Control of cheese quality –

strategies and methods for delivery of consistent sensory product quality at the point of consumption

Styring av hvitost-kvalitet –

strategier og metoder for å oppnå riktig sensorisk kvalitet på produktene ved forbruk

Philosophiae doctor (PhD) thesis

Hilde Kraggerud

Department of Chemistry, Biotechnology and Food Science
Norwegian University of Life Sciences

Ås 2011

Thesis nr.: 2011:68

ISBN-nr.: 978-82-575-1031-2

ISSN-nr.: 1503-1667

CONTENTS

		Page
Table of contents		
Acknowledgements		
Abstract		
Sammendrag		
List of original papers		
1	Introduction	1
2	A brief theoretical review	3
2.1	Quality control	3
2.2	Sensory quality	7
2.3	Cheese production	9
2.3.1.	Raw milk	10
2.3.2.	Bacterial cultures	12
2.3.3.	Coagulation	12
2.3.4.	Syneresis and post-coagulation	13
2.3.5.	Forming, pressing and salting	14
2.3.6.	Cheese ripening	15
2.4	Methods of analysis	16
2.4.1.	Compositional analysis	17
2.4.2.	Monitoring cheese ripening	17
2.4.3.	Rapid analytical methods	18
2.5.	Mathematical modelling	19
2.5.1.	Multivariate data modelling	19
2.5.2.	Predictive modelling of manufacturing processes	21
3	Main results	21
3.1	Paper 1: Season and ripening temperature influence fatty acid composition and sensory properties of semi-hard cheese during ripening.	22
3.2	Paper 2: X-ray images for the control of eye formation in cheese.	23
3.3	Paper 3: Quality scoring - a tool for sensory evaluation of cheese ?	23
3.4	Paper 4: Prediction of sensory quality of cheese during ripening from chemical and spectroscopy measurements.	24
4	Discussion	25
5	Challenges and future perspectives	27
6	Reference list	30
	Appendix (individually paginated)	
	Paper 1	
	Paper 2	
	Paper 3	
	Paper 4	

Acknowledgements

From the start this work was a part of the IBION project (Industrial BIOstatistics Network), a joint research programme with several partners from industry and academia– with the focus on optimal utilization of raw materials, mathematical modelling and spectroscopic methods of analysis. The project was partly financed by the Norwegian Research Council, through grant number 145456, and partly by the partners from industry, including my employer for 25 years, TINE BA.

I want to thank TINE for giving me the opportunity to spend time on this work. I especially want to thank my manager, Research Director Johanne Brendehaug, for encouraging me to start and to finish this work. This thesis would not have been possible without the help of many people at TINE: everybody employed in the pilot plant at Voll, the sensory panel, the laboratory and the participants in the TINE IBION project. I also want to thank my supervisor Professor Roger Abrahamsen for his patience and his thorough work throughout this 8-year-long process. I also want to thank the other co-authors of the articles: Siv Skeie (UMB), Martin Høy (Nofima), Jens Petter Wold (Nofima), Steffen Solem (Vinmomopolet) and Lars Røkke (TINE) for their cooperation. Edward Hopkin has provided support in improving my English language in the written work.

Last but not least I want to thank my family for their patience and support through this long period – especially my husband for 25 years – Stig, but also my children Sven, Lina and Anna. And thanks to my mother, Louise, for accommodation and sister-in–law, Anne, for transport when I have been working at Ås. I would not have been able to work through all this without you!

Abstract

An effective approach to quality control is an important issue for a food producer, as consumers expect consistent delivery of products. Two important strategies for control of food product end quality was discussed in this thesis: Process regulation and Statistical Process control (SPC). Extensive experiments were carried out in order to demonstrate different aspects of control of cheese quality.

The sensory quality of a product is of great importance, as it is directly perceived at consumption. How do we define, and measure cheese quality? In paper 3 this subject was discussed, and quality scoring was found appropriate as a methodology for sensory quality, provided consumer input in definition of product specifications.

Rapid, particularly non-destructive measurements are important in control strategies. In paper 4 spectroscopic methods were found promising for fast and reliable results. Spectroscopy was found to be able to substitute chemical measurements for the purpose of measuring relevant sensory attributes of cheese. In paper 2 X-ray methodology, found suitable for non-destructive on-line measurements of eye formation in cheese during ripening, was developed.

In statistical process control, SPC, it is essential to understand the influence of all relevant factors from raw material through process to product. In paper 1 the effect of variation through all seasons of the year, as well as the effect of maturation after different ripening temperatures on sensory properties of cheese was examined. This gives us a better platform for adjustments with the aim of variability reduction for the actual cheese varieties.

Sammendrag

En effektiv tilnærming til kvalitetsstyring er viktig for næringsmiddelprodusenter, da forbrukerne forventer levering av produkter med jevn kvalitet. To viktige strategier for styring av matprodukters sluttkvalitet ble diskutert: Prosessregulering, og statistisk prosesskontroll (SPC). Omfattende forsøk ble utført for å vise ulike aspekter av styringen av ostekvalitet.

Den sensoriske kvaliteten til produktene er av stor betydning, da den blir direkte oppfattet ved forbruk. Men hvordan definerer vi og måler ostekvalitet? I artikkel 3 diskuteres dette emnet, og kvalitetsbedømmelse med poeng ble funnet å være en metode som passer for formålet, forutsatt at resultater fra forbrukerundersøkelser brukes som grunnlag for produktspesifikasjonene.

Hurtigmetoder, spesielt ikke-destruktive målinger, er viktige styringsverktøy. I artikkel 4 vurderte man spektroskopiske metoder som lovende for raske og pålitelige analyser av ost. Spektroskopi ble funnet å kunne erstatte kjemiske målinger I forhold til å måle relevante sensoriske egenskaper I ost. I artikkel 2 ble det utviklet en røntgenmetode som passet for måling av hullsetting I ost under modning.

I statistisk prosesskontroll, SPC, er det grunnleggende å ha forståelse for innvirkningen av alle relevante faktorer, fra råmaterialer gjennom prosessen til ferdig ost. I artikkel 1 ble effekter av variasjoner gjennom året, og ulike modnings-temperaturer undersøkt, I forhold til påvirkning på sensorisk kvalitet. Dette gir oss en bedre plattform for justeringer med henblikk på reduksjon av variasjon for de undersøkte norske ostetypene.

List of original papers

Paper no.: 1

Kraggerud, H., Skeie, S., Høy, M., Røkke, L. & Abrahamsen, R.K. (2008)

Season and ripening temperature influence fatty acid composition and sensory properties of semi-hard cheese during ripening.

International Dairy Journal, 18, 801-810.

Paper no.: 2

Kraggerud, H., Wold, J.P., Høy, M., & Abrahamsen, R.K. (2009)

X-ray images for the control of eye formation in cheese.

International Journal of Dairy Technology, 62, 147-153.

Paper no.: 3

Kraggerud, H., Solem, S., T. , & Abrahamsen, R.K.

Quality scoring – a tool for sensory evaluation of cheese?

Submitted to Food Quality and Preference June 2011, resubmitted Oct 2011.

Paper no.: 4

Kraggerud, H., Næs, T. , & Abrahamsen, R.K.

Prediction of sensory quality of cheese during ripening from chemical and spectroscopy measurements.

Submitted to International Dairy Journal.

1 Introduction

Food quality refers to all the attributes that influence the value of a product for the consumer and comprises intrinsic product attributes like safety, sensory properties, convenience and health, and extrinsic attributes like how it is produced (Luning & Marcelis, 2007). An effective approach to ensuring consistent delivery of products of defined quality is very important for a food producer. Quality control (QC) throughout the production chain from raw materials to final cheese product is a challenge. Furthermore, as maturation of the cheese continues from cheese making right through to the sales period for cheese, additional product variations are introduced before it reaches the consumer's plate. The sensory quality of a product is perceived directly at consumption, making relevant measures of sensory quality a prerequisite.

There are two important strategies for ensuring quality during a food production process:

1. Process regulation – in which the actual process is regulated according to input from measurements of raw materials, process or product – which is frequently used for automation
2. Statistical Process Control (SPC) – with continuous improvement and reduction in variability as the main goals

The two strategies are often combined in the control of cheese production and will be further discussed.

Time and cost issues are important in the choice of analytical methods for quality control of cheese. Sensory methods are often time-consuming and they depend on human senses. Individual variations among product quality assessors and consumer target quality, and variability in measurements are important issues. Many attempts have been made to replace sensory analysis with “objective” measurements, such as chemical analysis, for higher precision of analytical results, and spectroscopic methods in recent decades for rapid results and often also non-destructive analysis. In-line and online analytical methods in production lines enable automation of a process, as time lag may make immediate regulation of a process impossible.

The goal of this thesis has been to examine methods useful for control of cheese quality. Two Norwegian cheese varieties have been used as models. Evaluation of different analytical methods for use during cheese ripening was an important part of the task. Emphasis was laid on measurement of the sensory quality of mature cheese as this is the most relevant factor from a consumer’s point of view. A better understanding of ripening and maturation processes of the same cheese varieties has also been an important element. This is especially useful for the strategy of continuous improvement, which requires insights into the all aspects of the subject area and a holistic approach.

2 A brief theoretical review

2.1 Quality control

Food quality management is important for a food producer. One of the most frequently used standards for quality control, ISO 9001, (International Organization for Standardization, 2008) requires that “top management shall ensure that customer requirements are determined and are met with the aim of enhancing customer satisfaction”. Furthermore, according to ISO 9001, “the organization shall plan and implement the monitoring, measurement, analysis and improvement processes needed to demonstrate conformity to product requirements”. Quality management is often solved by applying control systems and procedures. A holistic approach to quality control is necessary, and Figure 1 presents an approach applicable to food quality management, showing how many functions interact, and illustrating how it must be given careful attention. The human role in decisionmaking and as important stakeholders in the processes must be underlined and given special attention.

Process control is extensively used in the dairy industry. It enables automation of processes and interactive decision support throughout a process. In cheese production examples of process steps using automatic regulation are the temperature control of milk pasteurization, fat content standardization and temperature control during the various steps in cheesemaking. A controlled variable is kept constant at a given setpoint. An input variable is measured and used to decide what actions must be taken to reach the target setpoint. To be able to control a process, knowledge about relationships between raw materials, process parameters and resulting end-product attributes is necessary (Jørgensen & Næs,

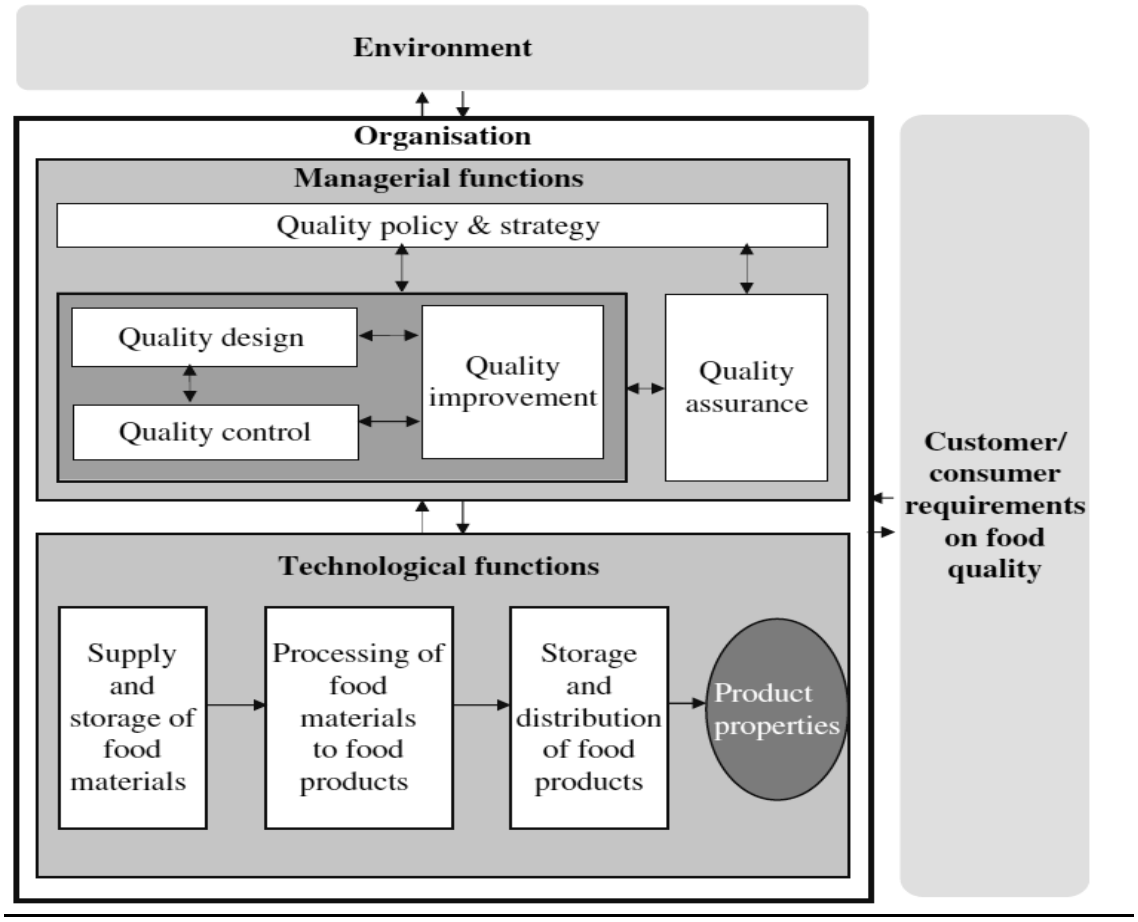

Figure 1 Food quality management functions model (Luning & Marcelis, 2007)

2004). Predictive modelling relates target responses back to input factor settings, using different mathematical models such as statistical modelling, fuzzy systems and artificial neural networks. These provide efficient ways of studying the complexities and interactions in production of dairy products (Roupas, 2008). Five strategies for reduction of variability are suggested by MacKay & Steiner, (1998) and illustrated in Figure 2, comprising: Strategy 1 Output inspection/sorting, Strategy 2 Feedback control, Strategy 3 Reducing variation in process input, Strategy 4 Feed forward control, Strategy 5 Making the process less sensitive to variation in input.

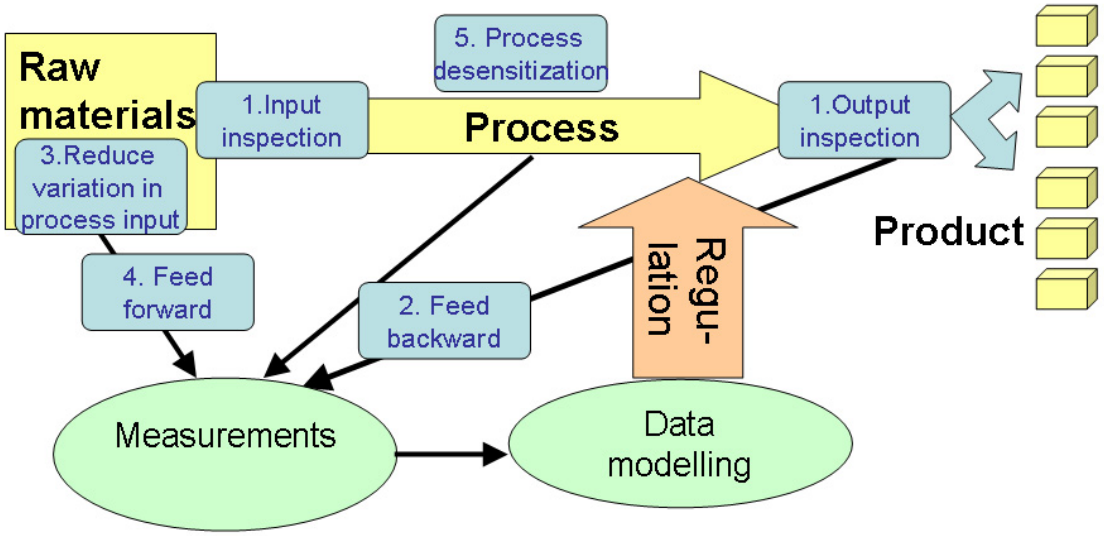

Figure 2: Process control through the process from milk to cheese consumption. Strategies for variability reduction illustrated

Statistical process control (SPC) is the application of statistical methods for monitoring and control of a process, both the target value of a process and the variation of the process about that value. The focus is on continuous improvement and reduction of variation in the various unit operations in a processing line and in the properties of the end product. By collecting data from samples at various points within the process, variations in the process that may affect the quality of the end product can be detected and corrected. Early detection and prevention of problems is emphasized. Key tools in SPC are control charts and designed experiments. Figure 3 illustrates stepwise improvements in a process, resulting in lower variation in the measured variable.

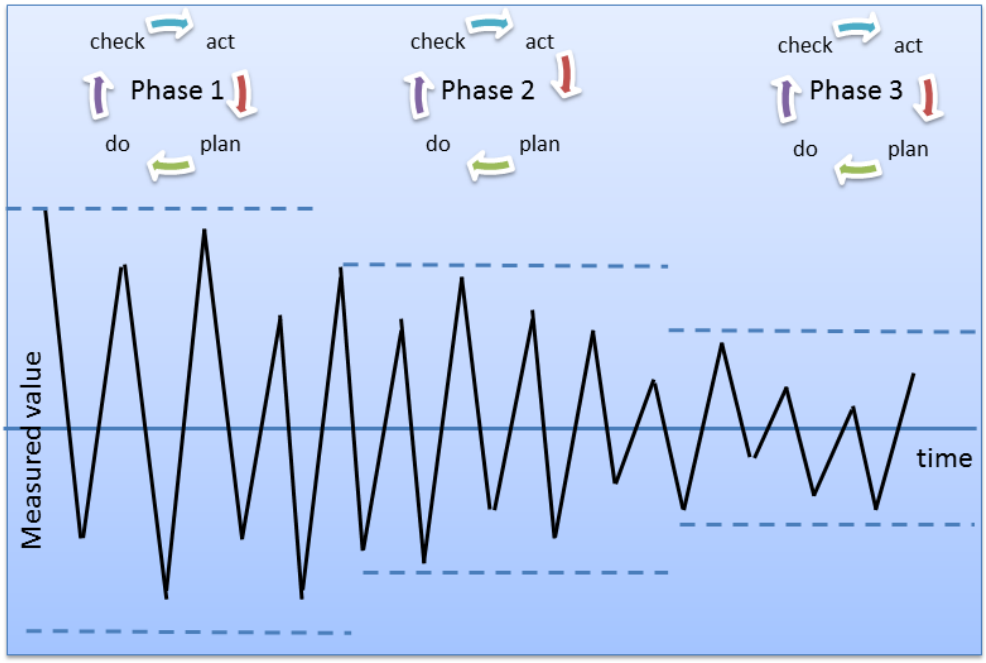

Figure 3 Statistical process control (SPC) with continuous improvement. Consecutive phases resulting in reduction in variability after each phase.

Homogeneity of dairy product samples and variability of analytical results are issues not very often addressed in research papers. But from our experience these factors are important in industrial cheese production and make interpretation and use of analytical results challenging. Statistical methods are thus necessary to analyse the results. Milk is a complex raw material of variable composition and quality which can introduce significant variation into the properties of the products made from it (Roupas, 2008), for example, through changes from season to season (Allais, Perrot, Curt, & Trystram, 2007).

The operators working in food processing premises traditionally play an important role in food manufacturing. They make online evaluations of product properties during production and they adjust processes according to experience to ensure quality and smooth running of the unit processes and the whole processing line.

Methodological guidelines on how to develop a decision support system based on expert know-how have been published (Allais, Perrot, Curt, & Trystram, 2007). This is interesting with respect to traceability, safer measurements, training and instruction of inexperienced operators and increase in reliability of decisions (Roupas, 2008).

2.2 Sensory quality

Since late in the 19th century scoring methods have been used by the dairy industry for sensory evaluation of products. There are three basic categories of sensory tests: 1) traditional judging/grading, 2) consumer tests (affective), and 3) analytical sensory tests (Bodyfelt, Drake, & Rankin, 2008). Selection and training of panel members is especially important for methods in categories 1) and 3) (Delahunty & Drake, 2004). Sensory evaluation methods are very different in the amount of time they consume. Descriptive analysis, a sensory analytical method much used in research and development, is much more time-consuming than the quality scoring method. Our experience is that a proportion of at least 10:1 is realistic timewise, in favour of quality scoring. This makes conventional descriptive methods less relevant for quality classification for regular use in the industry.

Product specifications are essential to a food producer. It is important to include consumers' input in establishing and evaluating sensory product specifications in order to ensure that consumers' expectations are met. Several methods are suggested in the literature. Consumer acceptance limited to evaluation of defects can be determined by so-called survival analysis (Hough, Sanchez, Garbarini de Pablo, Sanchez, Calderon Villaplana, Gimenez, & Gambarot, 2002). Development of a consumer-preference-based scoring guide has been described for a total quality

scoring system (Ismail, Haffar, Baalbaki, & Henry, 2001). Some authors have used a descriptive analysis method throughout product development and quality control, with use of consumer responses to determine target ranges of intensity and limits for each sensory attribute (Pecore & Kellen, 2002; Weller & Stanton, 2002). The use of preference mapping techniques is widespread and is applied in this paper for specification of target quality, notwithstanding an example from Norway using preference mapping which showed lack of agreement between quality specifications, assessment and consumers' preferences (Hersleth, Ilseng, Martens, & Næs, 2005).

In contrast to sensory intensity, quality is more elusive. Considerable difficulty is involved in establishment of a frame of reference, a definition, measurements and interpretation of results (Bodyfelt, 1981). Absence of defects is important as well as the descriptive definition of quality (Amerine, Pangborn, & Roessler, 1965). Traditional quality evaluation methods for dairy products are based on the use of expert assessors and are defect-oriented (Bodyfelt et al., 2008). Daily grading at the manufacturing location based on deviation from a reference scale has been recommended (Pecore & Kellen, 2002; Weller & Stanton, 2002). For quality classification, there is need for a determinative term to make the sorting task easy in practice. This term could either be calculated from a number of separate parameters, as in Quality Index Methodology (QIM) (Martinsdóttir, Sveinsdóttir, Luten, Schelvis-Smit, & Hyldig, 2001) or it could be executed directly by the assessors using overall quality terms (Elortondo, Ojeda, Albisu, Salmeron, Etayo, & Molina, 2007; Etaio, Albisu, Ojeda, Gil, Salmerón, & Elortondo, 2010; International Organization for Standardization & International Dairy Federation, 2009; King, Gillette, Titman, Adams, & Ridgely, 2002; Pecore & Kellen, 2002).

2.3 Cheese production

Cheesemaking is an ancient method of conserving and dehydrating milk. Basically the same raw materials – milk of different species – are turned into a large range of cheese varieties, using the same production principles. During the past two centuries cheesemaking has changes from a craft activity into an industrial one, with automation and extensive control systems and greater uniformity of production. Still there is a lot of variability to be dealt with, originating from biological raw materials and microbial conditions during cheesemaking and ripening. Throughout manufacture and ripening, a series of finely-tuned biochemical steps occur, some in succession and some simultaneously and lead to high quality products when in balance However, imbalance can lead to off-flavours and off-odours. No two batches of the same variety, and probably no two cheeses, are identical (Fox & McSweeney, 2004). Factors affecting cheese quality are illustrated in Figure 4, which also shows the main steps of cheesemaking (Fox & Cogan, 2004). This gives an idea of the complexity of the issue of cheese quality. The quality of cheese is influenced by the gross composition, especially moisture content (moisture-in-non-fat-solids(MNFS)), NaCl concentration (S/M), pH and fat / fat-in-dry-matter(F/DM). Several authors agree that moisture content, pH and S/M are key determinants in cheese of the Cheddar type (Fox, 1975; Gilles & Lawrence, 1973; Pearce & Gilles, 1979). In a very extensive study of New Zealand commercial cheeses, the following conclusions were reached: 1) Within the given compositional range (e.g. 52-56% MNFS), composition did not have decisive influence on the quality grade, which decreased outside this range of MNFS, 2) Composition alone does not provide an exclusive basis for grading, 3) MNFS was found to be the principal factor affecting quality (Lelievre & Gilles, 1982).

2.3.1 Raw milk

Milk is the main raw material of cheese and its microbial, enzymatic and chemical status are of great importance to the end product. Major constituents of milk are water, fat, protein, lactose, organic acids and minerals. Variation in the composition of milk can be due to a number of factors, among them breed, feeding and season, animal health and stage of lactation. Milk from cows in very early and late lactation should not be used, nor milk from cows with mastitis. The milk should also be free from antibiotics that may inhibit bacterial growth and possibly cause allergic reactions for some consumers. Chemical taints and free fatty acids which can cause off-flavours in cheese should also be avoided (Fox & Cogan, 2004). Strategy 1 Output/input inspection/sorting, from chapter 2.1 (MacKay & Steiner, 1998), should be used.

Many model experiments and single factor cheesemaking experiments have been conducted studying the factors affecting renneting, such as protein content and protein composition of the milk, pH and Ca content. But there is a lack of information from cheesemaking experiments involving several simultaneous changes in such factors. Standardization of milk composition before the actual cheesemaking process makes reduction of variability in milk composition possible. Examples can be the concentrations in milk to a predetermined level of total solids, fat or protein, standardization of the ratio of fat and protein/casein, adjusting the pH and the calcium content in the milk by adding $CaCl_2$. Typically strategy 3, Feedback control (chapter 2.1.) for reducing variation in process input, (MacKay & Steiner, 1998).

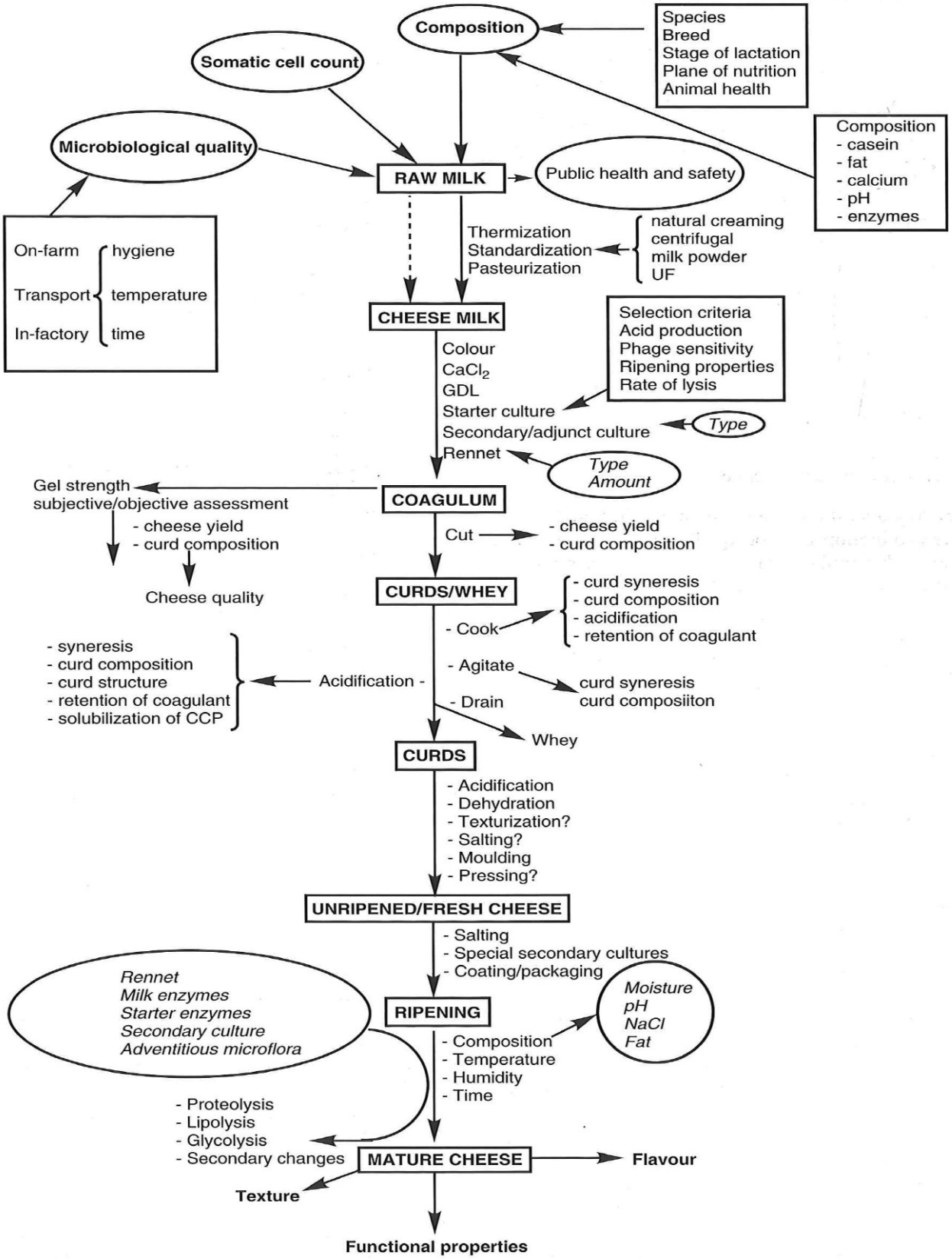

Figure 4: Interaction of compositional and technological factors that affect the quality of cheese (Fox & Cogan, 2004)

2.3.2 Bacterial cultures

Acidification by means of the growth of lactic acid bacteria (LAB) is an important step in cheesemaking, and is essential for the characteristics of a cheese. In pasteurized, and also in some cases unpasteurized cheesemilk, LAB are added (and known as the starter culture), often as mixed strain starters with multiple bacterial types and strains. As renneting is highly pH dependent, the attributes of the coagulum are strongly affected by the growth of the LAB culture. The starter culture is also very important for cheese ripening because of the activity of various enzymes produced by and released from the starter bacteria during cheesemaking and ripening of the cheese. Adjunct cultures of microorganisms other than LAB are also used for some cheese varieties. Control of the process is frequently based on the results of analyses of the starter culture, such as activity measurements, for adjustment of the amount of starter added and parameters like the time and temperatures used in the cheesemaking.

2.3.3 Coagulation

An important step in cheesemaking is the coagulation of cheese using rennet as coagulant which, with its proteinase activity in the presence of calcium, results in gel formation at temperatures above 20°C. Standardizing the conditions of coagulation and control of the coagulation step are important in order to obtain consistent quality of the cheese. Firmness of the gel is traditionally checked by manual procedures and many attempts have been made to make this checking less dependent on human intervention. Spectroscopic methods are among the most frequently used for automatic control of gel firmness or as guidance to the staff in their determination of

the correct level of firmness for cutting the gel, which is the next process step (Callaghan, 2011; Dal Zotto, De Marchi, Cecchinato, Penasa, Cassandro, Carnier, Gallo, & Bittante, 2008; De Marchi, Fagan, Donnell, Cecchinato, Dal Zotto, Cassandro, Penasa, & Bittante, 2009; Sandra, Alexander, & Dalgleish, 2007; Sandra, Cooper, Alexander, & Corredig, 2011). These methods can be used to establish predictive models for feedback control (strategy 2 Feedback control in 2.1.), and make possible automatic activation of the cutting of the gel.

2.3.4 Syneresis and post-coagulation

A milk rennet coagulum shows strong syneresis if cut or broken. When an optimal level of gel firmness is obtained the coagulum is cut and stirred. For washed curd cheeses, like the Dutch type, some of the expelled whey is removed and water is added and again removed by draining off some of the whey before pre-pressing the cheese which at this stage is still immersed in whey. Temperature fluctuations are also a part of the post-coagulation process. With a number of factors which may be regulated, like temperature, acidity, stirring velocity, cut size of the curd particles, amount of whey removed and amount of water added, it is possible to exert considerable influence on the acidity of the curd. These factors indicate a very complex set of processes and regulations and may create a lot of possible variations within the same cheese variety. However, the water content of the resulting pressed cheese is the key parameter. Correct regulation of these parameters is essential for the quality of the cheese and regulation by feed-forward / feed-backward strategies can be used with various methods of measuring the degree of syneresis from the gel. Methods used include determining the amount of whey expelled and measuring dry matter or density in curd pieces (Walstra, van Dijk, & Geurts, 1985). Feed-forward predictive modelling using nine input process variables and neural network methodology has been found useful for predicting pH (Paquet, Lacroix, & Thibault, 2000). Statistical modelling methods were used in a similar approach measuring pH and moisture (Perrot, Agioux, Ioannou, Mauris, Corrieu, & Trystram, 2004). A model has also been developed for measuring cheese fines in whey (Jørgensen, Segtnan,

Thyholt, & Næs, 2004). Computer visualization has also been utilized for cheese curd syneresis measurements (Everard, Callaghan, Fagan, Donnell, Castillo, & Payne, 2007; Everard, Callaghan, Mateo, Castillo, Payne, & Donnell, 2009).

2.3.5 Forming, pressing and salting

After syneresis of the coagulum, the cheese grains are pre-pressed, formed and pressed in moulds. The accuracy of portioning of the quantity of cheese per mould is another key performance parameter. Several other factors that are susceptible to regulation occur at this point in the process. Having the weight as constant as possible is very important for minimizing losses in the packaging process. The transition from a batch process to a continuous process, where continuous systems for pre-pressing and forming are used, has to be handled properly. Post-acidification and stirring before pre-pressing of one batch of cheese over time is a challenge in order to obtain the same cheese composition and weight throughout the whole batch. In systems of batch-wise pre-pressing an even distribution of the curd over the whole area for pressing and avoidance of intake of air into the mixture of cheese and whey are important.

Measuring the moisture content of the cheese as soon as possible in the process is important in order to be able to adjust cheesemaking parameters as early as possible for the succeeding cheese vats. For this purpose, online NIR reflectance measurements have been implemented directly after the cheese has been put in the moulds, with successful calibration results in some dairies in Norway (pers.comm. TINE). Procedures for pressing and draining off the whey are mostly well standardized. Another possible control point is after pressing, but the delay in obtaining the proper results at this point has been found too long to utilize them for any corrections necessary to the ongoing cheesemaking process. Furthermore, the accuracy of the results has not been much better than at the control point directly after moulding (pers.comm.,TINE).

Salting is also an important part of the process. For the cheese varieties in question brining for 1-3 days is used. The temperature of the brine and NaCl concentration in the brine are of course important to check in order to obtain a stable and evenly

distributed salt content in the final cheese matrix. In this important part of the cheesemaking process both temperature and salt content strongly affect the growth of starter and non-starter bacteria in the cheese and this, again, is important for the end-product quality of the ripened cheese.

2.3.6 Cheese ripening

The ripening of cheese is due to the activity of microorganisms and enzymes that come from various sources: the raw milk itself, rennet, primary starter, secondary cultures and non-starter bacteria. Breakdown of proteins, fats and carbohydrates are the main sources of typical flavour compounds and structure characteristics in ripened cheese, as illustrated in Figure 5.

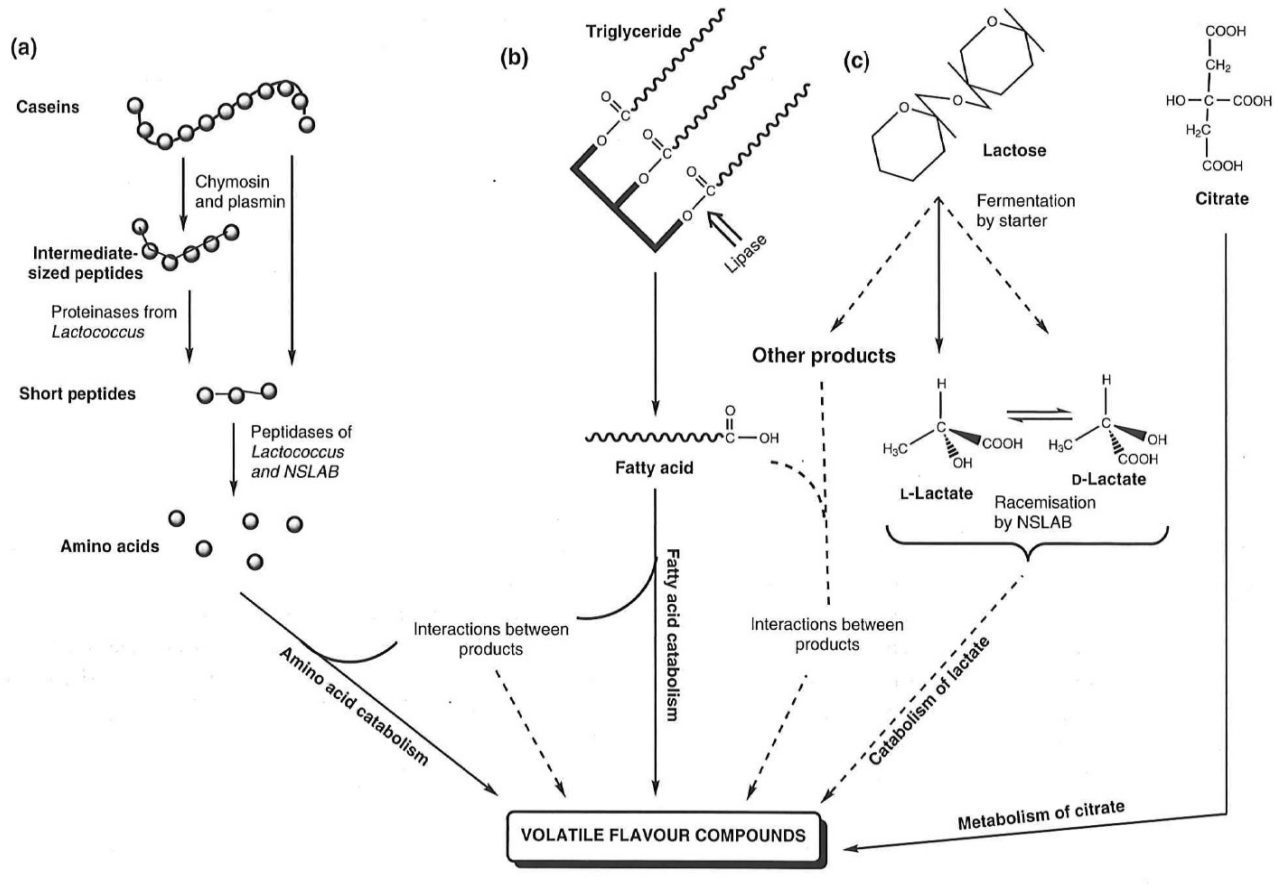

Figure 5: General overview of the biochemical pathways which operate in cheese during ripening. From (McSweeney, 2004).

Factors affecting microbial growth and enzyme activity directly influence the ripening process. Ripening temperature was shown to have systematic effects on sensory properties (Kraggerud, Skeie, Høy, Røkke, & Abrahamsen, 2008). Temperature is a very important factor. The time/temperature scheme used during the main steps of the cheesemaking, like pre-acidification of cheesemilk, renneting, post-coagulation, pressing/moulding, salting, pre-storage, the early stage of ripening and the later stage (maturation) is characteristic for the cheese variety in question. Ripening at elevated temperatures for shorter periods for eye formation is typical for the Dutch-type cheese varieties produced in Norway, but still there is considerable variation between dairies with respect to the total time/temperature scheme for the various steps after moulding, even with the production of the same cheese variety.

Proper control and possible altering of the temperature scheme could be used as a way of affecting end product quality in a desired direction. This requires extensive knowledge about the effects on quality. The measurement of relevant input parameters as early as possible in the cheesemaking procedure is important in order to control and possibly change the temperature at relevant stages in the process. Although cheese ripening is continuously the subject of research (Collins, McSweeney, & Wilkinson, 2003; Fox & McSweeney, 2004; McSweeney & Sousa, 2000; Sousa, Ardö, & McSweeney, 2001; Yvon & Rijnen, 2001) there is still a lot of work to be done in order to be able to control the ripening of cheese completely.

2.4 Methods of analysis

Cheese analysis includes microbiological evaluations, compositional analysis and analysis of metabolic products formed during ripening of the cheese, in addition to sensory analysis. This thesis pays specific attention to rapid analytical methods. Furthermore, sampling of cheese requires knowledge and care, irrespective of the analysis in question. There are guidelines for sampling available published by various authors and standardization organizations (International Dairy Federation, 1995). Still, variations within and between cheeses of the same batch can be considerable in our experience.

2.4.1 Compositional analysis

Methods published by ISO, IDF and AOAC are those most frequently used for analysis of cheese composition, including moisture, protein, fat, ash and salt. They will not be further discussed here. These standard methods are generally labour-intensive and often time-consuming – with a delay of at least one day from sampling to the results being available, for instance for dry matter analysis by oven drying. This makes use of the results impossible during production – and these results can only be used in a retrospective way. For modern process control rapid methods are a prerequisite. Some more traditional methods have been developed to obtain a faster result that can be used for process control, e.g. dry matter using microwave oven instead of a conventional oven. Still the process of cutting the sample, grating and so on, takes considerable long time. The same applies to many of the spectroscopic methods which may involve pre-processing of the sample before the analysis itself can start.

2.4.2 Monitoring cheese ripening

Cheese ripening has been intensively studied for a number of years, monitoring primary metabolism (breakdown of carbohydrates, lipolysis and proteolysis) and secondary metabolism, including breakdown of fatty acids and amino acids. For this purpose, methods for investigation of biochemical changes in cheese and understanding of the ripening process need to be developed. These questions, too, have attracted much attention among scientists and have been reviewed by several authors (Collins, McSweeney, & Wilkinson, 2004; Upadhyay, McSweeney, Magboul, & Fox, 2004). Common methodology includes chromatography, electrophoresis, colorimetric and enzymatic methods. Preparation of samples often includes dilution, extraction, precipitation, separation, fractionation and/or liberation of compounds. A challenge is often that the complexity of cheese would require a wide range of analyses to describe sensory quality. To avoid the use of a plethora of analytical methods, instruments like electronic noses have been tried (Hansen, Petersen, & Byrne, 2005), but these instruments have not yet proven to be very useful for practical purposes. Chemical and instrumental methods used in the study of cheese

ripening and cheese quality have recently been reviewed by (Subramanian & Rodriguez-Saona, 2010).

2.4.3 Rapid analytical methods

The use of chemical and instrumental analysis involves several problems, among them: 1) use of solvents 2) requirement for specific accessories 3) extensive sample preparation 4) labour-intensive operations, 5) expensive equipment. Development and evaluation of new, rapid and simple methods have therefore been in focus (Subramanian & Rodriguez-Saona, 2010). Advances in spectroscopic instruments and data analysis have enabled the development of rapid and non-destructive methods of cheese analysis performed within a few seconds. Some of these methods may be used for measurments on cheese directly in the production line, as identified in Figure 6. Using inline/online instruments allows control of the production process using feed-forward/feed-back control strategies and predictive modelling methodology. (Roupas, 2008).

Spectroscopic methods are based on emission or absorption of electromagnetic radiation. Light is considered to be transmitted in photons and when light interacts with matter, it may stimulate transitions between energy levels, depending on the energy of the photon, which in turn is related to the frequency of the electromagnetic spectrum according to this equation:
E(energy)= *h*(Planck's constant)**v*(frequency) (Wilson, 2002)

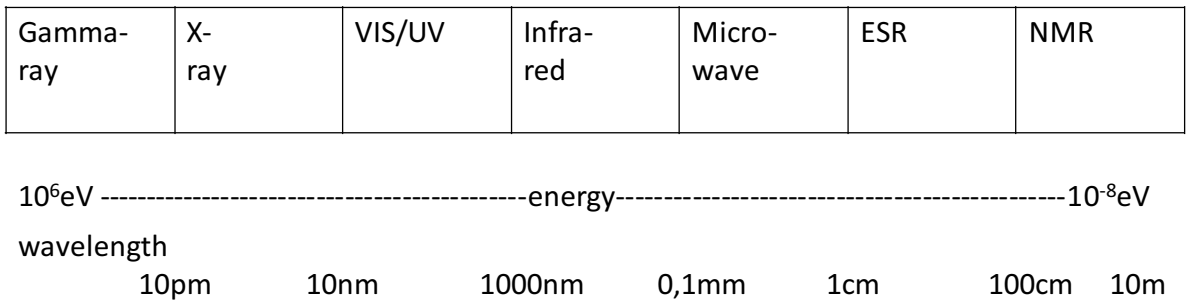

Figure 6: The electromagnetic spectrum, illustrating energy levels and wavelengths of the different parts of the spectrum. (NMR, nuclear magnetic resonance, ESR, electron spin resonance, VIS/UV, visible and Ultraviolet).

In UV/VIS spectroscopy, absorption of radiation is the result of excitation of bonding electrons in chromophores. If an electron is promoted to a higher energy state, it may lose that energy again and during relaxation to the lower electronic state, a photon is emitted, giving fluorescence.

Transitions between vibrational energy levels are the basic principles in infrared and Raman spectroscopies. The bonds between atoms are stretched and caused to oscillate at some natural frequency, dependent on the force constant of the bond and the masses of the actual atoms. Thereby different functional groups absorb light at different wavelengths of the electromagnetic spectrum (Wilson, 2002).

Several spectroscopy instruments have been developed during the recent decades and are widely in use also in dairy industry, for cheese especially for gross composition analysis (Müller & Steinhart, 2007).

2.5 Mathematical modelling

2.5.1 Multivariate data modelling

Cheese is a typical case of a complex material. Multiple variables are needed to be able to describe the nature of the sample in question. Today the a priori understanding of mechanisms and correlations is incomplete, as described earlier. It is therefore very difficult to establish detailed causal modelling to understand cheese quality. Thus data compression methods are needed to work on such a complex issue. Basically information on many variables is concentrated into a few underlying, latent variables, normally called components, scores or factors. Principal Component Analysis (PCA) is frequently used as a data compression method, enabling one to plot a concentrate of the information from many variables in one, two or three dimensions. The first dimension is the one that carries most information, the second PC will then carry the maximum share of the residual information (i.e. that not taken into account by the previous PC), and so on. Using a loading plot it is then possible to visualize which variables are important, which are correlated with each other and

how they relate to the samples. Using a score plot in the sample space it is possible to see which samples have the most in common and which are the most different and, when the corresponding loading plots are used, which variables describe which samples. The principles of extracting scores in two PCs is illustrated in Figure 7.

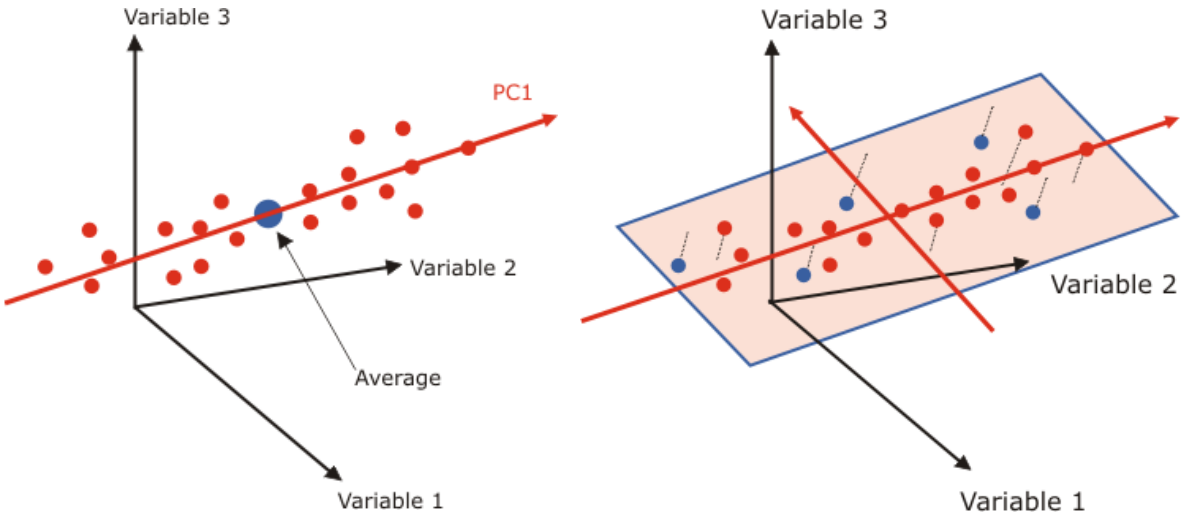

Figure 7: Scores from PCA. Left: The First Principal Component. Right: PCs 1 and 2. Illustration from Unscrambler 10.1. Help function (CAMO Software AS, Oslo Norway).

For regression purposes, factors from data compression are used as regressors when trying to model one or many regressands. Frequently used methods are Partial Least Squares regression (PLS) and Principal Components Regression (PCR).

PLS maximizes the covariance between X and Y. This is in contrast to PCR, which first performs Principal Component Analysis (PCA) on X and then regresses the scores (T) against the Y data. A conceptual illustration for PLS is shown graphically in Figure 8. (Allais et al., 2007; Martens & Martens, 2001)

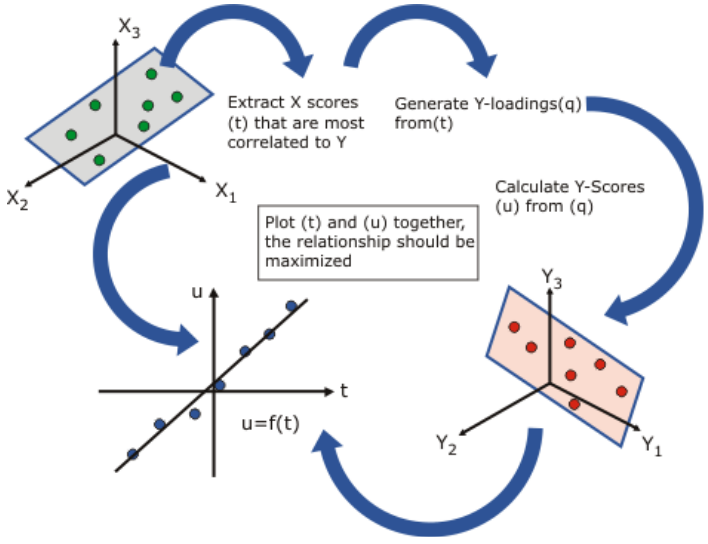

Figure 8: PLS procedure. Illustration from Unscrambler 10.1. Help function (CAMO Software AS, Oslo, Norway).

2.5.2 Predictive modelling of manufacturing processes

There are different approaches to mathematical modelling of processes and they can also be mutually interactive. The approaches can be divided into three main types according to Roupas (2008):

1) White-box-models. This approach is based on prior knowledge and fundamental principles from theory of the actual process and factors for modelling. It is also often described as mechanistic, first-principle or phenomenological.
2) Black-box models. They are applied to mechanisms and relationships that are poorly understood or too complex to model. Multiple regression models and neural networks are typical black-box models. They are also often described as empirical, inductive or input/output modelling.
3) Grey-box models. These combine the use of white-box and black-box modelling and are particularly useful when there is lack of fundamental theory to describe the process in question or when there is need to decrease the complexity of the model.

3 Main results

This work has had focus on control of sensory quality of cheese, and has has focus on different control strategies, as well as analytical methods. Most of the work was organised in three different experiments on two different semi-hard, washed curd, commercial Norwegian cheese varieties with similar gross composition. A broad range of chemical, sensory and spectroscopy methods were made on 244 cheese samples at three times during ripening: 8, 24 and 40 weeks, making 732 samples altogether. Results from these experiments were published in paper 1, 3 and 4. Experiments in paper 2 was carried out independently.

3.1 Paper 1: Season and ripening temperature influence fatty acid composition and sensory properties of semi-hard cheese during ripening.

The experience of Norwegian cheesemakers is that the speed of ripening and quality of cheese are affected by the season. Maintaining even cheese quality throughout the year is therefore a challenge for the cheesemakers. The objective of this paper was to study how seasonal variation of raw milk and different ripening conditions influence the sensory attributes of cheese during maturation. To study general effects in cheese of this type, two different semi-hard, washed curd, commercial Norwegian cheese varieties with similar gross composition were examined.

Multivariate models derived from sensory attributes of cheese demonstrated that ripening temperature and maturation time had systematic effects on the sensory properties of cheese of the varieties examined. The effects of the two factors were independent. In one of the cheese varieties sulphurous aroma occurred, and decrease in sulphurous aroma was observed during maturation. In the other cheese variety sulphurous aroma was hardly registered at all, probably because a different adjunct culture was used.

Fatty acid composition of raw milk varied systematically with season, showing a continuous trend throughout the year. The main differences were found between the indoor and outdoor feeding regimes applied according to the seasons. Saturated fat was higher with indoor feeding and unsaturated fat higher with outdoor feeding. Correlations were also observed between fatty acid composition of the raw milk and sensory properties of cheese. High firmness was correlated with indoor feeding and

saturated fatty acids. Flavour intensity of the cheese was found to be higher with outdoor feeding.

3.2 Paper 2: X-ray images for the control of eye formation in cheese.

Appearance is particularly important for cheese types with eyes and checking eye-formation is normally done by splitting cheese manually and this makes the actual cheese blocks unusable for normal commercial purposes. Therefore non-destructive monitoring of eye formation in cheese during ripening is desirable. A simple method was developed, based on existing equipment in the dairy industry that is normally used for metal detection. Images were acquired using a conventional, low resolution online X-ray instrument. Semi-hard cheese with propionibacteria that had been ripened under different conditions was analysed. Image processing methods were developed for detecting eyes in the cheese and measuring size distribution and eye volume. Overlapping eyes might be problematic but sufficient detection of overlapping eyes was successfully obtained. The method was found promising for quality control as it will make possible non-destructive monitoring of eye formation in cheese throughout the ripening period. This method can enable reduction of variablility with respect to appearance.

3.3 Paper 3: Quality scoring – a tool for sensory evaluation of cheese ?

The objective of this paper was to evaluate the relevance of data from quality scoring methodology of ISO/IDF (2009) performed by expert assessors for the sensory quality control of cheese. The approach to this evaluation was comparison of quality scoring with sensory quantitative descriptive data from a trained panel and consumer preference data. Significant regression correlations were found between quality

scoring and descriptive data in a data set obtained from evaluation of Norwegian semi-hard Dutch-type cheese at 8, 24 and 40 weeks of age (n=459). However, the level of explained variance was low.

In a smaller set of data aimed at preference mapping, higher correlations were found between quality scoring and descriptive data. Preference mapping showed that the average consumer and the quality scoring expert assessors disagreed in particular on the consistency properties of cheese. External preference mapping after segmentation of consumers by hierarchical clustering was found useful. Consumers could be divided into 3 main clusters. One of these clusters mainly agreed with the expert assessors, while the cheese preferences of the two other clusters were in disagreement with expert assessor approval. Thus it would be possible to suggest various product specifications, highly approved by consumers, for the variety of cheese investigated. A high level of explained variance was found between consumers' overall preference scores and overall quality scores and this could indicate that quality scoring is a relevant sensory quality measure.

3.4 Paper 4: Prediction of sensory quality of cheese during ripening from chemical and spectroscopy measurements.

The extensive material of 459 samples of Norvegia, the Dutch type of semi-hard cheese was analysed using a number of chemical, chromatographic, sensory and spectroscopic methods during maturation. From 8 to 24 and 40 weeks there was a highly systematic development in chemical and sensory attributes. Modelling with multivariate regression, PLS, gave relatively low correlation coefficients between sensory and other analytical methods, probably due to high standard error in the sensory data. Chemical data and FTIR measurements gave almost equivalent validation results in prediction of sensory data. Fluorescence spectroscopy and spectroscopy with NIR between 400m and 1100nm, both performed on the surface of

cheese, showed slightly less valid results for measurement of sensory variables in this experiment. Using a combination of spectra from all instruments gave a higher correlation than spectra from instruments taken separately.

Sensory characteristics at the greater age (40 weeks) were not very well forecasted by early measurements on cheese (8 weeks), when examined by sensory, chemical and spectroscopic methods. This could be partly due to the noise in sensory data. Cheese producers who would like to predict quality development during maturation and the period of sale would have a considerable benefit from applying this kind of technique.

The results from spectroscopic measurements were promising regarding possible use in control of cheese quality, especially FTIR in combination with a mixture of different spectroscopic data. Nevertheless, it is difficult to envisage that sensory measurements could be replaced completely, but as vibrational spectroscopy could well be used as supplementary analysis. Chemical analysis can to a large extent be replaced by spectroscopy with the advantages of fast results and low variable unit cost. The possibility to analyse more samples in order to cover variability within and between cheese batches is an important potential improvement.

4 Discussion

Customer satisfaction is an important goal of all food producers. One important aspect contributing to this goal is consistent quality of the product, so that the expectations of the consumers are met every time they buy and consume the actual product variety. For cheese, this is even more of a challenge than for other food products, as ripening during storage and the period of sales affects the sensory quality in addition to the potential differences between production batches. Two important strategies for control of food product end quality are discussed in this

thesis: process regulation and Statistical Process Control (SPC). But first of all: how do we define and measure cheese quality? In paper 3 this subject was discussed. The importance of gaining insights into consumers' preferences for the actual product was demonstrated using preference mapping methodology. Segmentation of consumers on the basis of sensory preference patterns is very important in this aspect and also makes diversification of the product portfolio a possibility. Product specifications prepared on the basis of sensory consumer testing comprise an important task together with the possibility to evaluate the conformity between products and product specifications. For this purpose the standardized quality scoring methodology currently in use in Norwegian dairy industry was found suitable. The continuous work on assessor training and coordination which is a prerequisite for the suitability for purpose of the methodology should be emphasized.

Process automation – where the process is regulated according to input from measurements of process and products during production – relies on rapid measurements to ensure minimum delay. In paper 3 spectroscopic methods were found promising as a source of fast and reliable measurements of important parameters in cheese. Spectroscopy was found to be an adequate substitute for chemical measurements, for the purpose of measuring relevant sensory attributes of cheese. There is also a lot of work done by other authors and suppliers of equipment in this field, enabling more and more automation in the dairy industry. In paper 2 the development of X-ray methodology suitable for non-destructive online measurements of eye formation in cheese during ripening was described. This method could be used as a tool for checking the time for transferring cheese from the warm room to a lower temperature. This is an important manual control point in the cheese ripening

process at which cheese is divided to evaluate eye formation. As this causes loss of cheese, only very few cheeses are checked and this introduces uncertainty as variation between single cheeses can be considerable. A non-destructive method can therefore be of great value for industry.

In statistical process control, SPC, reduction of variability through continuous improvement of each process step and the process as a whole is important. The ultimate measure is final product quality, as this is what the consumers encounter when they consume, and sensory quality is the quality perceived by them. For this purpose, thoroughgoing knowledge of the process and causes of variation, as well as understanding of the influence of all relevant factors from the raw material through to the cheesemaking process and the product itself on cheese quality is essential. This is not an easy task. Cheese reseach has been going on for more than a century and there will continue to be a lot of research activities aiming to understand the depth of variability of cheese. In paper 1 the effect of variation throughout the four seasons of the year, as well as the effect of maturation in different ripening temperatures, on sensory properties of cheese were examined. This gives us a better platform for adjustments with the aim of reducing variability in today's cheese varieties.

5 Challenges and future perspectives

Food quality control is a complex area, and a continuous challenge, especially owing to the biological nature of the raw materials and their sources of supply. Changing requirements due to increased consumer awareness on subjects like food safety and environmental issues, as well as food law requirements, and increased international

competition, result in quality management acquiring much more attention in the food industry. In this work a holistic approach is necessary with emphasis both on managerial and technological functions. The dynamics between humans and food technology, quality systems and the know-how of the employees are important in decision-making processes and implementation of systems. Decision support systems will probably be of greater importance in the future, either based on automated or manually-obtained information, looking at the food chain as a whole

Cheese, with its enzymatic and microbial processes and milk as a quite variable biological raw material, is one of the most complicated products to make and to ensure that the right quality factors play an important role. To our knowledge all cheese producers struggle with variability in quality, at least from time to time. The ability to measure relevant information from raw materials and the process, in order to control and predict future quality, will certainly be further developed in future. The drivers of this development will be evolution in sensors and instruments, in computer technology and mathematical modelling of the data. Genetic tools will probably be important in a biological process, as well as enhanced understanding of cheese technology. Already a wide range of analytical techniques is available and can be exploited, but this situation is often ignored in existing routines in production plants which are often hard to change. So-called expert systems, or intelligent systems, which started to be developed in the 1970-80's, have not been the success that had been expected, probably owing to their complexity and opposition from operators. But availability of new technology will probably encourage the use of such systems in time.

Another driver of development, is economics. Obtaining only small gains in product yield, which for example can be made by targeting gross composition more precisely in cheese, has been the main driver in our company's development in this area.

Dairy companies all over the world have visions regarding technological development. One example is from Fonterra (New Zealand), one of the largest dairy companies in the world. Their vision is of "Lights out manufacturing units": a vision of fully automated factories, to such an extent that there are no human operators present and the lights can be turned off. Fonterra has also implemented an automatic control system claimed to "think" for itself. It is linked to instruments in the plant that supply measurement information from critical control points. The system can run the plant like the best operator and obtain optimal quality (Mills, 2006).

No doubt, automation of the control of cheese quality will be a hot topic for the future!

6 Reference list

Allais, I., Perrot, N., Curt, C., & Trystram, G. (2007). Modelling the operator know-how to control sensory quality in traditional processes. *Journal of Food Engineering, 83*, 156-166.

Amerine, M. A., Pangborn, R. M., & Roessler, E. B. (1965). *Principles of sensory evaluation of food*. New York: Academic Press, Inc.

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

Bodyfelt, F. W. (1981). Dairy Product Score Cards: Are They Consistent with Principles of Sensory Evaluation? *Journal of Dairy Science, 64*, 2303-2308.

Callaghan, D. J. (2011). CHEESE ; Gel Firmness and Its Measurement. In H. Roginski, J. W. Fuquay, & P. F. Fox (Eds.), *Encyclopedia of Dairy Sciences* (pp. 585-590)

Collins, Y. F., McSweeney, P. L. H., & Wilkinson, M. G. (2004). Lipolysis and catabolism of fatty acids in cheese. In F.F. Patrick (Ed.), *Cheese Chemistry, Physics and Microbiology* (Volume 1 ed., pp. 373-389). London, UK: Academic Press.

Collins, Y. F., McSweeney, P. L. H., & Wilkinson, M. G. (2003). Lipolysis and free fatty acid catabolism in cheese: a review of current knowledge. *International Dairy Journal, 13*, 841-866.

Dal Zotto, R., De Marchi, M., Cecchinato, A., Penasa, M., Cassandro, M., Carnier, P., Gallo, L., & Bittante, G. (2008). Reproducibility and Repeatability of Measures of Milk Coagulation Properties and Predictive Ability of Mid-Infrared Reflectance Spectroscopy. *Journal of Dairy Science, 91*, 4103-4112.

De Marchi, M., Fagan, C. C., Donnell, C. P., Cecchinato, A., Dal Zotto, R., Cassandro, M., Penasa, M., & Bittante, G. (2009). Prediction of coagulation properties, titratable acidity, and pH of bovine milk using mid-infrared spectroscopy. *Journal of Dairy Science, 92*, 423-432.

Delahunty, C. M., & Drake, M. A. (2004). Sensory Character of Cheese and its Evaluation. In P.F. Fox, P. L. H. McSweeney, T. M. Cogan, & T. P. Guinee (Eds.), *Cheese. Chemistry, Physics and Microbiology* (Third ed., pp. 455-487). London: Elsevier Academic Press.

Elortondo, F. J. P., Ojeda, M., Albisu, M., Salmeron, J., Etayo, I., & Molina, M. (2007). Food quality certification: An approach for the development of accredited sensory evaluation methods. *Food Quality and Preference, 18*, 425-439.

Etaio, I., Albisu, M., Ojeda, M., Gil, P. F., Salmerón, J., & Elortondo, F. J. P. (2010). Sensory quality control for food certification: A case study on wine. Method development. *Food Control, 21*, 533-541.

Everard, C. D., Callaghan, D. J., Fagan, C. C., Donnell, C. P., Castillo, M., & Payne, F. A. (2007). Computer Vision and Color Measurement Techniques for Inline Monitoring of Cheese Curd Syneresis. *Journal of Dairy Science, 90*, 3162-3170.

Everard, C. D., Callaghan, D. J., Mateo, M. J., Castillo, M., Payne, F. A., & Donnell, C. P. (2009). The use of colour parameters derived from an online fibre-optic sensor to monitor curd syneresis during cheese making. *Journal of Food Engineering, 94*, 1-6.

Fox, P. F. (1975). Influence of cheese composition on quality. *Irish Journal of Agricultural Research, 14*, 33-42.

Fox, P. F., & Cogan, T. M. (2004). Factors that Affect the Quality of Cheese. In P.F. Fox, P. L. H. McSweeney, T. M. Cogan, & T. P. Guinee (Eds.), *Cheese: Chemistry, Physics and Microbiology* (3 ed., pp. 583). London, UK: Elsevier Academic Press.

Fox, P. F., & McSweeney, P. L. H. (2004). Cheese: An overview. In P.F. Fox, P. L. H. McSweeney, T. M. Cogan, & T. P. Guinee (Eds.), *Cheese: Chemistry, Physics and Microbiology* (3 ed., pp. 1-18)

Gilles, J., & Lawrence, R. C. (1973). The assessment of Cheddar cheese quality by compositional analysis. *New Zealand Journal of Dairy Science and Technology, 8*, 148-151.

Hansen, T., Petersen, M. A., & Byrne, D. V. (2005). Sensory based quality control utilising an electronic nose and GC-MS analyses to predict end-product quality from raw materials. *Meat Science, 69*, 621-634.

Hersleth, M., Ilseng, M. A., Martens, M., & Næs, T. (2005). Perception of cheese: A comparison of quality scoring, descriptive analysis and consumer responses. *Journal of Food Quality, 28*, 333-349.

Hough, G., Sanchez, R. H., Garbarini de Pablo, G., Sanchez, R. G., Calderon Villaplana, S., Gimenez, A. M., & Gambarot, A. (2002). Consumer Acceptability Versus Trained Sensory Panel Scores of Powdered Milk Shelf-Life Defects. *Journal of Dairy Science, 85*, 2075-2080.

International Dairy Federation. (1995). IDF 50C Milk and milk products - Guidance on sampling. International Dairy Federation. Brussels, Belgium, International Dairy Federation.

International Organization for Standardization. (2008). ISO 9001 Quality management systems - Requirements. Geneva, Switzerland

International Organization for Standardization, & International Dairy Federation. (2009). ISO 22935-3 / IDF 99-3 Milk and milk products - Sensory analysis - Part 3: Guidance on a method for evaluation of compliance with product specifications for sensory properties by scoring. Geneva, Switzerland

Ismail, B., Haffar, I., Baalbaki, R., & Henry, J. (2001). Development of a total quality scoring system based on consumer preference weightings and sensory profiles: application to fruit dates (Tamr). *Food Quality and Preference, 12*, 499-506.

Jørgensen, K., & Næs, T. (2004). A design and analysis strategy for situations with uncontrolled raw material variation. *Journal of Chemometrics, 18*, 45-52.

Jørgensen, K., Segtnan, V., Thyholt, K., & Næs, T. (2004). A comparison of methods for analysing regression models with both spectral and designed variables. *Journal of Chemometrics, 18*, 451-464.

King, S., Gillette, M., Titman, D., Adams, J., & Ridgely, M. (2002). The Sensory Quality System: a global quality control solution. *Food Quality and Preference, 13*, 385-395.

Kraggerud, H., Skeie, S., Høy, M., Røkke, L., & Abrahamsen, R. K. (2008). Season and ripening temperature influence fatty acid composition and sensory properties of semi-hard cheese during cheese maturation. *International Dairy Journal, 18*, 801-810.

Lelievre, J., & Gilles, J. (1982). The relationship between the grade (product value) and composition of young commercial Cheddar cheese. *New Zealand Journal of Dairy Science and Technology, 49*, 1098-1101.

Luning, P. A., & Marcelis, W. J. (2007). A conceptual model of food quality management functions based on a techno-managerial approach. *Trends in Food Science & Technology, 18*, 159-166.

MacKay, R. J., & Steiner, S. H. (1998). Strategies for variability reduction. *Quality Engineering, 10*, 125-136.

Martens, H., & Martens, M. (2001). *Multivariate analysis of Quality – an introduction.* Chichester, UK: Wiley.

Martinsdóttir, E., Sveinsdóttir, K., Luten, J., Schelvis-Smit, J. B., & Hyldig, G. (2001). *QIM Eurofish Sensory evaluation of fish freshness. Reference manual for the fish sector.* IJmuiden, Netherlands: QIM Eurofish.

McSweeney, P. L., & Sousa, M. J. (2000). Biochemical pathways for the production of flavour compounds in cheeses during ripening: A review. *Lait, 80*, 293-324.

McSweeney, P. L. H. (2004). Biochemistry of cheese ripening: Introduction and overview. In P.F. Fox (Ed.), *Cheese: Chemistry, Physics and Microbiology* (Volume 1 ed., pp. 347-360). London, UK: Academic Press.

Mills, R. (2006). Dairy manufacturing with the lights off. *In Focus, 30*, 12-13.

Müller, A., & Steinhart, H. (2007). Recent developments in instrumental analysis for food quality. *Food Chemistry, 102*, 436-444.

Paquet, J., Lacroix, C., Audet, P., & Thibault, J. (2000). Electrical conductivity as a tool for analysing fermentation processes for production of cheese starters. *International Dairy Journal, 10*, 391-399.

Paquet, J., Lacroix, C., & Thibault, J. (2000). Modeling of pH and acidity for industrial cheese production. *Journal of Dairy Science, 83*, 2393-2409.

Pearce, K. N., & Gilles, J. (1979). Composition and grade of Cheddar cheese manufactured over three seasons. *New Zealand Journal of Dairy Science and Technology, 14*, 63-71.

Pecore, S., & Kellen, L. (2002). A consumer-focused QC/sensory program in the food industry. *Food Quality and Preference, 13*, 369-374.

Perrot, N., Agioux, L., Ioannou, I., Mauris, G., Corrieu, G., & Trystram, G. (2004). Decision support system design using the operator skill to control cheese ripening--application of the fuzzy symbolic.. *Journal of Food Engineering, 64*, 321-333.

Roupas, P. (2008). Predictive modelling of dairy manufacturing processes. *International.Dairy Journal, 18*, 741-753.

Sandra, S., Alexander, M., & Dalgleish, D. G. (2007). The rennet coagulation mechanism of skim milk as observed by transmission diffusing wave spectroscopy. *Journal of Colloid and Interface Science, 308*, 364-373.

Sandra, S., Cooper, C., Alexander, M., & Corredig, M. (2011). Coagulation properties of ultrafiltered milk retentates measured using rheology and diffusing wave spectroscopy. *Food Research International, 44*, 951-956.

Sousa, M. J., Ardö, Y., & McSweeney, P. L. H. (2001). Advances in the study of proteolysis during cheese ripening. *International Dairy Journal, 11*, 327-345.

Subramanian, A., & Rodriguez-Saona, L. (2010). Chemical and Instrumental Approaches to Cheese Analysis. In S. Taylor (Ed.), *Advances in Food and Nutrition Research* (pp. 167-213). Advances in Food and Nutrition Research. London, UK: Elsevier.

Upadhyay, V. K., McSweeney, P. L. H., Magboul, A. A. A., & Fox, P. F. (2004). Proteolysis in cheese during ripening. In F.F. Patrick (Ed.), *Cheese Chemistry, Physics and Microbiology* (Volume 1 ed., pp. 391-VIII). London, UK: Academic Press.

Walstra, P., van Dijk, H. J. M., & Geurts, T. J. (1985). Rennet coagulation and coagulants in cheese manufacture. *Neth.Milk Dairy J., 39*, 209-246.

Weller, J. N., & Stanton, K. J. (2002). The establishment and use of a QC analytical/ descriptive/ consumer measurement model for the routine evaluation of products at manufacturing facilities. *Food Quality and Preference, 13*, 375-383.

Wilson, R. (2002). ANALYSIS/ Spectroscopy ,Overview. In H. Roginski, J. W. Fuquay, & P. F. Fox (Eds.), *Encyclopedia of Dairy Sciences* (First ed., pp. 40-45). London, UK: Academic Press.

Yvon, M., & Rijnen, L. (2001). Cheese flavour formation by amino acid catabolism. *International Dairy Journal, 11*, 185-201.

Paper I

Available online at www.sciencedirect.com
ScienceDirect
International Dairy Journal 18 (2008) 801–810
INTERNATIONAL DAIRY JOURNAL
www.elsevier.com/locate/idairyj

Season and ripening temperature influence fatty acid composition and sensory properties of semi-hard cheese during maturation

H. Kraggerud[a,b,*], S. Skeie[a], M. Høy[c], L. Røkke[b], R.K. Abrahamsen[a]

[a]*Department of Chemistry, Norwegian University of Life Sciences, Biotechnology and Food Science, N-1432 Ås, Norway*
[b]*TINE R&D Center, Box 50, N-4358 Kleppe, Norway*
[c]*Matforsk, Norwegian Food Research Institute N-1432 Ås, Norway*

Received 1 May 2007; accepted 6 November 2007

Abstract

The influence of ripening temperature and seasonality on sensory attributes during maturation of cheese was studied in two different semi-hard, washed curd, commercial Norwegian cheese varieties with similar gross composition. Multivariate models derived from sensory attributes of cheese demonstrated that ripening temperature and maturation time had systematic, independent effects on the sensory properties of cheese of the varieties examined. Decrease in sulphurous aroma was observed during maturation. Fatty acid composition of raw milk varied with season, showing a continuous trend throughout the year with the main differences found between indoor and outdoor feeding seasons. Correlations were observed between fatty acid composition of the raw milk and sensory properties of cheese. Firmness was correlated with indoor feeding and saturated fatty acids. In contrast, flavour intensity of the cheese was found to be higher in the grazing season.

1. Introduction

During ripening of cheese, numerous biochemical changes occur that lead to the development of the characteristic flavour and texture for different cheese varieties (Fox & McSweeney, 2004). As temperature is increased, enzymatic and chemical reactions occur at faster rates. Accordingly, one of the major strategies for accelerating cheese ripening has been the use of elevated ripening temperature (Wilkinson, 1993). Temperature effects related to Cheddar ripening have been extensively studied. Grazier, Bodyfelt, McDaniel, and Torres (1991) found that the intensity of most flavour characteristics of experimental cheese increased as a function of time and temperature, while buttery aroma and flavour tended to decrease. Folkertsma, Fox, and McSweeney (1996) found accelerated proteolysis and lipolysis by increasing the ripening temperature of Cheddar cheese. Hannon et al. (2005) ripened Cheddar at elevated temperatures for short periods in the early stages of the ripening process, followed by maturation at 8 °C for the remainder of the ripening period. They found acceleration of the mature flavour attributes in cheeses ripened at elevated temperature compared to control cheeses ripened at 8 °C during the entire ripening period. The casein degradation of Fynbo cheese was characterised for cheeses ripened at 5, 12, and 16 °C (Sihufe, Zorrilla, & Rubilio, 2003). Index of maturation (Water soluble N/Total N) and proteolysis kinetics were highly affected by the ripening temperature. Lawlor, Delahunty, Wilkinson, and Sheehan (2003) analysed the sensory attributes of Swiss-type and Swiss–Cheddar hybrid-type cheeses made using different manufacturing conditions, among them curd wash and ripening temperatures. A difference in ripening temperature (9 and 12 °C) resulted in cheeses with differences in flavour, odour, and texture attributes.

In addition to differences due to different ripening conditions, other phenomena influence the flavour and texture that develop in a cheese as it ripens. In the first place, the fatty acid composition of milk varies throughout the year owing to variation in feed factors. Chilliard and

*Corresponding author. Tel.: +47 51787577; fax: +47 51787501.
E-mail address: hilde.kraggerud@tine.no (H. Kraggerud).

0958-6946/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.idairyj.2007.11.003

Ferlay (2004) reviewed several forage factors that influence the fatty acid composition of cows and goats milk. They particularly emphasised interactions between these factors and found considerable variability in the fatty acid composition. They concluded that there is a need to evaluate more deeply how different feeding strategies influence the nutritional, sensory, and technological aspects of milk fat quality in spite of the fact that several studies on the effects of feed factors on the sensory quality of milk and dairy products had already been published.

The concentration and composition of various milk components differs according to season. In a study by Auldist, Walsh, and Thomson (1988), important properties like protein:fat ratio and casein:whey protein, and solid fat content of bovine New Zealand milk were shown to be affected by time of the year. The composition of New Zealand milk is also affected by very concentrated calving. In contrast, Aigster, Sims, Staples, Schmidt, and O'Keefe (2000) found no differences in the sensory or rheological properties of cheese made from milk with a diet-based increase of oleic acid, compared with normal milk. In Abondance cheese, Bugaud et al. (2001) found that making cheese from milk from different seasons or from animals kept on mountain or valley affected the cheese texture.

In Norway, the diet of cows is determined by the management practice of the farm. During the year, Norwegian cows are obliged to have a minimum of 8 weeks of grazing. In practice, cows will be pasture-fed for a much longer period—varying throughout the country. In the southwest of Norway where this experiment took place, the grazing season would normally be in the period between May and September/October. In the intermediate periods between indoor and outdoor feeding, the cows are given transition feeding to avoid a sudden change from one feeding regime to another. Pasture silage is given indoors in addition to feed concentrate which is given all year. In the cheese-making region where this experiment was conducted, the calving period is spread throughout the year.

The experience of Norwegian cheesemakers is that the speed of ripening and quality of cheese are affected by the season. Maintaining even cheese quality throughout the year is therefore a challenge for the cheesemakers. The objective of this paper was to study how seasonal variation of raw milk and different ripening conditions influence the sensory properties of two Norwegian cheese varieties.

2. Materials and methods

2.1. Experimental design

Two cheese varieties, A and B, were made in duplicate ($2 \times 2 = 4$), over six seasons of the year ($6 \times 4 = 24$). Duplicate samples of each batch were stored at three ripening temperatures ($24 \times 2 \times 3 = 144$), and assessed at three ripening stages ($144 \times 3 = 432$).

Both cheese varieties are washed curd, semi-hard, Dutch-type cheese with similar fat (27%) and dry matter (59%) content. Both cheese varieties were produced from pasteurised milk in two different commercial dairy plants located within a distance of 20 km. The cheeses were made with the same type of DL-starter. For variety A, an adjunct culture of propionic acid bacteria was added.

Cheeses were sampled six times during 1 year, every 8 weeks, in the period from December 2004 to September 2005: (1) December, (2) February, (3) March/April, (4) May, (5) July, and (6) September. The first three samplings were made during the period of indoor feeding and the last three samplings during the grazing period. At each sampling, cheese from two cheese vats from different raw milk bulk tanks were sampled.

The cheese samples were ripened at three different temperatures in the warm room, one of the early stages of the ripening, with a ΔT of 3 °C between each level: $T(-1) < T(0) < T(1)$. All samples were taken out of the warm room around 1 month after production and stored at <4 °C for further maturation. Each sample consisted of three 5-kg cheeses wrapped in plastic foil. After 8, 24, and 40 weeks of maturation, one cheese was analysed. Sampling was carried out in accordance with IDF Standard 50C (IDF, 1995).

2.2. Fatty acid composition

Samples of cream separated from the cheese-milk were analysed in duplicate. Fat was extracted from the samples as described in IDF Standard 5 (IDF, 2004). Triglycerides were methylated with alkaline methanol, according to IDF Standard 182 (IDF, 2002a), and methyl esters analysed according to IDF Standard 184 (IDF, 2002b). The results were given in percent (w/w) of total fatty acids in the triglycerides.

2.3. Sensory analysis

Descriptive sensory analysis was carried out by a panel of six selected, trained assessors; with very few exceptions all of them attended all sessions.

The vocabulary of sensory attributes used in this experiment was developed to obtain the shortest possible list of attributes which would give the most complete description of the semi-hard cheeses. The attributes presented in Table 1 have been in use in this panel for around 20 years. The vocabulary is in accordance with ISO Standard 5492 (ISO, 1992). Each sensory attribute was evaluated on a discrete interval scale from 1 to 9 points, as defined in ISO Standard 4121 (ISO, 2003). Before each session, the panel members participated in a calibration session in order to agree on the use of attributes and scales. Two different cheese samples were used as calibration samples. All assessments were conducted in individual booths at the sensory laboratory, which complies with international standards for design of test rooms, ISO 8589 (ISO, 1988). Samples were tempered to 14 ± 2 °C prior to assessment. In each session, the order of assessment of the

Table 1
Vocabulary of sensory attributes, listed in the order of appearance by the sensory assessment of a product

Class of attributes	Attribute	Explanation	Scale extremes (1–9)
Texture—mechanical	Pressure firmness	Force perceived by compressing the cheese with a finger	Low–high
Texture—mechanical	Shear firmness	Force perceived by parting the cheese with a knife	Low–high
Odour	Odour intensity	Total intensity of odours	Low–high
Odour	Acetic odour	Intensity of acetic odour	None–much
Texture—mechanical	Elasticity	Rapidity and degree of recovery from a deforming force	Low–high
Texture—mechanical	Cohesiveness	Degree to which a substance can be deformed before it breaks	Low–high
Texture—mechanical	Firmness on chewing	Force perceived on compressing product between the teeth	Low–high
Texture—mechanical	Pasty	Force required to remove material that adheres to the mouth.	Not–very
Texture—mechanical	Solubility	Effort to disintegrate product until ready for swallowing	Low–high
Texture—surface	Dry	Water absorbed by or released from the product	Not–very
Texture—geometrical	Floury	Perception of small particles in a texture	Not–very
Texture—geometrical	Grainy	Perception of moderate sized particles in a texture	Not–very
Flavour	Flavour intensity	Perceived total intensity of flavour	Low–high
Flavour	Aromatic	Flavour with a pleasant annotation	Not–very
Flavour	Nutty	Flavour of nuts	Not–very
Flavour	Malty	Flavour of malt	Not–very
Taste	Sweet	Sweet taste	Not–very
Taste	Sour	Complex Sensation, generally due to the presence of organic acids.	Not–very
Taste	Salty	Taste of salt	Not–very
Trigeminal	Pungent	Sharp sensation of the buccal and nasal mucous membrane	Not–very
Taste	Bitter	Taste of bitter	Not–very
Flavour	Sulphurous	Sulphurous flavour	Not–very

samples was randomised, also with respect to the age of the cheese. The scores of each sample were averaged over all assessors.

2.4. Analysis of data

Analysis of variance (ANOVA) was carried out using a general linear model procedure of the software package Minitab® 15.1. (MINITAB Inc., State College, PA, USA).

Multivariate data analyses were performed using Unscrambler version 9.6 (CAMO Process AS, Oslo, Norway) using methods like principal component analysis (PCA) and partial least squares regression (PLS) (Martens & Næs, 1989). In general, all the variables in the multivariate analysis were given the same weight because the data sets used had variables with uniform scales. Design variables included in the multivariate analysis were treated as dummy variables where each level of a factor was represented by one variable with level 0 or 1. An exception from this was age, which was treated as a continuous variable. Design variables included in the analysis were passified, given a very low weight, in order not to influence the results (Martens & Martens, 2001). Validation of models was done using cross validation (Martens & Næs, 1989).

Scores from the PCA/PLS analysis describe the data structure in terms of sample patterns and more generally show sample differences or similarities. Each sample has a score on each principal component (PC). It reflects the sample location along that PC; i.e., the coordinate of the sample on the PC. Loadings from PCA/PLS describe the data structure in terms of variable correlations. The loading vectors are linear combinations of the original variables and each variable is assigned a loading value for each loading vector. This value reflects both how much the variable contributed to that particular component and correspondingly how much of the original variation is explained by that PC. In a correlation loadings plot (Martens & Martens, 2001), the 50% and 100% explained variance limits are marked with circles.

3. Results and discussion

3.1. Fatty acid composition of milk

The fatty acid composition of the triglycerides was measured in raw milk from all batches from which cheese was made, in total 24 samples. Fig. 1a shows the distribution of the samples from seasons 1 to 6 in a PCA score plot. The first two PCs explained 98% of the variation in the 24 samples. The difference between pasture feeding (seasons 4, 5, 6) vs indoor feeding (seasons 1, 2, 3) was the most important factor associated with a distinction between the samples. Although the second PC did not explain more than 3% of the variation, it revealed a year-round variation in which seasons were ordered in sequence around the year with respect to the fatty acid composition. The transition feeding between feeding regimes and changes in the pasture quality during the outdoor season may explain this structure.

Table 2 shows the average fatty acid composition for indoor and outdoor feeding seasons; there were significant differences for most of the fatty acids. The main differences were found in C16:0 and C18:1. The correlation loadings

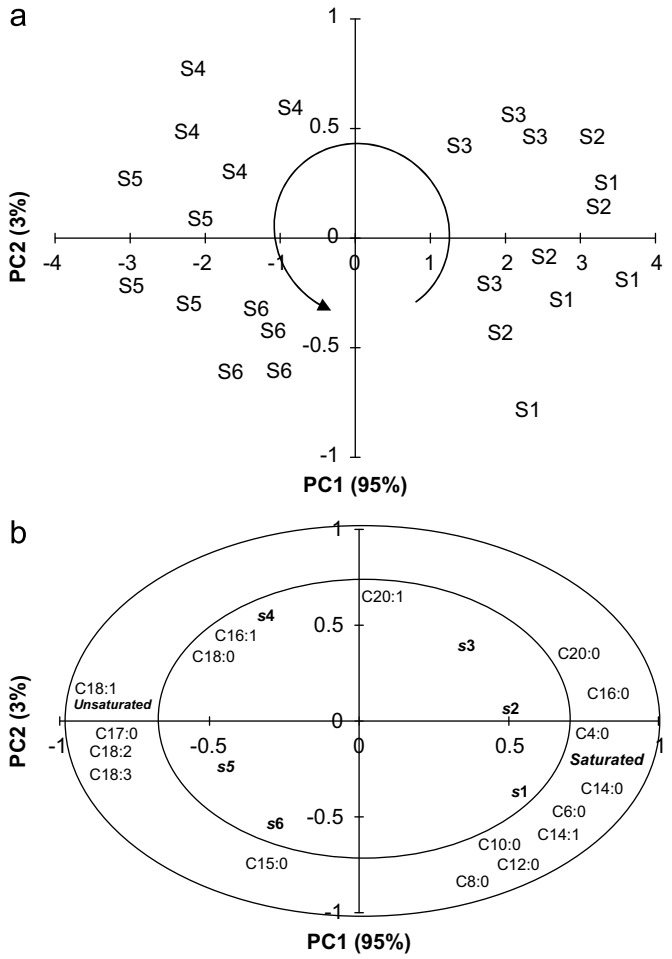

Fig. 1. (a) PCA scores PC1 vs PC2 for fatty acid distribution from the 24 raw milk batches used in cheesemaking. Each point is marked with a season: S1, December; S2, February; S3, March/April; S4, May; S5, July; S6, September. (b) PCA correlation loadings PC1 vs PC2 for fatty acid distribution from the 24 raw milk batches used in cheesemaking. In addition, the following passified variables are shown: *s*1, December; *s*2, February; *s*3, March/April; *s*4, May; *s*5, July; *s*6, September and sum of saturated and unsaturated fatty acids, respectively.

corresponding to the score plot of Fig. 1a are shown in Fig. 1b. The sum content of saturated and unsaturated fatty acids, respectively, was computed, included as passified variables in the model and, as shown in Fig. 1b, saturated vs unsaturated fatty acids represent the main variation along PC 1. Samples from indoor feeding were dominated by the shorter chain saturated fatty acids C4:0–C16:0, and also C14:1, while milk from the grazing season was dominated by fatty acids C18 (both saturated and unsaturated), C16:1, and C17:0. PC 2 was distinguished by C20:1 (late spring/early summer), as opposed to C15:0, C8:0–C12:0, and C14:1 (late autumn/early winter). Chilliard and Ferlay (2004) reported that intake of pasture results in an effect similar to C18 fatty acid supplementation. The effects give an increase in C18 fatty acids, both saturated and unsaturated, and a simultaneous decrease of C6–C14 fatty acids. This was confirmed by the present data. In another study (Rego et al., 2004), pasture feeding was compared with total mixed ration feeding and a similar pattern was found as in our data: pasture gave an increase in C17:0 and C18 in the milk, while the fatty acids C10–C16 decreased in milk with grazing. As in the present study, no differences in the content of C15:0 and C16:1 were found in the milk.

Table 2
Fatty acid distribution in raw milk samples during indoor (seasons 1–3) and outdoor (seasons 4–6) feeding in percentage (w/w) of total ($n = 24$)

Fatty acid	Feeding season		Significance[a]
	Indoor	Outdoor	
C04:0	4.21	4.09	***
C06:0	2.43	2.35	***
C08:0	1.31	1.29	n.s.
C10:0	2.77	2.74	n.s.
C12:0	3.12	3.06	n.s.
C14:0	10.77	10.07	***
C14:1	0.91	0.84	**
C15:0	1.45	1.51	*
C16:0	29.48	26.25	***
C16:1	1.30	1.36	***
C17:0	1.30	1.41	***
C18:0	11.49	11.80	*
C18:1	23.63	26.43	***
C18:2	1.90	2.35	***
C18:3	0.43	0.57	***
C20:0	0.18	0.17	***
C20:1	0.07	0.07	n.s.

[a]n.s.—not significant; $^{*}p = 0.05$; $^{**}p = 0.01$; $^{***}p = 0.001$.

3.2. Sensory attributes in cheese related to temperature and age

Principal component analysis was performed using sensory attributes as variables and with all 432 samples included. In addition, the factors of the experiment were included as passified variables to make interpretation of the effects easier. The sample scores and loadings from the two first components are shown in Fig. 2a and 2b, respectively. PC 1 explained 58% of the variation in sensory properties while PC 2 explained 20%. The first PC mainly explained the difference between cheese varieties, while PC 2 mainly explained the effect of maturation time. In other words, the cheese variety influenced the sensory quality more than age. The difference between the two varieties apparently had a large effect on the sensory attributes. Fig. 2a also shows that variety A had a much broader sensory variation than variety B within each maturation time. After 8 weeks of maturation, a slight overlapping between the cheese varieties was observed, in which a few cheeses of variety A had the same sensory characteristics as B. After 24 and 40 weeks of maturation, the cheese varieties were completely separated from each other. The correlation loading plot (Fig. 2b) shows that the experimental factors cheese variety and age were well explained by the model. The effect of ripening temperature seemed to vary in a systematic

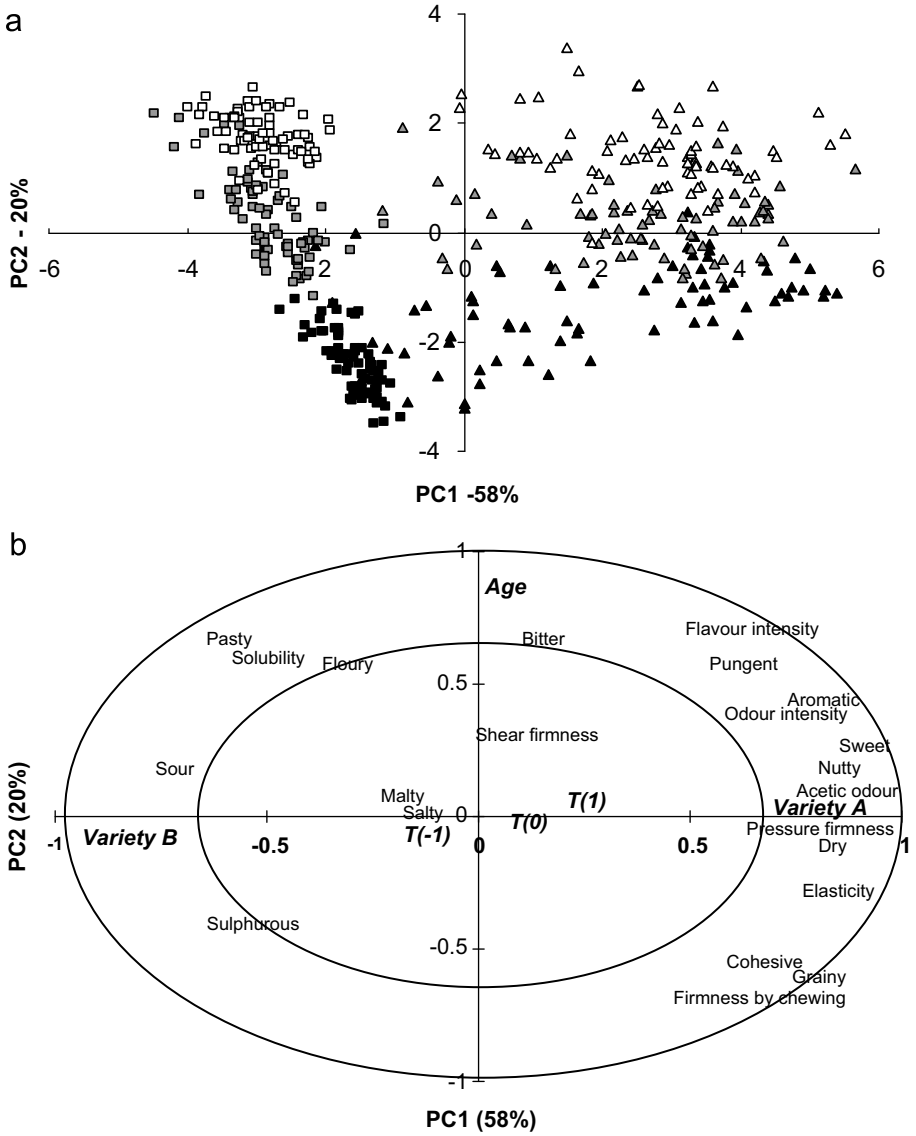

Fig. 2. (a) PCA score plot PC1 vs PC2 for all 22 sensory descriptors of all 432 samples. Markers shape cheese variety: ▲, A; ■, B. Colours of the symbols represent maturation time: black, 8 weeks; grey, 24 weeks; and white, 40 weeks. (b) PCA correlation loadings PC1 vs PC2 for the same sample and variable set as in part (a). In addition, the following passified variables are shown: maturation time = age; temperatures, $T-1$, $T0$, $T1$; and cheese varieties, A and B.

fashion, but was not as well explained by the model. Cheeses of variety B were characterised by higher scores in the flavour attributes sulphurous and sour. Variety A had higher scores on aromatic, sweet, nutty, and acetic odour, as could be expected when adding propionibacteria (Frölich-Wyder & Bachmann, 2004). With respect to the texture attributes, variety B was more pasty, soluble, and floury, while A was more dry, cohesive, grainy, elastic, and firm on chewing. The flavour properties increasing with age were flavour intensity, pungent, and bitter, while the attribute sulphurous decreased. The general increase of flavour attributes with age is caused by the production of a wide range of volatile compounds during maturation, by metabolism of triglycerides and proteins (McSweeney, 2004). Proteolysis contributes to changes in texture of cheese during maturation, due to the breakdown of the protein network, reduced water activity, and increase in pH, and could also lead to increased bitterness due to bitter peptide formation (Sousa, Ardö, & McSweeney, 2001). During ripening, mechanical properties like firmness by chewing, cohesiveness, and elasticity decreased, while solubility increased. Sensory attributes involving small particles floury and pasty increased, while grainy, involving medium-sized particles, decreased.

Since the dominating variation was related to cheese variety, the two cheese varieties were studied in more detail, as illustrated in Figs. 3 and 4 for varieties A and B, respectively. For variety A, 69% of the sensory variation in the 214 samples was explained by the first two PCs. Fig. 3a shows the average scores for each time/temperature combination; each point is the average of 24 samples. The ripening temperature was the main factor influencing

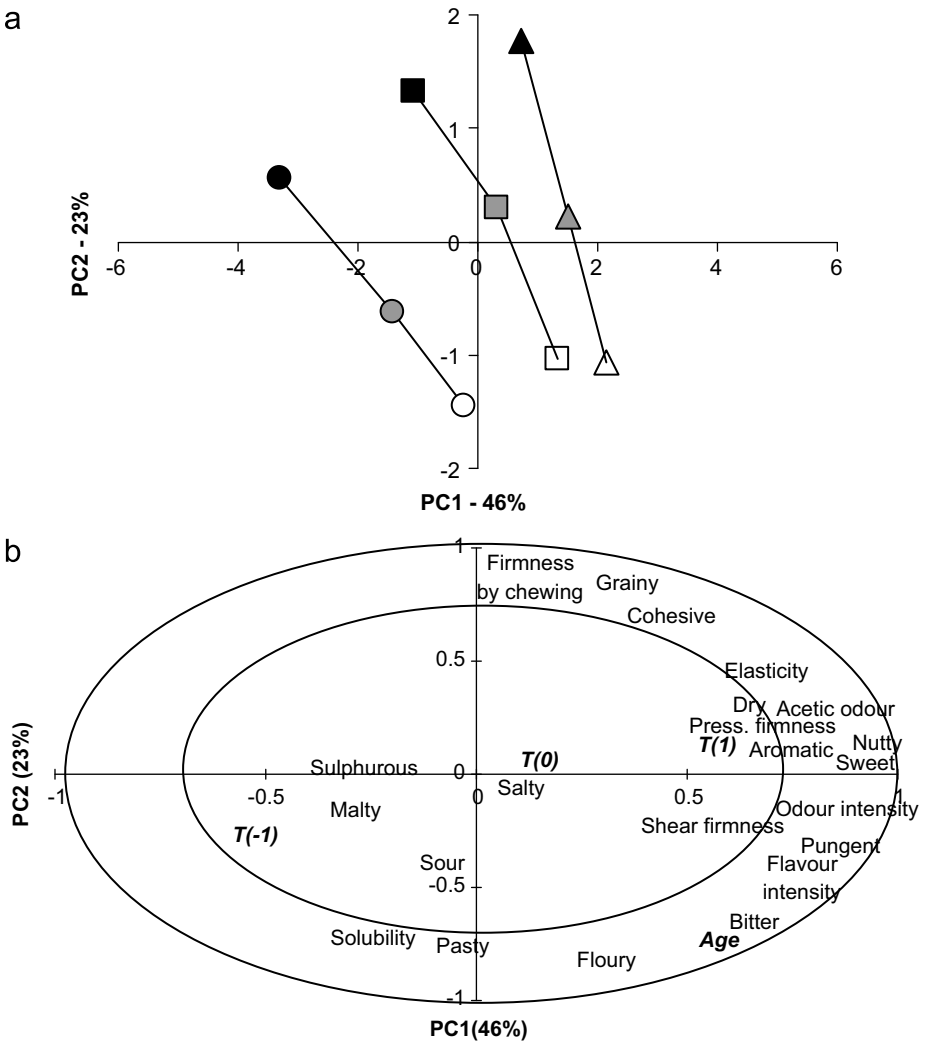

Fig. 3. (a) PCA score plot for cheese variety A. Each point is the average of 24 samples. Symbols representing ripening temperatures: ●, $T(-1)$; ■, $T(0)$; ▲, $T(1)$. Colours of the symbols represent maturation stage: black, 8 weeks; grey, 24 weeks; and white, 40 weeks. (b) PCA correlation loading plot for cheese variety A, showing loadings of all sensory descriptors in principal components 1 (PC1) and 2 (PC2). In addition, the following passified variables are shown: maturation time = age, and temperatures: $T-1$, $T0$, $T1$.

the variation of the sensory properties of variety A (PC 1 = 46%), while maturation time dominated in PC 2. An increased ripening temperature was positively correlated with odour and flavour intensity, aromatic, nutty, sweet, pungent, and acetic odour (Fig. 3b). Cheeses ripened at higher temperatures were more dry, elastic, and had higher pressure and shear firmness. With age, the attributes solubility, pasty, and floury increased, while firmness on chewing, cohesion, and grainy decreased with age. Fig. 3b shows that regarding total sensory properties, the ripening temperature effect was nearly perpendicular to the age effect. This demonstrated that the two effects were independent of each other. With longer maturation time, the effect of the ripening temperature was more and more equalised. From Fig. 2, we saw that the differences between cheeses at the start of ripening of this particular cheese variety were large, compared to cheese B.

Results from PCA of sensory properties for the 216 samples of cheese B are shown in Fig. 4. Of the variation, 72% was explained in the two first PCs. Referring to Fig. 2, the total variation in this cheese variety was lower than in variety A. The maturation time was the dominating factor for distinguishing between the samples on the sensory properties of this cheese variety, as seen from the scales of the score plot, Fig. 4a. The influence of ripening temperature on the sensory attributes increased with age, although the different temperatures were applied at an early stage of the ripening process. This is in contrast with variety A. The scattering of the cheese B samples at 8 weeks was much less (Fig. 2a) which may explain this phenomenon. The flavour attributes—flavour intensity, bitter, aromatic, and pungent—increased with age, while sulphurous decreased with longer maturation time. The samples tended to be less cohesive, elastic, grainy, dry, and firm on chewing with increasing maturation time as opposed to the

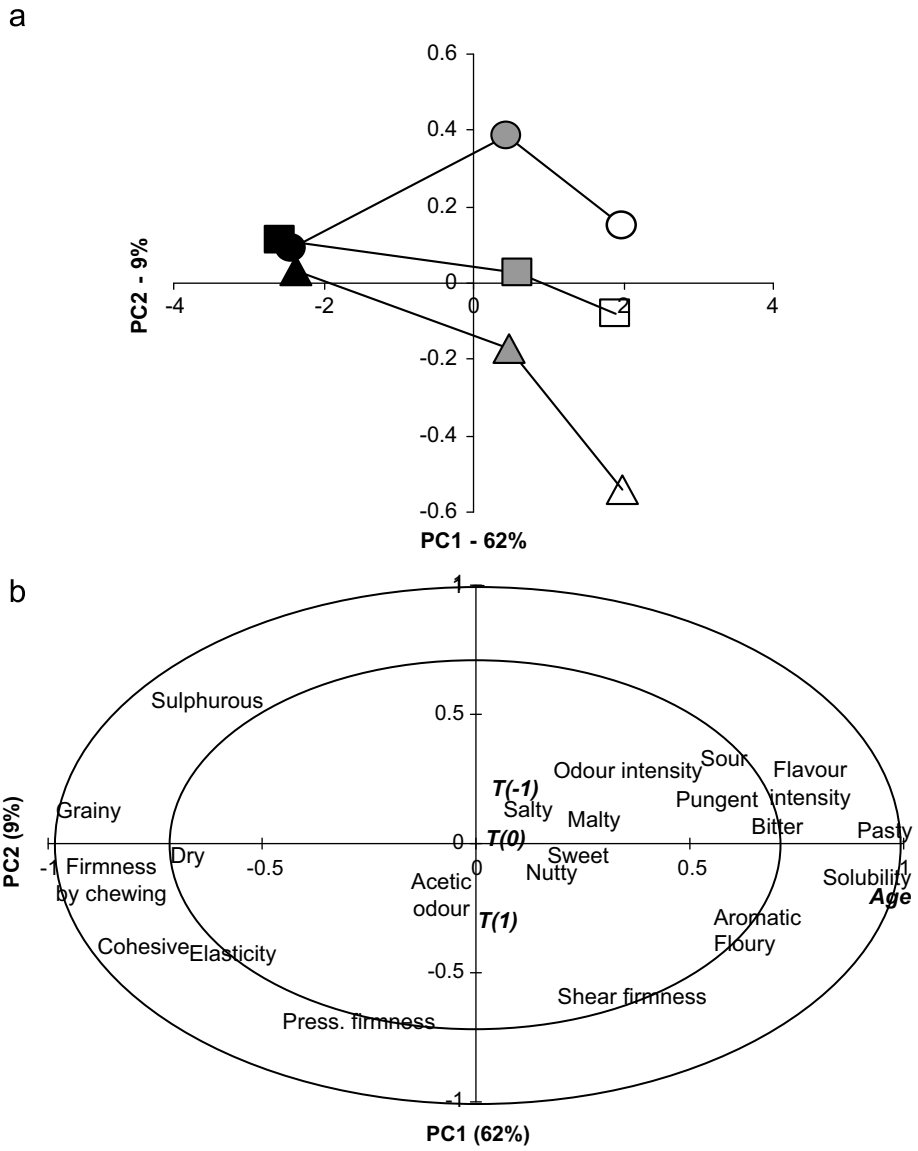

Fig. 4. (a) PCA score plot for cheese variety B. Each point is the average of 24 samples. Symbols representing ripening temperatures: ●, $T(-1)$; ■, $T(0)$; ▲, $T(1)$. Colours of the symbols represent age/maturation time: black, 8 weeks; grey, 24 weeks; and white, 40 weeks. (b) PCA correlation loading plot for cheese variety B, showing loadings of all sensory descriptors in principal components 1 (PC1) and 2 (PC2). In addition, the following passified variables are shown: maturation time = age, and ripening temperatures: $T-1$, $T0$, $T1$.

attributes pasty, solubility, and floury which were increased by age. Pressure and shear firmness were positively correlated to ripening temperature. These changes are in accordance both with the general effects from Fig. 2 and the effects in cheese A, Fig. 3. Also, in this case we see that age and ripening temperature effects were independent.

Increased flavour intensities with higher ripening temperature were also observed by Shakeel-Ur-Rehman et al. (2000). They found higher concentrations of most volatile compounds with increased temperature. Hannon et al. (2005) performed an experiment similar to that in this study in which temperatures were maintained for a shorter period at an early stage of ripening before maturation at uniform temperature. They found that after 8 months, the cheeses were associated with typically mature attributes like rancid, bitter, astringent, high strength of flavour, and a pungent odour, which is in good accordance with our findings. In contrast to our results, they also found that ripening could be accelerated by 2 months by the elevation of temperature for a shorter period of time, showing similar flavour profiles after 6 months as the control cheeses after 8 months. Bertola, Califano, Bevilacqua, and Zaritzky (2000) found that the development of different instrumental rheology measurements in Gouda cheese was accelerated by increasing the storage temperature. This would also often be the case if we look at sensory attributes one by one. But, looking at the total sensory properties as in our experiment, differences originating from the ripening temperature were found to be independent of the age effects. Temperature differences were maintained or even intensified during maturation, which makes temperature elevation unsuitable as the only means of acceleration of the ripening.

Amárita et al. (2002) used a cheese model slurry with different strains of *Lactococcus lactis* to study the production of different sulphur compounds from methionine. Sensory analysis indicated a characteristic methional aroma (like cooked potato) in the beginning, but as incubation proceeded, the intensity of methional aroma decreased and samples were judged by the panel tasters as developing a cheese-like flavour. Fig. 5 illustrates the decrease of sulphurous flavour with age in variety B cheeses. This may indicate the presence of sulphur-tasting intermediate compounds which were further degraded with the maturation of the cheese. Shakeel-Ur-Rehman et al. (2000) found a reduction of sulphur/eggy flavour with increasing age. This corresponds well with our findings.

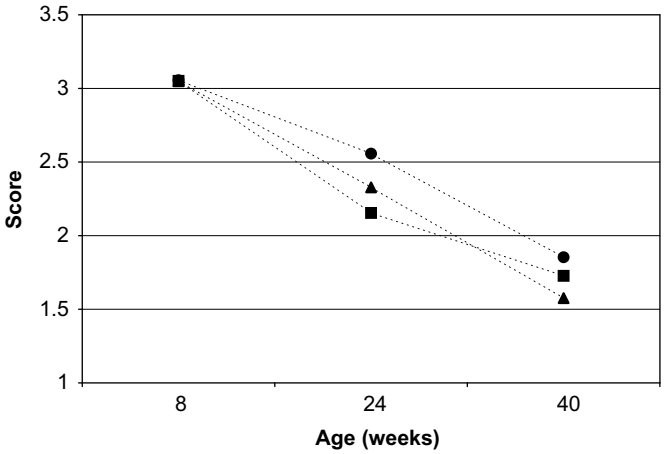

Fig. 5. Average sensory score of the sensory flavour attribute sulphurous in cheese variety B during maturation for each ripening temperature: symbols ●, $T(-1)$; ■, $T(0)$; ▲, $T(1)$.

3.3. Correlations between fatty acids and sensory attributes of cheese

To avoid interference from the effects of ripening temperature, age and cheese variety, the correlation between the fatty acid profiles of milk with sensory quality of cheese was made with a selection of 72 cheeses. We chose variety B at 8 weeks of age since the temperature differences were minimal at that maturation time. Fig. 6 shows a PLS2 loading plot in which fatty acids were used as explanatory variables (X-matrix), while the 22 sensory attributes were chosen as explained variables (Y-matrix). In the two first PCs, 27% of the total sensory variation was explained by the fatty acid composition. Shear firmness was the only sensory variable explained to a level more than 50% in the first two PCs. This sensory attribute was related to seasons 1 and 2 and indoor feeding and associated with saturated fatty acids C6-C14, and with C14:1 and C20. There are also other variables fairly well explained in the first two factors; for example, flavour intensity on the opposite side of the plot with respect to seasons. This sensory attribute was closely related to seasons 4 and 5 and associated with C18 saturated and unsaturated fatty acids and with C16:1 and C17:0. Modelling these two sensory variables one by one with PLS1, illustrated in Figs. 7a and 7b, provided fairly good regression models. The correlation coefficient of the sensory attribute shear firmness vs the fatty acids was 0.91 (Fig. 7a), while the corresponding correlation coefficient for flavour intensity was 0.86 (Fig. 7b).

The differences indicated in seasonal variation are well in accordance with earlier findings. Bugaud et al. (2001) also found differences related to forage in Abondance cheese but no differences related to sensory attributes were found to correlate with fatty acids. However, they found a

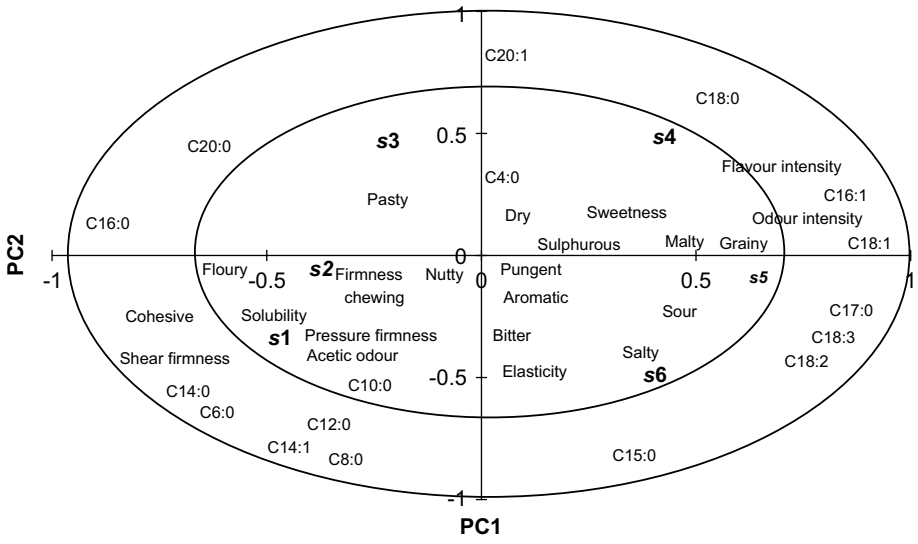

Fig. 6. PLS correlation loadings plot. Selected samples in the regression: cheese variety B; maturation time, 8 weeks (72 samples). *X*-variables were fatty acid distribution, *Y*-variables the 22 sensory variables. Dummy variables shown are seasons: *s*1, December; *s*2, February; *s*3, March/April; *s*4, May; *s*5, July; *s*6, September.

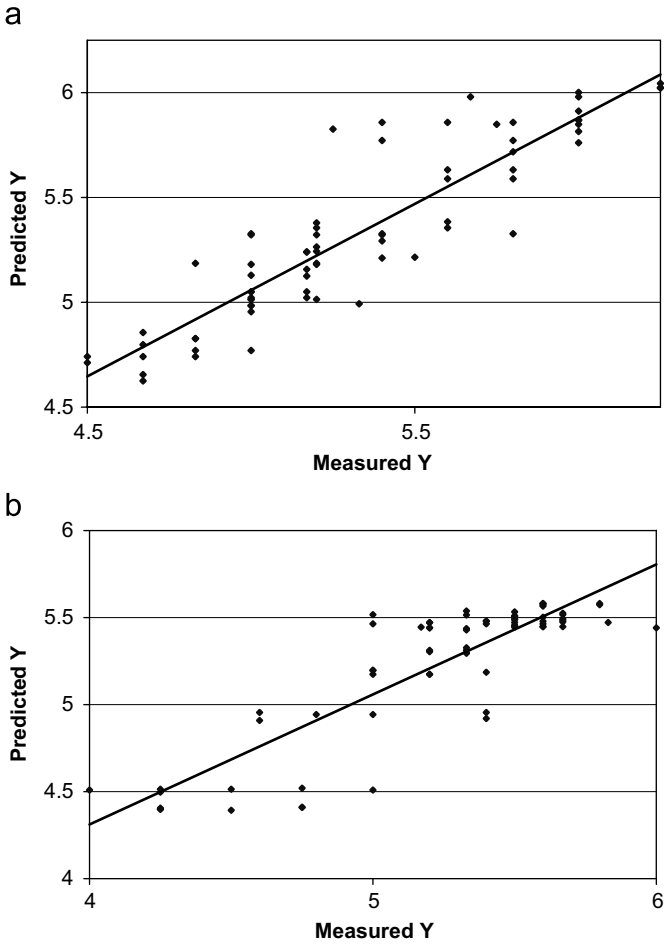

Fig. 7. (a) PLS regression cross-validated prediction model, based on cheese variety B and maturation time 8 weeks. *X*-variables: fatty acid distribution. *Y*-variable: sensory attribute shear firmness. Correlation coefficients: calibration, 0.91; validation, 0.88. (b) PLS regression cross-validated prediction model, based on cheese variety type2; maturation time, 8 weeks. *X*-variables: fatty acid distribution. *Y*-variable: sensory attribute odour intensity. Correlation coefficients: calibration, 0.86; validation, 0.82.

correlation between fracture strain and proportion of C18 unsaturated fatty acids. The fracture modulus (fracture stress/fracture strain) has been found to correlate with firmness, springiness, and adhesiveness (Foegeding, Brown, Drake, & Daubert, 2003). Fracture strain would then be expected to be negatively correlated with those sensory attributes which makes sense in the data shown in Fig. 7a where attributes like shear and pressure firmness, cohesion, and pasty are opposite to the C18 unsaturated fatty acids.

The influence of the composition of the three different types of Alpine highland pasture with respect to rheological, chemical, and sensory properties of cheese was studied by Buchin, Martin, Dupont, Bornard, and Achilleos (1999). They found that the type of pasture did not influence fatty acid composition of the milk. However, they found differences in texture and flavour properties explained by primary proteolysis which they assumed could be due to different levels of plasmin/plasminogen content of the batches of milk used. The volatile compounds were consistent with the flavour of the cheeses and were assumed to have their origin in terpenes and xylenes from the plants in the diet, or from other sources that we could not identify. This may indicate that the differences in flavour attributes in the present experiment according to seasons could also originate from other sources than the fatty acids, but may be correlated with them during the grazing season.

4. Conclusions

The PCA models derived from the sensory attributes of the cheeses demonstrated that the factors ripening temperature and maturation time had systematic, independent effects on the sensory attributes of both cheese varieties examined. Effects of increased ripening temperature on sensory properties were different for the two cheese varieties. In variety A, temperature influenced the sensory properties more than in variety B, probably associated with the addition of propionibacteria culture to variety A. With increasing age, cheeses were typically more bitter, pungent, and intense in flavour. In variety B, we found a linear decrease of sulphurous aroma intensity during maturation from 8 to 40 weeks. Texture properties associated with casein breakdown, like firmness by chewing, cohesiveness, and elasticity, decreased with age. Sensory attributes like floury and grainy indicated a decrease in particle size during maturation.

Fatty acid composition of raw milk varied with seasons, showing a continuous trend throughout the year, with the main differences found between indoor feeding and outdoor feeding seasons. Correlations between fatty acid composition of the raw milk and some of the sensory properties of cheese were observed. Firmness was correlated with winter season and unsaturated fatty acids; in contrast, flavour intensity was found to be higher in the grazing season.

References

Aigster, A., Sims, C., Staples, C., Schmidt, R., & O'Keefe, S. F. (2000). Comparison of cheeses made from milk having normal and high oleic fatty acid compositions. *Journal of Food Science*, *65*, 920–924.

Amárita, F., Martínez-Cuesta, C. M., Taborda, G., Soto-Yárritu, P. L., Requena, T., & Peláez, C. (2002). Formation of methional by *Lactococcus lactis* IFPL730 under cheese model conditions. *European Food Research and Technology*, *214*, 58–62.

Auldist, M. J., Walsh, B. J., & Thomson, N. A. (1988). Seasonal and lactational influences on bovine milk composition in New Zealand. *Journal of Dairy Research*, *65*, 401–411.

Bertola, N. C., Califano, A. N., Bevilacqua, A. E., & Zaritzky, N. E. (2000). Effects of ripening conditions on the cheese texture of Gouda cheese. *International Journal of Food Science and Technology*, *35*, 207–214.

Buchin, S., Martin, B., Dupont, D., Bornard, A., & Achilleos, C. (1999). Influence of the composition of Alpine highland pasture on the chemical, rheological and sensory properties of cheese. *Journal of Dairy Research*, *66*, 579–588.

Bugaud, C., Buchin, S., Nõel, Y., Tessier, L., Pochet, S., Martin, B., et al. (2001). Relationships between Abondance cheese texture, its

composition and that of milk produced by cows grazing different types of pastures. *Le Lait*, *81*, 593–607.

Chilliard, Y., & Ferlay, A. (2004). Dietary lipids and forages interactions on cow and goat milk fatty acid composition and sensory properties. *Reproduction Nutrition Development*, *44*, 467–471.

Foegeding, E. A., Brown, J., Drake, M. A., & Daubert, C. R. (2003). Sensory and mechanical aspects of cheese texture. *International Dairy Journal*, *13*, 585–591.

Folkertsma, B., Fox, P. F., & McSweeney, P. L. H. (1996). Accelerated ripening of cheddar cheese at elevated temperatures. *International Dairy Journal*, *6*, 1117–1134.

Fox, P. F., & McSweeney, P. L. H. (2004). Cheese, an overview. In P. F. Fox, P. L. H. McSweeney, T. M. Cogan, & T. P. Guinee (Eds.), *Cheese: Chemistry, Physics and Microbiology: General Aspects*, (3rd edn., Vol. 1, pp. 1–18). London, UK: Elsevier.

Frölich-Wyder, M. T., & Bachmann, H. P. (2004). Cheeses with propionic acid fermentation. In P. F. Fox, P. L. H. McSweeney, T. M. Cogan, & T. P. Guinee (Eds.), *Cheese: Chemistry, Physics and Microbiology: Major Cheese Groups*, (3rd edn., Vol. 2, pp. 141–156) London, UK: Elsevier.

Grazier, C. L., Bodyfelt, F. W., McDaniel, M. R., & Torres, J. A. (1991). Temperature effects on the development of Cheddar cheese flavor and aroma. *Journal of Dairy Science*, *74*, 3656–3668.

Hannon, J. A., Wilkinson, M. G., Delahunty, C. M., Wallace, J. M., Morrissey, P. A., & Beresford, T. P. (2005). Application of descriptive sensory analysis and key chemical indices to assess the impact of elevated ripening temperatures on the acceleration of Cheddar cheese ripening. *International Dairy Journal*, *16*, 263–273.

IDF (1995). *Milk and milk products—guidance on sampling*. IDF Standard 50C. Brussels, Belgium: International Dairy Federation.

IDF (2002a). *Milk fat—preparation of fatty acid methyl esters.* IDF Standard 182. Brussels, Belgium: International Dairy Federation.

IDF (2002b). *Milk fat—the determination of the fatty acid composition by gas–liquid chromatography*. IDF Standard 184. Brussels, Belgium: International Dairy Federation.

IDF (2004). *Cheese and cheese products—determination of fat content*. IDF Standard 5C. Brussels, Belgium: International Dairy Federation.

ISO (1988). *Sensory analysis—general guidance for the design of test rooms.* ISO Standard 8589. Geneva, Switzerland: International Organization for Standardization.

ISO (1992). *Sensory analysis—vocabulary.* ISO Standard 5492. Geneva, Switzerland: International Organization for Standardization.

ISO (2003). *Sensory analysis—guidelines for the use of quantitative response scales.* ISO Standard 4121. Geneva, Switzerland: International Organization for Standardization.

Lawlor, J. B., Delahunty, C. M., Wilkinson, M. G., & Sheehan, J. (2003). Swiss-type and Swiss–Cheddar hybrid-type cheeses: Effects of manufacture on sensory character and relationships between sensory attributes and volatile compounds and gross compositional constituents. *International Journal of Dairy Technology*, *56*, 39–51.

Martens, H., & Martens, M. (2001). *Multivariate analysis of quality—an introduction*. Chichester, UK: Wiley.

Martens, H., & Næs, T. (1989). *Multivariate calibration*. Chichester, UK: Wiley.

McSweeney, P. L. H. (2004). Biochemistry of cheese ripening. Introduction and overview. In P. F. Fox, P. L. H. McSweeney, T. M. Cogan, & T. P. Guinee, (Eds.), *Cheese: Chemistry, physics and microbiology: General aspects*, (3rd edn., Vol. 1, pp. 47–360). London, UK: Elsevier.

Rego, O. A., Rosa, H. J. D., Portugal, P. V., Franco, T., Vouzela, C. M., Borba, A. E. S., et al. (2004). The effects of supplementation with sunflower and soybean oils on the fatty acid profile of milk fat from grazing dairy cows. *Animal Research*, *54*, 17–24.

Shakeel-Ur-Rehman, Banks, J., Brechany, E. Y., Muir, D., McSweeney, P. L. H., & Fox, P. F. (2000). Influence of ripening temperature on the volatiles profile and flavour of Cheddar cheese made from raw or pasteurized milk. *International Dairy Journal, 10*, 55–65.

Sihufe, G. A., Zorrilla, S. E., & Rubilio, A. C. (2003). Casein degradation of Fynbo cheese salted with NaCl/KCl brine and ripened at various temperatures. *Journal of Food Science*, *68*, 117–123.

Sousa, M. J., Ardö, Y., & McSweeney, P. L. H. (2001). Advances in the study of proteolysis during cheese ripening. *International Dairy Journal*, *11*, 327–345.

Wilkinson, M. G. (1993). Acceleration of cheese ripening. In P. F. Fox (Ed.), *Cheese: Chemistry, physics and microbiology: General aspects*, (2nd edn., Vol. 1, pp. 523–556). London, UK: Elsevier.

Paper II

doi: 10.1111/j.1471-0307.2009.00478.x

ORIGINAL RESEARCH

X-ray images for the control of eye formation in cheese

H KRAGGERUD,*[1,2] J P WOLD,[3] M HØY[3] and R K ABRAHAMSEN[1]

[1]*Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, N-1432 Ås, Norway,* [2]*TINE R & D Center, Box 50, N-4358 Kleppe, Norway, and* [3]*Nofima Mat, Norwegian Food Research Institute N-1432 Ås, Norway*

There is demand for non-destructive monitoring of eye formation in cheese during ripening. The objective of this work was to develop a simple method based on existing equipment in the dairy industry. Images were acquired using a conventional, low resolution online X-ray instrument. Image processing methods for detecting eyes of cheese and measuring volume and size distribution were developed. Sufficient detection of overlapping eyes was obtained. Semihard cheese with propionibacteria ripened under different conditions was analysed. The method was found promising for quality control as it will make possible non-destructive monitoring of eye formation of cheese throughout the ripening period.

Keywords X-ray, Image processing, Eye formation, Cheese, Ripening temperature.

INTRODUCTION

The size of the eyes is an important quality parameter for large-eyed cheese varieties. Product specifications typically contain measures of the number and the size of the eyes. Cheese makers try to control this quality parameter by evaluating the cheese during warm room ripening, traditionally done by listening to the sound made while tapping the surface of the cheese, in combination with 'looking inside'. This can be done by cutting the cheese in two halves, or by taking out a small cylinder from the cheese using a cheese trier. In both cases the cheese cannot be used for normal commercial sale after examination. A major disadvantage is also that only a small part of the cheese volume is examined.

Fermentation of lactose by heterofermentative lactic acid bacteria in the starter culture is a source of CO_2 formation in the cheese during cheese-making and ripening. The quantity of gas formed in this type of fermentation is relatively small. In cheese varieties with propionic acid bacteria as an adjuct culture, the classic propionic acid fermentation was described as essential for the formation of CO_2 over 100 years ago, and thus essential for the development of characteristic eyes in such cheese varieties (von Freudenreich and Orla-Jensen 1906). However, the metabolism of propionibacteria in cheese is complicated and not yet fully understood. The two main metabolic pathways of CO_2 formation from lactate by propionibacteria are the classic propionic acid fermentation, in which lactate is converted to propionate, acetate and CO_2, and amino acid catabolism, a pathway that was discovered later (Brendehaug and Langsrud 1985: Frölich-Wyder and Bachmann 2004).

Supersaturation with CO_2 gas is needed for eye formation and can be achieved when the rate of CO_2 production is relatively fast and the rate of diffusion out of the cheese is slow. Part of the CO_2 gas will remain in the eyes of the cheese, some CO_2 will remain dissolved in the body of the cheese and some will diffuse out of the cheese loaf (van den Berg *et al*. 2004). When vacuum packaging in plastic films is used during ripening of rindless cheese, the diffusion rate of gas through the film will affect the gas pressure and thereby possibly influence eye formation.

Gas formation by propionibacteria primarily takes place during a warm room stage of ripening. When the development of eyes is sufficient, propionic acid formation is retarded by cooling the cheese to a lower temperature (Frölich-Wyder and Bachmann 2004).

X-rays have been used for imaging for many years. The best known use of X-rays is within medical diagnostics but they are also used extensively in industry and other areas. The intensity of the X-rays is modified by absorption by the material they are passing through and the resulting energy is captured by a detector to form an image in grey scales (Gonzales and Woods 2001).

Image processing techniques have been used increasingly for food quality evaluation in recent years. Image features like colour, size, shape and texture have been applied in various food-monitoring applications. Different methods of image acquisition, using digital cameras, ultrasound,

*Author for correspondence.
E-mail: hilde.kraggerud@tine.no

magnetic resonance imaging, computed tomography (CT), X-rays and other methods, are applied. There are different approaches for preprocessing, various segmentation techniques and statistical methods to be used for different purposes (Gonzales and Woods 2001; Du and Sun 2004; Zheng *et al*. 2006).

Non-destructive analysis of eye formation in cheese would be useful for cheese-makers, enabling them to make better decisions on the time for transferring cheese from the warm room to the cold room for further ripening. Thus, one will be able to measure the growth rate of the eyes with the possibility of making measurements on the same cheese loaf several times without disrupting the ripening. X-ray instruments are frequently used in the food industry, for instance for detection of foreign bodies and detection of unusual shapes and colours (Haff and Toyofuko 2008).

Strand (1985) used X-ray CT eight times during the ripening period of cheese in both the warm room and the cold room, saved images of the inner layer of the cheeses and measured the area of the eyes on this inner layer. Images of the cheeses showed the development of eyes during ripening very clearly and calculation of the area of the eyes for a given layer was possible. CT equipment, however, is still very expensive and complicated in use, which makes it impractical to use other than for research purposes.

Eskelinen *et al*. (2007) have carried out a feasibility study on ultrasonics for non-destructive testing in structural quality control of Finnish Emmental cheese. They demonstrated that monitoring of gas–solid structure and characterization of cheese eyes was possible with the use of an ultrasonic method. Magnetic resonance imaging (MRI) was evaluated for measuring eye formation and structural quality of US Swiss-type cheese by Rosenberg *et al*. (1992). The possibility of following eye formation and development of structural defects by MRI was demonstrated. The possibility of finding cracks and assessing maturation time in cheese by an acoustic impulse-response technique and ultrasound has also been described (Benedito *et al*. 2006; Conde *et al*. 2008).

The objective of this work was to develop a simple method for non-destructive monitoring of eye formation in cheese during ripening with use of a conventional online X-ray instrument. This involves image processing methods for measuring the development in number and volume of eyes during ripening of cheese. In particular, the challenge of separating overlapping eyes in the X-ray images was addressed. To evaluate the potential of the method, cheese was exposed to different ripening conditions. Ripening temperature and permeability of plastic film were varied in order to introduce systematic differences.

MATERIALS AND METHODS

Cheese sampling

Cheese was produced in a commercial dairy plant and sampled directly after brining at the time of packaging in plastic film. The product studied was a variety of Norwegian washed curd, semihard cheese with 27% fat in the product, made from pasteurized milk. The starter culture used was DL-culture with an adjunct culture of propionic acid bacteria. Cheese blocks of around 10 kg were vacuum wrapped in plastic bags after brining. Each cheese was then packed in a cardboard container and pre-ripened for a short period at relatively low temperature. The cheeses were then transferred to the warm room where the main eye formation took place and where our measurements were conducted.

Experimental design

The cheeses were exposed to factors which were expected to result in differences in the speed of eye formation. Cheeses were wrapped in plastic bags with four different levels of gas permeability. Film 1 had the lowest gas permeability and more permeable films had progressively increasing numbers (2, 3 and 4). All cheeses were ripened at three different temperatures in the warm room: T-1 < T0 < T1, with a ΔT of 3°C between each level. Twelve cheeses from each of two cheese vats were sampled and analysed (4 films × 3 temperatures). In each of the two cheese vats, milk from different raw milk tanks (silos) was used.

Image acquisition

X-ray images were recorded with the Safeline AVS X-ray inspection system (Safeline AVS Ltd, Ashwell, Herts, UK), normally used in this dairy as a metal detector. The X-ray detector was an integrated photodiode array and pre-processing unit. Images were stored on the PC in Bitmap format in grey scales.

Cheeses were taken directly from the ripening room still packed in their plastic bags and cardboard containers. The cheeses passed through the instrument on a conveyor belt, lying flat on the belt as illustrated in Figure 1. X-ray beams passed through the cheese while passing on the conveyor. After imaging, the cheeses were returned to the ripening room.

Statistical treatment

Analysis of variance (ANOVA) was carried out using a general linear model procedure of the software package Minitabs 15.1. (MINITAB Inc., State College, PA, USA).

Image processing

The original X-ray images contained 'noise' and other effects which were necessary to remove in

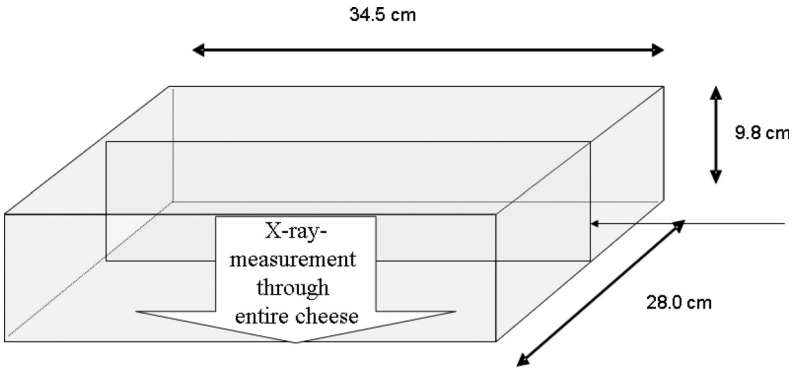

Figure 1 Schematic view of X-ray scanning of a cheese.

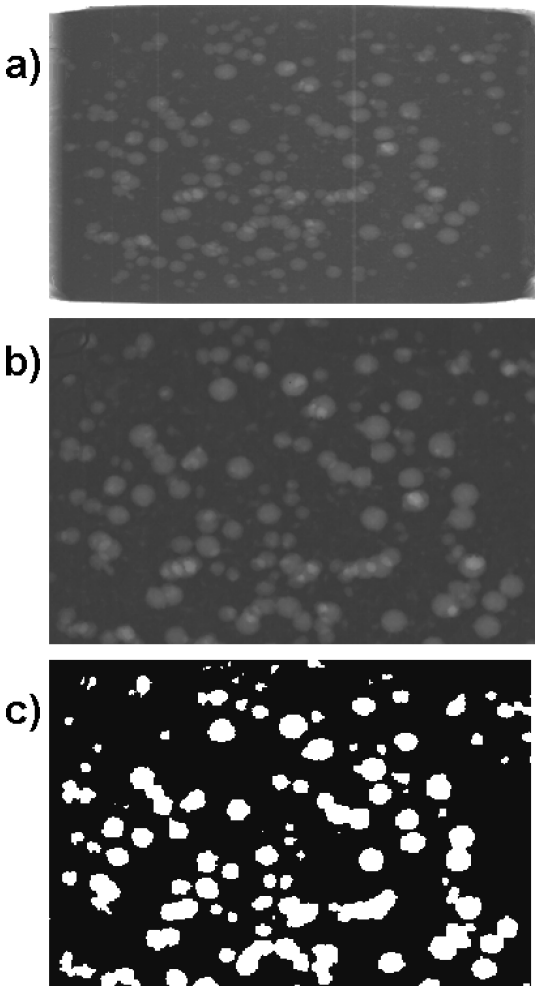

Figure 2 Pre-processing of images: (a) Original image; (b) selected area of the image; (c) binary image.

order to automate the image analysis. Thus, the image processing was divided into two main parts:

1. Pre-processing of images, which results in a binary image of a section of each cheese, as illustrated in Figure 2.

2. A routine for detecting circular eyes in the cheese, even if these eyes were overlapping. Numbers and sizes of eyes were estimated. Based on this, the gas volume of the eyes and a histogram of the size distribution were computed, as illustrated in Figure 3.

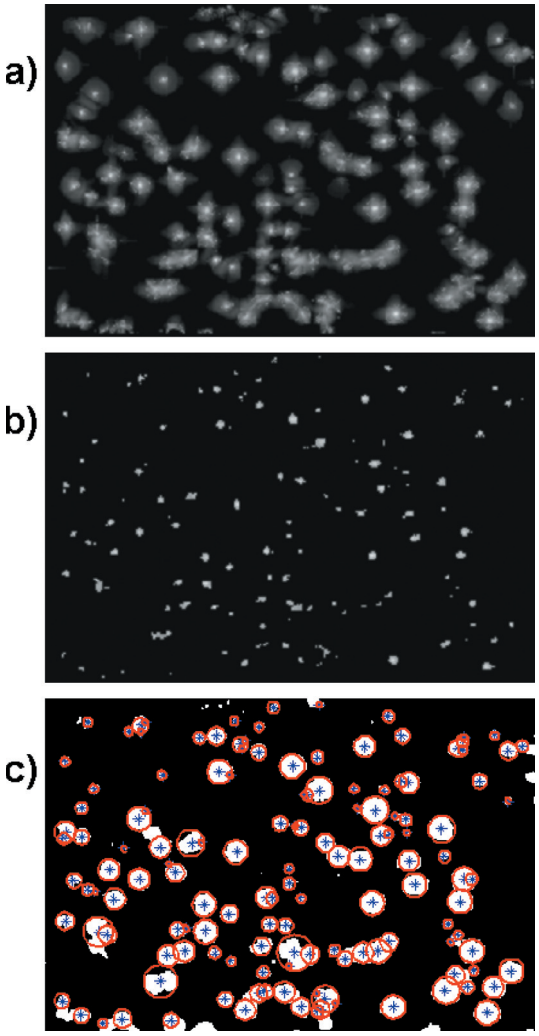

Figure 3 Detection of eyes: (a) Max correlation image; (b) thresholded correlation image; (c) image with calculated centroids and circumferences.

Pre-processing

The raw image is illustrated in Figure 2(a). A vertical artefact of 4 pixels' width was first removed and substituted by pixels more in accordance with the ambient pixels. Then a section of 220 × 400 pixels was extracted (Figure 2b), to remove edges and disturbing parts, such as the green lines and red figures. The next step was to apply a median filter (Tukey 1971) to remove black ('dead') pixels and weak signal pixels.

For effective detection of the eyes, it is necessary to eliminate low-frequency variations in the background, as such variation complicates simple thresholding. This was done by subtracting a fitted 2D polynomial from the image. The remaining variation in the image was then mainly caused by the eyes. As images varied individually, a thresholding criterion for each image was chosen to be the median of all pixel values + 8. After all these steps, binary images are obtained in which the eyes are white and the background is black (Figure 2c). The aim of making binary images was to establish

a good basis for detection and measurement of the eyes.

Eye detection

Most image segmentation methods work best when the objects to be detected are separated from each other. The main challenge in this work was due to the fact that the image combines all the layers through the cheese, which often results in overlapping eyes in the image. The goal was to obtain an automated method that could detect and separate all eyes, including all those which overlapped. A dedicated algorithm was developed which solved this problem to a satisfactory degree. The basis for this algorithm was the pre-assumption that the eyes are circular, and the approach used utilizes so-called matched filtering based on cross-correlation (Pratt 1991). The following steps were involved:

1. Thirteen circular white templates of radius 2–14 pixels were generated.

2. Each template was scanned through the binary image (Figure 2c), pixel by pixel. The cross-correlation value between the circular template and the underlying image pixels was computed for every pixel in the image.

3. For each radius, a cross-correlation image was obtained in which high correlations gave white areas and less-correlated points, darker areas. High correlations were obtained when the underlying image pixels were similar in size and shape to the circular template.

4. From the 13 correlation images, generated for the 13 template radiuses, the highest correlation values were extracted for each pixel and a maximum-correlation image was constructed (Figure 3a).

5. This image was thresholded at a suitable value to give small defined spots (Figure 3b).

6. The centroid of each spot was detected and chosen as the centre of a circle.

7. The radius that resulted in the highest correlation for each circle centre was defined as the radius of the eye.

On this basis, it was possible to find the number of eyes of a given size and position (Figure 3c). It is then simple to compute the volumes based on an assumption that the eyes are globular. The result will not be 100% correct since some eyes will be missed in very overlapped situations and two or more smaller overlapped eyes can be mistaken for one bigger one, but, all in all, the algorithm produces a good estimate. Refining both the image quality and the processing is possible but this relatively simple method was regarded as sufficient for estimation of eye volumes in the cheeses. All image processing routines were developed in Matlab version 7.3 with the Image Toolbox (The Mathworks Inc., Natick, MA, USA).

The volume was computed for each eye with the simple formula for the volume of a sphere:

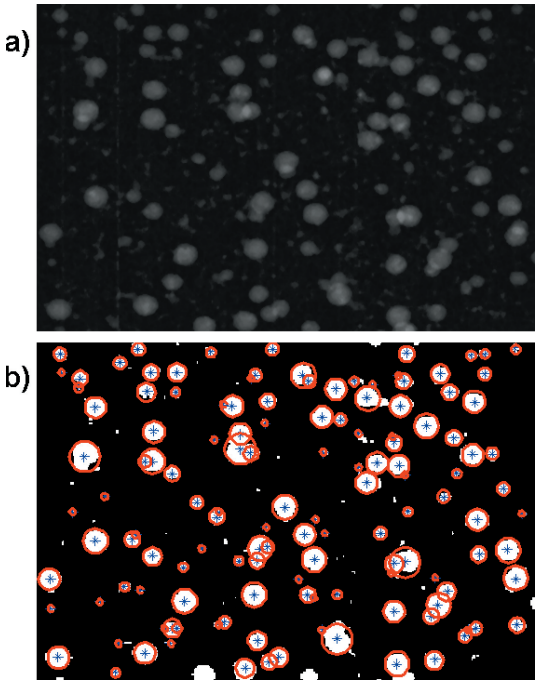

Figure 4 (a) Selected area of the original image; (b) image with calculated centroids and circumferences.

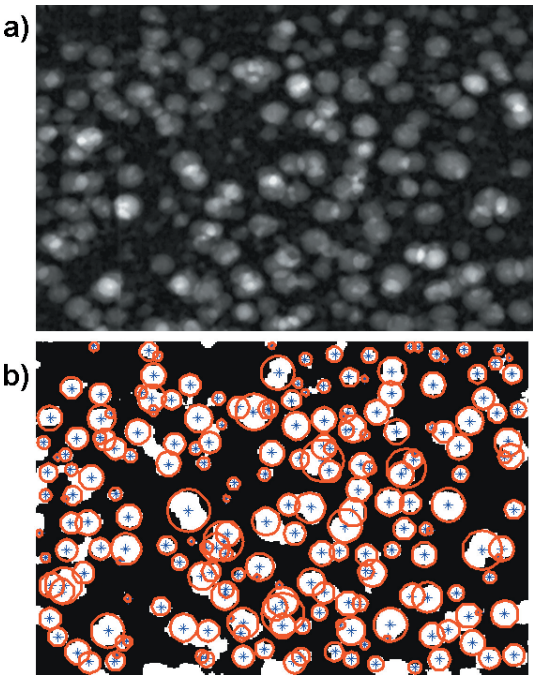

Figure 5 (a) Selected area of the original image; (b) image with calculated centroids and circumferences.

$v = 4/3\pi r^3$. The sum of volumes of all the eyes was then computed by addition.

RESULTS AND DISCUSSION

Figures 4 and 5 show examples of a raw image with the corresponding image with calculated centroids and circumferences. Figure 4 has a low degree of overlapping eyes, and it is therefore easy

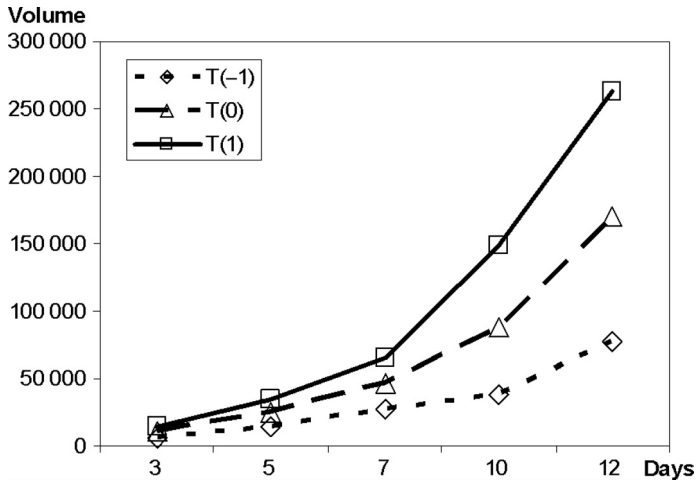

Figure 6 Relative eye volume development during warm room ripening at temperatures T(–1), T(0) and T(1) with a ΔT of 3°C between each level.

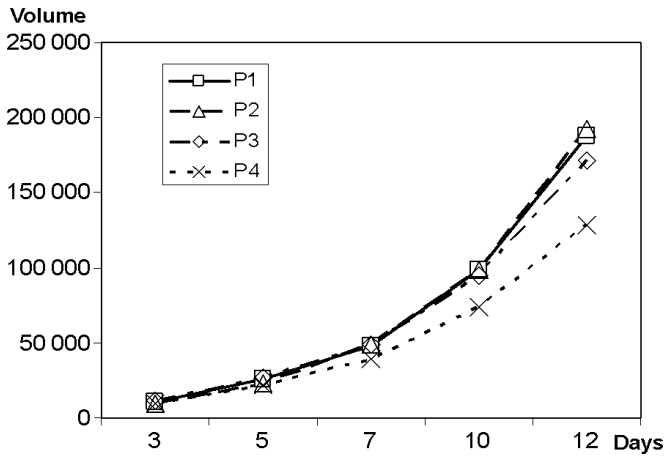

Figure 7 Relative eye volume development during warm room ripening with plastic films P1, P2, P3, P4, film P1 with the lowest gas permeability, and films P2, P3 and P4 increasingly permeable with increasing number.

to calculate. In Figure 5 there are a lot of overlapping eyes but it is clear that the detection algorithm gives a fairly good estimate of the number and size of the eyes, thus producing a sufficient basis for calculation of the volume and size distribution. The segmentation result is not 100% correct since some eyes will be missed if they overlap a great deal, and two or more smaller overlapped eyes can be mistaken for one bigger one; but as already noted, the algorithm produces a good estimate.

However, more accurate segmentation of the eyes can probably be obtained. The output of the routine technique is rather sensitive to the choice of threshold values when generating the first binary images (Figure 2c). Then the thresholding of the correlation image (Figure 3a,b) is critical to find a level that preserves the degree of detail needed to detect small and big eyes at the same time. We used a global threshold value for the correlation images, while a more sophisticated local threshold criterion can probably improve the segmentation, reducing the risk of missing overlapping eyes.

Table 1 Analysis of variance for eye volume at 12 days ripening

Source	*Degrees of freedom*	*P*
Vat	1	0.186
Film	3	0.266
Ripening temperature	2	0.047
Vat*Film	3	0.022
Vat*Ripening temperature	2	0.011
Film*Ripening temperature	6	0.093

Figure 6 shows the calculated total eye volume during ripening in the warm room as an average for the three different temperatures T(–1), T(0) and T(1). The eye volume was highly dependent on the temperature in the ripening room. With ΔT of 3°C between the levels, the eye volume clearly increased with increasing temperature. After 12 days of ripening, the effect of the ripening temperature was significant (Table 1).

Numerous biochemical changes occur during cheese ripening, leading to the development of the characteristic flavour and texture of the cheese varieties. As temperature is increased, enzymatic and chemical reactions occur at faster rates (Fox and McSweeney 2004). Ripening temperature has been shown to have a great influence on the sensory quality of cheese of the type investigated in this experiment (Kraggerud *et al*. 2008). The formation of CO_2 is expected to increase with temperatures up to the optimum growth temperature of propionibacteria at around 30°C (Frölich-Wyder and Bachmann 2004). On the other hand, the resistance of the cheese matrix is lower at higher temperatures, which contributes to keeping the CO_2 pressure relatively low in the cheese (van den Berg *et al*. 2004).

In Figure 7, the calculated average eye volumes in cheese packed with the four different packaging films are shown. In cheese packed in the film with highest permeability to CO_2 (P4) the eye volume development seems to be somewhat slower. However, this effect was not significant. It is possible that the diffusion of CO_2 in this case has led to evacuation of CO_2 from the package. This will reduce the pressure of CO_2 in the cheese and thus decrease eye development.

The number of eyes of different sizes was calculated, from a radius of 2 to 14 pixels. This is presented visually in histograms in Figure 8, with the total number of eyes of each size for each combination of ripening time (days) and temperature. The first row of histograms in Figure 8 shows the results from the lowest temperature, T(–1), with

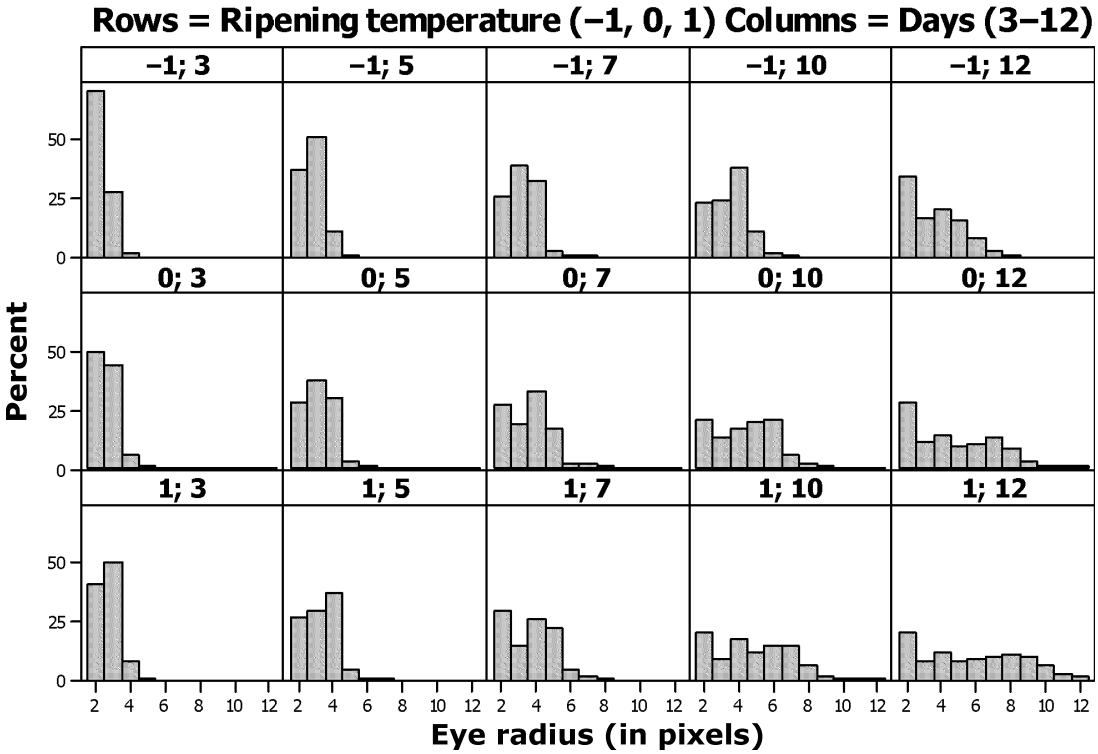

Figure 8 Histograms showing frequency distribution by increasing eye radius (measured in pixels). The rows of histograms represent temperatures T(−1), T(0) and T(1) with a ΔT of 3°C between each level. The columns represent ripening time in days: 3, 5, 7, 10 and 12 days in the warm room.

increasing ripening time from left to right. The following rows show the results from temperatures T(0) and T(1). The figure shows that the number of smaller eyes decreased and the number of larger eyes increased with the ripening time. Cheeses ripened at the higher temperatures achieved higher shares of larger eyes.

The most important quality parameter is likely to be the size of the largest eyes of the cheese. On the basis of the calculations from the X-ray images, it appears that starting to cool the cheese when a given proportion of eyes larger than a set limit is achieved will control the eye size. The limit should be set in accordance with established product specifications of the particular brand of cheese.

CONCLUSIONS

A method for non-destructive monitoring of eye formation in cheese during ripening was developed based on an online X-ray scanner. Automated image processing routines were able to detect the eyes, even when they overlapped, and thereby enabled measurement of the volume and the number of eyes during the ripening of cheese. Systematic differences in the speed of eye formation due to ripening temperature were demonstrated in semihard cheese with propionibacteria. The permeability of the plastic films used for the packaging of the cheese had no significant influence on eye formation in this experiment.

The X-ray image method used could be applicable to non-destructive monitoring of cheese during ripening as a tool to determine the right time to move the cheese from the warm room to the cold room for further maturation. This non-destructive method would also be very useful for quality improvement and research and development purposes, as it makes it possible to follow developments in the same individual cheeses throughout the entire ripening period without disturbing the ripening process.

REFERENCES

Benedito J, Simal S, Clemente G and Mulet A (2006) Manchego cheese texture evaluation by ultrasonics and surface probes. *International Dairy Journal* **16** 431–438.

Brendehaug J and Langsrud T (1985) Amino acid metabolism in propionibacteria: resting cells experiments with four strains. *Journal of Dairy Science* **68** 281–289.

Conde T, Mulet A, Clemente G and Benedito J (2008) Detection of internal cracks in Manchego cheese using the acoustic impulse-response technique and ultrasounds. *Journal of Dairy Science* **91** 918–927.

Du C J and Sun D W (2004) Recent developments in the applications of image processing techniques for food quality evaluation. *Trends in Food Science and Technology* **15** 230–249.

Eskelinen J J, Alavuotunki A P, Haeggstrom E and Alatossava T (2007) Preliminary study of ultrasonic structural quality control of Swiss-type cheese. *Journal of Dairy Science* **90** 4071–4077.

Fox P F and McSweeney P L H (2004) Cheese an overview. In *Cheese: Chemistry Physics and Microbiology*, 3rd edn, Vol. 1, pp 1–18. Fox P F, McSweeney P L H, Cogan T M, Guinee T P, eds. London: Elsevier.

Frölich-Wyder M T and Bachmann H P (2004) Cheeses with propionic acid fermentation. In *Cheese: Chemistry Physics and Microbiology*, 3rd edn, Vol. 2, pp 141–156. Fox P F, McSweeney P L H, Cogan T M, Guinee T P, eds. London: Elsevier.

Haff R P and Toyofuko N (2008) X-ray detection of defects and contaminants in the food industry. *Sensing and Instrumentation for Food Quality and Safety* **2** 262–273.

Gonzales R C and Woods R E (2001) *Digital Image Processing*, 2nd edn. Upper Saddle River, NJ: Prentice Hall Inc.

Kraggerud H, Skeie S, Høy M, Røkke L and Abrahamsen R (2008) Season and ripening temperature influence fatty acid composition and sensory properties of semi-hard cheese during maturation. *International Dairy Journal* **18** 801–810.

Pratt W K (1991) *Digital Image Processing*, 2nd edn, pp 651–668. New York, NY: John Wiley & Sons, Inc.

Rosenberg M, McCarthy M and Kauten R (1992) Evaluation of eye formation and structural quality of Swiss-type cheese by magnetic resonance imaging. *Journal of Dairy Science* **75** 2083–2091.

Strand A H (1985) Scanning of Jarlsberg using data-tomograph. *Meieriteknikk* **39** 2–7.

Tukey J W (1971) *Exploratory Data Analysis*. Redding, MA: Addison-Wesley.

van den Berg G, Meijer W C, Düsterhöft E M and Smit G (2004) Gouda and related cheeses. In *Cheese: Chemistry Physics and Microbiology*, 3rd edn, Vol. 1, pp 103–140. Fox P F, McSweeney P L H, Cogan T M, Guinee T P, eds. London: Elsevier.

von Freudenreich E and Orla-Jensen S (1906) Über die in emmentalerkäse stattfindende propionsäuregärung. *Zentralblatt für Bakteriologie, Parasitenkunde, Infectionskrankheiten und Hygiene* **17** 529–543.

Zheng C, Sun D W and Zheng L (2006) Recent developments and applications of image features for food quality evaluation and inspection—A review. *Trends in Food Science and Technology* **17** 642–655.

Paper III

Quality scoring – a tool for sensory evaluation of cheese ?

Hilde Kraggerud*[1 2], Steffen Solem[3], and Roger K Abrahamsen[1]

[1] Norwegian University of Life Sciences, Department of Chemistry, Biotechnology and Food Science, N-1432 Ås, Norway, [2] TINE R&D Center, Box 8034, N-4068 Stavanger, Norway, [3] Vinmonopolet, Box 6953 St. Olavs plass, N-0130 Oslo , Norway

* Corresponding author: *Tel*: +4795230454, *Fax*: +4722967281.
E-mail address: hilde.kraggerud@tine.no (H. Kraggerud)

Keywords:

Cheese, sensory quality control, quality scoring, sensory descriptive analysis, preference mapping

Abstract

The objective of this paper was to evaluate the relevance of data from quality scoring methodology of ISO/IDF performed by expert assessors for the sensory quality control of cheese. The approach to this evaluation was comparison of quality scoring with sensory quantitative descriptive data from a trained panel and consumer preference data. Significant regression correlations were found between quality scoring and descriptive data in a data set obtained from evaluation of 8, 24 and 40 week old Norwegian semi hard Gouda type cheese (n=459). However, the level of explained variance was low.

In a smaller set of data aimed at preference mapping, higher correlations were found between quality scoring and descriptive data. Preference mapping showed that the average consumer and the quality scoring expert assessors disagreed in particular on consistency properties. External preference mapping after segmentation of consumers by hierarchical clustering was found useful. Consumers could be divided into 3 main clusters. One of these clusters mainly agreed with the expert assessors; the cheese preferences of the two other clusters were in disagreement with expert assessor approval. Thus it would be possible to suggest new product specifications, highly approved by consumers, for the variety of cheese investigated. Adjustments in product specifications acoording to consumers preferences would enhance relevance of the quality scoring. A high level of explained variance was found between consumers' overall preference scores and overall quality scores, which could indicate that quality scoring is a relevant, holistic sensory quality measure.

1. Introduction

An effective approach to ensure consistent delivery of products of defined quality is very important for a food producer. According to ISO 9001, (International Standardisation Organisation, 2008), one of the most frequently used quality management systems, “top management shall ensure that customer requirements are determined and are met with the aim of enhancing customer satisfaction”. Also according to ISO 9001 “the organization shall plan and implement the monitoring, measurement, analysis and improvement processes needed to demonstrate conformity to product requirements”.

Quality control (QC) throughout the production chain from raw materials to final cheese product is a challenge. In addition to the task of market-sorting of products, correct identification of sensory quality characteristics is important for continuous improvement of the cheese making process. The sensory quality is what consumers perceive directly, which also is in accordance with the claim: “The ultimate measure of product quality and success is sensory quality” (Drake, 2007). Relevant measures of sensory quality are a prerequisite for diagnosis of the causes of deviations from specifications and for application of appropriate corrective actions and improvements in the cheese making process and the ripening and storage conditions.

Three basic categories of sensory tests can be applied to dairy products: 1) traditional judging/grading, 2) affective (consumer) tests, and 3) analytical sensory tests (Bodyfelt, Drake, & Rankin, 2008). For methods in categories 1) and 3), selection and training of panel members is very important (Delahunty & Drake, 2004).

According to (Bodyfelt, 1981) sensory intensity is generally not difficult to assess. Quality is more elusive and therefore definition, measurement and interpretation involve considerable difficulty. Since late in the 19th century scoring methods have been used by the dairy industry for sensory evaluation of products. Absence of defects is often used rather than a descriptive definition of quality (Amerine, Pangborn, & Roessler, 1965). Traditional quality evaluation methods for dairy products are defect-oriented and based on the use of expert assessors (Bodyfelt et al., 2008). Daily grading at the manufacturing location based on deviation from a reference scale, very similar to the quality scoring method used in this paper, has been recommended (Weller & Stanton, 2002a). For classification purposes, there is need for a determinative term to make the sorting task easy to practice. This term could either be calculated from a number of separate parameters, as in the Quality Index Methodology (QIM) (Martinsdóttir, Sveinsdóttir, Luten, Schelvis-Smit, & Hyldig, 2001) or it could be executed directly by the assessors using overall quality terms (Elortondo, Ojeda, Albisu, Salmeron, Etayo, & Molina, 2007; Etaio, Albisu, Ojeda, Gil, Salmerón, & Elortondo, 2010; International Standardisation Organisation, 2009; King, Gillette, Titman, Adams, & Ridgely, 2002; Pecore & Kellen, 2002).

Time and cost issues are important for food producers in the choice of method for quality control. Sensory evaluation methods are very different in time consumption. Descriptive analysis is much more time-consuming than the quality scoring method. A proportion of at least 10:1 is realistic in our experience. This makes conventional descriptive methods less relevant to implement as a quality classification method for regular use in the industry.

It is important to include consumers' input in establishing and evaluation of product specifications in order to ensure that consumers' expectations are met. Several methods are suggested in the literature. Consumer acceptance limited to evaluation of defects can be determined by so-called survival analysis (Hough, Sanchez, Garbarini de Pablo, Sanchez, Calderon Villaplana, Gimenez, & Gambarot, 2002). A method has been described for development of a consumer-preference-based scoring guide in a total quality scoring system (Ismail, Haffar, Baalbaki, & Henry, 2001). The methodology for using consumer responses to determine target ranges of intensity and limits for each attribute, throughout product development and the QC programme has been described (Pecore & Kellen, 2002; Weller & Stanton, 2002b; Weller & Stanton, 2002a). The use of preference mapping techniques is widespread and is applied in this paper for specification of target quality.

Although quality scoring (QS) according to standardized methodology is and has been widely in use in industry, this methodology is rarely discussed in international research publications. The objective of this paper was to evaluate the relevance of data from the quality scoring methodology of ISO/IDF (International Dairy Federation, 1997; International Standardisation Organisation, 2009) as a measure for sensory quality of cheese. The approach to this evaluation was in the first place comparison of quality scoring data with sensory quantitative descriptive data. And then comparison of quality scoring and sensory descriptive data with consumer liking.

2. Materials and methods

2.1. Sample sets

All cheese evaluated was rindless, 27% fat, semi-hard Gouda type cheese – the most common cheese variety in the Norwegian market. Two data sets were collected.

Data set 1: 153 cheese samples were evaluated at three ages of ripening: 8, 24, and 40 weeks, making 459 samples altogether. The cheeses evaluated had a quite large variation in sensory quality. They were part of different parallel experiments, with sources of variation like raw milk treatment, different ripening temperatures and production at different dairy plants. To make a robust set of data, intended also to cover seasonal variation, 12 successive samples were collected at intervals of 4 weeks in the course of almost one year . Sensory evaluation sessions were held every 4 weeks throughout a period of 1½ years, in total 20 assessments. At each session, cheese of either; 8, 8 + 24, 8 + 24 + 40, 24 + 40 or 40 weeks of age were analysed, as illustrated in table 1.

Data set 2: All cheese samples were commercial production samples collected from wholesalers or retailers, a total of 17 samples from 5 different dairy plants – with different ages of ripening. From the 17 samples, 7 were selected for preference mapping. All sensory evaluations were conducted within a period of 1 month.

2.2. Quality scoring

2.2.1. Method

Internal TINE-method based on the reference method IDF 99C:1997, Sensory evaluation of dairy products by scoring, Part I and Part IV, was used (International Dairy Federation, 1997). Further specifications are given in the text below.

2.2.2. Sensory quality definition of the product

The product specifications include a verbal definition of the sensory quality. Expertise in the sensory specification of the product is necessary for the assessors, in addition to proper training and regular participation in sensory assessments over time. The sensory product specifications for these cheese varieties do not necessarily reflect consumers' views because they have been established in the framework of experts' definitions of cheese quality. For the actual cheese variety the verbal sensory product specifications are as follows (based on an age of the cheese of 7-10 weeks):

Outer appearance: Even and straight edges, with closed, dry surface.

Inner appearance: Evenly distributed round eyes (5-20 in a cross-section) with diameter 3-8 mm.

Consistency: Semi-firm, flexible and easily sliceable with cheese slicer.

Flavour: Pure, mild, aromatic with weak acidity.

2.2.3. Selection and training of assessors

Authorized expert assessors were used. To be authorized as an assessor in TINE, a candidate has to participate in at least 20-30 assessment sessions in the course of about one year, before participation in an Authorization Test, in which agreement with reference assessors and the candidate's repeatability are the most important criteria used for evaluation of the suitability of the candidate as an assessor. Assessors used in the experiments have been evaluating cheese for more than 20 years, although one had been active for only 2-3 years.

Data set 1: Three authorized expert assessors participated in each quality scoring session, picked from a team of seven assessors. Each assessor evaluated from 25 to 280 samples in the total experimental period.

Data set 2: Five authorized expert assessors participated in this evaluation. The evaluation took place in the course of one day.

2.2.4. Scales of evaluation

Appearance, consistency and odour/flavour as well as overall quality score were assessed, using a numerical interval scale from 1 to 5, with 0.5 point intervals. The scale is defined according to IDF 99C, Part 1 clause 9.2 (International Dairy Federation, 1997) as follows: 5 = conformity with the pre-established sensory specification (PS), 4 = minimal deviation from PS, 3 = noticeable deviation from PS, 2 = considerable deviation from PS, 1 = very considerable deviation from PS. The score 0 – which was defined in IDF99C (unfit for human consumption) was not used in our method. Score 0 has been removed from the scale in the newer version ISO 22935-3 / IDF 99-3 (International Standardisation Organisation, 2009). When an assessor scored <4 points for a sample, a description of the deviation was given, using predefined nomenclature of defect terms.

The company quality policy is that a product is classified for sale on the ordinary market if the average score (over assessors) is >2.7 (class 1) Cheese with lower scores (class 2) is utilized for other purposes, e.g. grated or processed cheese, depending on the type of deviation.

2.2.5. Procedure

Cheese was presented to the assessors in random order and placed in a row on a long table. Before sensory assessment, each cheese block (5 kg) was cut in two. Sensory assessment was carried out using a standard cheese slicer for cutting. Each session started with calibration of the assessors with two different cheeses as reference samples. The assessors got no information about the samples, and were not allowed to communicate during the following assessment. In Data set 1 there was one replicate of each sample, due to the high number of samples. In data set 2 there were two replicates per assessor. Assessments were made in two successive sessions with a pause in between.

2.3. Descriptive sensory analysis

2.3.1. Vocabulary

The vocabulary of sensory attributes used in this experiment was developed to obtain the most complete description of the semi-hard cheeses with the shortest possible list of attributes. The attributes have been in use by this panel for many years. The vocabulary used is in accordance with ISO 5492 (International Standardisation Organisation, 1992). The terms to be included or excluded were evaluated for each data set, depending on the needs of the specific experiments. These attributes have earlier been described in detail (Kraggerud, Skeie, Høy, Røkke, & Abrahamsen, 2008). Each sensory attribute was evaluated using an interval scale from 1 to 9 points, as defined in ISO 4121 (International Standardisation Organisation, 2003). In data set 1, a discrete scale was used in a manual paper form. Data set 2 was

collected with Eye Question (Logic8, Wageningen, Netherlands) using a continuous 1-9 point scale.

2.3.2. Assessors

For descriptive sensory analysis, an internal laboratory panel was used comprising six assessors, well trained on cheese for several years. With very few exceptions they all attended all sessions.

2.3.3. Procedure

Before each session, a calibration session was performed in order to obtain an agreement on the use of attributes and scales, using two different cheeses as test samples. All assessments were conducted in individual booths at the sensory laboratory, designed in compliance with international standards for test rooms, ISO 8589 (International Standardisation Organisation, 1988). Cheese samples were tempered to 14 ºC ± 2ºC prior to assessment. In each session the order of assessment of the samples was randomized for each assessor. In Data set 1 there was one replicate of each sample, due to a very high number of samples, while there were two replicates in experiment 2. The scores of each sample were averaged over assessors.

2.4. Consumer testing

Seven samples were tested, selected from Data set 2. Samples were tested in two different laboratories in different parts of the country. Consumer selection criteria were over 15 years of age, and frequent user of cheese (>once a week). The

questions asked and procedure of testing were identical for the two laboratories. In both cases a warm-up sample was served before evaluation of the series of 7 samples, served in randomized, monadic sequence. The temperature of the samples was +4°C when served. Consumers were presented with an ordinary 1-kg piece of the cheese and they used the cheese-slicer to serve themselves with an adequate amount of cheese for testing. This procedure was chosen to imitate the normal cheese eating situation for a consumer in Norway, using cheese directly from refrigerator, with cheese slicer as equipment. 180 and 170 respondents participated in the two laboratories, 350 consumers altogether. Data from the two laboratories were handled as one data set in the data analysis. Incomplete data sets were removed from the data analysis, resulting in 342 respondents.

Among questions asked were overall liking on a 9 point discrete interval scale with defined end points: 1= Dislike extremely, 9 = like extremely. In this paper only this overall liking parameter is analysed and presented.

2.5. Analysis of data

2.5.1. Multivariate data analysis

For multivariate data analysis the Unscrambler version 9.8/10.1. (CAMO Software AS, Oslo, Norway) and XLSTAT ver 2008.5.02 (Addinsoft SARL, Paris, France) were used. Principal component analysis (PCA), principal component regression (PCR) and partial least squares regression (PLS) are methods suitable for studying large quantities of data. Linear combinations of the original variables in terms of underlying latent variables (principal components, PC – numbered PC1, PC2… with PC1

explaining the most) are computed and often plotted to study effects (Martens & Næs, 1989). Scores describe sample patterns and show differences or similarities between samples. Loadings describe variable correlations and patterns. Limits of 50% and 100% explained variance are marked with circles in correlation loadings plots (Martens & Martens, 2001). For regression modelling of response variables from a set of predictor variables, PLS/PCR was used.

2.5.2. Preference mapping techniques

Preference mapping are techniques for illustrating the relationships between sensory and consumer data. In external preference mapping data from consumers are projected on to a sensory descriptive map, in our study PCA (McEwan.J.A., 1996). External preference mapping (PREFMAP) technique was performed using XLSTAT ver 2008.5.02 (Addinsoft SARL, Paris, France), and also Unscrambler . A polynomial regression is calculated, adding linear and quadratic terms and the best fit models chosen, vector (linear) or circular (quadratic), corresponding to PCR. Contour plots show the probability of having a score greater than medium, in our case score 5, in a given area of the perceptual map. Extended internal preference mapping technique (MDPREF) (Næs, Brockhoff, & Tomic, 2010) was also performed, with quality scoring and sensory descriptive data as dependent variables and principal components from consumer liking data as predictors.

2.5.3. Analysis of variance

Analysis of variance (ANOVA) was carried out using software package Minitab® 15.1. (MINITAB Inc., State College, PA, USA) and XLSTAT ver 2008.5.02 (Addinsoft SARL, Paris, France).

2.5.4. Agglomerative Hierarchical Clustering (AHC)

Agglomerative Hierarchical Clustering (AHC) is an iterative classification method, in our case performed in XLSTAT ver 2008.5.02 (Addinsoft SARL, Paris, France). Successive clustering operations produce a binary clustering tree (dendrogram). The proximity between two objects is determined by measuring at what point they are similar (similarity clustering), using Pearson's correlation coefficients. The agglomeration method used was unweighted pair-group average single linkage.

3. Results and discussion

3.1. Data set 1

3.1.1. Sensory quality scoring

In Table 2 average scores for the quality scoring attributes of cheese of different ages are presented. Also the number of samples in each of the two quality classes is given, one below 2.75 points overall quality score (class 1) and one above 2.75 points overall quality score (class 2). In sample class 1, which covered cheese from 8 to 40 weeks of age, a general tendency for quality scores to decrease during storage of the cheese was observed (Table 2). Scores for 40 week old cheeses were significantly lower than for cheeses after 8 and 24 weeks of ripening for all quality scoring properties assessed. In class 2 the number of the scores were almost

doubled from 8 (N=58) and 24 (N=61) week old cheese to 40 weeks of ripening (N=111).

Sensory product specifications are among the most important elements in quality control programmes. One way of implementing specifications in sensory evaluation, is to establish product standards for training (King et al., 2002). But this is quite difficult with a biological material like cheese, with large variations in sensory characteristics originating from raw material and process variables as well as the effects of ripening during storage (Kraggerud et al., 2008). The changes in sensory properties during cheese ripening described in 3.1.2. are in accordance with earlier observations on various cheese types (Hickey, Kilcawley, Beresford, Sheehan, & Wilkinson, 2006; Kraggerud et al., 2008; Muir, Hunter, Banks, & Horne, 1995).
The sensory specifications used for all cheeses in these experiments are based on cheese of normal age at the time of packaging for this variety of cheese - 7-20 weeks.

For extra mature cheese in the market, however, the specifications are different. It is very difficult to maintain a sharp limit when quality deviation occurs in relation to degree of ripening in cheese. The longer the cheese was stored, the more the cheese deviates from the sensory specifications used in this study. The tendency towards lower quality found after 40 weeks of ripening has to be considered as an innate quality deviation from product specification.

3.1.2. Sensory descriptive analysis versus quality scoring

The sensory variations of data set 1 are illustrated in a PCA/PCR score plot (Figure 1). Of the variation 61% is explained in PC1, 10% in PC2. From this score plot it is obvious that sensory properties change with age, as PC1 is dominated by age differences. The correlation loadings plot is given in Figure 2. This figure shows that PC1 was dominated by consistency attributes. Chewing firmness, elasticity, cohesion, grainy and dryness were reduced, while attributes like soluble and floury increased during ripening (to the right in the plot). Taste/flavour attributes like bitter, pungent, sour and flavour intensity were increased in intensity during storage. The PCR regression models for quality scoring (QS) variables vs sensory descriptive variables were significant, with an optimal number of 5 factors. Explained variation in Y (quality variables) was however low, with 15% explained after 2 factors. Almost all variables had significant regression coefficients, the only exceptions were shear firmness, aromatic, sweetness and malty. Low QS scores were correlated with consistency variables such as pressure firmness, grainy, elastic and pasty, and to flavour/taste characteristics such as odour and flavour intensity, acidic, pungent and bitter. Quality scoring attributes were positioned in the opposite direction of ripening time in the correlation loadings plot (Figure 2), which illustrates the tendency of quality to be reduced during extended ripening, as described earlier in section 3.1.1.

The low explained variance in the relationship between sensory panel and quality scoring might originate from low signal to noise ratio in both sensory and quality scores, as the total variation was relatively low in this data set. Total standard deviation of the four QS variables were between 0.37 and 0.45, while average standard error of the mean over assessors (SE Mean) for the 458 samples was

between 0.15 and 0.18. SE Mean was for most cheese consistency variables of the profile around 0.2, while flavour variables were a little higher, near 0.3 on average.

The non-linear nature of quality scoring is also a plausible explanation. ‘Just about right’(JAR) would give a high quality score, whereas ‘too much’ or ‘too little’ of many “typical” sensory properties would both result in lower scores. This is for example the case for the texture properties, where JAR for attributes like firmness would give the highest quality score, while firmer and less firm cheeses would both get lower scores. When it comes to defect properties, the nature of the quality scale is more like the sensory attributes, with a lower quality score because defects are higher in intensity. This is in contrast to sensory descriptive linear scales, where all attributes are measured on linear intensity scales.

3.2. Data set 2

3.2.1. Sensory descriptive analysis versus quality scoring

Seven samples were selected for preference mapping out of a total of 17 samples analysed with sensory descriptive analysis. A qualitative approach to 3 PCA dimensions from the sensory attributes was used to find samples covering most of the sensory space. In addition one sample (B) with deviation in appearance, which was not covered in the sensory description, was chosen. The number of samples was chosen to be able to analyse a manageable number of samples in the consumer test, which was eight, owing to the planned use of one warm-up sample.

Principal components regression was performed, with sensory descriptive data as X and quality scoring data as Y, on the data set from 7 cheeses. The PC modelling of sensory descriptive analysis explained 89% of the variance in PC1+2 (Figure 3). PCR correlation loadings showed that QS variables were quite well explained in the calibration set (around 60% after 3 PC). At an optimum of 3 factors, validation root mean square error (RMSE) of overall quality score was 0.4. Overall quality scores (presented in Table 3) had significant positive correlations with the consistency variables shear firmness and solubility, and were negatively correlated with grainy. The attribute aromatic was the only flavour/taste attribute with a significant regression coefficient. The attribute was positively correlated with overall quality score.

3.2.2. Sensory descriptive analysis vs consumer testing

Average liking scores for overall liking of the 342 consumers participating in the test are presented in Table 3. According to Tukey's test Sample C was significantly less liked than the rest of the samples, while sample G was significantly less liked than samples B and F.

External preference mapping (PREFMAP) was performed, regressing overall liking responses of consumers into the PCA sensory map. PREFMAP based on PCR, was chosen as method because it can be directly related to the sensory descriptive analysis: the sensory PCA act as a stable reference throughout the analysis. This is particularly useful when there are several groups of consumers (the different clusters found here, for example). Then they can all be related to the same PCA results. This is also the reason for choosing PCR instead of PLS regression for the external

preference mapping. Figure 4 shows the positioning of the 342 consumers and the 7 samples in the sensory perceptual map.

In the external preference map most of the consumers' scores are positive in PC1, with 151 consumers in the upper right quadrant, and 81 in the lower right of Figure 4. Of the variation in the consumers' overall liking, 41% was explained by the sensory descriptive characteristics of the samples in PC1+PC2 and another 17% of the variation was explained in PC3. In dimension 3 (PC3) sample G was differentiated from the other samples, in particular with respect to the taste attributes like salty, acidity and bitter, but other attributes also characterized this sample as different from the others.

3.2.3. Clustering of consumers

Consumers were evenly distributed in age (around 50% over 40 years and 50% under) and sex, (around 50% men and 50% women). They were also evenly distributed between two laboratories, Oslo and Stavanger. Almost all were heavy users (> once a week) of this type of cheese. A clustering procedure was performed to examine if there were typical clusters of consumers with similar preference patterns. Hierarchical clustering was performed on the dataset of 342 consumers. The dendrogram in Figure 5 shows six clusters of consumers. Three main clusters were found, with 96, 76 and 89 consumers respectively. Attention will be paid to these three main clusters. The three remaining clusters had only 18, 22 and 42 respondents respectively. Clusters with less than 30 consumers are too uncertain to pay attention to. Cluster number 6 proved hard to interpret, as consumers' preferences of that cluster were diverging. This phenomenon of the last clusters

having very divergent consumers' preferences is in accordance with our own experience from other datasets.

External preference maps were then generated with only consumers from each of the three clusters. These maps were based on the same sensory perceptual space of 7 samples as illustrated in Figures 3 and 4. In the contour plots in Figure 6, optimal areas for each of the clusters correspond to the highest values of probability (red colours), while blue colours illustrate the least liked areas of the map. The upper row shows PC1 vs PC2, while the lower row illustrates PC1 vs PC3. Cluster 1 was characterized by high values in PC2, corresponding to flavour characteristics like ripened, aromatic and sweet. This cluster apparently was not very much affected by PC1 – dominated by consistency attributes. Clusters 2 and 3 overlap to a high degree in PC1/PC2 – they are both characterized by high values in PC1, which means that cheese consistency was of major interest. To separate Cluster 2 from Cluster 3, contour plots for PC1 vs PC3 were made (Figure 6). Cluster 2 was positive in the direction of Sample G. Typical attributes in this direction were salty, bitter, flavour intensity and acidity. Cluster 3 preferred the opposite direction in PC3: typical attributes were pasty and soluble, and low in pressure firmness, elastic, ripened and aromatic.

We also compared patterns of overall liking scores between the two geographical regions (different laboratories) in which we tested, and found essentially no difference in the two preference maps. They are therefore not presented here.

3.2.4. Extended internal preference mapping

Extended internal preference mapping (MDPREF) was also performed on data set 2. Using PCR, all consumers‘ overall liking scores versus all sensory descriptive attributes and quality scores were modelled. The results are shown in Figure 7. The explained variance of sensory variables was 64% after 2 factors. Most of the consumers were to the left in PC1. Many sensory variables were explained to a very high degree. It can be seen that there of course was good agreement with external preference mapping: favoured attributes were pasty, soluble, malty and floury, and low in firmness, grainy, sulphurous and cohesive.

All the quality scoring variables were explained to a high degree in PC2, while hardly at all in factor 1. Overall quality score had, as an example, R^2=0.998, RMSEC 0.02 (calibration set) with 5 factors. After 2 factors, R^2 was 0.78, RMSEC 0.22. With as few as 7 samples, there was the maximum divergence in sample space. Calibration is the relevant model to interpret here.

3.2.5. Quality scoring vs consumer testing

When looking at the average consumer, there was a disagreement between quality scoring results and consumer preferences. By clustering of consumers, a relatively large cluster (Cluster 1) was observed more or less agreeing with the QS assessors. However the two other large clusters were in disagreement with the QS assessors. This may be interpreted in terms of the way that relatively large groups of consumers may want cheese with other characteristics than the present specifications of the cheese variety being tested. From this segmentation of the consumers some

important observations could be made regarding the sensory perception of this type of cheese by consumers:

- Consistency specifications should generally be adjusted in the direction of less firmness / more solubility.
- There is room for at least 3 different product directions within this variety of cheese.
    1. A ripened, aromatic product (represented by sample D)
    2. A softer and yet aromatic variety (not represented by any particular samples in our experiment)
    3. A variety in the softer direction, in which attributes like bitterness, salty and doughy are tolerated by the consumers (represented by sample E and F).

The preference mapping method is widely used for determining product specifications. The results in this paper indicate that preference mapping in combination with clustering of consumers has a good potential for use in the industry in their product development and in their adjustment or differentiation of the characteristics of products available today.

In this work a fairly high degree of explained variance between consumer liking and quality scores was observed in section 3.2.4. Overall scores of quality and liking are both holistic and optimum-orientated. Consumers' main overall liking criteria were apparently somewhat different from QS overall quality criteria: almost nothing was explained in PC1. In the 2nd PC as much as 78% was explained, and after 3 factors >99% in total. This indicates that there were probably common criteria for the parameters of overall quality and consumers' preferences, even though in section 3.2

the findings indicated that their main directions were opposite. This indicates that the suitability of methodology for sensory quality control is highly dependent on the interpretation and use made of the results. Quality scoring methods could be appropriate for measuring sensory quality of cheese relevant for the consumers' point of view. This phenomenon would be very interesting to study in more detail. More experimental work is therefore needed, with a higher number of samples evaluated by consumers and by quality assessors, as well as sensory descriptive panel assessment.

4. Quality scoring / grading of cheese – a continuous discussion

In the course of the years there has been much discussion around cheese assessment / scoring / grading. In 1979 an article from Australia (McBride & Hall, 1979) was published with the title: "Cheese grading versus consumer acceptability: An inevitable discrepancy". In a study 12 cheeses were graded by official cheese graders. Consumers also tested the same cheese, one pair of cheeses per consumer. 17 different pairs were constructed over the 12 cheeses (which seems to be far too few combinations for a statistically valid test). They found no or little correlation between consumer and grader assessment. The implication drawn from this experiment was that grading was of dubious value. Instead the authors suggested a system of descriptive sensory analysis using a small number of flavour attributes and judging the degree of each attribute on a category scale. They also suggested that this would enable cheese packagers to classify and pack cheese according to their flavour characteristics.

In a critique, (O'Mahony, 1979) discussed, among other methods, the ASDA score card methodology, especially with respect to the use of scales. The ASDA differs quite a lot from the IDF scoring methodology (International Dairy Federation, 1997; International Organization for Standardization & International Dairy Federation, 2009a). But still there is much in common. O'Mahoney suggested that the conversion from a degree of one or more off-flavours to a score on one quality scale is less informative than flavour profiling. He also suggested that such a score cannot be handled statistically as it is derived from different ordinal scales. In our view an overall quality score is very much related to an overall liking score, in terms of being a holistic quality score, either objective or subjective. The linearity of the scales can, of course, be discussed.

"Use and misuse of sensory evaluation in quality control" (Sidel J.L., Stone, & Bloomquist, 1981), discussed different aspects of sensory quality control. The authors underlined the difference between a technical expert making a quality judgment and a sensory panel. One important misuse identified was more than one test objective in a single test. Quality tests should not combine affective judgments and excellence with identification or descriptive defects. They also suggested that the expert should be trained as a vital resource for determining product proximity to a consumer-defined standard. The selection and training of assessors was also emphasized, and the requirement for an objective statistical approach to the evaluation of the readiness of a judge. The use of scales and score sheets and other suggestions for improvements of quality scoring was also discussed.

It was proposed that a Committee on Sensory Data should develop a statement of policy about sensory data, primarily data about taste, to guide authors and reviewers for the Journal of Dairy Science. Among their statements were that quality scores are not appropriate or adequate for research. Traditional quality scales may obscure other important sensory differences as they emphasize the presence or absence of given defects. Each flavour attribute of a product should be separately scaled and analysed statistically (Hammond, Dunkley, Bodyfelt, Larmond, & Lindsay, 1986).

An interesting essay about the complexity of sensory quality has been published (Lawless, 1995). He claims that consumers' opinions must be used as benchmarks – as they perceive products in an integrated fashion, product quality being multidimensional, and overall quality can only be measured with great difficulty. An example given is that of the defect term oxidized used by judges, which will change from sensations like cardboard to tallowy and painty, and then to fishy in a descriptive analysis. It has been shown that such complex defect-oriented terms are less correlated with consumer opinion and not as discriminating as simpler associative terms (Claassen & Lawless, 1992; Lawless & Claassen, 1993).

It has also previously been shown that consumers' preferences were not always in accordance with the expert assessors' scores for cheeses similar to the cheese variety assessed in this work (Hersleth, Ilseng, Martens, & Næs, 2005). They found that quality scoring assessors disagreed with consumers in their judgment of cheese with similar quality defects to those in our experiment.

The IDF (and later ISO/IDF) method has been under continuous development through the years. In all versions, the first came in 1987 (International Dairy Federation, 1987), the 5 point scale for each attribute (appearance, consistency, flavour) has been defined with reference to pre-established sensory standard, 5 points corresponding to no deviation from specifications. Further grading of the products are then based on the points achieved in the evaluation. A standard nomenclature of defects has been suggested for each product group. The newest version of standards has been substantially improved, and is in accordance with Good Laboratory Practice. It now consist of three parts, one with general guidance for recruitment, selection, training and monitoring of assessors (International Organization for Standardization & International Dairy Federation, 2009c). The second contain recommended methods for sensory evaluation, with general recommendations about samples, equipment etc (International Organization for Standardization & International Dairy Federation, 2009b). The third part is guidance on quality scoring – a method for evaluation of compliance with product specifications (International Organization for Standardization & International Dairy Federation, 2009a). It seems to us these methods are unknown to most of the sensory community, and are very seldom referenced in publications. These methods and principles could also well be used for other food industries, as they are general in their nature.

As the principle of quality scoring methodology is compliance with product specifications, many of the subjects discussed around grading of dairy products in former published articles are not directly applicable for these methods. When some of us started working in the dairy industry more than 30 years ago, the situation was

quite different, as there were a set of “official judges” who worked across 180 dairy companies in Norway. Nowadays this is not the case, as all assessors belong to the few (2 cheese producing) companies left. We think this also applies to most countries. It is, of course in the company’s own interest that the product specifications are in accordance with consumer preferences, which is also suggested in the ISO/IDF method. For this reason, regular consumer studies, often using preference mapping techniques, are conducted in order to adjust product specifications. So the inevitable discrepancy between cheese grading and consumer acceptability (McBride & Hall, 1979) should generally not be the case using the QS methodology.

Nevertheless, it is an important task to train and monitor assessors, this is not easy, but we use instance interlaboratory studies comparing panels, and performance measurements for each assessor. In our company the assessor monitoring will now be drastically changed, as electronic registration of each score and defect given will be stored in a database. Development of monitoring performance measures for running analysis will be developed.

4. Conclusions

Significant regression correlations were found between QD and QS data in a data set obtained from Norwegian semi-hard cheese (n=459), with 8, 24 and 40 week old cheese. However explained variance in QS data was low – only 15% in the first two principal components (PC1&PC2). The focus in quality scoring is traditionally on defects and on deviations from product specification. Variations in age or in degree of maturation of the cheese made interpretation of scores difficult. The same product

specification was used for all samples and the differences in age of the cheese were unknown to the assessors. There was a general tendency that scores were lower for more mature cheeses, the product specifications used being those for less mature cheeses.

In a smaller set of data with 7 samples used for preference mapping fairly high explained variance between QD and QS data was found (61% in PC1+2). Preference mapping showed that the average consumer and the quality scoring expert assessors disagreed in particular on consistency properties. Preference mapping after segmentation of consumers by hierarchical clustering was found useful. Three main clusters were found among the consumers. One of these clusters mainly agreed with the expert assessors, and preferred aromatic flavours. The two other clusters preferred cheese with a consistency which deviated from the consistency highly approved by the expert assessors. From this, it is possible to derive three different desirable product specifications

Defining, measuring and interpreting quality involve considerable difficulty and time and cost issues are important for food manufacturers. Descriptive analytical methods are considerably more time consuming than quality scoring and are thus less appropriate for regular use in business practice. Overall quality score was, to a high degree, explained from consumer liking scores, which can be interpreted in such a way that there might be innate common aspects for the two methods. Quality scoring is thus shown to be a relevant method for sensory quality measurement. In theory there should be full agreement between quality assessors and consumers, as long as

product specifications are based on consumer views. In addition to unambiguous product specifications, training and monitoring of assesors is important tasks.

The demonstration of clusters of consumers can suggests that relatively large groups of consumers may want cheese with other specifications than the present characteristics of the cheese variety being tested. In particular consistency specifications should for this cheese variety be adjusted in the direction of less firmness / more solubility.

Reference List

Amerine, M. A., Pangborn, R. M., & Roessler, E. B. (1965). *Principles of sensory evaluation of food*. New York: Academic Press, Inc.

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

Bodyfelt, F. W. (1981). Dairy Product Score Cards: Are They Consistent with Principles of Sensory Evaluation? *Journal of Dairy Science, 64*, 2303-2308.

Claassen, M. R., & Lawless, H. (1992). Comparison of Descriptive Terminology Systems for Sensory Evaluation of Fluid Milk. *Jorunal of Food Science, 57*, 596-600.

Delahunty, C. M., & Drake, M. A. (2004). Sensory Character of Cheese and its Evaluation. In P.F. Fox, P. L. H. McSweeney, T. M. Cogan, & T. P. Guinee (Eds.), *Cheese. Chemistry, Physics and Microbiology* (Third ed., pp. 455-487). London: Elsevier Academic Press.

Drake, M. A. (2007). Invited Review: Sensory Analysis of Dairy Foods. *Journal of Dairy Science, 90*, 4925-4937.

Elortondo, F. J. P., Ojeda, M., Albisu, M., Salmeron, J., Etayo, I., & Molina, M. (2007). Food quality certification: An approach for the development of accredited sensory evaluation methods. *Food Quality and Preference, 18*, 425-439.

Etaio, I., Albisu, M., Ojeda, M., Gil, P. F., Salmerón, J., & Elortondo, F. J. P. (2010). Sensory quality control for food certification: A case study on wine. Method development. *Food Control, 21*, 533-541.

Hammond, E., Dunkley, W., Bodyfelt, F., Larmond, E., & Lindsay, R. (1986). Report of the Committee on Sensory Data to the Journal Management Committee of the American Dairy Science Association. *Journal of Dairy Science, 69*, 298.

Hersleth, M., Ilseng, M. A., Martens, M., & Næs, T. (2005). Perception of cheese: A comparison of quality scoring, descriptive analysis and consumer responses. *Journal of Food Quality, 28*, 333-349.

Hickey, D. K., Kilcawley, K. N., Beresford, T. P., Sheehan, E. M., & Wilkinson, M. G. (2006). The influence of a seasonal milk supply on the biochemical and sensory properties of Cheddar cheese. *International Dairy Journal, 16*, 679-690.

Hough, G., Sanchez, R. H., Garbarini de Pablo, G., Sanchez, R. G., Calderon Villaplana, S., Gimenez, A. M., & Gambarot, A. (2002). Consumer Acceptability Versus Trained Sensory Panel Scores of Powdered Milk Shelf-Life Defects. *Journal of Dairy Science, 85*, 2075-2080.

International Dairy Federation. (1987). IDF 99A Sensory evaluation of dairy products. International Dairy Federation. Brussels, Belgium, International Dairy Federation.

International Dairy Federation. (1997). IDF 99C Sensory evaluation of dairy products by scoring. International Dairy Federation. Brussels, Belgium, International Dairy Federation.

International Organization for Standardization, & International Dairy Federation. (2009a). ISO 22935-3 / IDF 99-3 Milk and milk products - Sensory analysis - Part 3: Guidance on a method for evaluation of compliance with product specifications for sensory properties by scoring. Geneva, Switzerland

International Organization for Standardization, & International Dairy Federation. (2009b). ISO 22935-2/ IDF 99-2 Milk and milk products - Sensory analysis -Part 2: Recommended methods for sensory evaluation. Geneva, Switzerland

International Organization for Standardization, & International Dairy Federation. (2009c). ISO 22935-1 / IDF 99-1 Milk and milk products - Sensory analysis - Part 1: General guidance for the recruitment, selection, training and monitoring of assessors. Geneva, Switzerland

International Standardisation Organisation. (2008). ISO 9001 Quality management systems - Requirements. Geneva, Switzerland

International Standardisation Organisation. (2009). ISO 22935-3 / IDF 99-3 Milk and milk products - Sensory analysis - Part 3: Guidance on a method for evaluation of compliance with product specifications for sensory properties by scoring. Geneva, Switzerland

International Standardisation Organisation. (1992). ISO 5492 Sensory analysis - Vocabulary. Geneva, Switzerland

International Standardisation Organisation. (2003). ISO 5492 Sensory analysis — Guidelines for the
use of quantitative response scales. Geneva, Switzerland

International Standardisation Organisation. (1988). ISO 8589 Sensory analysis - General guidance for the design
of test rooms. Geneva, Switzerland

Ismail, B., Haffar, I., Baalbaki, R., & Henry, J. (2001). Development of a total quality scoring system based on consumer preference weightings and sensory profiles: application to fruit dates (Tamr). *Food Quality and Preference, 12*, 499-506.

King, S., Gillette, M., Titman, D., Adams, J., & Ridgely, M. (2002). The Sensory Quality System: a global quality control solution. *Food Quality and Preference, 13*, 385-395.

Kraggerud, H., Skeie, S., Høy, M., Røkke, L., & Abrahamsen, R. K. (2008). Season and ripening temperature influence fatty acid composition and sensory properties of semi-hard cheese during cheese maturation. *International Dairy Journal, 18*, 801-810.

Lawless, H. (1995). Dimensions of sensory quality: A critique. *Food Quality and Preference, 6*, 191.

Lawless, H., & Claassen, M. R. (1993). Validity of Descriptive and Defect-oriented Terminology Systems for Sensory Analysis of Fluid Milk. *Jorunal of Food Science, 58*, 108-112.

Martens, H., & Martens, M. (2001). *Multivariate analysis of Quality – an introduction.* Chichester, UK: Wiley.

Martens, H., & Næs, T. (1989). *Multivariate Calibration*. Chichester, UK: Wiley.

Martinsdóttir, E., Sveinsdóttir, K., Luten, J., Schelvis-Smit, J. B., & Hyldig, G. (2001). *QIM Eurofish Sensory evaluation of fish freshness. Reference manual for the fish sector.* IJmuiden, Netherlands: QIM Eurofish.

McBride, R. L., & Hall, C. (1979). Cheese grading versus consumer acceptability: An evitable discrepancy. *The Australian Journal of Dairy Technology,* 66-68.

McEwan.J.A. (1996). Preference mapping for product optimization. In T. Næs & E. Risvik (Eds.), *Multivariate Analysis of Data in Sensory Science* (pp. 71-102). Amsterdam: Elsevier Science.

Muir, D. D., Hunter, E. A., Banks, J. M., & Horne, D. S. (1995). Sensory properties of Cheddar cheese: changes during maturation. *Food Research International, 28*, 561-568.

Næs, T., Brockhoff, P., & Tomic, O. (2010). *Statistics for Sensory and Consumer Science* (1 ed.). Chichester, UK: Wiley.

O'Mahony, M. (1979). Psychophysical Aspects of Sensory Analysis of Dairy Product: A Critique. *Journal of Dairy Science, 62*, 1954-1962.

Pecore, S., & Kellen, L. (2002). A consumer-focused QC/sensory program in the food industry. *Food Quality and Preference, 13*, 369-374.

Sidel J.L., Stone, H., & Bloomquist, J. (1981). Use and Misuse of Sensory Evaluation in Research and Quality Control. *Journal of Dairy Science, 64*, 2296-2302.

Weller, J. N., & Stanton, K. J. (2002a). The establishment and use of a QC analytical/ descriptive/ consumer measurement model for the routine evaluation of products at manufacturing facilities. *Food Quality and Preference, 13*, 375-383.

Weller, J. N., & Stanton, K. J. (2002b). The establishment and use of a QC analytical/ descriptive/ consumer measurement model for the routine evaluation of.. *Food Quality and Preference, 13*, 375-383.

Table 1:
Time schedule for cheese samples in data set 1. Number of cheese samples at different times (counted in weeks from the start of the experiment).

	Weeks from start																						
	0	**4**	**8**	**12**	**16**	**20**	**24**	**28**	**32**	**36**	**40**	**44**	**48**	**52**	**56**	**60**	**64**	**68**	**72**	**76**	**80**	**84**	Sum
Production of cheese	12	16	12	13	12	14	12	12	12	12	12	14											153
Sampling 8 weeks			12	16	12	13	12	14	12	12	12	12	12	14									153
Sampling 24 weeks							12	16	12	13	12	14	12	12	12	12	12	14					153
Sampling 40 weeks											12	16	12	13	12	14	12	12	12	12	12	14	153
Sum samples for analysis			12	16	12	13	24	30	24	25	36	42	36	39	24	26	24	26	12	12	12	14	459

Table 2
Data set 1. Average quality scores for each cheese storage group: 8, 24 and 40 weeks. Significance level for difference between age groups – ANOVA main factors sample and age. (*** p<0.001, ** p<0.01, * p<0.05). Pairwise comparisons marked with letters (Tukey Simultaneous Tests, p<0.05). Number of samples in Class 1 and 2, and the proportion og class 2 samples are given.

Weeks	Overall Quality ***	Appearance ***	Consistency ***	Flavour ***	N ≥2,75p (Class 1)	N <2,75p (Class 2)	Percentage Class 2
8	3.0 [a]	3.3 [a]	3.2 [b]	3.2 [a]	401	58	**!Syntaksfeil, :**
24	3.0 [a]	3.4 [a]	3.3 [a]	3.2 [a]	397	61	**!Syntaksfeil, :**
40	2.8 [b]	3.2 [b]	3.1 [c]	3.0 [b]	348	111	**!Syntaksfeil, :**

Table 3:
Data set 2. Average consumer Overall liking score for each sample (n=342). Tukey's significant difference test, p<0.05 marked with letters in superscript (same letter = no significant difference). Average expert assessor quality score (QS) – overall quality, with defect specification for samples with score <4.

Sample	Consumers Overall liking	QS Overall quality	QS Defects
A	5.69 [ab]	3.5	Bitter
B	6.11 [a]	2.9	Many small eyes, sour
C	4.27 [c]	2.8	Firm, non-typical flavour
D	5.65 [ab]	4.0	
E	5.71 [ab]	2.5	Doughy, sour, bitter
F	5.91 [a]	3.3	Doughy, sour, bitter
G	5.33 [b]	2.9	Doughy, sour, bitter, salty

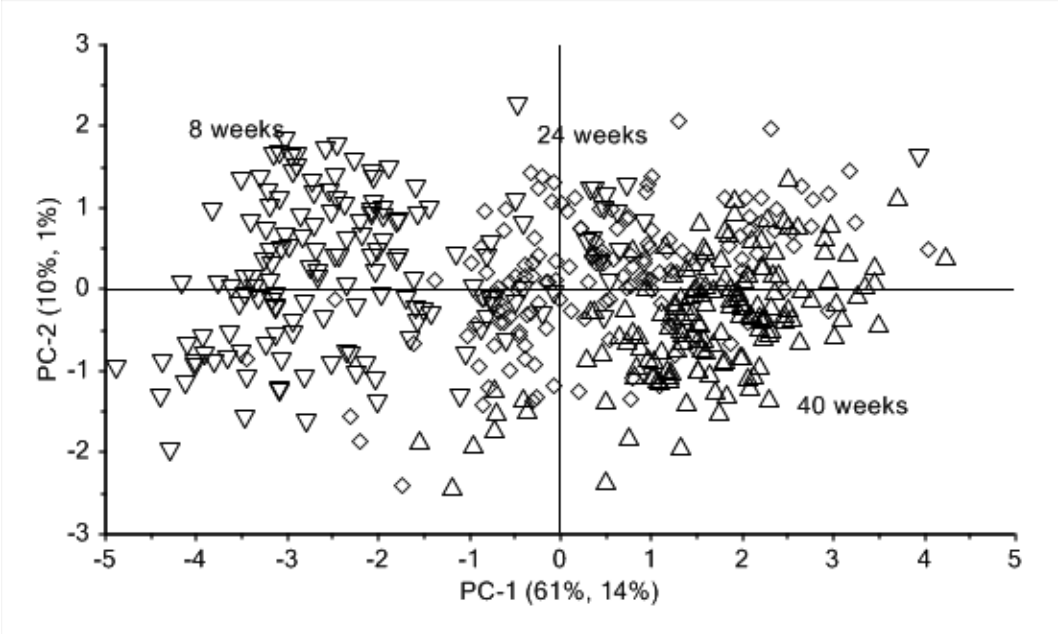

Fig 1: Data set 1. PCA score plot, 459 samples. The symbols represent stage of maturation: Inverted triangle 8 weeks, diamond 24 weeks, and triangle 40 weeks. Explained variance PC1: 61%, PC2: 10%.

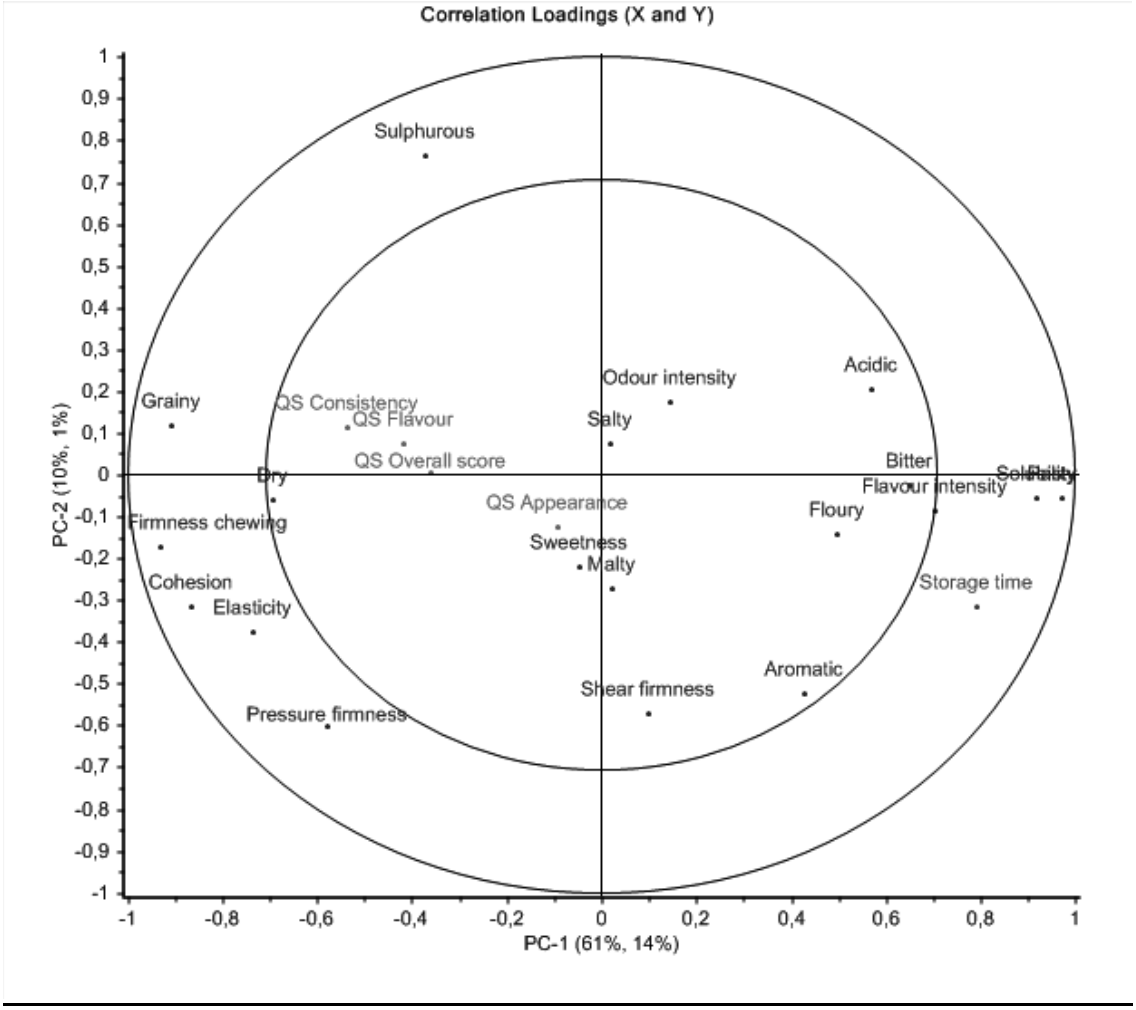

Fig 2: Data set 1. Correlation loadings from Principal component regression, PCR, sensory description(X) vs quality scoring(Y). The inner circle correspond to 50%, and outer circle 100% explained variance. Storage time is included as a passified variable. Explained variance in the 2 first components: X: 61%+10%, Y: 14%+1% .

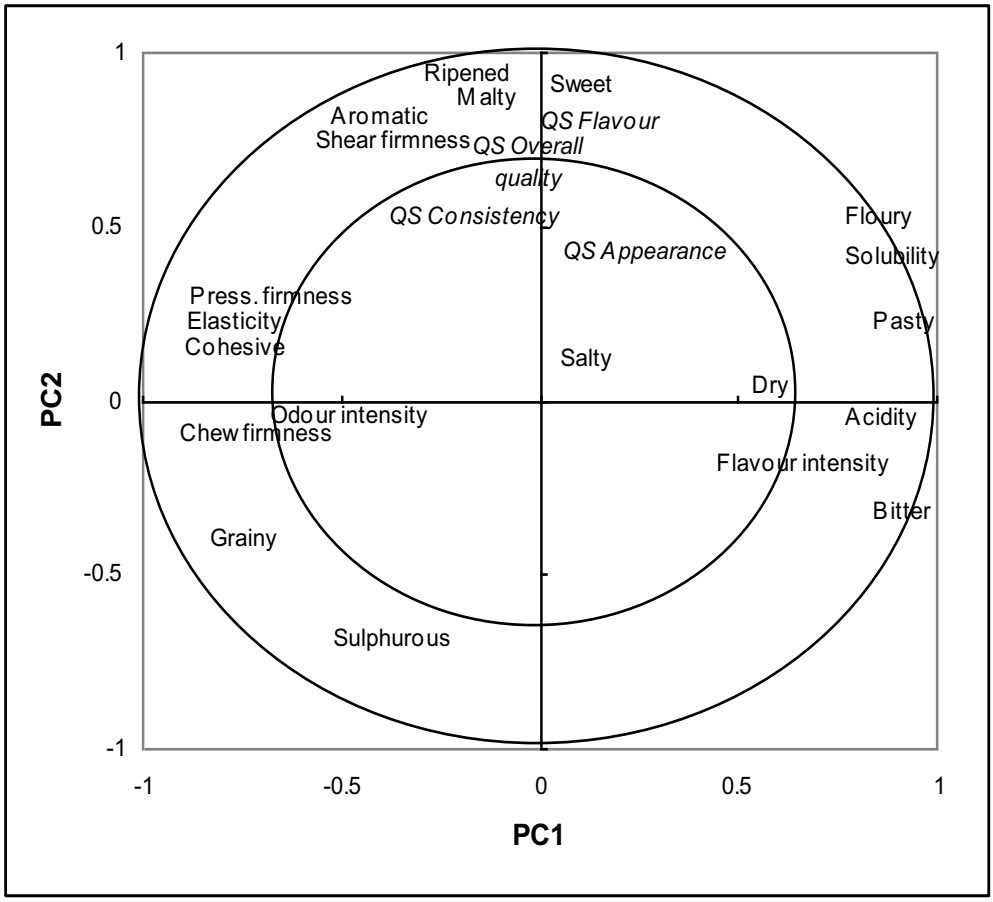

Fig 3: Data set 2 . Correlation loadings from PCR. PC1+2. Sensory description (X) vs Quality scoring (Y) , 7 samples. The inner circle correspond to 50%, and outer circle 100% explained variance. Explained 89% in X, 61% in Y after 2 factors.

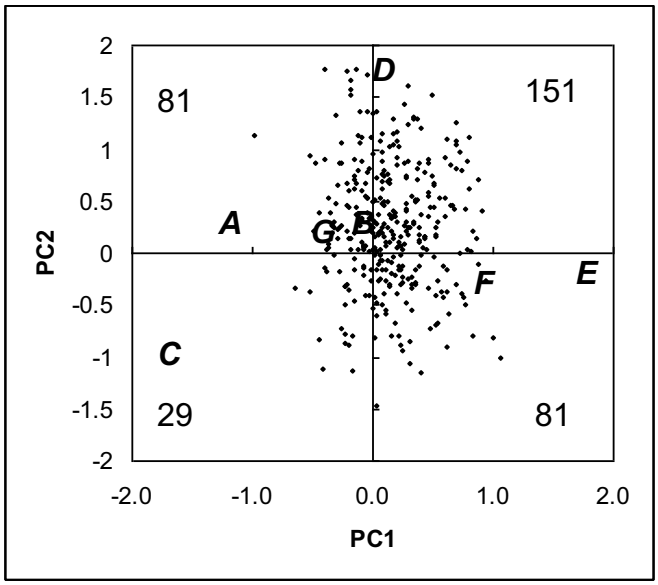

Fig 4: Data set 2. External preference map – PCR with 342 consumers. Explained variance in Y (consumer scores) PC1 23%+ PC2 18%+ PC3 17%. Each consumer shown as a dot, sample positioning marked with letters A-G.

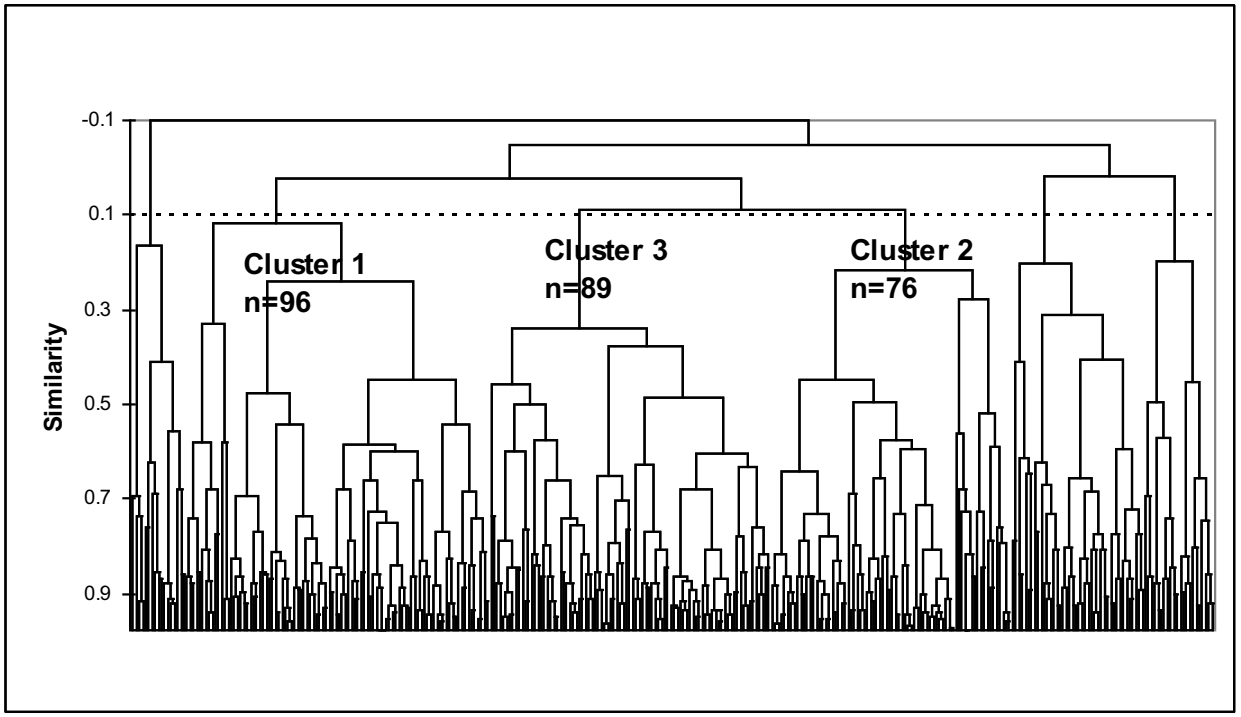

Fig 5: Similarity dendrogram, made by Hierarchical clustering. Number of consumers in each of the three largest clusters are 96, 76, and 89 respectively. Number of consumers in each of the three largest clusters are 18, 22, and 41.

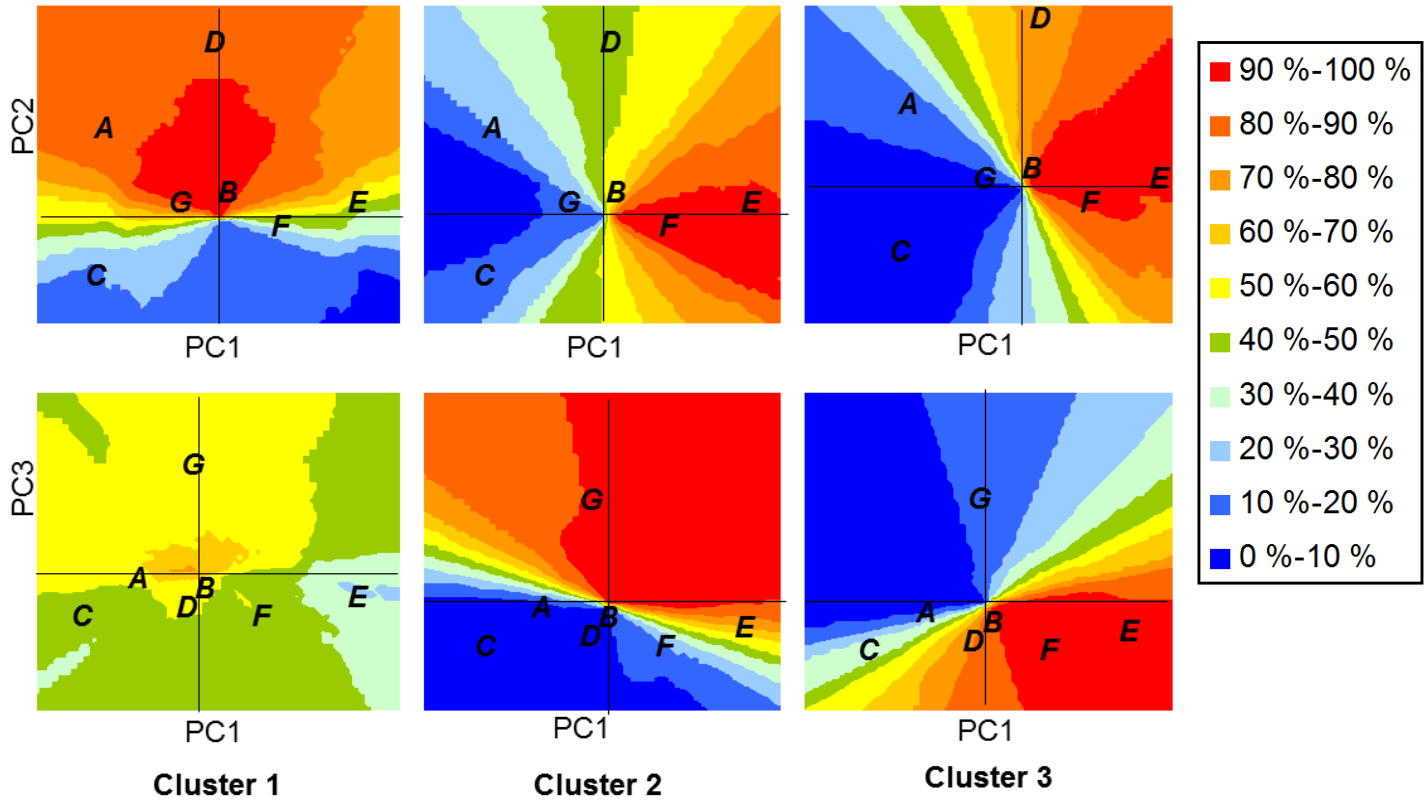

Fig 6: External Preference map (PREFMAP). Contour plots for clusters 1-3. Upper row: PC1 vs PC2. Lower row PC1 vs PC3. Colours correspond to Percentages of probability for overall liking score > 5 on the 1-9 hedonic scale.

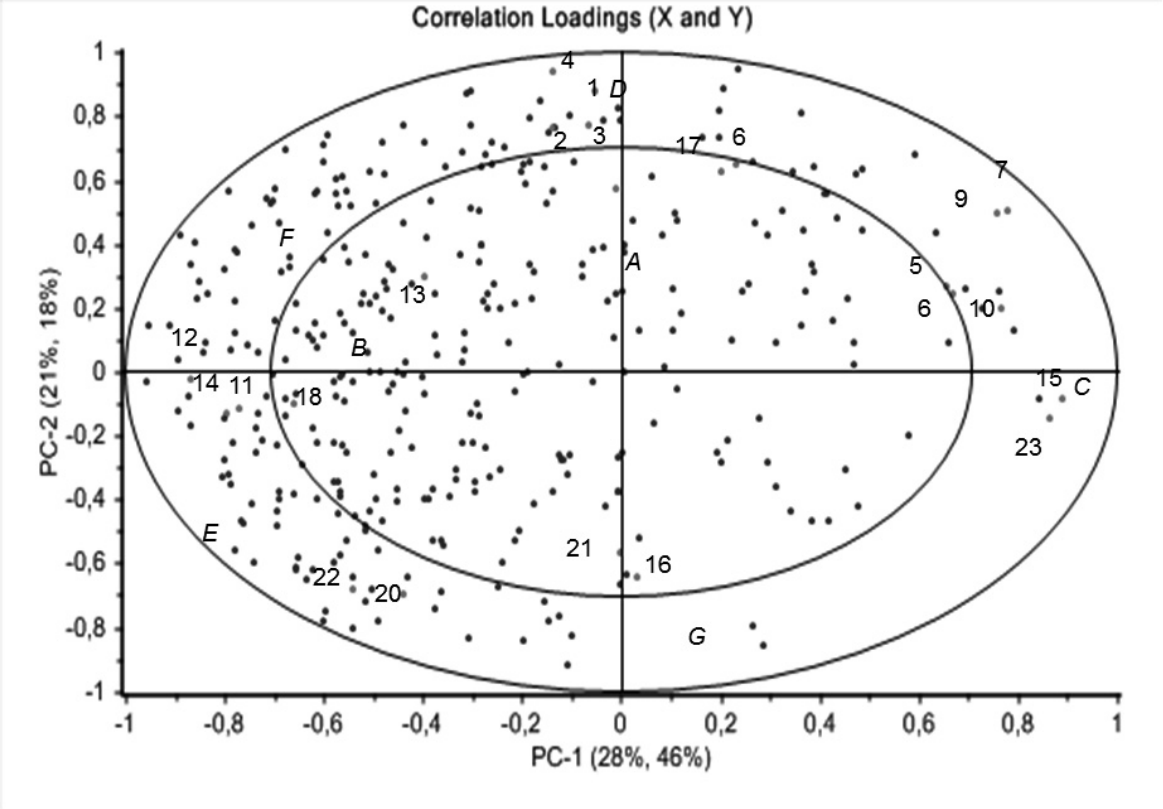

Fig 7: Internal Preference map (MDPREF) PC1 vs PC2. Data set 2, correlation loadings from PCR, X: Overall liking for each consumer (X) vs Y: Quality scores + sensory dectriptive terms. Explained variance X: 25%+20%, Y: 36%+28% . Each consumer shown as a dot, and sensory variables with numbers: 1:QS Overall quality 2:QS Appearance 3:QS Consistency 4:QS Flavour 5:Pressure firmness 6:Shear firmness 7:Odour intensity 8:Elasticity 9:Cohesive 10:Chew firmness 11:Pasty 12:Soluble 13:Dry 14:Floury 15:Grainy 16:Flavour intensity 17:Aromatic 18:Malty 19:Sweeet 20:Acidity 21:Salty 22: Bitter 23:Sulphurous 24:Ripened. Corresponding sample positioning with letters A-G.

Paper IV

Prediction of sensory quality of cheese during ripening from chemical and spectroscopy measurements

Hilde Kraggerud*[1 2], Tormod Næs[3], and Roger K Abrahamsen[1]

[1] Norwegian University of Life Sciences, Department of Chemistry, Biotechnology and Food Science, N-1432 Ås, Norway, [2] TINE R&D Center, Box 8034, N-4068 Stavanger, Norway, [3] Nofima, N-1430 Ås, Norway

* Corresponding author: *Tel*: +4795230454, *Fax*: +4722967281.
E-mail address: hilde.kraggerud@tine.no (H. Kraggerud)

ABSTRACT

An extensive material of 459 samples of Norvegia cheese was analysed during maturation using a broad range of analytical methods. From 8 to 24 and 40 weeks there was a highly systematic development in chemical and sensory attributes. Chemical data and FTIR measurements gave almost equivalent validation results in prediction of sensory data. Fluorescence spectroscopy and NIR spectroscopy, both performed on the surface of cheese, gave slightly less valid results than FTIR. Using a combination of spectra from all instruments gave higher correlations than spectra from instruments taken one by one. Replacement of sensory measurements would hardly be possible, but chemical analysis can be replaced by spectroscopy to a large extent. Sensory characteristics at a greater age (40 weeks) were not very well forecasted from early measurements on cheese (8 weeks), which could have been useful for prediction of quality development during cheese maturation.

Keywords: Cheese, Sensory analysis, chemical analysis, spectroscopy

1. INTRODUCTION

For quality control predicting product quality as early as possible in the production process is of major interest for all products. For a product like cheese, with a long ripening period before the product is ready for sale, it is especially interesting to be able to predict the product quality at an early stage, months before the actual consumption of the food. The ultimate goal would be to be able to predict the sensory quality - which is the nearest one could come to consumers' experience of the product - with non-destructive rapid methods.

The sensory characteristics of cheese are very complex and are, among other factors, a result of the ripening process of the particular cheese variety. Initial milk quality and composition and the particular cheese making process used result in a certain chemical composition of the fresh cheese, and biochemical processes like proteolysis, glycolysis and lipolysis are of importance. Factors that affect the quality of cheese have been extensively reviewed (Collins, McSweeney, & Wilkinson, 2003; Delahunty & Drake, 2004; Fox & Cogan, 2004; McSweeney & Sousa, 2000).

The complex task of analysing cheese to determine composition and monitor ripening by chemical and instrumental approaches has recently been updated in a review (Subramanian & Rodriguez-Saona, 2010). Many attempts have been made to predict chemical or sensory properties of cheese by spectroscopic methods. A number of these have examined relatively few samples, often with large variation in the sample set. Rather few studies are based on comparison with reference measurements, while many studies are based on classifications and detailed examination of electromagnetic spectra. Often samples have considerable differences in the major chemical components known to affect the ripening process

(Beresford, Fitzsimons, Brennan, & Cogan, 2001; Gilles & Lawrence, 1973; Johnson & Law, 1999). For use in industrial quality control, the methods must be able to show small deviations from normal composition.

Characterisation of cheese during ripening using spectroscopic methods has been reviewed by several authors (Karoui & De Baerdemaeker, 2007; Karoui, Mazerolles, & Dufour, 2003). Fourier Transform Infrared (FTIR) has become more or less a "standard method" for rapid and accurate determination of major chemical components in the dairy industry. The focus has been on analysis of macromolecules in cheese (Chen, Irudayaraj, & McMahon, 1998; McQueen, Wilson, Kinnunen, & Jensen, 1995; Rodriguez-Saona, Koca, Harper, & Alvarez, 2006). More recently characterisation of ripening has also been the subject of research (Cattaneo, Giardina, Sinelli, Riva, & Giangiacomo, 2005; Chen et al., 1998; Fagan, Everard, O'Donnell, Downey, Sheehan, Delahunty, & O'Callaghan, 2007; Fagan, O'Donnell, O'Callaghan, Downey, Sheehan, Delahunty, Everard, Guinee, & Howard, 2007; Karoui, Mazerolles, Bosset, De Baerdemaeker, & Dufour, 2007; Koca, Rodriguez-Saona, Harper, & Alvarez, 2007; Kocaoglu-Vurma, Eliardi, Drake, Rodriguez-Saona, & Harper, 2009; Martín-del-Campo, Picque, Cosío-Ramírez, & Corrieu, 2007; Rodriguez-Saona et al., 2006; Subramanian, Harper, & Rodriguez-Saona, 2009a; Subramanian, Harper, & Rodriguez-Saona, 2009b; Subramanian, Alvarez, Harper, & Rodriguez-Saona, 2011)

Near infrared (NIR) spectroscopy has also been used for different components and sensory attributes of cheese (Adamopoulos, Goula, & Petropakis, 2001; Blazquez, Downey, O'Callaghan, Howard, Delahunty, Sheehan, Everard, & O'Donnell, 2006;

Cattaneo et al., 2005; Downey, Sheehan, Delahunty, Callaghan, Guinee, & Howard, 2005; González-Martín, González-Pérez, Hernández-Hierro, & González-Cabrera, 2008; Karoui, Mouazen, Dufour, Pillonel, Schaller, De Baerdemaeker, & Bosset, 2006; Skeie, Feten, Almøy, Østlie, & Isaksson, 2006; Wittrup & Nørgaard, 1998)

Fluorescence spectroscopy has been widely used for prediction of light-induced oxidation with reference measurements including sensory analysis, reviewed by Andersen & Mortensen (2008). But fluorescence spectroscopy has also been regarded as a promising tool for exploring texture development during ripening (Dufour, Devaux, Fortier, & Herbert, 2001; Dufour, Mazerolles, Devaux, Duboz, Duployer, & Mouhous Riou, 2000; Garimella Purna, Prow, & Metzger, 2005; Karoui & De Baerdemaeker, 2007; Karoui & Dufour, 2006; Karoui & Dufour, 2003; Karoui, Dufour, & De Baerdemaeker, 2007; Karoui, Dufour, Pillonel, Schaller, Picque, Cattenoz, & Bosset, 2005; Karoui, Dufour, Schoonheydt, & Baerdemaeker, 2007; Karoui et al., 2003).

Still there is a long way to go to have established, valid methods of analysis for predicting cheese quality for practical use in the dairy industry, where variation between productions normally is relatively small. Our approach attempted to simulate commercial production conditions and a variety of situations. A large number of cheese samples (153) were analysed at 3 points during maturation (8, 24 and 40 weeks) making 459 samples altogether. Variations in major chemical composition were limited. Production was made during all seasons of the year, with sampling and analysis at 20 points during a period of 18 months. A broad range of chemical,

sensory and spectroscopic methods was applied in order to be able to compare a number of possible methods.

The objectives were to study the development in sensory and chemical parameters during cheese ripening, and to examine the feasibility of different spectroscopic techniques (NIR, FTIR, fluorescence) for the monitoring of sensory attributes. Relationships between chemical and sensory measurements, and the possibility of early prediction of sensory quality development during ripening of cheese was also examined. Using realistic conditions with small variations in major components and a high number of samples was essential in this work.

2. MATERIALS AND METHODS

2.1. Cheese sample selection

A Dutch-type, semi-hard, rindless cheese variety, Norvegia, with approximately 46% fat in dry matter, was evaluated. During the first 3-4 weeks after cheesemaking, all cheeses undergo ripening stages at set temperatures – varying from plant to plant - and adjusted to acquire the desired quality. After 3-4 weeks from cheese making, the cheese was matured at ≤ 4°C.

Cheese included in this study came from three coordinated experiments. Experiment 1 was an experiment with varying raw milk quality and age in our large scale pilot plant (120 kg cheese per batch). In experiment 2 samples from commercial production were collected on the first day and manipulated with respect to ripening temperatures (Kraggerud, Skeie, Høy, Røkke, & Abrahamsen, 2008). In experiment 3 random samples were collected from commercial production at 6 different cheesemaking plants. Each sample consisted of four vacuum packaged 5 kg

cheeses, 3 of them used for analysis at different ages. The samples were produced during one year, and there were 6 samplings with 8 weeks between each. Throughout a period of 18 months a number of samples were analysed every 4 weeks. A production and sampling scheme is shown in Table 1.

2.2. Sampling procedures

Sampling of cheese for analysis was made in accordance with IDF Standard 50C (International Dairy Federation, 1995). Spectroscopy, pH measurement and sensory analysis were performed within a few days at each sampling point. Fat, dry matter and salt were determined only at one ripening point, after 8 weeks, as they are considered not to change during ripening. Other chemical analyses (not all analyses were performed at all sampling points) were conducted on grated samples frozen at ≤-20ºC until analysis.

2.3. Spectroscopy

NIR spectroscopy

NIRSystems 6500 scanning instrument was used for measuring NIR spectra (FOSS NIRSystems Inc., Silver Spring, MD, USA). The samples were scanned in a reflectance mode using a fibre-optic probe directly in contact with polyethylene covered surface of cheese samples. The instrument was operated in the spectral region 400–1100 nm and measured at 2 nm steps. The instrument's ceramic plate was used as reference for calibration. Sample temperature was between 4 and 8 °C. Each cheese sample was scanned three times at different spots of the cheese surface, and the average sprectrum was used for all further data analysis. Spectra were treated with Multiplicative Scatter Correction (MSC) (Martens & Næs, 1989), a

transformation method used to compensate for additive and/or multiplicative effects in spectral data.

FT-IR spectroscopy

Pre-processing of samples was done according to standard methods, using solvents and equipment delivered by the instrument supplier, Foss Instruments (Foss, Hillerød, Denmark). Two parallel samples of grated cheese were dissolved in a proportion of 1/10 in the solvent LOSsolver Cheese tempered in a waterbath at 43°C, and using a LOSmixer with a cooling device, regulated at 45°C, to homogenize the mixture for 1 minute. Antifoam-Y30 was used to prevent foaming. Sample bottles were then placed in the waterbath at 43°C for 14 minutes in order to remove air bubbles.

FTIR transmittance spectra were the recorded using a Milkoscan FT120 spectrophotometer (Foss, Hillerød, Denmark). The instrument was equipped with a fat homogenization module, a flow cell of path length 50μm and an automated sampler. Spectra were recorded in the range of 974 to 5000 cm^{-1} with a resolution of 3.858 cm^{-1}. Duplicate spectra of each sample were recorded on each of two parallel samples, and the average of these four spectra used for further analysis. Transmittance values were converted to absorbance values by the transformation A = log (1/T) where A is absorbance, and T transmittance. The spectral regions retained for data analysis were 974-1593 cm^{-1} and 1700-2986cm^{-1}, because of heavy interference from water in the region 1600-1700 cm^{-1}. Eight samples' spectra (analyzed in sequence on the same day) were found to be outliers, and were omitted from the dataset, leaving 451 samples with FTIR spectra for further modelling.

Fluorescence spectroscopy

Fluorescence was measured on the surface of cheese, taking the average signal over a surface of approximately 20 cm^2. The equipment for measuring consisted of xenon lamp (Oriel 6258, Oriel Corporation, Stratford, CT), the excitation light was generated by a 300 WXenon light source and a 10-nm bandwidth interference filter (Oriel 59920). An optical ”cut-off” filter (Melles Griot 03FCG049) for filtering the excitation signal and a CCD camera (Princeton TEA/CCD-512-TKBM1, Princeton Instruments Inc., Trenton, NJ) with spectrograph for measuring the emission spectra (Acton SP-150, Acton Research Corp., Acton, MA). Emission spectra in the region 287-796 nm were measured for each of the excitation wavelengths, 280, 325, 382 & 450 nm. Exposure time was 1 s for all spectroscopic measurements. Two parallel measurements per sample were averaged in further analysis, then normalized to a total area of 1 below each spectrum.

2.4. Chemical analysis

Dry matter was determined after 8 weeks according to IDF Standard 4A (International Dairy Federation, 1982), salt determined according to IDF Standard 88A (International Dairy Federation, 1988) and fat according to IDF Standard 5B (International Dairy Federation, 1986). pH was measured with calibrated PH-meter Radiometer PHM 210 with electrode GK2401c (Nerliens Mezansky AS, Oslo, Norway) on 25g of grated cheese with 4-5ml of deionized water added. For determination of total nitrogen IDF Standard 20B (International Dairy Federation, 1993) was used with Kjeldahl apparatus Foss Kjeltech 2400 and Foss Tecator Digestor Auto (Foss Tecator, Höganäs, Sweden). Amino nitrogen(AN) and Soluble

nitrogen(SN) were analysed according to (Mogensen, 1948). Amino nitrogen was analysed using a formol titration method with the following steps: 1. Neutralization of cheese solution with 0,25M NaOH with phenolphthalein as indicator. 2. Addition of 35% formaldehyde in the neutralized mixture, and retitration with 0,1M NaOH. Soluble nitrogen was determined using Kjeldahl apparatus (Tecator, Höganäs, Sweden) after precipitation with 1M HCl and filtration.

Organic acids were analysed with HPLC as described by (Skeie, Lindberg, & Narvhus, 2001). Perkin Elmer Series 200 equipment with UV Spectrophotometric Detector was used (Perkin Elmer, Norwalk, USA). The following components were determined in g kg^{-1}: pyruvate, succinic acid, lactic acid. Volatile compounds were separated from samples at 90º C in a gas-tight bottle, and the headspace analysed using Perkin Elmer Autosystem XL Gas Chromatography with Headspace Sampler HS40, headspace injector HS-101, column 6'1/8" 0.2 % Carbowax 1500 80/100 Carbopack C with FID detector (Perkin Elmer, Norwalk, USA). The carrier gas was nitrogen and the detector gas hydrogen and air. The following components were determined (in relative units): acetaldehyde, ethanol, acetone, i-propanol, n-propanol, 2-butanone, diacetyl,2-butanol and acetoin. Volatiles were extracted from the samples in ether and the ether phase analysed with gas chromatography in a packed glass column with FID detection, GC Perkin Elmer Autosystem gas chromatography, column Supelco 2m ¼inch YD 2mm glass column, with Supelco GP 10% SP 1000/1% H_3PO_4 on 100/120 Chromosorb WAW, and FID detector (Perkin Elmer, Norwalk, USA). The carrier gas was nitrogen and the detector gas hydrogen and air. The components determined in mmol kg^{-1} were: acetoin, acetic acid, propionic acid and butyric acid. Fatty acid composition analysis by GC was applied as described by

Kraggerud et al. (2008). Analysis of free amino acids was done according to the procedure described in Skeie et al. (2006).

2.5. Sensory analysis

Quality scoring

Quality scoring (QS) was performed using an internal TINE-method based on the reference method IDF 99C:1997, Sensory evaluation of dairy products by scoring, Part I and Part IV (International Dairy Federation, 1997). At least 3 authorized expert assessors were used per session. To be authorized as an assessor in TINE, a candidate has to participate in at least 20-30 assessment sessions in the course of about one year. In Authorization Test congruence and correlation with reference assessors and the candidate's reproducibility and repeatability are the most important criteria used. Appearance, consistency and odour/flavour as well as overall quality score were assessed, using a numerical interval scale from 1 to 5, with 0.5 point intervals. The scale is defined according to IDF 99C, Part 1 clause 9.2 (International Dairy Federation, 1997) as follows: 5 = conformity with the pre-established sensory specification (PS), 4 = minimal deviation from PS, 3 = noticeable deviation from PS, 2 = considerable deviation from PS, 1 = very considerable deviation from PS. The score 0 – which was defined in IDF99C (unfit for human consumption) was not used. When an assessor scored <4 points for a sample, a description of the deviation was given, using predefined nomenclature of defect terms.

Blocks of cheese (5 kg each) at a temperature of 14 ºC ± 2ºC were presented to the assessors in random order and placed in a row on a long table. Before sensory assessment, each cheese block was cut in two. Sensory assessment was carried out

using a standard cheese slicer for cutting. Each session started with calibration of the assessors with two different cheeses as reference samples. The assessors got no information about the samples and were not allowed to communicate during the following assessment. There was one replicate of each cheese sample, owing to the great number of samples.

Descriptive sensory analysis

The vocabulary of sensory attributes used in this experiment is in accordance with ISO 5492 (International Organization for Standardization, 1992). The terms to be included or excluded is evaluated for each data set, depending on the needs of the specific experiments. The attributes have earlier been described in more detail (Kraggerud et al., 2008). An interval scale from 1 to 9 points was used for each sensory attribute, as defined in ISO 4121 (International Organization for Standardization, 2003). An internal laboratory panel was used comprising six assessors, well trained on cheese for several years. Before each session, a calibration session was performed in order to obtain an agreement on the use of attributes and scales, using two different cheeses as test samples. All assessments were conducted in individual booths at the sensory laboratory, designed in compliance with international standards for test rooms, ISO 8589 (International Organization for Standardization, 1988). Cheese samples were held at 14 ºC ± 2ºC during to assessment. Order of assessment of the samples was randomized for each assessor. Each sample was assessed in one replicate, owing to the great number of samples. The scores of each sample were averaged over the assessors.

2.6. Analysis of data

Basic statistics and analysis of variance (ANOVA) were carried out using software package Minitab® 15.1. (MINITAB Inc., State College, PA, USA). Multivariate analysis was performed using Unscrambler version 10.1. (CAMO Software AS, Oslo, Norway), with several “standard” methods like principal components analysis (PCA) and partial least squares regression (PLS) using NIPALS algorithm (Martens & Næs, 1989). Validation of models and choice of optimal number of components were done using Cross Validation with uncertainty testing (Martens & Martens, 2000). Correlation coefficient R and root mean square error of prediction (RMSEP) were used to evaluate the goodness of fit of models (values from validation used in all cases) (Martens & Næs, 1989). Standard Error of Cross Validation (SECV), frequently used by other authors is essentially the same parameter as RMSEP.

Standard error, s_{error}, was computed for each sensory variable as standard error of the mean ($s/\sqrt{n}$) averaged over all samples assessed, using MINITAB.
Ratio Performance Deviation (RPD), (Williams & Sobering, 1993) was calculated according to this equation:

RPD = StDev(ref)/RMSEP

where StDev(ref) is the standard deviation of the reference method results.
Range error ratio,

RER = (Max value-Min value)/RMSEP

(Williams, 1987) was also used for evaluating model performance.

3. RESULTS AND DISCUSSION

3.1. Changes in sensory and chemical characteristics during cheese maturation

In these experiments different analyses relevant for the characterization of cheese maturation were performed at 3 points: 8, 24 and 40 weeks of age. In addition some analyses were performed during processing and on fresh cheese. Table 2 shows basic statistics for measured sensory and chemical parameters at 8, 24 and 40 weeks. The variation in major chemical components like salt, fat, protein and dry matter was relatively narrow. The samples analysed should therefore be appropriate for the purpose of testing methods for use in ordinary production where there will be interest in detecting minor variations. Figure 1 shows the developments in sensory attributes during maturation. A systematic increase or decrease in almost all measurements during the maturation period investigated was observed. Anova was performed with sample no and age as factors and the results are presented in Table 2. Highly significant differences ($p<0,001$) in both factors were found for almost all variables. Consistency variables, such as firmness chewing, grainy, cohesion and elasticity decreased, while solubility and pasty, as well as flavour variables like flavour intensity, pungent and bitter typically increased, during maturation.

In figure 2 relative development in various chemical measurements at 8, 24, and 40 weeks age are illustrated. Anova with age and sample as factors showed that there were no significant differences for the variable propionic acid (which is not a typical component in Norvegia, and hence varied between 0.0 and 0.2 mmol kg^{-1}). All the other variables had significant age differences with $p<0.001$, except total N (Kjeldahl) which had $p=0.088$. Some of the components increased, and others decreased

during 8-40 weeks of ripening. A few components were practically constant, like total nitrogen content, lactic acid and acetic acid. During ripening the soluble N and amino acid containing (amino N) fractions increased gradually as protein was degraded during ripening. Free amino acids, illustrated in figure 3, all demonstrated a near linear increase with age. Cross correlation between all amino acids, and also with amino N, were significant, almost all of them very high (table 3). This can indicate that the pattern of protein degradation was relatively similar during maturation for these samples, even though they originated from 5 different production plants. These findings in general correspond well with what was expected, from cheese ripening biochemistry theory (McSweeney, 2004).

3.2. Regression modelling and validation

Predicting sensory variables, with their relatively low precision compared to the total variation range, is not an easy task. For calibration of models for use in industry, it would probably be best to calibrate on a set of data with larger variation, then use the calibration for prediction on subsequent samples.

There is a number of different validation parameters which can be used, among them validated RMSEP is considered to be relevant. RMSEP can be compared with the precision of the reference method, and has the same unit as the reference values, in contrast to regression coefficient, which will be related to the actual sample space. The other used validation measures in this article, RPD and RER, are also very dependent of the total sample space, but were chosen as they have been used by other authors. When it comes to comparison of results between authors, interpretation is always difficult, because number of samples, breadth of range,

standard error of the reference method and other factors affect all these validation parameters. Furthermore, there have been very few experiments with high numbers of samples which could also be an important factor influencing the values of the validation parameters.

Cross-validation was performed in this case. Explained variance with full crossvalidation (leave one out) was compared to segmented cross-validation with 20 random segments (23 samples per segment in our case). An example of explained variance results, for PLS1 regression of amino nitrogen from FTIR spectra, is illustrated in figure 4. Calibration lines for the same samples (459) are of course identical. Explained validation variance showed out to be almost identical, it is impossible to separate the two lines from each other, average difference between the two lines for the 20 components was only -0.19%.

Outliers is also an important challenge to be addressed. Eight samples' FTIR spectra, analyzed in sequence, were found to be outliers. Results from modelling AN from FTIR with and without these outliers in the dataset are also illustrated in figure 4. The explained variance was slightly higher when removing the 8 outliers, both in calibration and validation, validation variance was 74.5% vs 71.5%. The actual 8 outliers were excluded from all our modelling based on FTIR. With lower number of samples, outliers will be even more important to handle.

The number of factors in the chosen model can be of importance for the results. In PLS model components are extracted in such a way that the first factor/PC explains the largest amount of variation, followed by the second factor, etc. At a certain point,

the variation modeled by any new PC is mostly noise. The optimal number of factors - modeling useful information, but avoiding overfitting - is determined with the help of the residual validation variances. If less than 3% of the residual variance is explained in the successor factor, the optimum number of factors is chosen in Unscrambler. In our case in figure 4, optimal number of factors were 5 for the 459 sample, and 6 for the 451 sample segmented crossvalidation models. It is also necessary to view the actual explained variance curve, to see whether the chosen optimal number of factors is proper, as there can in some cases be temporary drops in the curve. When validation residual variance is minimal, RMSEP also is, and the model with an optimal number of components will have the lowest expected prediction error.

With PLS regression it is possible to model dependent variables one ny one (PLS1), or many y-variables at the same time (PLS2). Table 4 compares PLS2 results with PLS1 results for two of the sensory attributes. The same applies to table 5 and 6, for two variables in each table. PLS1 compared to PLS2 had only minor differences in RMSEP and correlation for most of the models. Some models were better with PLS1 – as could be expected. As the differences were small, we chose to use PLS2 for groups of y-variables for our main results and discussion, as we had a large number of y-variables to be modelled.

3.3. Prediction of sensory attributes

Table 4 compares validation results from PLS2 modelling of all sensory variables modelled from several chemical variables and from FTIR, NIR and fluorescence spectroscopy of the same samples. RMSEP was found to be only slightly higher than

the standard error of the reference method for many of the sensory variables. Chemical variables and FTIR gave the best, and almost the same, prediction results for many sensory variables. Correlation coefficients were relatively low, but, as shown in Figure 5, this can be caused by noisy reference data as the correlations seem clear. Fluorescence and NIR spectroscopy also showed reasonable results, but generally not as high as FTIR. This might be due to that both these methods were performed on surface of cheese, which can be a problem if the cheese is not homogeneous. FTIR was performed on ground and pre-processed cheese samples, and involve a lot more labour on each analysis.

Publications comparing FTIR and sensory variables of cheese are quite few. Kocaoglu-Vurma et al. (2009) analysed 15 samples from 3 manufacturers of Swiss cheese with FTIR and sensory descriptive analysis. They found correlations between 0.69 and 0.96 when predicting sensory variables from FTIR using PLS. Fagan et al. (2007) concluded that mid-infrared spectroscopy has the potential to predict age, SN, and several sensory texture attributes of cheddar cheese – with R up to 0.8 for different sensory attributes, in the same area as the models found here. The same group used FTIR for processed cheese and found RER values between 5 and 10 for sensory descriptive attributes, indicating that the models had good practical utility. They found RER values between 6 and 12 for texture attributes of processed cheese analysed with FTIR (Fagan, Everard, O'Donnell, Downey, Sheehan, Delahunty, Callaghan, & Howard, 2007). In the data presented here – even with the relatively narrow range of data - RER values were in the range 5-10 for almost all sensory variables, and some variables showed RER >10, which indicates a high utility value (Williams, 1987). Subramanian et al. (2009b) developed an extraction method which,

in combination with measurements with FTIR, showed promising results based on clustering of samples. The experiment consisted of 15 Cheddar samples, with no reference data except for quality classification, and no key validation data, which makes comparison with the present results difficult.

Lerma-García, Gori, Cerretani, Simó-Alfonso, & Caboni (2010) were able to classify Pecorino cheese according to both ripening time and production technique from FTIR results. It should be added that differences between individual cheeses seem relatively large, and that reference analysis measurements are not reported. Martín-del-Campo et al. (2007) used ATR-FTIR to predict ripening date of Camembert cheese, with a precision of ±1 day. FTIR has also been used to study Crescenza cheese during 20 days, and it was possible to define the critical day during shelf-life of this fresh cheese (Cattaneo et al., 2005).

NIR analysis has been used for assessment of sensory properties in a semi-hard cheese variety (Sorensen & Jepsen, 1998). Thirty-two batches of cheese were made, some with the addition of undesired bacteria, covering a wide range of pH and moisture content, and measured 4 times during ripening from 5 to 11 weeks. Squared correlation coefficients of the best predicted sensory attributes were 0.7–0.8, a little higher than our R^2 from spectroscopy. This can be due to wider ranges within the dataset which tend to generate higher correlation coefficients. There was no other parameter directly comparable with our validation parameters. Twenty Emmenthal cheeses were evaluated by NIR and sensory panel and the method found feasible for prediction of some sensory attributes (Karoui et al., 2006). Eight attributes obtained R^2 higher than 0.5. RER was between 5 and 7 for most attributes,

while one was 10 and one 3.5. RPD was mostly between 1.5 and 2. All these results were quite comparable with our results from different spectroscopic measurements of sensory attributes. Experimental Cheddar cheeses were measured by NIR and sensory evaluation and models were evaluated to be sufficiently accurate for industrial use (Downey et al., 2005). Calculated RPD and RER were comparable to our results from spectroscopic data. Blazquez et al. (2006) modelled sensory and texture parameters with NIR in experimental processed cheese with fairly high variation between samples and obtained RER values between 8 and 12.

Changes in cheese during ripening have also been described with fluorescence spectra – some of them with sensory, chemical and rheological measurements as reference - in work by Dufour et al. (2001) who found squared correlations between fluorescence and sensory texture attributes of 0.22-0.69, and for pastiness 0.89, in soft cheese. This observation also corresponds well with our findings of pasty as one of the best predicted sensory variables by all our spectroscopic measurements (Table 3), as also underlined by Lebecque, Laguet, Deveaux, & Dufour (2001) who obtained R^2 below 0.5 in prediction of sensory attributes from fluorescence tryptophane spectra, and somewhat higher using fluorescence vitamin A spectra.

3.4. Prediction of chemical parameters from spectroscopic measurements

Major components of cheese, like fat, protein and water/dry matter, are commonly measured with spectroscopic methods like FTIR and NIR (Adamopoulos et al., 2001; Chen, Kocaoglu-Vurma, Harper, & Rodriguez-Saona, 2009; González-Martín et al., 2008; Karoui et al., 2006; Wittrup & Nørgaard, 1998). They will not be discussed

here, as the variation in our dataset was limited in relation to major components. In Table 6, validation results from prediction of different minor chemical attributes are given. In PLS2 model the highest values of correlation, RPD and RER, were obtained for Amino N and Soluble N measured with FTIR, and with mixed spectra. This indicates a high capacity of prediction of important ripening characteristics. Amino acids (table 7) also showed high correlation coefficients, most in the area 0.8-0.9. All have 1.5<RPD<2 and 5<RER<10 which should indicate that they can be useful (Williams & Sobering, 1993). Modelling with PLS1, one variable at a time, might be useful, for the example Lactic acid in table 6, the highest R value raised from 0.70 to 0.88 – for the models from mixed spectra. This might also be typical for other variables, which we have not tested, as prediction of chemical variables were not the main issue.

Other authors have made attempts to measure chemical variables in addition to the major components in cheese. Skeie et al. (2006) used NIR to predict selected amino acids in Norvegia and Präst cheese. They achieved very high correlations (>0.9 for many variables) compared with HPLC/standard method. Our data all over showed a higher RMSEP values, the best fit, from mixed spectra, slightly higher. The correlations in our dataset were also lowe, but all higher than 0.74 with mixed spectra. In Skeie's data the optimum number of PLS components was much higher. Our results in this case were lowest for NIR. In our experiment we used NIR measurements from 400-1100 nm, whereas Skeie used the region from 780-2500 nm. This indicates that the higher part of the NIR spectrum could be used for higher prediction ability. Another difference in the NIR measurements was that scanning was made directly on the cheese surface – non-destructive – while Skeie made

measurements on grated cheese. Grated cheese might result in lower variation in the sample between reference measurement and spectroscopy, as the sample is then mixed and its parts are as uniform as possible. It is known that the variation inside one cheese can be of importance. On the other hand, non-destructive measurements would be a great advantage, and we have reason to believe from other results that it is possible to improve the method to obtain acceptable results also with limited variation in the samples.

Karoui et al. (2006) measured NPN, TN, and SN by NIR. RMSP and RPD values were comparable with values obtained in our experiment but, in this case also, correlation coefficients in the present experiments were lower – probably the much greater number of samples can be a reason for this.

Koca et al. (2007) applied FTIR for monitoring of short-chain free fatty acids (FFA) in Swiss cheese. The range of FFA was much higher than in our experiment. SECV, recalculated to mmol/100g, was quite comparable for acetic acid measured with FTIR. (See our results expressed as RMSEP values in table 6). FTIR was used for monitoring of amino acids, organic acids and ripening changes in water-soluble fractions of 12 Cheddar samples at 5 points during ripening (7-73 days) (Subramanian et al., 2011). They could predict amino acids and organic acids, lactic, formic and oxalic, with correlation coefficients of 0.89 and higher. RER, based on range were, in their case, between 10 and 20 in most components which is higher than in our experimental data. We had trouble comparing our amino acid range and RMSEP values with the SECV quoted in Table 1 in the reference article, but given the concentrations in Table 3 in the same article the results seem more comparable

with our data. An example of this contradiction: In table 1, the range of Alanine was stated 43.4-363.0 nmol g^{-1} cheese, whereas in table 3 Alanine average day 3 was 4.3 nmol g^{-1} cheese and day 73 14.5 nmol g^{-1} cheese, both outside the given range in table 1. It is again difficult to compare validation parameters directly – as our dataset is much more comprehensive when it comes to number of samples.

3.5. Early prediction of sensory quality

In Table 5, validation results are given from prediction of sensory attributes at 40 weeks of age from different measurements at 8 weeks (number of samples 153, as there is only one age group). Correlation coefficients were generally lovwer, but RMSEP results were comparable, even a little better for some attributes compared to predictions on the 8 week samples (table 4). This is consistent with what was found in 3.1 (above), development during cheese ripening is on average linear. But the individual differences per sample is more difficult to predict, making correlations lower. The potential with respect to early prediction of sensory quality is therefore uncertain. This can also be interpreted from the correlation loading plot in figure 9. Quality scoring attributes from 8 weeks variables are found in dimension 2, while the corresponding 40 weeks QS variables are found to the left (dimension 1), which indicate these variables are non-correlated. We see that most of the sensory descriptive terms are in the same direction both at 8 and 40 weeks, which indicates that they were better correlated, even though all correlation coefficients in genereal were low, which might also be caused by noise in the sensory measurements.

4. CONCLUSIONS

An extensive assembly of 459 samples of Norvegia, a Dutch-type semi-hard cheese, was analysed using a number of chemical, chromatographic, sensory and

spectroscopic methods during maturation. From 8 to 24 and 40 weeks of age there was a highly systematic development in chemical and sensory parameters.

Sensory attributes pasty, grainy, solubility, cohesion, firmness on chewing, flavour intensity, pungent and bitter obtained the highest correlations with chemical variables modelled by PLS2. Noise in the sensory data and the very high number of samples seem to be the cause of relatively low correlation coefficients. The sensory attributes with the best validation results provided slightly lower correlations with FTIR measurements than with chemical analysis. Fluorescence spectroscopy and spectroscopy using NIR between 400nm and 1100nm, performed on the surface of cheese, gave slightly less valid results than FTIR for measurement of sensory variables in this experiment. Using a combination of spectra from all instruments gave a higher correlation than spectra from instruments taken separately.

Results from a mixture of spectroscopic measurements, and also FTIR, NIR and fluorescence alone, were promising in order to replace the more onerous sensory measurements, but there is still some way to go with our methods. However, many authors has shown very promising results, which indicate spectroscopy as a relevant supplement. On the other hand chemical results can be replaced to a large extent by spectroscopy, with accurate results. The advantages will be fast results, which makes direct utilization of results in production possible. The spectroscopy methods are also normally labour-saving, both compared to chemical and sensory methods. The benefit will be the possibility to analyse more samples in order to cover variabililty within and between cheese batches.

Sensory characteristics at a greater age were not very well forecasted by early measurements on cheese, whether sensory, chemical or spectroscopic. This would have been very useful for cheese producers who would like to predict quality development during maturation and the period of sale, but will need more research before it is applicable.

Reference List

Adamopoulos, K. G., Goula, A. M., & Petropakis, H. J. (2001). Quality Control During Processing of Feta Cheese. NIR Application. *Journal of Food Composition and Analysis, 14*, 431-440.

Andersen, C. M., & Mortensen, G. (2008). Fluorescence spectroscopy: a rapid tool for analyzing dairy products. *Journal of Agricultural and Food Chemistry, 56*, 720-729.

Beresford, T. P., Fitzsimons, N. A., Brennan, N. L., & Cogan, T. M. (2001). Recent advances in cheese microbiology. *International Dairy Journal, 11*, 259-274.

Blazquez, C., Downey, G., O'Callaghan, D. J., Howard, V., Delahunty, C., Sheehan, E., Everard, C., & O'Donnell, C. P. (2006). Modelling of sensory and instrumental texture parameters in processed cheese by near infrared reflectance spectroscopy. *Journal of Dairy Research, 73*, 58-69.

Cattaneo, T. M. P., Giardina, C., Sinelli, N., Riva, M., & Giangiacomo, R. (2005). Application of FT-NIR and FT-IR spectroscopy to study the shelf-life of Crescenza cheese. *International Dairy Journal, 15*, 693-700.

Chen, G., Kocaoglu-Vurma, N. A., Harper, W. J., & Rodriguez-Saona, L. E. (2009). Application of infrared microspectroscopy and multivariate analysis for monitoring the effect of adjunct cultures during Swiss cheese ripening. *Journal of Dairy Science, 92*, 3575-3584.

Chen, M., Irudayaraj, J., & McMahon, D. J. (1998). Examination of full fat and reduced fat Cheddar cheese during ripening by Fourier transform infrared spectroscopy. *Journal of Dairy Science, 81*, 2791-2797.

Collins, Y. F., McSweeney, P. L. H., & Wilkinson, M. G. (2003). Lipolysis and free fatty acid catabolism in cheese: a review of current knowledge. *International Dairy Journal, 13*, 841-866.

Delahunty, C. M., & Drake, M. A. (2004). Sensory Character of Cheese and its Evaluation. In P.F. Fox, P. L. H. McSweeney, T. M. Cogan, & T. P. Guinee (Eds.), *Cheese. Chemistry, Physics and Microbiology* (Third ed., pp. 455-487). London: Elsevier Academic Press.

Downey, G., Sheehan, E., Delahunty, C., Callaghan, D., Guinee, T., & Howard, V. (2005). Prediction of maturity and sensory attributes of Cheddar cheese using near-infrared spectroscopy. *International Dairy Journal, 15*, 701-709.

Dufour, E., Devaux, M. F., Fortier, P., & Herbert, S. (2001). Delineation of the structure of soft cheeses at the molecular level by fluorescence spectroscopy-relationship with.. *International Dairy Journal, 11*, 465-473.

Dufour, E., Mazerolles, G., Devaux, M. F., Duboz, G., Duployer, M. H., & Mouhous Riou, N. (2000). Phase transition of triglycerides during semi-hard cheese ripening. *International Dairy Journal, 10*, 81-93.

Fagan, C. C., Everard, C., O'Donnell, C. P., Downey, G., Sheehan, E. M., Delahunty, C. M., Callaghan, D. J., & Howard, V. (2007). Prediction of processed cheese instrumental texture and meltability by mid-infrared spectroscopy coupled with chemometric tools. *Journal of Food Engineering, 80*, 1068-1077.

Fagan, C. C., Everard, C., O'Donnell, C. P., Downey, G., Sheehan, E. M., Delahunty, C. M., & O'Callaghan, D. J. (2007). Evaluating Mid-infrared Spectroscopy as a New Technique for Predicting Sensory Texture Attributes of Processed Cheese. *Journal of Dairy Science, 90*, 1122-1132.

Fagan, C. C., O'Donnell, C. P., O'Callaghan, D. J., Downey, G., Sheehan, E. M., Delahunty, C. M., Everard, C., Guinee, T. P., & Howard, V. (2007). Application of mid-infrared spectroscopy to the prediction of maturity and sensory texture attributes of cheddar cheese. *Journal of food science, 72*, 130-137.

Fox, P. F., & Cogan, T. M. (2004). Factors that Affect the Quality of Cheese. In P.F. Fox, P. L. H. McSweeney, T. M. Cogan, & T. P. Guinee (Eds.), *Cheese: Chemistry, Physics and Microbiology* (3 ed., pp. 583). London, UK: Elsevier Academic Press.

Garimella Purna, S. K., Prow, L. A., & Metzger, L. E. (2005). Utilization of Front-Face Fluorescence Spectroscopy for Analysis of Process Cheese Functionality. *Journal of Dairy Science, 88*, 470-477.

Gilles, J., & Lawrence, R. C. (1973). The assessment of Cheddar cheese quality by compositional analysis. *New Zealand Journal of Dairy Science and Technology, 8*, 148-151.

González-Martín, I., González-Pérez, C., Hernández-Hierro, J. M., & González-Cabrera, J. M. (2008). Use of NIRS technology with a remote reflectance fibre-optic probe for predicting major components in cheese. *Talanta, 75*, 351-355.

International Dairy Federation. (1995). IDF 50C Milk and milk products - Guidance on sampling. International Dairy Federation. Brussels, Belgium, International Dairy Federation.

International Dairy Federation. (1997). IDF 99C Sensory evaluation of dairy products by scoring. International Dairy Federation. Brussels, Belgium, International Dairy Federation.

International Dairy Federation. (1986). IDF 5B Cheese and processed cheese products. Determination of fat content - gravimetric method (reference method). International Dairy Federation. Brussels, Belgium, International Dairy Federation.

International Dairy Federation. (1988). IDF 88A Cheese and processed cheese products. Determination of chloride content - potentiometric titration method. International Dairy Federation. Brussels, Belgium, International Dairy Federation.

International Dairy Federation. (1993). IDF 20B Milk - Determination of nitrogen content. International Dairy Federation. Brussels, Belgium, International Dairy Federation.

International Dairy Federation. (1982). IDF 4A Cheese - Determination of dry matter. International Dairy Federation. Brussels, Belgium, International Dairy Federation.

International Organization for Standardization. (1988). ISO 8589 Sensory analysis - General guidance for the design
of test rooms. Geneva, Switzerland

International Organization for Standardization. (2003). ISO 4121 Sensory analysis - Guidelines for the use of quantitative response scales. Geneva, Switzerland

International Organization for Standardization. (1992). ISO 5492 Sensory analysis - Vocabulary. Geneva, Switzerland

Johnson, M., & Law, B. (1999). The origins, development and basic operations of cheesemaking technology. In B. Law (Ed.), *Technology of cheesemaking* (pp. 1-32). Sheffield, England: Sheffield Academic Press Ltd.

Karoui, R., & De Baerdemaeker, J. (2007). A review of the analytical methods coupled with chemometric tools for the determination of the quality and identity of dairy products. *Food Chemistry, 102*, 621-640.

Karoui, R., & Dufour, E. (2003). Dynamic testing rheology and fluorescence spectroscopy investigations of surface to centre differences in ripened soft.. *International Dairy Journal, 13*, 973-985.

Karoui, R., & Dufour, E. (2006). Prediction of the rheology parameters of ripened semi-hard cheeses using fluorescence spectra in the UV and visible.. *International Dairy Journal, 16*, 1490-1497.

Karoui, R., Dufour, E., & De Baerdemaeker, J. (2007). Monitoring the molecular changes by front face fluorescence spectroscopy throughout ripening of a semi-hard cheese. *Food Chemistry, 104*, 409-420.

Karoui, R., Dufour, E., Pillonel, L., Schaller, E., Picque, D., Cattenoz, T., & Bosset, J. O. (2005). The potential of combined infrared and fluorescence

spectroscopies as a method of determination of the geographic origin.. *International Dairy Journal, 15*, 287-298.

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

Karoui, R., Mazerolles, G., Bosset, J. O., De Baerdemaeker, J., & Dufour, E. (2007). Utilisation of mid-infrared spectroscopy for determination of the geographic origin of Gruyere PDO and L'Etivaz PDO.. *Food Chemistry, 105*, 847-854.

Karoui, R., Mouazen, A. M., Dufour, E., Pillonel, L., Schaller, E., De Baerdemaeker, J., & Bosset, J. O. (2006). Chemical characterisation of European Emmental cheeses by near infrared spectroscopy using chemometric tools. *International Dairy Journal, 16*, 1211-1217.

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

Koca, N., Rodriguez-Saona, L. E., Harper, W. J., & Alvarez, V. B. (2007). Application of Fourier Transform Infrared Spectroscopy for Monitoring Short-Chain Free Fatty Acids in Swiss Cheese. *Journal of Dairy Science, 90*, 3596-3603.

Kocaoglu-Vurma, N. A., Eliardi, A., Drake, M. A., Rodriguez-Saona, L. E., & Harper, W. J. (2009). Rapid profiling of Swiss cheese by attenuated total reflectance (ATR) infrared spectroscopy and descriptive sensory analysis. *Journal of food science, 74*, 232-239.

Kraggerud, H., Skeie, S., Høy, M., Røkke, L., & Abrahamsen, R. K. (2008). Season and ripening temperature influence fatty acid composition and sensory properties of semi-hard cheese during cheese maturation. *International Dairy Journal, 18*, 801-810.

Lebecque, A., Laguet, A., Deveaux, M. T., & Dufour, É. (2001). Delineation of the texture of Salers cheese by sensory analysis and physical methods. *Lait, 81*, 609-624.

Lerma-García, M. J., Gori, A., Cerretani, L., Simó-Alfonso, E. F., & Caboni, M. F. (2010). Classification of Pecorino cheeses produced in Italy according to their ripening time and manufacturing technique using Fourier transform infrared spectroscopy. *Journal of Dairy Science, 93*, 4490-4496.

Martens, H., & Martens, M. (2000). Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Quality and Preference 11*, 5-16.

Martens, H., & Næs, T. (1989). *Multivariate Calibration*. Chichester, UK: Wiley.

Martín-del-Campo, S. T., Picque, D., Cosío-Ramírez, R., & Corrieu, G. (2007). Evaluation of Chemical Parameters in Soft Mold-Ripened Cheese During Ripening by Mid-Infrared Spectroscopy. *Journal of Dairy Science, 90*, 3018-3027.

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

McSweeney, P. L., & Sousa, M. J. (2000). Biochemical pathways for the production of flavour compounds in cheeses during ripening: A review. *Lait, 80*, 293-324.

McSweeney, P. L. H. (2004). Biochemistry of cheese ripening: Introduction and overview. In P.F. Fox (Ed.), *Cheese: Chemistry, Physics and Microbiology* (Volume 1 ed., pp. 347-360). London, UK: Academic Press.

Mogensen, M. T. S. (1948). Bestemmelse af ostens proteolytiske spaltningsgrad med særligt henblik på formoltitreringen. Malmö, Sweden, Statens Mejeriförsök.

Rodriguez-Saona, L. E., Koca, N., Harper, W. J., & Alvarez, V. B. (2006). Rapid determination of Swiss cheese composition by Fourier transform infrared/attenuated total reflectance spectroscopy. *Journal of Dairy Science, 89*, 1407-1412.

Skeie, S., Lindberg, C., & Narvhus, J. (2001). Development of amino acids and organic acids in Norvegia, influence of milk treatment and adjunct Lactobacillus. *International Dairy Journal, 11*, 399-411.

Skeie, S., Feten, G., Almøy, T., Østlie, H., & Isaksson, T. (2006). The use of near infrared spectroscopy to predict selected free amino acids during cheese ripening. *International Dairy Journal, 16*, 236-242.

Sorensen, L. K., & Jepsen, R. (1998). Assessment of Sensory Properties of Cheese by Near-infrared Spectroscopy. *International Dairy Journal, 8*, 863-871.

Subramanian, A., Harper, W. J., & Rodriguez-Saona, L. E. (2009a). Rapid Prediction of Composition and Flavor Quality of Cheddar Cheese Using ATR-FTIR Spectroscopy. *Journal of food science, 74*, C292-C297.

Subramanian, A., Harper, W. J., & Rodriguez-Saona, L. E. (2009b). Cheddar cheese classification based on flavor quality using a novel extraction method and Fourier transform infrared spectroscopy. *Journal of Dairy Science, 92*, 87-94.

Subramanian, A., Alvarez, V. B., Harper, W. J., & Rodriguez-Saona, L. E. (2011). Monitoring amino acids, organic acids, and ripening changes in Cheddar cheese using Fourier-transform infrared spectroscopy. *International Dairy Journal, 21*, 434-440.

Subramanian, A., & Rodriguez-Saona, L. (2010). Chemical and Instrumental Approaches to Cheese Analysis. In S. Taylor (Ed.), *Advances in Food and Nutrition Research* (pp. 167-213). Advances in Food and Nutrition Research. London, UK: Elsevier.

Williams, P. C. (1987). Implementation of near-infrared technology. In P.C. Williams & K. Norris (Eds.), *Near-infrared technology in the agricultural and food industries* (pp. 143-167). St. Paul, MN: American Association of Cereal Chemists International.

Williams, P. C., & Sobering, D. C. (1993). Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *Journal of Near Infrared Spectroscopy, 1*, 25-32.

Wittrup, C., & Nørgaard, L. (1998). Rapid Near Infrared Spectroscopic Screening of Chemical Parameters in Semi-hard Cheese Using Chemometrics. *Journal of Dairy Science, 81*, 1803-1809.

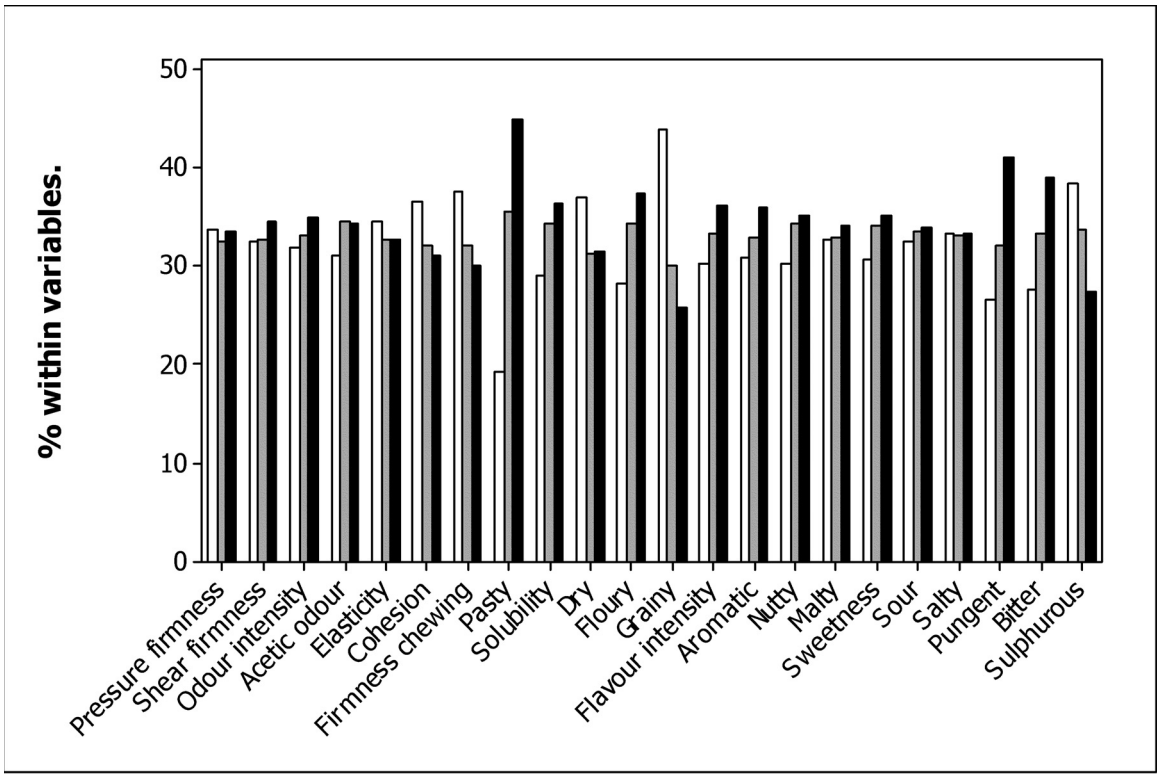

Figure 1:
Average values for sensory attributes at each stage of maturity: white 8 weeks, grey 24 weeks, and black 40 weeks. Values are adjusted to % within each variable in order to fit into the same scale.

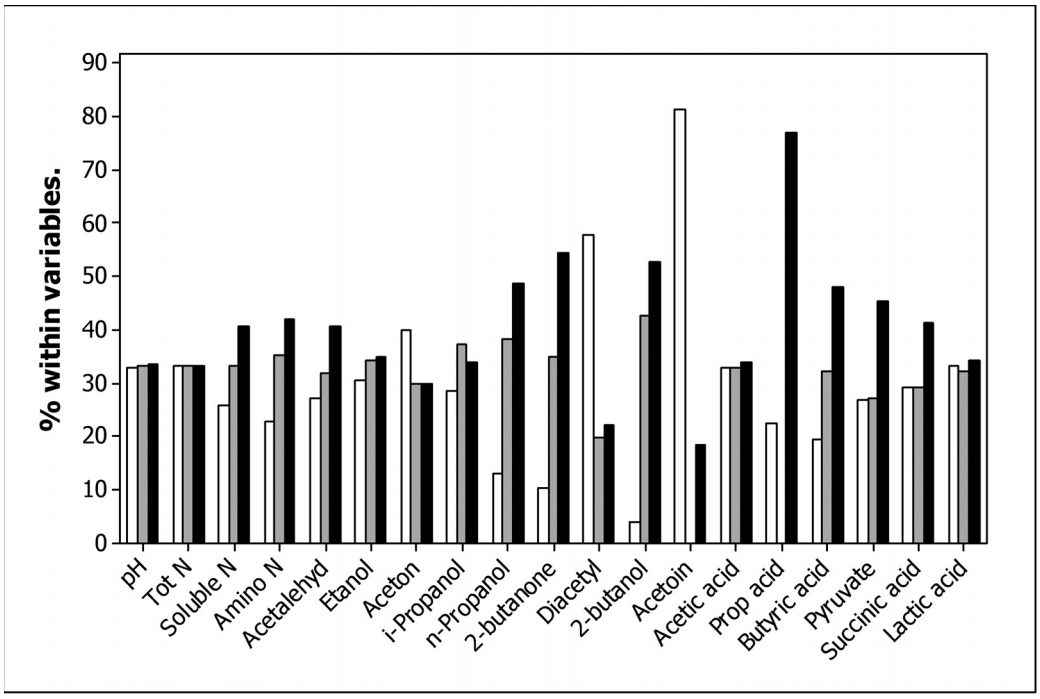

Figure 2:
Average values for chemical analysis results at each stage of maturity: white 8 weeks, grey 24 weeks, and black 40 weeks. Values are adjusted to % within each variable in order to fit into the same scale.

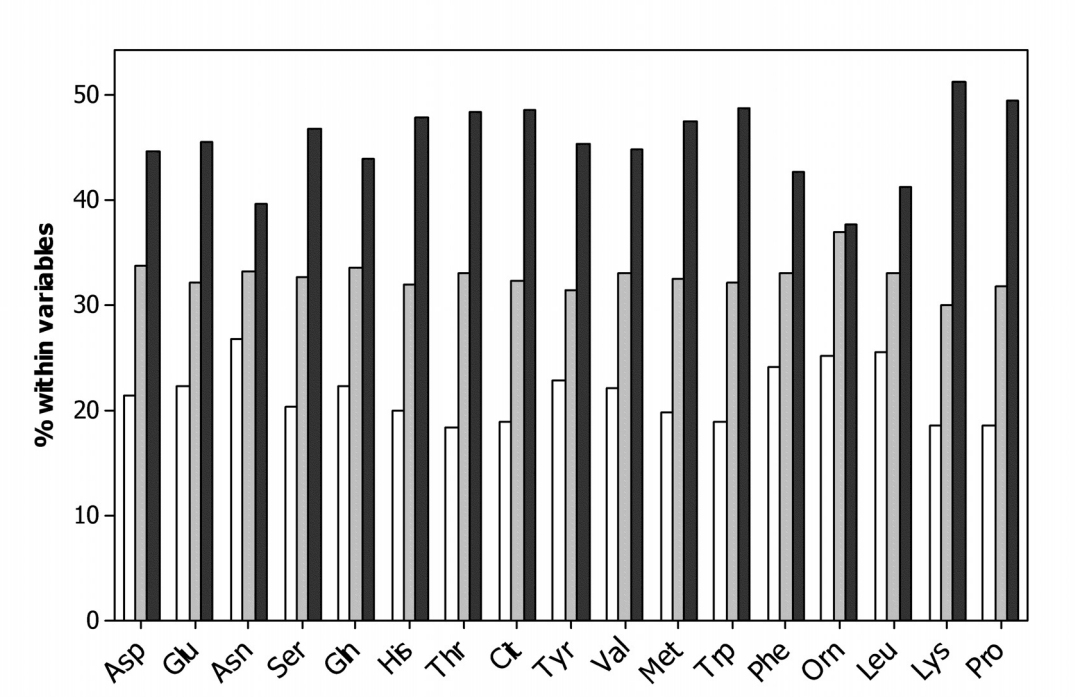

Figure 3: Average values for amino acids measured by HPLC at each stage of maturity: white 8 weeks, grey 24 weeks, and black 40 weeks. Values are adjusted to % within each variable in order to fit into the same scale.

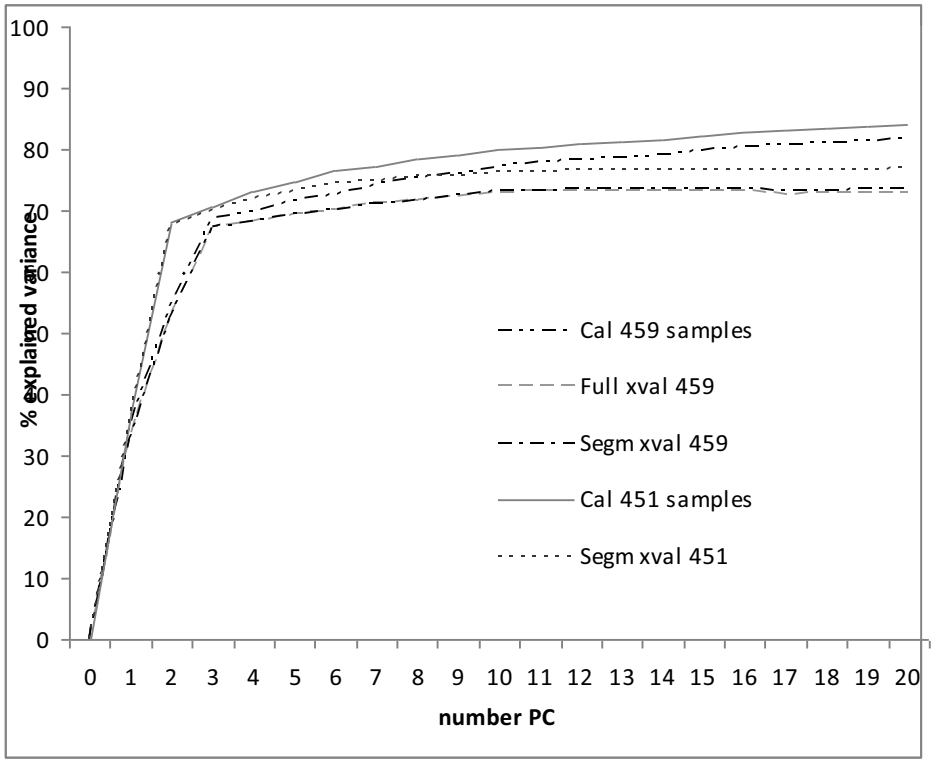

Figure 4:PLS1 regression validation: Explained variance in Y for 20 PLS1 factors with variables FTIR as X, amino nitrogen (AN) as Y. Line plot show calibration variance for all samples (Cal 459 samples) and sample set with 8 outlier samples omitted (Cal 451 samples). Validation variance for Full cross-validation (Full xval 459) and segmented cross-validation with 20 random segments (Segm xval 459) as well as for the outlier reduced data set (Segm xval 451).

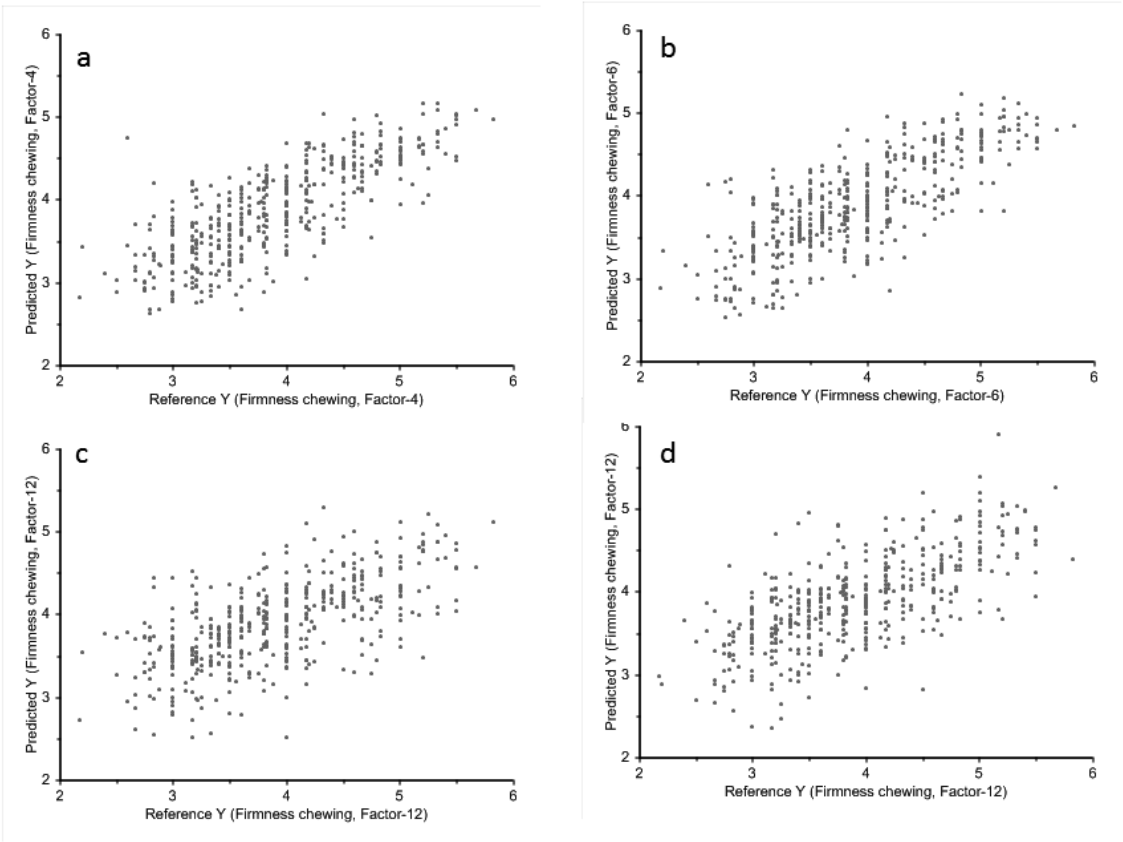

Figure 5: PLS1 Regression plot with measured value as x, predicted value as y for the dependent variable sensory descriptive attribute Firmness chewing. Independent variable sets were: a) Chemical variables. R=0.78 b) FTIR spectra. R=0.77 c) NIR spectra R=0.64 d) Fluorescence spectra. R=0.67

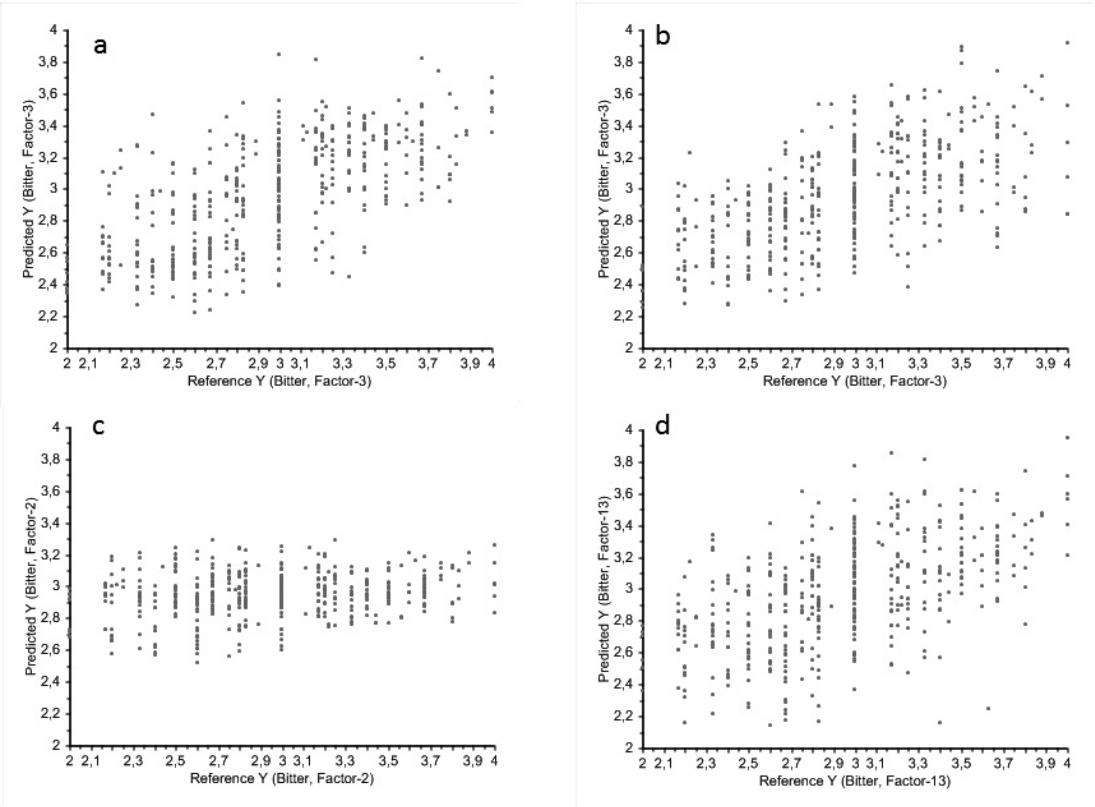

Figure 6: PLS1 Regression plot with measured value as x, predicted value as y for the dependent variable sensory descriptive attribute bitter. Independent variable sets were: a) Chemical variables. R=0.64 b) FTIR spectra. R=0.61 c) NIR spectra R=0.22 d) Fluorescence spectra. R=0.58

7

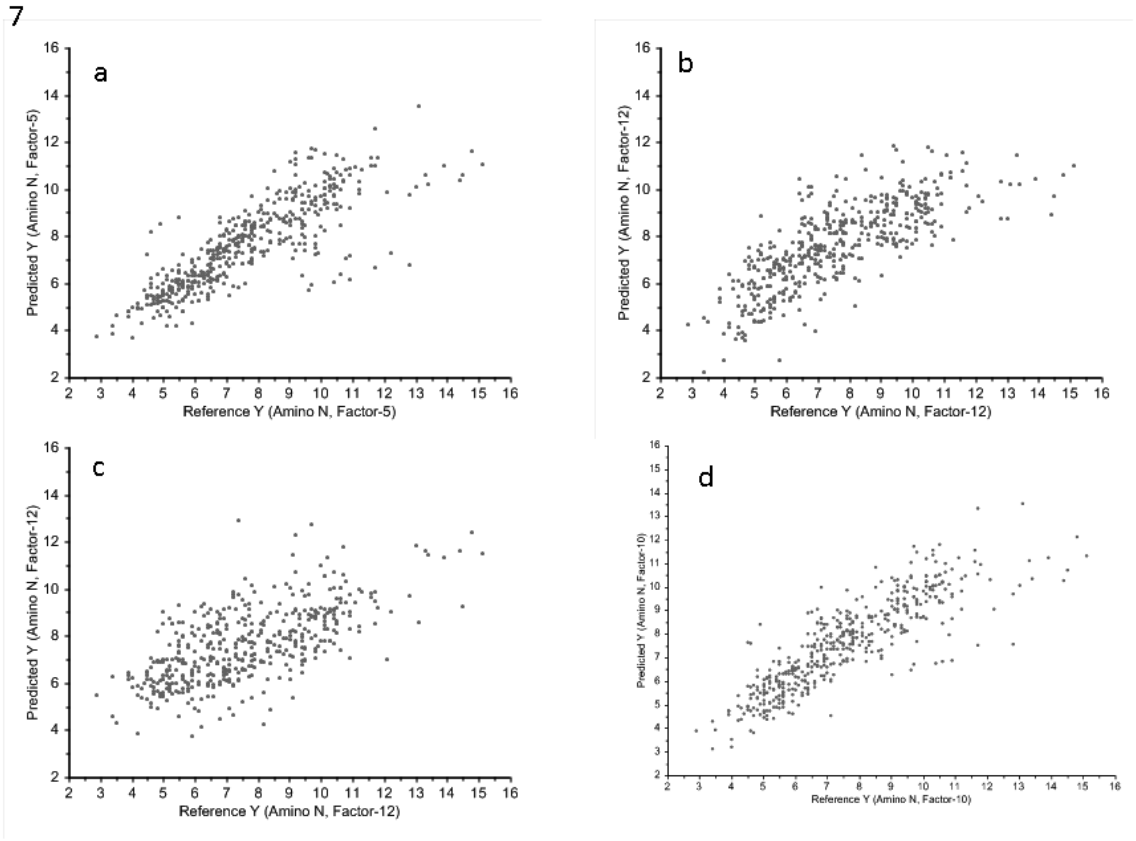

Figure 7: PLS1 Regression plot with measured value as x, predicted value as y for the dependent variable Amino Nitrogen. Independent variable sets were: a) FTIR spectra. R=0.83 c) NIR spectra R=0.65 d) Fluorescence spectra. R=0.77

8

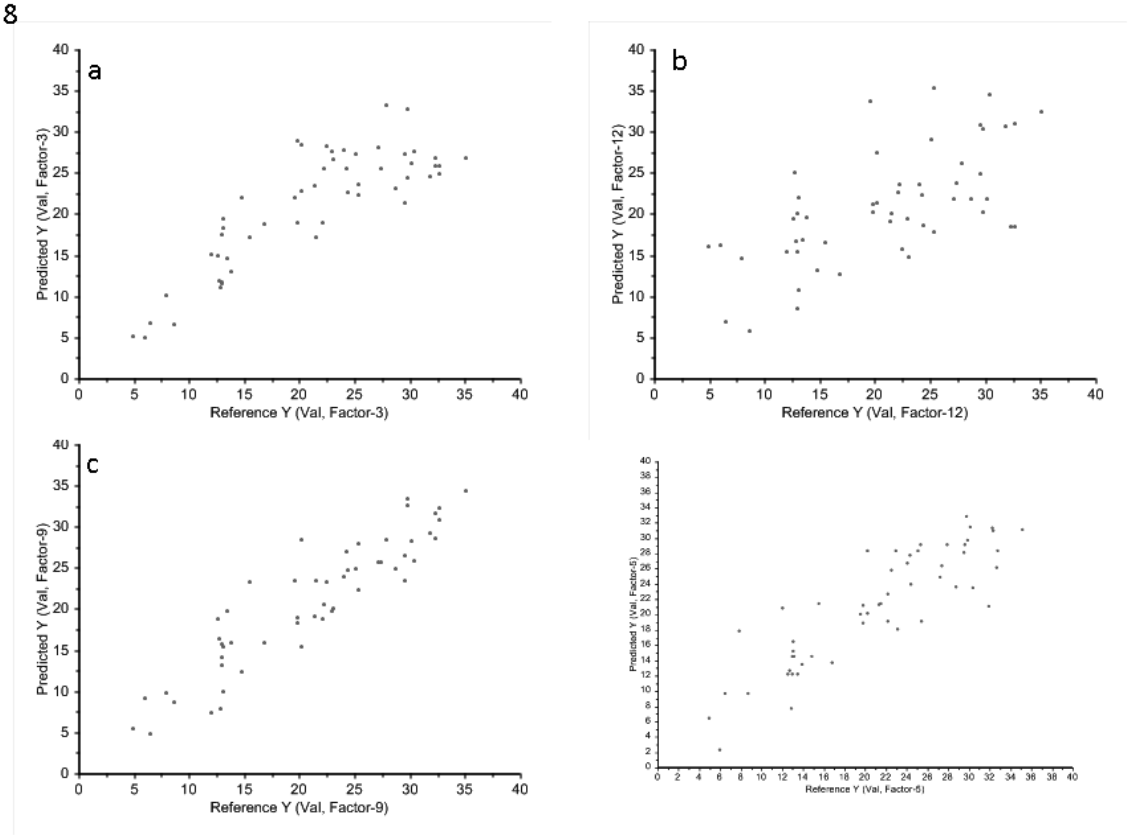

Figure 8:PLS1 Regression plot with measured value as x, predicted value as y for the dependent variable amino acid Val. Independent variable sets were: a) FTIR spectra. R=0.84 c) NIR spectra R=0.66 d) Fluorescence spectra. R=0.92

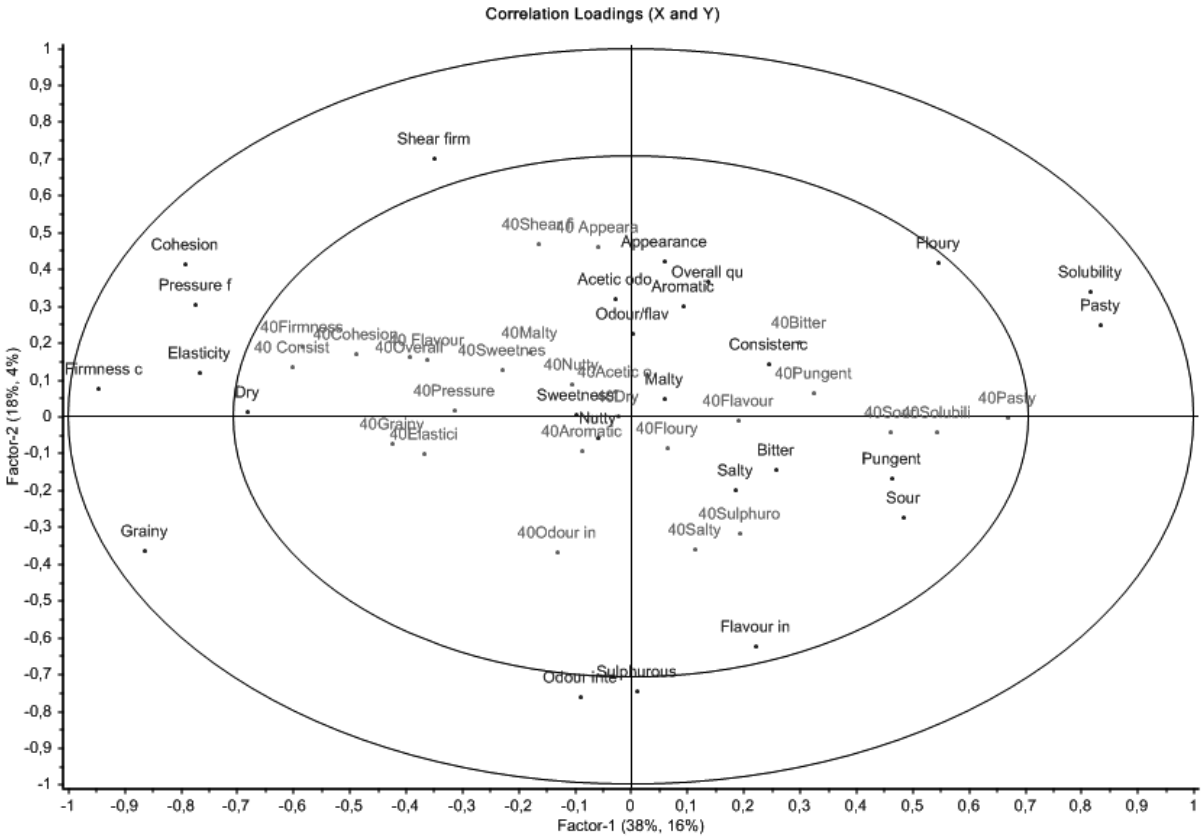

Figure 9:
PLS2 regression correlation loadings, with variables from sensory assessment at 8 weeks as X, corresponding variables at 40 weeks as Y. 153 cheese samples. Explained variance in Y: 20% after 2 factors.

Table 1: Overview of sampling and analysis throughout the experimental period. Number of samples given for each sampling point throughout the experimental period.

Weeks from start	**0**	**4**	**8**	**12**	**16**	**20**	**24**	**28**	**32**	**36**	**40**	**44**	**48**	**52**	**56**	**60**	**64**	**68**	**72**	**76**	**80**	**84**	Sum
Cheese production	12	16	12	13	12	14	12	12	12	12	12	14											153
Sampling 8 weeks			12	16	12	13	12	14	12	12	12	12	12	14									153
Sampling 24 weeks							12	16	12	13	12	14	12	12	12	12	12	14					153
Sampling 40 weeks											12	16	12	13	12	14	12	12	12	12	12	14	153
Samples analysed	0	0	12	16	12	13	24	30	24	25	36	42	36	39	24	26	24	26	12	12	12	14	459

Table 2: Basic statistics for all measured variables. A) Sensory variables B) Chemical variables p from ANOVA with 2 factors : Sample (n=153) and Age (n=3)

A. Variable	Unit	N	Mean	StDev	Minimum	Maximum	p Sample	p Age
Quality scoring								
Quality – Overall	1-5 scale	458	3.02	0.40	1.00	4.00	0.093	0.078
Quality - Appearance	1-5 scale	458	3.37	0.45	2.00	4.25	0.005	0.754
Quality - consistency	1-5 scale	458	3.24	0.37	2.33	4.00	0.521	0.146
Quality – Flavour	1-5 scale	458	3.17	0.42	1.00	4.00	0.019	0.006
Sensory descriptive								
Pressure firmness	1-9 scale	459	5.51	0.59	3.80	7.50	<0.001	<0.001
Shear firmness	1-9 scale	459	5.51	0.41	4.50	6.80	<0.001	<0.001
Odour intensity	1-9 scale	459	5.20	0.39	4.17	6.33	<0.001	<0.001
Acetic odour	1-9 scale	459	1.04	0.12	1.00	2.00	0.408	0.062
Elasticity	1-9 scale	459	6.15	0.51	4.80	7.67	<0.001	<0.001
Cohesion	1-9 scale	459	4.37	0.67	2.83	6.17	<0.001	<0.001
Firmness chewing	1-9 scale	459	3.89	0.72	2.17	5.83	<0.001	<0.001
Pasty	1-9 scale	459	2.92	1.31	1.00	5.80	<0.001	<0.001
Solubility	1-9 scale	459	4.75	0.69	3.00	6.67	<0.001	<0.001
Dry	1-9 scale	459	1.30	0.36	1.00	2.80	0.019	<0.001
Floury	1-9 scale	459	2.90	0.48	1.33	4.20	<0.001	<0.001
Grainy	1-9 scale	459	1.75	0.74	1.00	4.33	<0.001	<0.001
Flavour intensity	1-9 scale	459	5.69	0.50	4.00	7.00	<0.001	<0.001
Aromatic	1-9 scale	459	3.79	0.53	2.60	5.20	0.009	<0.001
Nutty	1-9 scale	459	1.02	0.09	1.00	1.60	<0.001	<0.001
Malty	1-9 scale	459	1.18	0.26	1.00	3.25	<0.001	0.007
Sweetness	1-9 scale	459	1.28	0.27	1.00	2.71	0.001	0.009
Sour	1-9 scale	459	5.89	0.43	4.75	7.50	<0.001	<0.001
Salty	1-9 scale	459	5.10	0.22	4.60	6.33	<0.001	<0.001
Pungent	1-9 scale	459	2.06	0.55	1.00	4.00	<0.001	<0.001
Bitter	1-9 scale	459	2.97	0.55	1.50	5.40	<0.001	<0.001
Sulphurous	1-9 scale	459	2.08	0.72	1.00	3.80	<0.001	<0.001

Table 3. Cross correlation coefficient (Pearson) for chemical analytes: Soluble N (SN), amino N (AN) and amino acids in cheese. 53 samples 8, 24, and 40 weeks age.

	SN	AN	Asp	Glu	Asn	Ser	Gln	Thr	Cit	Tyr	Val	Met	Phe	Orn	Leu	Lys	Pro
SN	1.00																
AN	0.34	1.00															
Asp	0.83	0.89	1.00														
Glu	0.87	0.88	0.96	1.00													
Asn	0.92	0.94	0.93	0.94	1.00												
Ser	0.89	0.93	0.94	0.93	0.95	1.00											
Gln	0.78	0.93	0.88	0.86	0.92	0.92	1.00										
Thr	0.89	0.96	0.92	0.93	0.97	0.97	0.95	1.00									
Cit	0.73	0.73	0.66	0.72	0.73	0.74	0.69	0.79	1.00								
Tyr	0.85	0.86	0.96	0.94	0.89	0.92	0.84	0.88	0.64	1.00							
Val	0.88	0.97	0.94	0.93	0.97	0.97	0.96	0.99	0.74	0.91	1.00						
Met	0.92	0.95	0.94	0.94	0.98	0.97	0.93	0.98	0.77	0.91	0.99	1.00					
Phe	0.90	0.96	0.93	0.93	0.98	0.97	0.96	0.98	0.74	0.91	1.00	0.99	1.00				
Orn	0.73	0.79	0.80	0.76	0.85	0.81	0.77	0.78	0.46	0.75	0.81	0.83	0.81	1.00			
Leu	0.89	0.96	0.92	0.91	0.97	0.95	0.97	0.97	0.72	0.90	0.99	0.97	0.99	0.78	1.00		
Lys	0.84	0.92	0.89	0.87	0.92	0.94	0.93	0.96	0.74	0.85	0.95	0.95	0.95	0.78	0.94	1.00	
Pro	0.83	0.96	0.90	0.88	0.93	0.95	0.96	0.96	0.71	0.87	0.98	0.96	0.97	0.80	0.96	0.95	1.00

Table 4: Validation measures for sensory variables modelled from chemical and spectroscopy data. SEMean: Standard error of the mean (between assessors), X values in the calibrations are Chemical variables, FTIR, NIR, Fluorescence and all spectral data. RMSEP: Root mean square error of Prediction - validated. Explained variance is measured in % of total Y-variance. Correlation coefficient R. RPD Ratio Performance Deviation = StDev(ref)/RMSEP, and RER, Range error ratio=(Max value-Min value)/RMSEP. Opt n PLSn: Optimal number of factors in PLS1/PLS2 modelling.

	SE	RMSEP					Correlation coefficient R					RPD					RER				
	Mean	Chem	FTIR	NIR	Fluor	Spectr	Chem	FTIR	NIR	Fluor	Spectr	Chem	FTIR	NIR	Fluor	Spectr	Chem	FTIR	NIR	Fluor	Spectr
PLS2 Opt n PC		4	5	15	6	11															
Overall quality	0.15	0.36	0.36	0.37	0.37	0.35	0.41	0.44	0.36	0.39	0.45	1.10	1.11	1.07	1.08	1.12	8.30	8.39	8.07	8.14	8.46
Appearance	0.14	0.44	0.44	0.42	0.44	0.43	0.09	0.13	0.36	0.16	0.25	1.00	1.01	1.07	1.02	1.03	5.06	5.11	5.39	5.12	5.20
Consistency	0.17	0.30	0.29	0.32	0.32	0.29	0.58	0.61	0.47	0.52	0.61	1.23	1.26	1.14	1.15	1.26	5.63	5.74	5.18	5.27	5.75
Odour/flavour	0.18	0.37	0.36	0.38	0.37	0.36	0.48	0.51	0.44	0.45	0.50	1.14	1.16	1.11	1.11	1.15	8.21	8.35	7.99	8.01	8.32
Pressure firmness	0.23	0.53	0.51	0.57	0.58	0.50	0.43	0.49	0.37	0.32	0.52	1.11	1.14	1.03	1.02	1.17	6.97	7.20	6.52	6.40	7.35
Shear firmness	0.21	0.36	0.39	0.38	0.38	0.35	0.48	0.29	0.38	0.52	0.50	1.14	1.06	1.08	1.06	1.16	6.44	5.96	6.09	6.01	6.53
Odour intensity	0.26	0.36	0.37	0.36	0.32	0.31	0.31	0.27	0.20	0.53	0.60	1.06	1.04	1.09	1.19	1.26	5.96	5.86	6.10	6.68	7.05
Elasticity	0.23	0.44	0.46	0.47	0.47	0.41	0.53	0.49	0.42	0.45	0.61	1.17	1.13	1.09	1.09	1.25	6.51	6.30	6.06	6.10	6.97
Cohesion	0.25	0.45	0.48	0.53	0.57	0.45	0.74	0.69	0.60	0.55	0.73	1.49	1.37	1.25	1.18	1.46	7.45	6.88	6.25	5.89	7.34
Firmness chewing	0.22	0.45	0.50	0.54	0.60	0.47	0.79	0.72	0.67	0.58	0.76	1.61	1.43	1.33	1.20	1.52	8.24	7.28	6.80	6.12	7.76
Pasty	0.27	0.73	0.85	0.98	1.03	0.80	0.83	0.76	0.67	0.62	0.79	1.78	1.55	1.33	1.27	1.63	6.53	5.68	4.89	4.68	5.98
Solubility	0.25	0.43	0.51	0.53	0.58	0.47	0.78	0.68	0.65	0.56	0.74	1.60	1.36	1.31	1.21	1.49	8.45	7.19	6.91	6.37	7.88
Dry	0.22	0.30	0.30	0.32	0.33	0.30	0.57	0.55	0.49	0.40	0.56	1.22	1.19	1.14	1.09	1.21	6.06	5.91	5.67	5.43	6.00
Floury	0.30	0.42	0.44	0.44	0.45	0.42	0.49	0.37	0.42	0.35	0.49	1.15	1.09	1.10	1.07	1.15	6.89	6.52	6.57	6.39	6.86
Grainy	0.25	0.47	0.54	0.58	0.62	0.52	0.78	0.70	0.63	0.56	0.72	1.57	1.39	1.28	1.19	1.42	7.05	6.21	5.72	5.35	6.38
Flavour intensity	0.24	0.35	0.35	0.39	0.36	0.32	0.72	0.71	0.63	0.70	0.78	1.42	1.42	1.28	1.39	1.58	8.49	8.46	7.62	8.29	9.44
Aromatic	0.37	0.45	0.44	0.46	0.47	0.43	0.54	0.55	0.52	0.51	0.60	1.16	1.18	1.15	1.11	1.22	5.76	5.86	5.68	5.49	6.04
Malty	0.10	0.26	0.26	0.26	0.25	0.25	0.21	0.22	0.22	0.25	0.35	1.00	1.00	1.00	1.02	1.05	8.72	8.71	8.70	8.84	9.10
Sweetness	0.25	0.25	0.25	0.26	0.25	0.25	0.23	0.09	0.12	0.24	0.22	1.07	1.05	1.03	1.07	1.06	6.91	6.79	6.69	6.92	6.86
Acidic	0.22	0.35	0.36	0.38	0.39	0.35	0.56	0.53	0.48	0.45	0.58	1.21	1.18	1.14	1.10	1.23	7.77	7.59	7.32	7.10	7.91
Salty	0.08	0.21	0.21	0.21	0.21	0.21	0.18	0.24	0.28	0.24	0.29	1.02	1.03	1.04	1.02	1.04	8.08	8.18	8.25	8.08	8.28
Pungent	0.35	0.42	0.43	0.46	0.47	0.42	0.63	0.61	0.53	0.50	0.64	1.32	1.28	1.21	1.19	1.33	7.12	6.93	6.52	6.42	7.21
Bitter	0.31	0.42	0.42	0.48	0.49	0.43	0.62	0.62	0.44	0.40	0.60	1.31	1.30	1.13	1.12	1.28	9.31	9.26	8.09	7.94	9.10
Sulphurous	0.30	0.61	0.64	0.63	0.64	0.57	0.52	0.47	0.50	0.54	0.61	1.17	1.13	1.15	1.13	1.26	4.55	4.41	4.47	4.39	4.90
PLS1 Opt n PC		4	6	13	13	9															
Firmness chewing	0.22	0.45	0.46	0.54	0.47	0.49	0.78	0.77	0.67	0.76	0.74	1.60	1.56	1.15	1.53	1.47	8.15	7.97	5.85	7.80	7.48
PLS1 Opt n PC		3	3	14	16	8															
Bitter	0.31	0.41	0.43	0.49	0.36	0.43	0.64	0.61	0.42	0.61	0.60	1.34	1.27	1.12	1.52	1.27	9.51	9.07	7.96	10.83	9.07

Table 5: Validation measures for sensory variables at 40 weeks age modelled from chemical and spectroscopy data. SEMean: Standard error of the mean (between assessors), X values in the calibrations are Chemical variables, FTIR, NIR, Fluorescence,all spectral data, and 8 weeks sensory. RMSEP: Root mean square error of Prediction - validated. Explained variance is measured in % of total Y-variance. Correlation coefficient R. RPD Ratio Performance Deviation = StDev(ref)/RMSEP. Opt n PLSn: Optimal number of factors in PLS1/PLS2 modelling.

			RMSEP						Correlation coefficient R						RPD					
	StDev	Range	Chem	FTIR	NIR	Fluor	Spectr	8w sens	Chem	FTIR	NIR	Fluor	Spectr	8w sens	Chem	FTIR	NIR	Fluor	Spectr	8w sens
Opt n PC PLS2			6	4	10	6	7	7												
Overall quality	0,32	2,17	0,29	0,29	0,30	0,28	0,28	0,28	0,38	0,38	0,35	0,45	0,46	0,43	1,08	1,07	1,05	1,12	1,13	1,12
Appearance	0,44	2,00	0,41	0,44	0,42	0,38	0,39	0,38	0,34	0,14	0,30	0,47	0,45	0,43	1,06	1,00	1,03	1,14	1,12	1,15
Consistency	0,24	1,33	0,21	0,22	0,21	0,21	0,21	0,19	0,54	0,48	0,54	0,50	0,52	0,61	1,19	1,13	1,18	1,15	1,17	1,28
Odour/flavour	0,33	2,33	0,31	0,31	0,31	0,28	0,28	0,31	0,41	0,41	0,40	0,53	0,54	0,43	1,09	1,09	1,08	1,18	1,19	1,09
Pressure firmness	0,57	2,75	0,47	0,49	0,57	0,53	0,47	0,50	0,56	0,51	0,24	0,37	0,55	0,44	1,20	1,17	1,01	1,08	1,20	1,14
Shear firmness	0,37	1,80	0,29	0,35	0,33	0,32	0,32	0,32	0,60	0,30	0,45	0,49	0,49	0,49	1,25	1,05	1,11	1,14	1,14	1,16
Odour intensity	0,32	1,83	0,29	0,33	0,32	0,30	0,30	0,30	0,39	0,01	0,16	0,32	0,31	0,35	1,10	0,99	1,00	1,07	1,06	1,08
Elasticity	0,42	2,25	0,39	0,40	0,41	0,41	0,38	0,39	0,45	0,36	0,37	0,35	0,51	0,38	1,06	1,03	1,02	1,02	1,11	1,07
Cohesion	0,46	2,50	0,40	0,39	0,43	0,44	0,39	0,38	0,51	0,54	0,38	0,28	0,53	0,53	1,15	1,18	1,07	1,04	1,18	1,21
Firmness chewing	0,47	2,83	0,39	0,36	0,41	0,45	0,37	0,35	0,57	0,67	0,52	0,35	0,63	0,66	1,20	1,33	1,15	1,05	1,27	1,36
Pasty	0,72	4,00	0,62	0,65	0,70	0,67	0,60	0,56	0,51	0,44	0,31	0,37	0,55	0,64	1,16	1,11	1,04	1,07	1,20	1,29
Solubility	0,33	2,67	0,30	0,32	0,32	0,31	0,30	0,30	0,46	0,35	0,37	0,38	0,44	0,51	1,12	1,05	1,05	1,07	1,11	1,13
Dry	0,17	1,00	0,18	0,18	0,17	0,18	0,17	0,18	0,08	-0,05	0,27	0,01	0,18	-0,14	0,98	0,96	1,02	0,98	1,00	0,95
Floury	0,44	2,34	0,35	0,38	0,35	0,34	0,31	0,39	0,59	0,47	0,59	0,62	0,70	0,45	1,27	1,17	1,27	1,30	1,43	1,14
Grainy	0,26	1,43	0,24	0,24	0,24	0,25	0,23	0,22	0,21	0,21	0,26	-0,05	0,36	0,43	1,07	1,08	1,08	1,04	1,14	1,19
Flavour intensity	0,27	1,60	0,27	0,26	0,26	0,28	0,28	0,28	0,27	0,27	0,34	-0,01	0,16	0,02	0,99	1,00	1,01	0,94	0,96	0,97
Aromatic	0,51	2,20	0,50	0,53	0,53	0,51	0,52	0,51	0,28	0,07	0,13	0,25	0,16	0,16	1,00	0,95	0,96	1,00	0,96	0,99
Malty	0,29	2,25	0,30	0,30	0,30	0,29	0,29	0,30	0,16	0,26	0,22	0,26	0,31	0,25	0,96	0,98	0,97	0,99	1,00	0,99
Sweetness	0,33	1,71	0,30	0,32	0,31	0,30	0,28	0,31	0,31	0,14	0,22	0,32	0,45	0,19	1,12	1,06	1,08	1,12	1,20	1,08
Acidic	0,4	2,50	0,36	0,36	0,38	0,38	0,33	0,35	0,41	0,43	0,26	0,25	0,54	0,42	1,12	1,12	1,05	1,05	1,21	1,16
Salty	0,18	1,08	0,17	0,16	0,16	0,17	0,16	0,17	0,43	0,41	0,51	0,42	0,47	0,37	1,10	1,15	1,15	1,10	1,13	1,10
Pungent	0,55	3,00	0,45	0,48	0,49	0,50	0,50	0,51	0,53	0,44	0,41	0,35	0,35	0,26	1,23	1,15	1,13	1,11	1,10	1,08
Bitter	0,48	3,15	0,43	0,49	0,46	0,43	0,45	0,46	0,45	0,13	0,36	0,43	0,38	0,33	1,12	1,00	1,06	1,12	1,08	1,05
Sulphurous	0,47	2,00	0,49	0,47	0,49	0,48	0,48	0,46	0,14	0,27	0,15	0,13	0,21	0,29	0,97	1,01	0,97	0,98	0,99	1,04

Table 6: Basic statistics and validation measures for chemical variables modelled from spectroscopy data. StDev: Total standard deviation (between samples- reference method). X values in the calibrations are FTIR. NIR and Fluorescence. RMSEP: Root mean square error of Prediction - validated. Correlation coefficient R. RPD Ratio Performance Deviation=StDev/RMSEP. RER Range error ratio=(Max-Min)/RMSEP. Opt n PLSn: Optimal number of factors in PLS1/PLS2 modelling.

		Basic statistics			RMSEP				Correlation coeff. R				RPD				RER			
	Unit	N	Min-Max	StDev	FTIR	NIR	Fluor	Spectra	FTIR	NIR	Fluor	Spectra	FTIR	NIR	Fluor	Spectra	FTIR	NIR	Fluor	Spectra
Opt n PC PLS2					10	9	5	10												
pH		459	5.39 - 5.81	0,07	0,05	0,07	0,06	0,05	0,77	0,12	0,48	0,69	1,57	1,02	1,17	1,39	9,11	5,90	6,79	8,08
Tot N	mg g-1	459	39.3 - 45.2	1,05	0,69	1,00	1,01	0,85	0,72	0,14	0,08	0,58	1,51	1,04	1,04	1,23	8,49	5,88	5,84	6,94
Soluble N	% of Tot N	459	7.4 - 30.6	4,61	1,86	3,12	2,87	2,20	0,84	0,42	0,55	0,85	2,48	1,48	1,60	2,10	12,49	7,44	8,08	10,56
Amino N	% of Tot N	459	2.9 - 23.3	2,95	0,68	1,30	1,12	1,11	0,87	0,35	0,57	0,86	4,34	2,28	2,62	2,66	30,03	15,75	18,15	18,37
Acetalehyd	rel unit	441	0.7 - 5.1	0,52	0,37	0,42	0,38	0,39	0,53	0,28	0,51	0,64	1,39	1,23	1,38	1,33	11,73	10,41	11,68	11,25
Etanol	rel unit	441	69.5 - 1289.6	232	206	204	214	183	0,51	0,51	0,42	0,59	1,13	1,14	1,08	1,27	5,93	5,99	5,70	6,68
Aceton	rel unit	440	0.4 - 3.4	0,48	0,52	0,45	0,53	0,38	0,52	0,68	0,50	0,57	0,91	1,07	0,91	1,25	5,73	6,71	5,71	7,84
i-Propanol	rel unit	440	0.2 - 7.2	1,19	0,88	0,96	0,87	0,99	0,50	0,33	0,51	0,53	1,35	1,24	1,37	1,20	7,91	7,29	8,05	7,05
n-Propanol	rel unit	440	0 - 3.3	0,61	0,27	0,26	0,28	0,44	0,58	0,60	0,50	0,68	2,28	2,35	2,17	1,38	12,15	12,49	11,57	7,37
2-butanone	rel unit	440	0 - 10.3	1,44	0,92	0,87	0,90	1,23	0,29	0,38	0,24	0,50	1,57	1,66	1,60	1,18	11,21	11,84	11,39	8,39
Diacetyl	rel unit	440	0.1 - 4.6	0,70	0,73	0,74	0,72	0,55	0,50	0,45	0,50	0,61	0,97	0,95	0,98	1,29	6,17	6,06	6,23	8,25
2-butanol	rel unit	440	0 - 32.5	4,78	2,52	2,60	2,59	4,11	0,38	0,26	0,24	0,51	1,89	1,84	1,85	1,16	12,89	12,51	12,56	7,91
Acetoin	mmol kg-1	326	0 - 0.5	0,12	0,11	0,10	0,11	0,08	0,61	0,67	0,54	0,65	1,11	1,19	1,05	1,55	4,72	5,09	4,48	6,61
Acetic acid	mmol kg-1	326	7.2 - 12.8	0,78	0,65	0,68	0,67	0,53	0,63	0,59	0,60	0,59	1,20	1,15	1,16	1,47	8,60	8,22	8,35	10,54
Butyric acid	mmol kg-1	326	0 - 1.2	0,13	0,09	0,10	0,10	0,08	0,51	0,35	0,47	0,68	1,42	1,31	1,40	1,61	12,81	11,80	12,59	14,45
Pyruvate	g kg-1	312	0 - 0.3	0,05	0,02	0,02	0,03	0,03	0,38	0,32	0,14	0,55	1,94	1,92	1,84	1,44	12,44	12,29	11,78	9,24
Succinic acid	g kg-1	312	0.1 - 1	0,17	0,14	0,14	0,13	0,11	0,53	0,48	0,62	0,68	1,26	1,22	1,38	1,64	6,51	6,34	7,15	8,51
Lactic acid	g kg-1	312	8.9 - 14.6	0,99	0,61	0,78	0,75	0,58	0,74	0,49	0,55	0,70	1,63	1,27	1,33	1,71	9,34	7,27	7,61	9,79
Opt n PC PLS1					6	16	16	6												
Amino N	% of Tot N	459	2.9 - 23.3	2,95	1,11	1,48	1,20	1,10	0,86	0,74	0,84	0,86	2,66	1,99	2,46	2,68	18,38	13,78	17,00	18,55
Opt n PC PLS1					13	10	10	18												
Lactic acid	g kg-1	312	8.9 - 14.6	0,99	0,47	0,71	0,69	0,46	0,89	0,71	0,72	0,88	2,11	1,40	1,44	2,16	12,13	8,03	8,26	12,39

Table 7: Basic statistics and validation measures for chemical variables modelled from spectroscopy data. StDev: Total standard deviation (between samples- reference method). X values in the calibrations are FTIR. NIR and Fluorescence. RMSEP: Root mean square error of Prediction - validated. Correlation coefficient R. RPD Ratio Performance Deviation=StDev/RMSEP. RER Range error ratio=(Max-Min)/RMSEP. Opt n PLSn: Optimal number of factors in PLS1/PLS2 modelling.

	Basic statistics				RMSEP				Correlation coefficient R				RPD				RER			
	Unit	N	Min-Max	StDev	FTIR	NIR	Fluor	Spectra	FTIR	NIR	Fluor	Spectra	FTIR	NIR	Fluor	Spectra	FTIR	NIR	Fluor	Spectra
PLS2 Opt n PC					3	10	6	9												
Asp	mmol kg-1	53	1.42 - 8.74	1.96	1.23	1.53	0.90	1.08	0.78	0.65	0.89	0.84	1.60	1.28	2.17	1.82	5.97	4.80	8.11	6.80
Glu	mmol kg-1	53	11.3 - 60.5	12.61	8.29	9.55	6.48	7.27	0.75	0.67	0.86	0.82	1.52	1.32	1.95	1.74	5.94	5.15	7.59	6.77
Asn	mmol kg-1	53	9.0 - 29.8	5.19	2.90	3.88	2.46	2.36	0.83	0.68	0.88	0.89	1.79	1.34	2.11	2.20	7.16	5.34	8.46	8.81
Ser	mmol kg-1	53	2.46 - 17.3	3.95	2.42	3.13	1.82	1.97	0.79	0.64	0.89	0.87	1.63	1.26	2.17	2.00	6.14	4.76	8.18	7.55
Gln	mmol kg-1	53	3.04 - 27.5	6.90	3.84	5.63	3.58	2.92	0.83	0.62	0.86	0.90	1.79	1.23	1.93	2.36	6.39	4.36	6.85	8.39
His +Gly	mmol kg-1	53	2.64 - 26.4	6.34	4.08	4.81	3.18	3.06	0.76	0.67	0.87	0.88	1.55	1.32	2.00	2.08	5.83	4.95	7.50	7.79
Thr	mmol kg-1	53	1.74 - 12.7	3.01	1.75	2.26	1.34	1.33	0.81	0.68	0.90	0.90	1.72	1.33	2.24	2.27	6.32	4.89	8.23	8.32
Cit	mmol kg-1	53	1.24 - 11.1	2.08	1.69	1.71	1.46	1.40	0.59	0.59	0.71	0.74	1.23	1.21	1.42	1.49	5.88	5.79	6.77	7.09
Tyr	mmol kg-1	53	1.5 - 12.3	2.63	1.74	2.15	1.59	1.63	0.75	0.61	0.81	0.79	1.51	1.22	1.65	1.61	6.25	5.05	6.82	6.67
Val	mmol kg-1	53	4.94 - 35.1	8.12	4.38	6.05	3.65	3.37	0.84	0.69	0.90	0.91	1.85	1.34	2.23	2.41	6.89	4.99	8.27	8.97
Met	mmol kg-1	53	1.3 - 10.3	2.59	1.51	1.87	1.14	1.09	0.81	0.71	0.90	0.91	1.72	1.39	2.28	2.37	6.00	4.84	7.93	8.26
Trp+Ile	mmol kg-1	53	1.54 - 21.8	5.28	3.24	3.86	2.38	2.58	0.79	0.70	0.89	0.87	1.63	1.37	2.22	2.05	6.26	5.25	8.52	7.86
Phe	mmol kg-1	53	5.4 - 23.0	4.65	2.50	3.51	2.23	2.00	0.84	0.68	0.88	0.90	1.86	1.32	2.08	2.32	7.07	5.04	7.93	8.82
Orn	mmol kg-1	52	4.86 - 32.5	6.92	5.32	5.40	4.46	4.22	0.63	0.64	0.76	0.79	1.30	1.28	1.55	1.64	5.20	5.12	6.20	6.55
Leu	mmol kg-1	53	13.9 - 58.8	11.98	6.14	9.28	6.07	5.13	0.86	0.66	0.87	0.90	1.95	1.29	1.97	2.33	7.32	4.84	7.41	8.75
Lys	mmol kg-1	53	0.04 - 37.6	10.44	6.46	8.16	5.75	5.32	0.78	0.64	0.84	0.86	1.62	1.28	1.81	1.96	5.81	4.60	6.52	7.05
Pro	mmol kg-1	53	3.46 - 33.1	8.10	4.71	6.34	3.93	3.39	0.81	0.65	0.88	0.91	1.72	1.28	2.06	2.39	6.29	4.67	7.53	8.73
PLS1 Opt n PC					3	12	9	10												
Val	mmol kg-1	53	4.94 - 35.1	8.12	4.32	6.30	3.26	3.33	0.84	0.66	0.92	0.91	1.88	1.29	2.49	2.44	6.99	4.79	9.26	9.06
PLS1 Opt n PC					3	6	5	7												
Cit	mmol kg-1	53	1.24 - 11.1	2.08	1.69	1.71	1.37	1.23	0.58	0.58	0.76	0.80	1.23	1.22	1.52	1.69	5.87	5.80	7.24	8.07