UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Agronomia Animali Alimenti Risorse Naturali e Ambiente

Second Cycle Degree (M.Sc.)

in

Biotechnologies for Food Science

Classification and prediction of fresh cheeses using NIR Spectroscopy

Supervisor
Prof. Giovanni Bittante

Co-supervisor
Dr. Nageshvar Patel
Department: DAFNAE

Student
Kristjana Naka

Matriculation
No.: 1172799

ACADEMIC YEAR 2019 – 2020

COPYRIGHT

The author and the promoter give permission to use this thesis for consultation and to copy parts of it for personal use. Any other use is subject to the Laws of Copyright. Permission to produce any material contained in this work should be obtained from the author.

ACKNOWLEDGEMENT

My appreciation goes to Prof. Giovanni Bittante for accepting to be the supervisor of this master thesis. I want to thank the professor for giving me the opportunity to carry out of the work related to this thesis and project, using the DAFNAE'S well equipped laboratory and facilities.

My deepest gratitude goes also to my thesis advisor, Nageshvar Patel, whose knowledge, advice and support helped me to gain an understanding of the project related to cheese quality and to complete the thesis successfully.

I am thankful to my family for all the support and encouragement in this period.

To all the people who believed in me, Thank you.

ABSTRACT

Cheese is a dairy product with a high nutritional value, playing an important role in food industry with an economic importance and consumed widely all over the globe. Nowadays is important to use different techniques to make sure that consumers get a quality product, encountering consumers requirements for value, health conditions and transparency.

Although, there are several techniques used for dairy product quality, there are still some limitations to those analytical techniques. Moreover, cheeses industry is constantly looking for fast, reliable and economical techniques to analyze the cheeses in order to control the qualities of cheeses. In recent past, near-infrared spectroscopy (NIRS) has gain quite popularity and become a reliable tool for fast analysis of cheeses from quality prediction to classification and authentication.

Therefore, the aim of our study was to study the potential use of NIR spectroscopy to classify two similar but differently produced fresh cheeses such as Paneer (acidic coagulation) and Tosella (enzymatic coagulation), produced by using both buffalo and cow milk. In addition to the classification, physical qualities of theses cheeses are also predicted using NIR spectroscopy, in order to see the potential and applicability of NIR spectroscopy in cheeses classification and quality prediction. Thus, we have produced 75 model cheeses (37 paneer and 38 tosella) in the lab and performed the scanning of cheeses using NIR instrument (Labspec, 2500) in the reflection mode in the wavelength range from 350 to 1830 nm. Physical qualities such as color, texture and cooking losses of cheeses also were performed.

Spectral data were pre-processed using SNV, Savitzky Golay first derivative and mean centering. Preprocessed spectra were decomposed by Principal Component Analysis (PCA) for qualitative analysis of spectra. Furthermore, cheeses were clustered and classified using unsupervised and supervised chemometric techniques. Hierarchical cluster analysis was used for clustering while linear discriminant analysis (LDA) based on principle component called PCA-DA and Soft independent

modelling of class analogies (SIMCA) were used as supervised techniques for classification of cheeses. In addition to the classification of cheeses, physical qualities of cheeses were predicted using Partial least square regression (PLSR) technique.

Results of PCA show that PC1, PC2 and PC3 together accounted for 85 % of total spectra variability. Clustering of four types of cheeses (TB, TC, PB and PB) by PCA was based in their differences in physical values but still the separation between cheeses was not very clear. HCA, based on the Ward method classified cheeses in terms of manufacturing process, coagulation conditions, cheese composition and origin of milk with a good clustering of cheeses based on their species and a blur separation based on their type. From the other side, PCA-DA model was used on the basis of PCA showing a good classification with 99.35 % of total variance and a good accuracy of calibration and prediction, 86 % and 91 % respectively. Also, SIMCA showed good values of classification and calibration around 89 % of safe. Among other cheeses TB was separated with the highest level of specificity and accuracy 100%. These results showed that PCADA and SIMCA could me more efficient than PCA and HCA for classification of cheeses type within the same milk used for its production.

PLSR was used for prediction of physicochemical qualities, reporting; coefficient of determination (R2) values in prediction ranging from 0.45 to 0.71 and root mean square errors of prediction (RMSEP) ranging from 0.62 to 1.45. In sum, it can be concluded that results of this study illustrate the potential use of this NIR spectroscopic technique as a useful and low-cost monitoring tool for quality of cheese analysis in routine laboratories. Moreover, this study also highlighted that in future some advance chemometrics tool such as artificial neural network, support vector machine etc. could be explored for better separation of cheeses and prediction of cheeses qualities within same species of milk.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVATIONS

CHAPTER 1: INTRODUCTION

Cheeses are the curd or substance formed by the coagulation of milk by acids or rennet or similar enzymes and they are one of the oldest and complex food products in the world (Settani and Moschetti, 2010). Cheesemaking is an ancient biotechnology that dates back to domestication of animals. It is thought that it was probably first discovered by observing the accidental souring of milk and then pressing and salting the solid curd to preserve it for later consumption (Powell and Limsowtin, 2016). With the time cheese production evolved in so many ways, in result to this, these days, cheeses are produced in a wide range of flavors, textures, and forms and their importance can be realized by the fact that among all the milk products, largest proportion ~ 30% of total milk production is used in cheese production worldwide (Eurostat, 2018).

Cheeses are mainly produced in three main stages: the coagulation of the milk which creates the curd, the expulsion of the whey and the maturing of the curd (Legg and Johnston 2017). Certainly, the characteristics and activity of starters, the variability among fundamental processing and aging processes influence both chemical and sensorial qualities of cheeses (Almena-Aliste and Mietton 2014). Different methods and practices produce different cheeses; therefore, cheeses can be classified based on several criteria. One of the important criteria for cheeses classification is based on the method of manufacturing the cheese and the length of time it is aged (McSweeney et al.,2017). Based on this method cheeses are grouped below.

Classifications of Cheese by Texture:

- Hard Grating Cheeses (Parmesan, Grana Padano, Sbrinz etc.)

- Firm/Hard (Emmental, Cheddar, Provolone etc.)

- Semisoft (Brick, Muenster, Roquefort, Taleggio etc.)

- Soft (Camembert, Brie, Asiago etc.)

- Fresh (Mozzarella, Ricotta, Cottage cheese, Tosella, Paneer etc.)

- <u>Processed</u> (smooth cheeses made from mixing several cheeses or adding other ingredients: American, cheese spreads etc.)

Among all the above cheeses, fresh cheeses are important for nutritional value in daily diet and they have to be consumed in a short time, because of low shelf life (5-9 days), after this period the cheese begins to lose moisture and increases its microbiological values making it unhealthy for consumption (Perry, 2004). Furthermore, fresh cheeses are peculiar in terms of their utility, taste and the way they are produced for example fresh cheeses using enzymes or acids or cheeses produced with high heat or low heat. Additionally, fresh cheeses are consumed vastly because of their soft texture and mild flavor because in some part of the world aged cheeses are not preferred because of strong aged flavor and some other prevalent reasons.

Among the fresh cheeses, two of the cheeses, paneer and tosella are highly consumed because of their texture and flavour. However, these cheeses are not widely known and they are produced regionally. These two cheeses are similar in characteristics but their production is completely different. Paneer is among one of the most consumed fresh cheeses which is produced at high heat using acidic coagulants such as citric acid. In addition to paneer, Tosella is another cheese which is consumed widely and have similar appearance (Fig. 1), texture and flavor like paneer but produced with low heat and using rennet.

Paneer is marble white in appearance, having firm, symmetric and spongy body with a compact texture and a sweetish-acidic flavor. According to Prevention of Food Adulteration Rules (PFA <u>2010</u>), *Paneer* means the product obtained from cow or buffalo milk or a combination there of by precipitation with sour milk, lactic acid or citric acid. It shall not contain more than 70% moisture and the milk fat content shall not be less than 50% of the dry matter (Srivastava and Goyal 2007). It is estimated that about 5% of total milk produced in India is converted into *paneer* (Sahu et al., 2017).

Tosella also called Schiz, is another fresh cheese which is a traditional cheese produced in the eastern part of Trentino region (Settanni and Moschetti, 2010). It has high nutritional value such as high protein and minerals. In the past, Tosella was made from raw cows' milk but, to avoid hygienic implications, it is nowadays produced with pasteurized milk, despite the fact that cheeses made from raw milk are well known to have more intense flavours than cheeses of the same age and type made from pasteurized milk (Albenzio et al., 2001; Beuvier et al., 1997; Demarigny et al., 1997).
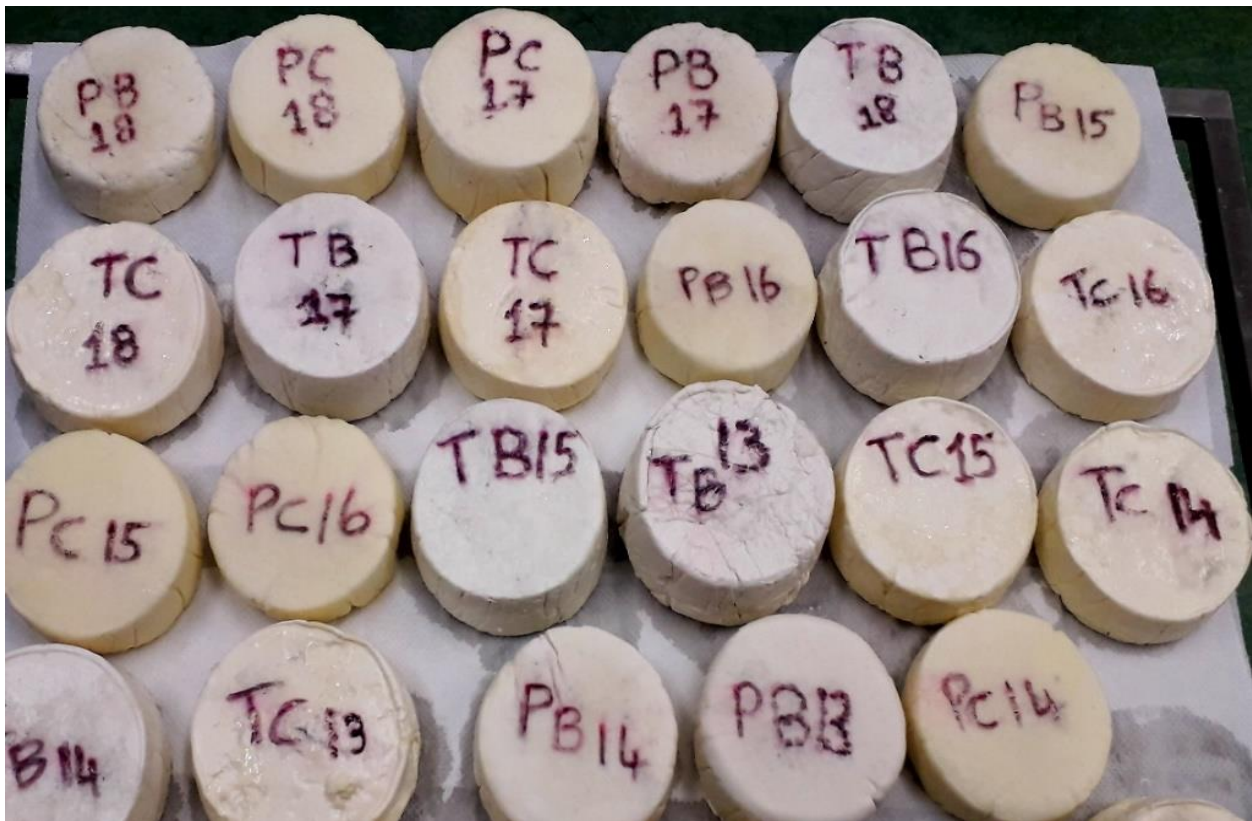
Figure 1. Paneer and Tosella produced in DAFNAE laboratory.The ID of samples are; TC-Tosella cow, TB-Tosella buffalo, PC-Paneer cow, PB-Paneer buffalo and the number in the end corresponding to the number of samples produced.

Since milk is the raw material for paneer and tosella production, it has been shown that fat, fatty acids and other important components of milk can show significant variability depending on the animal's breed, genetic variable, feeding diet (Arnould and Soyeurt 2009; Stocco et al.,2017). Moreover, milk

used from different species of animals such as buffalos and cows may have different impact on paneer and tosella qualities. Moreover, types of milk used in production of these cheeses affect the final price and preferences from the consumer. Additionally, some consumers refrain from consuming the tosella for socio-cultural reasons e.g., in some culture consumers do not buy food products which contain rennet from animals or calve stomach. Therefore, it is important to make sure that consumers get whatever they asked for, but not an adulterated product. This is because it is an economic offence and may hurt consumer sentiments or health.

To give the consumer the right information about the types and quality of the cheeses it is very important to perform the analysis of cheeses in order to ensure the authenticity of cheeses. In general cheeses qualities can be analyzed using conventional methods. Presently, the conventional methods for determination of different cheeses qualities are generally solvent extraction for lipids, Kjeldahl methods for protein, oven drying method for moisture/dry matter content, titration method for NaCl analysis and many more other methods based on these principles. These methods are accurate and produce precise results.

However, these techniques require a lot of time, specialized material, they are mostly expensive, and destroy the cheeses samples which is economically disadvantage to the producer (Karoui et al.,2006; Lenart et al.,2012; Tao and Ngadi 2007). Critical study of research works done from past 25-30 year can be concluded that near infrared spectroscopy has proven its worth and applicability in measuring the different qualities of cheese. Since, efficient routine analysis of cheeses composition is important for dairy farms, cheeses producers, industry and consumers, a rapid, cheaper and non-destructive spectral methods were developed to measure main components of cheeses such as fat, protein and moisture (Karoui et al., 2006; McQueen et al., 1995; Rodriguez-Otero et al., 1996). Therefore, this study focuses on cheese analysis using NIR spectroscopy for cheeses classification and prediction of paneer and tosella qualities.

CHAPTER 2: LITERATURE REVIEW

This chapter is a review of state of art pertaining to the cheese quality analysis and NIR spectroscopy in relation to the application to the cheese analysis. Table 1 shows a list of different methods generally used for evaluation of cheeses. Several methods, destructive and non-destructive (causing no damage to the original sample), including chemical enzymatic assays, and instrumental techniques have been developed over the decades.

2.1 Primary methods to measure the cheeses qualities

2.1.1 MALDI-MS is a non-destructive, potentially useful technique in protein identification and quantification of dairy products (De Noo, 2005). MALDI consist in the mix of sample with a suitable matrix material and applied to e metal plate. Then a pulsed laser irradiates the sample and the analyte molecules are ionized by being protonated or deprotonated in the hot plume of ablated gases and then they can be accelerated into whichever mass spectrometer is used to analyze them (Poonia et al., 2017). The technique has the advantage of requiring a small quantity of sample and can be used for the analysis of heterogeneous biological samples such as cheese as demonstrated by (Liland et al., 2009). The initial high cost of procedure and requirement of trained equipment are the two main limitations in the wide use of this technique.

2.1.2 Chromatographic techniques offer a possibility for the rapid and reliable discrimination and quantitative determination of similar chemical molecules in complex matrices. They have been applied for the characterization of different proteins in cheese, with the aim of identifying and semi-quantifying peptides and peptide-like molecules with potential technological and biological properties. Proteins are separated based on their affinity for either the stationary or mobile phase. Based on the fact that the amino acidic sequences are not identical, they can be easily discriminated by HPLC-MS after applying a simple extraction procedure (McSweeney et al., 2011). The advantages of this HPLC coupled to mass spectrometry are sample handling together with the

high selectivity, robustness and sensitivity, but there also some limitations such as high cost and requirement for an experienced equipment.

2.1.3 ELISA is a common method used when evaluating the quality of dairy products, based on a highly specific interaction of an antigen and an antibody, with one of them carrying a covalently bound enzyme (Dvorak et al., 2016). The assay's enzyme catalyzes the chemical transformation of a substrate (added to the reaction mixture) into a colored product (Poonia et al., 2017). This assay can be used in quantitative and qualitative test of dairy products to quantify or detect different proteins in fresh cheeses. It can also be stored for a long time in the form of a ready-to-use kit for routine testing of milk or dairy products. ELISA is characterized by the following advantages: (i) simple procedure, (ii) high specificity and sensitivity, (iii) high efficiency as analyses can be performed without complicated sample pre-treatment, (iv) safe and eco-friendly and (v) cost effective procedure. However, it also exhibits some disadvantages such as (i) labor-intensive, (ii) high possibility of false negative or positive and (iii) antibody instability (Sakamoto et al., 2018).

2.1.4 Capillary electrophoretic methods have been in use for the detection and the determination of different proteins in dairy products (Recio et al., 2000). After staining with Coomassie Blue R, protein bands can be quantified by densitometry. SDS-PAGE on miniature gels with the automated system requires less than three hours and enables rapid and reliable separation of proteins with high resolution and good quantification using only a small amount of sample and buffers (Poonia et al., 2017). Besides, the protocol does not require coated capillaries that are necessary in methods using alkaline buffers; hence, the analytical costs are substantially lower. However, this method has the shortcomings of poor reproducibility, low sample loading capacity and low throughput due to ineffective interfaces between the two separation dimensions (Huang et al., 2006).

2.1.5 The polymerase chain reaction (PCR), based on the amplification of species-specific DNA sequences, can successfully discriminate between the milks of different species in a mixture (Plath et al., 1997). It has been proposed as a useful technique for classification of species in dairy products based on the analysis of protein fraction (Lopes-Calleja et al., 2007). It is largely applied thanks to the high sensitivity and specificity as well as rapid processing time and low cost, but it has also some limitations such as requirement of certain conditions for its use and it shows considerable problems with post-PCR contamination (Bottero et al., 2003).

Spectroscopic methods basically analyze the interaction of electromagnetic waves with matter. There is wide range of spectroscopic electromagnetic spectrum; these include X-ray, ultra violet, visible, infrared, microwave, and radio-frequency (Straughan and Walker 1976).

• X-ray region (wavelengths between 0.5 and 10 nm) is involved in energy changes of electrons of the internal layers of atoms and molecules. Guggisberg et al., (2013) investigated the measuring of holes of semi-hard Tilsit-type cheese using X-ray computed tomography analysis. By their study was achieved a better image of holes location by using X-ray generator than other imaging techniques.

• Ultraviolet region (10-350 nm) is involved in electronic transitions from the excited state to the ground state are observed. In this wavelength range, luminescence (fluorescence and phosphorescence) may also be observed. Karoui et al., (2005), predicted the rheology parameters of ripened semi-hard cheeses from spectra recorded on these cheeses at a young stage. They used UV spectroscopy and dependent only on visible light, the riboflavin fluorescence spectra potentially provided viable and economic prediction of the rheology of ripen cheese.

• The visible region (350-800 nm) is another zone where electronic transitions occur. Molecules exhibiting a large number of conjugated double bonds absorb energy in this region.

Table 1. Methods used for cheese quality analysis

Technique	System	Advantages	Disadvantages
HPCL-MS	Destructive	High sensitivity Robustness high selectivity	very expensive not portable experienced technician
MALDI-MS	Non-destructive	high accuracy small quantity of sample	specialized equipment high cost
ELISA	Destructive	simple, eco-friendly high specificity and sensitivity Cost effective	labor-intensive, high possibility of false positive or negative antibody instability
Capillary electrophoresis	Destructive	sample small amount high resolution good quantification	poor reproducibility low sample loading capacity
PCR	Non-destructive	high sensitivity and specificity costly effective	post-PCR contamination certain conditions of operation
LIBS	Non-destructive	Minimal sample preparation Chemical free Rapid detection	highly cost poor precision system complexity
Raman Spectroscopy	Non-destructive	no sample preparation highly specific, fast Not interfaced by water	results covered by impurities can destroy the sample heating through laser radiation
X-ray	Non-destructive	improved image quality and detection capabilities	high cost need for high voltage power risk for food irradiation
CT-scan	Non-destructive	short time analysis accurate extraction of depth and distance information	high cost, specialized equipment long computation time and low speed
NIRS	Non-destructive	No sample preparation Very specific Quick, fast Economic, eco-friendly	different calibrations needed for different materials complex spectra interpretation

• The near infrared (NIR) region (800-2500 nm) is the first spectral region exhibiting absorption bands related to molecule vibrations. This region is characterized by harmonics and combination bands and is widely used for composition analyses of food products.

• The mid-infrared (MIR) region (2500-25000 nm) is the main region of vibrational spectroscopy. This region retains information, allowing organic molecules to be identified and the structure and conformation of molecules such as proteins and lipids to be characterized.

Most of the above-mentioned techniques are slow and expensive and require sophisticated laboratory facilities and highly specialized equipment, therefore there is a need for alternative method which would permit a reduction in analytical workload without sacrificing important information. In this regard, near infrared spectroscopic (NIRS) method that has already been documented for the rapid, precise and non-destructive analyses of a wide range of dairy products, that we will mention later, can provide a solution for this challenge.

2.2 NIR spectroscopy methods principle

NIR spectroscopy analysis is based on the principle that certain types of chemical bonds absorb specific frequencies of light energy in the NIR region. If a photons energy matches the molecular energy levels difference it will be absorbed and will raise the energy of a molecule from a ground state to a specific excited state as illustrated in Figure 2 (Kumar et al., 2011).

NIRS provide the information about structure and physical properties of food materials. When light hits a sample the incident radiation may be reflected, absorbed or transmitted and the relative contribution of each phenomenon depends on the chemical and physical nature of the sample (Osborne et al., 1993)
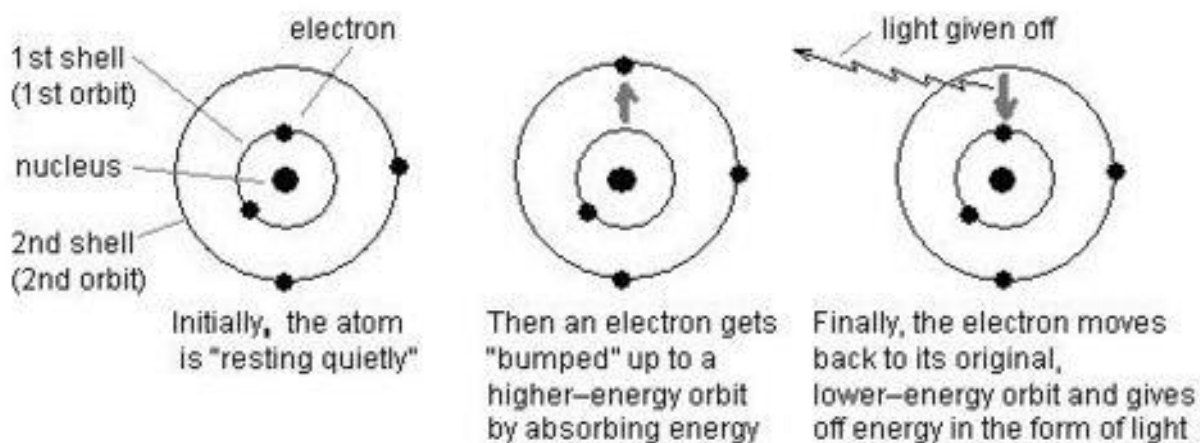
Figure 2. Absorption of electromagnetic radiation.

2.3 Scientific work done on cheeses quality evaluation through NIRS

Several authors have reported about the evolution and the use of NIR for dairy products analysis through years. The application of near-infrared spectroscopy to cheese analysis is noticed first by the study of Frank and Birth (1982), by determining the fat, protein, and moisture contents in a total of 30 samples of different cheese varieties (Cheddar, Colby, Parmesan, Romano and Gruyere). For the determination of the composition they previously grated the sample, freeze-dried a subsample of the grated cheese and made measurements of the grated and freeze-dried subsamples. For calibration the second derivative and multiple linear regression with three wavelengths for each component were used. Better results were obtained with the freeze-dried subsample. Finally, in their study only fat and protein were determined and water detection was missing (Frank and Birth 1982), giving the idea that technique improvements were needed.

Frankhuizen and Van der Veen (1985) applied near-infrared spectroscopy to the quality analysis of different Dutch cheeses (Edam, Gouda and processed cheese). For that, they grated the sample through a Hobart mill and made the measurement with a Technicon optical filters Infra- Analyzer 400 instrument, at 19 different wavelengths. They used calibrating by MLR with six wavelengths for moisture determination, the most significant one corresponding to 1940 nm (characteristic band of

10

water absorption); for protein, between six and ten wavelengths were used, including the characteristic protein bands of 2139 and 1734 nm; for fat, seven wavelengths between 2310 and 1759 nm were included. Upon using a group of different samples for prediction, they concluded that better results are obtained when the instrument is calibrated in a specific way for each variety of cheese.

Wehling and Pierce (1988) used the reflectance NIR spectroscopy for the determination of Cheddar cheese moisture content. The sample was introduced in the instrument in a plastic bag, after having been crumbled in cylinders of 1 mm in diameter and 20 mm in length, and the calibration was made by MLR. From their study were specified the spectra peaks related to different contents of the cheese. The measurements at 1818 and 1445 nm were proportional to the moisture content and the measurement at 1734 nm were directly related to the fat and protein contents.

Years later, Wehling and Pierce (1994) studied various mathematical treatments for the determination of cheese composition. They collected the spectra by transmittance and reflectance of previously grated samples. By the final results, the smaller SEP were obtained when using the PLS of the first derivative in the case of moisture and the PLS of the direct spectra in the case of fat. In the same way, the use of optical fiber on intact samples was tested, thus obtaining worse results than with grated samples.

Rodriguez-Otero and Hermida (1996) used with success near-infrared reflectance spectroscopy to determine fat, protein, and total solids of fermented milks avoiding any pretreatment of the sample. However, better results were obtained by the same authors when they analyzed the samples by near infrared transflectance spectroscopy (Rodriguez-Otero et al., 1997).

Other authors have made attempts to measure chemical variables in addition to the major components in cheese. Skeie et al., 2006, used NIR spectroscopy to predict selected aminoacids in Prast cheese, common in Norvegia. By the study were achieved high correlations, more than 0.9 for many

variables. They used the region from 780-2500 nm which indicates that the higher part of the NIR spectrum could be used for higher prediction ability.

Cevoli et al.,2013 investigated the potential of NIR coupled with different statistical methods to estimate the authenticity of grated Parmigiano Reggiano cheese (PDO). NIR spectra were obtained using a spectrophotometer Vector 22/N (Bruker Optics, Milan, Italy) in the diffuse reflectance mode. Principal components analysis (PCA) was used for an explorative spectra analysis, while the Artificial Neural Networks (ANN) were used to classify spectra, according to different cheese categories. The classification performance for all sample classes, was higher of 90% in test set validation.

Another study in the last decade, was focused in the feasibility of NIR transmittance spectroscopy to predict cheese ripeness using the ratio of water-soluble nitrogen to total nitrogen as an index of cheese maturity (WSN/TN). Curro (2017) used 52 Protected Designation of Origin (PDO) cow milk cheeses of 5 varieties (Asiago, Grana Padano, Montasio, Parmigiano Reggiano and Piave) provided for laboratory and chemometric analysis. NIR instruments operated in transmittance mode for the wavelengths from 850-1050nm. Prediction equations for WSN and TN matched with cheese, were developed using cross validation the whole data set and external validation on a subset of the entire data. The coefficients of determination for WSN and TN were more than 0.85 both in cross validation and external validation, showing a good prediction by NIR spectroscopy technique.

2.4   Accuracy, repeatability and reproducibility of NIR

The accuracy, repeatability (precision) and reproducibility of NIR analysis have been extensively examined. Angelino (1996), Osborne and Fearn (1983) and Valdes and Summers (1986) reported that the NIR analysis has produced a comparable or with better than those of reference methods of analysis. The accuracy of NIR analysis refers to the closeness of measured values to reference values. Parameters measuring the accuracy of NIR determination include root mean standard error of

prediction (RMSEP), root mean square of cross validation (RMSECV), root mean square error of calibration (RMSEC). Although many NIR analyses were reported to be satisfactory, there are no critical levels of RMSECV, RMSEP or RMSEP for an acceptable or satisfactory NIR analysis. The acceptability of NIR analysis is based on the tolerance of the error in practical situations.

2.5 Advantages and disadvantages

Analysis using NIR method has several advantages. For example, NIR spectroscopy is quick, simple sample or no sample preparations and it take multicomponent measurement. Also, this method is ecofriendly, noninvasive and economical (Osborne et al.,1993). Compared with other rapid or noninvasive techniques such as NMR (Nuclear Magnetic Resonance), FTIR (Fourier Transform Infrared spectroscopy) and FT Raman, NIR is especially suitable for materials of high water content and requires no chemical pretreatment of samples, which favors a clean environment. Although, NIRS has these good qualities, it has some limitations too. First of all, it should be kept in mind that it is secondary methods of analysis. Hence, it needs to standardize against primary methods. Therefore, it results depend on accuracy of primary methods of result which is used for calibrating the NIRS. Any error in reference method will result an error in this method too. Furthermore, NIR is specific to material. It means that different calibration is needed for different materials. For this reason, the NIR technique was first used in the milling, brewing and dairy industries where a large number of samples are of a similar type (Osborne et al, 1993). In addition, when using NIR reflectance spectra, only information of exterior parts of samples are extracted hence error might occur when internal content or structure is different from that of the exterior. Finally, NIR spectral interpretation is very complex (Osborne et al, 1993).
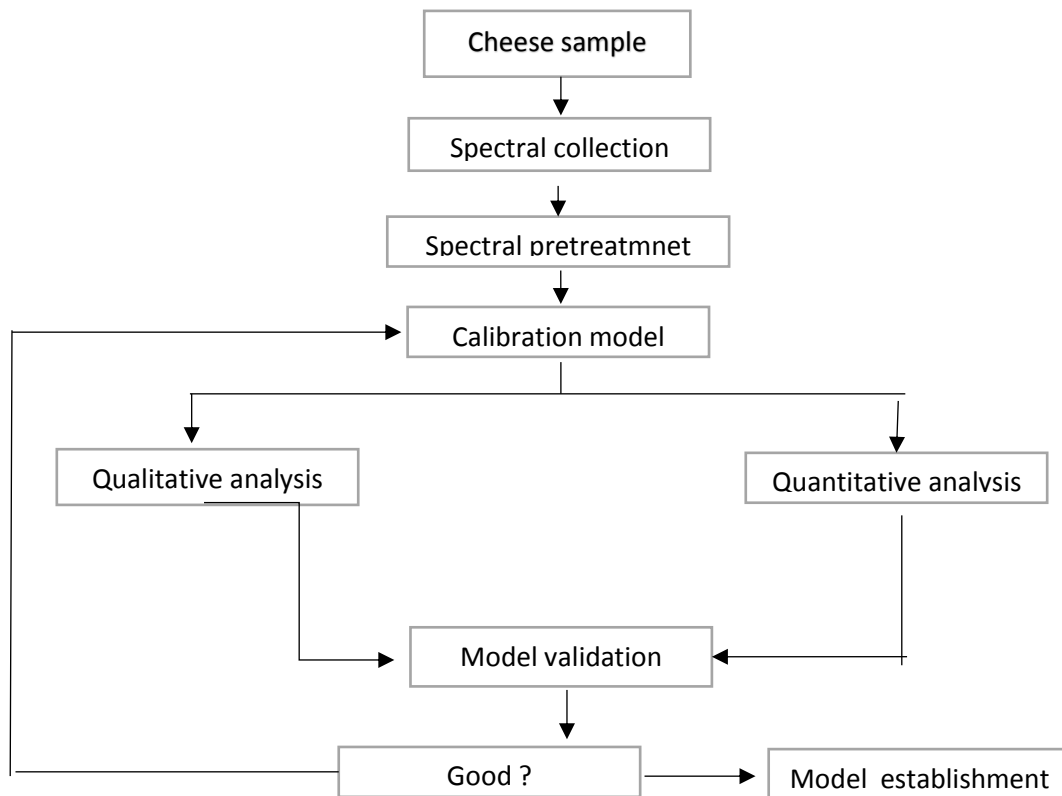
Figure 3. Flow diagram of NIR analysis.

Conclusion

Based on above review it can be concluded that there are numerous methods for measuring cheese composition and qualities. However, most of them are either time consuming, expensive, and non-ecofriendly. As consumer awareness and demand are increasing for high quality food such as cheese, food industry needs a tool/ technique which can provide them multiple measurements, takes less time to measure and preferably nondestructive in order to ensure food quality and to check whether products have been produced according to specification. NIR spectroscopy is a technique which can provide a rapid and nondestructive determination of cheeses qualities as explained above and seems to be a good tool for quality analysis as mentioned by different authors through years.

2.6 Hypothesis and Objectives

It is well established that the demand for cheeses is high and in future it will grow more. Food industry needs to brace up themselves for this increasing demand so that they can serve good quality of cheeses and cheese products to the consumers. With consumers' growing requirement for food nutrition composition, guaranteeing and improving the quality of milk and dairy products becomes more and more important. It is thus necessary to analyze the varieties of milk and milk products.

Therefore, the aim of this work is to explore the potential use of NIR spectroscopy as a rapid, a chemical free technique for discrimination of cheeses and to predict physical qualities of Tosella and Paneer.

We have hypothesized that:

- NIR Spectroscopy can be used in classification of two different fresh cheeses which look alike but have different production methods
- Near infrared spectroscopy can be used as a potent method to predict the physical qualities of paneer and tosella.

Specific objectives in this experiment are:

1- To explore qualitatively the cheeses spectra using principle component analysis (PCA)
2- To classify two different fresh cheeses, Paneer and Tosella from cow and buffalo milk using unsupervised and supervised chemometric techniques.
3- To predict the physical qualities of Tosella and Paneer using NIR spectroscopy.

CHAPTER 3: MATERIALS AND METHODS

In this chapter, a brief description of experimental site, production of fresh cheeses, NIR analysis to the multivariate methods used for the achievement of objectives, are given.

3.1 Cheese production

The cow milk for production of Paneer and Tosella cheeses, was obtained from the experimental farm of Agripolis (Legnaro, Italy), while buffalo milk was collected from a farm in Piove di Sacco, Italy. Milk was collected from 6 different cows and buffalos and was transported into the lab with plastic bottles in a portable cooling box to ensure proper temperature conditions. Then, it was stored in fridge overnight.

3.1.1 Production of Paneer

Paneer was produced using 1.5 L cow milk and 1 L buffalo milk. Four cheeses vats were filled with milk for a single batch cheeses production in total 3 batches of cheeses were produced in a single day and overall cheeses was produced for three different days. Cheese production process is shown in Figure 4.A. The production process was preceded by the heating of milk at 82°C for 5 minutes and after heating, the milk was cooled at 70°C. Milk was coagulated by addition of 175 ml of 2 % citric acid solution for cow milk and 120 ml of 2 % citric acid for buffalo milk. Similar approach was suggested by Sachdeva and Singh (1988). After coagulation of milk, the curd was allowed to settle down for 10 min without stirring. Thereafter, the curd along with the whey was transferred in a hoop lined with cheese cloth to remove the whey (Figure 4.B). After the draining of the whey, the curds of cheeses were subjected to pressing to obtain a compact block of paneer. Pressing occurred in 250 atm for 10 min in two runs for both sides. The cheeses were signed with specific numbers as in their vats, as reported in Table 3 and the final weight was taken to see the water lost. Then they were chilled in water.
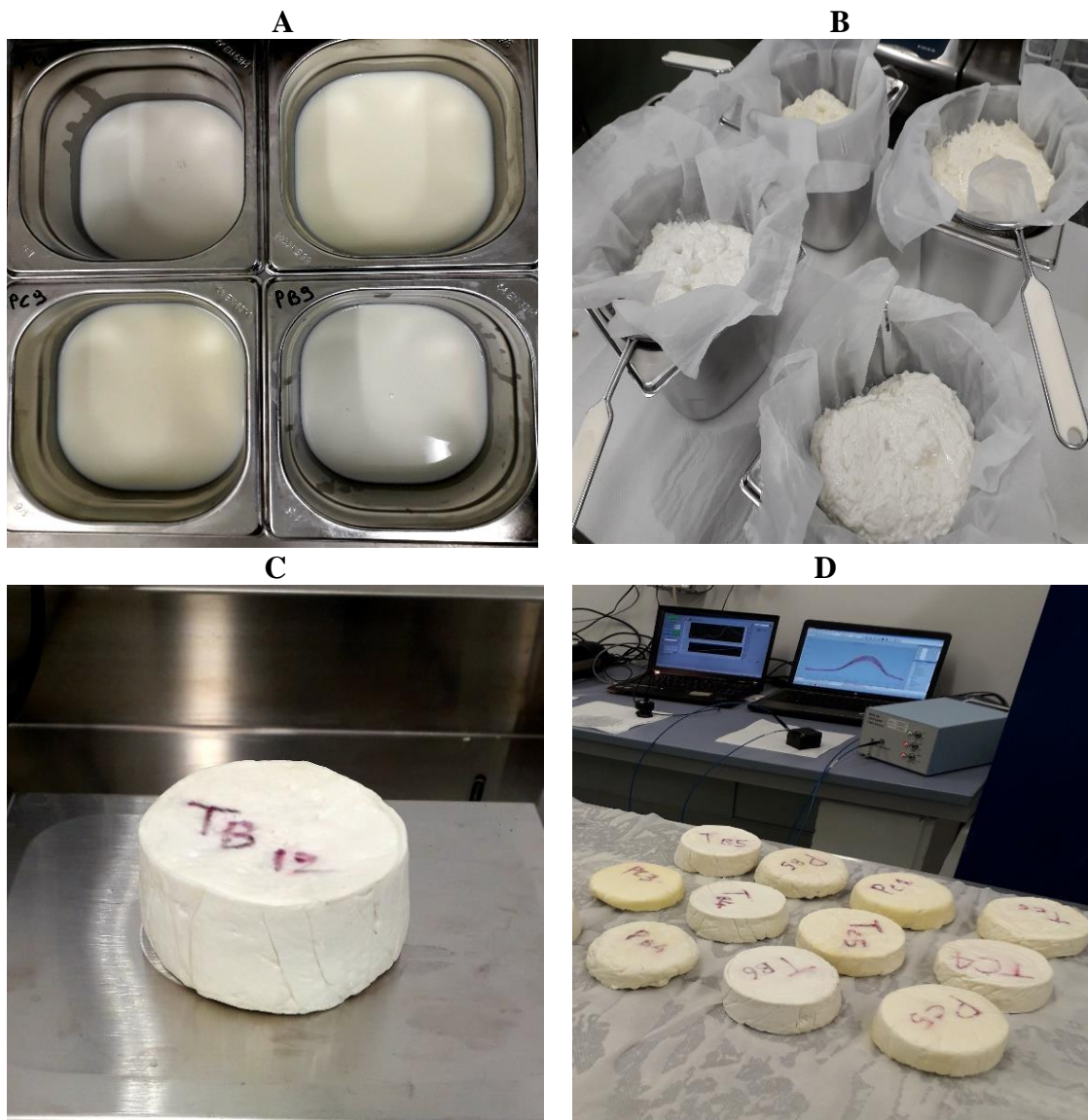
Figure 4. (A) Heating of buffalo and cow milk in vats (B) curd ready to be pressed, (C) weighting of cheese, (D) NIR instrument used for fresh cheeses analysis.

3.1.2 Production of Tosella

Similar to paneer, the same amount of milk was used for tosella production. The heating and coagulation of milk was done at 35°C and the cooking of curd was performed at 40°C. Initially milk was heated till its temperature arrived 35°C, then milk coagulants both 82.5 ml of 2 % citric acid

solution and 8 ml rennet enzyme in cow milk, and 65 ml of 2 % citric acid and 8 ml rennet in buffalo milk, were added both at the same time. After coagulation, the curd was cooked for 10 minutes and then first cut was performed. After 10 minutes of first cut, second cut was performed and then curd was cooked for 10 more minutes. Then, the cheeses were filtered and pressed at 250 atm for 10 min in 2 runs and then the curd was weighted to measure the water losses (Figure 4.C). The final step for cheeses was transferred to the 20% brine solution which had temperature of 4°C. Cheeses stayed there for one hour. Then their weight was recorded and they were stored in refrigerator for 48 hours for further analysis. In the end of the cheeses production, 75 cheeses, 37 Paneer and 38 Tosella were produced.

Table 2. Cheese markings used to identify Paneer and Tosella cheese produced by cow`s and buffalo`s milk. The ID of samples is equivalent to: first letter corresponding to Paneer /Tosella, the second one related to Buffalo/cow and the number in the end corresponding to the number of samples produced.

Buffalo milk		Cow milk	
Paneer	Tosella	Paneer	Tosella
PB1	TB1	PC1	TC1
PB2	TB2	PC2	TC2
PB3	TB3	PC3	TC3
PB4	TB4	PC4	TC4
PB5	TB5	PC5	TC5
PB6	TB6	failed	TC6
PB7	TB7	PC7	TC7
PB8	TB8	PC8	TC8
PB9	TB9	PC9	TC9
PB10	TB10	PC10	TC10
PB11	TB11	PC11	TC11
PB12	TB12	PC12	TC12
PB13	TB13	PC13	TC13
PB14	TB14	PC14	TC14
PB15	TB15	PC15	TC15
PB16	TB16	PC16	TC16
PB17	TB17	PC17	TC17
PB18	TB18	PC18	TC18
		PC19	TC19
		PC20	TC20

3.2 Spectral acquisition

Tosella and Paneer were analyzed using a portable NIR spectrophotometer working in reflectance mode (LabSpec2000, QualitySpec® Pro, Analytical Spectral Devices Inc., Boulder, CO, USA). Spectra were collected in the range 350-1830 nm at data-point intervals of 1 nm and coupled to a high-intensity probe (Analytical Spectral Devices Inc.). The LabSpec2000 model is powered directly from a standard AC connection and includes an internal broad-spectrum halogen light source that eliminates the need for an external fiber optic light source. All samples (in the form of cheese block) were scanned at three different position of same cheese at the room temperature, by contacting the scanning head of the fiber-optic contact probe (10-mm diameter) over on the fresh cheeses surface, without any sample modification. In order to achieve better signal, temperature of sample was ensure to stay around 23°C, by managing the heat of instrument with turning off the light when it was not necessary. It was also ensured that the instrument was clean for spectra acquisition of each samples.

3.3 Quality analysis of cheeses

3.3.1 Color analysis

Cheese were cut in half and kept at room temperature to get the equilibrium and then color was measured at 3 different positions on the surface of all raw cheese using a portable spectro colorimeter (CM 508, Minolta Co. Ltd., Osaka, Japan). The average of three measured color was used in final results for further analysis. Color traits were expressed according to the Commission Internationale de l'Eclairage colorimetric system (CIE L*a*b*), using primary illuminant D65 (standard daylight) with a 10° observer. Briefly, the L* value is the lightness coefficient (ranging from black = 0 to white = 100), the a* value indicates the position on the green (−)-to-red (+) axis, b* the position on the blue (−)-to-yellow (+) axis. The instrument also directly calculates C* (chroma) and h° (hue angle) values: C* is calculated as (a*2 + b*2)1/2, h° is calculated from the arctangent of b*/a*, and together they

represent the polar coordinates of the cheese in the color plane in which a* and b* represent the Cartesian (rectangular) coordinates. The 3 acquisitions were averaged before statistical analysis.

3.3.2 Texture analysis

Texture of all the cheeses were determined using a texture analyzer (XT2i, Stable Micro Systems Ltd., Godalming, Surrey, UK) with a Warner–Bratzler shear device (50-N load cell; 2 mm/s crosshead speed). For each cheese, 2 cylinder-shaped core samples were taken (1 cm^2 cross sectional area; 3 cm long), and their textural values were averaged before statistical analysis. Texture data were reported as hardness (defined as the maximum shear force, expressed in Newtons), and work (defined as the working shear force, expressed in 10^{-3} J).

3.3.3 Cooking losses

The fresh cheeses were cooked on top of an electric hot plate (Tristar BP-2970 Piastra 70 x 23 cm, thermostatic temperature control, Tristar srl, Italy)) with a voltage and power of 120V~60Hz, 2000W. Initially plate was heated on a pre settled temperature setting. Once it arrives the temperature cheeses are cooked for 4 minutes on each side. After cooking cheeses weight were taken. Cooking loss was calculated by using the following formula. Initial weight of cheeses- final weight of cheeses, divided by initial weight of chesses and final results was multiplied by 100 in order to express the weight loss in percentage.

3.4 Statistical analysis of reference and spectral data

3.4.1 Statistical analysis of cheese traits

Data editing: all the data were edited and combined (reference and spectral data) to create a data matrix. After this reference method result were analyzed using SAS MIXED procedure (SAS Institute Inc., Cary, NC, USA) according to the following hierarchical mixed model:

$$y_{ijklm} = \mu + species_i + animal\,(species)_j + session_k + Cheese_l\,(animals) + animals{\times}Cheese_{il} + e_{ijklm}$$

where y_{ijklm} is the trait analyzed; μ is the overall intercept of the model; $species_i$ is the fixed effect of the ith type of species (i = 2 levels, cow's and buffalo's milk); $animal(species)_j$ is the random effect of the jth milk producing animal (j = 8 cows and 7 buffaloes); $session_k$ is the random effect of session of cheese-making (k = 1 to 4); $Cheese_l(animal)$ is the fixed effect of type of cheese within type of milk (l = 1 for Paneer and 2 for Tosella/Schiz); $animal{\times}Cheese_{il}$ is the fixed effect of interaction between type of milk and type of cheese; e_{ijklmn} is the random residual $\sim N\,(0, \sigma_e^2)$.

Using this model outlier data or observations were detected and removed. Similarly spectral data were also inspected using Mahalanobis distances. All the spectra which were outside $\pm$ 3 mean distances were discarded. Finally, a clean data matrix was created by matching the spectral and references data.

3.4.2 Spectral data analysis

3.4.2.1 reflectance to absorbance

Spectra analysis was performed using the R studio version 4.0. Spectral data were first converted from reflectance to absorbance by using (log(1/R)) formula where R is reflectance value. Then spectra were plotted as absorbance vs wavelengths, Figure 5.

3.4.2.2 Preprocessing of spectra

Spectra data were pre-processed prior to any type of spectral quantitative or qualitative evaluation to make the extracted information more accessible and to improve the goodness of fit in the chemometric modelling approaches (Nache et al.,2015). Different types and combinations of spectral pretreatment were used to simplify the spectra and develop robust models. These treatments were Savitzky Golay first-derivative, second-derivative with second order of polynomial and 15 points in each interval. Standard normal variate (SNV), and mean centering were other pre-processing techniques utilized for analysis of spectra.

By calculating the first and second derivatives, sample-to-sample baseline variations are eliminated and also absorption peaks are enhanced (Savitsky and Golay, 1964). SNV is a mathematical transformation method used to remove slope variation and to correct for scatter effects (Barnes et al., 1989). Mean centering calculates the mean of each variable and subtracts this from all measured values for the variable. Another way of interpreting mean-centered data is that, after mean-centering, each row of the mean-centered data includes only how that row differs from the average sample in the original data matrix.

3.4.2.3 Qualitative analysis (Principle component analysis)

Principal Component Analysis (PCA) was applied to reduce the spectral data dimension and explore the structure of the data and to identify any possible separation among mixed groups of samples. For this, it defines new variables, the so-called Principal Components (PC), consisting of linear combinations of the original ones (Kamruzzaman et al., 2012). PCA was applied after pre-processing. PCA returned a compact representation of the raw data and highlighted the existing correlations among samples and variables through the combined plots of the model scores and loadings.

3.4.2.4 Unsupervised technique (Hierarchical cluster)

Unlike supervised methods, unsupervised methods do not require any labeled sensor data. Instead, they try to automatically find interesting activity patterns in unlabeled sensor data. A number of different unsupervised approaches have been proposed for discovering activity patterns. Hierarchical cluster analysis (HCA) was used as an unsupervised method to cluster the cheeses data. In the hierarchical cluster analysis, the spectral distances between all spectra are computed hierarchically. First and foremost, two spectra with the highest similarity (smallest spectral distance spectra), are join together to form a cluster. Then the distances of the cluster formed and all other spectra are calculated, and two spectrum or clusters with the smallest distance again are merged to form a new cluster. The procedure is duplicated until only one bug cluster remains (Gurbanov et al.,2016).

3.4.2.5 Supervised technique

3.4.2.5.1 PCA-DA

Spectral data have many more variables (predictors) than observation. Meaning that a lot of information is available for each sample. Finding relevant differences between classes of samples in such a situation is difficult. The number of parameters to estimate in regular forms of discriminant analysis far exceeds the number of independent samples available. So, in order to achieve the classification of such data often-used strategy is to compress the information in a much smaller number of variables, usually linear combinations of the original set, and perform simple methods like LDA on the new, small, data matrix (principle components). There is such technique called PCA-DA which is basically doing LDA with PCs extracted from the PCA analysis.

3.4.2.5.2 SIMCA

Classification models were developed using SIMCA. SIMCA is a method of supervised data classification that requires a training data set consisting of samples with attributes and their class membership. In order to build the classification models, the samples belonging to each class need to be analyzed using principal components analysis (PCA). New observations are projected into each PC model and the residual distances calculated. The observation may be found to belong to multiple classes and a measure of goodness of the model can be found from the number of cases where the observations are classified into multiple classes.

3.4.2.6 Prediction model using PLSR

Partial least squares regression is based on linear transition from a large number of original descriptors to a new variable space based on small number of orthogonal factors (called latent variables). PLS model was built on the obtained spectral data of fresh cheeses and spectral pretreatments such as standard normal variate (SNV) and SG first derivative were applied. Then data were mean centered and divided into calibration (80%) and prediction (20%) sets.

3.4.3 Model evaluation

The best prediction model was selected based on root mean square error of prediction (RMSEP), the number of the latent variables (LVs), at the minimum values of root mean square error of calibration (RMSEC) and the maximum value of determination coefficient in calibration (R2C). In general, as low as the RMSEC value can be, and R2 as close as possible to 1, the better will be the model predictions (Andrade et al., 2019).

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Physical quality of cheeses

Descriptive statistics of all physical qualities such as color, texture and cooking losses of both Paneer and Tosella are presented in Table 3. The results show that in general, Tosella has higher L mean values than paneer. It means that Tosella is white in color than Paneer and this can also be seen in Figure 1. However, other mean color parameters such as a, b, c and h were higher for paneer. For a* was negative side meaning that Paneer was slightly greenish and b* was positive, which tell that cheeses was light yellow. Furthermore, Paneer has higher variability in comparison to Tosella which can be seen in the values of SD, min and maximum range for all color parameters, except H value of Tosella from Table 1. In addition to the color of cheeses, cooking losses in Paneer was way lower than Tosella. This is justifiable in the sense that Tosella has higher moisture and soft texture which can be seen in texture result of cheeses, where is reported that texture of Paneer has high value in terms of maximum work required to cut the cheeses.

There is complete lack of physical qualities results in case of Tosella. However, for Paneer results related to color and texture are comparable to the work of Jotarkar et al., (2017) who have reported similar values of L, a, b color parameter for Paneer. Paneer texture value is low in comparison to Ahmed and Bajwa (2019) work. This could be due to the fact that they have used different type of coagulant and coagulants strength and Paneer production temperature.

4.2 NIR spectra of cheeses

Spectra of cheeses collected from all the samples in the range of 350-1830 nm were converted from reflectance to absorbance. Both reflectance and absorbance spectra are presented in Figure 5. The first point to highlight is spectral range and wavelength of spectra. It could also be seen that, spectra have visible range spectrum, all the area of 3rd overtone and the last part of 1st overtone. As water and fat are the main components in cheese, the absorbance spectra clearly exhibit absorbance peaks of fat

980 and 1450 nm for moisture respectively. The absorption peak at 980 nm due to $2v1 + v3$ (v1; symmetric stretching, v3; antisymmetric stretching) vibration of water is quite broad and dominates the spectrum (Sasic and Ozaki, 2001). Large baseline variation can be observed, which can be attributed to variation in light scattering by the fat globules and casein micelles 970 nm and 1450 nm (second and first overtone of the O–H stretch vibration).

Table 3. Descriptive statistics of physical attributes and quality index of fresh cheeses (Paneer n=35 and Tosella, n=37).

Type of cheese	vars	Mean	SD	min	Max
Paneer	L	85.0	2.46	77.9	89.9
	a	-1.7	1.14	-3.95	0.08
	b	16.9	3.85	10.4	25.4
	c	17.1	3.79	10.5	25.4
	H	96.4	4.5	89.8	103.7
	Work (J)	5.5	2.49	1.08	12.2
	Cooking losses	5.32	3.98	1.3	15.3
	Shear force (n/cm^2)	0.08	0.03	0.02	0.18
Tosella	L	89.1	2.05	84.5	92.8
	a	-0.52	0.84	-1.87	1.33
	b	11.1	3.05	7.41	18.4
	c	11.2	3.01	7.51	18.5
	H	93.9	5.26	85.9	103.2
	Work (J)	3.65	1.23	0.01	6.12
	Cooking losses	20.1	6.82	5.81	33.2
	Shear force (n/cm^2)	0.05	0.01	0	0.07

Mean spectrum of each group of cheeses are also plot and presented in Figure 6. Mean spectrum shows that paneer spectra from buffalo and cow are closer to each other. Also, for Tosella spectra is achieved the same result for two species which are closer to each other. Moreover, we can also observe that paneer has low reflectance than tosella and this can be expected because we have seen that tosella is slightly lighter or white than paneer (Table 3). A clear separation of paneer and tosella can be seen while separation between different species in same type of cheese is not so evident.

Preprocessed spectra are shown in Figure 7. Raw spectra were converted into pseudo-absorbance (Log (1/R)) and Standard Normal Variate (SNV), Savitsky Golay first derivate and then mean centering were used as preprocessing methods. Comparison of raw spectra and preprocessed spectra shows that the spectral preprocessing reduces the baseline variation and highlights the absorption peaks. (Wu et al., 2008) and (Ding et al., 2014) reported similar effects on the spectral response values in the process of measurement, including high frequency random noise, baseline drift, signal-to-background ratio, uneven concentration, light scattering, optical path change and others.

Cheeses spectra in general had a pattern similar to those previously reported for other studies related to cheeses. Osborne and Fearn (1986) reported some spectra ascribed to water : 970 nm and 1450 nm (second and first overtone of the O–H stretch vibration) and also bands at 1208 nm (a –CH stretch second overtone), 1728 and 1765 nm (first overtone of the –CH stretch) that arise from fat components. Spectra has information pertaining to the main regions, overtones and combinations of fundamental vibrations of C–H, N–H and O–H functional groups that are most prominent absorption bands in the NIR region.

Similar evidences are shown by Wehling and Pierce (1988), that determined cheese constituents by using reflectance spectroscopy. They reported the measurements in 3 wavelengths; 1818, 1734, and 1445 nm that are related respectively to the fat, protein contents and inversely related to moisture. Figure 7 shows the preprocessed spectra of cheese with SNV first derivate, which compared to original spectra became much more concentrated after pretreatments that eliminated the influence of baseline drift and particle scattering and also the signal-to- noise has been improved.
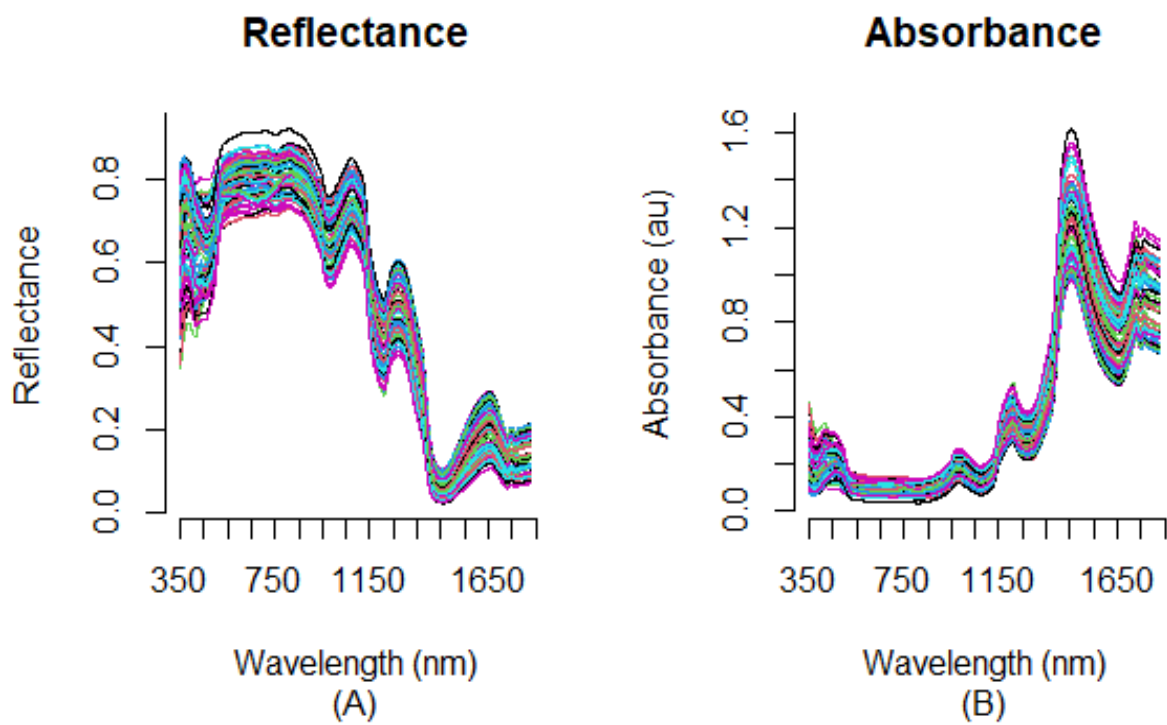
Figure 5. (A) NIR raw spectra of cheeses, (B) Reflectance to absorbance converted spectra.
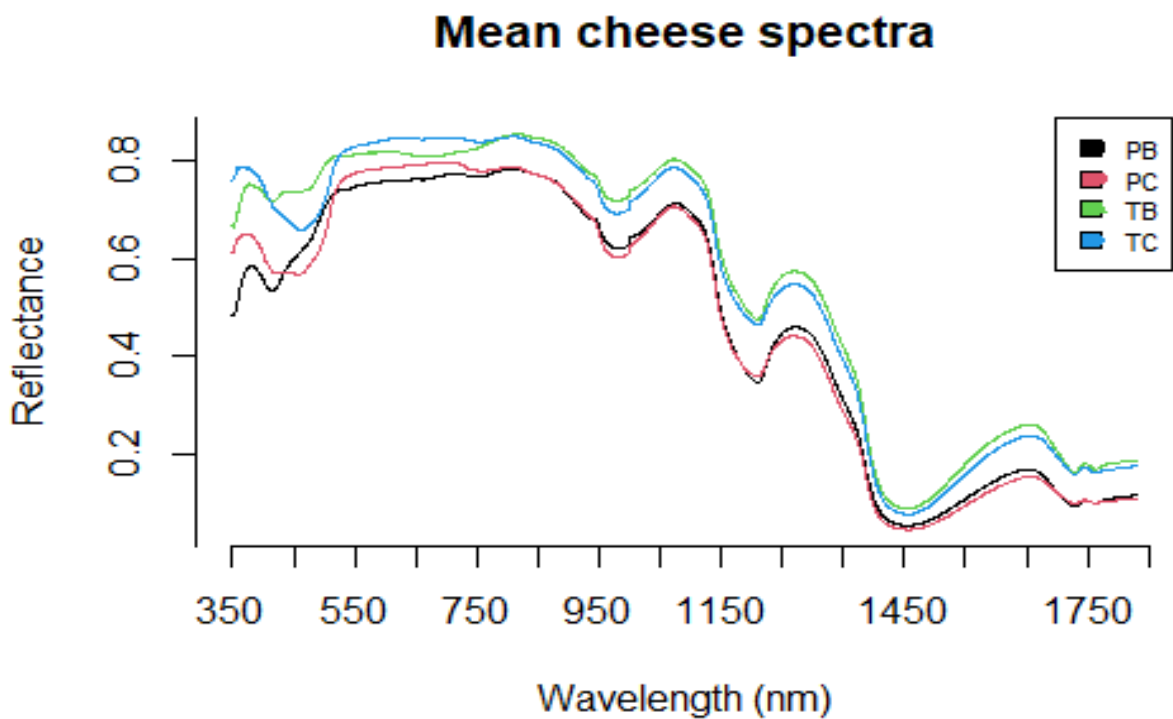

Figure 6. Mean spectra of different cheeses in reflectance mode.
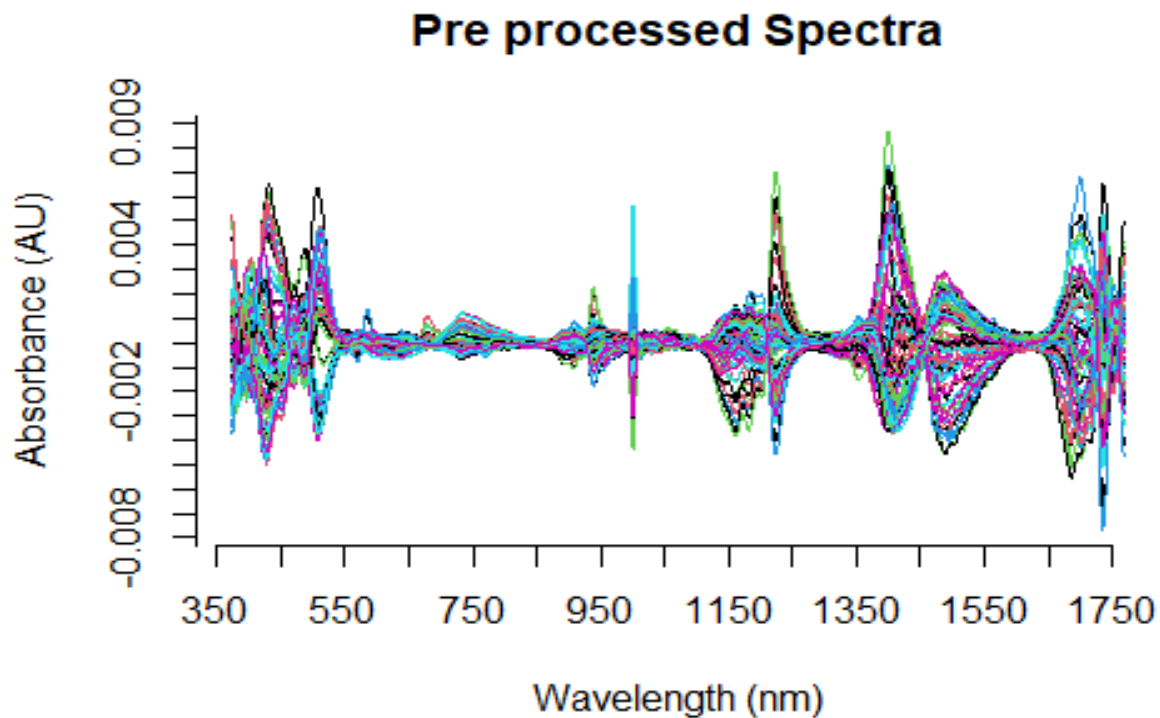
Figure 7. NIR preprocessed spectra of fresh cheeses

4.3 Principle component analysis

Cheeses spectra were analyzed using Principle component analysis. Initially cheeses spectra have high dimensionality which is reduced by PCA. PCA analysis results are shown in the table 4 which contains information about the proportional variability explained by each principle components (PC). Table 4 shows that PC1 explained 46% of variability from original spectra data. Moreover, it can be seen that 8 PCS have explained over 98% of variability of original spectra. Also, a scree plot in figure 8 between components and variability shows that 4-5 pcs are enough to explained the major variability of original spectra. Therefore, we have selected 5 pcs for further analysis.

The PCA scores plots and loading in figure 9 are shown which explain that PC1 vs PC2, PC2 vs PC3 and PC1 vs PC3. PC1 and PC2 in total explained over 71 % of variability while PC1, PC2 and PC3 together accounted for 85 % of total variability. Score plot results contain the four different groups of cheeses called paneer from buffalo milk (PB), Paneer from cow milk (PC), Tosella from buffalo

milk (TB) and tosella from cow milk (TC). PC1 is mainly separating types of cheeses but not very clearly. We can see that tosella is on positive side of PC1 while paneer is on the other side with slightly less evident. However, PC2 has a clear separation among two species of milk used in production of different cheeses. It can be seen from score plot that cheeses produced from buffalo milk are nicely separated from cow milk along the PC2 axis. Paneer buffalo (PB) and Tosella buffalo (TB) samples were well separated from the rest, forming two clusters and the remaining clusters was common for Paneer cow (PC) and Tosella cow samples. Furthermore, a plot between PC2 vs PC3 shows that PC2 is explaining the species differences while PC3 is about species differentiation. Although, a plot between PC1 vs PC3 shows the same differentiation.

The clustering of four types of cheeses at different positions in PCA score plot depicts that cheese samples (PB, PC, TB and TC) are slightly different in physical values from one another. However, separation between similar cheeses but different species are little blur. This is also evident in the raw mean spectra of cheeses. Moreover, cheeses from different species appeared similar in visible appearances. In order to explain the spectroscopic reasoning of classification (score plot PCA) and in order to see that what part of the spectra contributed to the PCA model, the loading plots for five PCs are shown in Figure 9. Inspection of loading plot shows that PC1 which explains mainly differences between different types of cheeses are mainly in visible area of spectrum as it can be seen in loading that PC1 from 550 nm to 750 nm are different than PC2. This is also in line with the fact that both the cheeses were different in terms of color appearances. Moreover, for PC2 which differentiate mainly the milk species, could be explained by loading probably at 860-870 which is lipid and very distinct peak at 1200 nm which is also a lipid peak. Similarly, PC3 which explain the species differentiation is characterized by loading at 1445 nm which is related to water and shows the differences due to the moisture content for type of cheeses. A survey of literature explains similar views such as wavelengths at 1818, 1734, and 1445 nm may be related respectively to the fat, protein contents and inversely related to moisture (Wehling and Pierce 1988). Thus, these components contributed largely to PC1, PC2 and PC3 loadings plot, differentiating the spectral profiles of these

by PC4 and PC5 types. They reveal that PC1, PC2 and PC3 are positive in the most intense bands of the whole spectra. PC3 and PC4 spectra ascribed to water and fat 1450 nm (first overtone of the O–H stretch vibration) and also PC4 bands at 1208 nm (a –CH stretch second overtone). The peaks and troughs of the loadings for PC 1 and PC 2 were at the same wavelengths at 1728 and 1765 nm (first overtone of the –CH stretch) that arise from fat components (Osborne and Fearn 1986). Loading plots also hint at how variables correlate with one another: a small angle implies positive correlation, a large one suggests negative correlation, and a 90° angle indicates no correlation between two characteristics.

Table 4. Results of PCA analysis showing the proportional variability explained by each principle components (PC).

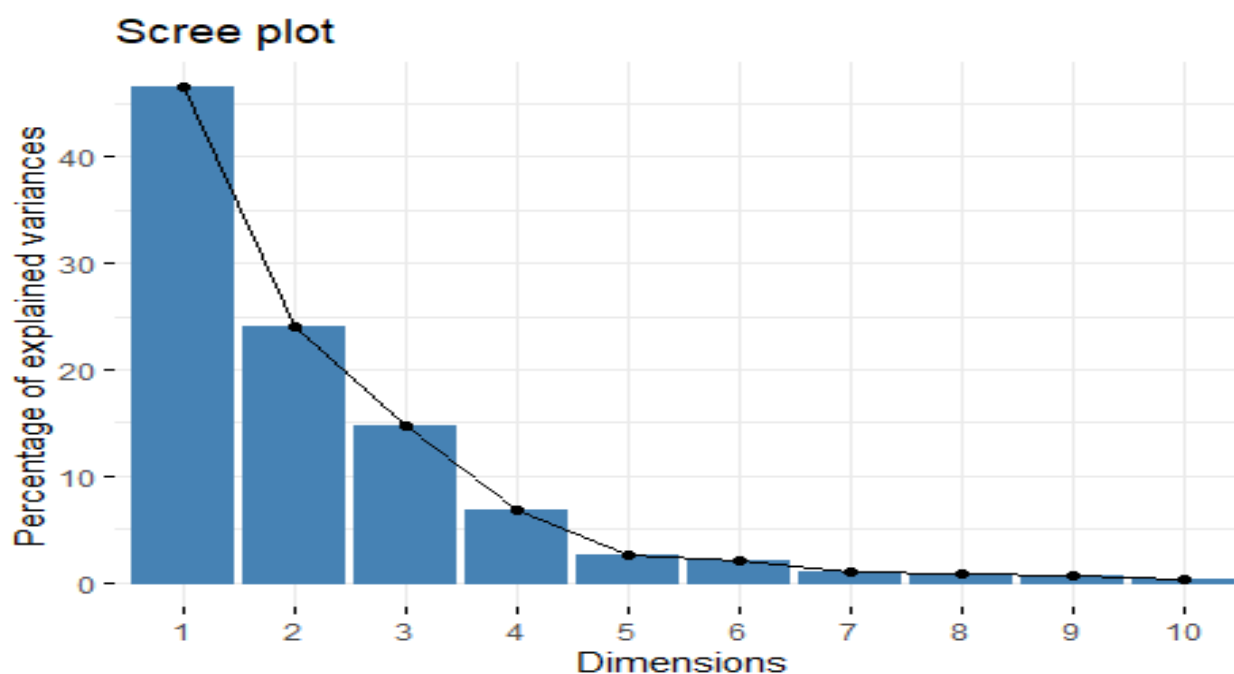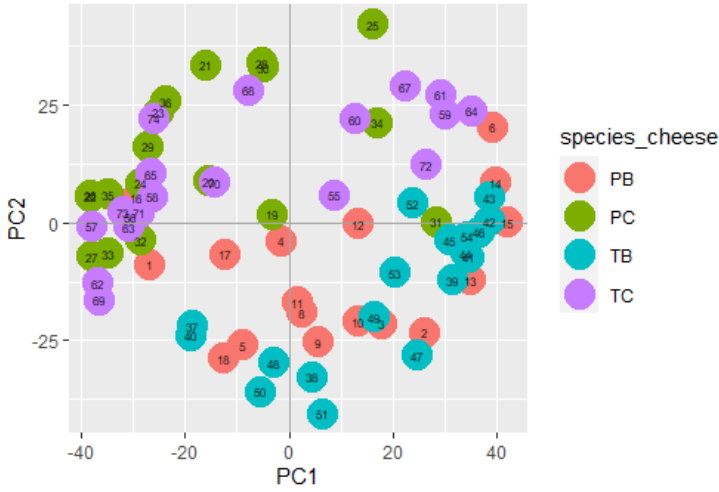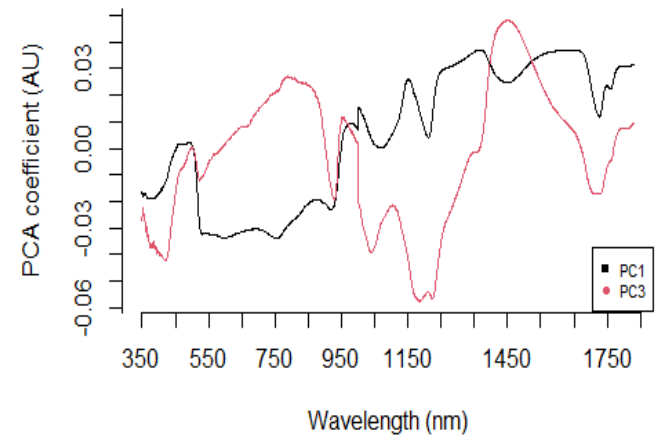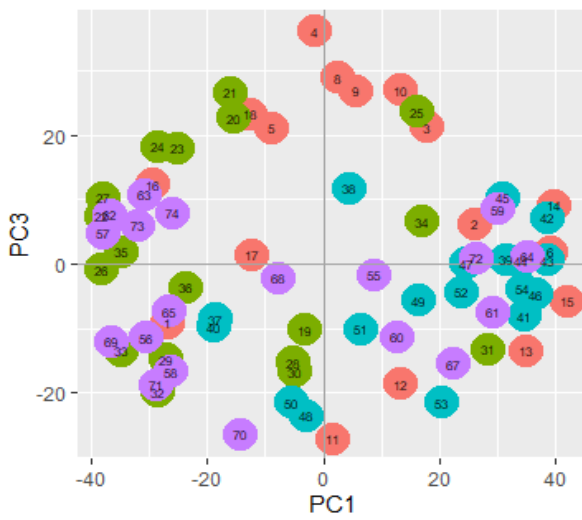|	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Standard deviation	26.24	18.88	14.79	10.05	6.27	5.46	3.70	3.37
Proportion of Variance	0.46	0.24	0.15	0.07	0.03	0.02	0.01	0.01
Cumulative Proportion	0.46	0.71	0.85	0.92	0.95	0.97	0.98	0.98

Figure 8. PCA scree plot.

PCA Fresh_cheeses

Loadings
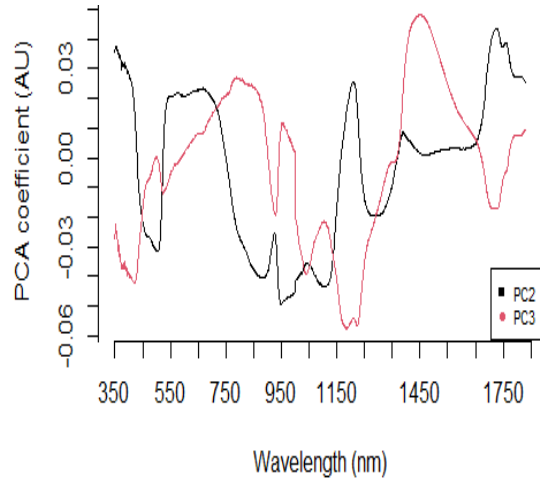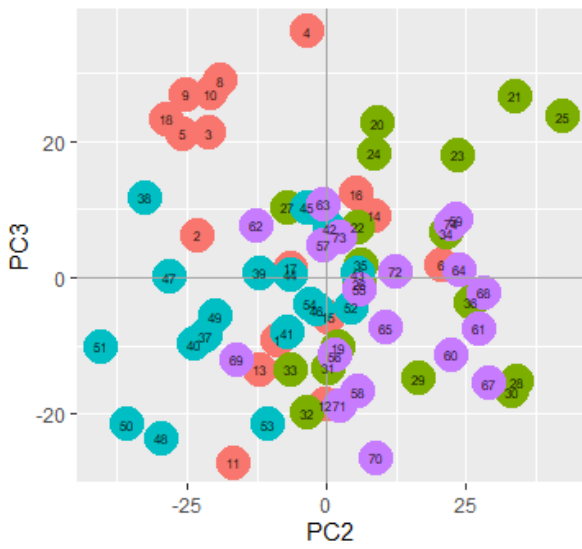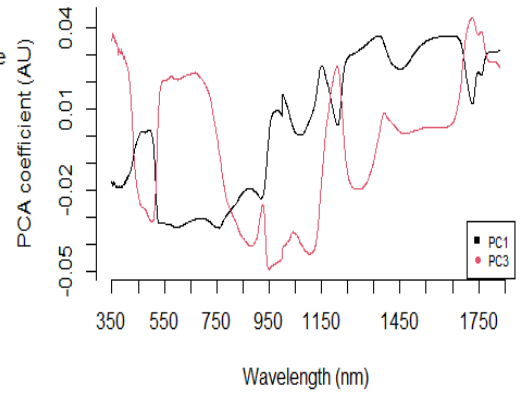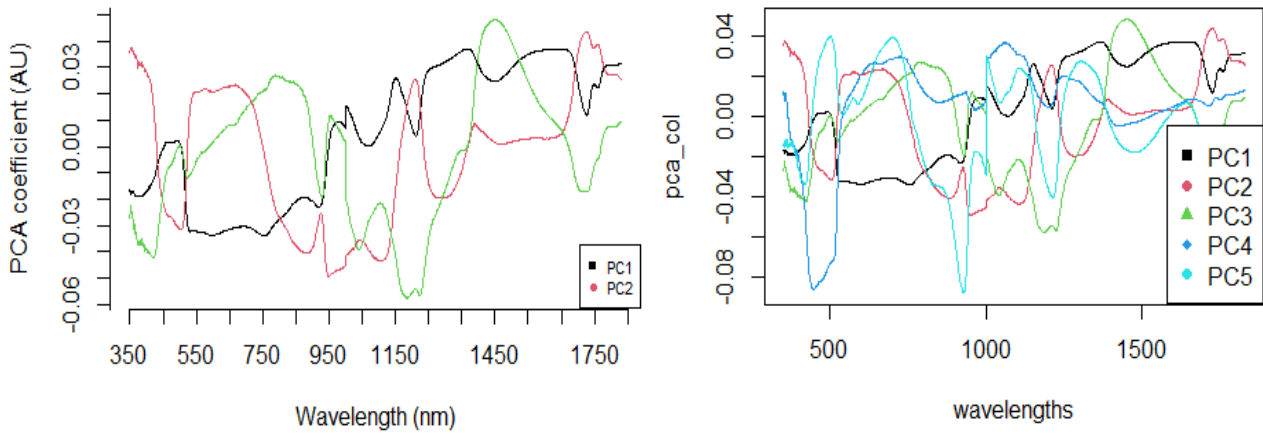
Figure 9. PCA score plot and PCA loadings of cheeses spectra.

4.4 Hierarchical cluster analysis

HCA was performed to classify the cheeses based on their dissimilarities without considering information about the class membership. As can be seen from the figure 10, optimum number could be 4 clusters. Table 5 shows the clustering accuracy using different methods to calculate the clusters. We can see that Ward methods has yield better results than other three methods. Therefore, this method was used for clustering the cheeses. Based on the results, cluster membership can be seen and visualized. A dendrogram plotted with four clusters is shown in Fig. 11. Hierarchical cluster analysis on the fresh cheese data was used to cluster closely related cheese samples in terms of the manufacturing process, coagulation conditions, cheese composition and origin of milk. The first cluster identified by HCA grouped 18 Paneer samples prepared from buffalo milk. Cluster 2 contained 16 Paneer cheeses produced by cow milk. Cluster 3 grouped 23 Tosella buffalo milk cheese and the fourth cluster had 15 Tosella cheese samples manufactured from cow milk. The first 2 and last 2 clusters were grouped in 2 main consortiums based on the way of manufacturing (use of different coagulants, temperature of coagulation etc.). The PCA and HCA results were similar regarding classification of cheeses based on their species and types.

Application of Hierarchical analysis was used on Kashar cheeses by Eroglu et al., (2013). As in our study even their results were proved by the HCA, confirming the PCA correlations. Their results were useful to understand which physicochemical properties would contribute to the texture of cheese samples by providing a basis for the development of a clear classification of cheeses.

We concluded, in a similar way to Murtaza et al., 2013, that grouping and subgrouping on the dendrogram and the positions of cheese samples on PCA plot indicated that all parameters significantly influenced the physical characteristics. The differences in these parameters are mainly depending on the manufacturing process, coagulation conditions, cheese composition and origin of milk.
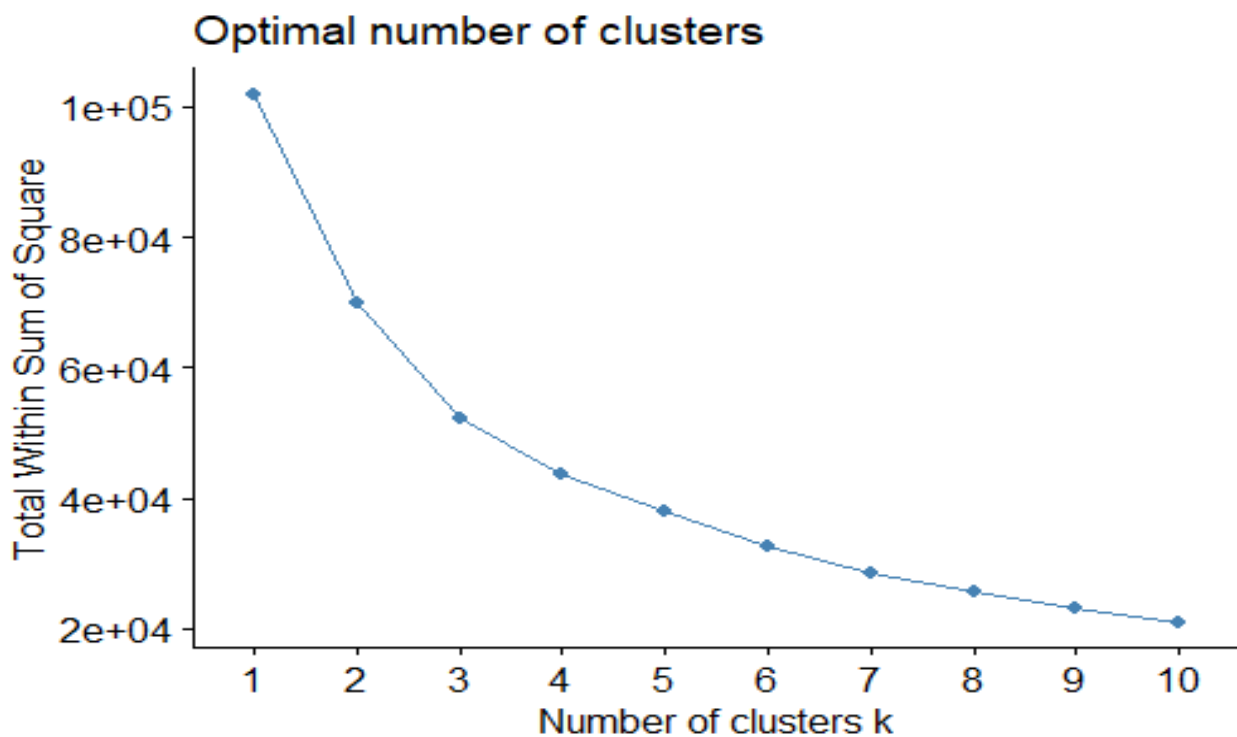
Figure 10. Scree plot for HCA.

Table 5. Hierarchical analysis of fresh cheeses.

Map dbl	Different methods to calculate accuracy		
average	single	complete	ward
0.725	0.552	0.818	0.932
Cluster membership	Calculated by ward method		
1 (PB)	2(PC)	3(TB)	4(TC)
			15
18	16	23	
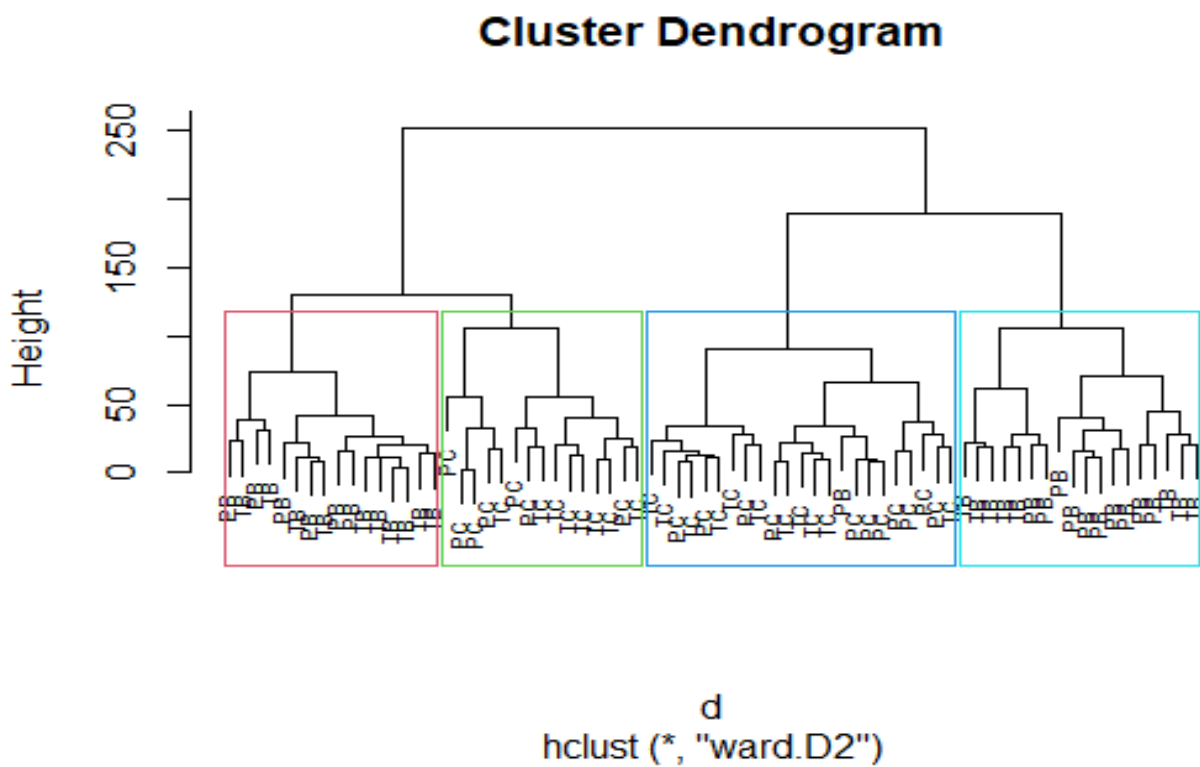
Figure 11. Dendrogram obtained from hierarchical cluster analysis (HCA) of physical data of fresh cheeses using Ward's clustering algorithm. PB-Paneer Buffalo; PC-Paneer Cow; TB-Tosella Buffalo and TC- Tosella Cow.

4.5 PCADA

PCADA method was used to classify the cheese samples according to their milk origin (buffalo or cow) and their manufacturing technique (Paneer and Tosella). The classification of clusters TB, PB, TC and PC has been demonstrated in figure 12. PCA with linear discriminant analysis (PCA-LDA) was applied to the scaled and mean centered NIR spectra. To perform the discrimination, PCA-LDA was used with five PCs because they accounted for 95% of the data variance (table of PCs). The final number of PCs was defined according to the distribution of the variance for each PC, such that the minimum number of PCs accounted for the maximum variance, which occurred before the variance reached a small and constant trend, was selected.
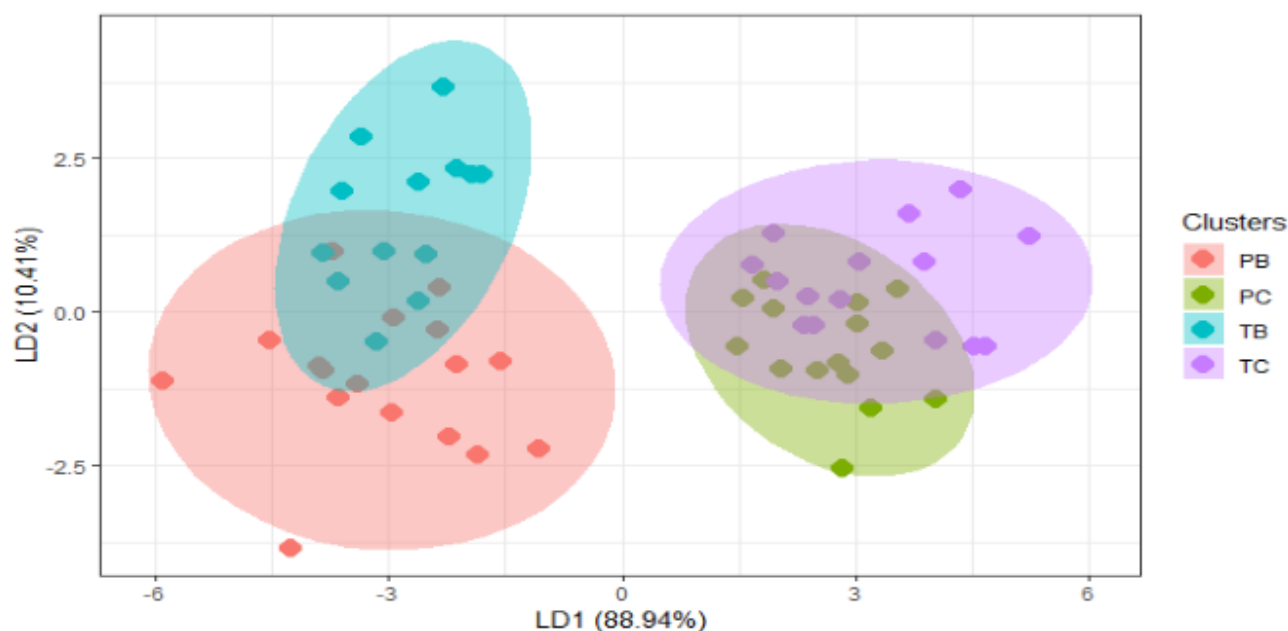
Figure 12. Score plot of classification of fresh cheese samples of the PCDA, according to cheese milk origin (composition) and manufacturing technique.

By applying the PCA-LDA it was possible to obtain satisfactory discrimination between all four different groups, and the classification rates for the calibration and prediction sets were 0.867 and 0.917%, respectively (Table 6). It was possible to visualize a clear separation between two species of

milk used cows and buffalos. However, separation of different cheeses within same types of milk is little overlapping or not very clear. Moreover, histogram in fig.13 also shows that PB is well separated from TC and PC, however PB is similar to TB and PC is overlapping with TC. Using the same chemometrics method, Pillonel et al. (2003) achieved good results for classification of the geographical origin of Emmental cheeses.
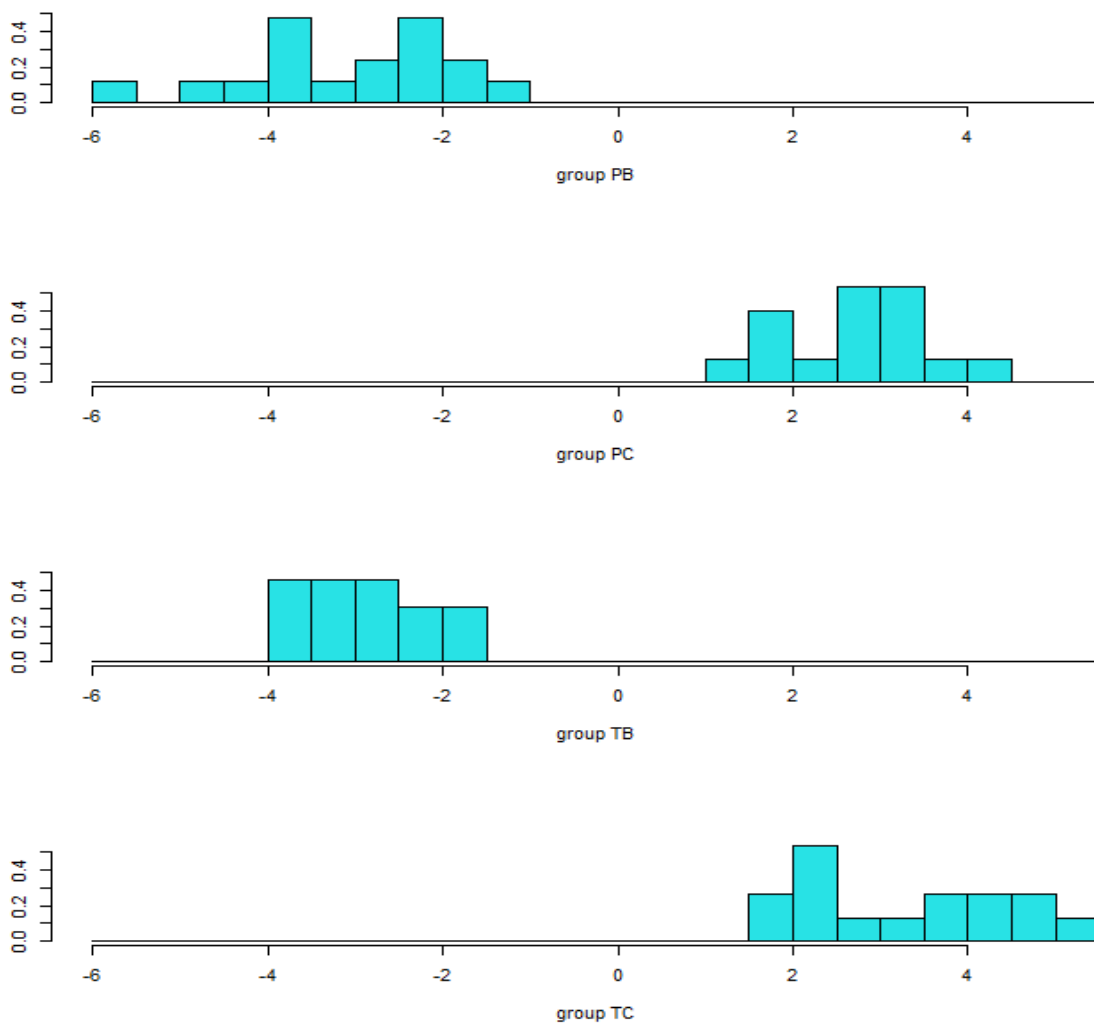
Figure 13. PCDA histogram showing the location of fresh cheeses groups PB-Paneer Buffalo; PC-Paneer Cow; TB-Tosella Buffalo and TC- Tosella Cow, in plot score

Table 6. PCDA analysis for classification of fresh cheeses.

Actual predicted		PCDA calibration				PCDA prediction			
		PB	PC	TB	TC	PB	PC	TB	TC
	PB	15	0	2	0	0	0	0	0
	PC	0	13	0	2	0	2	0	0
	TB	2	0	11	0	0	0	5	0
	TC	0	2	0	13	0	1	0	4
Accuracy		0.867				0.917			

4.6 SIMCA

In the present study, the single PCA model for SIMCA classification was developed with 72 cheese samples after removing outliers. SIMCA as a method of classification was used to classify and also predict memberships of each class. Each defined group in dataset was designated using PCA model which was established by calibration samples. Then, the developed models were applied to classify of validation set members.

After developing a SIMCA model on the basis of measured spectra, the number of correct classifications per total number of each type of validation sample is listed in Table 7. From the four clusters TB yielded the highest level of specificities and accuracy 100% , as seen by the data points being well-separated from the TB cluster (Fig. 14A and B) and the associated high prediction capability (Table 7).The specificity of the other clusters is dependent on the type of cheese sample and is associated with NIR activity showing a unique and/or intense NIR spectral features from that of TB that resulted in a higher specificity.

However, SIMCA showed good values of classification and calibration, as observed in the table 7 around 89% of safe. Fig.14 displays the Coomans' plot which show the distance from each sample

of developed model in each class (PB, PC, TB and TC). Coomans' distance shows that cow and buffalos are well separated while distances between within species for different cheeses were close.

SIMCA can be used to classify the cheeses based on milk species and types with a rational level of precision. Although, additional investigations with other different cheeses should be examined to verify this result. This is an outcome we might have expected because the clear separation between cheese with different types and buffalo classes were seen in the overall PCA and HCA analysis. Therefore, supervised methods (PCA-LDA and SIMCA) could be more reliable than Unsupervised methods (PCA and HCA) to classify the cheeses types within the same milk used for its production.

Table 7. Summary for calibration and classification of fresh cheeses. PB-Paneer Buffalo; PC-Paneer Cow; TB-Tosella Buffalo and TC- Tosella Cow.TP-True positive; FP-False positive; TN- True negative; FN-False negative; Spec- specificity; Sens- sensitivity.

Calibration
Summary for calibration results

	N.comp	TP	FP	TN	FN	Spec.	Sens.	Accuracy
PB	1	13	22	22	0	0.5	1	0.614
PC	1	13	13	31	0	0.705	1	0.772
TB	1	13	5	38	1	0.884	0.929	0.895
TC	1	17	15	25	0	0.625	1	0.737

Prediction
Summary for SIMCA multiple classes classification result

	N.comp	TP	FP	TN	FN	Spec.	Sens.	Accuracy
PB	1	4	5	6	0	0.545	1	0.667
PC	1	4	3	7	1	0.7	0.8	0.733
TB	1	4	0	11	0	1	1	1
TC	1	2	7	6	0	0.462	1	0.533
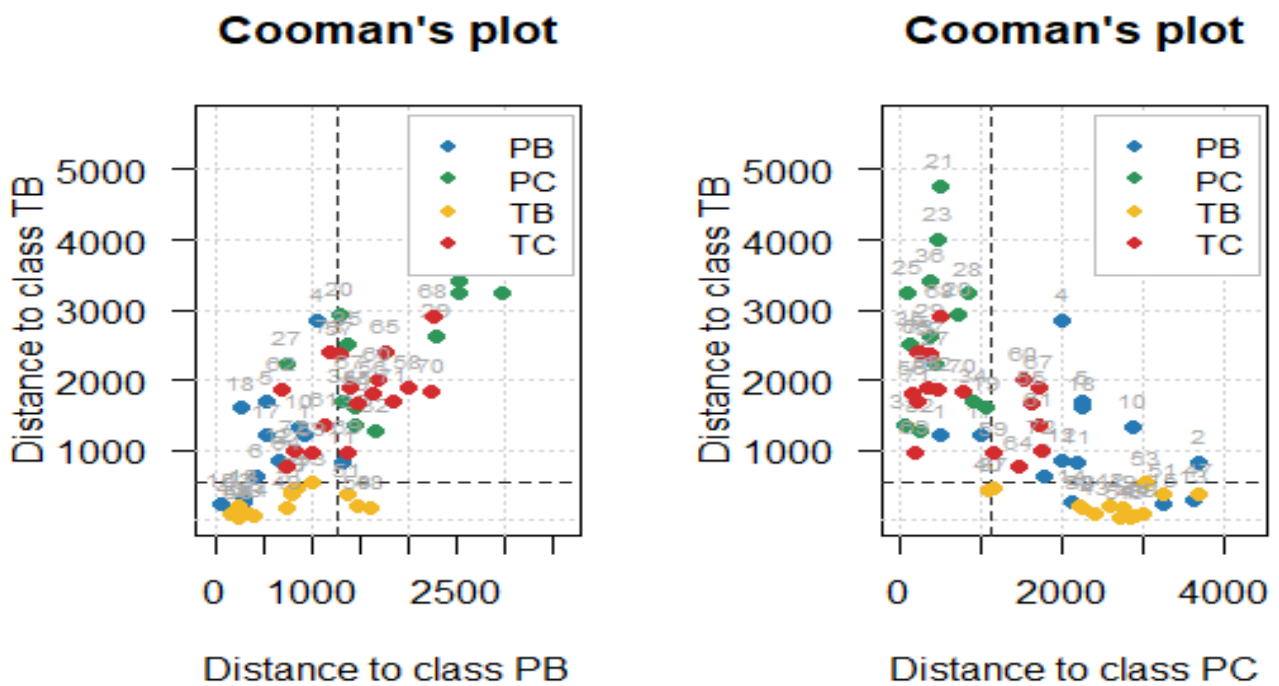
Figure 14 (A) Coomans plot showing spectroscopic separation of groups PB and TB, into the corresponding class PB and class TB by SIMCA. The abscissa denotes the distance to class PB. The ordinate denotes the distance to class TB. (B) Coomans plot showing spectroscopic separation of groups PC and TC into the corresponding class PC and class TC by SIMCA. The abscissa denotes the distance to class PC while the ordinate denotes the distance to class TB.

4.7 Prediction of cheese qualities using PLSR

The optimal calibration models for determining the color, texture and cooking losses of fresh cheeses are constructed based on the lowest RMSEC and the highest correlation coefficient in calibration (RC2). The values of these statistical parameters as well as the number of latent variables of the optimal calibration models are presented in Table 8. High performance was calculated with coefficient of determination (R2) values in prediction ranging from 0.45 to 0.71 and root mean square errors of prediction (RMSEP) ranging from 0.62 to 1.45. The model with the highest R2 with a value of 0.71 as well as the lowest RMSEP with a value of 0.62 was PLS model of predicting the color (a*)

of fresh (raw) cheeses. Besides the prediction, even R2C (0.93) and RMSEC (0.24) showed good values, using 9 LVs, for predicting (a*) color attribute. The lowest performance was for cooking losses prediction where R2C showed a low value of 0.45 and RMSEC a high value of 0.72 using only 2 LVs.

These results can be compared to other works done by several authors in the last decade for prediction of different cheese compounds. Stocco et al., (2019) used PLS regression method to predict the texture and color traits of 1050 different cheeses. Unlike our study, they used three instruments Vis NIRS-R, NIRS-R and NIRS-T for the acquisition of raw spectra. Based on the reported results; texture traits had a low fit (R2 cal = from 0.37 to 0.47), although it was better than work (R2 cal = from 0.14 to 0.17). Anyway, both these traits had lower prediction values compared to ours. The lightness color trait was predicted with a similar moderate accuracy as our study (R2 cal= from 0.62 to 0.72) by all the instruments and all the spectral ranges, with the best result from Vis NIRS-R, followed by NIRS-R and NIRS-T. Accuracy of prediction of all the other cheese color traits by all the instruments and spectral intervals was very low (R2 cal = from 0.04 to 0.65), compared to our results ranging from 0.77 to 0.93 for other color traits.

Different results from ours, were obtained by Kraggerud et al., (2014) who reported R2 cv lower than 0.70 for prediction models of most FAA determined on intact samples of Norwegian cheese using NIR spectroscopy in reflectance mode. In their study NIR was not an effective technique for prediction of characteristics of cheeses. He used then a Fourier-transform MIR spectroscopy in transmittance mode, which performed generally better (R2 cv between 0.75 and 0.85) than NIR spectroscopy.

Therefore, it can be concluded that comparing to other literature, the model for prediction of physical properties of fresh cheeses fitted well the experimental values and made good predictions for the percentage of physical attributes. These results are due to the fact, that few laboratory steps are involved in reference laboratory analysis because if more steps are involved in analysis, the error of

laboratory increases and this will directly affect the accuracy of the prediction models (De Marchi et al., 2014). Another explanation of these results would be the high variability of physical parameters in fresh cheeses produced by two different type of milk, collected by different breeds.

Table 8. Figures of merit of the PLS prediction models for physical attributes in fresh cheeses. LVs, latent variables; R2P, coefficient of determination in prediction; RMSEC, root mean square error of calibration; RMSEP, root mean square error of prediction.

	Calibration			Prediction	
	LVs	R2c	RMSEC	R2P	RMSEP
L	4	0.654	0.518	0.606	0.737
a	9	0.933	0.24	0.718	0.623
b	5	0.78	0.485	0.451	0.635
c	5	0.779	0.485	0.457	0.634
h	8	0.876	0.336	0.554	0.741
Max_work	9	0.86	0.334	0.633	0.787
Work to cut	4	0.73	0.445	0.632	0.847
Total work	6	0.77	0.412	0.646	0.819
Cooking losses	2	0.46	0.728	0.6	1.45

5. Conclusions

The main goal of this study was to show the potential use of NIR spectroscopy as a rapid, non-destructive and chemical free technique for discrimination of fresh cheeses and to predict physical qualities of Tosella and Paneer. We concluded that NIR spectroscopy together with chemometric techniques such as supervised and unsupervised methods showed a good ability to achieve our objectives. In general, Principal Component Analysis (PCA) for qualitative analysis of spectra showed a clear separation among two species of milk (buffalo and cow) used in production of fresh cheeses and Hierarchical cluster analysis was used with a high efficacy for clustering similar cheeses. In the end was illustrated that PCA and HCA results were similar regarding classification of cheeses based on their species and types. PCA-DA model was used on the basis of PCA showing a good classification with 99.35 % of total variance and a good accuracy of calibration and prediction, 86 % and 91 % respectively. We noticed that supervised methods (PCA-LDA and SIMCA) could be more reliable than Unsupervised methods (PCA and HCA) for classification of cheeses types within the same milk used for its production. In addition to the classification of cheeses, PLSR was used for prediction of physical quality with coefficient of determination ($R2$) values in prediction ranging from 0.45 to 0.71 and root mean square errors of prediction (RMSEP) ranging from 0.62 to 1.45. To conclude, we can say that the explanation of these results, illustrate the good potential use of this spectroscopic technology as a useful and a low-cost laboratory tool for monitoring the quality of cheese analysis, giving the consumer the right information related to fresh cheeses.

6. NIRS limitations

NIR spectroscopy technique is quick and easy tool for classification and quality prediction of different dairy products. However, this tool like many other spectral analysis techniques, has some technical limitations which should be mentioned before applying this technique. For example, sample processing is needed in some cases and the error in the sample loading is also a factor which can affect the spectral features of similar samples. In most of the cases, the calibration is required from the references which is a time-consuming process because a specific calibration is needed for each type of species. In our opinion the main limitation to this technique is the instability of calibration because samples change their characteristics over time as well as a uniform instrumentation is not available around the world. Anyway, besides of some technical limitations, NIRS is still a promising alternative to the costly food analysis techniques.

7. Recommendation for future work

Cheese, with its enzymatic and production processes and milk as a quite variable raw material, is one of the most complicated products to be analyzed for its quality. More attempts are still required for the validation of the developed model using high numbers of samples. But it is expected that the ability to measure relevant information from raw materials and the production process, in order to control and predict cheese quality, will certainly be further developed in future. Some more advance chemometrics tools can be utilized to see the better classification and prediction of cheeses qualities.

REFERENCES

Ahmed, A., and Bajwa, U. (2019). Composition and microstructure appraisal of paneer coagulated with sour fruit juices. J. Food Sci Technol vol 56, 253-261.

Albenzio, M., Corbo, M.R., Rehman, S.U., Fox P.F., De Angelis, M., Corsetti A and Gobbetti, M. (2001). Microbiological and biochemical characteristics of Canestrato Pugliese cheese made from raw milk, pasteurized milk, or by heating the curd in hot whey. Int. J. Food Microbiol vol 67, 35-48.

Almena-Aliste, M., and Mietton, B. (2014). The microbiology of traditional hard and semi-hard cooked mountain cheeses. Microbiol Spectrum vol 2 (1), CM -0003-2012.

Andrade, Pereira J., C. G., Almeida Junior., de Viana J. C., C. C. R., Neves, and de O.,Silva L. N (2019). FTIR-ATR determination of protein content to evaluate whey protein concentrate adulteration. Food Science and Technology vol 99, 166–172.

Angelino, Steven, A.G.F. (1996). Determination of the moisture and nitrogen contents of barley and malt by NIRS. Journal of the institute of Brewing, vol 102, 73-74.

Arnould and Soyeurt 2009 - Arnould, R., and Soyeurt, H. (2009). Genetic variability of milk fatty acids. J Appl Genet 50(1), 29-39.

Barnes et al., 1989 - Barnes, R. J., Dhanoa, M. S., and Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Applied Spectroscopy vol 43(5), 772–777.

Beuvier, E., Berthaud, K., Cegarra, S., Dasen, A., Poche, S., Buchin, S and Buboz, G (1997). Ripening and quality of Swiss-type cheese made from raw, pasteurized or microfiltered milk. Int. Dairy J. vol 7, 311-323.

Bottero MT, Dalmasso IA, Nucera D, et al. . (2003). Development of a PCR assay for the detection of animal tissues in ruminant feeds. J Food Prot vol 66(12), 2307-2312.

Cevoli Chiara, Fabbri Angelo, Gori Alessandro, Caboni Maria Fiorenza and Guarnieri Adriano (2013). Screening of grated cheese authenticity by NIR spectroscopy. Journal of Agricultural Engineering, vol XLIV(s2): e53.

Curro, S., Manuelian, C.L., Penasa, M., Cassandro, M., and De Marchi, M. (2017). Technical note: Feasibility of near infrared transmittance spectroscopy to predict cheese ripeness. J. Dairy Sci. vol 100, 8759-8763.

De Marchi, M., Penasa, M., Zidi, A.,† and Manuelian, C. L., (2018). Invited review: Use of infrared technologies for the assessment of dairy products. Application and perspectives. J. Dairy Sci. vol 101, 10589-10604.

De Noo, Mirre, E. (2005). Reliability of human serum protein profiles generated with C8 magnetic beads assisted MALDI-TOF Mass spectroscopy. Anal chem. vol 77(22), 7232-7241.

Demarigny, Y., Beuvier, E., Buchin, S., Pochet, S., and Grappin, R., (1997). Influence of raw milk microflora on the characteristics of Swiss-type cheeses: II. Biochemical and sensory characteristics vol 77, 151-167.

Ding, R., Huang,X, Han, F., Dai, H., Teye, E., and Xu, F., (2014). Rapid and nondestructive evaluation of fish Freshness by near infrared reflectance spectroscopy combined with chemometrics analysis. Analytical Methods vol 6, 9675-9683.

Dvorak Lukas, Mlcek Jiri and Sustova Kvetoslava (2016). Comparison of FT-NIR Spectroscopy and ELISA for Detection of Adulteration of Goat Cheeses with Cow's Milk. Journal of AOAC International vol 99, 180-186.

Eroglu Ali, Dogan Mahmut, Toker Omer Said and Yilmaz Mustafa Tahsin (2015). Classification of Kashar Cheeses Based on Their Chemical, Color and Instrumental Textural Characteristics Using Principal Component and Hierarchical Cluster Analysis. International Journal of Food Properties vol 18, 909–921.

Eurostat, 2018- https://ec.europa.eu/eurostat/statistics explained/index.php/Milk_and_milk_product_statistics#Milk_production (November 2019).

Frank, J. F., and Birth, G. S., (1982). Applications of near infrared reflectance spectroscopy to cheese analysis. J. Dairy Sci. vol 65, 1110-1116.

Frankhuizen, R., and Van der Veen, N. G., (1985). Determination of major and minor constituents in milk powders and cheese by near infrared reflectance spectroscopy. Neth. Milk Dairy J. vol 39, 191-207.

Guggisberg Dominik, Fröhlich-Wyder Marie-Therese, Irmler Stefan, Greco Mark, Wechsler Daniel and Schuetz Philipp (2013). Eye formation in semi-hard cheese: X-ray computed tomography as a non-invasive tool for assessing the influence of adjunct lactic acid bacteria. Dairy Sci. and Technol. vol 93, 135–149.

Gurbanov, R., Bilgin, M., and Severcan, F., (2016). Restoring effect of selenium on the molecular content, structure and fluidity of diabetic rat kidney brush border cell membrane. Biochimica et Biophysica Acta (BBA)-Biomembranes vol 4, 845-854.

Huang, Y.F., Huang, C.C., Hu, C.C., and Chang, H.T. (2006). Capillary electrophoresis- based separation techniques for the analysis of proteins. Electrophoresis vol 27, 3503–3522.

Jotarkar Parameshwar, Panjagari Narender Raju, Singh Ashish Kumar and Arora Sumit (2017). Effect of whey protein-iron based edible coating on the quality of Paneer and process optimization. International Journal of Dairy Technology, vol 71, 395-407.

Kamruzzaman, M., Barbin, D., ElMasry, G., Sun D.W., and Allen, P. (2012). Potential of hyperspectral imaging and pattern recognition for categorization and authentication of red meat. Innovative Food Science and Emerging Technologies vol 16, 316-325.

Karoui, R., and Dufuor, E., (2005). Prediction of the rheology parameters of ripened semi-hard cheeses using fluorescence spectra in the UV and visible ranges recorded at a young stage. Int, Journal Dairy vol 16, 1490-1497.

Karoui Romdhane, Mouazen Abdul Mounem, Dufour Éric, Pillonel Laurent, Emmanuelle Josse Jacques-Olivier et al. (2006). Chemical characterization of European Emmental cheeses by near infrared spectroscopy using chemometric tools. International Dairy Journal vol 16, 1211-1217.

Kraggerud, Hilde, Næs, Tormod and Abrahamsen, Roge (2014). Prediction of sensory quality of cheese during ripening from chemical and spectroscopy measurements. International Dairy Journal vol 34, 6-18.

Kumar Sunil, Rai, D. C.,Niranjan K. and Bhat Zuhaib F., (2011). Paneer, an Indian soft cheese variant: a review. J Food Sci Technology vol 51, 821–831.

Legg K.Andrew and Johnston A. Keith (2017). General aspects of cheese technology. Cheese fourth edition vol 26, 643-675.

Lenart, J., Szigedi, T., Dernovics, M., and Fodor, M., (2012). Application of FT-NIR spectroscopy on the determination of the fat and protein contents of lyophilized cheeses. Acta Alimentaria, vol 41, 351–362.

Liland Kristian Hovde, Mevik Bjørn-Helge, Rukke Elling-Olav, Almøy Trygve and Isaksson Tomas (2009). Quantitative whole spectrum analysis with MALDI-TOF MS. vol 99, 39-48.

Lopes-Calleja I.M., Gonzales, I., Fajardo, V., Hernandez, P.E., Garsia, T., Martin, R. (2007). Application of an indirect ELISA and a PCR technique for detection of cow milk in sheep`s and goat`s milk cheeses. International Dairy Journal vol 17, 87-93.

McQueen, D. H., Wilson, R., Kinnunen, A., and Jensen, E. P (1995). Comparison of two infrared spectroscopic methods for cheese analysis. Talanta vol 42, 2007–2015.

McSweeney, P.L.H., and Fuquay, J.W., (2011) Biochemistry of Cheese Ripening. Encyclopedia of Dairy Sciences (Second Edition) vol 1, 667-674.

McSweeney Paul L.H., Ottogalli Giorgio and Fox Patrick.F (2017). Diversity and Classification of cheese varieties: An overview: Cheese, Chemistry, Physics and Microbiology vol 4, 781-808.

Murtaza, M. A., Rehman S. U., † Anjum, F. M. † and Huma N. (2013). Descriptive sensory profile of cow and buffalo milk Cheddar cheese prepared using indigenous cultures. J. Dairy Sci. vol 96, 1380–1386.

Nache, M., Scheier, R., Schmidt, H., and Hitzmann, B. (2015). Non-invasive lactate- and pH monitoring in porcine meat using Raman spectroscopy and chemometrics. Chemometrics and Intelligent Laboratory Systems vol 142, 197-205.

Osborne, B. G., and Fearn, T., (1983). Collaborative evaluation of near infrared reflectance analysis for the determination of protein, moisture and hardness in wheat. J of food and science vol 34, 1011-1017.

Osborne, B. G., and T. Fearn. (1986). Near Infrared Spectroscopy in Food Analysis. Longman Scientific and Technical, Essex, England.

Osborne, N. G., Fearn, T, and Hindle, P. H, (1993). Practical NIR Spectroscopy with Applications in Food and Beverage Analysis. American Journal of Analytical Chemistry vol 8, 227.

Perry, K. S. P. (2004). Cheese: Chemical, Biochemical and Microbiological aspects. New chemistry vol 27, 293-300.

Pillonel, L., Ampuero, S., Tabacchi, R., et al. (2003). Analytical methods for the determination of the geographic origin of Emmental cheese: volatile compounds by GC/MS-FID and electronic nose. Eur Food Res Technol vol 216, 179-183.

Plath, A., Krause, I. and Einspanier (1997). Species identification in dairy products by three different DNA-based techniques. Z. Lebensm Unters Forsch vol 205, 437-441.

Poonia Amrita, JHA Alok, Sharma Rajan, Singh Harikesh Bahadur, Kumar Ashwini, Sharma Rai and Sharma Nitya (2017). Detection of adulteration in milk: A review. International Journal of Dairy Technology Vol 70, 23-42.

Powell I.B., and Limsowtin M.C., (2016). Starter cultures: general aspects. Encyclopedia in dairy science, 552-558.

Recio, I., Garcıa-Risco, M.R., Lopez-Fandino, R., Olano, A., and Ramos, M., (2000). Detection of rennet whey solids in UHT milk by capillary electrophoresis. International Dairy Journal vol 10, 333-338.

Rodriguez-Otero, J. L., and Hermida (1996). M. Analysis of fermented milk products by near infrared reflectance spectroscopy. J. AOAC Int. vol 79, 817-821.

Rodriguz-Otero, J.L., Veiga, M.J., Castano, L., Centeno, J.A., and Hermida, M. (1997). Application and performance of the Meatspec for cheese analysis. Milchwissenschaft vol 53, 503-505.

Sachdeva, S., and Singh, S., (1988). Optimization of processing parameters in the manufacture of paneer. J Food Sci Technol vol 25, 142-145.

Sahu Pranita, Agrawal, A. K., Sinha Geetesh, Sahu C and Sahu Ruchi (2017). Effect of temperature and time on some physical characteristics of Paneer whey. Asian Journal of Dairy and Food research vol 36, 132-136.

Sakamoto Seiichi, Putalon Waraporn and Morimoto Satoshi (2018). Enzyme -linked immunosorbent assay for the quantitative/qualitative analysis of plant secondary metabolites. Journal of natural medicines vol 72(1):43.

Sasic, S., and 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. Anal. Chem. vol 73, 64-71.

Savitzky, A., and Golay., M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry vol 36, 1627-1639.

Settanni, L., and Moschetti, G. (2010). Non-starter lactic acid bacteria used to improve cheese quality and provide health benefits. Food Microbiol. vol 27, 691-697.

Skeie Siv, Feten Guri, Almoy Trygve, Ostlie Hilde and Isaksson Tomas (2006). The use of near infrared spectroscopy to predict selected free amino acids during cheese ripening. International Dairy Journal vol 16, 236-242.

Stocco Giorgia, Cipolat-Gotet Claudio, Ferragina Alessandro, Berzaghi Paolo, and Bittante Giovanni (2019). Accuracy and biases in predicting the chemical and physical traits of many types of cheeses using different visible and near-infrared spectroscopic techniques and spectrum intervals. J. Dairy Sci. vol 102, 9622-9638.

Srivastava, S., and Goyal, GK. (2007). Preparation of paneer—A review. Indian J Dairy Sci vol 60(6):377–388.

Straughan, B.P., and Walker, S (1976). Electronic spectra of polyatomic molecules. Spectroscopy, 120-160.

Tao Feifei and Ngadi Michael (2017). Applications of spectroscopic techniques for fat and fatty acids analysis of dairy foods. Current opinion in Food Science vol 17, 100-112.

Valdes, E. V., and Summers, J. D. (1986). Determination of Crude Protein and Fat in Carcass and Breast Muscle Samples of Poultry by Near Infrared Reflectance Spectroscopy. Poultry Science vol 65, 485-490.

Wehling, R. L., and Pierce, M. M., (1988). Determination of moisture in Cheddar cheese by near infrared reflectance spectroscopy. J. Assoc. Off. Anal. Chem. vol 71, 571-574.

Wehling, R. L., and Pierce, M. M., (1994). Comparison of sample handling and data treatment methods for determining moisture and fat in Cheddar cheese by near-infrared spectroscopy. Agric. Food Chem vol 12, 2830-2835.

Wu, D,.He, Y., and Feng S (2008). Short-wave near-infrared spectroscopy analysis of major compounds in milk powder and wavelength assignment. Analytica chimica acta vol 610, 232-242.