9999866, 0.9999753, 2.5017e-5, 1.389e-5, 0.0001667)	1
[6,1,20,1]	(0.9964271, 0.9930909, 0.0141434, 0.0040654, 0.2001896)	(0.9989643, 0.9979106, 0.0031256, 0.001325, 0.0159593)	0.1
[5,1,18,1]	(0.9999942, 0.9999886, 1.8054e-5, 6.8569e-6, 0.0003291)	(0.9999971, 0.9999941, 5.9131e-6, 3.7858e-6, 4.543e-5)	1
[5,1,20,1]	(0.9981026, 0.9962464, 0.0087614, 0.0022919, 0.111902)	(0.9999907, 0.9999812, 0.0000402, 1.1956e-5, 0.0001435)	0.1
[4,1,20,1]	(0.9932946, 0.9868123, 0.0231138, 0.0079061, 0.3931549)	(0.999599, 0.9992589, 0.0013233, 0.0004158, 0.0050007)	1
[4,1,20,1]	(0.0001435, 0.7331663, 0.1921746, 0.1423995, 7.9548858)	(0.892038, 0.8199366, 0.1268502, 0.0919453, 1.2150718)	0.1
[3,1,16,1]	(0.9983588, 0.9968192, 0.0056505, 0.0019039, 0.0921631)	(0.9984608, 0.9968967, 0.9968967, 0.0019678, 0.0237034)	1
[2,1,20,1]	(0.9477477, 0.9042832, 0.110271, 0.049711, 2.7733561)	(0.9168911, 0.8478464, 0.1246853, 0.0875216, 1.1621754)	1

Figure 4.48 illustrates the band profile of the whole spectral range of freeze dried milk samples. Major peaks are seen at 1200 nm (fat) due to the 2nd overtone of –CH2 stretching, 1780 nm (fat and protein), 1800 nm (lactose, casein), 1950 nm (casein), 2310 nm (fat, casein) and 2350 nm (casein, fat).

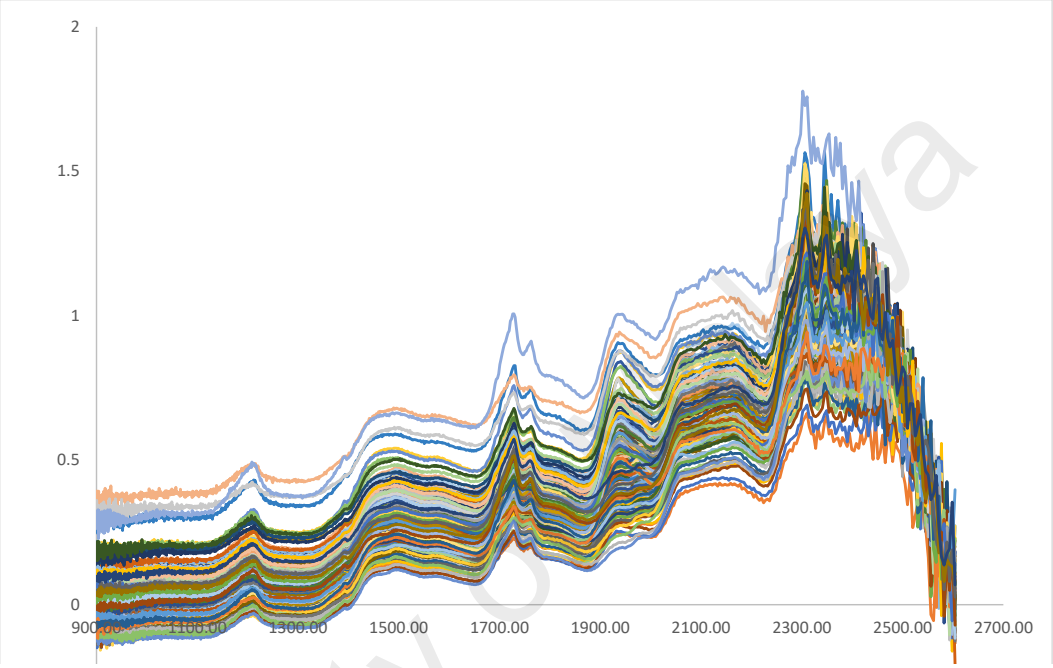

Figure 4.48: Spectrum of raw cow milk samples using a single photodiode FT-NIR spectrometer.

4.4.3.1 PCA

PCA was applied to the FT-NIR spectra to reduce the number of variables in the data matrix by choosing the PCs that had eigenvalues>1. In this work FT-NIR spectra of raw cow milk samples from 9 origins in Peninsular Malaysia namely Kedah, Perlis, Perak, Pinang, Pahang, Johor, Melaka and Terengganu is used for PCA and is observed that PCA doesn't give a clear clustering. 10 PC's with eigenvalue>1 is extracted from 173 samples.

Furthermore, factory milk samples from some selected continents of Asia, America, Europe and Middle East is used for PCA and is noted that PCA couldn't give clear mapping. 12 PCs with eigenvalues >1 is extracted from 144 factory milk samples. Moreover, non-supervised method of PCA is applied to remove out the unnecessary

information and keep the relevant ones so that similar ones are grouped closer to each other in one group.

4.4.3.2 Artificial neural network models built

In this research MLP-ANN is used as classification model for raw cow milk samples from various regions of Peninsular Malaysia and factory cow milk samples from some selected continents of the world. PCA as variable reduction method was used to determine the appropriate PC scores which will be used as inputs, for building the models (Palacios-Morillo *et al*., 2014).

(a) ***ANNs built for raw cow milk samples from Peninsular Malaysia using FT-NIR***

For raw cow milk samples there were 10 PCs which were retained with eigenvalues >1 accounting for more than 90% of the total variance. 60% of the data set was used as training, 25% as testing and 15% as validation. Training data is the data used for adjusting the predicted model to experimental data whereas validation data is another set of data used to check if the model is well designed. The predictor model should correctly estimate the geographical origin of a new set of data. Two nodes were used in the output layer corresponding to northern and southern sampling regions. The JMP Boosting was set to 2 base models, 10 boosting models and a learning rate of 0.1 and 1 under feed forward configuration. Generalized R-square, entropy R-square, RMSE misclassification rate, Med Abs Dev and -Log likelihood were used as selection factors to ascertain the best model. Different numbers of input variables were used to see if a reduced number of PCs as input variables can still give a good prediction.

It was noted that by using four PCs as input variables the prediction rate can achieve 100% classification and the R-square was 0.97 for training and 0.98 for validation with a learning rate of 1. In the case of a learning rate of 0.1, 5 PCs as input variables did not give a prediction rate of 100%. The best ANN model obtained is that

which has 10 input neurons, one hidden layer, ten hidden nodes and one output neuron as this is the architecture that yields the highest values of generalized R-Square and entropy R-Square while having the smallest values for RMSE, Mean Abs Dev and –Log likelihood. The prediction performance were compared to one another and investigated by misclassification rate and other statistical information as reported in Table 4.16. The graphical representation of the best ANN models obtained for the prediction of milk in Malaysia based on the FT-NIR data is illustrated in Figure 4.49. The definition of the format of data used in the table is as explained in section 4.4.2.4.

(b) ***ANNs built for factory cow milk samples from some selected countries using FT-NIR data***

For factory cow milk samples from various regions of the world, 12 PCs with eigenvalue >1 were retained. 60% of the data set was used as training, 25% as testing and 15% as validation. The JMP Boosting was set to 2 base models, 10 boosting models and a learning rate of 0.1 and 1 under feed forward configuration. Generalized R-square, entropy R-square, RMSE misclassification rate, Med Abs Dev and -Log likelihood were used as selection factors to ascertain the best model as reported in Table 4.17. The model with 12 input variables, one hidden layer, 16 hidden node and 1 output representing the origins of the milk samples (Canada, U.S.A, Belgium, Iran, Turkey, Azerbaijan, Australia and Malaysia) gives 100% prediction performance and zero misclassification rate as shown in Figure 4.50.

Table 4.16: ANN models built for raw cow milk samples based on northern and southern sampling regions in Peninsular Malaysia.

Network	Training	Validation	Learning rate
[10,1,20,1]	(0.9990942, 0.998233, 0.0032222, 0.0010698, 0.1483537)	(0.9990479, 0.998165, 0.0022614, 0.0010952, 0.0384232)	0.1
[10,1,10,1]	(0.9999998, 0.9999996, 7.6445e-7, 2.3271e-7, 3.2114e-5)	(0.9999998, 0.9999995, 8.2902e-7, 2.8121e-7, 9.8423e-6)	1
[9,1,20,1]	(0.9911425, 0.9828814, 0.0280911, 0.0099869, 1.4372346)	(0.9912069, 0.9832041, 0.0206327, 0.009827, 0.3516965)	0.1
[9,1,14,1]	(0.999999, 0.9999981, 4.009e-6, 1.1607e-6, 0.0001602)	(0.9999975, 0.9999951, 1.0437e-5, 2.9226e-6, 0.0001023)	1
[8,1,20,1]	(0.9881985, 0.9774856, 0.036194, 0.012783, 1.878145)	(0.9906989, 0.9815841, 0.0452151, 0.0103674, 0.3941562)	0.1
[8,1,20,1]	(0.9999872, 0.9999749, 4.782e-5, 1.5251e-5, 0.0021048)	(0.9999735, 0.9999489, 7.7863e-5, 3.0582e-5, 0.0010705)	1
[7,1,20,1]	(0.9886091, 0.97805, 0.0497649, 0.0118269, 1.8428722)	(0.9972079, 0.9946301, 0.0103135, 0.0031581, 0.1124428)	0.1
[7,1,20,1]	(0.9999991, 0.9999983, 4.0451e-6, 1.0331e-6, 0.0001426)	(0.9999991, 0.9999983, 3.1233e-6, 1.0194e-6, 3.568e-5)	1
[6,1,20,1]	(0.9850738, 0.9713551, 0.0456045, 0.0162389, 2.4049569)	(0.9893933, 0.9797815, 0.0266345, 0.0117201, 0.4233639)	0.1
[6,1,20,1]	(0.9998909, 0.9998909, 0.0004213, 0.0001265, 0.0001265)	(0.9998587, 0.999717, 0.000735, 0.0001779, 0.0060564)	1
[5,1,16,1]	(0.9999917, 0.9999837, 4.6333e-5, 9.8902e-6, 0.001365)	(0.9999758, 0.9999534, 7.9716e-5, 2.7874e-5, 0.0009757)	1
[4,1,20,1]	(0.9752884, 0.9531037, 0.0705174, 0.025427, 3.9373011)	(0.9783131, 0.9591704, 0.0703538, 0.0212067, 0.8549497)	1

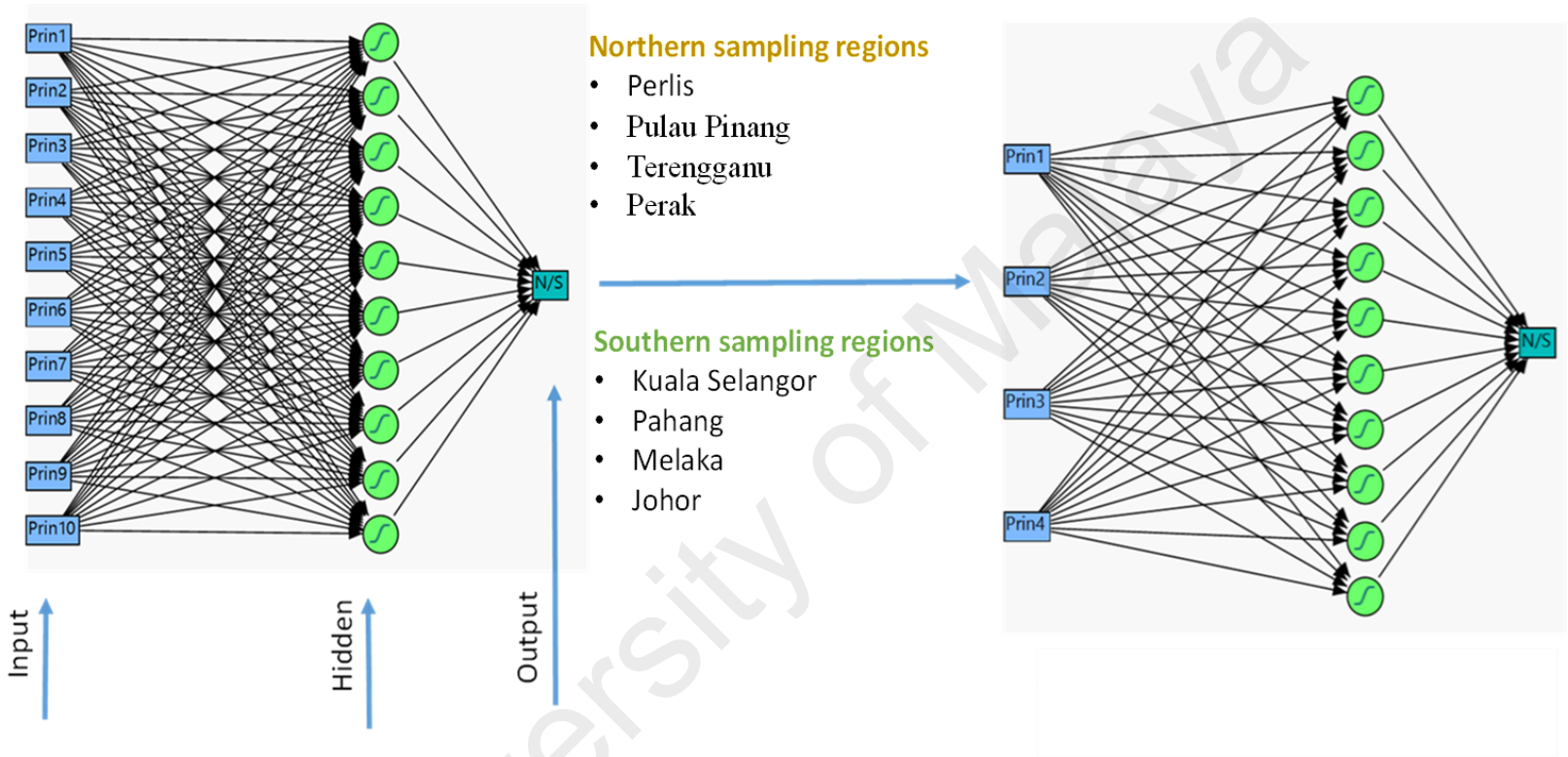

Figure 4.49: The best ANN models using FTNIR data a) before reducing the inputs; b) after reducing the inputs on the bases of northern and southern sampling regions.

Table 4.17: ANN models based on the data of factory cow milk samples obtained from various countries of the world.

Network	Training	Validation	Learning rate
[12,1,16,1]	(0.9999016, 0.9993939, 0.0022902, 0.0008935, 0.0914047)	(0.9999949, 0.9999714, 7.5671e-5, 4.0513e-5, 0.0006077)	1
[11,1,10,1]	(0.9954642, 0.9731213, 0.0759547, 0.0362359, 4.053257)	(0.9959329, 0.9778215, 0.0454855, 0.0302673, 0.4704653)	1
[10,1, 10,1]	(0.9990394,0.9941266,0.0163551, 0.0085454, 0.8857009)	(0.9997477, 0.9985829, 0.0032362, 0.0019987, 0.0300599)	1
[9,1,6,1]	(0.9951346, 0.9712493, 0.0775826, 0.0389648, 4.3355519)	(0.9953154, 0.9745723, 0.0542845, 0.0343744, 0.5393886)	1

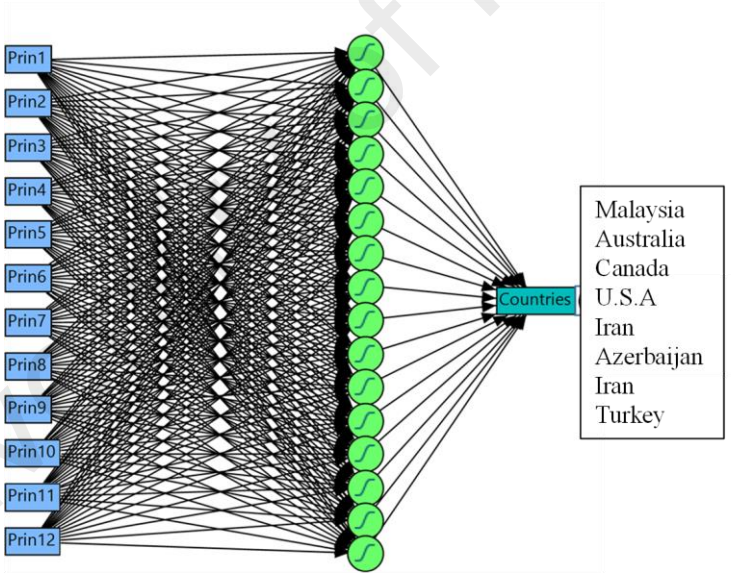

Figure 4.50: The best ANN model using FTNIR data on the bases of country of origin.

Figure 4.51 shows the spectra of factory milk samples from some selected countries of the world. As observed, the average spectrum of Iran milk has peaks around 1213.56, 1489.83, 1733.66, 1755.57, 1931.95, 2069.48, 2170.08, 2310.48, 2349.57, 2353.99, 2380.89, 2455.77, 2441.35, 2465.42, 2520.1, 2530.03 nm. As mentioned before the peaks are due to protein, fat and lactose.

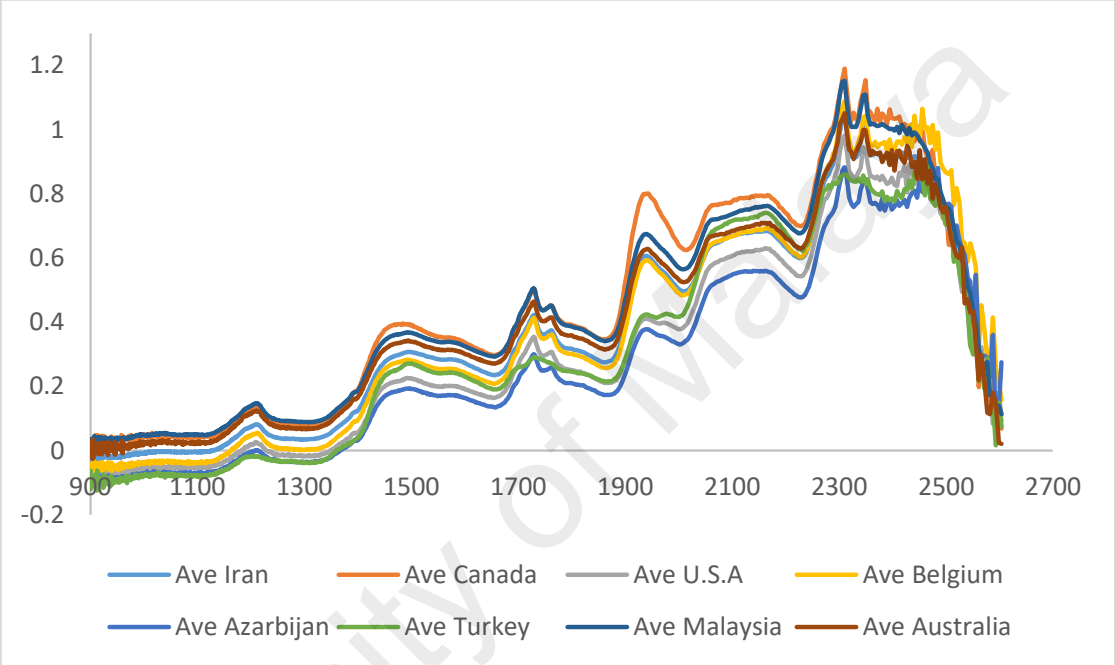

Figure 4.51: Spectrum of factory milk samples of some selected countries using a single photodiode FT-NIR spectrometer.

(c) Artificial neural network built for factory milk samples from different continents using FT-NIR

Figure 4.52 shows the average spectra of milks from continents namely America, Europe, Asia and Middle East. The peaks are around 1206.53, 1470.55, 1588.27, 1724.09, 1758.04, 1934.94, 2055.87, 2158.84, 2306.22, 2349.57, 2450.92 nm. As it has been mentioned previously these peaks are due to lactose, protein and fat which in different wavelengths have different assignments.

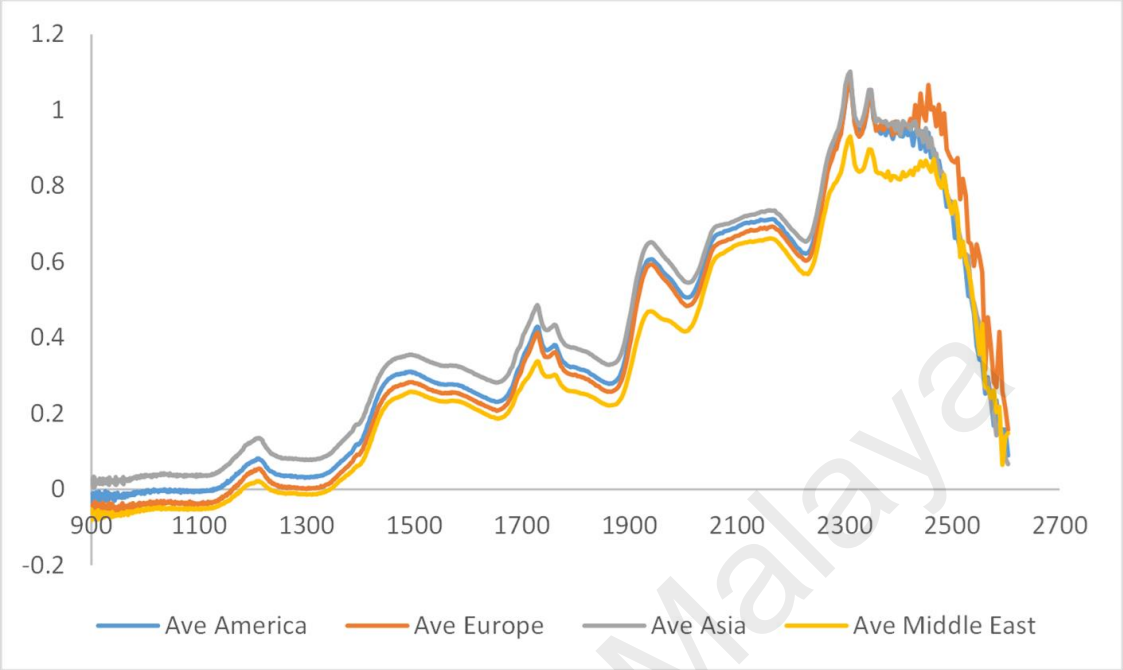

Figure 4.52: Spectra of factory milk samples of some selected continents using a single photodiode FT-NIR spectrometer.

In order to build a model that would be able to predict milks from different continents, 12 PCs with eigenvalue >1 were retained. The JMP Boosting was set to 2 base models, 10 boosting models and a learning rate of 0.1 and 1 under feed forward configuration. Generalized R-square, entropy R-square, RMSE misclassification rate, Med Abs Dev and -Log likelihood were used as selection factors to ascertain the best model. The best ANN model is shown Table 4.18. The model is that of 12 input variables, 1 hidden layer, 8 hidden nodes and 1 output, which gave 100% prediction performance with zero misclassification rate as shown in Figure 4.53. It was observed that reducing the number of input variables does not result in 100% classification.

Table 4.18: ANN models of factory cow milk samples built based on continent of origin.

Network	Training	Validation	Learning rate
[12,1,8,1]	(0.9975957, 0.9923285, 0.0288886, 0.0072219, 0.7863897)	(0.999007, 0.9970242, 0.0063581, 0.0028163, 0.0425511)	1

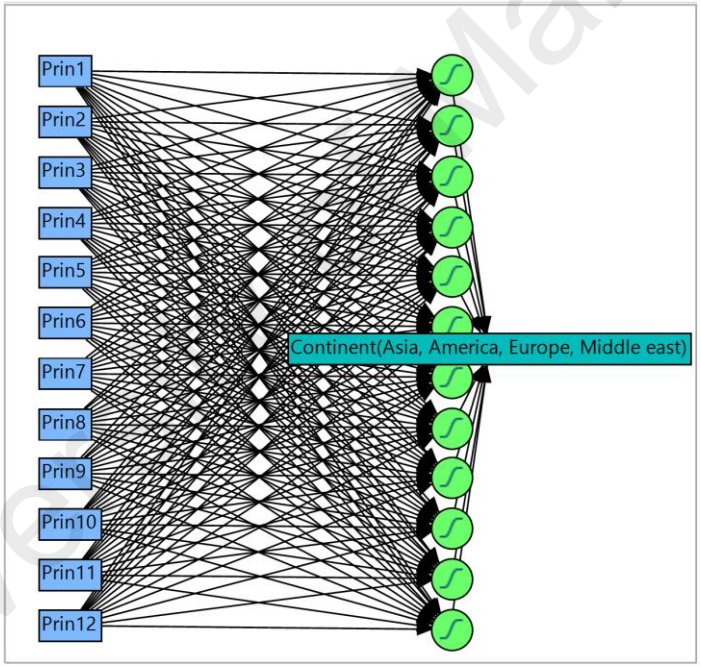

Figure 4.53: The best model using FTNIR data on the bases of continent of origin.

4.4.4 Milk spectra obtained from Shimadzu UV-2600

The spectra of 118 plasma of milk samples had been taken in absorbance mode for samples obtained from the northern and southern regions of Peninsular Malaysia, namely Perlis, Perak, Kedah, Terengganu and Pulau Pinang from the north and Johor, Melaka, Pahang and Kuala Selangor from the southern part of the country. Each spectrum was analyzed three times. Figure 4.54 shows the spectra for the plasma of milk samples in the UV range from 190-400 nm. Clear variation can be observed within the whole spectral range. The peaks in the range of 200-240 nm are generally due to the presence of lipids and the peaks within the range of 240-320 are generally due to proteins (Forcato *al*., 2005).

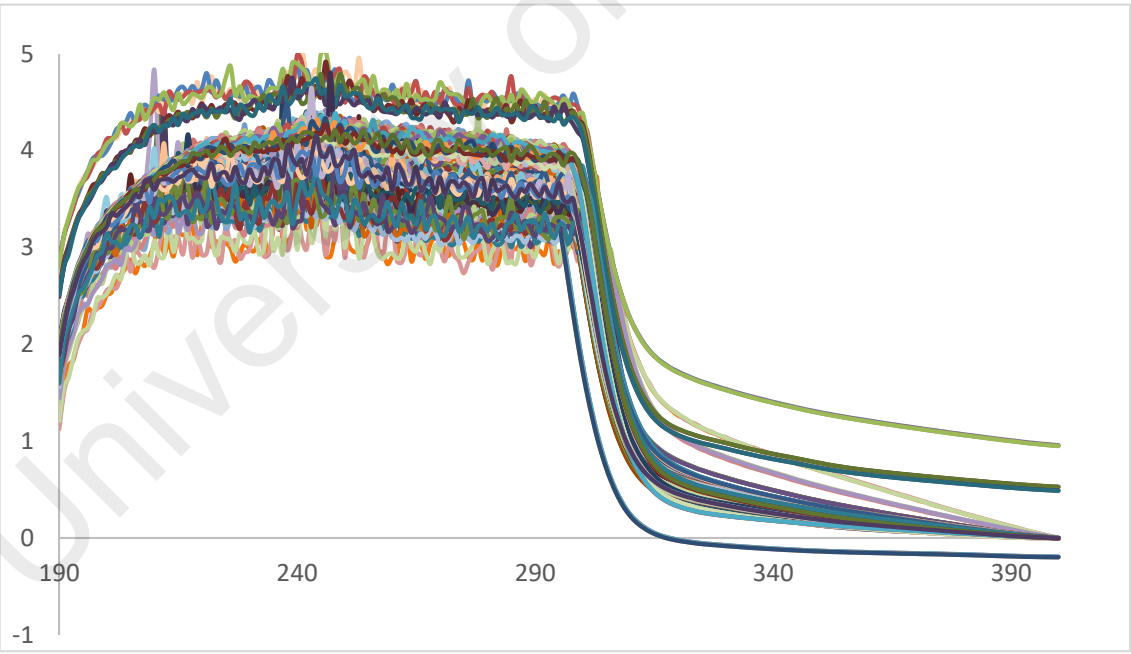

Figure 4.54: Spectra of the plasma of raw cow milk samples from northern and southern regions in Peninsular Malaysia using UV-2600.

ANN models were built based on the data obtained to predict the origin of sampling sites. Based on the 118 spectra obtained using the Shimadzu UV-2600, a PCA analysis was carried out to determine the PC scores which will be used as input

variables for building the ANN models. In order to reduce the variables, PCs with eigenvalues >1 were selected. There are a total of 6 PCs which are retained explaining a total variance of 90.8 %. Data was then divided into training (60%), validation (25%) and testing (15%). For validation purpose, K-fold cross validation was used by dividing the data to K- dataset (in our case K=5) for building k- models with different verification dataset. The classification performances of the models were checked by R-Square, R-entropy square, RMSE, Mean St Dev, -Loglikelihood.

Again, the ANN models were built using the boosting method available in JMP. In order to determine the best architecture the simplest model using only one hidden layer and two learning rates of 1 and 0.1. The input variables were also varied in order to observe variation in prediction performance.

The results of the ANN model building are reported in Table 4.19 where the best model is graphically presented in Figure 4.55. It was observed that the model with 4 input variables, one hidden layer, twelve hidden nodes and one output neuron exhibited the best R-Square and entropy R-Square. The misclassification rate for validation was zero and 100% classification performance was obtained but the misclassification rate for testing was 0.01 in all of the obtained architectures. Reducing the number of input variables (PCs) to three could still provide satisfactory prediction performance.

Henceforth, another architecture was built to see if the obtained model would give zero misclassification for training data as well as validation. The initial architecture was built with two hidden layers where in the first layer, 2 of the nodes have TanH transfer functions and 1 linear function and in the second layer 1 TanH and 1 linear transfer transfer functions were used.

Table 4.19: ANN models using UV data for factory cow milk samples built based on northern and southern sampling regions of Peninsular Malaysia.

Network	Training	Validation	Learning rate
[6, 1, 20, 1]	(0.986599, 0.971688, 0.0740336, 0.015317, 1.8350418)	(0.9969159, 0.9933845, 0.0143645, 0.0044546, 0.1095006)	1
[6, 1 ,20, 1]	(0.9301436, 0.8633601, 0.1484013, 0.0791519, 8.9392612)	(0.9591879, 0.9167679, 0.116572, 0.0481577, 1.3251085)	0.1
[5,1, 12, 1]	(0.9861898, 0.9708406, 0.0796521, 0.0150809, 1.8899689)	(0.9882696, 0.9751533, 0.0652629, 0.0144074, 0.4112649)	1
[5,1, 20, 1]	(0.9681104, 0.9344209, 0.1020468, 0.0356233, 4.2903193)	(0.9803842, 0.9587975, 0.0610038, 0.0263294, 0.6559702)	0.1
[4, 1 ,16, 1]	(0.9837237, 0.9657547, 0.0829884, 0.0177048, 2.2196073)	(0.998838, 0.9975004, 0.0052434, 0.00171, 0.0413736)	1
[4, 1, 12, 1]	(0.9867503, 0.9720016, 0.0785017, 0.0141862, 1.8147153)	(0.99927, 0.9984287, 0.0034607, 0.0010776, 0.0260081)	0.1
[3, 1, 20,1]	(0.9807316, 0.9596317, 0.0890728, 0.0205748, 2.6164723)	(0.9959006, 0.9912198, 0.010918, 0.0059949, 0.1453311)	1
[3,1, 20 , 1]	(0.9526382, 0.9046179, 0.1285037, 0.0537637, 0.0105263)	(0.9355755, 0.8726583, 0.1655382, 0.067933, 2.0273615)	0.1

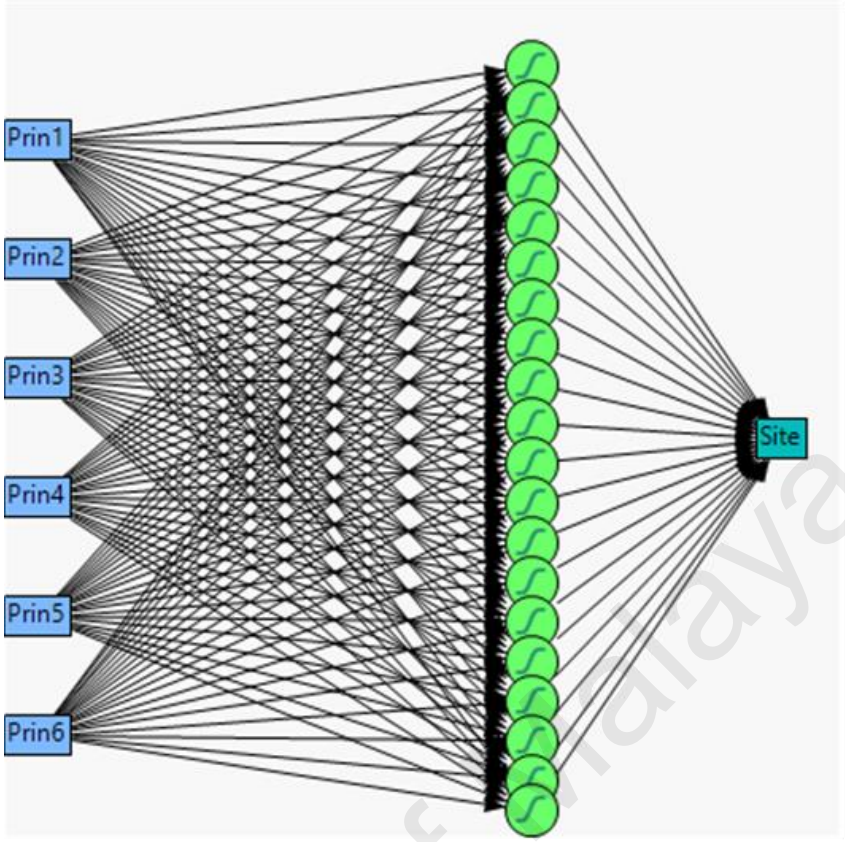

Figure 4.55: The best ANN model using UV data on the bases of northern and southern sampling sites in Peninsular Malaysia.

The boosting method of JMP was used for two learning rates of 0.1(min) and 1(max). The number of input variables were also changed in order to observe for changes in prediction performance. As observed in Table 4.20 the best model obtained still gave zero misclassification rate for validation but not for testing data. From the models observed, the one with 6 PCs, two hidden layer and learning rate of one is preferred due to the fact that the generalized R-Square and entropy R-Square of this model are closer to one and the other factors are low. The misclassification rate for validation is zero with 100% of classification rate while for testing it gives 0.01 misclassification rate. It is also noted that reducing the number of input variables (PCs) to two could still give a good prediction performance. In conclusion, the new architecture presented in Figure 4.56 did not give a better performance compared to the architecture with one hidden layer.

Table 4.20: ANN models with two hidden layers of factory cow milk samples built based on the northern and southern sampling regions of Peninsular Malaysia.

Network	Training	Validation	Learning rate
[6, 2, 8, 12,1]	(0.9881588, 0.9749271, 0.0731811, 0.0131601, 1.6251022)	(0.9986112, 0.9970135, 0.0044691, 0.0020496, 0.0494329)	1
[6, 2, 20, 30,1]	(0.9842428, 0.9668223, 0.077188, 0.0183905, 2.1504113)	(0.9954217, 0.990201, 0.017637, 0.006596, 0.1621941)	0.1
[5, 2, 4, 6,1]	(0.9881149, 0.9748358, 0.0734121, 0.0131827, 1.6310169)	(0.9973314, 0.9942723, 0.009055, 0.0039085, 0.0948052)	1
[5, 2, 20, 30,1]	(0.9825741, 0.9633962, 0.080794, 0.0200647, 2.3724783)	(0.9786012, 0.9552956, 0.0700958, 0.0279124, 0.7399511)	0.1
[4, 2, 8, 12,1]	(0.9933311, 0.9857571, 0.050888, 0.0078864, 0.9342048)	(0.9767112, 0.9517542, 0.1071278, 0.0240314, 0.7596866)	1
[4, 2, 20, 30,1]	(0.9812936, 0.9607779, 0.0809928, 0.0218379, 2.5421802)	(0.9961369, 0.9917229, 0.0135064, 0.0056147, 0.1370031)	0.1
[3, 2, 20, 30,1]	(0.9867741, 0.9720508, 0.0764071, 0.0146696, 1.8115262)	(0.9984858, 0.9967445, 0.0057813, 0.0022283, 0.0538849)	1
[3, 2, 20,30,1]	(0.9661228, 0.930523, 0.102194, 0.040354, 4.5453242)	(0.9604382, 0.9191803, 0.1160909, 0.0467866, 1.2867012)	0.1
[2, 2, 8, 12,1]	(0.9756061, 0.9492616, 0.0939933, 0.025486, 3.2886124)	(0.9859398, 0.9703183, 0.0531343, 0.0187969, 0.4912944)	1
[2, 2, 20, 30,1]	(0.9750767, 0.9481988, 0.0968392, 0.0276868, 3.3574976)	(0.9909584, 0.9807739, 0.0287761, 0.0128215, 0.3182322)	0.1

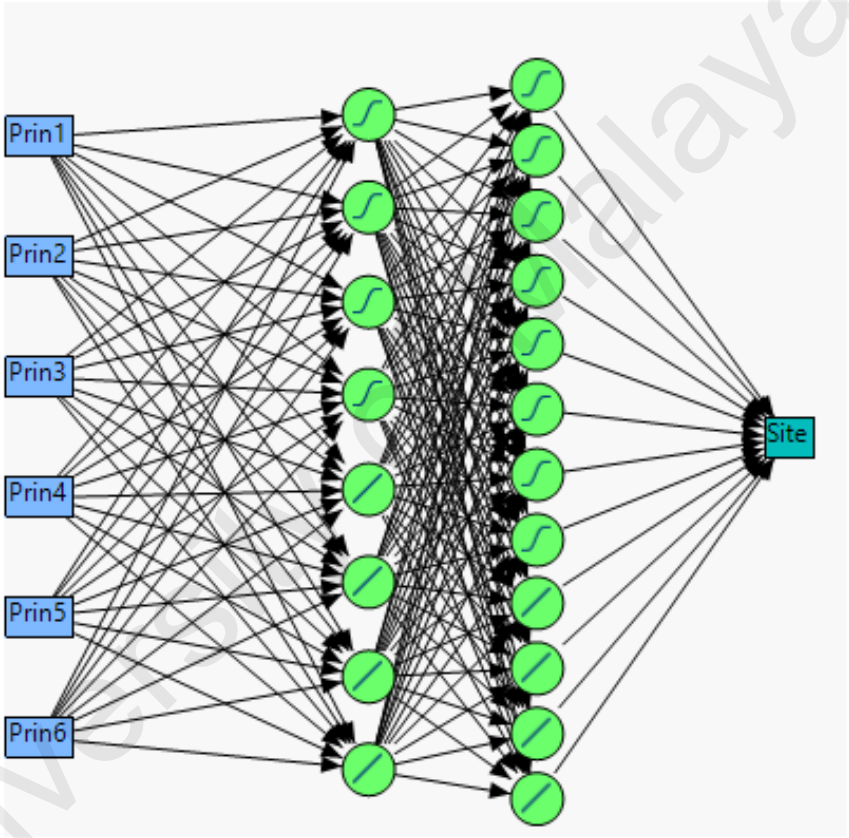

Figure 4.56: The best ANN model using UV data based on northern and southern sampling regions in Peninsular Malaysia.

CHAPTER 5: CONCLUSION

The main motivation of this thesis is to study the different methodologies of analysing cow milk in order to ascertain ways to determine its geographical origin. Several analytical methods were employed coupled with various chemometric methods in order to achieve the desired objectives. In this chapter, the conclusion of each of the studied methodologies will first be discussed and later, the overall conclusion if this thesis will be made.

5.1 ICP-MS results

From the results of ICP-MS it can be concluded that the two sample student t-test confirms the significant differences in composition of milk samples which allows the separation of milk samples possible on the basis of their geographical origin. This is entirely possible due to the different environmental and farming conditions of each individual region.

Moreover, it is observed that milk in Malaysia is significantly different ($p<0.05$) in the concentrations of Mn, Mg, Ba, Fe, K, Ca, Mo and Se compared to some of the selected regions of the world. It is also observed that concentrations of Cu, Na and Zn are similar between Malaysia and other countries studied in this research.

As far as PCA and clustering is concerned, Malaysian milk clustered based on the significant loadings of K, Ba, Mg, Fe, Na, K, Mn and Ca while milk from the selected countries of the world clusters mainly by virtue of the loadings on Se, Cu and Mo.

Based on the different loading of metals that act to separate the milk samples for different geographical locations in the PCA, it is clear that concentrations of essential and trace elements can be used to identify the origins and authenticity of milk, which

form the basis in ascertaining the quality of milk. By employing the constellation plot on the results of HCA, a clearer picture of milk sample clustering allows a better view of how the content of metals in milk samples can be used to discriminate milks from different geographical regions. Heat maps have also been used to illustrate the variation in the elemental concentration for each individual region. Discriminant analysis was used as an alternative method to cluster the samples based on the detected elements and it was observed that the model prediction gave 100% correct classification of samples based on their geographical origin.

5.2 IRMS results

We have determined carbon, nitrogen and oxygen isotopic ratios ($\delta^{13}C$, $\delta^{15}N$, $\delta^{18}O$) of 69 cow milk samples from various farms in Peninsular Malaysia as well as 21 cow tail hair samples in an attempt to ascertain geographical origins of these samples. Results obtained show that there is a high positive correlation between isotopic ratios of cow hair and milk suggesting that hair could be used as a substitute for milk in building the database to ascertain origins of milk. It is also observed that there are significant differences in the mean values of isotopic ratio for samples from different sampling regions which allow their discrimination based on geographical origins. We also noticed that the isotopic ratio of carbon in Malaysian milk is related to the abundance of C4 plants near the equator. The value of $\delta^{13}C$ in milk decreases in samples collected as we move from the equator towards the poles. The $\delta^{15}N$ values in both cow's milk and hair of peninsular Malaysia are higher in the northern region compared to the south and this is most probably due to the fact that cows in the north are mostly grass fed and organic fertilisers are used more frequently compared to chemical fertilisers. The isotopic ratio of oxygen is seen to be dependent on the latitude of each region and the samples from the south are seen to be richer in heavy oxygen isotope. Simple correlation plots of relevant isotopic ratios indicate that geographical clustering between samples of the

northern and southern regions naturally arise. PCA and HCA also show clustering due to geographical regions. Moreover, heat maps have also been used to illustrate the variation in isotopic ratios based on color variation. Factorial analysis was able to cluster the samples based on their production region by reducing the number of original variables of the samples to the relative factors. Based on the results of factorial analysis the variables that influence the clustering on the basis of factors 1 and 2 were noted.

5.3 Ultrasonic results

The nutritional and physical information of raw cow milk from different regions of Peninsular Malaysia show different physical and nutritional values which could be due to various factors such as different environmental condition, differences in breed and the food consumed by the animal or even the lactation state of each farm. We can relate this information to the ones we have obtained from elemental composition of milk, soil, mixed food, Napier grass, pallet and water.

Furthermore, the milk produced from different breeds of cow living in the same environment and consuming the same food was analyzed to see if there are any notable differences in the milk produced. It was noted that differences in the breed can cause differences in physical and nutritional information of milk as well as its elemental characteristics even if different breeds of cow are living in the same farm and are fed with the same food.

5.4 Spectroscopic results

Analysis of milk samples from different regions of Peninsular Malaysia was compared between three instruments, namely a Shimadzu UV-3600, single photodiode FT-NIR and Shimadzu UV-2600. From the analysis carried out on the Shimadzu UV-3600 which was in the range of 200-2500 nm several spectra were collected. Samples from northern sampling region are significantly different ($p<0.05$) from south in the

ranges of UV (200-400nm), Vis (400-800nm) and some parts of NIR (1500-2500nm). Differences in the spectra of milk are mainly due to the difference in the composition of the milk, namely the protein, fat and lipid. In UV range differences in the intensity of the peaks are due to the difference of composition in protein and fat between samples from the north and the south. In the Vis and NIR regions, different compositions of fat, protein and lipid also cause separation between the northern and southern samples.

PCA as a variable reduction method was then applied to the full range spectra. 14 PCs having eigenvalues >1 were retained. From the loading plot almost all wavelengths of the full spectral range loads PC1 and PC3 positively. PC1 explains 71.9% of the total variance and PC3 explains 6.29 % of the total variance. PCA scores plot does not show linear separation of the regions based on the full spectral range data. Due to this, the data were split based on the different spectral ranges. PCA for each range was carried out.

Then using the full spectral range, ANN models were built. Initially, 14 PCs were used as input variables for the training and testing of ANN models. Then, the number of input variables were reduced and the boosting method of JMP was used to optimize the ANN models. It was found that the model with 7 input neurons, 1 hidden layer, 10 hidden node and 1 output neuron gave the best performance where milk samples from the north and south of Peninsular Malaysia are classified correctly by the model and is characterized by the most optimum statistics. However, reducing the input variables (no of PCs) to 3 still give 100% classification indicating that three PCs are also sufficient to be used for building a suitable model in this case.

Several spectra in the range of 900-2600 nm were collected from the FT-NIR. For this data, PCA did not give a clear separation of samples based on northern and southern sampling regions. Consequently, the data was analyzed using ANN. The input

data for the ANN was chosen based on 10 PCs that had eigenvalues > 1. The best ANN model for clustering raw cow milk samples based on the northern and southern sampling regions of Peninsular Malaysia is the one with 10 input neurons, one hidden layer, 10 hidden nodes and one output node. The model is preferred from a number of different architectures as it gave 100% prediction rate and highest R-Square and entropy R-Square with low RMSE, Mean Abs Dev, and –Logliklihood as well as zero misclassification rate. By reducing the input variables, it was noted that with four PCs as input variables the prediction rate can achieve 100% classification and the R-square was 0.97 for training and 0.98 for validation with a learning rate of 1. In the case of a learning rate of 0.1, 5 PCs as input variables did not give a prediction rate of 100%.

Factory milk samples are best differentiated based on the country of their origin, namely U.S.A, Canada, Belgium, Iran, Azerbaijan, Turkey, Australia and Malaysia, 12 PCs as input variables, one hidden layer, 16 hidden nodes and one output neuron is seen to perform best R-Square values closest to one. By reducing the number of input variables to 9, with one hidden layer, 6 hidden nodes and one output neuron the classification rate of 100% with zero misclassification can still be achieved. Separating factory cow milk samples based on the continent of origin was also carried out in this work. The model with 12 input variables, one hidden layer, 8 hidden nodes and one output neuron performs best in separating milk based on the 4 continents of America, Europe, Middle East and Asia.

A Shimadzu UV-2600 was used for analyzing the plasma of milk samples that were extracted after centrifuging the samples. 6 PCs with eigenvalues > 1 presenting 90.8% of total variance were initially selected for the ANN input. ANN model with 4 input variables, one hidden node, 12 hidden layer, and one output neuron with 0.1 learning rate performed best. Reducing the input variables in this case to 3 input

variables, with one hidden layer, 12 hidden nodes and one output neuron still gives a 100% prediction but not with the most optimum statistics.

5.5 Overall conclusion

Milk samples from different regions of Peninsular Malaysia had been analyzed by ICP-MS, IRMS, milko analyser and UV/Vis/NIR spectrometers. The information obtained from these instruments are respectively reported as elemental, isotopic, nutritional, physical and spectrophotometric. On the basis of geographical origin we are clearly able to differentiate factory milk samples of Malaysia from other selected regions by their elemental composition. On the isotopic ratio information, we observe that the separation of samples on the basis of their individual region is also possible. By applying various chemometric methods it is generally observed that two major clusters consisting of the northern sampling regions of Perak, Perlis, Terengganu, Kedah and Pulau Pinang and the southern region of Johor, Melaka, Kuala Selangor and Pahang are formed. Analyses of spectroscopic information of milk samples also show clustering of raw cow milk to two classes of northern and southern sampling regions which are due to the difference in the content of protein, fat and lactose for the different regions. ANN models were also developed based on the spectroscopic data. Best ANN models developed based on the data from different spectroscopic ranges all gave 100% classification rates for both testing and training data with zero misclassification. In addition, study of the nutritional and physical characteristics of milk samples have also been carried out in this thesis.

REFERENCES

Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature, 422*(6928), 198-207.

Almeida, C. M., & Vasconcelos, M. T. S. D. (2001). ICP-MS determination of strontium isotope ratio in wine in order to be used as a fingerprint of its regional origin. *Journal of Analytical Atomic Spectrometry, 16*(6), 607-611.

Alonso, M. L., Benedito, J. L., Miranda, M., Castillo, C., Hernandez, J., & Shore, R. F. (2002). Interactions between toxic and essential trace metals in cattle from a region with low levels of pollution. *Archives of Environmental Contamination and Toxicology, 42*(2), 165-172.

Alzate, A., Perez-Conde, M. C., Gutierrez, A. M., & Camara, C. (2010). Selenium-enriched fermented milk: A suitable dairy product to improve selenium intake in humans. *International dairy journal, 20*(11), 761-769.

Ambrose, S. H., & Deniro, M. J. (1986). Reconstruction of African Human Diet Using Bone-Collagen Carbon and Nitrogen Isotope Ratios. *Nature, 319*(6051), 321-324.

Andueza, D., Agabriel, C., Constant, I., Lucas, A., & Martin, B. (2013). Using visible or near infrared spectroscopy (NIRS) on cheese to authenticate cow feeding regimes. *Food Chem, 141*(1), 209-214.

Anjos, O., Iglesias, C., Peres, F., Martínez, J., García, Á., & Taboada, J. (2015). Neural networks applied to discriminate botanical origin of honeys. *Food Chemistry, 175*, 128-136.

Ariyama, K., & Yasui, A. (2006). The determination technique of the geographic origin of Welsh onions by mineral composition and perspectives for the future. *Jarq-Japan Agricultural Research Quarterly, 40*(4), 333-339.

Ataro, A., McCrindle, R. I., Botha, B. M., McCrindle, C. M. E., & Ndibewu, P. P. (2008). Quantification of trace elements in raw cow's milk by inductively coupled plasma mass spectrometry (ICP-MS). *Food Chemistry, 111*(1), 243-248.

Awad, T., Moharram, H., Shaltout, O., Asker, D., & Youssef, M. (2012). Applications of ultrasound in analysis, processing and quality control of food: A review. *Food Research International, 48*(2), 410-427.

Barbosa, R. M., Batista, B. L., Varrique, R. M., Coelho, V. A., Campiglia, A. D., & Barbosa, F. (2014). The use of advanced chemometric techniques and trace element levels for controlling the authenticity of organic coffee. *Food Research International, 61*, 246-251.

Barile, D., Coı, J., Arlorio, M., & Rinaldi, M. (2006). Identification of production area of Ossolano Italian cheese with chemometric complex approach. *Food Control, 17*(3), 197-206.

Baroni, M. V., Podio, N. S., Badini, R. G., Inga, M., Ostera, H. A., Cagnoni, M., Hoogewerff, J. (2011). How much do soil and water contribute to the composition of meat? A case study: Meat from three areas of Argentina. *Journal of agricultural and food chemistry, 59*(20), 11117-11128.

Baroni, M. V., Podio, N. S., Badini, R. G., Inga, M., Ostera, H. A., Cagnoni, M., Wunderlin, D. A. (2015). Linking soil, water, and honey composition to assess the geographical origin of argentinean honey by multielemental and isotopic analyses. *J Agric Food Chem, 63*(18), 4638-4645.

Basheer, I., & Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of microbiological methods, 43*(1), 3-31.

Bassbasi, M., De Luca, M., Ioele, G., Oussama, A., & Ragno, G. (2014). Prediction of the geographical origin of butters by partial least square discriminant analysis (PLS-DA) applied to infrared spectroscopy (FTIR) data. *Journal of Food Composition and Analysis, 33*(2), 210-215.

Batista, B. L., da Silva, L. R. S., Rocha, B. A., Rodrigues, J. L., Berretta-Silva, A. A., Bonates, T. O., . . . Barbosa, F. (2012). Multi-element determination in Brazilian honey samples by inductively coupled plasma mass spectrometry and estimation of geographic origin with data mining techniques. *Food Research International, 49*(1), 209-215. doi:10.1016/j.foodres.2012.07.015

Benincasa, C., Lewis, J., Sindona, G., & Tagarelli, A. (2008). The use of multi element profiling to differentiate between cowand buffalo milk. *Food Chemistry, 110*(1), 257-262.

Bernuy, B., Meurens, M., Mignolet, E., & Larondelle, Y. (2008). Performance comparison of UV and FT-Raman spectroscopy in the determination of conjugated linoleic acids in cow milk fat. *Journal of agricultural and food chemistry, 56*(4), 1159-1163.

Berrueta, L. A., Alonso-Salces, R. M., & Héberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatography A, 1158*(1), 196-214.

Bilandžić, N., Sedak, M., Đokić, M., & Božić, Đ. (2015). Determination of Macro- and Microelements in Cow, Goat, and Human Milk Using Inductively Coupled Plasma Optical Emission Spectrometry. *Spectroscopy Letters, 48*(9), 677-684.

Bilandžić, N., Sedak, M., Đokić, M., Božić, Đ., Solomun-Kolanović, B., & Varenina, I. (2015). Differences in macro-and microelement contents in milk and yoghurt. *Archives of Biological Sciences*(00), 117-117.

Bodyfelt, F. W., Drake, M. A., & Rankin, S. A. (2008). Developments in dairy foods sensory science and education: From student contests to impact on product quality. *International dairy journal, 18*(7), 729-734.

Bontempo, L., Larcher, R., Camin, F., Hölzl, S., Rossmann, A., Horn, P., & Nicolini, G. (2011). Elemental and isotopic characterisation of typical Italian alpine cheeses. *International dairy journal, 21*(6), 441-446.

Bontempo, L., Lombardi, G., Paoletti, R., Ziller, L., & Camin, F. (2012). H, C, N and O stable isotope characteristics of alpine forage, milk and cheese. *International dairy journal, 23*(2), 99-104.

Borges, E. M., Gelinski, J. M. L. N., de Oliveira Souza, V. C., Barbosa Jr, F., & Batista, B. L. (2015). Monitoring the authenticity of organic rice via chemometric analysis of elemental data. *Food Research International, 77*, 299-309.

Borin, A., Ferrão, M. F., Mello, C., Maretto, D. A., & Poppi, R. J. (2006). Least-squares support vector machines and near infrared spectroscopy for quantification of common adulterants in powdered milk. *Analytica Chimica Acta, 579*(1), 25-32.

Brandão, M. C., Carmo, A. P., Bell, M. J. V., & Anjos, V. C. (2010). Caracterização de leite por espectroscopia infra-vermelho. *Revista do Instituto de Laticínios Cândido Tostes, 65*(373), 30-33.

Brereton, R. G. (2003). *Chemometrics: data analysis for the laboratory and chemical plant*: John Wiley & Sons.

Brescia, M., Monfreda, M., Buccolieri, A., & Carrino, C. (2005). Characterisation of the geographical origin of buffalo milk and mozzarella cheese by means of analytical and spectroscopic determinations. *Food Chemistry, 89*(1), 139-147.

Bressy, F. C., Brito, G. B., Barbosa, I. S., Teixeira, L. S. G., & Korn, M. G. A. (2013). Determination of trace element concentrations in tomato samples at different stages of maturation by ICP OES and ICP-MS following microwave-assisted digestion. *Microchemical Journal, 109*, 145-149.

Brighenti, M., Govindasamy-Lucey, S., Lim, K., Nelson, K., & Lucey, J. A. (2008). Characterization of the rheological, textural, and sensory properties of samples of commercial US cream cheese with different fat contents. *J Dairy Sci, 91*(12), 4501-4517.

Bunaciu, A. A., Aboul-Enein, H. Y., & Hoang, V. D. (2016). Vibrational spectroscopy used in milk products analysis: A review. *Food Chem, 196*, 877-884.

Cai, Q., Long, M. L., Zhu, M., Zhou, Q. Z., Zhang, L., & Liu, J. (2009). Food chain transfer of cadmium and lead to cattle in a lead-zinc smelter in Guizhou, China. *Environ Pollut, 157*(11), 3078-3082.

Camin, F., Perini, M., Colombari, G., Bontempo, L., & Versini, G. (2008). Influence of dietary composition on the carbon, nitrogen, oxygen and hydrogen stable isotope ratios of milk. *Rapid Commun Mass Spectrom, 22*(11), 1690-1696.

Camin, F., Wehrens, R., Bertoldi, D., Bontempo, L., Ziller, L., Perini, M.,Larcher, R. (2012). H, C, N and S stable isotopes and mineral profiles to objectively guarantee the authenticity of grated hard cheeses. *Anal Chim Acta, 711*, 54-59.

Camin, F., Wietzerbin, K., Cortes, A. B., Haberhauer, G., Lees, M., & Versini, G. (2004). Application of multielement stable isotope ratio analysis to the characterization of French, italian, and spanish cheeses. *J Agric Food Chem, 52*(21), 6592-6601. doi:10.1021/jf040062z

Cancilla, J. C., Perez, A., Wierzchos, K., & Torrecilla, J. S. (2016). Neural Networks Applied to Determine Thermophysical Properties of Amino Acid Based Ionic Liquids. *Physical Chemistry Chemical Physics*.

Capici, C., Mimmo, T., Kerschbaumer, L., Cesco, S., & Scampicchio, M. (2015). Determination of Cheese Authenticity by Carbon and Nitrogen Isotope Analysis: Stelvio Cheese as a Case Study. *Food Analytical Methods, 8*(8), 2157-2162.

Capuano, E., Rademaker, J., van den Bijgaart, H., & M. van Ruth, S. (2014). Verification of fresh grass feeding, pasture grazing and organic farming by FTIR spectroscopy analysis of bovine milk. *Food Research International, 60*, 59-65.

Caredda, M., Addis, M., Ibba, I., Leardi, R., Scintu, M. F., Piredda, G., & Sanna, G. (2016). Prediction of fatty acid content in sheep milk by Mid-Infrared spectrometry with a selection of wavelengths by Genetic Algorithms. *LWT - Food Science and Technology, 65*, 503-510.

Caroli, A., Chessa, S., & Erhardt, G. (2009). Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. *Journal of dairy science, 92*(11), 5335-5352.

Casale, M., Bagnasco, L., Giordani, P., Mariotti, M. G., & Malaspina, P. (2015). NIR spectroscopy as a tool for discriminating between lichens exposed to air pollution. *Chemosphere, 134*(0), 355-360.

Cashman, K. D. (2002). Calcium intake, calcium bioavailability and bone health. *Br J Nutr, 87 Suppl 2*(2), S169-177.

Cheajesadagul, P., Arnaudguilhem, C., Shiowatana, J., Siripinyanond, A., & Szpunar, J. (2013). Discrimination of geographical origin of rice based on multi-element fingerprinting by high resolution inductively coupled plasma mass spectrometry. *Food Chem, 141*(4), 3504-3509.

Chesson, L. A., Valenzuela, L. O., O'Grady, S. P., Cerling, T. E., & Ehleringer, J. R. (2010). Hydrogen and oxygen stable isotope ratios of milk in the United States. *J Agric Food Chem, 58*(4), 2358-2363.

Chevallier, E., Chekri, R., Zinck, J., Guérin, T., & Noël, L. (2015). Simultaneous determination of 31 elements in foodstuffs by ICP-MS after closed-vessel microwave digestion: Method validation based on the accuracy profile. *Journal of Food Composition and Analysis, 41*(0), 35-41.

Chudzinska, M., & Baralkiewicz, D. (2011). Application of ICP-MS method of determination of 15 elements in honey with chemometric approach for the verification of their authenticity. *Food Chem Toxicol, 49*(11), 2741-2749.

Chung, I.-M., Park, I., Yoon, J.-Y., Yang, Y.-S., & Kim, S.-H. (2014). Determination of organic milk authenticity using carbon and nitrogen natural isotopes. *Food Chemistry, 160*, 214-218.

Clark, L. C., Combs, G. F., Jr., Turnbull, B. W., Slate, E. H., Chalker, D. K., Chow, J., Taylor, J. R. (1996). Effects of selenium supplementation for cancer prevention

in patients with carcinoma of the skin. A randomized controlled trial. Nutritional Prevention of Cancer Study Group. *Jama, 276*(24), 1957-1963.

Codina-Torrella, I., Guamis, B., & Trujillo, A. J. (2015). Characterization and comparison of tiger nuts (Cyperus esculentus L.) from different geographical origin. *Industrial Crops and Products, 65*, 406-414.

Coppa, M., Chassaing, C., Ferlay, A., Agabriel, C., Laurent, C., Borreani, G.,Martin, B. (2015). Potential of milk fatty acid composition to predict diet composition and authenticate feeding systems and altitude origin of European bulk milk. *J Dairy Sci, 98*(3), 1539-1551.

Coppa, M., Martin, B., Agabriel, C., Chassaing, C., Sibra, C., Constant, I.,Andueza, D. (2012). Authentication of cow feeding and geographic origin on milk using visible and near-infrared spectroscopy. *Journal of dairy science, 95*(10), 5544-5551.

Craig, H. (1961). Isotopic variations in meteoric waters. *Science, 133*(3465), 1702-1703.

Crittenden, R., Andrew, A., LeFournour, M., Young, M., Middleton, H., & Stockmann, R. (2007). Determining the geographic origin of milk in Australasia using multi-element stable isotope ratio analysis. *International dairy journal, 17*(5), 421-428.

Crittenden, R. G., Andrew, A. S., LeFournour, M., Young, M. D., Middleton, H., & Stockmann, R. (2007). Determining the geographic origin of milk in Australasia using multi-element stable isotope ratio analysis. *International dairy journal, 17*(5), 421-428.

Cubero-Leon, E., Peñalver, R., & Maquet, A. (2014). Review on metabolomics for food authentication. *Food Research International, 60*(0), 95-107.

da Rocha, R. A., Paiva, I. M., Anjos, V., Furtado, M. A. M., & Bell, M. J. V. (2015). Quantification of whey in fluid milk using confocal Raman microscopy and artificial neural network. *Journal of dairy science, 98*(6), 3559-3567.

de la Fuente, M. A., & Juarez, M. (2005). Authenticity assessment of dairy products. *Crit Rev Food Sci Nutr, 45*(7-8), 563-585.

de la Guardia, M., & Illueca, A. G. (2013). *Food protected designation of origin: methodologies and applications* (Vol. 60): Newnes.

De Luca, M., Santonico, M., Pennazza, G., & Iarossi, S. (2014). Ultrasound Based Sensor for Fat Detection in Fresh Milk *Sensors* (pp. 499-502): Springer.

Dervisoglu, M., Gul, O., Yazici, F., Guvenc, D., Atmaca, E., & Aksoy, A. (2014). Toxic and essential elements in butter from the Black Sea region, Turkey. *Food Addit Contam Part B Surveill, 7*(1), 49-53.

Di Paola-Naranjo, R. D., Baroni, M. V., Podio, N. S., Rubinstein, H. R., Fabani, M. P., Badini, R. G., Wunderlin, D. A. (2011). Fingerprints for main varieties of

argentinean wines: terroir differentiation by inorganic, organic, and stable isotopic analyses coupled to chemometrics. *J Agric Food Chem, 59*(14), 7854-7865.

Diomande, D., Antheaume, I., Leroux, M., Lalande, J., Balayssac, S., Remaud, G. S., & Tea, I. (2015). Multi-element, multi-compound isotope profiling as a means to distinguish the geographical and varietal origin of fermented cocoa (Theobroma cacao L.) beans. *Food Chemistry, 188*, 576-582.

Domingo, E., Tirelli, A. A., Nunes, C. A., Guerreiro, M. C., & Pinto, S. M. (2014). Melamine detection in milk using vibrational spectroscopy and chemometrics analysis: A review. *Food Research International, 60*(0), 131-139.

dos Santos, L. G. C., De Nadai Fernandes, E. A., Bacchi, M. A., Sarriés, G. A., Blumer, L., & Júnior, F. B. (2009). Chemical composition of bovine milk from Minas Gerais State, Brazil. *Journal of Radioanalytical and Nuclear Chemistry, 282*(2), 493-496.

Drivelos, S. A., & Georgiou, C. A. (2012). Multi-element and multi-isotope-ratio analysis to determine the geographical origin of foods in the European Union. *TrAC Trends in Analytical Chemistry, 40*, 38-51.

Ehtesham, E., Baisden, W., Keller, E., Hayman, A., Van Hale, R., & Frew, R. (2013). Correlation between precipitation and geographical location of the δ 2 H values of the fatty acids in milk and bulk milk powder. *Geochimica et Cosmochimica Acta, 111*, 105-116.

Ehtesham, E., Hayman, A., Van Hale, R., & Frew, R. (2015). Influence of feed and water on the stable isotopic composition of dairy milk. *International dairy journal, 47*, 37-45.

Elflein, L., & Raezke, K.-P. (2008). Improved detection of honey adulteration by measuring differences between 13C/12C stable carbon isotope ratios of protein and sugar compounds with a combination of elemental analyzer—isotope ratio mass spectrometry and liquid chromatography—isotope ratio mass spectrometry (δ13C-EA/LC-IRMS). *Apidologie, 39*(5), 574-587.

Elgersma, A., Tamminga, S., & Ellen, G. (2006). Modifying milk composition through forage. *Animal Feed Science and Technology, 131*(3-4), 207-225.

Elmer, P. (2001). The 30-minute Guide to ICP-MS. *Perkin Elmer, Shelton CT*.

Erich, S., Schill, S., Annweiler, E., Waiblinger, H.-U., Kuballa, T., Lachenmeier, D. W., & Monakhova, Y. B. (2015). Combined chemometric analysis of 1 H NMR, 13 C NMR and stable isotope data to differentiate organic and conventional milk. *Food Chemistry, 188*, 1-7.

Fallah, A. A., Saei-Dehkordi, S. S., Nematollahi, A., & Jafari, T. (2011). Comparative study of heavy metal and trace element accumulation in edible tissues of farmed and wild rainbow trout (Oncorhynchus mykiss) using ICP-OES technique. *Microchemical Journal, 98*(2), 275-279.

Fausett, L. (1994). Fundamentals of neural networks: architectures, algorithms, and applications.

Feng, X.-d., Su, R., Xu, N., Wang, X.-h., Yu, A.-m., Zhang, H.-q., & Cao, Y.-b. (2013). Portable analyzer for rapid analysis of total protein, fat and lactose contents in raw milk measured by non-dispersive short-wave near-infrared spectrometry. *Chemical Research in Chinese Universities, 29*(1), 15-19.

Forcato, D., Carmine, M., Echeverria, G., Pécora, R., & Kivatinitz, S. (2005). Milk fat content measurement by a simple UV spectrophotometric method: An alternative screening method. *Journal of dairy science, 88*(2), 478-481.

Forcato, D. O., Carmine, M. P., Echeverría, G. E., Pécora, R. P., & Kivatinitz, S. C. (2005). Milk Fat Content Measurement by a Simple UV Spectrophotometric Method: An Alternative Screening Method. *Journal of Dairy Science, 88*(2), 478-481.

Fraga, C. G. (2005). Relevance, essentiality and toxicity of trace elements in human health. *Mol Aspects Med, 26*(4-5), 235-244.

Gabbi, A., McManus, C., Silva, A., Marques, L., Zanela, M., Stumpf, M., & Fischer, V. (2013). Typology and physical–chemical characterization of bovine milk produced with different productions strategies. *Agricultural Systems, 121*, 130-134.

Galeano Diaz, T., Durán Merás, I., Sánchez Casas, J., & Alexandre Franco, M. F. (2005). Characterization of virgin olive oils according to its triglycerides and sterols composition by chemometric methods. *Food Control, 16*(4), 339-347.

Giannenas, I., Nisianakis, P., Gavriil, A., Kontopidis, G., & Kyriazakis, I. (2009). Trace mineral content of conventional, organic and courtyard eggs analysed by inductively coupled plasma mass spectrometry (ICP-MS). *Food Chemistry, 114*(2), 706-711.

Gnanadesikan, R. (2011). *Methods for statistical data analysis of multivariate observations* (Vol. 321): John Wiley & Sons.

Gonzalez-Montana, J. R., Senis, E., Gutierrez, A., & Prieto, F. (2012). Cadmium and lead in bovine milk in the mining area of the Caudal River (Spain). *Environmental Monitoring and Assessment, 184*(7), 4029-4034.

Gonzalez-Weller, D., Karlsson, L., Caballero, A., Hernandez, F., Gutierrez, A., Gonzalez-Iglesias, T., Hardisson, A. (2006). Lead and cadmium in meat and meat products consumed by the population in Tenerife Island, Spain. *Food Addit Contam, 23*(8), 757-763.

Gonzalvez, A., Armenta, S., & de la Guardia, M. (2009). Trace-element composition and stable-isotope ratio for discrimination of foods with Protected Designation of Origin. *TrAC Trends in Analytical Chemistry, 28*(11), 1295-1311.

Grace, J. (2014). changes in nitrogen content and isotopic composition in subarctic Empetrum nigrum seeds in the period 1976–2010. *Boreal environment research, 19*, 209-221.

Granato, D., Katayama, F., & Castro, I. (2010). Assessing the association between phenolic compounds and the antioxidant activity of Brazilian red wines using chemometrics. *Lwt-Food Science and Technology, 43*(10), 1542-1549.

Gray, A. L., & Date, A. R. (1983). Inductively coupled plasma source mass spectrometry using continuum flow ion extraction. *The Analyst, 108*(1290), 1033.

Gray, J., & Thompson, P. (1976). Climatic information from 18O/16O ratios of cellulose in tree rings. *Nature, 262*, 481-482.

Guerreiro, J. S., Barros, M., Fernandes, P., Pires, P., & Bardsley, R. (2013). Principal component analysis of proteolytic profiles as markers of authenticity of PDO cheeses. *Food Chemistry, 136*(3), 1526-1532.

Guler, A., Kocaokutgen, H., Garipoglu, A. V., Onder, H., Ekinci, D., & Biyik, S. (2014). Detection of adulterated honey produced by honeybee (Apis mellifera L.) colonies fed with different levels of commercial industrial sugar (C 3 and C 4 plants) syrups by the carbon isotope ratio analysis. *Food Chemistry, 155*, 155-160.

Guo, B., Wei, Y., Pan, J., & Li, Y. (2010). Stable C and N isotope ratio analysis for regional geographical traceability of cattle in China. *Food Chemistry, 118*(4), 915-920.

Guo, J., Yue, T., Yuan, Y., & Wang, Y. (2013). Chemometric classification of apple juices according to variety and geographical origin based on polyphenolic profiles. *J Agric Food Chem, 61*(28), 6949-6963.

Härdle, W. K., & Simar, L. (2012). *Applied multivariate statistical analysis*: Springer Science & Business Media.

Hawke, J., & Taylor, M. (1983). Influence of nutritional factors on the yield, composition and physical properties of milk fat *Developments in Dairy Chemistry—2* (pp. 37-81): Springer.

Hayes, J. M., Freeman, K. H., Popp, B. N., & Hoham, C. H. (1990). Compound-specific isotopic analyses: a novel tool for reconstruction of ancient biogeochemical processes. *Org Geochem, 16*(4-6), 1115-1128.

Henton, E., MCorriston, J., Martin, L., & Oches, E. A. (2014). Seasonal aggregation and ritual slaughter: Isotopic and dental microwear evidence for cattle herder mobility in the Arabian Neolithic. *Journal of Anthropological Archaeology, 33*, 119-131.

Hilding-Ohlsson, A., Fauerbach, J. A., Sacco, N. J., Bonetto, M. C., & Corton, E. (2012). Voltamperometric discrimination of urea and melamine adulterated

skimmed milk powder. *Sensors (Basel), 12*(9), 12220-12234. doi:10.3390/s120912220

Houk, R. S. (1986). Mass spectrometry of inductively coupled plasmas. *Analytical Chemistry, 58*(1), 97A-105A. doi:10.1021/ac00292a003

Houk, R. S., Fassel, V. A., Flesch, G. D., Svec, H. J., Gray, A. L., & Taylor, C. E. (1980). Inductively coupled argon plasma as an ion source for mass spectrometric determination of trace elements. *Analytical Chemistry, 52*(14), 2283-2289.

Hrbek, V., Vaclavik, L., Elich, O., & Hajslova, J. (2014). Authentication of milk and milk-based foods by direct analysis in real time ionization–high resolution mass spectrometry (DART–HRMS) technique: A critical assessment. *Food Control, 36*(1), 138-145.

Hsieh, B.-T., Chang, C.-Y., Chang, Y.-C., & Cheng, K.-Y. (2011). Relationship between the level of essential metal elements in human hair and coronary heart disease. *Journal of Radioanalytical and Nuclear Chemistry, 290*(1), 165-169. doi:10.1007/s10967-011-1174-z

Huang, M., Kim, M. S., Chao, K., Qin, J., Mo, C., Esquerre, C., Zhu, Q. (2016). Penetration Depth Measurement of Near-Infrared Hyperspectral Imaging Light for Milk Powder. *Sensors, 16*(4), 441.

Huck-Pezzei, V. A., Seitz, I., Karer, R., Schmutzler, M., De Benedictis, L., Wild, B., & Huck, C. W. (2014). Alps food authentication, typicality and intrinsic quality by near infrared spectroscopy. *Food Research International, 62*, 984-990.

Jeon, H., Lee, S.-C., Cho, Y.-J., Oh, J.-H., Kwon, K., & Kim, B. H. (2015). A triple-isotope approach for discriminating the geographic origin of Asian sesame oils. *Food Chemistry, 167*, 363-369.

Jiang, J., Lu, S., Zhang, H., Liu, G., Lin, K., Huang, W., Yu, Y. (2015). Dietary intake of human essential elements from a Total Diet Study in Shenzhen, Guangdong Province, China. *Journal of Food Composition and Analysis, 39*(0), 1-7.

Joint, F. A. O. W. H. O. E. C. o. F. A., Bend, J., Bolger, M., Knaap, A. G., Kuznesof, P. M., Larsen, J. C., . . . Williams, G. M. (2007). Evaluation of certain food additives and contaminants. *World Health Organization technical report series*(947), 1-225, back cover.

Kalač, P., & Samková, E. (2010). The effects of feeding various forages on fatty acid composition of bovine milk fat: A review. *Czech J. Anim. Sci, 55*(12), 521-537.

Kamal, M., & Karoui, R. (2015). Analytical methods coupled with chemometric tools for determining the authenticity and detecting the adulteration of dairy products: A review. *Trends in Food Science & Technology, 46*(1), 27-48.

Kamizake, N. K., Gonçalves, M. M., Zaia, C. T., & Zaia, D. A. (2003). Determination of total proteins in cow milk powder samples: a comparative study between the

Kjeldahl method and spectrophotometric methods. *Journal of Food Composition and Analysis, 16*(4), 507-516.

Karabagias, I., Michos, C., Badeka, A., Kontakos, S., Stratis, I., & Kontominas, M. G. (2013). Classification of Western Greek virgin olive oils according to geographical origin based on chromatographic, spectroscopic, conventional and chemometric analyses. *Food Research International, 54*(2), 1950-1958.

Kartheek, M., Smith, A. A., Muthu, A. K., & Manavalan, R. (2011). Determination of adulterants in food: A review. *Journal of Chemical and Pharmaceutical Research, 3*(2), 629-636.

Khan, N., Choi, J. Y., Nho, E. Y., Hwang, I. M., Habte, G., Khan, M. A., . . . Kim, K. S. (2014). Determination of Mineral Elements in Milk Products by Inductively Coupled Plasma-Optical Emission Spectrometry. *Analytical Letters, 47*(9), 1606-1613.

Khan, N., Jeong, I. S., Hwang, I. M., Kim, J. S., Choi, S. H., Nho, E. Y., . . . Kim, K. S. (2014). Analysis of minor and trace elements in milk and yogurts by inductively coupled plasma-mass spectrometry (ICP-MS). *Food Chem, 147*, 220-224.

Khanmohammadi, M., Karami, F., Mir-Marqués, A., Bagheri Garmarudi, A., Garrigues, S., & de la Guardia, M. (2014). Classification of persimmon fruit origin by near infrared spectrometry and least squares-support vector machines. *Journal of Food Engineering, 142*, 17-22.

Kim, H., Suresh Kumar, K., & Shin, K. H. (2015). Applicability of stable C and N isotope analysis in inferring the geographical origin and authentication of commercial fish (Mackerel, Yellow Croaker and Pollock). *Food Chem, 172*, 523-527.

Kira, C. S., & Maihara, V. A. (2007). Determination of major and minor elements in dairy products through inductively coupled plasma optical emission spectrometry after wet partial digestion and neutron activation analysis. *Food Chemistry, 100*(1), 390-395.

Komatsu, F., Kagawa, Y., Kawabata, T., Kaneko, Y., Kudoh, H., Purvee, B., Chimedregzen, U. (2012). Influence of essential trace minerals and micronutrient insufficiencies on harmful metal overload in a Mongolian patient with multiple sclerosis. *Curr Aging Sci, 5*(2), 112-125.

Kropf, U., Korošec, M., Bertoncelj, J., Ogrinc, N., Nečemer, M., Kump, P., & Golob, T. (2010). Determination of the geographical origin of Slovenian black locust, lime and chestnut honey. *Food Chemistry, 121*(3), 839-846.

Kruzlicova, D., Fiket, Ž., & Kniewald, G. (2013). Classification of Croatian wine varieties using multivariate analysis of data obtained by high resolution ICP-MS analysis. *Food Research International, 54*(1), 621-626.

Lachas, H., Richaud, R., Herod, A., Dugwell, D., & Kandiyoti, R. (2000). Determination of trace elements by ICP-MS of biomass and fuel oil reference

materials using milligram sample sizes. *Rapid communications in mass spectrometry, 14*(5), 335-343.

Laporte, M.-F., & Paquin, P. (1999). Near-infrared analysis of fat, protein, and casein in cow's milk. *Journal of agricultural and food chemistry, 47*(7), 2600-2605.

Larrea-Marín, M. T., Pomares-Alfonso, M. S., Gómez-Juaristi, M., Sánchez-Muniz, F. J., & de la Rocha, S. R. (2010). Validation of an ICP-OES method for macro and trace element determination in Laminaria and Porphyra seaweeds from four different countries. *Journal of Food Composition and Analysis, 23*(8), 814-820.

Lei, Y., Zhou, Q., Zhang, Y.-l., Chen, J.-b., Sun, S.-q., & Noda, I. (2010). Analysis of crystallized lactose in milk powder by Fourier-transform infrared spectroscopy combined with two-dimensional correlation infrared spectroscopy. *Journal of Molecular Structure, 974*(1–3), 88-93.

Li, C., Yang, S.-C., Guo, Q.-S., Zheng, K.-Y., Shi, Y.-F., Xiao, X.-F., & Long, G.-Q. (2014). Determining the geographical origin of the medicinal plant Marsdenia tenacissima with multi-element analysis and data mining techniques. *Chemometrics and Intelligent Laboratory Systems, 136*, 115-120.

Licata, P., Di Bella, G., Potorti, A. G., Lo Turco, V., Salvo, A., & Dugo, G. M. (2012). Determination of trace elements in goat and ovine milk from Calabria (Italy) by ICP-AES. *Food Addit Contam Part B Surveill, 5*(4), 268-271.

Liu, D., Zeng, X.-A., & Sun, D.-W. (2013). NIR spectroscopy and imaging techniques for evaluation of fish quality—a review. *Applied Spectroscopy Reviews, 48*(8), 609-628.

Liu, R., Wu, P., Yang, L., Hou, X., & Lv, Y. (2014). Inductively coupled plasma mass spectrometry-based immunoassay: a review. *Mass Spectrom Rev, 33*(5), 373-393.

Liu, X., Guo, B., Wei, Y., Shi, J., & Sun, S. (2013). Stable isotope analysis of cattle tail hair: A potential tool for verifying the geographical origin of beef. *Food Chemistry, 140*(1), 135-140.

Longobardi, F., Casiello, G., Cortese, M., Perini, M., Camin, F., Catucci, L., & Agostiano, A. (2015). Discrimination of geographical origin of lentils (Lens culinaris Medik.) using isotope ratio mass spectrometry combined with chemometrics. *Food Chemistry, 188*, 343-349.

Longobardi, F., Casiello, G., Cortese, M., Perini, M., Camin, F., Catucci, L., & Agostiano, A. (2015). Discrimination of geographical origin of lentils (Lens culinaris Medik.) using isotope ratio mass spectrometry combined with chemometrics. *Food Chem, 188*, 343-349.

Longobardi, F., Casiello, G., Sacco, D., Tedone, L., & Sacco, A. (2011). Characterisation of the geographical origin of Italian potatoes, based on stable isotope and volatile compound analyses. *Food Chemistry, 124*(4), 1708-1713.

Luis, G., Rubio, C., Revert, C., Espinosa, A., González-Weller, D., Gutiérrez, A. J., & Hardisson, A. (2015). Dietary intake of metals from yogurts analyzed by inductively coupled plasma optical emission spectrometry (ICP-OES). *Journal of Food Composition and Analysis, 39*(0), 48-54.

Luo, D., Dong, H., Luo, H., Xian, Y., Guo, X., & Wu, Y. (2015). Multi-Element (C, N, H, O) Stable Isotope Ratio Analysis for Determining the Geographical Origin of Pure Milk from Different Regions. *Food Analytical Methods*, 1-6.

Luo, D., Dong, H., Luo, H., Xian, Y., Wan, J., Guo, X., & Wu, Y. (2015). The application of stable isotope ratio analysis to determine the geographical origin of wheat. *Food Chem, 174*, 197-201.

Lüthi-Peng, Q., & Puhan, Z. (1999). Determination of protein and casein in milk by fourth derivative UV spectrophotometry. *Analytica chimica acta, 393*(1), 227-234.

Luykx, D. M., & Van Ruth, S. M. (2008). An overview of analytical methods for determining the geographical origin of food products. *Food Chemistry, 107*(2), 897-911.

Maçatelli, M., Akkermans, W., Koot, A., Buchgraber, M., Paterson, A., & van Ruth, S. (2009). Verification of the geographical origin of European butters using PTR-MS. *Journal of Food Composition and Analysis, 22*(2), 169-175.

Magdas, D.-A., Dehelean, A., Feher, I., Cristea, G., Puscas, R., Dan, S.-D., & Cordea, D.-V. Discrimination markers for the geographical and species origin of raw milk within Romania. *International dairy journal*.

Magdas, D.-A., Dehelean, A., Feher, I., Cristea, G., Puscas, R., Dan, S.-D., & Cordea, D.-V. (2016). Discrimination markers for the geographical and species origin of raw milk within Romania. *International dairy journal, 61*, 135-141.

Magdas, D. A., & Puscas, R. (2011). Stable isotopes determination in some Romanian fruit juices. *Isotopes in environmental and health studies, 47*(3), 372-378.

Malhat, F., Hagag, M., Saber, A., & Fayz, A. E. (2012). Contamination of cows milk by heavy metal in Egypt. *Bull Environ Contam Toxicol, 88*(4), 611-613.

Manning, L., & Soon, J. M. (2014). Developing systems to control food adulteration. *Food Policy, 49*, 23-32.

Massart, D. L., Vandeginste, B. G., Buydens, L., Lewi, P., & Smeyers-Verbeke, J. (1997). *Handbook of chemometrics and qualimetrics: Part A*: Elsevier Science Inc.

May, T. W., & Wiedmeyer, R. H. (1998). A table of polyatomic interferences in ICP-MS. *ATOMIC SPECTROSCOPY-NORWALK CONNECTICUT-, 19*, 150-155.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics, 5*(4), 115-133.

Miller-Ihli, N. J., & Baker, S. A. (2001). Trace Element Composition of Municipal Waters in the United States: A Comparison of ICP-AES and ICP-MS Methods. *Journal of Food Composition and Analysis, 14*(6), 619-629. doi:10.1006/jfca.2001.1024

Millour, S., Noël, L., Kadar, A., Chekri, R., Vastel, C., & Guérin, T. (2011). Simultaneous analysis of 21 elements in foodstuffs by ICP-MS after closed-vessel microwave digestion: Method validation. *Journal of Food Composition and Analysis, 24*(1), 111-120.

Miranda, M., Lopez-Alonso, M., Castillo, C., Hernandez, J., & Benedito, J. L. (2005). Effects of moderate pollution on toxic and trace metal levels in calves from a polluted area of northern Spain. *Environ Int, 31*(4), 543-548.

Molkentin, J. (2013). Applicability of organic milk indicators to the authentication of processed products. *Food Chem, 137*(1-4), 25-30.

Molkentin, J., & Giesemann, A. (2007). Differentiation of organically and conventionally produced milk by stable isotope and fatty acid analysis. *Anal Bioanal Chem, 388*(1), 297-305.

Molkentin, J., & Giesemann, A. (2010). Follow-up of stable isotope analysis of organic versus conventional milk. *Analytical and Bioanalytical Chemistry, 398*(3), 1493-1500.

Monakhova, Y. B., Godelmann, R., Andlauer, C., Kuballa, T., & Lachenmeier, D. W. (2013). Identification of imitation cheese and imitation ice cream based on vegetable fat using NMR spectroscopy and chemometrics. *International Journal of Food Science, 2013*.

Monfreda, M. (2012). *Principal Component Analysis: A Powerful Interpretative Tool at the Service of Analytical Methodology*: INTECH Open Access Publisher.

Moore, J. C., Spink, J., & Lipp, M. (2012). Development and application of a database of food ingredient fraud and economically motivated adulteration from 1980 to 2010. *J Food Sci, 77*(4), R118-126.

Muñiz-Valencia, R., Jurado, J. M., Ceballos-Magaña, S. G., Alcázar, Á., & Hernández-Díaz, J. (2014). Characterization of Mexican coffee according to mineral contents by means of multilayer perceptrons artificial neural networks. *Journal of Food Composition and Analysis, 34*(1), 7-11. doi:http://dx.

Nagpal, R., Behare, P. V., Kumar, M., Mohania, D., Yadav, M., Jain, S.,Yadav, H. (2012). Milk, milk products, and disease free health: an updated overview. *Crit Rev Food Sci Nutr, 52*(4), 321-333.

Nardi, E. P., Evangelista, F. S., Tormen, L., Saint´Pierre, T. D., Curtius, A. J., Souza, S. S. d., & Barbosa, F. (2009). The use of inductively coupled plasma mass spectrometry (ICP-MS) for the determination of toxic and essential elements in different types of food samples. *Food Chemistry, 112*(3), 727-732.

Nóbrega, J. A., Gélinas, Y., Krushevska, A., & Barnes, R. M. (1997). Direct determination of major and trace elements in milk by inductively coupled plasma atomic emission and mass spectrometry. *Journal of Analytical Atomic Spectrometry, 12*(10), 1243-1246.

Nunez-Sanchez, N., Martinez-Marin, A. L., Polvillo, O., Fernandez-Cabanas, V. M., Carrizosa, J., Urrutia, B., & Serradilla, J. M. (2016). Near Infrared Spectroscopy (NIRS) for the determination of the milk fat fatty acid profile of goats. *Food Chem, 190*(0), 244-252.

O'Brien, B., Mehra, R., Connolly, J., & Harrington, D. (1999). Seasonal variation in the composition of Irish manufacturing and retail milks: 4. Minerals and trace elements. *Irish journal of agricultural and food research*, 87-99.

Oliveira, M., Ramos, S., Delerue-Matos, C., & Morais, S. (2015). Espresso beverages of pure origin coffee: mineral characterization, contribution for mineral intake and geographical discrimination. *Food Chem, 177*, 330-338.

Olszewski, T., Ryniecki, A., & Boniecki, P. (2008). Neural network development for automatic identification of the endpoint of drying barley in bulk. *Journal of Research and Applications in Agricultural Engineering, 53*(1), 26-31.

Organization, W. H. (1996). Trace elements in human nutrition and health.

Orun, E., Yalcin, S. S., Aykut, O., Orhan, G., Morgil, G. K., Yurdakok, K., & Uzun, R. (2011). Breast milk lead and cadmium levels from suburban areas of Ankara. *Sci Total Environ, 409*(13), 2467-2472.

Osborne, B. G., & Fearn, T. (1986). *Near infrared spectroscopy in food analysis*: Longman.

Osorio, M. T., Koidis, A., & Papademas, P. (2015). Major and trace elements in milk and Halloumi cheese as markers for authentication of goat feeding regimes and geographical origin. *International Journal of Dairy Technology*.

Palacios-Morillo, A., Alcázar, Á., de Pablos, F., & Jurado, J. M. (2013). Differentiation of tea varieties using UV–Vis spectra and pattern recognition techniques. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 103*, 79-83.

Palacios-Morillo, A., Jurado, J. M., Alcázar, Á., & de Pablos, F. (2014). Geographical characterization of Spanish PDO paprika by multivariate analysis of multielemental content. *Talanta, 128*, 15-22.

Patra, A. K., Wagh, S., Jain, A., & Hegde, A. (2010). Assessment of daily intake of trace elements by Kakrapar adult population through ingestion pathway. *Environmental Monitoring and Assessment, 169*(1-4), 267-272.

Pellerano, R. G., Uñates, M. A., Cantarelli, M. A., Camiña, J. M., & Marchevsky, E. J. (2012). Analysis of trace elements in multifloral Argentine honeys and their classification according to provenance. *Food Chemistry, 134*(1), 578-582.

Pereira, P. C. (2014). Milk nutritional composition and its role in human health. *Nutrition, 30*(6), 619-627.

Pilarczyk, R., Wojcik, J., Czerniak, P., Sablik, P., Pilarczyk, B., & Tomza-Marciniak, A. (2013). Concentrations of toxic heavy metals and trace elements in raw milk of Simmental and Holstein-Friesian cows from organic farm. *Environmental Monitoring and Assessment, 185*(10), 8383-8392.

Podio, N. S., Baroni, M. V., Badini, R. G., Inga, M., Ostera, H. A., Cagnoni, M., Wunderlin, D. A. (2013). Elemental and isotopic fingerprint of Argentinean wheat. Matching soil, water, and crop composition to differentiate provenance. *J Agric Food Chem, 61*(16), 3763-3773.

Potortì, A. G., Di Bella, G., Lo Turco, V., Rando, R., & Dugo, G. (2013). Non-toxic and potentially toxic elements in Italian donkey milk by ICP-MS and multivariate analysis. *Journal of Food Composition and Analysis, 31*(1), 161-172.

Profrock, D., & Prange, A. (2012). Inductively coupled plasma-mass spectrometry (ICP-MS) for quantitative analysis in environmental and life sciences: a review of challenges, solutions, and trends. *Appl Spectrosc, 66*(8), 843-868. doi:10.1366/12-06681

Raty, J. A., & Peiponen, K.-E. (1999). Reflectance study of milk in the UV-visible range. *Applied Spectroscopy, 53*(9), 1123-1127.

Rayman, M. P. (2000). The importance of selenium to human health. *Lancet, 356*(9225), 233-241.

Rayman, M. P. (2008). Food-chain selenium and human health: emphasis on intake. *Br J Nutr, 100*(2), 254-268.

Rebechi, S. R., Velez, M. A., Vaira, S., & Perotti, M. C. (2016). Adulteration of Argentinean milk fats with animal fats: Detection by fatty acids analysis and multivariate regression techniques. *Food Chem, 192*, 1025-1032.

Rees, G., Kelly, S. D., Cairns, P., Ueckermann, H., Hoelzl, S., Rossmann, A., & Scotter, M. J. (2016). Verifying the geographical origin of poultry: The application of stable isotope and trace element (SITE) analysis. *Food Control, 67*, 144-154.

Reinholds, I., Bartkevics, V., Silvis, I. C. J., van Ruth, S. M., & Esslinger, S. (2015). Analytical techniques combined with chemometrics for authentication and determination of contaminants in condiments: A review. *Journal of Food Composition and Analysis, 44*(0), 56-72.

Renou, J.-P., Bielicki, G., Deponge, C., Gachon, P., Micol, D., & Ritz, P. (2004). Characterization of animal products according to geographic origin and feeding diet using nuclear magnetic resonance and isotope ratio mass spectrometry. Part II: Beef meat. *Food Chemistry, 86*(2), 251-256.

Rey-Crespo, F., Miranda, M., & López-Alonso, M. (2013). Essential trace and toxic element concentrations in organic and conventional milk in NW Spain. *Food and Chemical Toxicology, 55*(0), 513-518.

Ripley, B. D. (2007). *Pattern recognition and neural networks*: Cambridge university press.

Ródenas de la Rocha, S., Sánchez-Muniz, F. J., Gómez-Juaristi, M., & Marín, M. T. L. (2009). Trace elements determination in edible seaweeds by an optimized and validated ICP-MS method. *Journal of Food Composition and Analysis, 22*(4), 330-336.

Rodriguez, E. M. R., Alaejos, M. S., & Romero, C. D. (2001). Mineral concentrations in cow's milk from the Canary Island. *Journal of Food Composition and Analysis, 14*(4), 419-430.

Rozanski, K., Araguás-Araguás, L., & Gonfiantini, R. (1992). Relation between long-term trends of oxygen-18 isotope composition of precipitation and climate. *Science, 258*(5084), 981-985.

Rutkowska, J., Bialek, M., Adamska, A., & Zbikowska, A. (2015). Differentiation of geographical origin of cream products in Poland according to their fatty acid profile. *Food Chem, 178*, 26-31.

Sacco, D., Brescia, M., Sgaramella, A., Casiello, G., Buccolieri, A., Ogrinc, N., & Sacco, A. (2009). Discrimination between Southern Italy and foreign milk samples using spectroscopic and analytical data. *Food Chemistry, 114*(4), 1559-1563.

Saei-Dehkordi, S. S., & Fallah, A. A. (2011). Determination of copper, lead, cadmium and zinc content in commercially valuable fish species from the Persian Gulf using derivative potentiometric stripping analysis. *Microchemical Journal, 98*(1), 156-162.

Sall, J., Lehman, A., Stephens, M. L., & Creighton, L. (2012). *JMP start statistics: a guide to statistics and data analysis using JMP*: SAS Institute.

Šašic, S., & Ozaki, Y. (2001). Short-wave near-infrared spectroscopy of biological fluids. 1. Quantitative analysis of fat, protein, and lactose in raw milk by partial least-squares regression and band assignment. *Analytical Chemistry, 73*(1), 64-71.

Scampicchio, M., Eisenstecken, D., De Benedictis, L., Capici, C., Ballabio, D., Mimmo, T., Kaser, A. (2015). Multi-method Approach to Trace the Geographical Origin of Alpine Milk: a Case Study of Tyrol Region. *Food Analytical Methods*, 1-12.

Scampicchio, M., Eisenstecken, D., De Benedictis, L., Capici, C., Ballabio, D., Mimmo, T., Kaser, A. (2016). Multi-method Approach to Trace the Geographical Origin of Alpine Milk: a Case Study of Tyrol Region. *Food Analytical Methods, 9*(5), 1262-1273.

Scampicchio, M., Mimmo, T., Capici, C., Huck, C., Innocente, N., Drusch, S., & Cesco, S. (2012). Identification of milk origin and process-induced changes in milk by stable isotope ratio mass spectrometry. *J Agric Food Chem, 60*(45), 11268-11273.

Schroeder, H. A., Balassa, J. J., & Tipton, I. H. (1966). Essential trace metals in man: manganese. A study in homeostasis. *J Chronic Dis, 19*(5), 545-571.

Schuhmacher, M., Bosque, M. A., Domingo, J. L., & Corbella, J. (1991). Dietary intake of lead and cadmium from foods in Tarragons Province, Spain. *Bulletin of environmental contamination and toxicology, 46*(2), 320-328.

Schwertl, M., Auerswald, K., Schäufele, R., & Schnyder, H. (2005). Carbon and nitrogen stable isotope composition of cattle hair: ecological fingerprints of production systems? *Agriculture, ecosystems & environment, 109*(1), 153-165.

Shahbazi, Y., Ahmadi, F., & Fakhari, F. (2016). Voltammetric determination of Pb, Cd, Zn, Cu and Se in milk and dairy products collected from Iran: An emphasis on permissible limits and risk assessment of exposure to heavy metals. *Food Chem, 192*, 1060-1067. doi:10.1016/j.foodchem.2015.07.123

Simsek, O., Gültekin, R., Öksüz, O., & Kurultay, S. (2000). The effect of environmental pollution on the heavy metal content of raw milk. *Food / Nahrung, 44*(5), 360-363.

Siripatrawan, U., & Harte, B. R. (2007). Solid phase microextraction/gas chromatography/mass spectrometry integrated with chemometrics for detection of Salmonella typhimurium contamination in a packaged fresh vegetable. *Analytica chimica acta, 581*(1), 63-70.

Soares, V. A., Kus, M. M. M., Peixoto, A. L. C., Carrocci, J. S., Salazar, R. F. S., & Izário Filho, H. J. (2010). Determination of nutritional and toxic elements in pasteurized bovine milk from Vale do Paraiba region (Brazil). *Food Control, 21*(1), 45-49.

Sola-Larrañaga, C., & Navarro-Blasco, I. (2009). Chemometric analysis of minerals and trace elements in raw cow milk from the community of Navarra, Spain. *Food Chemistry, 112*(1), 189-196.

Souza, S. S., Cruz, A. G., Walter, E. H., Faria, J. A., Celeghini, R. M., Ferreira, M. M., Sant'Ana, A. d. S. (2011). Monitoring the authenticity of Brazilian UHT milk: A chemometric approach. *Food Chemistry, 124*(2), 692-695.

Stawarz, R., Formicki, G., & Massanyi, P. (2007). Daily fluctuations and distribution of xenobiotics, nutritional and biogenic elements in human milk in Southern Poland. *J Environ Sci Health A Tox Hazard Subst Environ Eng, 42*(8), 1169-1175.

Suzuki, Y., Chikaraishi, Y., Ogawa, N. O., Ohkouchi, N., & Korenaga, T. (2008). Geographical origin of polished rice based on multiple element and stable isotope analyses. *Food Chemistry, 109*(2), 470-475.

Swarup, D., Patra, R., Naresh, R., Kumar, P., & Shekhar, P. (2005). Blood lead levels in lactating cows reared around polluted localities; transfer of lead into milk. *Science of the Total Environment, 349*(1), 67-71.

Thomas, R. (2001). Spectroscopy tutorial-A beginner's guide to ICP-MS-Part II: The sample-introduction system. *Spectroscopy, 16*(5), 56-+.

Thomas, R. (2013). *Practical guide to ICP-MS: a tutorial for beginners*: CRC press.

Trienekens, J., & Zuurbier, P. (2008). Quality and safety standards in the food industry, developments and challenges. *International Journal of Production Economics, 113*(1), 107-122.

Tripathi, R., Raghunath, R., Sastry, V., & Krishnamoorthy, T. (1999). Daily intake of heavy metals by infants through milk and milk products. *Science of the Total Environment, 227*(2), 229-235.

Tsenkova, R., Atanassova, S., Toyoda, K., Ozaki, Y., Itoh, K., & Fearn, T. (1999). Near-infrared spectroscopy for dairy management: measurement of unhomogenized milk composition. *Journal of dairy science, 82*(11), 2344-2351.

Tuzen, M., Sesli, E., & Soylak, M. (2007). Trace element levels of mushroom species from East Black Sea region of Turkey. *Food Control, 18*(7), 806-810.

Valenti, B., Martin, B., Andueza, D., Leroux, C., Labonne, C., Lahalle, F., Ferlay, A. (2013). Infrared spectroscopic methods for the discrimination of cows' milk according to the feeding system, cow breed and altitude of the dairy farm. *International dairy journal, 32*(1), 26-32.

van Leeuwen, K. A., Prenzler, P. D., Ryan, D., & Camin, F. (2014). Gas Chromatography-Combustion-Isotope Ratio Mass Spectrometry for Traceability and Authenticity in Foods and Beverages. *Comprehensive Reviews in Food Science and Food Safety, 13*(5), 814-837.

Van Ruth, S., Bremer, M., & Frankhuizen, R. (2009). Detection of Adulterations: Addition of Foreign Lipids and Proteins.

Vanhoe, H. (1993). A review of the capabilities of ICP-MS for trace element analysis in body fluids and tissues. *J Trace Elem Electrolytes Health Dis, 7*(3), 131-139.

Varela, P., & Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Research International, 48*(2), 893-908.

Vromman, V., Saegerman, C., Pussemier, L., Huyghebaert, A., De Temmerman, L., Pizzolon, J. C., & Waegeneers, N. (2008). Cadmium in the food chain near non-ferrous metal production sites. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess, 25*(3), 293-301.

Wang, M., Feng, W. Y., Zhao, Y. L., & Chai, Z. F. (2010). ICP-MS-based strategies for protein quantification. *Mass Spectrom Rev, 29*(2), 326-348.

Wang, S., Guo, Q., Wang, L., Lin, L., Shi, H., Cao, H., & Cao, B. (2015). Detection of honey adulteration with starch syrup by high performance liquid chromatography. *Food Chemistry, 172*, 669-674.

Workman Jr, J. J. (1996). Interpretive spectroscopy for near infrared. *Applied Spectroscopy Reviews, 31*(3), 251-320.

Wu, D., He, Y., Shi, J., & Feng, S. (2009). Exploring near and midinfrared spectroscopy to predict trace iron and zinc contents in powdered milk. *J Agric Food Chem, 57*(5), 1697-1704.

Wu, Y., Luo, D., Dong, H., Wan, J., Luo, H., Xian, Y., Wang, B. (2015). Geographical origin of cereal grains based on element analyser-stable isotope ratio mass spectrometry (EA-SIRMS). *Food Chem, 174*, 553-557.

Yanagi, Y., Hirooka, H., Oishi, K., Choumei, Y., Hata, H., Arai, M., Kumagai, H. (2012). Stable carbon and nitrogen isotope analysis as a tool for inferring beef cattle feeding systems in Japan. *Food Chemistry, 134*(1), 502-506.

Yang, C. J., Ding, W., Ma, L. J., & Jia, R. (2015). Discrimination and characterization of different intensities of goaty flavor in goat milk by means of an electronic nose. *J Dairy Sci, 98*(1), 55-67.

Yilmazcan, O., Erdemir, U. S., Izgi, B., Ozer, E. T., & Gucer, S. (2014). Lead Fractionation Analysis in Lipstick Samples by Inductively Coupled Plasma-Mass Spectrometry. *Spectroscopy Letters, 48*(4), 290-295.

Zamberlin, S., Antunac, N., Havranek, J., & Samarzija, D. (2012). Mineral elements in milk and dairy products. *Mljekarstvo, 62*(2), 111-125.

Zazzo, A., Harrison, S., Bahar, B., Moloney, A., Monahan, F., Scrimgeour, C., & Schmidt, O. (2007). Experimental determination of dietary carbon turnover in bovine hair and hoof. *Canadian Journal of Zoology, 85*(12), 1239-1248.

Zhang, A., Sun, H., Dou, S., Sun, W., Wu, X., Wang, P., & Wang, X. (2013). Metabolomics study on the hepatoprotective effect of scoparone using ultra-performance liquid chromatography/electrospray ionization quadruple time-of-flight mass spectrometry. *Analyst, 138*(1), 353-361.

Zhao, H., Guo, B., Wei, Y., & Zhang, B. (2013). Multi-element composition of wheat grain and provenance soil and their potentialities as fingerprints of geographical origin. *Journal of Cereal Science, 57*(3), 391-397.

Zhao, H., Guo, B., Wei, Y., Zhang, B., Sun, S., Zhang, L., & Yan, J. (2011). Determining the geographic origin of wheat using multielement analysis and multivariate statistics. *Journal of agricultural and food chemistry, 59*(9), 4397-4402.

Zhao, Y., Zhang, B., Chen, G., Chen, A., Yang, S., & Ye, Z. (2014). Recent developments in application of stable isotope analysis on agro-product authenticity and traceability. *Food Chem, 145*, 300-305.

Zheng, N., Wang, Q., Zhang, X., Zheng, D., Zhang, Z., & Zhang, S. (2007). Population health risk due to dietary intake of heavy metals in the industrial area of Huludao city, China. *Science of the Total Environment, 387*(1), 96-104.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

1. Zain, S.M., Behkami, S., Bakirdere, S. (2016). Milk authentication and discrimination via metal content clustering- A case of comparing milk from Malaysia and selected countries of the world. Food control, 66, 306-314. (*Impact of 3.38*)

2. Behkami, S.,Zain, S.M., Gholami, M., Bakirdere, S., (2017). Isotopic ratio analysis of cattle tail hair: A potential tool in building the database for cow milk geographical traceability. Food Chemistry, 217,438-444. (*Impact of 4.05*)

3. Gholami, M.,Behkami, S.,Zain, S.M.,& Bakirdere,S. (2016). A simple design for microwave assisted digestion vessel with low reagent consumption suitable for food and environmental samples. *Scientific Reports, 6*, 37186. (*Impact of 5.22*)

4. Behkami, S., Zain, S.M. Artificial neural network modeling based on UV-Vis/NIR spectral data: A method used to ascertain authenticity of freeze dried cow milk. (Under editor evaluation in food control)

5. Behkami, S., Zain, S.M. Single photodiode FT-NIR spectrometer used to verify protected designation of origin of solid state cow milk by artificial neural network modeling. (Under review in food control)

6. Behkami, S., Zain, S.M. Eelemental composition of tropical raw cow milk: A tool for determining its authenticity and traceability. (Under editor evaluation in food control)

7. Distribution of light isotopes of C,N, O and H in cow milk: A correlation with its regional level enviromental input (precipitation) and local variables (latitude). submitted

Conferences:

1. Behkami, S., Zain, S.M, Determination of geographical origin of milk in Malaysia using isotopic ratio mass spectroscopy and chemometric tools. International conference on chemical and medical science (CAM-2013) Dec. 29-30, 2013 Kuala Lumpur (Malaysia).

2. Behkami, S., Zain, S.M, Chemometric determination of geographical origin of milk samples in Malaysia. International science conference (Waset -2014) June 16-17 Toronto, (Canada).

3. Gholami, M., Behkami, S., Zain, S.M, As simple, precise and cost effective PTFE container design capable to work in domestic microwave oven. International science conference (Waset -2014) June 16-17 Toronto, (Canada).

4. Zain, S.M., Behkami, S. Tracing the geographical origin of cow milk using spectroscopic data and chemometric methods. International symposium on pure and applied chemistry (ISPAC-2016) August 15-18 Kuching, Sarawak, (Malaya). Submitted for presentation.