Application of fluorescence spectroscopy and chemometrics to classify the spore level of nonfat dry milk and predict the process cheese emulsion properties

by

Chenhao Qian

B.S., University of California, Davis, 2018

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Food Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Approved by:

Major Professor
Jayendra K. Amamcharla

Copyright

Abstract

Spore contamination is one of the major quality concerns in nonfat dry milk (NDM). Some thermophilic and mesophilic spores can survive severe heat treatment, leading to problems in food products that uses NDM as an ingredient. Powder manufacturers need to monitor the spore level throughout production and cease processing in the case of intolerable contamination caused by biofilm attachment and development on the equipment. Current spore enumeration methods are either time-consuming, labor intensive, or expensive, and thus are often not practical in common production facilities. Fluorescence spectroscopy, due to its high specificity and accuracy, can target dipicolinic acid (DPA), which is an intrinsic fluorescent compound in spores, to predict the spore level in contaminated powders. In this study, a total of 40 NDM samples were procured from commercial sources. Traditional plating methods were performed to obtain the reference values of spore count. The results ranged from 1.7 to 5.0 log CFU/g NDM. To enhance the fluorescence signal, pre-treatment steps included autoclaving, acidification, and centrifugation to extract the available DPA from spores present in the reconstituted NDM at 10% concentration. Classification models were constructed using partial least square discriminant analysis (PLSDA), random forest (RF), and forward selection quadratic discriminant analysis (FS-QDA), and models were validated by bootstrapping techniques. The highest classification accuracy was observed for random forest at 87% success rate. The fluorescence-based classification models can provide a rapid tool for industry to quickly determine the quality of processed milk powders.

In another study, fluorescence spectroscopy was used to predict the emulsion characteristics of process cheese made by bench-top Thermomix^{TM}. Emulsion characteristics are the functional properties of process cheese in specific applications, and can be assessed through

texture, fat droplet size, viscoelasticity, and other attributes related to its performance in final products. These properties vary from product to product, and are crucial to maintain in order to ensure consistent quality. However, this is difficult for process cheese manufacturers as they are dealt with varying natural cheese depending on the availability. Fluorescence spectroscopy and near infrared, as rapid and noninvasive techniques, are suitable for on-line measurement of the key emulsion attributes. Essential properties, such as texture profile, rheological properties, and particle size were analyzed by textural profile analyzer (TPA), dynamic stress rheometer (DSR), and dynamic light scattering analyzer, respectively, on 40 process cheese samples produced at the K-State Dairy Products Lab. These samples consisted of 5 batches of 8 combinations of two cheese ages, two mixing speeds, and two holding times. ANOVA results showed significant differences in functional properties existed between process cheese samples. Principle component analysis (PCA) projected the samples into two-dimensions to identify the important variation among samples and variables. Quantitative models based on partial least square regression (PLSR) were developed using tryptophan fluorescence emission spectra, vitamin A excitation spectra, and NIR short-wave region. The calibration models were validated by the leave-one-out method, and yielded the highest correlation coefficient of 0.73 between pre-processed NIR spectra and hardness, followed by particle size. Fluorescence showed limited success in predicting other emulsion attributes. The study showed that correlation existed between fluorescence and NIR spectra, and key emulsion characteristics in process cheese, yet the model was not accurate enough for implementation in the industry. Future study may utilize pilot-scale production unit and include more samples to improve the model performance.

Table of Contents

List of Figures

List of Tables

Acknowledgements

To my major advisor, Dr. Jayendra Amamcharla, I am sincerely grateful for your mentorship, patience, and trust. Thank you for introducing me to research on chemometrics and helping me develop the skills for my future goals. I wouldn't be here continuing my career in research if you did not consolidate my conviction at the first place. To my committee members Dr. Randall Phebus and Dr. Karen Schmidt, thank you for your insight and time. I would like to thank Dr. Tonya Schoenfuss and Lisa Chou from Department of Food Science at University of Minnesota for collaboration on the process cheese project. Your wisdom and help are appreciated. I would also like to thank my lab mates, Boheeja Zaitoun, Karthik Babu, Priya Thorrakkatu, Gunvant Rathod, for company and scientific discussion. I would like to extend my gratitude to Ben Ma, for providing the insight on research and offering help on coding. I appreciate Sam for patiently making 40 process cheese samples for me. I am thankful to Daniel Vega and Katia Pozuelo Bonilla for assistance on the microbiological methods, and to Bade Tonyali for leadership in college bowl team. I am sincerely grateful for my family who always support me unconditionally. Lastly, I would like to thank National Dairy Council for their financial support.

Chapter 1 - Introduction

In past decades, fluorescence spectroscopy has gained popularity as a potential tool for rapid and noninvasive measurement in the food industry. Due to its advantages, practical applications have been observed in food production facilities to implement this technology into quality assurance programs. It can provide real-time spectral data that can be analyzed by chemometrics, which utilizes mathematical and statistical transformation to obtain meaningful results from complex data. Milk contains intrinsic fluorophores, such as tryptophan and vitamin A. that can show fluorescence signals upon excitation by light at appropriate wavelengths. The dairy industry particularly can benefit from the fluorescence-based non-contact and non-destructive analytical technology. Applications of fluorescence spectroscopy have been reported in quantification of components, to monitor structural changes, and to authenticate products in the field of dairy science (Karoui et al., 2005; SádeCká and TóThoVá, 2007).

Nonfat dry milk (NDM) is a dairy-based ingredient used in many shelf-stable food applications, such as process cheese, soups, and UHT beverages. It can provide functionality for the end products including improved emulsion, boosted protein content, and enhanced foamability. Thus, any quality defects in NDM can potentially be carried over to the finished product. As one of the major issues associated with NDM, high spore counts can compromise the product's quality by introducing undesirable enzymes as well as forming ropiness and lactic acid (Chen et al., 2004; Heyndrickx and Scheldeman, 2008). These thermophilic and mesophilic spores can enter the milk system through the natural environment, feed, and milking facility. In addition, the thermophilic and mesophilic spores become concentrated as water is removed from skim milk during subsequent processing such as evaporation and spray drying. Further increase in spore counts during processing can be attributed to the development of biofilms on the

stainless-steel processing surfaces. Currently, no existing methods can accurately enumerate the spores in a timely manner to maintain the spore count under control. To address this issue, fluorescence spectroscopy can target an intrinsic fluorescent component, dipicolinic acid (DPA), present in spores. However, DPA is not readily available as a fluorescent compound of interest, so additional extraction steps are required to enhance the fluorescence signal before measurement.

Apart from NDM, process cheese is also a leading dairy product in U.S. supermarket sales (IDFA, 2006). The availability of process cheese in diverse forms make it a versatile ingredient in food preparation. End-use applications depend on the functional properties of process cheese including unmelt texture, melting point, fat droplet size, and microstructure. The functional properties of process cheese are influenced by multiple factors including raw ingredients (cheese maturity, intact casein, type and quantity of emulsifying salt) and manufacturing process parameters. However, it is a standard practice in the dairy industry to use re-work cheese in the production of process cheese. Thus, the recipe for process cheese can vary from batch to batch depending on the availability of natural cheeses. In order to accurately assess the properties of process cheese during manufacture, a rapid in-line or on-line method is necessary. Fluorescence spectroscopy combined with chemometrics could potentially provide the solution to meet this industry need.

Th objectives of this thesis research were to use fluorescence spectroscopy to a) calibrate and validate models for classification of NDM based on spore count, and b) predict emulsion characteristics of process cheese during the manufacturing. These techniques could potentially reduce the measurement time and labor for the dairy industry to maintain consistent quality products.

References

Chen, L., Coolbear, T., and Daniel, R. M. 2004. Characteristics of proteinases and lipases produced by seven *Bacillus* sp. isolated from milk powder production lines. Int Dairy J, 14(6), 495-504.

Heyndrickx, M., and Scheldeman, P. 2008. *Bacilli* associated with spoilage in dairy products and other food. Applications and systematics of *Bacillus* and relatives, page 64-82.

IDFA. 2006. Dairy facts. Washington, D.C.: Intl. Dairy Foods Assn.

Karoui, R., Bosset, J. O., Mazerolles, G., Kulmyrzaev, A., and Dufour, É. 2005. Monitoring the geographic origin of both experimental French Jura hard cheeses and Swiss Gruyère and L'Etivaz PDO cheeses using mid-infrared and fluorescence spectroscopies: a preliminary investigation. Int Dairy J, 15(3), 275-286.

SádeCká, J., and TóThoVá, J. 2007. Fluorescence spectroscopy and chemometrics in the food classification-a review. Czech J Food Scis, 25(4), 159-173.

Chapter 2 - Literature Review

Fluorescence Spectroscopy

Traditional methods to assess the chemical and physical properties of food products can be time-consuming, labor-intensive and destructive. As a result, a rapid detection method is highly needed in the food industry to improve the efficiency and accuracy of analysis. It can help to ensure product consistency and therefore improve consumer acceptability. Real-time measurements can also provide timely feedback to the manufacturer as a part of quality assurance system. Recent technological advancements allowed the implementation of in-line process analytical technology (PAT) using fluorescence spectroscopy (Panikuttira et al., 2020) and other portable devices (Henihan et al., 2019) to facilitate quality control during food manufacturing. Fluorescence spectroscopy and chemometrics have been widely applied to assess quality change in the food system (SádeCká and TóThoVá, 2007). It is advantageous over the traditional methods because it is rapid, sensitive, and non-destructive. The signature fluorescence spectrum of foods can be obtained in a few seconds and used to both qualitatively and quantitatively investigate characteristics and properties of the food product. Among other purposes, it is particularly useful to authenticate products from specific geographical origins (Karoui et al., 2005) as well as track the structural change in the food products (Karoui and Dufour, 2006; Karoui, Mazerolles, and Dufour, 2003). Fluorescence spectroscopy is more sensitive and specific than other spectroscopic methods because it targets only specific fluorophores (Karoui and Blecker, 2011).

Principles and Instrumentation

A fluorophore, by definition, is a molecule that can emit light upon absorption of energy in the form of light at a lower wavelength. The principle of fluorescence can be explained by a Jablonski diagram in Figure 2.1. The molecule first absorbs the light that excites the electron from a ground state S_0 to an excited state S_1 or S_2. Through internal conversion, the molecule is generally rapidly relaxed into a lower energy state S_1 without the emission of light. Usually after 10^{-8} s, the electron returns to its ground state (Lakowicz, 2013). The energy difference is emitted in the form of light. Some molecules, such as riboflavin, due to their polycyclic structures, can have multiple vibrational levels at the excited state, and therefore can have several excitation wavelength peaks. Due to the energy loss between excitation and emission, the emission wavelength is always lower than the excitation wavelength. The difference of these two wavelengths is called Stokes shift (Valeur and Brochon, 2012). By controlling two wavelength parameters, fluorescence spectroscopy is more specific than other spectroscopic methods that only record the absorption spectra. To characterize a fluorophore, it is common to collect either emission or excitation spectra while fixing the other one at a maximum wavelength. However, when dealing with the unidentified analyte with unknown maximum excitation or emission, excitation-emission matrix (EEM) is more favorable. It is achieved by performing continuous scans over a range of excitation wavelengths to collect multiple emission wavelengths. The result is a 3D representation of fluorophores contained in a mixture (Bahram et al., 2006).

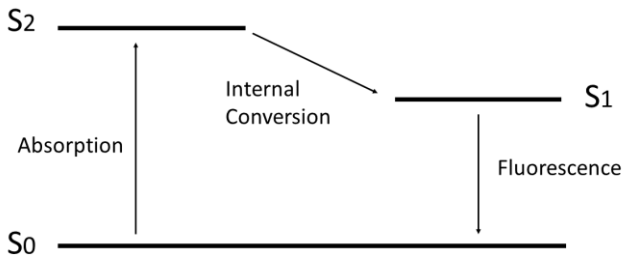

Figure 2.1. Jablonski diagram (adapted from Lakowicz, 2013)

A fluorescence spectrometer is the common tool to collect fluorescence spectral information. A simplified setup is displayed in Figure 2.2. The monochromator narrows the light from the lamp into certain wavelengths, and the in-coming light excites the fluorophore which emits the light at higher wavelength screened by another monochromator. The emitted light is then collected at photodetector and expressed as spectral data in computer software. In the classic device set-up, the emission light is collected at the right angle of the absorption light. The sample is usually diluted before measurement to avoid spectrum distortion that occurs if absorbance is greater than 0.1 (Karoui and Becker, 2011). However, diluting samples is not feasible for solid samples, and can potentially lower the concentration of analyte below the detection limits. The right-angle geometry also suffers from the scattering effect when applied to turbid samples. Since food matrices are complex and can contain multiple interference substances, front-face fluorescence spectroscopy (FFFS) is commonly used as an alternative to overcome the limitations. In this case, the penetration of light into the matrix is no longer a factor for the emission signal. However, it also means more reflected light from the surface of the cuvette will be detected, minimizing the sensitivity of the measurement (Lakowicz, 2013).

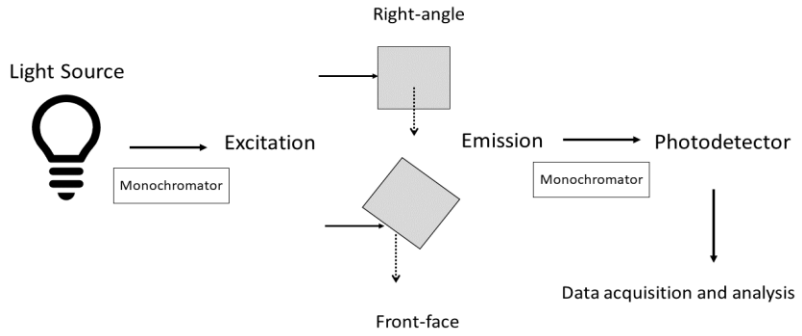

Figure 2.2. Simplified set-up of a fluorescence spectrometer (adapted from Karoui and Becker, 2011)

Applications in Dairy Products

Dairy products can contain several intrinsic fluorophores such as aromatic amino acids, nucleic acid, riboflavin, and vitamin A or storage-derived compounds such as oxidative and Maillard products. These molecules, even at low concentration, can produce fluorescence signals that help identify the compositional differences or processing changing in dairy products. Table 2.1 summarizes published research where fluorescence spectroscopy has been used to characterize fluorophores in dairy products. The applications are discrimination and identification of products and processing parameters (Birlouez-Aragon, Sabat, and Gouti, 2002; Granger et al., 2006; Karoui et al., 2007; Karoui, Martin, and Dufour, 2005), quantification of composition (Karoui et al., 2006; Becker et al., 2003), analysis of molecular structure (Dufour et al., 2001), monitoring change during processing and storage (Dufour and Riaublanc, 1997; Liu and Metzger, 2007; Wold et al., 2006), as well as predicting the functionality (Purna, Prow, and Metzger, 2005; Kulmyrzaev et al., 2005; Karoui and Dufour, 2003).

Table 2.1. Fluorophores of interest in dairy products

Fluorophores	Dairy products	Purpose	Spectral parameters	Reference
Tryptophan residues	Ice cream	Formulation discrimination	Ex 290 nm/Em 300-360 nm	(Granger et al., 2006)
	Cheese	Analysis of molecular interaction change	Ex 290 nm/Em 305-400 nm	(Dufour et al., 2001)
	Cheese	Identification of manufacturing process and sampling zones	Ex 290 nm/Em 305-450 nm	(Karoui et al., 2007)
	Milk	Determination of geographical origin	Ex 290 nm/Em 305-400 nm	(Karoui, Martin, and Dufour, 2005)
	Milk	Discriminating heat treatment	Ex 290 nm/Em 340 nm	(Birlouez-Aragon, Sabat, and Gouti, 2002)
	Process cheese	Predicting functionality	Ex 290 nm/Em 305-400 nm	(Purna, Prow, and Metzger, 2005)
Riboflavin	Cheese	Analyze light-induced change	Ex 380 nm/Em 400-750 nm	(Wold et al., 2006)

	Yogurt	Rapid determination of riboflavin content	Ex 270-550 nm/Em 310-590 nm	(Becker et al., 2003)
	Cheese	Discrimination of rheological properties	Ex 250-350 nm/Em 410 nm	(Kulmyrzaev et al., 2005; Karoui and Dufour, 2003)
	Cheese	Predicting chemical parameters	Ex 270-350 nm/Em 410 nm	(Karoui et al., 2006)
Vitamin A				
	Milk	Discrimination of homogenization and heat treatment	Ex 260-350 nm/Em 410 nm	(Dufour and Riaublanc, 1997)
	Milk	Estimation of heat treatment	Ex 350 nm/Em 440 nm	(Birlouez-Aragon et al., 1998; Brilouez-Aragon, Sabat, andGouti, 2002)
Maillard-reaction products	Nonfat dry milk	Monitoring storage change	Ex 360 nm/Em 380-480 nm	(Liu and Metzger, 2007)

Ex = excitation wavelength; Em = emission wavelength

One of the major components in milk and milk-derived products is protein, which contains aromatic amino acid residues that can fluoresce under UV light. Among three aromatic amino acids, tryptophan residue is extensively studied to differentiate dairy products as well as quantify their compositions. It is excited at 290 nm and has a maximum emission wavelength at about 345 nm (SádeCká and TóThoVá, 2007). Due to the sensitivity to local environment, tryptophan emission spectra can be used to reflect the average environment of these protein residues, and therefore provide information about the microstructure as well as molecular interactions (Lakowicz, 2013). Since microstructures are very specific to products and subject to change upon processing, studies have used tryptophan emission spectra to discriminate the products and processing conditions. Karoui et al. (2007) applied principle component analysis (PCA) and common components and specific weights analysis (CCSWA) to the spectra to identify three varieties of soft cheese as well as the sampling location within the block. A similar approach has been utilized on the milk to differentiate its geographical origins (Karoui, Martin, and Dufour, 2005). Factorial discrimination analysis was applied and the correct classification for validation set was 69.2% for tryptophan spectra, showing that compositional differences were retained in spectral information. Another use of the tryptophan residue is to predict functionality (Purna, Prow, and Metzger, 2005). The melting point of process cheese was measured using a dynamic stress rheometer (DSR) and a prediction model based on tryptophan emission spectra was developed using partial least square regression (PLSR) to predict the melting temperature with a correlation coefficient of 0.93. Higher meltability was observed with lower peak height in 335-350 nm region. It indicated that the emission peak and region of tryptophan residue can change according to the polarity of environment, and thus the emulsion.

Another fluorophore that has been associated with emulsions in dairy products is vitamin A, as its excitation spectra provides the information on the protein-lipid interaction. Vitamin A is a carotenoid naturally present in milk, and has excitation and emission maxima at 325 nm and 470 nm, respectively, when measured in pure solution (Duggan et al., 1957). As the milk is homogenized, the fat globule size is reduced and thus the local environment is more aqueous, inducing a change in signal pattern. One study found that applying principle component analysis (PCA) to normalized vitamin A emission spectra can discriminate the milk that undergoes heat treatment and homogenization (Dufour and Riaublanc, 1997). It demonstrated the promise to use vitamin A spectra for further investigating the emulsion system in other dairy products. Vitamin A is fat soluble and it can also be used to quantify the fat-related components. Using front-face fluorescence spectroscopy with partial least square regression (PLSR), Karoui et al. (2006) were able to predict the fat, dry matter, fat in dry matter, and water-soluble nitrogen (WSN) in soft cheese with high values of determination coefficient of 0.88, 0.86, 0.86 and 0.84 respectively. Apart from rapid determination of chemical composition, vitamin A spectra was studied for its correlation with rheological properties of soft cheese (Kulmyrzaev et al., 2005; Karoui and Dufour, 2003). One study characterized soft cheese by both dynamic oscillation and vitamin A excitation spectra (250-350 nm). The storage modulus G' was found to be highly correlated with vitamin A fluorescence spectra by the canonical correlation analysis (CCA) with all 4 squared canonical coefficients higher than 0.98, suggesting a solid representation of rheological properties by fluorescence spectra (Karoui and Dufour, 2003). Another study analyzed soft cheese by several infrared regions, tryptophan emission spectra, vitamin A excitation spectra, and uniaxial compression data. By using common components and specific weights analysis (CCSWA), the authors were able to find commonality of different data matrices. The results

showed that the common component 1 mainly explained infrared data while the common component 2 explained most of the vitamin A spectra. However, the rheology data showed good correlation with common component 1 instead of 2, suggesting that the ability of vitamin A excitation spectra to predict the rheological properties was limited.

Some fluorophores in dairy products are developed during storage and processing. One of them is riboflavin, which degrades upon exposure to light into two products: lumichrome and lumiflavin. The degradation is signified by the reduction of fluorescent signal at 525 nm (SádeCká and TóThoVá, 2007). Thus, riboflavin spectra are commonly used to investigate light-induced changes in dairy products, such as discoloration, formation of off-flavor and nutritional loss (Becker et al., 2003). Wold et al. (2006) stored two varieties of cheese up to 48 hours and compared the fluorescence data of riboflavin emission spectra with the odor intensity of several oxidized notes obtained from sensory panels. They developed a model using PLSR and good correlations were found between the spectra and oxidized odor, sun odor, and acidic odor, which developed upon photodegradation. However, not all photosensitizers can be identified from the spectra data. A similar approach was taken by Becker et al. (2003) to quantify the riboflavin content in plain yogurt. The regression model showed a high correlation coefficient ($R = 0.99$) and low prediction error (0.092 μg riboflavin/g).

Other processing and storage induced fluorophores are Maillard products. Since milk contains protein and lactose, these Maillard compounds can be formed during heating processes. Intermediate reaction derivatives including pyrrole and imidazole have cyclic structures that emit fluorescent light. Their formation can be used to indicate the heat treatment and storage of dairy products by FAST (fluorescence of advanced Maillard products and soluble tryptophan) index proposed by Brilouez-Aragon, Sabat, and Gouti (2002). It is calculated as the fluorescence of

advanced Maillard products divided by soluble tryptophan in a clear solution at pH 4.6 (Andersen and Mortensen, 2008). Since the Maillard products are associated with flavor and functionality changes in the products, FFFS has been investigated for the potential to monitor the Maillard reaction occurring in the nonfat dry milk during storage (Liu and Metzger, 2007).

FFFS Applications

This thesis covers two novel applications of FFFS on dairy products. In the first study of this thesis, the FFFS is used to provide a sensitive and rapid classification of dairy powders based on the spore count. In the second study, FFFS is utilized to monitor the emulsion stability of process cheese food (PCF). The literature reviews on the non-fat dry milk and process cheese food are discussed in detail below.

Nonfat Dry Milk Overview

According to the CFR, nonfat dry milk is "the product obtained by removal of water only from pasteurized skim milk. It contains not more than 5 percent by weight of moisture, and not more than 1 1/2 percent by weight of milkfat unless otherwise indicated" (21 CFR131.125). Although the minimum levels of other constituents are not set, a typical nonfat dry milk is composed of 34-37% protein, 49-52% lactose, 0.6-1.25% fat, 8.2-8.6% ash, 3-4% moisture (USDEC, 2018). The term nonfat dry milk is interchangeably used with skim milk powder (SMP), but the latter has a minimum requirement for protein content to be above 34%. Furthermore, the SMP is defined by CODEX Alimentarius, while the NDM is regulated by the United States Department of Agriculture (USDA). The USDA definition conforms to the same standards from the CFR and grades the spray-dried nonfat dry milk into two categories, which are U.S. Extra and U.S. Standard. The quality factors for the grading include flavor, physical

appearance, bacterial count, milkfat content, moisture content, scorched particle content, solubility index, titratable acidity (USDA, 2011). As the highest-grade quality, Extra nonfat dry milk is completely free of lumps, and the reconstituted skim milk could have flavor notes described as chalky, cooked, or flat while that of standard grade may contain mildly bitter, oxidized, or scorched flavor (Stathopoulos, 2018).

The Process of Manufacturing Nonfat Dry Milk

The process of manufacturing nonfat dry milk is well-explained in the book chapter Manufacture and Properties of Dairy Powders by Kelly and Fox (2016). The unit operations include collection of skim milk, pre-heating, concentration, and drying. Pre-heating occurs after fat is separated out from the whole milk and is followed by concentration of pre-heated skim milk to 42-48% total solids and then drying the concentrate into powder form. In some cases, instantization and agglomeration are involved in improving the solubility of the powder. Of all the heat treatment during the manufacture, pre-heating is the highest heat treatment for the powder, during which most of the whey protein denatures (Singh and Creamer, 1991; Oldfield, Taylor, and Singh, 2005). The severity of heat applied both influence both the protein functionality and microbiological quality of the final product (Kelly and Fox, 2016). Applied heating is categorized into three levels: low heat, medium heat, and high heat, as shown in Table 2.2. The heating is applied either directly by steam injection or indirectly by the plate heat exchanger. However, the fouling in the indirect heat exchanger can accumulate biofilm, which could contaminate the next batch of samples. The concentration of skim milk before spray drying is usually achieved through a multiple-stage falling film evaporator. The water is vaporized at 65 – 70 °C as the skim milk passes through the steam-heated tube under vacuum (Bylund, 2003). In

some cases, each evaporator performs at a lower pressure than the previous one so that the vapor can be recycled as the heating medium for the following stage. The economic cost of energy could be further minimized by the use of mechanical and thermal vapor recompression. The last step of manufacturing is the drying process. The majority of milk powder today is produced by spray-drying, in which the concentrate is sprayed into a chamber where droplets contact with hot air and instantly evaporate the water and form powder particles. Although the processing diagram is relatively standard, the specific manufacturing regimes can still significantly influence the quality of powder.

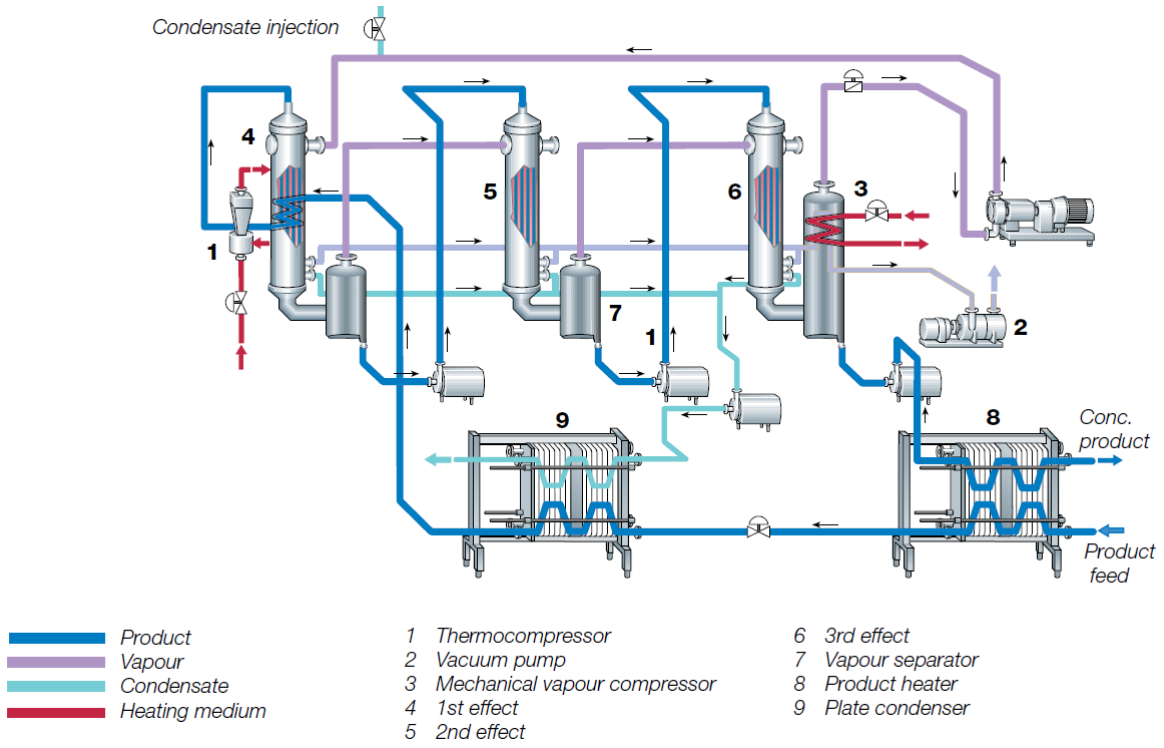

Figure 2.3. Three-effect evaporators (Bylund, 2003)

Powder Quality

The milk powder properties of interest to manufacturers can be physical or microbiological. Physical properties of powder include bulk density, reconstitution properties, heat classification, flowability, and heat stability (Kelly and Fox, 2016). Microbiological properties include standard plate count (SPC), coliform bacteria count, and spore count. The unit operations during manufacture have been shown to influence the physical quality of the final product.

Park and Drake (2016) compared vacuum evaporation and reverse osmosis in relation to flavor in powder and determined that vacuum evaporation produced more off-flavors due to associated heat treatment and reduced characteristic milk flavor by the vapor separation. Moreover, drying techniques can severely affect the powder's microstructure, which is highly correlated with desirable properties of powder such as bulk density, free-flowing, and solubility (Kelly, O'Connell and Fox, 2003). Compared to roller-dried milk powder, spray-dried milk powder typically is characterized by smooth particles with more regular shape. Heat classification is another important property of powder as the major application of nonfat dry milk as food ingredient is to provide various functionalities in processed food (Sharma, Jana, and Chavan, 2012). Depending on the end-use applications, functionality can be altered by changing the heat load during pre-heating. The level of whey protein denaturation induced by heat can influence the physical properties of milk powder. More severe heat treatment can lead to enhanced foaming capacity, poor solubility, and higher heat stability. To classify the heat treatment, whey protein nitrogen index developed by the American Dairy Products Institute is widely used before the drying process (ADPI, 2016). Other alternatives have been proposed to minimize the influence of protein concentration on the true value of whey protein nitrogen index

(WPNI), including Fourier transform near infra-red (FT-NIR) and SDS-PAGE that targets the disulfide crosslink formation in protein aggregates during heat treatment (Patel et al., 2007). The heat classification and associated functionality and application are summarized in Table 2.2.

Table 2.2. Heat classification, functionality, and application of NDM (Kelly and Fox, 2016)

Heat classification	Exemplary heat treatments	WPNI (mg/mL)	Functionalities	Applications
Low heat	70 °C for 15 s	> 6.0	Solubility, lack of cooked flavor	Recombined milk, milk standardization, cheese making
Medium heat	85 °C for 1 min 90 °C for 30 s 105 °C for 30 s	1.5 – 6.0	Emulsification, foaming, water absorption, viscosity, color, flavor	Ice cream, chocolate, confectionery
High heat	90 °C for 5 min 120 °C for 1 min 135 °C for 30 s	< 1.5	Heat stability, gelation, water absorption	Recombined evaporated milk

Microbiological contamination is another major defect in milk powder, and its occurrence can lead to severe public safety issues. The milk needs to have high microbiological quality before being processed into powders. Moreover, bactofugation or microfiltration is commonly implemented before pre-heating to ensure significantly low bacteria presence. Pre-heating is also critical to sterilize the product (Kelly and Fox, 2016). However, unlike vegetative cells, thermophilic bacteria spores can survive the heat treatment and be activated by sub-lethal heating, and therefore compromise the quality and safety of the product.

Spore in Dried Dairy Ingredient

Bacteria can contaminate the milk and milk-derived products via many pathways. Bacteria genera such as *Staphylococcus, Streptococcus, Bacillus, Micrococcus,* and *Corynebacterium* and sometimes coliforms, usually reside on the teat skin of cows and can get into the milk during excretion (Angulo, LeJeune, and Rajaha-Schultz, 2009). In addition, presence of *Campylobacter jejuni* (Humphrey and Beckett, 1987) and *Brucella* (Hamdy and Amin, 2002) have been tested positive in raw milk. Moreover, microorganisms from the environment and facilities can accumulate in the milk during collection, transportation, and storage. These include pathogenic bacteria from dairy farms, such as *Salmonella* species, *L. monocytogenes,* and Shiga toxin-producing *E.coli* (Oliver, Jayarao, and Almeida, 2005). However, most of these pathogens are effectively killed during the pre-heating process of dairy powder. Thus, due to the specific scope of this review, only spore-formers that can survive severe heat treatment and thus contaminate the dairy powder are discussed.

The notorious spore-formers mainly found in milk powder are thermophilic bacilli that can survive rigorous heat treatment. From a survey that examined 28 powders from 18 countries, *Anoxybacillus flavithermus, Bacillus licheniformis*, *Geobacillus stearothermophilus,* and *Bacillus subtilis* were found to have the highest occurrence in decreasing order (Rückert, Ronimus, and Morgan, 2004). These bacteria can be further divided into two groups: the obligate thermophiles and the facultative thermophiles. *Anoxybacillus flavithermus* and *Geobacillus* spp. belong to the obligate thermophiles and they grow only at relatively high temperatures between 40-68 °C (Ronimus et al.,2003; Scott et al., 2007). *Bacillus* genus are mostly facultative thermophiles that can tolerate both mesophilic and thermophilic conditions (Crielley et al., 1994; Ronimus et al., 2003; Scheldeman et al., 2006).

Although bacterial spores are initially present in milk, they are considerably more concentrated in milk powder after the multiple stage evaporation. These spore-formers that contaminate the milk and originate from the cow, milking equipment, and natural environment (Coorevits et al., 2008). Milk powder is one of many milk products that have issues with thermophile growth during processing. By attaching to the processing equipment as biofilms, they can cause secondary contamination (Flint, Bremer, and Brooks, 1997). Biofilms development differs depending on the stages of processing. Flint et al. (2001) reported that on the plate heat exchanger surfaces that are regularly cleaned and undergo high shear, thermophilic bacilli form a monolayer, whereas underneath the distribution plates of a plate heat exchanger multi-layer extracellular polymeric substances (EPS) matrix is more likely to form due to inconsistent flow. Many researchers concluded that processing units operate at high temperatures (40-65 °C) are more susceptible to biofilm formation (Flint et al., 1997; Murphy et al., 1999; Scott et al., 2007). Attachment of both vegetative cells and spores is believed to initiate biofilm formation (Parkar et al., 2001). Once attached, the spores can germinate, and biofilms grow through the reproduction of vegetative cells (Burgess et al., 2009). Spore development in biofilms is less studied. One study suggested that the formation of biofilms and spores occurred at the same time for *A. flavithermus* (Burgess et al., 2009) while the processes were separate for *Bacillus subtilis* in another study (Lindsay, Brözel, and Von Holy, 2005).

Although the presence of some *B. cereus* strains in dairy products was shown to produce toxins, the quantity is not appreciable until vegetative cell growth reaches10^7 cfu/ml (Griffiths, 1990), which is unlikely for most UHT beverages. Thus, food safety is less of a concern than spoilage (Cosentino et al., 1997). Instead, the compromise of product quality is the major issue for utilizing the contaminated powder in further processing. Due to mineralization and

dehydration, these spore-formers are highly protected by a spore coat, and thus have strong heat resistance (Beaman et al., 1982). Many strains were reported to be able to survive the UHT treatment (Schwarzenbach and Hill, 1999; Mostert, Luck, and Husmann,1979). Spoilage of UHT and canned milk products has been associated with the *Bacillus* genus that produces ropiness as well as lactic acid (Heyndrickx and Scheldeman, 2008; Gilmour and Rowe, 1990). Spores could also compromise the product by inducing lipase and protease activity (Chen, Coolbear, and Daniel, 2004; Chopra and Mathur, 1984; Cosentino et al., 1997). Thus, the presence of thermophilic spores provides the potential to germinate and spoil the final product if the contaminated milk powder is used as an ingredient when the processing or storage conditions are favorable. For this reason, low spore NDM is desired in the market to ensure consistent quality of final products.

To minimize the spore count in final powders, different methods for controlling contamination have been suggested. Practices such as shorter production cycles, the use of disinfectants, altering temperatures, decreasing surface area at susceptible sections, as well as the application of dual equipment have been proposed to effectively reduce biofilm development during the manufacturing (Burgess, Lindsay, and Flint, 2010). As good hygiene can significantly reduce the biofilm formation, the spore count could be used to determine the hygiene quality of the dairy processing facility. If the production cycle is too long or the cleaning process is not thorough enough, high spore accumulation as biofilm can occur and negatively affect the powder quality. A quantification method is needed to determine the extent of sporulation during processing, and therefore cease the operation for sanitation when the spore number is exceedingly high.

Spore Detection Techniques

Plate count is traditionally used for spore enumeration. Two most common methods are a total thermophile plate count (TPC) and a thermophilic spore count (TSC). Both methods use milk plate count agar (MPCA) and incubate the plate at 55°C for 48 hours. The difference is that TSC requires pre-heating the sample at 100°C for 30 mins to kill the vegetative cells while activate the spores, and 0.2% starch is added to the medium to aid germination (Burgess et al., 2010). Other time and temperature are reported in published study to select for more heat resistant spores (Coorevits et al., 2008; McGuiggan et al., 2002; Murphy et al., 1999; Rueckert et al., 2004). An ISO (2009) method is also developed to enumerate the spores in dairy powders as ingredients for UHT-treated products. It requires heating the sample at 106°C for 30 mins, and subsequently following the same procedure as TSC. However, heating above 100°C is not feasible in most labs and is hardly reproducible. To address this difficulty, a new method was proposed and shown to have the same predicating power. It is involved with heating the reconstituted skim milk at 100°C at 30 mins, and subsequently incubate the sample in tryptic soy agar (TSA) at 55°C for 48 hours (Eijlander et al., 2019).

Other rapid alternatives have been developed to reduce the labor cost as well as obtain more control over the manufacturing. Flow cytometry and real-time polymerase chain reaction (PCR) are two mostly implemented methods. The BactiFlow^{TM} flow cytometer measures the esterase activity in milk powders, and the method was proposed (Flint et al., 2006) and updated later to select for only mesophiles (Flint et al., 2007). The result was highly correlated with a TPC. However, the limitations exist as the technical difficulty to operate the cytometer as well as the high detection limit for high quality powders ($< 10^3$ CFU/g). A real-time PCR assay targeting the *spo0A* gene was developed by Rueckert, Ronimus, and Morgan (2006) to enumerate the

thermophilic bacilli in milk powders. The method can successfully quantify *B. licheniformis* and *A. flavithermus* present in the powder to the level of 6400 spores/g in less than an hour. However, these two methods both require technical expertise and high expense. Thus, spectrometric methods that are both rapid and easy to operate have been extensively studied to quantify the spores, although not in the context of dairy processing.

DPA-based assays

These methods are based on quantification of the biomarker dipicolinic acid (DPA) that are responsible for the heat resistance of spores. The use of Fourier-transform infrared spectroscopy (FTIR) in combination with Curie-point pyrolysis mass spectrometry (PyMS) was able to distinguish spores and vegetative cells in a variety of species with pyridine ring vibration at 1447-1439 cm^{-1} (Goodacre et al., 2000). The DPA also produce a strong fluorescence signal when bind to lanthanide ions, and thus fluorescence has been widely reviewed for detection of spores (Hindle and Hall, 1999; Pellegrino et al., 1998; Rosen, Sharpless, and McGown, 1997). However, all of these studies used pure spore suspension and the release of DPA from spores was not optimized. Thus, Pellegrino, Fell Jr, and Gillespie (2002) compared different extraction techniques and concluded the detection capacity could be enhanced by two-order of magnitude and the limit could be lowered to 1000 CFU/ml when utilizing heated dodecylamine (dda) as the extraction method. This showed promising for potential application in dairy powders. However, the studies mentioned above only used the maximum signal at emission peak as the prediction value for DPA or spore count. To include more features from the spectra and enhance the robustness of the prediction model, the use of the entire emission spectra in combination with chemometrics is preferred. Thus, this review proposes a method that classifies the spore count based on the emission spectra of DPA in pre-processed non-fat dry milk.

Nonfat Dry Milk as an Ingredient in Process Cheese

Although nonfat dry milk can improve the functionality of food products, it is mainly used in process cheese for increasing protein content while reducing the cost. The permitted level is not legally defined for process cheese (PC), process cheese food (PCF), and process cheese spread (PCS) but total lactose and whey protein content should be considered for the formula. The excessive amount of lactose content can lead to crystallization and undesirable Maillard browning (Berger et al, 1998; Kapoor and Metzger, 2008). The general rule for preventing the lactose crystallization is to maintain the lactose content below its solubility limit, which is 17% at 20 °C (Harper, 1992). Therefore, depending on the moisture content of specific process cheese product, lactose level as well as nonfat dry milk should be adjusted accordingly. The addition of skim milk powder has been shown to facilitate the Maillard reaction in process cheese stored at warm temperature due to the introduction of lactose (Thomas, 1969). The reaction can develop objectionable color and flavor, and thus lead to defects in the final product. Apart from lactose, whey protein content from nonfat dry milk could also induce changes in process cheese as whey can denature during the manufacture and associate within itself or with casein micelles via disulfide crosslinking between β-lactoglobulin and κ-casein (Wong et al, 1996). These interactions have a profound impact on the sensory as well as functional properties of process cheese. Many studies confirmed that the increase of whey protein concentration would result in higher firmness and lower meltability of process cheese (Thapa and Gupta, 1992; Mleko and Foegeding, 2000;Sołowiej, Cheung, and Li-Chan, 2014).

Process Cheese Components

Process cheese is a dairy product produced by blending and heating natural cheese, emulsifying agent, and other dairy or nondairy ingredients to a homogenous mixture (Kapoor and Metzger, 2008). It can be further defined into three categories, depending on the moisture, fat content, and pH of the final product (21CFR133.169). The allowable levels of natural cheese, emulsifying agent, and other optional ingredients are also determined by CFR and vary between three categories. Those are legally defined as pasteurized process cheese (PC), pasteurized process cheese food (PCF), and pasteurized process cheese spread (PSC).

Table 2.3. CFR definition for major characteristics of three process cheese categories (21CFR133.169)

Category	Moisture (% w/w)	Fat (% w/w)	pH
PC	≤40	≥30	≥5.3
PCF	≤44	≥23	≥5.0
PCS	44 to 60	≥20	≥4.0

The most principle component in any process cheese is natural cheese. It is commonly used solely or in combination with other natural cheese of different types and aging. Varying in cheese sources can lead to different intact casein content and total calcium content and thus can greatly impact the functional properties. The second essential ingredient in process cheese manufacture is emulsifying agent. It is usually a mono- or poly- citrate and phosphate salt with monovalent anions. The purpose of incorporating emulsifying salts is to both disrupt the protein network originally in the natural cheese and change the pH level in products. Hydrated by these agents, casein could interact with fat and water more easily, achieving a better emulsion stability (Meyer, 1973). Other dairy ingredients that are functional in process cheese production are butter

and non-fat dry milk (NDM). They can both improve the emulsion in process cheese and NDM is also a good protein source to boost the casein content. Non-dairy ingredients that could be used include salts, to enhance flavor and prevent spoilage, and food colloids to improve the functional properties (Simeone, Alfani, and Guido, 2004).

Process Cheese Manufacture

A standard flow chart for manufacturing the process cheese is shown in Figure 2.5 (Kapoor and Metzger, 2008). The selection of raw ingredients was discussed in the last section. After the choice of ingredients is made, the appropriate composition needs to be formulated depending on the product category as well as desired flavor and texture. Fat and moisture analysis are performed to ensure the product's composition meets the legal standard. The formulation for the product is not always fixed. Sometimes depending on the availability of natural cheese and introduction of re-work cheese, the recipe must be corrected regularly to achieve consistency (Tamime, 2011). Once the raw materials are prepared, natural cheese is first grinded, possibly with other dry ingredients using high-speed grinder to a homogenous mix before transferring to the cooker. During the cooking process, other ingredients that have not been mixed in the last step are incorporated, and then the pre-mix is heated while mixed. The minimum cooking temperature and time is legally defined (21CFR133.169) as 65.5°C for 30 seconds. Manufacturer adjust the cooking regime depending on the agitation parameters, heating methods, as well as production mode (Berger et al., 1998; Zehren and Nusbaum, 2000). These processing conditions can have a significant impact on the process cheese properties. Hot process cheese is then directly filled into packages and molded into blocks before it is cooled and stored. The rate and condition of cooling process also influence the product's properties. Slow cooling

can enhance the Maillard product's formation and promote the sporulation (Fox, 1993). Moreover, it can result in firmer and more adhesive textures (Piska and Štětina, 2003).

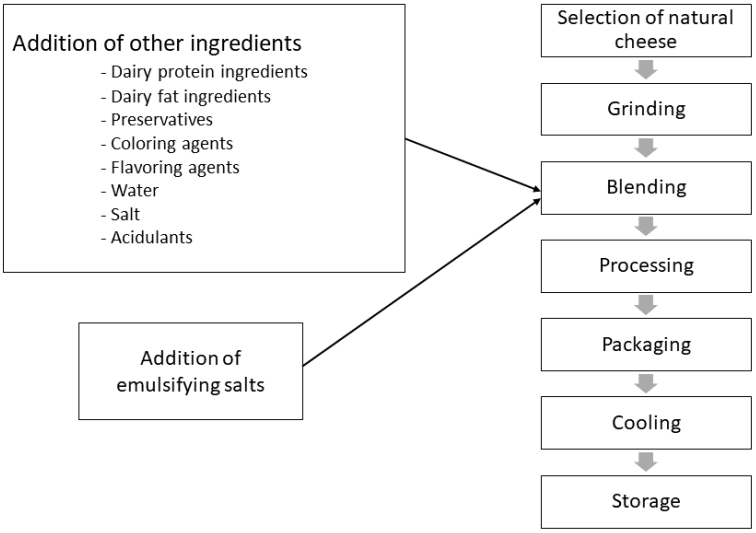

Figure 2.4. Flow diagram of process cheese manufacturing (adapted from Kapoor and Metzger, 2008)

Process Cheese Functional Properties

Changing the formulation, ingredients, and processing conditions can considerably influence the functional properties of the final product. Since process cheese is mainly used as an ingredient in other food products, such as pizza, sandwiches, cheesecakes, their performance to achieve the desirable attributes is important. A functional property, thus, is defined as its behavior in food preparation and consumption (Guinee, 2002). A group of functional properties are generalized into two types, which are unmelted texture and melted texture properties. Table 2.4 and 2.5 summarize the unmelted and melted properties. Due to various end-use applications of process cheese, their desirable attributes differ. For example, a cheese slice for pizzas should have enough firmness and cohesiveness to maintain the shape, limited stickiness to be sliceable,

as well as good meltability and stretchability after cooking. Given its importance to products, functional properties were studied in numerous researches. In their book, Gunasekaran and Ak (2002) viewed the cheese as a viscoelastic material that exhibit solid and liquid characteristics. Thus, one general approach to quantify the properties of process cheese is through rheological point of view by understanding its resistance to deformation and flow. For this reason, many measuring techniques and instruments have been proposed to characterize different properties.

Table 2.4. Unmelted properties, definition, and measurement techniques (Kapoor and Metzger, 2008)

Properties	Definitions	Measurement techniques
Firmness	Tendency to resist deformation upon external force	1. "Thumb print" test 2. "Sliceability" test 3. Textural profile analysis (TPA) 4. Penetrometry 5. Dynamic stress rheometry (DSR)
Brittleness/ fractureability	Tendency to fracture upon external force	1. Textural profile analysis (TPA)
Springiness/ resilience	Tendency to recover to original shape upon removal of external force	1. Textural profile analysis (TPA) 2. "Roll" test
Adhesiveness/ Stickiness	Tendency to separate from a contacting material	1. Textural profile analysis (TPA) 2. "Slice separation" test

Table 2.5. Melted properties, definition, and measurement techniques (Kapoor and Metzger, 2008)

Properties	Definitions	Measurement techniques
Meltability	Tendency to soften upon heating	1. Arnott test 2. Schreiber melt test 3. Dynamic stress rheometry (DSR) 4. Melt profile analysis 5. Rapid visco analyzer (RVA)
Viscosity	Tendency to spread and flow upon complete melting	1. Arnott test 2. Schreiber melt test 3. Tube melt test 4. Dynamic stress rheometry (DSR) 5. Melt profile analysis 6. Rapid visco analyzer (RVA)
Stretchability	Tendency to form strings when extended after heated	1. Pizza “fork” test

Techniques for Measuring Functional Properties

One of the most common techniques to measure the unmelted properties is texture profile analysis (TPA). By performing a uniaxial compression test, TPA instrument lower the probe to compress the cheese and then release in order to obtain the force and work of area needed for the deformation, and thus enable the characterization of multiple properties in a single measurement. (Breene, 1975; Peleg, 1976; Gunasekaran and Ak, 2003). These properties include hardness, brittleness, stickiness, and springiness. Another common tool to measure the viscoelasticity of process cheese is via dynamic stress rheometry (DSR). By applying a sweep test with fixed frequency and stress, the rheology instrument can determine the storage modulus (G’), which represents the solid-like property, and the loss modulus (G”), which represents the liquid-like

property. The values of G' and G" were found to be correlated with TPA hardness (Drake et al., 1999). In addition to measure unmelted properties, DSR is also capable of measuring melted properties. By ramping the temperature during the sweep test, DSR can also measure the melting temperature of process cheese as the temperature where tan δ (which is G"/G') equals 1 (Sutheerawattananonda and Bastian, 1998). The G" at 85°C was also proven to be a good indicator of meltability measured from Tube Melt Test (Prow and Metzger, 2005). Apart from DSR, many empirical techniques have been utilized to measure the meltability of cheese, including Arnott Test (Arnott, Morris, and Combs, 1957), Tube Melt Test (Olson and Price, 1958), Schreiber Melt Test (Kosikowski and Mistry, 1997). These tests aimed to measure the different perspective of change in cheese at certain shapes, and thus the results obtained could not be directly compared.

Chemometrics

Due to the complex composition of food matrix and multidimensionality of spectra data, useful information could not be directly extracted for interpretation without multivariate statistical analysis. This analysis technique, also known as chemometrics, can express the food properties as a function of fluorescence signals from excitation or emission spectra (Karoui and Becker, 2011). Prior to data analysis, the spectra data is commonly pre-processed in order to reduce the impact of noise, enhance features, and obtain a cleaner data structure. Data cleaning is then followed by dimension reduction and application of specific statistical models that are further calibrated and validated by the sample set. The use of chemometrics can be categorized into two approaches, which are supervised and unsupervised analysis. The unsupervised analysis is exploratory in nature, helping visualize the spatial distance between samples and identify the natural grouping. Since the unsupervised analysis does not require pre-existing knowledge about

the sample difference, it is usually performed first to obtain a general view of subset structure (SádeCká and TóThoVá, 2007). The purpose for the supervised analysis, however, is to discriminate the data set or quantify the intrinsic properties of samples using complex spectra data. It is involved with calibration and validation to enhance the accuracy and reproducibility of the model. A list of specific techniques is summarized in Figure 2.6.

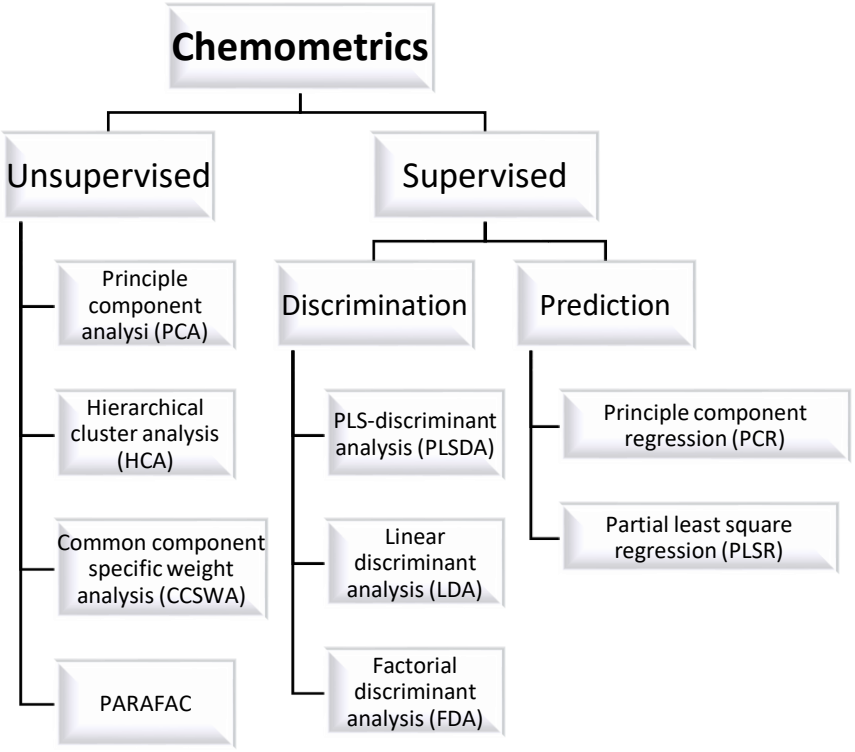

Figure 2.5. Diagram of chemometrics

Unsupervised Methods

The unsupervised methods are usually performed at the first step of data analysis to explore the similarity between samples or variables. The most widely used method is principle component analysis (PCA). It can reduce the highly dimensional data set into two to three principle components that explain the most variation between samples, and visualize the similarity between different sample groups in a 2-dimension map by plotting the loading score of

first and second principle components. Since the fluorescence spectra data can contain up to hundreds of variables, PCA is an effective way to select the most important variables while retaining most information. By correlating the principle components to the processing conditions or compositional difference, researchers have been applied the PCA to discriminate the dairy products with different formulation (Granger et al., 2006), rheological properties (Kulmyrzaev et al., 2005; Karoui and Dufour, 2003), heat treatment and homogenization (Dufour and Riaublanc, 1997), as well as geographical locations (Karoui, Martin and Dufour, 2005).

Another statistical analysis to visualize the natural grouping within data set is clustering analysis. Hierarchical clustering analysis (HCA) is one of the most common analysis applied to spectra data. It can calculate the distance between samples at multidimensional space and display the similarity in a dendrogram where samples that resemble are placed closer together (SádeCká and TóThoVá, 2007). The clustering is achieved by merging two similar observations or clusters into a new cluster until all observations are merged. Different linkage and distance criteria can be manipulated for the best results, granting the tool with good flexibility. Poulli, Mousdis, and Georgiou (2005) were able to classify the virgin olive oils based on the synchronous fluorescence spectra using HCA with 97.3% correct rate.

While PCA and HCA is suitable for differentiating one data set a time, sometimes multiple data sets are generated by several empirical techniques (chemical, physical, etc.) to characterize samples, calling for need to compare them in a unified analysis. This is when common component specific weight analysis (CCSWA) becomes useful. It can determine the common space of representation for all data sets, giving each of them different weight associated with each dimension in common space (Karoui et al., 2007). Thus, the specific weight can provide information on the correlation between the sample attributes or processing conditions

with the testing methods. The relationship between different data sets can also be concluded from the analysis. By applying CCSWA, Karoui et al. (2007) were able to interpret the spectroscopic information collected from riboflavin, tryptophan, and vitamin A in a systematic way that are not achievable by PCA.

Although the unsupervised methods discussed above can analyze two-dimension emission or excitation spectra, three-dimensional data such as excitation-emission matrix (EEM) could be collected and directly analyzed by parallel factor analysis (PARAFAC). The EEM is decomposed into a trilinear model that contains one score matrix and two loading matrices, which can be used to identify the fluorophore present in the sample (Munck et al., 1998). The classification of olive oil has been achieved by applying PARAFAC with Fisher's linear discriminant analysis (Guimet, Boqué, and Ferré, 2006). While PARAFAC is useful in identifying key fluorescence characteristics in a complex three-dimensional array, it is time-consuming to collect EEM and perform data analysis, making it less practical for rapid detection.

Supervised Methods

The unsupervised analysis enables the initial investigation of data set, which can be further improved in application by the model development via supervised methods. The methods can be divided into two categories, which are classification and regression. While the classification can discriminate observations into groups, the regression analysis can quantitatively predict a set of values such as chemical compositions or physical properties, from another set of data matrix.

Classification, as known as discriminant analysis, aims to maximize the spatial distance between groups in a hyper dimension. The most widely applied methods are partial least square

discriminant analysis (PLSDA), linear discriminant analysis (LDA), and factorial discriminant analysis (FDA). PLSDA and FDA are essentially the combination of dimension reduction method and discriminant analysis. While LDA directly projects the variable values on the hyperspace to minimize within group variance and maximize between group variance, PLSDA and FDA use the latent variables to predict the categorical data. The main difference between PLSDA and FDA is that, PLSDA finds the greatest variability by comparing the spectra data and reference value while FDA maximizes the total variance explained in spectra data and synthesize a new variable called "discriminant factors" from the selected PCs (SádeCká and TóThoVá, 2007). LDA assumes the number of samples is much larger than that of variables. Thus, for the highly dimensional fluorescence spectra, dimension reduction as a pre-processing tool is commonly needed, such as in PLSDA and FDA. By applying the first 10 PCs from FDA, Karoui and Dufour (2003) were able to classify three soft cheese variety from tryptophan and vitamin A spectra with correctness rate of 87.7% and 96% in validation set.

Sometimes when the quantification is required, classification might not be precise enough for the purpose of application. In such case, regression analysis can extend the use of spectra data. The most widely applied analysis is partial least square regression (PLSR). It finds the most relevant latent variables that correlate with the reference value to build a calibration model. The quality of the model is based on correlation coefficient (R), root mean square error of calibration (RMSEC) and root mean square error of prediction (RMSEP) (Karoui and Dufour, 2005). PLSR is suitable for predicting the chemical content of food. Becker et al. (2003) have applied PLSR to quantitatively determine the riboflavin content in yogurt with very high correlation coefficient of 0.99. PLSR is also a good tool to correlate spectra data to another set of data, such as rheological properties. From tryptophan emission spectra, Karoui and Dufour (2005) were able to predict

rheological values such as storage modulus (G'), loss modulus (G"), strain, tan(δ), and complex viscosity measured at 80 °C on the semi-hard cheese with high correlation coefficient of 0.98, 0.97, 0.98, 0.98, and 0.97 respectively.

References

ADPI - American Dairy Products Institute. 2016. What does WPNI tell me and how does it compare with other test methods? Accessed Nov.11, 2019. https://www.adpi.org/Portals/0/Academy/What%20does%20WPNI%20tell%20me.pdf

Andersen, C. M., and Mortensen, G. 2008. Fluorescence spectroscopy: a rapid tool for analyzing dairy products. J Agric Food Chem, 56(3), 720-729.

Angulo, F. J., LeJeune, J. T., and Rajala-Schultz, P. J. 2009. Unpasteurized milk: a continued public health threat. Clin Infect Dis, 48(1), 93-100.

Arnott, D. R., Morris, H. A., and Combs, W. B. 1957. Effect of certain chemical factors on the melting quality of process cheese. J Dairy Sci, 40(8), 957-963.

Bahram, M., Bro, R., Stedmon, C., and Afkhami, A. 2006. Handling of Rayleigh and Raman scatter for PARAFAC modeling of fluorescence data using interpolation. J Chemom, 20(3‐4), 99-105.

Beaman, T. C., Greenamyre, J. T., Corner, T. R., Pankratz, H. S., and Gerhardt, P. 1982. Bacterial spore heat resistance correlated with water content, wet density, and protoplast/sporoplast volume ratio. J Bacteriol, 150(2), 870-877.

Becker, E. M., Christensen, J., Frederiksen, C. S., and Haugaard, V. K. 2003. Front-face fluorescence spectroscopy and chemometrics in analysis of yogurt: rapid analysis of riboflavin. J Dairy Sci, 86(8), 2508-2515.

Berger, W., Klostermeyer, H., Merkenich, K., and Uhlmann, G. 1998. Process cheese manufacture. Ladenberg, Germany: BK Guilini Chemie GmbH and Co.

Birlouez-Aragon, I., Nicolas, M., Metais, A., Marchond, N., Grenier, J., and Calvo, D. 1998. A rapid fluorimetric method to estimate the heat treatment of liquid milk. Int Dairy J, 8(9), 771-777.

Birlouez-Aragon, I., Sabat, P., and Gouti, N. 2002. A new method for discriminating milk heat treatment. Int Dairy J, 12(1), 59-67.

Breene, W. M. 1975. Application of texture profile analysis to instrumental food texture evaluation. J Texture Stud, 6(1), 53-82.

Burgess, S. A., Brooks, J. D., Rakonjac, J., Walker, K. M., and Flint, S. H. 2009. The formation of spores in biofilms of *Anoxybacillus flavithermus*. J Appl Microbiol, 107(3), 1012-1018.

Burgess, S. A., Lindsay, D., and Flint, S. H. 2010. Thermophilic bacilli and their importance in dairy processing. Int J Food Microbiol, 144(2), 215-225.

Bylund, G. 2003. Dairy processing handbook. Tetra Pak Processing Systems AB.

CFR - Code of Federal Regulations Title 21. Accessed Nov 11, 2019. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=131.125.

CFR - Code of Federal Regulations Title 21. Accessed Nov 11, 2019. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=133.169

Chen, L., Coolbear, T., and Daniel, R. M. 2004. Characteristics of proteinases and lipases produced by seven Bacillus sp. isolated from milk powder production lines. Int Dairy J, 14(6), 495-504.

Chopra, A. K., and Mathur, D. K. 1984. Isolation, screening and characterization of thermophilic *Bacillus* species isolated from dairy products. J Appl Microbiol, 57(2), 263-271.

Coorevits, A., De Jonghe, V., Vandroemme, J., Reekmans, R., Heyrman, J., Messens, W., De Vos, P., and Heyndrickx, M. 2008. Comparative analysis of the diversity of aerobic spore-forming bacteria in raw milk from organic and conventional dairy farms. Syst Appl Microbiol, 31(2), 126-140.

Cosentino, S., Mulargia, A. F., Pisano, B., Tuveri, P., and Palmas, F. 1997. Incidence and biochemical characteristics of *Bacillus* flora in Sardinian dairy products. Int J Food Microbiol, 38(2-3), 235-238.

Crielly, E. M., Logan, N. A., and Anderton, A. 1994. Studies on the *Bacillus* flora of milk and milk products. J Appl Microbiol, 77(3), 256-263.

Drake, M. A., Gerard, P. D., Truong, V. D., and Daubert, C. R. 1999. Relationship between instrumental and sensory measurements of cheese texture. J Texture Stud, 30(4), 451-476.

Dufour, E., Devaux, M. F., Fortier, P., and Herbert, S. 2001. Delineation of the structure of soft cheeses at the molecular level by fluorescence spectroscopy—relationship with texture. Int Dairy J, 11(4-7), 465-473.

Dufour, E., and Riaublanc, A. 1997. Potentiality of spectroscopic methods for the characterisation of dairy products. I. Front-face fluorescence study of raw, heated and homogenised milks. Lait, 77(6), 657-670.

Duggan, D. E., Bowman, R. L., Brodie, B. B., and Udenfriend, S. 1957. A spectrophotofluorometric study of compounds of biological interest. Arch Biochem Biophys, 68(1), 1-14.

Eijlander, R. T., van Hekezen, R., Bienvenue, A., Girard, V., Hoornstra, E., Johnson, N. B., Meyer, R., Wagendrop, A., Walker, D. C., and Wells‐Bennik, M. H. 2019. Spores in dairy‐new insights in detection, enumeration and risk assessment. Int J Dairy Technol, 72(2), 303-315.

Flint, S. H., Bremer, P. J., and Brooks, J. D. 1997. Biofilms in dairy manufacturing plant‐description, current concerns and methods of control. Biofouling, 11(1), 81-97.

Flint, S., Palmer, J., Bloemen, K., Brooks, J., and Crawford, R. 2001. The growth of *Bacillus stearothermophilus* on stainless steel. J Appl Microbiol, 90(2), 151-157.

Flint, S., Drocourt, J. L., Walker, K., Stevenson, B., Dwyer, M., Clarke, I., and McGill, D. 2006. A rapid, two-hour method for the enumeration of total viable bacteria in samples from commercial milk powder and whey protein concentrate powder manufacturing plants. Int Dairy J, 16(4), 379-384.

Flint, S., Walker, K., Waters, B., and Crawford, R. 2007. Description and validation of a rapid (1 h) flow cytometry test for enumerating thermophilic bacteria in milk powders. J Appl Microbiol, 102(4), 909-915.

Fox, P. F. 1993. Cheese: an overview. In Cheese: chemistry, physics and microbiology (pp. 1-36). Springer, Boston, MA.

Gilmour, A., and Rowe, M. T. 1990. Micro-organisms associated with milk. Dairy Microbiol, 1, 37-75.

Goodacre, R., Shann, B., Gilbert, R. J., Timmins, E. M., McGovern, A. C., Alsberg, B. K., Bell, D. B., and Logan, N. A. 2000. Detection of the dipicolinic acid biomarker in *Bacillus* spores using Curie-point pyrolysis mass spectrometry and Fourier transform infrared spectroscopy. Anal Chem, 72(1), 119-127.

Granger, C., Da Costa, J. P., Toutain, J., Barey, P., and Cansell, M. 2006. Mapping of ice cream formulation using front-face fluorescence spectroscopy. Int Dairy J, 16(5), 489-496.

Griffiths, M. W. 1990. Toxin production by psychrotrophic *Bacillus* spp. present in milk. J Food Prot, 53(9), 790-792.

Gunasekaran, S., and Ak, M. M. 2002. Cheese rheology and texture. CRC press.

Guimet, F., Boqué, R., and Ferré, J. 2006. Application of non-negative matrix factorization combined with Fisher's linear discriminant analysis for classification of olive oil excitation–emission fluorescence spectra. Chemometr Intell Lab Syst, 81(1), 94-106.

Guinee, T. P. 2002. The functionality of cheese as an ingredient: a review. Aust J Dairy Technol, 57(2), 79.

Hamdy, M. E., and Amin, A. S. 2002. Detection of Brucella species in the milk of infected cattle, sheep, goats and camels by PCR. Vet J, 163(3), 299-305.

Harper, W. J. 1992. Lactose and lactose derivatives. In Whey and lactose processing (pp. 317-360). Springer, Dordrecht.

Henihan, L. E., O'Donnell, C. P., Esquerre, C., Murphy, E. G., and O'Callaghan, D. J. 2019. Fluorescence-based analyser as a rapid tool for determining soluble protein content in dairy ingredients and infant milk formula. Innov Food Sci Emerg Technol, 52, 75-79.

Heyndrickx, M., and Scheldeman, P. 2008. Bacilli associated with spoilage in dairy products and other food. Pages 64-82 in Applications and Systematics of Bacillus and Relatives.

Hindle, A., and Hall, E. H. 1999. Dipicolinic acid (DPA) assay revisited and appraised for spore detection. Analyst, 124(11), 1599-1604.

Humphrey, T. J., and Beckett, P. 1987. *Campylobacter jejuni* in dairy cows and raw milk. Epidemiol Infect, 98(3), 263-269.

ISO. 2009. ISO/TS 27265:2009 "Dried milk – Enumeration of the specially thermoresistant spores of thermophilic bacteria". In International Standard Organization (ISO) and International Dairy Federation (Ed.), Milk and processed milk products. Geneva, Switserland: International Standard Organization.

Kapoor, R., and Metzger, L. E. 2008. Process cheese: scientific and technological aspects—a review. Compr Rev Food Sci Food Saf, 7(2), 194-214.

Karoui, R., and Blecker, C. 2011. Fluorescence spectroscopy measurement for quality assessment of food systems—a review. Food Bioproc Tech, 4(3), 364-386.

Karoui, R., Bosset, J. O., Mazerolles, G., Kulmyrzaev, A., and Dufour, É. 2005. Monitoring the geographic origin of both experimental French Jura hard cheeses and Swiss Gruyère and L'Etivaz PDO cheeses using mid-infrared and fluorescence spectroscopies: a preliminary investigation. Int Dairy J, 15(3), 275-286.

Karoui, R., and Dufour, É. 2003. Dynamic testing rheology and fluorescence spectroscopy investigations of surface to centre differences in ripened soft cheeses. Int Dairy J, 13(12), 973-985.

Karoui, R., and Dufour, É. 2006. 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 Dairy J, 16(12), 1490-1497.

Karoui, R., Dufour, E., Schoonheydt, R., and De Baerdemaeker, J. 2007. Characterisation of soft cheese by front face fluorescence spectroscopy coupled with chemometric tools: Effect of the manufacturing process and sampling zone. Food Chem, 100(2), 632-642.

Karoui, R., Martin, B., and Dufour, É. 2005. Potentiality of front-face fluorescence spectroscopy to determine the geographic origin of milks from the Haute-Loire department (France). Lait, 85(3), 223-236.

Karoui, R., Mazerolles, G., and Dufour, É. 2003. Spectroscopic techniques coupled with chemometric tools for structure and texture determinations in dairy products. Int Dairy J, 13(8), 607-620.

Karoui, R., Mouazen, A. M., Dufour, E., Schoonheydt, R., and De Baerdemaeker, J. 2006. Utilisation of front-face fluorescence spectroscopy for the determination of some selected chemical parameters in soft cheeses. Lait, 86(2), 155-169.

Kelly, A. L., O'Connell, J. E., and Fox, P. F. 2003. Manufacture and properties of milk powders. In Advanced Dairy Chemistry—1 Proteins (pp. 1027-1061). Springer, Boston, MA.

Kelly, A. L., and Fox, P. F. 2016. Manufacture and properties of dairy powders. In Advanced Dairy Chemistry (pp. 1-33). Springer, New York, NY.

Kosikowski, F. V., and Mistry, V. V. 1997. Cheese and fermented milk foods. Volume 1: origins and principles. Cheese and fermented milk foods. Volume 1: origins and principles., (Ed. 3).

Kulmyrzaev, A., Dufour, E., Noël, Y., Hanafi, M., Karoui, R., Qannari, E. M., and Mazerolles, G. 2005. Investigation at the molecular level of soft cheese quality and ripening by infrared and fluorescence spectroscopies and chemometrics—relationships with rheology properties. Int Dairy J, 15(6-9), 669-678.

Lakowicz, J. R. (Ed.). 2013. Principles of fluorescence spectroscopy. Springer Science and Business Media.

Lindsay, D., Brözel, V. S., and Von Holy, A. 2005. Spore formation in *Bacillus subtilis* biofilms. J Food Prot, 68(4), 860-865.

Liu, X., and Metzger, L. E. 2007. Application of fluorescence spectroscopy for monitoring changes in nonfat dry milk during storage. J Dairy Sci, 90(1), 24-37.

McGuiggan, J. T., McCleery, D. R., Hannan, A., and Gilmour, A. 2002. Aerobic spore‐forming bacteria in bulk raw milk: factors influencing the numbers of psychrotrophic, mesophilic and thermophilic *Bacillus* spores. Int J Dairy Technol, 55(2), 100-107.

Meyer, A. 1973. Processed cheese manufacture.

Mleko, S., and Foegeding, E. A. 2000. Physical properties of rennet casein gels and processed cheese analogs containing whey proteins. Milchwissenschaft, 55(9), 513-516.

Mostert, J. F., Luck, H., and Husmann, R. A. 1979. Isolation, identification and practical properties of Bacillus species from UHT and sterilized milk. South African journal of dairy technology.= Suid-Afrikaanse tydskrif vir suiweltegnologie. 210 0 S Afr J Dairy Technol Suid Afrika Tydskrif Suiweltegnolog.

Munck, L., Nørgaard, L., Engelsen, S. B., Bro, R., and Andersson, C. A. 1998. Chemometrics in food science—a demonstration of the feasibility of a highly exploratory, inductive

evaluation strategy of fundamental scientific significance. Chemometr Intell Lab Syst, 44(1-2), 31-60.

Murphy, P. M., Lynch, D., and Kelly, P. M. 1999. Growth of thermophilic spore forming bacilli in milk during the manufacture of low heat powders. Int J Dairy Technol, 52(2), 45-50.

Oldfield, D. J., Taylor, M. W., and Singh, H. 2005. Effect of preheating and other process parameters on whey protein reactions during skim milk powder manufacture. Int Dairy J, 15(5), 501-511.

Oliver, S. P., Jayarao, B. M., and Almeida, R. A. 2005. Foodborne pathogens in milk and the dairy farm environment: food safety and public health implications. Foodborne Pathog Dis, 2(2), 115-129.

Olson, N. F., and Price, W. V. 1958. A melting test for pasteurized process cheese spreads. J Dairy Sci, 41(7), 999-1000.

Panikuttira, B., Payne, F. A., O'Shea, N., Tobin, J. T., O'Callaghan, D. J., and O'Donnell, C. P. 2020. Investigation of an in‐line prototype fluorescence and infrared backscatter sensor to monitor rennet‐induced coagulation of skim milk at different protein concentrations. Int J Food Sci Technol, 55(1), 175-182.

Parkar, S. G., Flint, S. H., Palmer, J. S., and Brooks, J. D. 2001. Factors influencing attachment of thermophilic bacilli to stainless steel. J Appl Microbiol, 90(6), 901-908.

Park, C. W., and Drake, M. 2016. Condensed milk storage and evaporation affect the flavor of nonfat dry milk. J Dairy Sci, 99(12), 9586-9597.

Patel, H.A. , Anema, S.G., Holroyd, S.E., Singh, H., and Creamer, L.K. 2007. Methods to determine denaturation and aggregation of proteins in low-, medium- and high-heat skim milk powders. Lait 87:251-268

Peleg, M. 1976. Texture profile analysis parameters obtained by an Instron universal testing machine. J Food Sci, 41(3), 721-722.

Pellegrino, P. M., Fell, N. F., Rosen, D. L., and Gillespie, J. B. 1998. Bacterial endospore detection using terbium dipicolinate photoluminescence in the presence of chemical and biological materials. Anal Chem, 70(9), 1755-1760.

Pellegrino, P. M., Fell Jr, N. F., and Gillespie, J. B. 2002. Enhanced spore detection using dipicolinate extraction techniques. Anal Chim Acta, 455(2), 167-177.

Piska, I., and Štětina, J. 2004. Influence of cheese ripening and rate of cooling of the processed cheese mixture on rheological properties of processed cheese. J Food Eng, 61(4), 551-555.

Poulli, K. I., Mousdis, G. A., and Georgiou, C. A. 2005. Classification of edible and lampante virgin olive oil based on synchronous fluorescence and total luminescence spectroscopy. Anal Chim Acta, 542(2), 151-156.

Prow, L. A., and Metzger, L. E. 2005. Melt analysis of process cheese spread or product using a rapid visco analyzer. J Dairy Sci, 88(4), 1277-1287.

Purna, S. G., Prow, L. A., and Metzger, L. E. 2005. Utilization of front-face fluorescence spectroscopy for analysis of process cheese functionality. J Dairy Sci, 88(2), 470-477.

Ronimus, R. S., Parker, L. E., Turner, N., Poudel, S., Rückert, A., and Morgan, H. W. 2003. A RAPD-based comparison of thermophilic bacilli from milk powders. Int J Food Microbiol, 85(1-2), 45-61.

Rosen, D. L., Sharpless, C., and McGown, L. B. 1997. Bacterial spore detection and determination by use of terbium dipicolinate photoluminescence. Anal Chem, 69(6), 1082-1085.

Rückert, A., Ronimus, R. S., and Morgan, H. W. 2004. A RAPD-based survey of thermophilic bacilli in milk powders from different countries. Int J Food Microbiol, 96(3), 263-272.

Rueckert, A., Ronimus, R. S., and Morgan, H. W. 2006. Development of a real-time PCR assay targeting the sporulation gene, spo0A, for the enumeration of thermophilic bacilli in milk powder. Food Microbiol, 23(3), 220-230.

SádeCká, J., and TóThoVá, J. 2007. Fluorescence spectroscopy and chemometrics in the food classification-a review. Czech J Food Scis, 25(4), 159-173.

Scheldeman, P., Herman, L., Foster, S., and Heyndrickx, M. 2006. *Bacillus sporothermodurans* and other highly heat‐resistant spore formers in milk. J Appl Microbiol, 101(3), 542-555.

Schwarzenbach, R., and Hill, B. M. 1999. Do bacterial spores in milk powders survive UHT processing?. In 3. International Symposium on recombined milk and milk products, Penang (Malaysia), 23-26 May 1999. Int Dairy Fed.

Scott, S. A., Brooks, J. D., Rakonjac, J., Walker, K. M., and Flint, S. H. 2007. The formation of thermophilic spores during the manufacture of whole milk powder. Int J Dairy Technol, 60(2), 109-117.

Sharma, A., Jana, A. H., and Chavan, R. S. 2012. Functionality of milk powders and milk‐based powders for end use applications—a review. Compr Rev Food Sci Food Saf, 11(5), 518-528.

Sikand, V., Tong, P. S., and Walker, J. 2008. Impact of protein standardization of milk powder with lactose or permeate on whey protein nitrogen index and heat classification. Dairy Sci Technol, 88(1), 105-120.

Simeone, M., Alfani, A., and Guido, S. 2004. Phase diagram, rheology and interfacial tension of aqueous mixtures of Na-caseinate and Na-alginate. Food Hydrocoll, 18, 463–470.

Singh, H., and Creamer, L. K. 1991. Denaturation, aggregation and heat stability of milk protein during the manufacture of skim milk powder. J Dairy Res, 58(3), 269-283.

Sołowiej, B., Cheung, I. W., and Li-Chan, E. C. 2014. Texture, rheology and meltability of processed cheese analogues prepared using rennet or acid casein with or without added whey proteins. Int Dairy J, 37(2), 87-94.

Stathopoulos, C. E. 2008. Dairy-based ingredients. In Gluten-free cereal products and beverages (pp. 217-236). Academic Press.

Sutheerawattananonda, M., and Bastian, E. D. 1998. Monitoring process cheese meltability using dynamic stress rheometry 1. J Texture Stud, 29(2), 169-183.

Tamime, A. Y. 2011. Processed cheese and analogues (Vol. 16). John Wiley and Sons.

Thapa, T. B., and Gupta, V. K. 1992. Rheology of processed cheese foods with added whey protein concentrates. Indian J. Dairy Sci, 45, 88-92.

Thomas, M. A. 1969. Browning reaction in Cheddar cheese. Aust J Dairy Technol, 24(4), 185.

USDA. 2011. United States Standards for Grades of Nonfat Dry Milk (Spray Process). Accessed Nov 11, 2019.

USDEC. 2018. Non-fat Dry Milk & Skim Milk Powder. Accessed Sept 5, 2020. https://www.thinkusadairy.org/products/milk-powders/milk-powder-categories/non-fat-dry-milk-and-skim-milk-powder

Valeur, B., and Brochon, J. C. (Eds.). 2012. New trends in fluorescence spectroscopy: applications to chemical and life sciences (Vol. 1). Springer Science and Business Media.

Wold, J. P., Veberg, A., Lundby, F., Nilsen, A. N., and Moan, J. 2006. Influence of storage time and color of light on photooxidation in cheese: A study based on sensory analysis and fluorescence spectroscopy. Int Dairy J, 16(10), 1218-1226.

Wong, D. W., Camirand, W. M., Pavlath, A. E., Parris, N., and Friedman, M. 1996. Structures and functionalities of milk proteins. Crit Rev Food Sci Nutr, 36(8), 807-844.

Zehren, V. L., and Nusbaum, D. D. 2000. Processed cheese. Cheese Reporter Publishing Co. Inc., Madison.

Chapter 3 - Research Objectives

This thesis is interested in applying fluorescence spectroscopy on the prediction of process cheese emulsion characteristics and classification of nonfat dry milk based on spore levels. The specific objectives include

- To build a classification model for differentiating nonfat dry milk based on the spore count categories using discriminant analysis and random forest
- To establish a regression model for quantifying the emulsion properties of process cheese using partial least square regression

Chapter 4 - Fluorescence Spectroscopy Combined with Chemometrics for Classification of Nonfat Dry Milk Based on Spore Counts

Abstract

Nonfat dry milk (NDM) is a popular ingredient in a wide range of shelf-stable food products. However, high spore containing NDM can lead to ropiness and introduce unwanted lipase and protease activity. Thermophilic and mesophilic spores can enter into raw milk through the cow, feed, and beddings at the farm level. Further, spore counts can increase during the manufacture of NDM due to the concentration factor as well as contamination from matured biofilms formed on the equipment. Therefore, the spore count is a critical quality indicator to be monitored during the production. Previous research suggests that dipicolinic acid (DPA) is present in the core of endospores and can be used as a fluorophore of interest for rapid detection of the presence of spores. The objective of this study was to use DPA fluorescence spectra and chemometrics to develop classification models based on the spore levels. Commercial NDM samples (n=40) were procured within the United States. The reference spore counts (CFU/g of NDM) were obtained by heating reconstituted NDM (10%) at 100 °C for 30 min, followed by plating on Tryptic Soy Agar, and incubating at 55 °C for 48 hours. In order to release all available DPA and to remove interferents, the reconstituted NDM (10%) was autoclaved at 121 °C for 30 min followed by acidification and centrifugation. Terbium chloride was added to the supernatant buffered to pH 5.6 to enhance the DPA fluorescence signal. Emission spectra of the terbium DPA complex were collected between 450-650 nm fixed at the excitation of 270 nm. Classification models were developed using partial least square quadratic discriminant analysis

(PLS-QDA), forward selection quadratic discriminant analysis (FS-QDA), and random forest (RF). It was found that random forest provided the highest mean classification accuracy of 87% while FS-QDA and PLS-QDA showed the mean accuracy at 84% and 83%, respectively (validated using bootstrapping technique). The results suggest the potential of using fluorescence spectroscopy to classify the NDM based on spore counts.

Introduction

Nonfat dry milk (NDM) is a dairy-based dry ingredient manufactured from vacuum evaporation and spray drying of skim milk. It is commonly used as an ingredient in many shelf-stable dairy and food products such as process cheese and UHT-treated beverages, as it provides functional and nutritional benefits depending on the application. Thus, to ensure food safety, skim milk must undergo a pre-heating treatment step to kill vegetative bacteria before the multi-stage evaporation and spray drying (Kelly and Fox, 2016). However, some thermophilic and mesophilic spores cannot be completely deactivated during traditional pasteurization, and therefore can remain in the final powder to cause major quality issues (Schwarzenbach and Hill, 1999; Mostert et al.,1979). These spores are not a food safety concern; however substantial quality loss can occur as a result of sporulation. The sporeforming with *Bacillus* genus has been reported to cause ropiness and produce lactic acid in UHT and canned milk products (Heyndrickx and Scheldeman, 2008; Gilmour and Rowe, 1990). In other cases, some spores can introduce unwanted enzymatic activity that can be a detriment to the flavor profile of finished products (Chen et al, 2004; Chopra and Mathur, 1984; Cosentino et al., 1997). When NDM was used as an ingredient in food products such as soups and beverages, poor quality may lead to decreased consumer acceptability of end products. Thus, spore level is an important factor among NDM quality parameters.

High spore counts in NDM can occur due to contamination from diverse sources including cow, feed, environment, and milking facilities (Coorevits et al., 2008). Subsequently, spores can germinate and lead to biofilm development on the surfaces of processing equipment during processing. The biofilm development usually occurs on processing units that operate at a relatively high temperature, such as heat exchange plates (Flint et al., 1997; Murphy et al., 1999). Mature biofilms can induce secondary contamination in the following batches. Concentration effect is another factor that contributes to the increased level of spores. Skim milk contains approximately 90% water. As the water evaporates throughout the process, the solid matter concentrates, leading to a significant increase in spore count (Buehner et al., 2014).

To monitor the spore level during the manufacture of NDM, it is a common practice to apply traditional plating methods, such as total thermophile plate count (TPC) and a thermophilic spore count (TSC). During laboratory spore analysis, the time and temperature combination used to kill vegetative cells before incubating the spores as well as the selection of incubation media often vary between methods, resulting in variations in spore counts. Moreover, these methods can take up to two days to complete. It could be too late before actions to address the high spore contamination are initiated. Previously, alternative analytical methods using fluorescence spectroscopy have been developed to estimate the spore counts. The measurement principle is based on the dipicolinic acid (DPA) content, a fluorophore of interest present in spores up to 15% of spore dry weight. Studies have been reported using DPA as a biomarker to quantify spores (Hindle and Hall, 1999; Pellegrino et al., 1998; Rosen et al., 1997). By applying different extraction techniques for DPA, Pellegrino et al. (2002) achieved the lowest detection limit at 10^3 CFU/ml in a spore suspension. Yet, no research has been reported evaluating the fluorescence based rapid method in the milk matrix. Therefore, the objective of the study was to evaluate the

fluorescence spectroscopy as a tool to classify NDM based on the spore level into different categories. It could be a useful tool for the dairy industry to make timely decision to take necessary actions to keep the spore levels under control.

Materials and Methods

Experimental design

Preliminary experiments were carried out to understand the fluorescence spectra of terbium dipicolinate spiked at different concentrations in a skim milk matrix. Skim milk without any added terbium dipicolinate was used as a control. In the next phase, commercial samples of NDM with varying degrees of spore presence were procured and spore counts in the NDM powders were determined using the reference method. Also, emission spectra of terbium dipicolinate were acquired for the NDM samples. Discriminant analysis was performed to develop classification model with emission spectra of terbium dipicolinate as input to predict the reference spore counts in the NDM samples.

Preliminary experiments

In order to prove that terbium dipicolinate can elicit a fluorescence signal in the skim milk matrix, an experimental protocol was developed in the lab. Samples of NDM were procured from the Kansas State University dairy plant. NDM powders were reconstituted to 10% total solids by dissolving 4 g of NDM into 36 g DI water under constant stirring for 30 min. Subsequently, 1 mL of 20% w/w trichloro acetic acid (TCA; Fisher Scientific, Pittsburgh, PA, USA) was added to precipitate the protein. The reconstituted skim milk and TCA mixture was centrifuged (Marathon 21000R model 120; Fisher Scientific, Pittsburgh, PA, USA) at 10,000 × g RCF (relative centrifugal force) for 10 min at 4 °C. The resultant supernatant was divided into 4

sub-samples in separate test tubes. Nine mL 1M sodium acetate buffer (pH = 5.6) and dipicolinic acid was added to each sub-sample (Samples A, B, C, and D). The sample A was the control group while samples B, C, D were dosed with 0.1 mL of $6X10^{-5}$, $6X10^{-4}$, $6X10^{-3}$ M dipicolinic acid solutions, respectively. The fluorescence measurement was set out to collect emission regions between 450 - 650 nm using a fluorescence spectrometer (LS-55; Perkin Elmer, Waltham, MA, USA). The 430 nm cut-off filter was applied, and excitation was fixed at 270 nm. The excitation and emission monochromator filters were set to 9 nm and 4 nm, respectively. For each sample, three scans were performed and averaged.

Nonfat dry milk samples

For the actual experiment, commercial NDM samples were procured from Dairy Farmers of America (DFA) (n=30) and from Cornell University (n=10). The powder was stored in whirl pak bags at -20 °C until the experiment. Samples included 2 high heat NDM, 1 medium heat NDM, 13 low heat NDM, and heat classification of the rest NDM was not provided. The heat classification was not used as a predictor variable because the aim of this study was to build a robust model applicable to all pre-heating levels experienced by NDM.

Thermophilic and mesophilic spore counts

The spore counts in NDM were obtained from using modified ISO method as described by Eijlander et al. (2019). Briefly, one gram of NDM was dissolved in 9 mL sterilized DI water in a whirl-pak detectable bag of 29 mL (Whirl-pak, Madison, MI, USA) and stomached in an AES Blue Line Smasher (Biomerieux, Marcy-l' Etoile, France) for 2 min. The reconstituted skim milk was further boiled in 15 ml conical centrifuge tube (VWR) at 100 °C in a water bath for 30 min to kill all vegetative cells while spores remained. After heating, samples were serially diluted using 0.1% peptone water and plated on TSA plates (Becton, Dickinson and Company).

These plates were incubated at 55 °C and counted after 48 hours for the best recovery of spores of interest according to Eijlander et al. (2019). The spore counts in NDM were reported as colony forming units (CFU)/g of NDM.

Right-angle fluorescence spectroscopy analysis

The NDM was pre-processed to enhance the signal for the fluorescence measurement. The powder was reconstituted in the same manner as for the spore count. Then, the skim milk in a 10 mL test tube was autoclaved at 121 °C for 30 min to completely release the available DPA in the spores. The sample was then acidified by adding 0.5 mL 1 N HCl (Certified ACS, Fisher Scientific, Hampton, NH, USA) to pH 4.6 and centrifuged at 10,000 × g RCF and 4 °C for 10 min. These two procedures were to ensure the removal of protein that can interfere with the fluorescence spectra. To enhance the signal, 2 mL supernatant was then buffered with 1 mL 1 M sodium acetate (pH = 5.6) and 0.1 mL $TbCl_3$ at concentration of $2.68X10^{-2}$ M. The $TbCl_3$ solution was made by dissolving the appropriate quantity of $TbCl_3$ crystals (Fisher Scientific, Pittsburgh, PA, USA) into DI water. After the mixing, the calibration sample was transferred into a Quartz cuvette (Starna Cells Inc., Atascadero, CA, USA). DPA in reconstituted NDM samples formed a terbium dipicolinate complex with Tb^{3+} ions in the $TbCl^3$ solution. The emission spectra for terbium dipicolinate from 450 - 650 nm were acquired using fluorescence spectroscopy (LS-55; Perkin Elmer, Waltham, MA, USA) with a 430 nm cut-off filter and excitation at 270 nm. The excitation and emission monochromator slit width were set at 9 nm and 4 nm, respectively for the best spectral results. Three scans were performed consecutively and averaged automatically by the built-in program to improve the signal-to-noise ratio.

Pre-processing of NDM fluorescence spectra

Pre-processing is common to spectra analysis prior to developing a calibration model. This treatment can reduce the impact of noise, enhance features, and obtain a cleaner data structure. For the spectra data obtained in this study, the standard normal variate (SNV) and second derivative based on Savitzky-Golay algorithm with 5-point neighbor values were performed on the emission spectra, respectively (Savitzky and Golay, 1964; Barnes et al., 1989). SNV allows the data to be compared at the same normalized scale while Savitzky-Golay second derivative smooths the spectra and enhances the wavelengths region responsible for highest intensity changes. The procedures are shown in Figure 4.1.

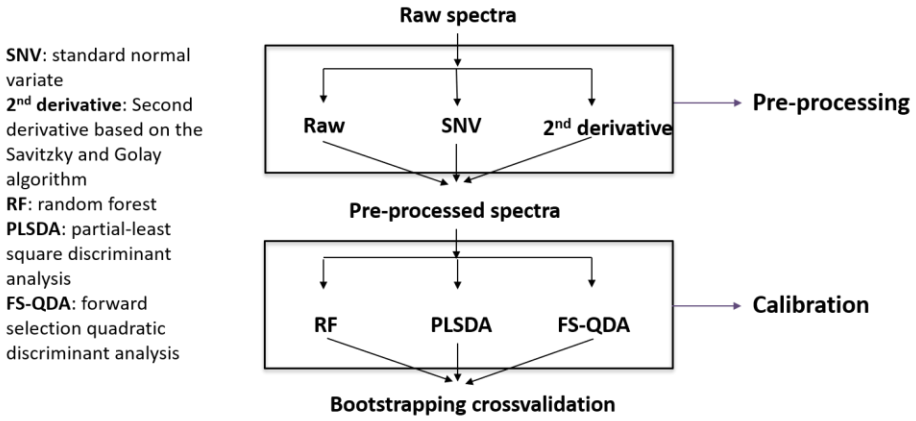

Figure 4.1. Procedures for pre-processing, model calibration and validation

Random forest

Random forest is a classification learning algorithm based on decision trees. It can handle high-dimensional data and ignore the irrelevant descriptors or variables (Svetnik et al., 2003). Rather than using all variables at once, it applies a subset of variables to split the samples into different nodes. The number of variables used to differentiate at each node is denoted as m_{try}. By default, m_{try} is equal to the square root of the number of all variables. Calibration with less m_{try} is

faster to perform. To train the data set, a sample is drawn from the bootstrapping data set with replacement. The classification tree is grown for each sample with best-fitting m_{try}. The process is repeated until a sufficient number of trees were grown. Finally, the error rate is calculated as the proportion of misclassified test samples from total samples, as shown in Equation 1 where $E_{X,Y}$ is the distribution of expected class Y from a feature vector X, and $\hat{Y}$ is the predicted class (Svetnik et al., 2003).

$$ER = E_{X,Y} I\left(\hat{Y} \neq Y\right) = \Pr\left\{\hat{Y} \neq Y\right\} \quad (1)$$

Forward selection quadratic discriminant analysis (FS-QDA)

Forward selection is the most straightforward method to choose the variables based on their capacity to improve the model performance (Nørgaard et al., 2000). First, all wavelengths were tested individually for the discriminating ability. Then, the most significant variable was tested in combination with the remaining variables one by one, until the maximum number of variables were retained. The maximum number of variables can be manually determined to be included in the quadratic discriminant analysis (QDA). QDA works by projecting the spectral data into a high dimension space and maximizing the similarity between samples within the same group (Amamcharla et al., 2010). A discriminating quadratic function was used to separate the groups. The quadratic function is displayed in Equation 2 and 3.

$$\Sigma_k = \frac{1}{N_k} \sum_{i=1}^{N_k} (x_i - \mu_k)(x_i - \mu_k)^T \quad (2)$$

$$\delta_k(x) = -\frac{1}{2} \log|\Sigma_k| - \frac{1}{2} (x - \mu_k)^T {\Sigma_k}^{-1} (x - \mu_k) + \log(\pi_k) \quad (3)$$

Partial least square discriminant analysis (PLS-DA)

Partial least square was used to select the principle components that have the highest prediction power. The same discriminant analysis was followed as in the FS-QDA section.

Bootstrapping validation

Bootstrapping is a resampling method in which replacement is allowed. The sample size generated using this technique is equal to the original sample size. The bootstrapping internal validation was selected due to the small size of the data set. To split the sample further into training and testing set would further decrease the sample size, and thus increase the expected error rate. To account for the small sample size, the bootstrapping is repeated numerous times. The principle of bootstrap can be explained in Equation 4 where the estimate $\hat{\theta}$ is obtained from applying function u to bootstrap data sets x, which follows the distribution $\hat{F}$ (Wehrens et al., 2020). For this study, RF and FS-QDA were performed on 25 bootstrap datasets using caret package in RStudio (Kuhn, 2008). This number was found to be sufficient as more replications did not improve the model performance. Whereas for PLSDA validation, bootstrapping was carried out 1000 iterations in SAS (version 9.4, SAS Institute, Inc., Cary, NC) to minimize the bias due to small sample size according to Balasubramanian et al. (2004). The average accuracy was calculated from the bootstrapping validation (Wehrens et al., 2000) and the model performance was compared through the averaged classification success rate.

$$\hat{\theta} = t(\hat{F}) = u(x) \quad (4)$$

Results and Discussion

Preliminary results

Control A was the unspiked reconstituted control NDM sample while another three samples were spiked with dipicolinic acid to achieve a final concentration of $6X10^{-5}$ M, $6X10^{-6}$ M and $6X10^{-7}$ M. Results shown in Figure 4.2 clearly show that with the addition of DPA, the fluorescence spectrum exhibited four peaks at 480-500 nm, 540-560 nm, 580-600 nm, and 620-630 nm. Along with the peaks, a decrease of intensity at other regions were also observed for the samples at higher concentrations of DPA. Concentration at $6X10^{-5}$ M and $6X10^{-6}$ can be visibly differentiated from the blank while that at $6X10^{-7}$ may require further processing to accentuate the spectral difference. The detection limit was comparable to the lowest value that other researchers have achieved, which was 2 nM DPA considering the matrix is different (Hindle and Hall, 1999). Thus, skim milk with varying DPA concentrations was expected to have different fluorescence fingerprint that can be used to classify the spore levels, given that DPA concentration was directly related to the spore count. Previous researcher that applied the enhanced extraction techniques to release the DPA from spores have achieved the lowest detection limit at 1,000 CFU/mL, which makes its application in milk matrix more promising.

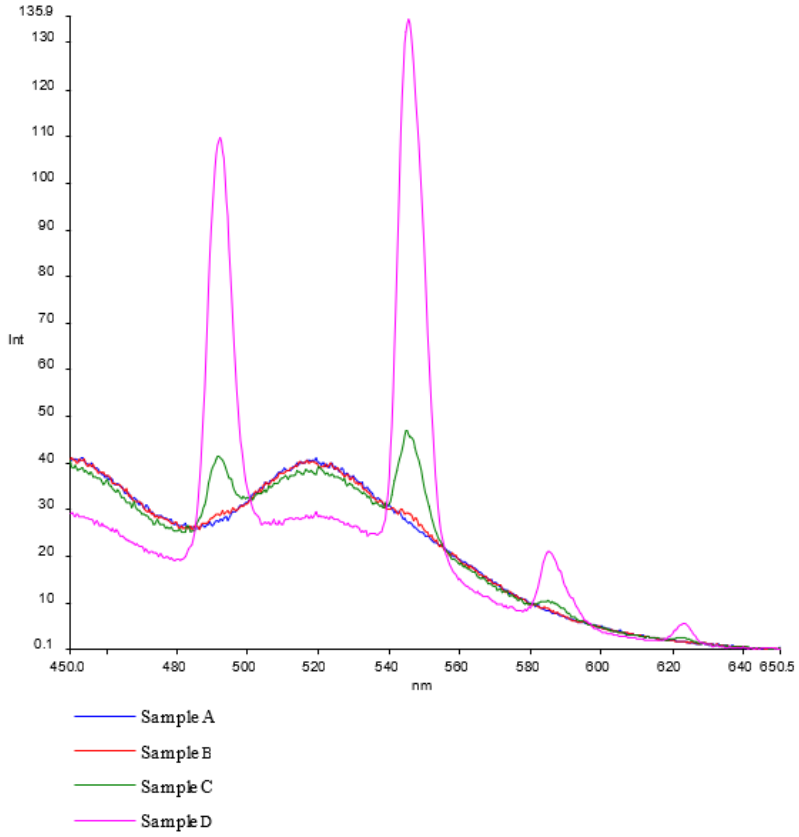

Figure 4.2. Fluorescence emission spectra of terbium dipicolinate at different concentrations in a skim milk matrix (Samples A = 0, B = $6X10^{-7}$ M, C = $6X10^{-6}$ M, and D = $6X10^{-5}$ M)

Spore count in NDM samples

Figure 4.3 shows the thermophilic and mesophilic spore counts obtained for the 40 NDM samples and fall within a range of 1.7 to 5.0 log CFU/g. This represented a 3.3-log span in spore counts, suitable for constructing a representative classification model for any reasonable spore counts in NDM. The median of this set of data was located at 2.6 log CFU/g, while 27 out of 40 samples demonstrated spore counts below 3.0 log CFU/g. The NDM samples with a spore level above 3 log CFU/g was designated as high spore NDM, while the rest of samples were grouped as low spore NDM. The 3 log CFU/g was set according to the industry recommendation as contaminated milk powder with higher than 3 log spores per gram NDM show strong heat

resistance, and thus applying effective heat treatment can further induce more quality problems (Kent et al., 2016). Overall, the distribution of spore counts indicated that the majority of NDM powder samples were produced under relatively acceptable sanitation conditions so that the spore counts ended up below the recommended level.

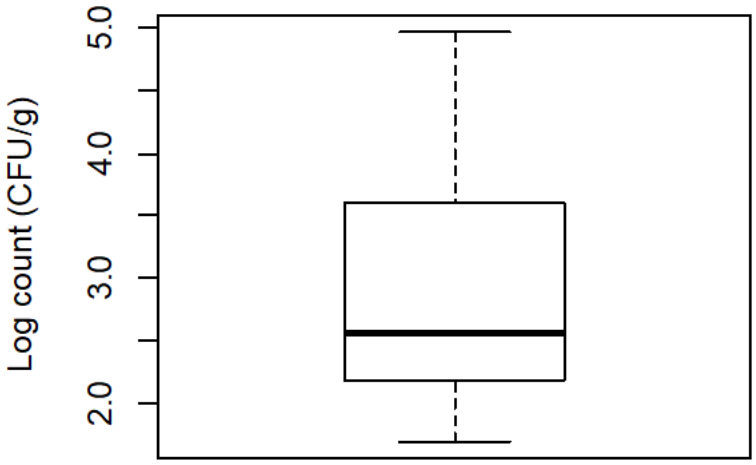

Figure 4.3. Boxplot of spore levels from 40 commercial NDM samples

Fluorescence spectroscopy analysis

Figure 4.4 shows a sample spectrum of terbium dipicolinate as the complex formed between DPA in spores and terbium ions provided in the solution. Terbium ions need to be provided in sufficient quantity to favor the formation of the terbium dipicolinate complex, and yet an excess amount can lead to reduced sensitivity (Hindle and Hall, 1999). Two visible peaks from terbium DPA were observed at 475 - 480 nm and 525 - 530 nm, despite the overlay with tryptophan emission spectra from free amino acids. The existence of more than one emission wavelength peak has been also reported in the literature (Rosen et al., 1997) and is due to multiple double bonds in the DPA chemical structure. Two peaks observed from this study can

be used to quantify the terbium dipicolinate complex that corresponds to the spore count in NDM. The remaining samples also produced spectra in a similar manner, but with a different extent of tryptophan interference. Thus, a multivariate analysis needed to be performed in order to distinguish the terbium dipicolinate peak from the baseline.

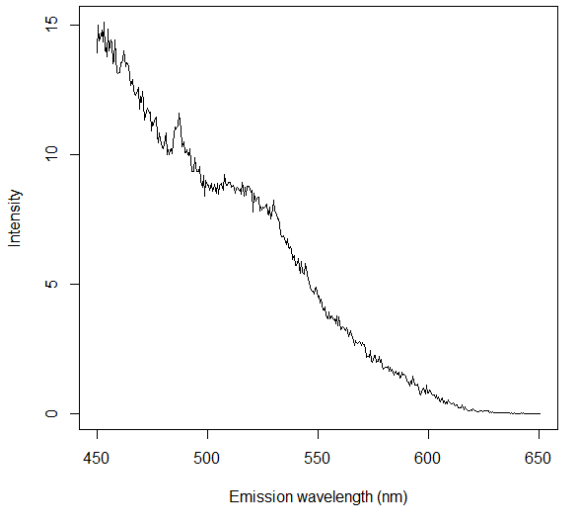

Figure 4.4. Sample emission spectra of terbium dipicolinate released from spores

Classification models

Table 4.1 shows the model performance for different pre-processing tools and classification techniques used in the present study. Pre-processing did not significantly improve the model performance. This is normal for fluorescence spectra since the specific wavelength peaks are emphasizing one particular analyte, rather than overlapping over multiple analytes at the same time. Overall, three classification techniques resulted in a reasonable success rate considering compositional variances from different NDM sources. Particularly, RF on the 2nd derivative had the highest average successful classification rate at 87%. The optimal number of variables used to split at each tree nodes was 64 with the highest kappa value at 0.66. The kappa value is defined as the accuracy normalized for the random chances on the dataset (Cohen,

1960). It is useful if the split of classes in the dataset is imbalanced. It is slightly lower than the observed accuracy of 87% because the dataset contains 67.5% high spore count, which skews the random chance of guessing correctly towards the class "high". Forward selection QDA and PLSDA showed success classification rates at 83-84% and 76-83%, respectively. The overall model performance was more sensitive towards high spore level (data not shown). Thus, the models had the tendency to overestimate the spore count, which translated to a higher chance of false positive determinations. High specificity means if the powder is tested as "low", it is less likely to be misclassified. Thus, using fluorescence is particularly suitable for the preliminary screening to select those classified as "high" for further testing.

Table 4.1. The average success rate from three calibration models (random forest, stepwise QDA, PLSDA) preprocessed by standard normal variate or Savitzky-Golay second derivative, and validated by bootstrapping method

Classification method	Pre-processing	Correctness (%)
Random forest	Raw	86
	SNV[1]	84
	2nd derivative SG[2]	87
Stepwise QDA[3] Forward selection	Raw	83
	SNV	84
	2nd derivative SG	84
PLSDA[4]	Raw	83
	SNV	76
	2nd derivative SG	77

[1]Standard normal variate
[2]Savitzky-Golay algorithm
[3]Quadratic discriminant analysis
[4]Partial least square discriminant analysis

This study builds on previous studies by attempting to classify the spore level in a milk matrix. Although no existing data on classification success rate could be found for comparison, it is worth noting that samples used in this study included 8 samples with very low spore counts

(<2 log CFU/g powder), which is out of range of most rapid detection methods. Other quantification methods such as real-time PCR (Rueckert, Ronimus, and Morgan 2006) has achieved spore enumeration in a similar time frame at a detection limit of 6,400 CFU/g, equivalently 3.8 log CFU/g, yet this limit is relatively high for the industry's NDM quality standard. At the same time, using fluorescence as an at-line process analytical technology would be simpler and more cost-effective, thus more competitive for the application purposes. Furthermore, this method only requires autoclave and centrifuge for sample preparation, conveniently purchasable from the industry standpoint. Thus, using fluorescence spectroscopy combined with chemometrics could be more practical than other methods.

Conclusions

In conclusion, this preliminary study showed the potential to use terbium dipicolinate as the fluorescence marker to quantify the spore level in NDM. The highest classification rate was obtained using random forest classification analysis on the pre-processed fluorescence spectra. Considering the sample preparation including autoclaving, centrifugation, and fluorescence measurement, the entire process takes less than an hour to complete. The classified results can advise the industry on the opportune time for hygiene activities that remove spore contamination from processing equipment. The tool could be also useful for pre-screening the NDM for further testing.

References

Amamcharla, J.K., S. Panigrahi, C.M. Logue, M. Marchello, and J.S. Sherwood. 2010. Fourier transform infrared spectroscopy (FTIR) as a tool for discriminating *Salmonella* typhimurium contaminated beef. Sens. Instrum. Food Qual. Saf.. doi:10.1007/s11694-009-9090-4.

Balasubramanian, S., Panigrahi, S., Logue, C. M., Marchello, M., Doetkott, C., Gu, H., Sherwood, J., and Nolan, L. (2004). Spoilage identification of beef using an electronic nose

system. Trans ASAE, 47(5), 1625.

Barnes, R.J., M.S. Dhanoa, and S.J. Lister. 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Appl. Spectrosc. doi:10.1366/0003702894202201.

Buehner, K. P., Anand, S., and Garcia, A. 2014. Prevalence of thermoduric bacteria and spores on 10 Midwest dairy farms. J Dairy Sci, 97(11), 6777-6784.

Chen, L., Coolbear, T., and Daniel, R. M. 2004. Characteristics of proteinases and lipases produced by seven *Bacillus sp*. isolated from milk powder production lines. Int Dairy J, 14(6), 495-504.

Chopra, A. K., and Mathur, D. K. 1984. Isolation, screening and characterization of thermophilic *Bacillus* species isolated from dairy products. J Appl Microbiol, 57(2), 263-271.

Cohen, J. 1960. A coefficient of agreement for nominal scales. Educ Psychol Meas, 20(1), 37-46.

Coorevits, A., De Jonghe, V., Vandroemme, J., Reekmans, R., Heyrman, J., Messens, W., De Vos, P., and Heyndrickx, M. 2008. Comparative analysis of the diversity of aerobic spore-forming bacteria in raw milk from organic and conventional dairy farms. Syst Appl Microbiol, 31(2), 126-140.

Cosentino, S., Mulargia, A. F., Pisano, B., Tuveri, P., and Palmas, F. 1997. Incidence and biochemical characteristics of *Bacillus* flora in Sardinian dairy products. Int J Food Microbiol, 38(2-3), 235-238.

Flint, S. H., Bremer, P. J., and Brooks, J. D. 1997. Biofilms in dairy manufacturing plant - description, current concerns and methods of control. Biofouling, 11(1), 81-97.

Gilmour, A., and Rowe, M. T. 1990. Micro-organisms associated with milk. Dairy Microbiol, 1, 37-75.

Heyndrickx, M., and Scheldeman, P. 2008. Bacilli associated with spoilage in dairy products and other food. Pages 64-82 in Applications and Systematics of Bacillus and Relatives.

Hindle, A.A., and E.A.H. Hall. 1999. Dipicolinic acid (DPA) assay revisited and appraised for spore detection. Analyst. doi:10.1039/a906846e.

Kent, D.J., K. Chauhan, K.J. Boor, M. Wiedmann, and N.H. Martin. 2016. Spore test parameters matter: Mesophilic and thermophilic spore counts detected in raw milk and dairy powders differ significantly by test method. J. Dairy Sci.. doi:10.3168/jds.2015-10283.

Kelly, A. L., and Fox, P. F. 2016. Manufacture and properties of dairy powders. In Advanced Dairy Chemistry (pp. 1-33). Springer, New York, NY.

Kuhn, M. 2008. caret Package. J. Stat. Softw.

Lee, S.K., S. Anema, and H. Klostermeyer. 2004. The influence of moisture content on the rheological properties of processed cheese spreads. Int. J. Food Sci. Technol.. doi:10.1111/j.1365-2621.2004.00842.x.

Mostert, J. F., Luck, H., and Husmann, R. A. 1979. Isolation, identification and practical properties of *Bacillus* species from UHT and sterilized milk. South African journal of dairy technology.= Suid-Afrikaanse tydskrif vir suiweltegnologie. 210 0 S Afr J Dairy Technol Suid Afrika Tydskrif Suiweltegnolog

Murphy, P. M., Lynch, D., and Kelly, P. M. 1999. Growth of thermophilic spore forming bacilli in milk during the manufacture of low heat powders. Int J Dairy Technol, 52(2), 45-50.

Nørgaard, L., A. Saudland, J. Wagner, J.P. Nielsen, L. Munck, and S.B. Engelsen. 2000. Interval partial least-squares regression (iPLS): A comparative chemometric study with an example from near-infrared spectroscopy. Appl. Spectrosc.. doi:10.1366/0003702001949500.

Pellegrino, P. M., Fell, N. F., Rosen, D. L., and Gillespie, J. B. 1998. Bacterial endospore detection using terbium dipicolinate photoluminescence in the presence of chemical and biological materials. Anal Chem, 70(9), 1755-1760.

Pellegrino, P. M., Fell Jr, N. F., and Gillespie, J. B. 2002. Enhanced spore detection using dipicolinate extraction techniques. Anal Chim Acta, 455(2), 167-177.

Rosen, D.L., C. Sharpless, and L.B. McGown. 1997. Bacterial Spore Detection and Determination by Use of Terbium Dipicolinate Photoluminescence. Anal. Chem.. doi:10.1021/ac960939w.

Savitzky, A., and M.J.E. Golay. 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal. Chem.. doi:10.1021/ac60214a047.

Schwarzenbach, R., and Hill, B. M. 1999. Do bacterial spores in milk powders survive UHT processing?. In 3. International Symposium on recombined milk and milk products, Penang (Malaysia), 23-26 May 1999. Int Dairy Fed.

Svetnik, V., A. Liaw, C. Tong, J. Christopher Culberson, R.P. Sheridan, and B.P. Feuston. 2003. Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. J. Chem. Inf. Comput. Sci.. doi:10.1021/ci034160g.

Wehrens, R., H. Putter, and L.M.C. Buydens. 2000. The bootstrap: A tutorial. Chemom. Intell. Lab. Syst.. doi:10.1016/S0169-7439(00)00102-7.

Chapter 5 - Prediction of Emulsion Characteristics of Process Cheese Food Using Fluorescence Spectroscopy Combined with Chemometrics

Abstract

The end-use applications of process cheese are highly dependent on its functionality. The functional properties including melted and unmelted characteristics are influenced by the natural cheese and processing conditions. However, the consistency in formulation of process cheese food is a challenge due to availability of raw ingredients, resulting in variations in the final quality. A rapid detection method, therefore, is needed in assessing the important emulsion characteristics during production, in order to make rational decisions on the readiness of the final process cheese food. Previously, fluorescence spectroscopy has been proven to be a useful tool in predicting the functionality, such as melting temperature. In addition, near infrared (NIR) spectroscopy has been applied to predict the properties of cheese, such as fat droplet size distribution. Hence, the objective of this study was to develop calibration models for predicting the emulsion characteristics of process cheese from fluorescence spectra of tryptophan and vitamin A, as well as short-wave NIR spectra using partial least square regression (PLSR). A factorial design (2×2×2) was followed with cheese age, mixing speed, and holding time as independent variables. Each treatment combination was produced 5 times at a lab-scale using a Thermomix^{TM} to obtain a total of 40 process cheese food samples. To characterize the functionality of process cheese food samples, texture profile analysis (TPA), dynamic stress rheology (DSR), and particle size analysis were performed. Spectra were collected from the process cheese by front-face fluorescence spectrometer and hand-held short wave NIR SCIO^{TM}.

Results from analysis of variance (ANOVA) and principle component analysis (PCA) showed a significant difference ($p < 0.05$) in emulsion properties among samples. Calibration models were developed based on pre-processed fluorescence and NIR spectra using PLSR and were validated by leave-one-out method. The model performance was measured in terms of correlation coefficient (R^2) and root mean square error (RMSE). Overall, NIR spectra provided the best results for hardness and particle size parameters. The highest R^2 at 0.73 for predicting the hardness was observed from NIR spectra preprocessed by the Savitzky-Golay first derivative. For a future study, commercial or pilot plant manufactured process cheese could be used to reduce between-batch variations and better represent the industrial settings. More samples should be included to increase the statistical power.

Introduction

Process cheese is a homogenous mixture of natural cheese, emulsifying agent, water, salt, and other ingredients achieved by heating the mixture under a constant shear. Due to its versatile forms in the supermarket, process cheese has been leading the cheese product category ahead of cheddar and mozzarella chesses (IDFA, 2006). A recent study found that in 2015, 70% Americans consume process cheese on a daily basis (Talbot-Walsh, Kannar, and Selomulya, 2018). As an ingredient in other food products, the specific usage of process cheese is influenced by the functional properties (melt and unmelt) which are in turn influenced by raw ingredients, formulation, and manufacturing conditions. However, the availability of natural cheese as well as introduction of re-work cheese can necessitate altering the processing parameters in order to produce process cheese of consistent functionality (Tamime, 2011). The functional properties can be assessed at melted and unmelted status and are traditionally measured by analytical

instruments such as a texture analyzer for texture profile analysis (TPA), dynamic stress rheometer (DSR) for melt temperature, rapid visco analyzer (RVA) for melt viscosity, and other microscopic methods. These attributes include hardness, adhesiveness, particle size, melting temperature, flowability, etc. Overall, measurements of these properties can indicate if the process cheese is homogenous to achieve the targeted functionality. A well-emulsified process cheese should have a higher hardness, smaller fat droplet size, and good meltability given the same starting materials (Savello et al., 1989). Unfortunately, the techniques employed for characterizing the emulsion status usually take a long time to perform, and thus cannot inform process cheese manufacturer about the quality of final product.

Recently, fluorescence spectroscopy is gaining popularity in monitoring molecular interactions between the major food components as well as tracking structural changes in dairy foods (Arroyo-Maya et al., 2016; Panikuttira et al., 2020). Due to the shift in intensity and peak regions of tryptophan residues under environments with different hydrophobicity, fluorescence spectroscopy has been applied for cheese products to predict their functional properties (Purma et al., 2005). Moreover, vitamin A fluorescence excitation spectra have been used to discriminate rheological properties (Dufour and Riaublanc, 1997; Karoui and Dufour, 2003; Kulmyrzaev et al., 2005). In addition to fluorescence, multiple studies have applied near infrared (NIR) spectroscopy in the field of dairy science, including monitoring of cheese coagulation (Lindgaard et al., 2012; Nicolau et al., 2015) and syneresis of coagulum after cheese cutting (Fagan et al., 2008), suggesting the potential to predict other functional properties related to molecular interactions. Compared to long-wave NIR (1100-2500 nm), short-wave NIR (700-1100 nm) can penetrate deeper into the sample and avoid water interference (Ma et al., 2019), which are particularly advantageous for solid samples with high water content, such as process cheese.

To obtain meaningful information by analyzing multiple variables simultaneously in a complex dataset such as spectral data, chemometrics, also known as multivariate statistical analysis, is necessary. This field of study contains tools that can reduce the dimensionality of data for more straight forward interpretation as well as provide prediction or classification on the reference values obtained from analytical measurements.

The objective of this study was to extend the efforts to predict more comprehensive aspects of process cheese emulsions using fluorescence- and NIR-based techniques. By shortening the measurement time and reducing the required labor, this study is intended to provide a useful tool for the industry to implement during process cheese manufacturing.

Materials and Methods

Process cheese samples

Process cheese samples (n=40) were manufactured on-site at lab-scale using a Thermomix (M5; Vorwerk, Thousand Oaks, CA, USA) based on the formulation given in Table 5.1. Butter and salt were purchased from a local supermarket (Walmart, Manhattan, KS, USA) while nonfat dry milk (NDM) was obtained from a commercial manufacturer. Mild and sharp cheddar cheese were provided by AMPI (Associated Milk Producers Inc., New Ulm, MN, USA). The proximate composition except protein was provided by AMPI. The total protein was measured by the Kjeldahl nitrogen method (AOAC International. 991.20). The intact casein content was analyzed as described by Ma et al. (2019).

Table 5.1. Formulation for process cheese production

	Mild		Sharp	
Ingredient	%	weight (g)	%	weight (g)
Natural Cheese	70.00	525.00	70.00	525.00
NDM	8.98	67.35	7.23	54.23
Butter	2.38	17.85	3.25	24.38
Salt	0.50	3.75	0.50	3.75
Disodium Phosphate	2.50	18.75	2.50	18.75
Water	15.64	117.30	16.52	123.90
Total	100	750	100	750

The formula for producing process cheese samples is provided in Table 5.1. It was designed to target a final composition of process cheese samples summarized in Table 5.2, in order to fit the criteria of process cheese food (PCF), which require moisture lower than 44%, fat higher than 23%, and pH exceeding 5.0 (21CFR133.169).

Table 5.2 Targeted composition of process cheese samples made from mild cheddar and sharp cheddar, calculated from ingredients' compositional analysis

	Mild (%)	Sharp (%)
Moisture	43.00	43.00
Fat	25.00	25.00
Total protein	18.01	18.35
Casein protein	15.76	15.03
Intact casein	87.50	81.88
Salt	1.68	1.58

A total of three factors with two levels each were used to design the formula as well as processing parameters combined to produce process cheese to achieve various degrees of emulsification. Three independent variables were the age of natural cheese (mild or sharp), mixing speed (speed 3 or 4 in Thermomix^{TM}), and holding time after the target cooking temperature was reached (30 or 180 seconds). Samples of process cheese were coded as shown in Table 5.3 for data organization.

Table 5.3. Sample code as (cheese age)-(mixing speed)-(holding time) (M = Mild cheddar, S = Sharp cheddar)

Sample No.	Sample code
Sample 1	M-3-30
Sample 2	M-3-180
Sample 3	M-4-30
Sample 4	M-4-180
Sample 5	S-3-30
Sample 6	S-3-180
Sample 7	S-4-30
Sample 8	S-4-180

The procedures for bench-top production of process cheese in a Thermomix are detailed as following. All ingredients were weighed according to the recipe before production (Table 5.1). First, butter was melted at the initial temperature of 40 °C in the Thermomix^{TM} at speed 3 for 10 seconds. Then, natural cheese was added to the container and shredded at speed 5 for 10 seconds. All dry ingredients were added subsequently and mixed at speed 5 for an additional 90 seconds to form a homogenous mixture. Finally, water was added into the mixture and temperature was set to 100 °C with mixing speed fixed at either 3 or 4 depending on the sample code. The mixture was cooked in the Thermomix while temperature was checked regularly during the heating process until it reached 80 °C. At that point, cook temperature was reduced from 100 °C to 80 °C and process cheese was held by continuous mixing at the same speed for either 30 seconds or 180 seconds based on the experimental design. Once the cooking was done, process cheese was evenly poured into two aluminum pans (14.2 cm × 11.5 cm × 4.3 cm) uncovered at room temperature (22 ± 1°C) for one hour before covered being with a lid and stored in a refrigerator at 4 °C. No additional water was added to samples as the water loss during cooking

was calculated to be less than 1%. After storing for 24 hours, one pan was vacuum sealed for future analysis and all analytical measurements were performed on the samples from the second pan.

Analysis of functional properties of process cheese food

Texture profile analysis. The process cheese was stored at 4 °C prior to the sampling. A 10 mm diameter cork borer was used to obtain a cylinder of sample from the block. The cylindrical process cheese sample was removed carefully from the cork borer and trimmed on both the sides to obtain a cylinder of 15 mm height. Triplicate samples were prepared from each process cheese block and immediately wrapped with cling wrap to prevent moisture loss. The samples were tempered at 22 °C for one hour until further analysis. The uniaxial double bite compression test was performed using a Stable Micro Systems Model TAXT2 Textural Profile Analyzer (Texture Technologies Corp, Hamilton, MA) with a 30 kg load cell. The test was based on the following conditions: 50 mm diameter cylinder probe (TA-25); strain, 80%; test speed 0.8 mm/s (Kapoor et al., 2004). TPA hardness and adhesiveness were analyzed according to the method described by Breene (1975).

Dynamic rheological analysis. This method was modified from a previous publication (Kommineni et al., 2012). The Anton Paar rheosystem (MCR 92; Anton Paar, Ashland, VA, USA) was used to perform the dynamic rheological analysis. A cylinder of process cheese was sampled from the block using 4 cm diameter cork borer and then a wire cutter was used to slice the sample into 2 mm thick disks. The process cheese disc was placed on the sample loading plate of the rheometer. The top plate was glued with 400-grit fine sandpaper to prevent the slippage of sample. Then the top plate was lowered to contact the sample until the gap distance between the two plates reached 2.0 mm. The rest of sample on the edges was trimmed and the

side was covered with vegetable oil to prevent moisture loss. The temperature was tempered at 20°C for 80 seconds before the initiation of a sweep test at a frequency of 1.0 Hz and stress of 400 Pa, which gradually increased the temperature from 20 to 90 °C at the rate of 1 °C/min. Measurements of elastic modulus (G'), viscous modulus (G") and phase degree (tan δ) were recorded and repeated for three times on each sample.

Fat droplet size. A 50 mL solution of 0.375% w/w ethylenediaminetetraacetic acid (EDTA disodium salt; Fisher Scientific, Pittsburgh, PA, USA) and 0.125 v/v Tween 20 (Tween^{TM} 20; Fisher Scientific, Pittsburgh, PA, USA) was adjusted to pH 10 using 1M sodium hydroxide. A representative sample of 0.5 g process cheese was dissolved in the solution to denature the protein and to eliminate the interference for the particle size analyzer (Lee et al., 2004). Samples were stored overnight at 4 °C. Before injection into the Beckman Coulter dynamic light scattering analyzer (DelsaMax Pro; Beckman Coulter, Indianapolis, IN, USA), the samples were tempered at room temperature (22 °C) for one hour. The fat droplet size was recorded as the average radius, D10, D50, and D90 in μm. The D10, D50, D90 diameters describes the diameter at the 10%, 50% and 90% cut-off of distribution for particle sizes from small to large.

Front-face fluorescence spectroscopy analysis

Emission spectra of tryptophan. The process cheese blocks at 4 °C were punched by a cork borer of 17 mm diameter and sliced to 5mm thickness to form 17 mm × 5 mm discs. The discs were covered with cling warp and tempered at room temperature (22 °C) before measurement. Then each disc was inserted into the powder cell with a quartz window and sealed with a screw. Once loaded with sample, the cell was held in a front face accessory attached in the LS55 fluorescence spectrometer. For each sample, 5 replicates were recorded at the fluorescence

spectra of tryptophan (λ_{Ex} at 290 nm; λ_{Em} 305-400 nm). The excitation slit width and emission slit width were 10 nm and 6 nm with the scanning speed set at 300 nm/min. The 1% attenuation filter was applied to the emission light. Spectral data from five replicates were combined and averaged on RStudio (version 1.2.1335; RStudio, Boston, MA) for further analysis.

Excitation spectra of vitamin A. Vitamin A provides insights into protein-lipid interaction at the molecular level. Vitamin A front-face excitation spectra (λ_{Em} at 410 nm; λ_{Ex} 250-350 nm) were collected at 22°C following the same sampling process. The emission and excitation slit width was set at 5 nm, the scan speed was set at 300 nm/min.

Near infrared (NIR) analysis

The short-wave NIR regions were collected by SCIO^{TM} , which is a handheld portable NIR spectrometer developed by Consumer Physics (Hod HaSharon, Israel). The process cheese was cut into five 3 cm × 5 cm rectangles for each sample. The samples were tempered at room temperature (25°C) for one hour before the measurement. The device was combined with a shade accessory that was directly placed on top of process cheese to restrict external light as well as to maintain the same distance from the lens and the subject. Random fluctuation known as instrumental error were unavoidable during measurements, so the collected spectra were averaged to reduce the noise before further analysis.

Statistical analysis

Three-way analysis of variance (ANOVA) was performed to compare the means of each response variable cross-classified by factor variables. Before the analysis, the normality test was checked using the Shapiro-wilk test and homogeneity of variance between groups was computed through Levene's test (data not shown).

Furthermore, principle component analysis (PCA) biplots were constructed from total attributes, fluorescence tryptophan emission spectra, vitamin A excitation spectra, and NIR spectra. By reducing the high dimensionality of spectra data, PCA allowed visualization of sample distribution in only two dimensions. The variability explained by each principle component (PC) was displayed on the plot. Variable plot was also overlaid with the individual plot for attributes PCA in order to see the relationship between each PC and emulsion attributes.

Spectral preprocessing and development of calibration models

Spectra data obtained from the fluorescence spectrometer and SCIO™ were first preprocessed using combination of various techniques, such as Savitzky-Golay smoothing (SG-S), first derivative (SG-1st) with 9-neighboring points (Savitzky and Golay, 1964), as well as zero minimum normalization as shown in Equation 5. These tools were able to reduce the innate noise from measurements and accentuate the features, and therefore improve the model performance.

$$X_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \quad (5)$$

Building upon the preprocessed spectra, partial least square regression (PLSR) was applied to develop a linear relationship between spectra data matrix and each emulsion attribute. PLSR is a useful regression model that find the best-fit linear relationship between the predictor and response variables by selecting the latent variables from both matrices. In this study, only one response variable was used at each time and the predictor variables were pre-processed fluorescence tryptophan emission spectra, vitamin A excitation spectra, and NIR spectra in each PLSR model.

Model performance was optimized based on the values of correlation coefficient (R^2) as well as root mean square error (RMSE; Equation 6).

$$RMSE = \sqrt{\frac{\sum(y_{pred} - y_{ref})^2}{N}} \quad (6)$$

Results and Discussion

Compositional analysis

As provided by the manufacturer, the mild cheddar cheese composition had 38.40% moisture, 33.00% fat, 1.68% salt and a pH of 5.13 while the sharp cheddar cheese contained 36.90% moisture, 32.00% fat, 1.54% salt, and pH of 5.3. The total protein compositions of natural cheese samples were 21.42% and 22.75% for mild and sharp cheddar, respectively. The mild cheddar had a slightly higher intact casein content (88.94%) than sharp cheddar (81.88%) as expressed in IC content percentage in total protein, suggesting a level of proteolysis occurred during the aging.

Fluorescence spectra of process cheese

The sample fluorescence spectra for tryptophan and vitamin A were displayed in Figures 5.1 and 5.2, respectively. In Figure 5.1, the fast-mixed sample represented by red had λ_{max} (wavelength of maximum intensity) at 341.5 nm, lower than λ_{max} (343nm) for slow-mixed sample shown as black line. The shift towards smaller λ_{max} indicated the samples was located in a less polar environment, suggesting trytophan residuals were surrounded by more fat dispersion (Purma et al., 2005). This phenomenon of λ_{max} deviating away from 350 nm is also known as the blue shift, which commonly occurs in environment with increased hydrophobicity (Lakowic, 1983). However, moving to a more nonpolar environment would also lower I_{max} (maximum intensity). In constrast, larger peak intensity was observed in well-mixed samples (red line), suggesting otherwise. The conflicting behavior of I_{max} and λ_{max} change could only be better

discerned and investigated through the application of preprocessing tools, which can enhance these differences in spectra to aid in the multivariate analysis.

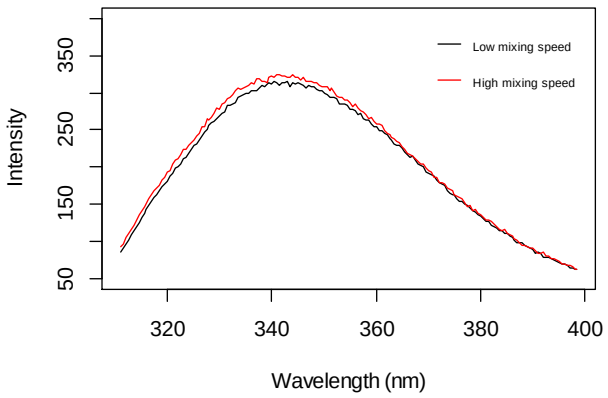

Figure 5.1. Representative tryptophan fluorescence emission spectra for two process cheese samples made by high and low mixing speed using the same cheese age and holding time

The sample vitamin A fluorescence spectra is shown in Figure 5.2. Two differnet peaks were observed at 300 nm and 320 nm. Samples made by varying mixing speed showed the maximum difference in excitation wavelength of 320 nm, and various intensity discrepency across the 250-300 nm regions. Previous studies have estabalished the potential of vitamin A spectra to investigate the physical state of triglycerides (Dufour et al., 2000) as well as the interaction of fat globules within the protein network (Herbert et al., 2000). Yet, the correlation of vitamin A spectral change to those microstructural variations were not specified. For this study, the emulsification level can alter the fat globle dispersion within the protein network, which could induce a shift in vitamin A intensity that can be captured by chemometrics.

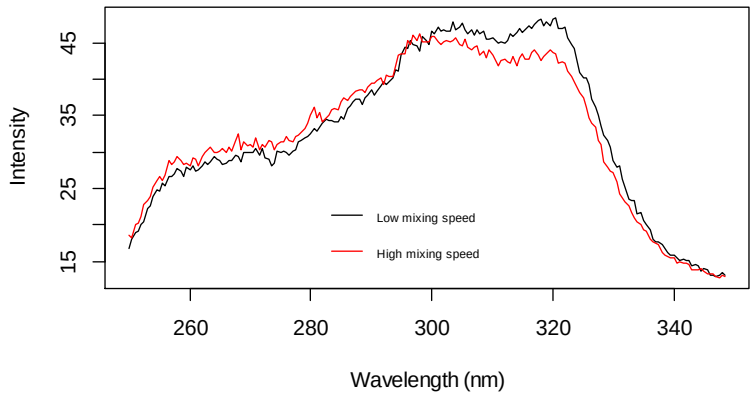

Figure 5.2. Representative fluorescence vitamin A excitation spectra for two process cheese samples made by high and low mixing speed using same cheese age and holding time

NIR spectra of process cheese

The raw spectra of two process cheese samples made using different mixing speeds are displayed in the graph A of Figure 5.3. This short-wave NIR region (740-1070 nm) is primarily the overlay of vibrations from different chemical bonds such as C-H, O-H, and N-H (Wu et al., 2008). Spectral variations were visible across the entire region. Specifically, fast-mixed samples had higher reflectance between 740-820 nm and lower reflectance between 850-1070 nm. Mathematical transformation is necessary for decomposing spectral overlay and assigning spectral variation to different functional groups. In this study, Savitzky-Golay second derivative was applied and resulting spectra are shown in the graph B in Figure 5.3. The spectral differences were located at around 820 nm, 900 nm, 965 nm, 975 nm, and 1000 nm. Corresponding to these findings, multiple wavelength regions related to this study have been reported for molecular bonds associated with major food components including carbohydrate (835 nm), protein (904, 987, 1002 nm), fat (992, 1018 nm), and water (968 nm) (Wu et al., 2008). The interaction between protein, fat, and water can significantly alter the microstructure

of process cheese, and thus result in varying functional properties. For this reason, this study applied quantitative models coupled with feature selection to minimize the number of wavelengths to only important ones, and correlate the intensity of these selected features to functional properties of process cheese.

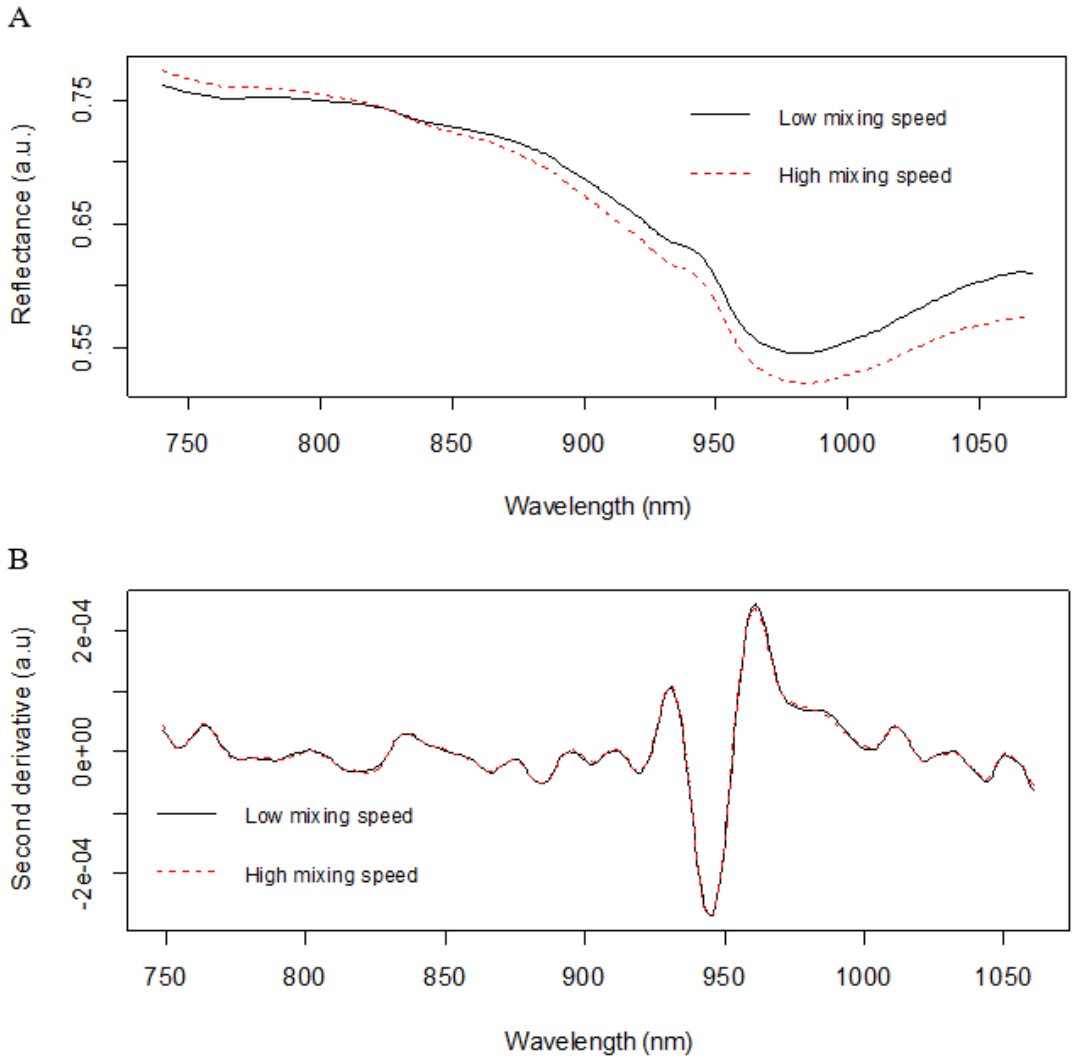

Figure 5.3 Representative raw spectra of NIR regions (740-1070 nm) for two process cheese samples made by high and low mixing speed with other variables fixed (A);Savitzky-Golay second derivative of NIR scans for the same samples (B)

Functional properties

Textural profile analysis

Hardness and adhesiveness are important textural attributes for process cheese. Different degrees of unmelted properties are expected from process cheese used in different food products (Kapoor and Metzger, 2008). A typical TPA graph is exhibited in Figure 5.4. The positive force at the first peak was considered as hardness while the area of the first dough (N×s) was analyzed as adhesiveness. Compared to the solid black line in the figure, the red dotted line represents a process cheese sample produced using higher mixing speed while maintaining other variables as constant. Aligned with observations by Rayan (1980), higher mixing speed result in finer particle size and higher level of emulsion, which correspond to firmer (high hardness) and more elastic behavior upon the application of force. Thus, the degree of emulsification can be directedly related to the hardness as well as adhesiveness.

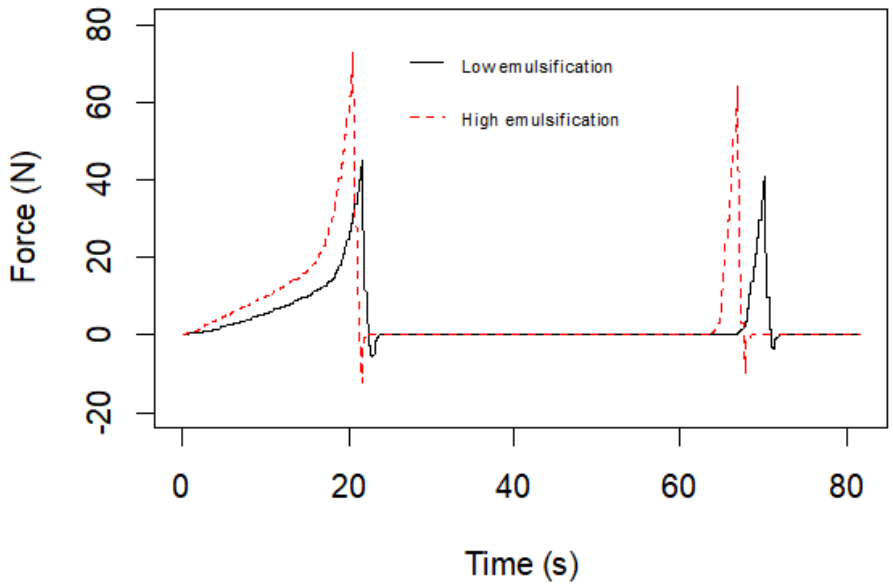

Figure 5.4 Sample TPA graph for process cheese with varying degrees of emulsification

The effects of cheese age, mixing speed, and holding time on the textural profile of process cheese was summarized in Figure 5.5. Multiple studies have shown that natural cheese

selection has considerable influence on the functionality of process cheese (Caric et al., 1985; Shimp, 1985; Zehren and Nusbaum, 2000). This is particularly true given that the amount of natural cheese used in the formulation could vary up to 80% for different process cheese products (FDA, 2006). Cheese age appeared to have little impact on hardness when the mixing speed was high, and yet sharp cheddar tended to have lower hardness at low mixing speed. This could suggest, if the mixing is not complete enough to ensure a well-emulsified structure, lower amount of intact casein in sharp cheddar, inducing the weaker protein interactions causing the low hardness. Compared to hardness, adhesiveness showed greater variation. It can also be concluded from the boxplots (Figure 5.5) that between batch variations were large given the range of each variable. Overall, mixing at higher speed would yield higher hardness, due to the stronger protein-lipid network established from more thorough mixing.

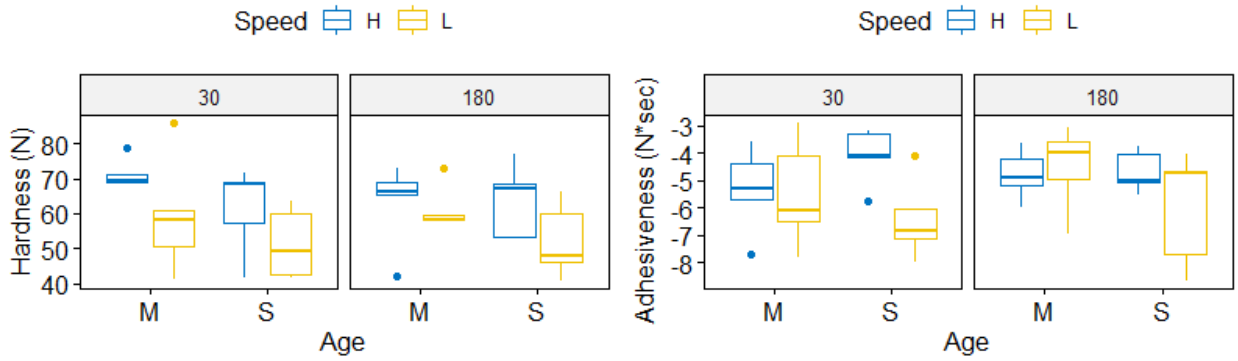

Figure 5.5. Boxplots for hardness and adhesiveness variables of process cheese samples differentiated by three processing parameters (M = mild cheddar, S = sharp cheddar, H = high, L = low, 30 = 30 second holding time, 180 = 180 second holding time)

Rheology profile

Storage modulus (G') and loss modulus (G") are measurements of elastic and viscous properties respectively. They have been proven to be associated with hardness (Drake et al.,

1999). Thus, it was observed from Figure 5.6 that both G' and G" decreased with increasing temperature as a result of thermal softening. Moreover, fast-mixed process cheese tended to have a higher storage and loss modulus.

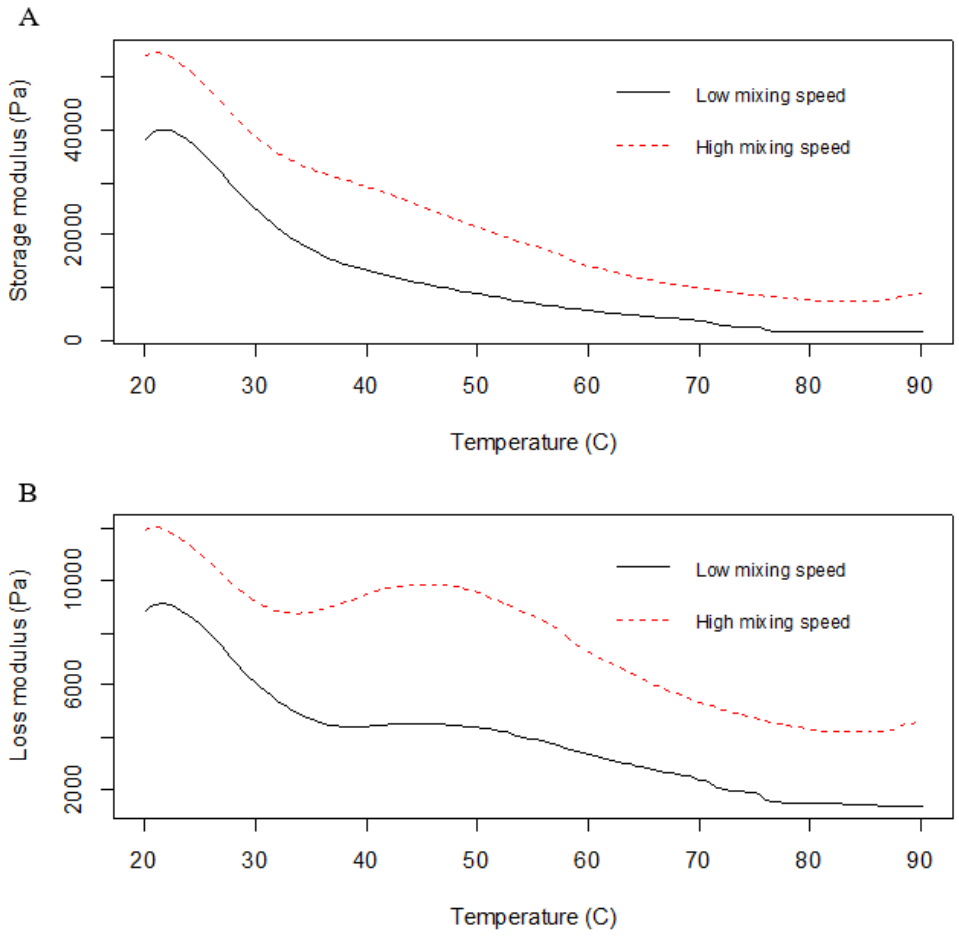

Figure 5.6. Sample DSR G' curves for two process cheese samples made by high and low mixing speed with other variables fixed (A); DSR G" curves for the same samples (B)

Based on Figure 5.7, the sample average demonstrated the same trend as the above graph. Compared to mixing speed, cheese age and holding time appeared to have mixed effects, presumably masked by the between-batch variation.

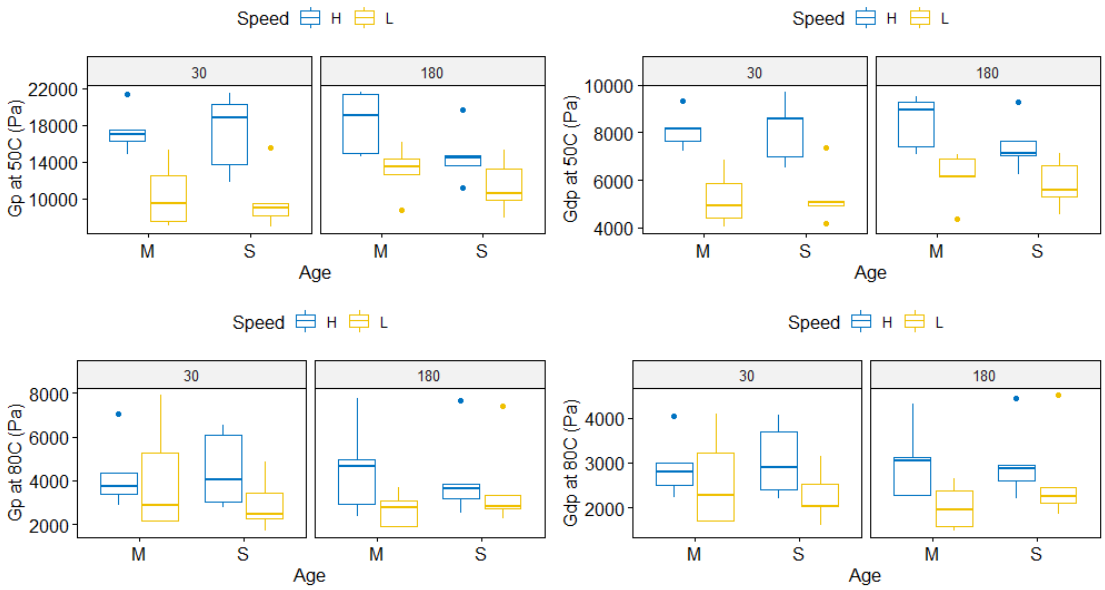

Figure 5.7. Boxplots for rheological variables of process cheese samples differentiated by three processing parameters (rheological variables: Gp50 = G' at 50 °C, Gdp50 = G" at 50 °C, Gp80 = G' at 80 °C, Gdp80 = G" at 80 °C; three parameters: M = mild cheddar, S = sharp cheddar, H = high, L = low, 30 = 30 second holding time, 180 = 180 second holding time)

Particle size

Fat droplet size is a good indiactor of emulsification degree of process cheese (Rayan, 1980). Small average particle size means the cheese is well-mixed, and thus fat globules are more dispersed around the protein to provide a stronger interaction (Kapoor and Metzger, 2008). In general, the higher mixing speed led to smaller fat droplet size as shown in Figure 5.8. Some outliers were represented as dots in the plot, possibly due to variations among batches and different sampling locations for less-emulsified cheese. It was also noticed that sharp cheddar tended to produce smaller fat droplets. This could be caused by natural lipolysis during the cheese aging.

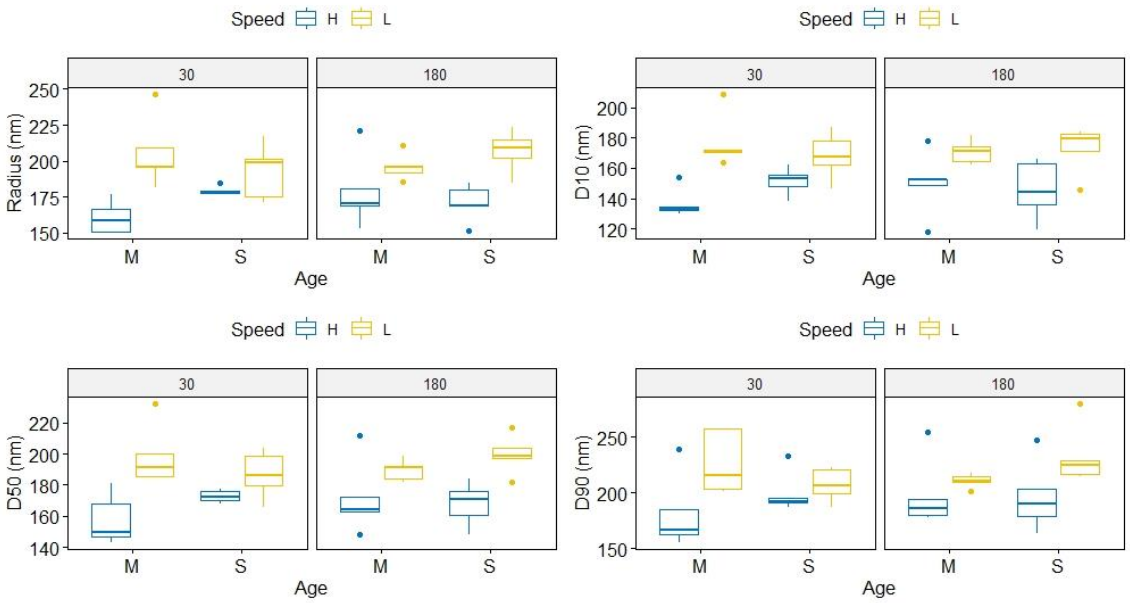

Figure 5.8. Boxplots for particle size variables of process cheese samples differentiated by three processing parameters(particle size variables: radius, D10, D50, D90; three processing parameters:M = mild cheddar, S = sharp cheddar, H = high, L = low, 30 = 30 second holding time, 180 = 180 second holding time)

Table 5.4. Summary of p-values from ANOVA table for the impact of three processing parameters (age = natural cheese age, speed = mixing speed, time = holding time) on emulsion attributes across all process cheese samples (* = 0.1%, ** = 1%, * = 5% statistical significance)**

	Hardness	Adhesiveness	Radius	D10	D90	G' at 50°C	G" at 50°C	G' at 80°C	G" at 80°C
Age	0.068	0.579	0.703	0.837	0.572	0.150	0.417	0.933	0.684
Speed	0.015 *	0.074	9.29e-06 ***	6.15e-06 ***	0.004 ***	3.32e-06 ***	1.04e-07 ***	0.087	0.023 *
Time	0.786	0.452	0.509	0.786	0.488	0.531	0.426	0.766	0.876
Age:Speed	0.588	0.057	0.561	0.364	0.678	0.720	0.554	0.977	0.886
Age:Time	0.509	0.386	0.893	0.655	0.500	0.277	0.340	0.501	0.443
Speed:Time	0.524	0.437	0.821	0.614	0.915	0.164	0.229	0.805	0.795
Age:Speed:Time	0.408	0.739	0.024 *	0.141	0.070	0.583	0.698	0.244	0.296

To further understand each manufacturing parameter on the final properties of process cheese, ANOVA was conducted. Prior to the analysis, all attributes variables passed the homogeneity of variance, among which particle size, D90, G' and G" at 80 °C did not meet the normality assumption. However, given the large sample size (n>30), this failure could be relieved. The ANOVA result from all variables is summarized in Table 5.4. From the table, the significance of each factor variable can be compared in terms of its effectiveness on the response variable. It can be concluded that mixing speed had a significant effect on most attributes, such as hardness, radius, D10, D90, G' and G" at 50 °C and G" at 80 °C. The difference between 30 second and 180 second could be too little to have an impact on the final properties because holding time made no statistical difference on all emulsion attributes. Although the results were not significant at 5% level, the cheese age did influence hardness and G' at 50°C at 6.8% and 15% level respectively, which supported claims by previous studies (Shimp, 1985; Zehren and Nusbaum, 2000). Overall, the ANOVA confirmed that significant differences existed between most emulsion characteristics of samples made using different speed as illustrated earlier. To visualize the difference, PCA was conducted.

PCA

PCA plots from objects and attributes variables were overlaid in Figure 5.9. In combination, PC1 and PC2 explained 64.2% of total variance between samples. Based on the arrow direction, particle size was oppositely related to hardness, adhesiveness, G' and G". Higher hardness, adhesiveness and smaller particle size were observed from well-emulsified process cheese. PC1 can be classified as the indicator for emulsification level. Lower scores represent better emulsification between protein and lipids.

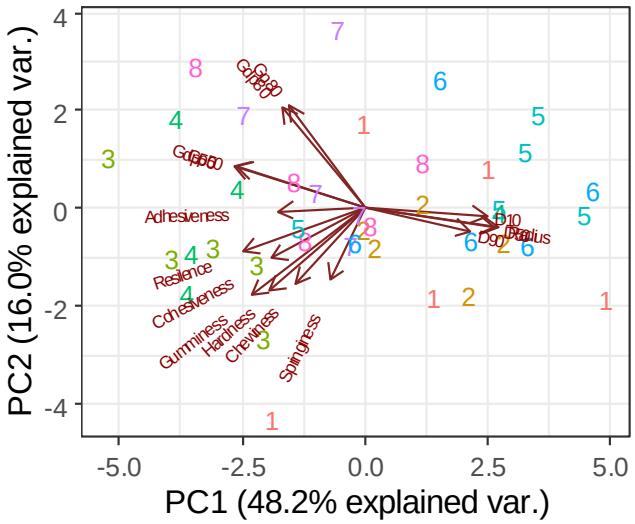

Figure 5.9. PCA individual and variable plot for emulsion attributes across all process cheese samples (sample-specific processing variables are as following: mild cheddar = 1,2,3,4 and sharp cheddar = 5,6,7,8; slow speed = 1,2,5,6 and fast speed = 3,4,7,8; short holding time = 1,3,5,7 and long holding time = 2,4,6,8)

In addition to the attributes, raw tryptophan, vitamin A and NIR spectra were also used to plot PCA in order to check if samples with various emulsification level exhibited significant difference in fluorescence as well as NIR signals. Unlike previous plots, PCA plots from spectra data were less sparse, and variability was mostly explained by a single PC. This was expected since in these spectra, one or two signal peaks were observed to differentiate between samples. From tryptophan PCA plot (Figure 5.10), it can be concluded that PC2 corresponded to the mixing speed, as samples 3,4,7,8, which were produced from high mixing speed, were clustered at the positive end of PC2. For these samples, PC1 accounted for the cheese age. Mild cheddar samples 3,4 were grouped at the positive side of PC1 while sharp cheddar samples 7,8 were on the opposite. Cheeses made with lower mixing speed 1,2,5,6 were scattered across PC1, suggesting various degree of emulsification, presumably caused by sampling locations.

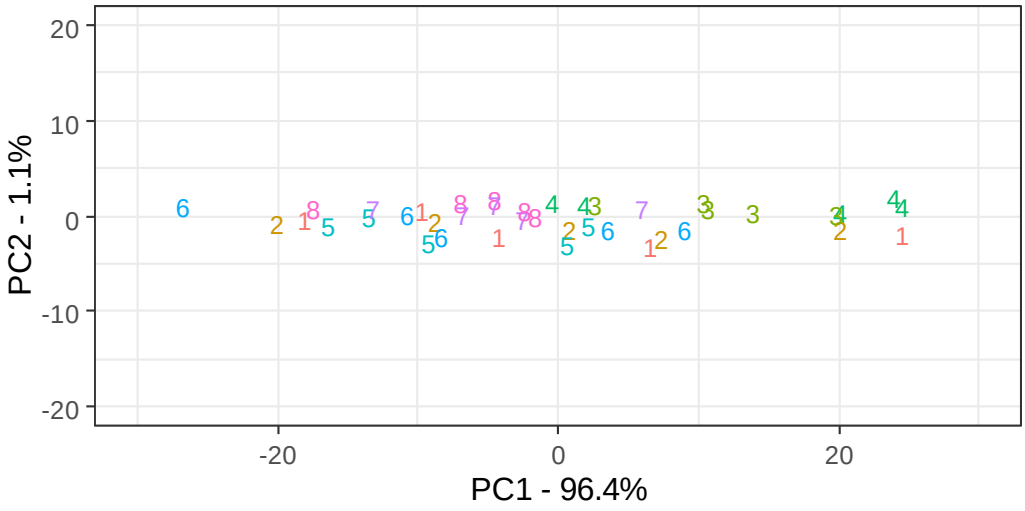

Figure 5.10. PCA individual plot for fluorescence tryptophan emission spectra across all process cheese samples (sample-specific processing variables are as following: mild cheddar = 1,2,3,4 and sharp cheddar = 5,6,7,8; slow speed = 1,2,5,6 and fast speed = 3,4,7,8; short holding time = 1,3,5,7 and long holding time = 2,4,6,8)

Samples' vitamin A PCA plot (Figure 5.11) was roughly segmented by cheese age. Mild cheddar was clustered at lower right corner while sharp cheddar was grouped at the opposite end. This was because vitamin A was correlated to cheese aging. As natural cheese stored for a longer time, vitamin A was photooxidized into forms that no longer emit fluorescence. Previous studies have used vitamin A fluorescence spectra to monitor the cheese storage time (Andersen, Vishart, and Holm, 2005; Christensen, Povlsen, and Sørensen, 2003). From the plot, it can be observed that vitamin A spectra could not be used to group samples based on the emulsification level indicated in the attributes plot. However, in the NIR spectra PCA plot, samples 3,4,7,8 were grouped on the positive side of PC2, suggesting the spectra could be useful in differentiating the emulsion system among process cheese samples. It was also noticed that due to poor mixing, samples 1,2 have two outliers while the rest were scattered across the PC1.

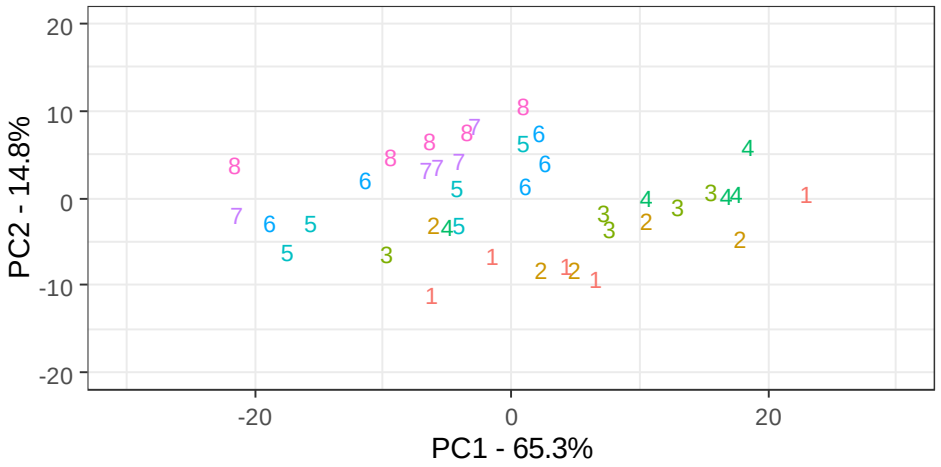

Figure 5.11. PCA individual plot for fluorescence vitamin A excitation spectra across all process cheese samples (sample-specific processing variables are as following: mild cheddar = 1,2,3,4 and sharp cheddar = 5,6,7,8; slow speed = 1,2,5,6 and fast speed = 3,4,7,8; short holding time = 1,3,5,7 and long holding time = 2,4,6,8)

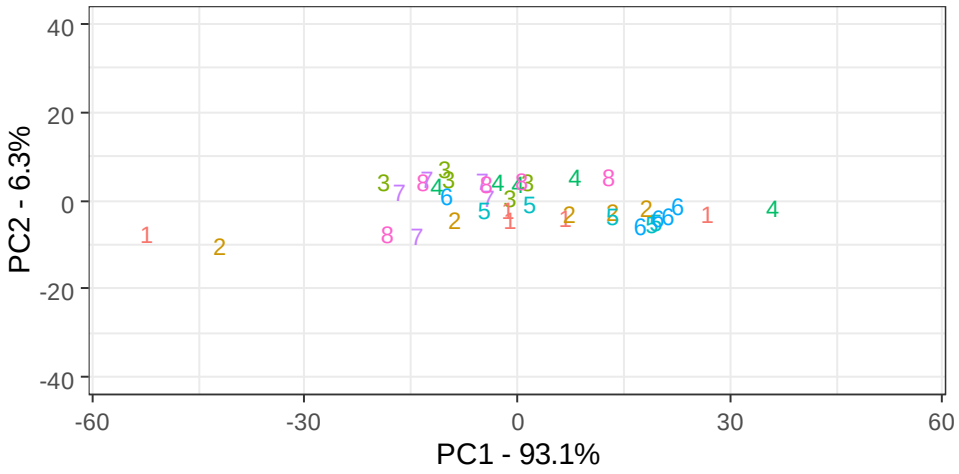

Figure 5.12. PCA individual plot for NIR spectra across all process cheese samples (sample-specific processing variables are as following: mild cheddar = 1,2,3,4 and sharp cheddar = 5,6,7,8; slow speed = 1,2,5,6 and fast speed = 3,4,7,8; short holding time = 1,3,5,7 and long holding time = 2,4,6,8)

Calibration models

The calibration results are summarized in Table 5.5. Overall, NIR spectra showed the highest correlation coefficient for hardness, radius, D10, and D50, which were 0.725, 0.461, 0.546, and 0.512 respectively. The correlations are weak and thus the model is not accurate enough for analytical purposes. In the past, long wave NIR (1100-2500 nm) has been used to predict the total particle size and distribution of Feret diameter in cheese (Alinovi et al., 2019) with good correlation coefficients between 0.504 and 0.828. A similar study also performed on using short-wave NIR to predict wheat flour particle size with R^2 of 0.55 (Zhu et al., 2017). Both studies had limited success on correlating NIR spectra with microscopic particle size, suggesting that NIR is more suitable for components quantification than the structural relationship between components. Apart from NIR, tryptophan showed the highest R^2 for G' and G" at 50°C. Previously, tryptophan emission spectra have been used to predict melting point obtained from DSR (Purna et al., 2005). The fluorescence signal from tryptophan is highly dependent on the hydrophobicity of the environment, and thus could potentially reflect the emulsion system of process cheese. However, the results from this study showed limited success. This could be due to the lower reproducibility of the Thermomix in making process cheese with constant emulsion attributes given the same processing parameters. Small sample is another limitation that reduced the model accuracy.

Table 5.5. PLSR calibration results using preprocessed fluorescence tryptophan spectra, vitamin A spectra and NIR spectra validated by leave-one-out method (SG-S = Savitzky Golay smoothing, N = Normalization, SG-1st = Savitzky Golay first derivative)

Attributes	Preprocessing	Tryptophan		Vitamin A		NIR	
		R^2	RMSE	R^2	RMSE	R^2	RMSE
Hardness	SG-S	0.32	9.7	0.15	12.0	0.73	6.1
	SG-S + N	0.25	10.1	0.00	12.0	0.72	6.3
	SG-1st	0.20	10.5	NA	NA	0.62	7.3
Adhesiveness	SG-S	0.05	1.5	0.41	1.3	0.06	1.5
	SG-S + N	0.08	1.5	0.27	1.4	0.06	1.5
	SG-1st	0.01	1.6	NA	NA	0.06	1.5
Radius	SG-S	0.14	20.3	0.03	21.7	0.45	16.2
	SG-S + N	0.20	19.5	0.05	21.3	0.45	16.2
	SG-1st	0.09	21.4	NA	NA	0.46	16.9
D10	SG-S	0.19	17.9	0.12	19.0	0.55	13.3
	SG-S + N	0.22	17.4	0.15	18.3	0.54	13.3
	SG-1st	0.23	19.0	NA	NA	0.55	13.2
D50	SG-S	0.19	18.0	0.04	19.7	0.50	14.1
	SG-S + N	0.25	17.2	0.07	19.2	0.50	14.1
	SG-1st	0.13	18.9	NA	NA	0.51	13.9
D90	SG-S	0.07	27.5	0.01	29.4	0.15	26.3
	SG-S + N	0.11	26.5	0.00	28.7	0.11	26.3
	SG-1st	0.07	28.1	NA	NA	0.16	25.9
G’ @50C	SG-S	0.39	3344.1	0.01	4403.0	0.35	3460.0
	SG-S + N	0.44	3197.0	0.01	4332.8	0.35	3469.0
	SG-1st	0.30	3608.0	NA	NA	0.35	3471.0
G” @50C	SG-S	0.43	1211.6	0.06	1578.9	0.47	1174.0
	SG-S + N	0.48	1152.1	0.07	1559.4	0.47	1177.0
	SG-1st	0.31	1334.9	NA	NA	0.47	1172.0
G’ @80C	SG-S	0.01	1846.4	0.19	1892.6	0.01	1835.0
	SG-S + N	0.00	1856.8	0.02	1860.1	0.01	1835.0
	SG-1st	0.00	1836.4	NA	NA	0.02	1899.0
G” @80C	SG-S	0.01	850.2	0.01	845.8	0.02	828.4
	SG-S + N	0.01	833.1	0.01	833.7	0.02	827.6
	SG-1st	0.00	845.1	NA	NA	0.01	842.0

Conclusion

This study validated the calibration models using tryptophan, vitamin A fluorescence spectra and NIR spectra on emulsion characteristics of process cheese using bench top Thermomix^{TM}. The preliminary results showed that significant differences in textures, particle size, and viscoelasticity existed from samples made with different processing parameters. The validation results showed the best correlation from NIR spectra to predict hardness and particle size. Although the accuracy was not high enough for implementation as a routine analytical method, the study provided the preliminary evidence that the use of chemometrics and spectroscopic method could be potentially used to predict the emulsion characteristics of process cheese.

References

Alinovi, M., Mucchetti, G., and Tidona, F. 2019. Application of NIR spectroscopy and image analysis for the characterisation of grated Parmigiano-Reggiano cheese. Int Dairy J, 92, 50-58.

Andersen, C. M., Vishart, M., and Holm, V. K. 2005. Application of fluorescence spectroscopy in the evaluation of light-induced oxidation in cheese. J Agric Food Chem, 53(26), 9985-9992.

Arroyo-Maya, I. J., Campos-Terán, J., Hernández-Arana, A., and McClements, D. J. 2016. Characterization of flavonoid-protein interactions using fluorescence spectroscopy: Binding of pelargonidin to dairy proteins. Food Chem, 213, 431-439.

Breene, W. M. 1975. Application of texture profile analysis to instrumental food texture evaluation. J Texture Stud, 6(1), 53-82.

Caric, M., Gantar, M. and Kalab, M. 1985. Effects of emulsifying agents on the microstructure and other characteristics of process cheese-a review. Food Struct, 4(2), 13.

Christensen, J., Povlsen, V. T., and Sørensen, J. 2003. Application of fluorescence spectroscopy and chemometrics in the evaluation of processed cheese during storage. J Dairy Sci, 86(4), 1101-1107.

Drake M.A., Gerard P.D., Truong V.D., Daubert C.R. 1999. Relationship between instrumental and sensory measurements of cheese texture. J Texture Stud 30:451–76

Dufour, E., Mazerolles, G., Devaux, M. F., Duboz, G., Duployer, M. H., and Riou, N. M. 2000. Phase transition of triglycerides during semi-hard cheese ripening. Int Dairy J, 10(1-2), 81-93.

Dufour, E., and Riaublanc, A. 1997. Potentiality of spectroscopic methods for the characterisation of dairy products. I. Front-face fluorescence study of raw, heated and homogenised milks. Lait, 77(6), 657-670.

Fagan, C. C., Castillo, M., O'Donnell, C. P., O'Callaghan, D. J., and Payne, F. A. 2008. On-line prediction of cheese making indices using backscatter of near infrared light. Int Dairy J, 18(2), 120-128.

Food and Drug Administration. 2006. 21 CFR, Part 133.169 to 133.180. Food and Drug Administration. Washington, D.C.: Dept. of Health and Human Services.

Herbert, S., Riou, N. M., Devaux, M. F., Riaublanc, A., Bouchet, B., Gallant, D. J., and Dufour, É. 2000. Monitoring the identity and the structure of soft cheeses by fluorescence spectroscopy. Lait, 80(6), 621-634.

IDFA. 2006. Dairy facts. Washington, D.C.: Intl. Dairy Foods Assn.

Karoui, R., and Dufour, É. 2003. Dynamic testing rheology and fluorescence spectroscopy investigations of surface to centre differences in ripened soft cheeses. Int Dairy J, 13(12), 973-985.

Kapoor, R., Lehtola, P., and Metzger, L. E. 2004. Comparison of pilot-scale and rapid visco analyzer process cheese manufacture. J Dairy Sci, 87(9), 2813-2821.

Kapoor, R. and Metzger, L. E. 2008. Process cheese: scientific and technological aspects—a review. Compr Rev Food Sci Food Saf, 7(2), 194-214.

Kommineni, A., J. Amamcharla, and L.E. Metzger. 2012. Effect of xylitol on the functional properties of low-fat process cheese. J. Dairy Sci.. doi:10.3168/jds.2012-5376.

Kulmyrzaev, A., Dufour, E., Noël, Y., Hanafi, M., Karoui, R., Qannari, E. M., and Mazerolles, G. 2005. Investigation at the molecular level of soft cheese quality and ripening by infrared and fluorescence spectroscopies and chemometrics—relationships with rheology properties. Int Dairy J, 15(6-9), 669-678.

Lakowicz, J. R. 1983. Protein fluorescence. In Principles of fluorescence spectroscopy (pp. 341-381). Springer, Boston, MA.

Lyndgaard, C. B., Engelsen, S. B., and van den Berg, F. W. 2012. Real-time modeling of milk coagulation using in-line near infrared spectroscopy. J Food Eng, 108(2), 345-352.

Ma, Y. B., Babu, K. S., and Amamcharla, J. K. 2019. Prediction of total protein and intact casein in cheddar cheese using a low-cost handheld short-wave near-infrared spectrometer. LWT, 109, 319-326.

Mazerolles, G., Devaux, M. F., Duboz, G., Duployer, M. H., Riou, N. M., and Dufour, É. 2001. Infrared and fluorescence spectroscopy for monitoring protein structure and interaction changes during cheese ripening. Lait, 81(4), 509-527.

Nicolau, N., Buffa, M., O'callaghan, D. J., Guamis, B., and Castillo, M. 2015. Estimation of clotting and cutting times in sheep cheese manufacture using NIR light backscatter. Dairy Sci Technol, 95(4), 495-507.

Panikuttira, B., Payne, F. A., O'Shea, N., Tobin, J. T., O'Callaghan, D. J., and O'Donnell, C. P. 2020. Investigation of an in‐line prototype fluorescence and infrared backscatter sensor to monitor rennet‐induced coagulation of skim milk at different protein concentrations. Int J Food Sci Technol, 55(1), 175-182.

Purna, S. G., Prow, L. A., and Metzger, L. E. 2005. Utilization of front-face fluorescence spectroscopy for analysis of process cheese functionality. J Dairy Sci, 88(2), 470-477.

Rayan, A. H. 1981. Microstructure and rheology of process cheese.

Savello, P. A., Ernstrom, C. A., and Kalab, M. 1989. Microstructure and meltability of model process cheese made with rennet and acid casein. J Dairy Sci, 72(1), 1-11.

Savitzky, A., and M.J.E. Golay. 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36:1627–1639.

Shimp, L. A. 1985. Process Cheese Principles. Food Technology (USA).

Talbot-Walsh, G., Kannar, D., and Selomulya, C. 2018. A review on technological parameters and recent advances in the fortification of processed cheese. Trends Food Sci Technol, 81, 193-202.

Tamime, A. Y. 2011. Processed Cheese and Analogues (Vol. 16). John Wiley and Sons.

Wu, D., He, Y., and Feng, S. 2008. Short-wave near-infrared spectroscopy analysis of major compounds in milk powder and wavelength assignment. Anal Chim Acta, 610(2), 232-242.

Zehren, V. L. and Nusbaum, D. D. 2000. Processed Cheese. Cheese Reporter Publishing Co. Inc., Madison.

Zhu, Q., Xing, Y., Lu, R., Huang, M., and Ng, P. K. 2017. Visible/shortwave near infrared spectroscopy and hyperspectral scattering for determining bulk density and particle size of wheat flour. J Near Infrared Spectrosc, 25(2), 116-126.

Chapter 6 - Conclusions

Fluorescence spectroscopy can identify intrinsic fluorescent compounds commonly found in dairy products, and provide spectra that can be analyzed by chemometrics to predict or classify components as well as properties associated with the product. The developed models can offer practical solutions to the industry related to process cheese and NDM manufacture. Chapter 4 developed a classification model for different spore levels in NDM. Based on the performance metrics, the results showed that:

- Different levels of thermophilic and mesophilic spore counts can be revealed in fluorescence spectral variation by the addition of $TbCl_3$ to pre-treated NDM.
- Right-angle fluorescence spectroscopy combined with random forest, forward selection quadratic discriminant analysis (FS-QDA), and partial least square discriminant analysis (PLSDA) can classify the NDM at 87%, 84%, and 83% accuracy, as validated by the bootstrapping method.
- The model performance suggests that the method can be used to pre-screen NDM for processing hygiene check.

Chapter 5 applied front-face fluorescence and a portable NIR scanner to process cheese samples with varying emulsification level. The tryptophan and vitamin A fluorescence spectra, as well as short-wave NIR spectra were collected, preprocessed and analyzed by partial least square regression (PLSR) to predict the functional properties related to process cheese. Overall, the results showed that:

- Changing natural cheese age, mixing speed, and holding time can produce process cheese by Thermomix^{TM} with different levels of emulsion properties.

- Mixing speed is the most differentiating factor in the functional properties of process cheese manufactured by Thermomix^{TM}.
- Highly emulsified process cheese samples tended to have higher hardness and elasticity, and smaller fat droplet size.
- PLSR built on NIR spectra had the highest prediction power at R^2 of 0.73 and 0.55 for hardness and D10 respectively.
- The study provided preliminary evidence that fluorescence and NIR spectra can be used to correlate the emulsion characteristics of process cheese.

Appendix A - R code for chapter 4

```
##Load packages
library(plyr)
library(tidyr)
library(dplyr)
library(varhandle)
library(lattice)
library(ggplot2)
library(caret)

##Data loading
temp <- list.files(pattern="*.csv")
class <- read.csv(file="...csv", header = T)

##Data cleaning
NDM<-join_all(lapply(temp, read.csv), "Created.As.New.DataSet" )
NDM<-NDM[-1,]
NDM <- unfactor(NDM)
names(NDM) <- NULL
names <- c("Em",temp)
names(NDM) <- names
NDM <- t(NDM)
data <- cbind(NDM,class)
colnames <- data[1,]
data <- data[-1,]
colnames(data) <- colnames
data <- data.frame(apply(data, 2, function(x)
as.numeric(as.character(x))))
row.names(data) <- temp


##Stepwise quadratic discriminant analysis
maxvar <- (10)
direction <- "forward"
tune <- data.frame(maxvar, direction)

sqda <- train(class~.,
              data = data,
              method = "stepQDA",
              trControl = trainControl(method = "optimism_boot"),
              tuneGrid = tune)

##Random forest
rf <- train(class~.,
            data = data,
            method = "rf",
            trControl = trainControl(method = "optimism_boot"))
```

Appendix B - R code for chapter 5

```
##Load pacakges
library(datarium)
library(magrittr)
library(tidyverse)
library(ggpubr)
library(rstatix)
library(dplyr)


##Data loading
trp<-read.csv(file = "...csv", header = T)
va<-read.csv(file = "...csv", header = T)
att<-read.csv(file = "...csv", header = T)
para<-read.csv(file = "...csv", header = T)
nir<-read.csv(file = "...csv", header = T)


##Boxplot
radius_plot <- ggboxplot(
  data, x = "Age", y = "Radius",
  color = "Speed", palette = "jco", facet.by = "Time"
)

ggarrange(radius_plot, D10_plot, D50_plot, D90_plot, ncol = 2, nrow
=2)
ggarrange(hardness_plot,adh_plot, ncol = 2, nrow =1)
ggarrange(gp50_plot, gdp50_plot, gp80_plot, gdp80_plot, ncol = 2, nrow
=2)

##Check normality assumption
model  <- lm(... ~ Age*Speed*Time, data = ...)
ggqqplot(residuals(model))
shapiro_test(residuals(model))

##Check covariance assumption
levene <- levene_test(... ~ Age*Speed*Time, data = ...)

##3-way anova
res.aov <-anova_test(... ~ Age*Speed*Time,data = ...)

##PCA
pca <- prcomp(att[,-1], scale. = TRUE)

ggbiplot(pca, ellipse = F, obs.scale = 1, var.scale = 1,labels =
att[,1],groups = as.factor(att[,1]), varname.size = 2) +
  theme_bw() +
  theme(legend.position = "none") +
  ggtitle("Attribute PCA graph")
```

```
##Data combination and cleaning
raw<-cbind(att, trp)
rawva<-cbind(att, va)
rawnir<-cbind(att, nir[,-c(1,2)])

##Spectra preprocessing
calsg<-savitzkyGolay(raw[,-c(1:17)], p = 2, w = 19 , m = 1)
calnorm<-data.Normalization(calsg, type = "n4", normalization = "row")

##Model calibration and validation
aa<-diag(1:10, 2:11)
pre_results<-matrix(ncol = 5, nrow = 16)
colnames(pre_results)<-colnames(aa)
for (n in 2:17){
  pls<-plsr(raw[,n] ~ as.matrix(calsg), ncomp = 10, validation =
"LOO")
  pre<-as.data.frame(pls$validation$pred)
  pre_results[(n-1), ]<-diag(pre[,which.min(pls$validation$PRESS)],
raw[,n])
}
rownames(pre_results)<-colnames(raw)[2:17]
pre_results
```