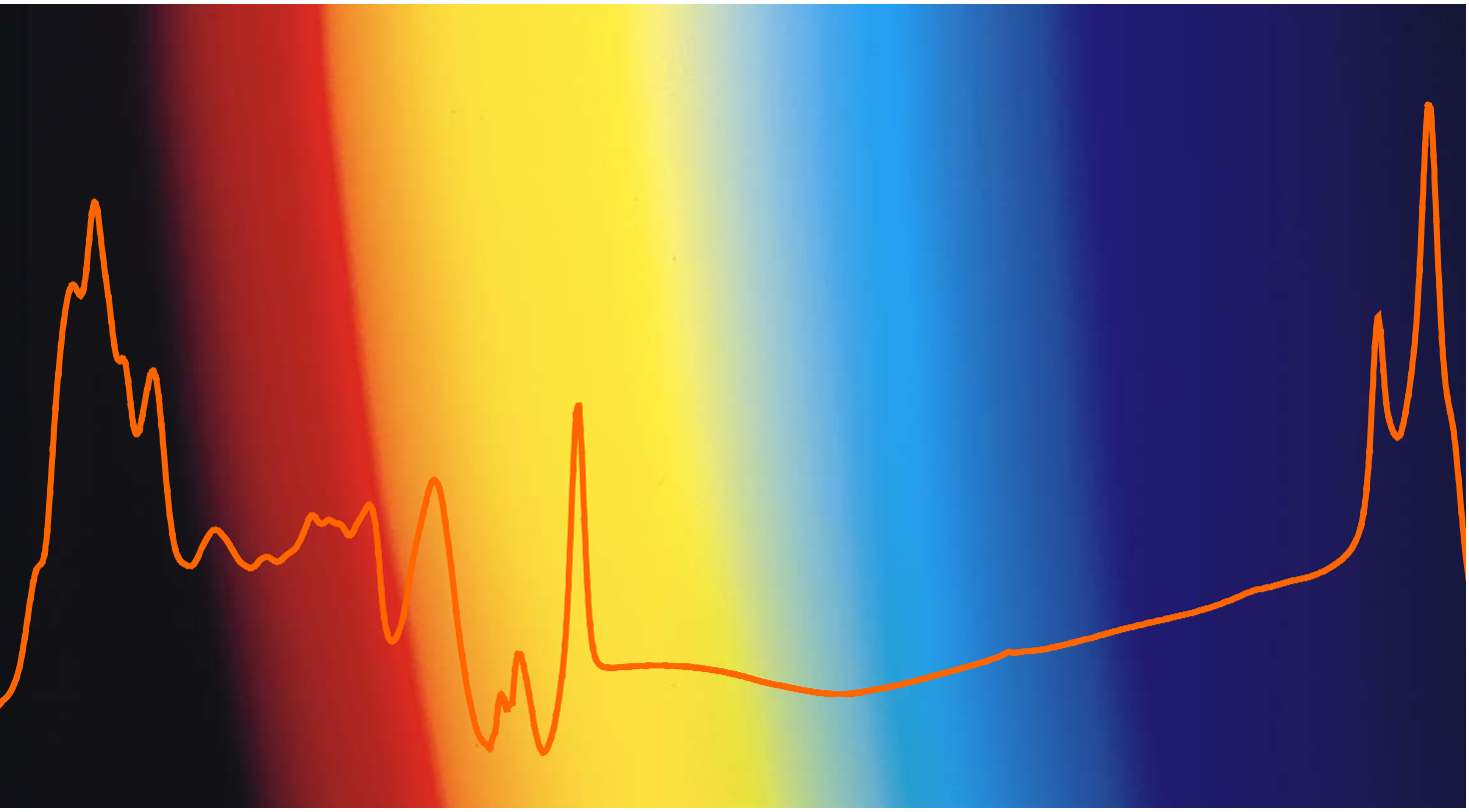


Spectroscopic Analyses on Dairy Products

Per Waaben Hansen

FOSS
FIRST IN FOOD ANALYSIS

Spectroscopic Analyses on Dairy Products

Per Waaben Hansen

PhD Thesis

Industrial PhD Project No. EF 601 (Lc 1575)

December 1998

The Royal Veterinary and Agricultural University
Department of Dairy and Food Science

The Danish Academy of Technical Sciences

Foss Electric A/S

Preface

The present report is the result of my work as an industrial PhD student from December 1995 until November 1998. It was carried out as a co-operation between The Danish Academy of Technical Sciences (ATV), Foss A/S, Denmark, and Food Technology at Department of Dairy and Food Science at The Royal Veterinary and Agricultural University (KVL), Denmark.

The report is aimed at researchers having a basic knowledge in spectroscopy (*e.g.* that absorption frequencies and intensities are intrinsic properties of all chemical substances), and who are familiar with the basic concepts of multivariate data analysis (*e.g.* that latent structures can be extracted from a data set). It provides the reader with an overview of the state-of-the-art, as well as some perspectives for the future within the area of spectroscopic analyses on dairy products. Therefore the report contains brief summaries with only few details. For detailed descriptions of the experiments and more thorough discussions, the reader is referred to the papers and patents following the report.

Many people have been involved in obtaining this final result. It is impossible to mention everyone, but I am particularly thankful to the following organisations and people:

- Foss A/S and The Danish Academy of Technical Sciences (ATV), for providing the funds for the work. Foss A/S is also thanked for providing me with this once-in-a-lifetime opportunity.

- The Food Technology group at the Department of Dairy and Food Science at The Royal Veterinary and Agricultural University (KVL), for letting me into their chemometrics research group, and for providing the equipment for my experiments.

- My supervisors, Tom Riisom (Foss) and Lars Nørgaard (KVL), for useful advice and discussions through three years. And, not least, for giving me the freedom to follow my ideas, while, at the same time, being there to give me constructive comments when I needed that.

- New Zealand Dairy Research Institute (NZDRI) for inviting me to work in New Zealand for seven months. My special thanks to my supervisor, Colin G. Hughes, for introducing me to people and companies in one of the world's most well-functioning dairy industries. My stay in New Zealand was of major importance to me professionally, as well as personally.

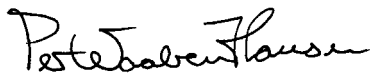
- Gunnar Mogensen, representing ATV, for useful practical advice.

- My colleagues at Foss, KVL and NZDRI, in particular Claus A. Andersson, Bram van Brakel, Rasmus Bro, Henrik V. Juhl, Rikke Nikolajsen, Carsten Ridder, and Henrik Stapelfeldt for useful discussions.

- My reviewers, (in addition to the people mentioned above) Dorrit Baunsgaard, Janette Flatman, Thomas Kjærsgaard Jørgensen and Poul-Erik Petersen, for useful comments and suggestions.

- My family and friends, without whom this had never been possible.

Lyngby, December 1998,



Contents

Publications

English Summary

Compositional analysis of milk and dairy products using spectroscopic methods – especially the use of mid-infrared spectroscopy – is well-established within the dairy industry. The success of these methods is closely related to the developments within the area of multivariate data analysis, or chemometrics. This is because spectroscopic data usually contains numerous overlapping bands, thus making traditional univariate linear regression impossible.

The present project relates to research on the boundary between spectroscopy and chemometrics, because: 1) improved spectroscopic data can result in improved chemometric models, and 2) advanced chemometric methods, on the other hand, are able to extract more information from the existing spectroscopic data. The project examines both situations: Firstly, a collection of spectroscopic methods (near infrared, Raman, ultraviolet/visible, fluorescence and nuclear magnetic resonance spectroscopy) were tested for use with milk. When these methods were compared to the existing mid-infrared method, the latter was the most versatile and reproducible method for use with milk. Secondly, the use of chemometric tools on mid-infrared data was tested. This resulted in application of this spectroscopic method for determination of the low concentration components urea and acetone in milk, as well as for differentiation between specific sugars in dairy process samples. The latter is an example of generation of higher order data, and how this adds new possibilities.

The present work consists of a summarising report followed by publications containing detailed descriptions of the results.

Danish Summary

Måling på sammensætningen af mælk og mejeriprodukter ved brug af spektroskopiske metoder – særligt anvendelsen af midt-infrarød spektroskopi – er veletableret indenfor mejeriindustrien. Metodernes succes er nært forbundet med udviklingen indenfor multivariat data analyse, eller kemometri. Dette skyldes, at spektroskopiske data typisk indeholder utallige overlappende absorptionsbånd, hvilket umuliggør traditionel univariat lineær regression.

Nærværende projekt omhandler forskning i grænseområdet mellem spektroskopi og kemometri, da: 1) forbedrede spektroskopiske data kan muliggøre forbedrede kemometriske modeller, og 2) avancerede kemometriske metoder på den anden side kan udtrække mere information fra allerede eksisterende spektroskopiske data. Projektet ser på begge muligheder: Først undersøgtes en række spektroskopiske metoder (nær-infrarød, Raman, ultraviolet/synlig, fluorescens og kernemagnetisk resonans spektroskopi) for deres anvendelighed i forbindelse med mælk. Sammenlignes disse metoder med den eksisterende midt-infrarøde metode, viste sidstnævnte sig som den bredest anvendelige og mest reproducerbare metode til måling på mælk. Dernæst blev brugen af kemometriske værktøjer i forbindelse med midt-infrarøde data undersøgt. Dette resulterede i anvendelse af denne spektroskopiske metode til bestemmelse af lav-koncentrations komponenterne urinstof og acetone i mælk, såvel som til skelnen mellem forskellige sukkerarter i procesprøver fra mejeriet. Sidstnævnte er et eksempel på måling af højere ordens data, og hvordan der på denne måde kan opnås nye egenskaber.

Nærværende arbejde består af en opsummerende rapport fulgt af publikationer, som indeholder detaljerede beskrivelser af resultaterne.

Chapter 1
Introduction

1.1 HISTORICAL BACKGROUND

The ability to measure the composition or properties of milk and dairy products has been an important tool for a long time. It makes the dairy industry able to pay the farmer for the milk on a fair basis, and to manufacture products of consistent quality.

1.1.1 Infrared spectroscopy

For the past two decades, mid-infrared (mid-IR) spectroscopy has been the most widespread method used for compositional analysis in the dairy industry. As early as in 1961 a patent application for a mid-IR method determining fat, protein and lactose in milk was filed (Goulden, 1964). This principle, based on a monochromator, was used in the IRMA instruments manufactured by Grubb Parson in Britain. However, it was not until instruments based on optical filters were introduced that the method became widespread (Ægidius, 1996). The first purpose-built mid-IR analyser built by Foss Electric based on optical filters was marketed in 1976 (MilkoScan instruments, see Figure 1.1), and since then the method has been developed for various purposes, the major areas being:

- Central laboratory testing, where both tank milk and individual cow samples are tested with regular intervals. The results form the basis for the payment of the farmer, as well as the advice given to him/her to make his/her production more efficient. Typical parameters are: fat, protein, lactose, solids and urea. The main issue in the central laboratories is the sample throughput: The performance of state-of-the-art instruments is 500 samples/hour (Foss MilkoScan 4000).

- Dairy laboratory testing, where raw, intermediate and final products are tested to ensure consistent quality. In such laboratories, the main issue is the versatility of the instruments, so both milk, cream, flavoured, cultured products etc. should be measurable. Filter-based instruments typically provide the parameters: fat, protein, lactose, and solids.

In 1993, the first purpose-built mid-IR instrument based on the Fourier transform infrared (FT-IR) principle was marketed (Anadis MI-200) (Asselain *et al.*, 1996). Contrary to filter-based instruments which measure the absorption at specific wavelengths, FT-IR equipment determines the full spectrum of the sample within the same period of time. This innovation opened for a wealth of possibilities, because suddenly more than 500 wavelengths were accessible. Previously only eleven wavelengths were available in the most sophisticated filter-based instruments. It is generally accepted that the results obtained from FT-IR instruments are comparable to (or better than) results from traditional filter-based equipment (van de Voort *et al.*, 1992b; Luinge *et al.*, 1993; Lefier *et al.*, 1996).

In addition, FT-IR instruments can be standardised (Andersen *et al.*, 1996). This means

Figure 1.1 The first purpose-built mid-IR analyser based on optical filters manufactured by Foss Electric – MilkoScan 203/300.

that by measuring a standard sample, spectra of the same sample from different instruments can be adjusted to become equal. This is an important property, as it makes calibrations transferable between instruments. Thus, it becomes possible to base calibrations on samples measured on different instruments at different locations.

1.1.2 Chemometrics

The existence of chemometrics (or multivariate data analysis) was early recognised to be useful when quantitative measurements based on mid-IR spectra had to be performed. This is because this type of spectral data contains many overlapping bands from all constituents. If a limited sample preparation is preferred, the problem of obtaining a result that is independent of interfering signals can only be solved in a mathematical way.

In the early days (*i.e.* until the mid-eighties), Classical Least Squares (CLS) or "direct calibration" (Martens and Næs, 1989) was used for establishing a relationship between spectral data and reference results. This method is based on the assumption that pure spectra of *all* signal-giving species in a sample are known, and that the signals are strictly proportional to the concentrations of the species (Beer-Lambert law). For the major components (fat, protein and lactose) in milk in low concentrations, these assumptions are reasonable. The composition of milk samples could, therefore, be determined with a relative accuracy of approx. 1 %. At higher concentrations, however, non-linearity became severe and various linearisation schemes were developed. Furthermore, the requirement that all signal-giving species should be known was a severe limitation *e.g.* when an estimate for a low-concentration component was required.

This problem was solved by use of Multiple Linear Regression (MLR) (Martens and Næs, 1989), which only requires knowledge of the concentration of the component of interest in the samples. This method, however, requires that absorption bands have low correlation. This requirement is not fulfilled with most spectroscopic methods, as they are co-linear by nature.

Multivariate methods based on latent vectors, such as Principal Component Regression (PCR) and Partial Least Squares (PLS) therefore became of interest, as they are able to take advantage of the structures in highly overlapping and co-linear data (Martens and Næs, 1989). The use of PLS with mid-IR spectroscopy made it possible to determine low-concentration components in or marginal properties of milk, such as urea (ranging from 0.01 to 0.08 %) (Nygaard *et al.*, 1993) or the freezing point depression (FPD) (Arnvidarson *et al.*, 1998). But it was not until the introduction of FT-IR technology that the full potential of these methods could be utilised.

1.2 SCOPE OF THE PROJECT

The present project focuses on the question: Which opportunities does today's spectroscopic methods and chemometric tools provide the dairy or dairy farmer with?

The dairy farmer will focus on parameters which allow him/her to monitor the cow's health status, while at the same time getting the most out of the resources he/she is investing. Parameters providing such feeding and health status information include urea and acetone, which are only found as trace amounts in milk. Thus, the main difficulty here is the concentration level.

At the dairy, on the other hand, the production engineers are mainly occupied with parameters relating to the properties of milk or the derived products. Milk with a given composition might be suited for a specific use, *e.g.* a milk high in casein is desirable for cheesemaking. Thus, fat, protein and lactose predictions are not sufficient. The dairy engineer is interested in knowing the more detailed composition of these three groups: fat fractions, specific proteins, or specific sugars. The difficulty is thus one of differentiation rather than the detection of low concentrations.

The solution to these problems is expected to be found by focusing on the following:

Use of spectroscopic methods other than FT-IR. One approach is to test other spectroscopic methods for use with milk. These include: near infrared (NIR), Raman, ultraviolet/visible (UV/VIS), fluorescence, and nuclear magnetic resonance (NMR) spectroscopy. They may provide information complementary to mid-IR, and thus the determination of parameters not obtainable by the present state-of-the-art methods. Optimal results might not be obtained in this area, as the methods will require further mechanical and electronic development to make them suitable for use with milk.

Optimised use of the FT-IR method. As mid-IR is a well-tested method in the area of milk analysis, FT-IR has already been optimised for use with milk. *E.g.* flow systems and temperature controls have been built to ensure a reproducible measurement not dependent on external factors. Thus, as the mechanical and electronic parts of the instruments have been improved almost to perfection, the data analysis is the area where major achievements are to be expected. An improved data analysis could provide a better extraction of minor spectral features informing about the parameters of interest to the dairy farmer and the dairy engineer. The fact that today's instruments obtain full spectrum information is particularly interesting, because this opens up for the use of methods handling such situations. Data pre-treatment (*e.g.* by use of derivatives or multiplicative scatter correction (MSC) (Geladi *et al.*, 1985)) and variable selection (*e.g.* by use of genetic algorithms (Leardi *et al.*, 1992) or principal variables (Höskuldsson, 1992)) are areas which require exploration. Further, if the FT-IR method can be used for generating higher-order data this opens up for the use of higher-order methods, such as Parallel Factor Analysis (PARAFAC) (Bro, 1997) or N-PLS (Bro, 1996).

The present project relates to a combination of these two approaches: Data from spectroscopic methods other than FT-IR will be used in combination with basic chemometric tools to evaluate their advantages and disadvantages. The intention is to do so using minimal sample pre-treatment. Only the major milk components (fat, protein and lactose) will be investigated in relation to these methods (Chapter 2).

The FT-IR data will mainly be used when the task is to test new chemometric methods or components occurring in low concentrations. The parameters under investigation with FT-IR are: urea (Chapter 3) and acetone (Chapter 4) in milk, as well as specific sugars in other dairy products (Chapter 5). In addition, the sampling method known as attenuated total reflectance (ATR) will be tested for use with milk in combination with FT-IR spectroscopy (Chapter 3).

Chapter 2

Chemometrics and Spectroscopy on Milk

2.1 INTRODUCTION

2.1.1 Spectroscopic methods

At present, mid-IR spectroscopy is the preferred method for milk and dairy product analysis. This is due to the fact that almost every chemical substance (apart from some salts and very simple chemical compounds) has its own distinctive spectrum. Only substances occurring in very low concentrations can be difficult to determine, as the noise level of the method might be encountered in such a case.

The main disadvantage of mid-IR spectroscopy on aqueous samples is the strong absorptions of water. The O-H bending band (at approx. 1650 cm^{-1}) effectively obscures potentially useful absorptions from *e.g.* protein, urea and acetone.

Therefore, it is of interest to know the advantages and disadvantages of other spectroscopic methods, as they might complement the mid-IR method. In the present study, the following methods have been chosen for further examination: NIR, Raman, UV/VIS, fluorescence and NMR spectroscopy. These methods have been selected, either because measurements are fast or because they contain spectral information which cannot be obtained by use of mid-IR spectroscopy.

2.1.2 Chemometric methods

Data sets, as well as chemometric methods, can be said to be of a certain *order* according to the dimensions of the data obtained from one sample. The following terms are used: zero, first, second and higher orders:

In case of *zero order data*, only one figure is obtained from each sample, *e.g.* the absorption at one specific wavelength. Collections of such data can be handled using a *first order method*, such as univariate linear regression, to establish a correlation. Interferences cannot be handled in such a case.

First order data is obtained when a whole spectrum, or a number of distinct absorptions, are collected from each sample. This is the most common data structure in spectroscopy. *Second order methods*, such as Principal Component Analysis (PCA), PLS, etc. (Martens and Næs, 1989) can perform analysis or regression on collections of such data. These methods can handle interferences, as long as they are known at the time of calibration.

Second (or higher) order data is obtained when a spectral three-dimensional landscape is collected for each sample. This is the case when a spectrum of the same sample is recorded under several different circumstances, such as varying pH, temperature, dilution, the course of a chemical reaction, or, in case of fluorescence, excitation wavelength (see Figure 2.4). *Third (or higher) order methods*, such as PARAFAC (Bro, 1997) or N-PLS (Bro, 1996) can be used to perform analysis or regression on collections of these data. With these methods, even interferences not present at the time of calibration can be handled. Collections of second order data can also be analysed by second order methods. To make the information useful for *e.g.* a PLS regression the landscapes must be unfolded, *i.e.* all spectra originating from the same sample are appended to each other to give a very wide "spectrum". In this way, however, the three-dimensional relations from the landscape disappear, and information is lost.

In the present study, most data are of first order (one spectrum per sample), and therefore full spectrum PLS regression will be performed on all data sets.

Fluorescence emission spectra are dependent on the excitation wavelength. If the fluorescence of the samples is scanned in both the excitation and emission mode (see section 2.5 and Figure 2.4), three-dimensional landscapes, *i.e.* second order data, will be pro-

	Sensitivity	Information content	Absence of interferences	Repeatability	Absence of light scatter	Sample handling	Other
NIR	**	**	*	**	**	***	
Raman	*	**	**	**	**	***	Weak water bands
UV/VIS	***	*	***	***	**	*	
Fluorescence	***	*	***	**	*	*	Higher order data structure
NMR	**	***	*	**	***	***	Parameter settings are crucial
FT-IR	***	***	*	***	**	**	

Table 2.1 Advantages and disadvantages of the spectroscopic methods studied when used for measuring undiluted milk. Dilution had to be performed prior to the UV/VIS experiments. The properties of FT-IR are shown for comparison.
Symbols: *Sensivity* and *Information content*: * = low, *** = high. *Absence of interferences*: * = many interferences, *** = few interferences. *Repeatability*: * = poor, *** = good. *Absence of light scatter*: * = severe light scatter, *** = no light scatter. *Sample handling* is an indication of whether sample handling is easy (*e.g.* no dilution required), and whether instrumentation is simple (*e.g.* whether conventional optics can be used). * = difficult, *** = easy.

duced. Therefore fluorescence data are best handled by a third order method, such as PARAFAC (for decomposition of the spectral matrix) or N-PLS (for predictive purposes). PARAFAC can resolve fluorescence landscapes into their basic excitation and emission spectra. This has been used for resolving the individual components contributing to the landscape obtained from sugar process samples (Bro, 1997). As will be shown, it is potentially useful for milk, as well.

2.1.3 Outline of the study

In the following, the different spectroscopic methods will be discussed one at a time. References to useful recent reviews or books will be given. The advantages and disadvantages of the methods will be highlighted, and important considerations when implementing the methods for use with milk and dairy products will be discussed. Focus will be on whether the methods complement the existing mid-IR equipment or not.

In Table 2.1 the methods are compared in terms of their sensitivity, information content, amount of interferences, repeatability, light scatter effects and ease of sample handling. These properties were chosen because they are important when fast spectroscopic methods for compositional measurements are considered.

Experiments have been carried out to assess the feasibility of measuring the major milk components, *i.e.* fat, protein and lactose, preferably in undiluted milk. NMR was not tested for lactose determination, as the experiments were not focused on this parameter. A summary of the results, in terms of accuracy and repeatability, is presented in Table 2.2. The experiments were set up in a way that would prove whether the principle is useful, not necessarily to get the optimal result. This means that improved experimental set-ups will be expected to give better results in terms of accuracy.

2.2 NIR SPECTROSCOPY

NIR spectroscopy is already recognised as a useful method for compositional analysis of food products, especially for meat and grain. It is based on vibrational overtones and combination bands observed in the mid-IR region. A review on dairy applications of NIR spectroscopy has been published recently (Rodriguez-Otero *et al.*, 1997).

		Fat			Protein			Lactose		
	[1]#samples	[2]#factors	RMSEP	s_r	[2]#factors	RMSEP	s_r	[2]#factors	RMSEP	s_r
NIR	49	6	0.085	0.027	8	0.121	0.066	6	0.105	0.050
Raman	15	4	0.135	0.070	5	0.321	0.193	-	-	-
UV/VIS	47	1	0.609	0.015	5	0.056	0.020	-	-	-
Fluorescence	50	2	0.590	0.030	3	0.220	0.030	-	-	-
NMR	15	3	0.267	0.159	2	0.176	0.048	[3]	[3]	[3]
[4]FT-IR (typical)			0.032	0.008		0.032	0.008		0.032	0.008

Table 2.2 Overview of results in terms of accuracy (RMSEP) and repeatability (s_r) for the various spectroscopic methods studied. The accuracy was calculated using four or six cross validation segments. For NMR, full cross validation was used. Blank entries indicate that no correlation could be established.
[1]#samples = number of samples used for the evaluation.
[2]#factors = number of PLS factors used in the model.
[3]Only fat and protein were tested in relation to NMR.
[4]The typical performance of a MilkoScan FT 120 FT-IR instrument is shown for comparison.

The advantages of the method is its ability to use longer pathlengths than mid-IR, and that the optical equipment used is much simpler. For example, optical fibres made from quartz glass can be used.

The main disadvantage is the low sensitivity of the signal compared to what is obtained from mid-IR. Thus, low concentration components cannot be expected to be determined by use of NIR.

2.2.1 Experimental set-up

The experiments were carried out using a Bomem MB Series FT-NIR instrument. It was equipped with a purpose-built temperature-controlled glass cuvette (fixed at 40 °C) with a pathlength of 1 mm connected to a flow system fed by a peristaltic pump. All spectra were ratioed against water, log-transformed to yield absorbencies, and spectral ranges obscured by water were removed.

Predictions for fat, protein and lactose obtained from mid-IR spectra (using the Foss MilkoScan FT 120 instrument) were used as reference results for the calibrations.

2.2.2 Results and discussion

50 individual cow samples were measured in triplicate. One was a spectral outlier (detected by PCA) and was removed from the data set. The average spectrum of the samples is presented in Figure 2.1. The results are shown in Table 2.2.

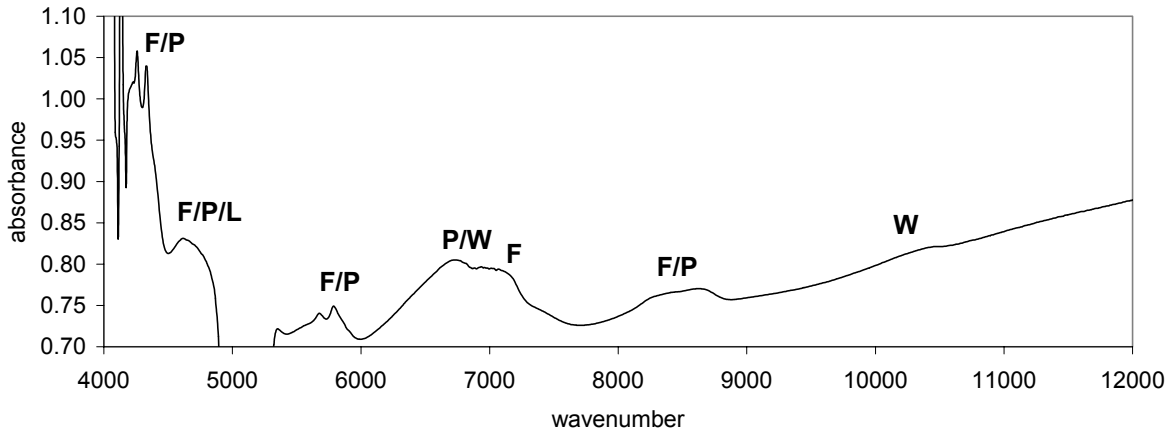


Figure 2.1 Average NIR spectrum of milk. The peaks originating from the major milk components are indicated, where possible. F = fat, P = protein, L = lactose, W = water. Assigned using data from Williams and Norris (1987).

Of the methods tested here, NIR has the best overall performance. It is actually the only method capable of predicting lactose in the samples. But, in the present context, where low-concentration components in milk are of interest, the NIR method does not provide *any* information which could not be obtained by mid-IR.

2.3 RAMAN SPECTROSCOPY

Raman spectroscopy (reviewed by Ellis *et al.*, 1989; Hendra *et al.*, 1995; Adar *et al.*, 1997) is a spectroscopic method which provides primary vibrational information similar to mid-IR spectroscopy. When monochromatic excitation light (*e.g.* generated by a laser) interacts with a molecule, the frequency of the emitted light might be shifted an amount corresponding to a vibrational transition in the molecule. When the intensity of the shifted light is plotted against the shift, a Raman spectrum is obtained. As the effect generating the signals is very different from mid-IR spectroscopy, the band intensities are not necessarily the same. This means that, although they both monitor molecular vibrations, bands which are weak in mid-IR might be strong in Raman and vice versa. Thus, Raman can provide information which is complementary to mid-IR. And, very importantly, the Raman spectrum of water is very weak, so all parts of the vibrational spectrum of milk can be determined.

Basically, Raman experiments can be carried out in two ways:

- CCD-Raman, using visible excitation (*e.g.* a HeNe laser emitting at 633 nm) and a charge-coupled device (CCD) for detection. Such measurements have the advantage of being fast (almost real time), and the equipment is relatively simple. Unfortunately, with visible excitation milk has a strong fluorescence peak which overlaps the weak Raman signal to an extent which cannot even be corrected for by use of chemometric tools.

- Fourier transform-Raman (FT-Raman), using NIR excitation (a Nd:YAG laser emitting at 1064 nm) and detection by an interferometer of the same type as is used for FT-IR and FT-NIR applications. The advantage of this method is the absence of fluorescence and the short time required for obtaining a full spectrum. The disadvantage is that a strong NIR absorption of water coincides with the range in which the shifted light is collected. This NIR band absorbs the shifted light, resulting in weak Raman signals.

In both situations, Raman spectroscopy has the advantage of being able to handle samples which cannot be introduced into the narrow mid-IR cuvette. Studies on milk powders and cream by use of FT-Raman have been carried out (Fehrmann *et al.*, 1995; Franz *et al.*, 1996).

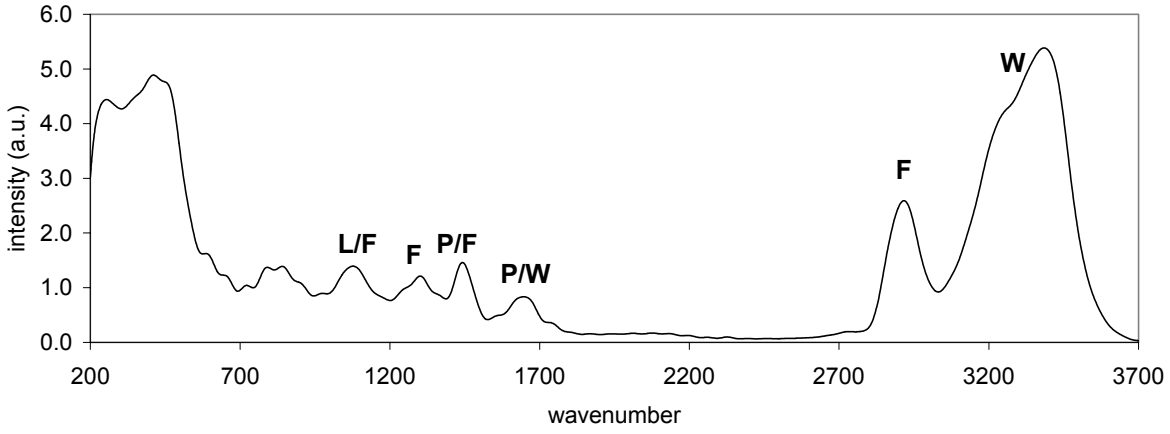


Figure 2.2 Average FT-Raman spectrum of milk. The peaks originating from the major milk components are indicated, where possible. F = fat, P = protein, L = lactose, W = water.

2.3.1 Experimental set-up

It was not possible to perform CCD-Raman measurements (using a Dilor LABRAM II instrument) due to the strong fluorescence of milk.

FT-Raman measurements were carried out using a Perkin-Elmer System 2000 with FT-Raman accessory. The instrument was equipped with a 4×7 mm quartz flow cuvette (fixed at 40 °C) connected to a flow system fed by a peristaltic pump. The collection angle was 180° (backscatter).

Mid-IR predictions for fat, protein and lactose were used as reference results for the calibration work.

2.3.2 Results and discussion

15 tank milk samples fortified with protein, lactose, citric acid and urea (in order to span the full range of normal samples) were measured in triplicate. The average spectrum of the samples is presented in Figure 2.2, and the results can be found in Table 2.2.

In spite of the low intensity due to the NIR water band absorbing the shifted light, reasonable correlations were obtained, especially for fat. This is not surprising, as the strongest Raman bands observed in the spectrum can be ascribed to fat and water.

Qualitative measurements of cream were also carried out. This gave very reproducible spectra being very similar to milk spectra. This also indicates that the Raman spectrum mainly detects fat and changes therein.

2.4 UV/VIS SPECTROSCOPY

UV/VIS spectroscopy (Silverstein *et al.*, 1991) is probably the spectroscopic method used most frequently in analytical chemistry, usually as the final detection of a compound after purification, separation, and/or after the course of a colour-producing reaction. One recent study has been directed towards analysis of whole milk with minimal sample pre-treatment (Reichardt *et al.*, 1995).

UV/VIS spectroscopy detects the transfer of electrons between electronic levels. Only for highly conjugated systems and carbonyl bonds are these transfers of an energy low enough to be seen in the visible or near-ultraviolet region (*i.e.* above 200 nm). Of the major components in milk, only proteins (due to the tyrosine, tryptophan and phenylalanine side chains, as well as the amide bond) are expected to contribute to the UV/VIS spectrum.

The advantages of UV/VIS spectroscopy is its high repeatability, sensitivity, and the almost complete absence of interferences, as few chemical compounds absorb in this region. This means that compounds showing UV/VIS absorptions will be detected with a high accuracy. At the same time the equipment is relatively simple and low-cost.

The main disadvantage of the method is its low information content, as it is difficult to base a multi-component analysis on such signals. In addition, light scattering might cause problems.

2.4.1 Experimental set-up

The experiments were carried out on a Perkin-Elmer Lambda 2 instrument. It was equipped with a purpose-built temperature-controlled CaF_2 cuvette (fixed at 40 °C) with a pathlength of 37 μm connected to a flow system fed by a peristaltic pump. The samples were diluted ten times to obtain absorbances in the range where the method performs optimally. All spectra were ratioed against water.

FT-IR predictions for fat, protein and lactose were used as reference results for the calibration work.

2.4.2 Results and discussion

50 individual cow samples were measured in triplicate. Three spectral outliers (detected by

PCA) were removed from the data set. The average spectrum of the samples is shown in Figure 2.3, and the results are shown in Table 2.2.

Only protein has significant UV/VIS absorptions, but fat could also be detected indirectly as a result of the light scatter effect and the inherent correlation between fat and protein. The fat results show a poor accuracy – a problem which might be solved by homogenisation of the samples before the measurement. This was done by Reichardt *et al.* (1995), and is the most likely reason for their better fat accuracy.

The UV range of the spectrum can provide an excellent protein determination, but apart from that, the value of UV/VIS spectroscopy for direct measurements on milk is limited, unless a milk additive showing UV/VIS absorptions is sought. Some antibiotics and additives fulfil this requirement, but such compounds are usually found in extremely low concentrations, and will therefore be difficult to detect.

2.5 FLUORESCENCE SPECTROSCOPY

Fluorescence spectroscopy (Lakowicz, 1983) is similar to UV/VIS spectroscopy in that it measures electronic transitions, but it is detecting emitted rather than transmitted light. This results in fewer interferences, as not all samples having a UV/VIS spectrum exhibit fluorescence. Until now, the method has been used mainly as a sensor for HPLC or for detection by use of dye-binding techniques.

Fluorescence is detected as the light emitted following excitation by monochromatic light. This is similar to Raman spectroscopy, but contrary to the latter method which measures vibrational transitions, fluorescence monitors electronic transitions. The effects are competitive, but usually fluorescence is much stronger than the Raman effect.

As the emission spectrum is highly dependent on the excitation wavelength, both an excitation *and* an emission spectrum can be obtained. If the emission spectrum is measured for various excitation wavelengths, a landscape similar to the one shown in Figure 2.4 will appear. The fact that landscapes can be generated gives fluorescence spectra a property which is not common in spectroscopy: a second order structure for each sample. This can be analysed using *e.g.* PARAFAC.

Thus, the advantages of fluorescence spectroscopy is the absence of interferences and that higher-order data can be obtained from it.

The most severe disadvantage is the strong dependence on light scatter, and there are no means for making mathematical corrections (as in *e.g.* NIR or UV/VIS), because no information on the amount of scatter is contained in the spectrum.

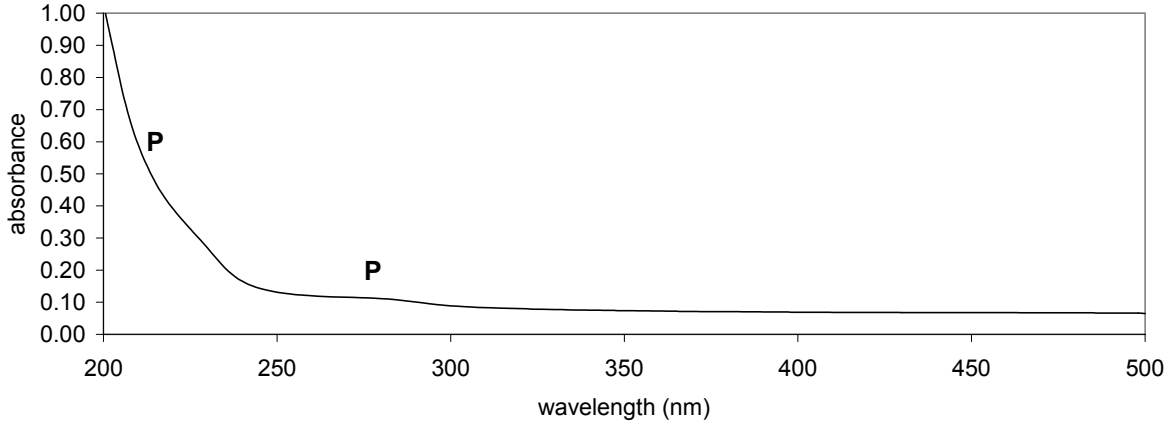


Figure 2.3 Average UV/VIS spectrum of milk. Only protein contributes with absorption bands (marked as "P"). Fat is seen indirectly as a baseline shift caused by light scatter.

2.5.1 Experimental set-up

The experiments were carried out on an Aminco-Bowman Series 2 fluorometer. A set-up with detection in a 90° angle to the excitation was used. The instrument was equipped with a 4×7 mm quartz flow cuvette (fixed at 40 °C) connected to a flow system fed by a peristaltic pump.

FT-IR predictions for fat, protein and lactose were used as reference results for the calibration work.

2.5.2 Results and discussion

50 individual cow samples were measured in triplicate. An ordinary PLS calibration, where the landscapes were unfolded to wide "spectra", was carried out. The results presented in Table 2.2 originate from these calculations.

The result is the poorest of all presented here. Normalisation (*i.e.* division of each spectrum by its vector norm) was applied to the spectra to obtain it. Protein information is clearly present in the spectra (see Figure 2.4), but it cannot be used. Even when N-PLS was used, no improvement was seen. This is due to the light scatter effect, which could not be corrected for. Thus, a quantitative measurement based on milk without any prior pre-treatment is not possible.

Even though a quantitative measurement is impossible, the qualitative resolution of the spectral landscapes by use of PARAFAC can still be performed.

The emission spectra were reduced in size by averaging groups of five points to speed up the calculations. PARAFAC was performed on the 50 samples, trying from two up to four contributing components in the fluorescence landscape. Three components were found to be optimal. Their resolved excitation and emission profiles are shown in Figure 2.5. From these, the excitation and emission maxima of the three components could be deducted. These values, together with their possible explanations, are shown in Table 2.3.

The three components were assigned to be (compare to Figure 2.4): Protein (peak at short wavelengths) (Lakowicz, 1983), riboflavin (peak at long wavelengths) (Lakowicz,

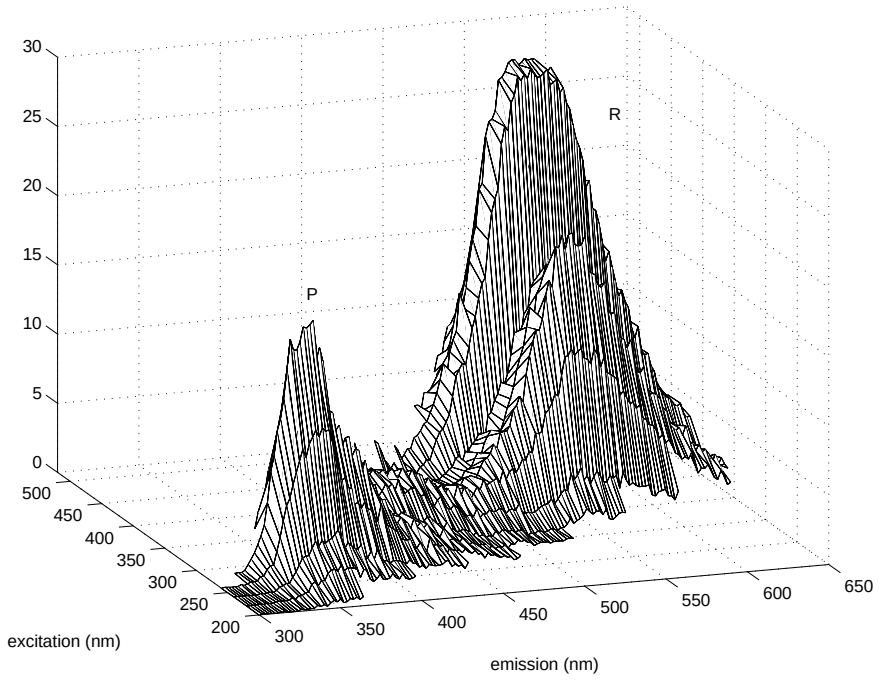


Figure 2.4 Fluorescence landscape of a milk sample containing 3.4 % fat, 3.3 % protein, and 4.8 % lactose. The fluorescence peak marked as "P" originates from protein (mainly tryptophan side chains), while the tallest ("R") is likely to originate from riboflavin (vitamin B_2).

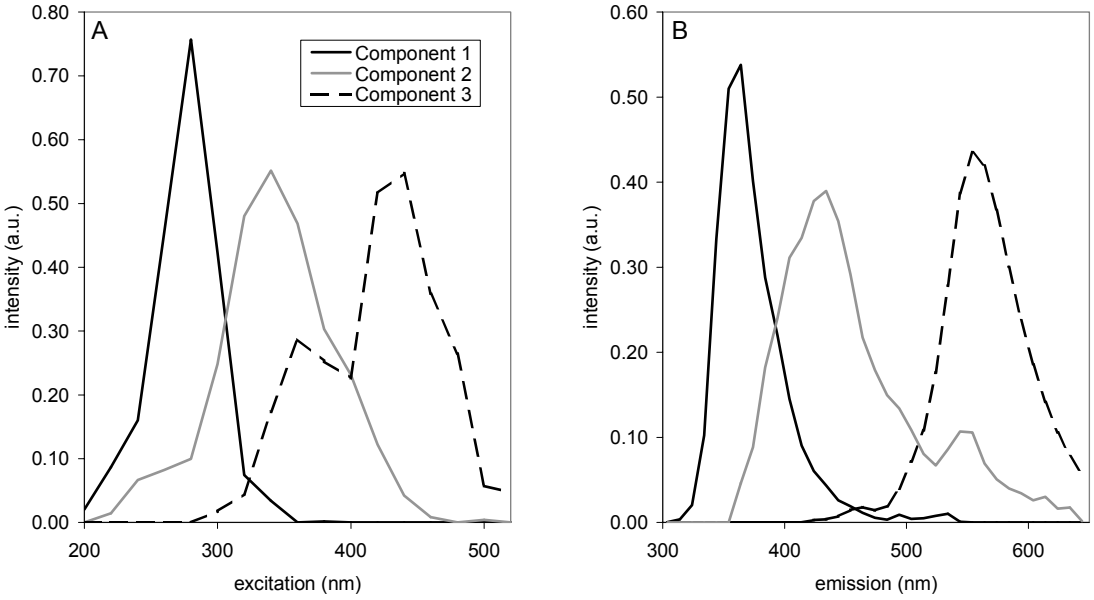


Figure 2.5 Resolved profiles obtained by use of PARAFAC. Part A shows the excitation mode. Part B shows the emission mode. The sample mode (scores) is not shown.

1983), and protein-aldehyde condensation products. The latter group can either originate from the Maillard reaction (Morales *et al.*, 1996) or from aldehydes (resulting from fat oxidation) reacting with milk proteins (Stapelfeldt and Skibsted, 1994). As the samples were fresh, and the temperature to which the samples were subject was relatively low, the latter reaction is the most likely explanation to this peak.

The emission maxima are generally at longer wavelengths (red-shifted) than the values reported in the literature. The phenomenon was investigated by dilution of the samples, and it was observed that the protein emission maximum shifted to approx. 345 nm (from 360 nm) on dilution, while the corresponding value for riboflavin was 535 nm (from 555 nm). This effect, called reabsorption, is usually seen for strongly absorbing samples. The maxima are shifted as a result of the newly emitted light being absorbed by the sample before it reaches the detection – an effect, which is wavelength dependent. In the present case the effect is mainly caused by light scatter rather than absorption, but the result is the same.

	$\lambda_{excitation}$	$\lambda_{emission}$	Possible assignment
Component 1	280	360	Protein (tryptophan side chain)
Component 2	340	435	Protein-aldehyde products
Component 3	430	555	Riboflavin (vitamin B_2)

Table 2.3 Components observed in the fluorescence landscape of milk with their associated excitation and emission maxima. The numbers are approximate only, as the spectral resolution is low in both directions. The assignments are based on results from other studies of milk (Lakowicz, 1983; Stapelfeldt and Skibsted, 1994; Morales *et al.*, 1996).

The method is highly sensitive: riboflavin is present in milk on a 0.2 mg/dl level. The concentration of protein-aldehyde condensation products is also expected to be low. Their production is promoted by heating.

Therefore, if the problems regarding light scattering are solved, fluorescence might prove to be a sensitive method for detection of the riboflavin content of milk as well as the deterioration of the samples (*e.g.* as a result of fat oxidation or heating).

2.6 NMR SPECTROSCOPY

NMR spectroscopy (reviewed by Brosio and Barbieri, 1996) detects the spin resonance of a given nucleus. Most common is the use of ^{1}H and ^{13}C resonances, but also ^{31}P is used. In quantitative applications, ^{1}H-NMR is the most frequently used method, as the signal is

strong due to the high abundance of this nucleus. Therefore such experiments are fast compared to other types of NMR. In the present context, only pulsed ^{1}H-NMR experiments are considered.

Pulsed NMR experiments uses a radio frequency pulse to bring the spins of the nuclei of the sample into the same state. Subsequently, the return of the spins back to normal is detected as the so-called free induction decay (FID). When a Fourier transform is applied to the FID, a spectrum with NMR intensities against frequency is obtained.

Apart from the nucleus chosen, another parameter which is important in determining the speed of the experiment is the strength of the permanent magnetic field in which the sample is situated. High fields (*e.g.* 500 or 600 MHz) provide highly resolved spectra with information on the various protons present in the sample. It is possible to measure high-quality NMR spectra of milk in this way. However, such equipment is expensive and requires skilled personnel. Thus, it is only useful for research purposes.

Low field NMR (field strengths in the range from 10 to 30 MHz), on the other hand, has proved very useful in the food industry, *e.g.* as a standard method for the determination of the solid fat content of fats and oils (Official methods and recommended practices of the American Oil Chemists' Society, 1996). The resolution is low, and therefore the FID is usually not Fourier transformed to give a spectrum. Instead, the analysis is performed directly on the FID (or relaxation) data.

The advantage of NMR spectroscopy is that the decay of a given proton is influenced by its surroundings. Other types of spectroscopy (*e.g.* mid-IR or fluorescence) are also sensitive to the molecular surroundings, but not to the same extent. For example, the relaxation of water is much faster when it is in a bound state than when it is free. NMR is rich in information of this type. Therefore NMR can provide information which cannot be gathered anywhere else. Another advantage is its capability to handle large samples: With the equipment used in the present study samples can have a diameter of up to 2.6 cm.

The disadvantage is that there is a wealth of parameters to optimise. The pulse sequence used is crucial to the result. Use of the wrong pulse sequence or parameter settings might result in a negative outcome of the experiments. This makes multivariate methods of utmost importance for building a useful method. In a multivariate context, the many adjustable parameters can even be seen as an advantage.

2.6.1 Experimental set-up

The NMR measurements were carried out using a Resonance Instruments MARAN instrument (field strength: 23.2 MHz). The probe was fixed at 35 °C, and the samples were presented in cylindrical tubes with a diameter of 2.6 cm. The relaxation data was corrected for sample weight.

Mid-IR predictions for fat and protein were used as reference results for the calibration work.

2.6.2 Results and discussion

15 tank milk samples fortified with protein, lactose, citric acid and urea (in order to span the full range of normal samples) were measured in triplicate. PLS calibrations against fat and protein were carried out using relaxation data from two pulse sequences: inversion recovery (T_1 determination) and Carr-Purcell-Meiboom-Gill (CPMG, T_2 determination). In addition, the exponentially fitted relaxation times T_1 and T_2 were used for univariate and bivariate calibration. The pulse sequences and relaxation times are described by Rutledge (1992) and will not be discussed here. These two sequences are known to be useful for fat and protein, and therefore only these parameters were tested. The results from the pulse sequences yielding the best results are presented in Table 2.2.

The best fat prediction was obtained when PLS calibration was performed directly in the inversion recovery relaxation profiles. Protein results were best if T_1 and T_2 were correlated to protein by use of MLR (or a full-factor PLS model). This shows that the result is highly dependent on the pulse sequence: a given component might not be measured when certain pulse sequences are used.

Thus, as a result of its ability to be tailor-made for a given application, NMR is a very versatile method. Furthermore, samples that are difficult to handle by other methods can be introduced into the NMR instrument.

2.7 CONCLUSIONS

In the present chapter, a brief summary of the results of simple experiments testing five very different spectroscopic methods has been given. The advantages and disadvantages of the methods are collected in Table 2.1. The major conclusions are:

- NIR does not provide information that is not already available in mid-IR spectroscopy, but it offers easier sample handling.

- Raman mainly provides information on fat. FT-Raman is the most useful experimental set-up for milk applications.

- UV/VIS provides an excellent protein determination, but requires sample dilution.

- Fluorescence requires removal of light scattering in order to be useful. If this is done, protein, riboflavin and sample deterioration might be determined.

- Of the methods studied, ^{1}H-NMR is the method that has the widest application range due to its flexibility. In addition, sample handling is simple.

In general, methods having a high information content suffer from the problem of many interferences. This is usually no problem when multivariate methods are used. Therefore, from the present preliminary investigations, it is clear that mid-IR is the most widely applicable method for milk analysis – especially when FT-IR technology is used. Most chemical compounds absorb in this range, and the method has been developed to a level where relatively low concentrations can be detected (*e.g.* 1 mM or 58 ppm in case of acetone in milk, see Chapter 4). This is supported by the results presented in the following chapters.

The major drawback of FT-IR when compared to NIR, Raman and NMR is the requirement that the sample has to be presented in a narrow cuvette, as this makes measurements on certain products impossible.

Chapter 3

Urea Determination in Milk

3.1 BACKGROUND

The urea content of milk is an important marker for the nutritional status of each individual cow. The amount of feed protein that is incorporated into milk proteins depends on the energy supplied (as carbohydrates) through the feed. In case of a low energy diet, the cow uses protein to produce milk sugars which is accompanied by production of urea. This means that:

- *If* the protein to energy ratio of the feed is high, the cow is using some protein for producing milk sugar, and the urea content of the milk will be high.

- *If* the protein to energy ratio of the feed is low, there is no protein for milk sugar production, and the cow's feed may be too low in protein for other purposes. In this case, the urea content of the milk will be low.

As a result, there is a good correlation between milk urea concentration and protein to energy ratio in the feed (Oltner and Wiktorsson, 1983).

Both extremes are undesirable, as the farmer does not obtain the full benefit from the the cows' feed. Therefore cows are classified into a "low", "normal" and "high" urea group, allowing the farmer to optimise the feeding of the individual cows. Thus, to reach the optimal protein to energy ratio, determination of urea is of value. Such a determination has to be fast and low-cost, as all cows should have their milk screened with regular intervals, to maintain useful surveillance.

The state-of-the-art for rapid urea determination is to use either a flow injection (FIA) method, which is capable of handling up to 100 samples per hour (Lefier, 1996), or a filter-based mid-IR instrument handling up to 500 samples per hour (Nygaard *et al.*, 1993). The latter method is, however, less accurate, and there is a need for a method which combines high speed with high accuracy. FT-IR could provide this.

Two experiments on urea determination in milk have been carried out: One where an ordinary transmission cuvette was used, and one where ATR sampling was used.

3.2 TRANSMISSION EXPERIMENTS

FT-IR instruments for compositional measurements on milk are generally using transmission cuvettes with pathlengths from 30 to 50 μm. This dimension is chosen as it combines a good signal-to-noise (S/N) ratio with a spacing that allows for replacement of the sample. Therefore this set-up is the obvious choice when determination of a new compositional parameter is the subject.

The work is described in Publication I (Hansen, 1998a), where FT-IR is tested as a method for urea determination in milk. The following is a summary of the results from this publication.

3.2.1 Description of the data set

The data set consisted of 900 individual cow samples measured in duplicate on an FT-IR instrument. After removal of 17 outliers, the samples were split into three sets: a calibration set (395 samples), a test set (292 samples) for finding the optimal number of PLS factors, and an independent test set (196 samples) for evaluation of the final model.

Only chemically meaningful ranges of the mid-IR spectrum (approx. 1000 to 3000 cm^{-1} with water bands and spectroscopically "inert" ranges removed) were used for calibration.

3.2.2 Result summary

Full spectrum PLS calibrations were carried out using various pre-processing methods, such as derivatives, MSC (Geladi *et al.,* 1985), and normalisation. The results found on the independent test are presented in Table 3.1.

Nothing was gained by subjecting the FT-IR data to various types of pre-processing: The best result (an accuracy (Root Mean Square Error of Prediction, RMSEP) of 3.9 mg/dl) was obtained by using the raw absorption matrix. The result using this model is shown in Figure 3.1.

Even though the filter-based instrument performs with an accuracy within the specifications of this type of equipment, there is still a large improvement in validation error (29 %) when using the FT-IR method. Thus, this method provides the farmer with a more accurate urea determination, which will be a better classification tool. The FT-IR method, however, is relatively slow (120 samples per hour) in its present configuration, so the issue of throughput will require attention to make it fully comparable to the filter-based mid-IR method.

A further important advantage of the FT-IR method is that it can be standardised (Andersen *et al.*, 1996), so once a robust calibration has been developed it should be globally transferable between instruments.

Treatment	[1]#fac	R^2	RMSEP	s_r
None	16	0.87	3.9	1.3
1st derivative	19	0.86	4.1	1.3
2nd derivative	16	0.84	4.3	1.7
Additive MSC	17	0.86	4.0	1.2
Full MSC	14	0.86	4.0	1.3
Normalisation	17	0.86	4.0	1.2
[2]Filter instrument	8	0.74	5.5	-

Table 3.1 Transmission experiments: Result of the validation on the independent test set (196 samples) using various types of pre-processing. Results are stated in the unit mg/dl.
[1]#fac = The number of PLS factors used in the model selected by the use of the test set (292 samples).
[2]The result from the traditional filter-based instrument (MilkoScan 4000 with 10 filters for urea determination) is shown for comparison. Exactly the same sample sets have been used in the calculations for this instrument.

3.3 ATR EXPERIMENTS

The principle behind ATR is that when a sample has an interface with a crystal of a high refractive index, most of the electromagnetic radiation reaching the interface from the crystal side will be reflected. However, a small amount of radiation will penetrate the interface and enter the sample. This radiation will interact with the sample, and in the case of mid-IR light, a mid-IR spectrum is produced. Thus, ATR is a sampling method based on reflection rather than transmission. The ATR set-up used for the present study is shown in Figure 3.2.

The main advantage of ATR sampling is that it can handle samples which cannot be introduced into an ordinary transmission cuvette, because sample thickness is no problem. Thus, pastes and viscous samples are easily measured. It is mainly in this area that ATR has been used for dairy product analysis. Cultured products (Steiner *et al.*, 1992; Nikolajsen, 1998), butter (van de Voort *et al.*, 1992a), cream (Kemsley *et al.*, 1994), sweetened condensed milk (Nathier-Dufour *et al.*, 1995) and cheese (McQueen *et al.*, 1995) have been measured using ATR sampling.

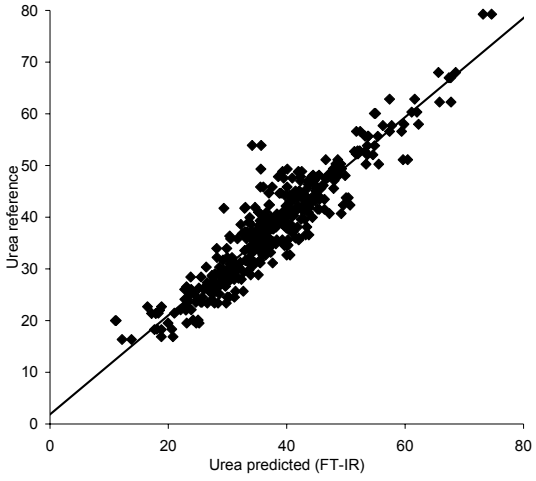


Figure 3.1 Transmission experiments: Reference vs. predicted urea content of the 196 independent test samples using the 16-factor PLS model based on the raw spectra. All replicates of all samples are shown. Accuracy (RMSEP) 3.9 mg/dl, repeatability (s_r) 1.3 mg/dl.

The general trend in these studies is the difficulty of performing reproducible measurements, either because of the heterogeneous nature of dairy products, or because of the tendency of fat to adhere to the crystal surface. The first problem might be difficult to solve, while the latter can be solved by a thorough cleaning of the crystal between measurements.

Another advantage of ATR is that it allows measurement of samples usually regarded as opaque in mid-IR terms. This is because the effective pathlength is less than what can be obtained by transmission spectroscopy. This pathlength is wavenumber dependent, and ranges from approx. 5 to 20 μm in the case of milk situated on a diamond crystal. In a milk context this means that it becomes possible to use information located under the water band at approx. 1650 cm^{-1} (O-H bend). This is of particular interest to a urea determination, as the strongest urea band (C=O stretch) is located at approx. 1640 cm^{-1}, and thus obscured by water in transmission experiments.

The ATR work is described in Publication II (Nikolajsen and Hansen, 1998, not previously published). The experiments were carried out in a fashion which was useful for indicating the possibilities of the method, so the actual accuracies found might be improved by technical means. A brief summary of the urea results is given below.

3.3.1 Description of the data set

The data set consisted of 105 samples, of which 29 were fortified with urea to ensure a good range of urea concentrations. Eight outliers were removed, and the remaining samples were divided into a calibration set (65 samples), used for calibration and cross validation, and an independent test set (32 samples) used for evaluation of the final model.

Only chemically meaningful ranges of the mid-IR spectrum (approx. 1000 to 3000 cm^{-1} with water and diamond bands removed) were used for calibration.

3.3.2 Result summary

Corresponding to the transmission experiments, various pre-processing methodologies (direct orthogonalisation (DO) (Andersson, 1998a), derivatives, standard normal variate (SNV) transformation, and use of a reference wavelength) were tested. In addition, two different variable selection methodologies: principal variables (Höskuldsson, 1992) and

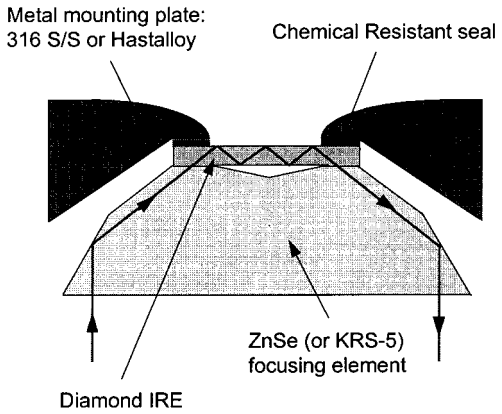


Figure 3.2 The ATR set-up used for the experiments. IR light is guided towards a diamond crystal inside which it is subject to multiple reflections (in the present case 18). The reflected light is collected and analysed using an ordinary FT-IR spectrometer. Reproduced from *The DurasamplIR: User's guide to effective sample handling*.

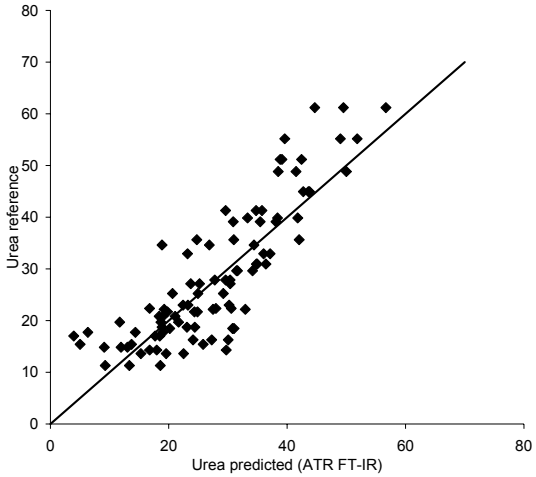


Figure 3.3 ATR experiments: Reference vs. predicted urea content of the 32 independent test samples using the 13-factor PLS model based on the raw spectra. All replicates of all samples are shown. Accuracy (RMSEP) 6.9 mg/dl, repeatability (s_r) 5.3 mg/dl.

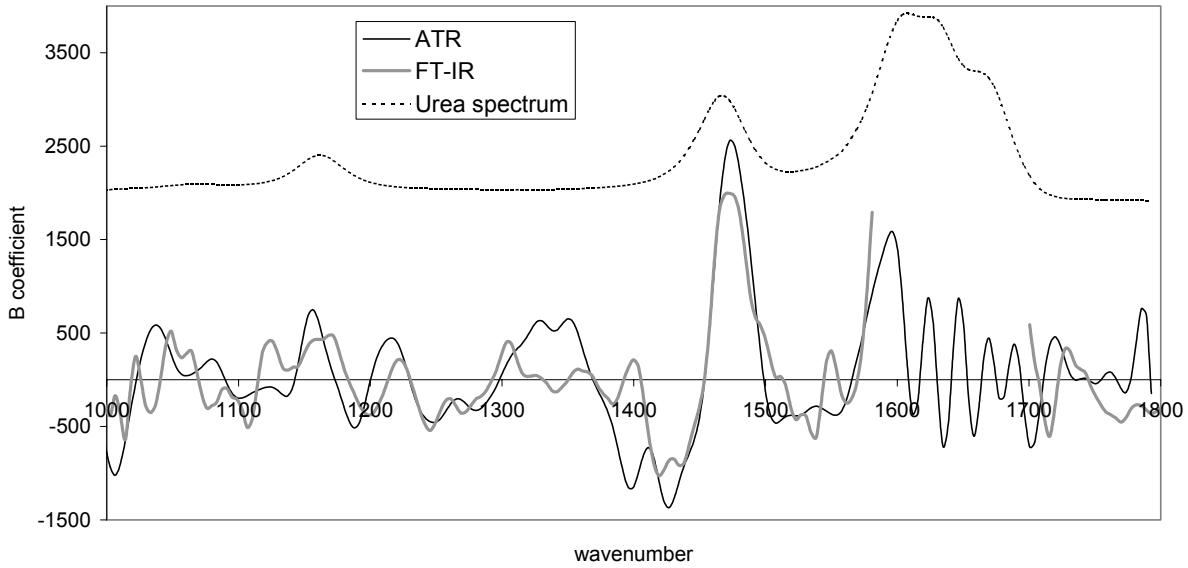


Figure 3.4 B coefficients from urea calibrations based on ATR (13 factors, no pre-processing) and transmission data (16 factors, no pre-processing). The ATR FT-IR spectrum of pure urea in water is overlaid.

the so-called "B filter method" (Andersson, 1998b), both combined with genetic algorithms (Leardi *et al.*, 1992)) were tried.

No improvement was obtained, neither by pre-processing nor by variable selection. The result on the independent test set was:

R^2	0.725
RMSEP	6.9 mg/dl
s_r	5.3 mg/dl

by use of 13 PLS factors and without any pre-processing. The reference vs. predicted plot is shown in Figure 3.3.

Compared to the result using a transmission cuvette this is poor. There are, however, some good indications:

- The repeatability (s_r) is very close to the accuracy (RMSEP). This means that the accuracy is mainly a result of a badly repeating experiment. This is usually a problem which can be solved by tightening the specifications for temperature control, humidity, cleaning and other external factors affecting the experiment. But *if* the repeatability problems are caused by the heterogeneous nature of milk, then they might be difficult to solve.

- The B coefficient spectrum (Figure 3.4) is similar to its counterpart obtained by transmission spectroscopy. The B coefficients peak at the exact positions of the major absorption bands of urea, meaning that urea *is* determined by the calibration. Very importantly: Information is being gathered under the water band at 1650 cm^{-1}. This means that *when* the effective pathlength is low, additional information from the strongest urea absorption bands can be used.

3.4 CONCLUSIONS

At present, transmission spectroscopy is the best method for urea determination. An accuracy (RMSEP) of 3.9 mg/dl was obtained.

Improved results due to pre-processing were not observed in the present studies.

Use of the ATR sampling technique did not yield better results than when using transmission spectroscopy. However, transmission experiments with shorter pathlengths, or an optimised ATR set-up *could* result in an even better FT-IR based urea determination. Furthermore, the range around 1650 cm^{-1}, where water is absorbing, could be used for determination of other milk components absorbing in this region (*e.g.* acetone or proteins).

Chapter 4

Acetone Determination in Milk

4.1 BACKGROUND

Contrary to the other milk components studied under the present project, acetone is usually not found in milk samples. It is only when a cow suffers from ketosis that it is observed in milk in significant amounts.

Ketosis is a metabolic disease found in high-yielding dairy cattle as a result of low feed intake compared to the amount of milk produced. The disease has its highest frequency between the third and sixth week of the lactation, and results in lowered milk yield (Gustafsson *et al.*, 1993) and lower fertility (Andersson *et al.*, 1991). The disease can be avoided by providing the cow with additional feed, and therefore it is of interest to detect both clinical and sub-clinical incidences of ketosis.

The compounds related to ketosis are called ketone bodies. Acetone is not the only ketone body: β-hydroxybutyrate and acetoacetate are also produced in significant amounts. They are released to the blood, from which they are transferred to milk, thus there exists a good correlation between the amount acetone in the cow's milk and the amount of ketone bodies in its blood (Steger *et. al*, 1972). Acetone is the most abundant of the three compounds in milk: it accounts for almost 70 % of the ketone bodies (Andersson, 1984), and is therefore likely to be the most detectable ketosis marker present in milk.

When a cow suffers from ketosis its milk acetone content rises to levels up to 2 mM, in some cases even higher. Gustafsson and Emanuelson (1996) have proposed the following classification limits for cows, in terms of the acetone content of their milk:

< 0.7 mM	The cow is healthy
0.7–1.4 mM	The cow might suffer from ketosis
> 1.4 mM	The cow suffers from ketosis

The milk yield was used as the symptom indicating ketosis.

The present state-of-the-art laboratory instrumentation for measurement of acetone in blood or milk is based on the FIA principle (Marstorp *et al.*, 1983). As with urea, this method is relatively slow compared to compositional measurements based on mid-IR spectroscopy. Therefore a faster method would be of interest.

The purpose of the present study was to see whether an FT-IR based acetone determination was possible, and – if so – to estimate the attainable accuracy. This was done by use of transmission experiments, as initial ATR studies gave negative results (Nikolajsen and Hansen, 1998 (Publication II)). The work has been reported in Publication III (Hansen, 1998c) and, as it was successful, patented (Hansen, 1998b (Publication IV)).

Furthermore, the special properties of the milk acetone data promoted the development of a pre-processing method: Independent Interference Reduction (IIR). This method is described in Publication V (Hansen, 1998d).

The following is a summary of the calibration work as well as a description of the IIR method.

4.2 ACETONE AND FT-IR

The experiments were carried out in order to estimate the accuracy and classification power of the transmission FT-IR method when used for acetone determination. As natural samples are difficult to get hold of, it was also tested whether fortified samples could be used for calibration. A detailed description is found in Publication III (Hansen, 1998c).

4.2.1 Description of the data set

Two calibration sets were measured on an FT-IR instrument: One consisting of 302 purely natural individual cow samples collected in Norway, Sweden and Denmark. The other set consisted of 125 individual cow samples of which some were fortified with acetone at two levels.

For independent testing of the calibrations, 58 samples from individual cows, collected and measured in New Zealand, were used. All samples were carefully selected to ensure a wide range in the acetone content (*e.g.* 0.0 to 2.8 mM for the independent test set).

4.2.2 Result summary

Full spectrum PLS calibrations were carried out using the natural and fortified samples, respectively. In addition, a calibration based on the combination of the two sets was performed. Cross validation was used for selecting the optimal number of PLS factors.

The models were tested using the independent test set. These results are shown in Table 4.1. Only when natural samples were used in the calibration step the accuracy on the test set was similar to that obtained on the calibration samples. Thus, it can be concluded that only natural samples can be used for calibration. The reference vs. predicted plot for the independent test samples using this calibration set is shown in Figure 4.1.

The number of PLS factors is very high: 21 for the model yielding the best result on the independent test set. This could suggest an overfit. However, as the test samples were collected and measured (using another instrument), in an area very different from the origin of the calibration samples, the model seems to be robust. Furthermore, the obtained B coefficients (Figure 4.2) support the assumption that acetone is actually measured by the PLS calibration, as they peak at the major absorptions of acetone.

The classification power of the FT-IR predictions was tested using the results from the independent test set. The accuracy was not good enough for the method to provide a useful classification between the three groups stated above. Therefore, a classification into the groups "well" and "possibly ill" was tried. When, for example, an FT-IR classification limit of 0.5 mM was set up, 53 % of the samples selected were false positives (*i.e.* having a true acetone content of 0.7 mM or less). On the other hand, no false negatives were found using the present data set. 53 % may seem high, but when the number of cows suffering from ketosis at a given time (estimated to be about 1 %) is taken into account, such an error level is acceptable.

Thus, FT-IR can provide an acetone determination in milk with an accuracy better than 0.3 mM. This accuracy enables the screening

Calibration samples	[1]#fac	R^2	RMSEP	s_r
Natural	21	0.81	0.27	0.12
Fortified	20	0.42	0.57	0.18
All	20	0.78	0.32	0.15

Table 4.1 Results of PLS calibration of acetone (in mM units) on the independent test set containing 58 samples from New Zealand.
[1]#fac = optimal number op PLS factors. It was selected using cross validation (with various numbers of cross validation segments) on the calibration set.

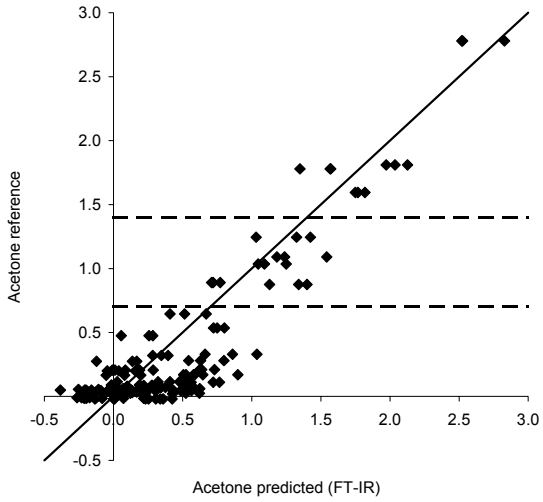


Figure 4.1 Result of the prediction of acetone in the independent test set (58 samples) using a 21-factor PLS model based on natural samples. The horizontal lines indicate the proposed classification limits at 0.7 and 1.4 mM (Gustafsson and Emanuelson, 1996). Correlation (R^2) 0.81, accuracy (RMSEP) 0.27 mM, and repeatability (s_r) 0.12 mM.

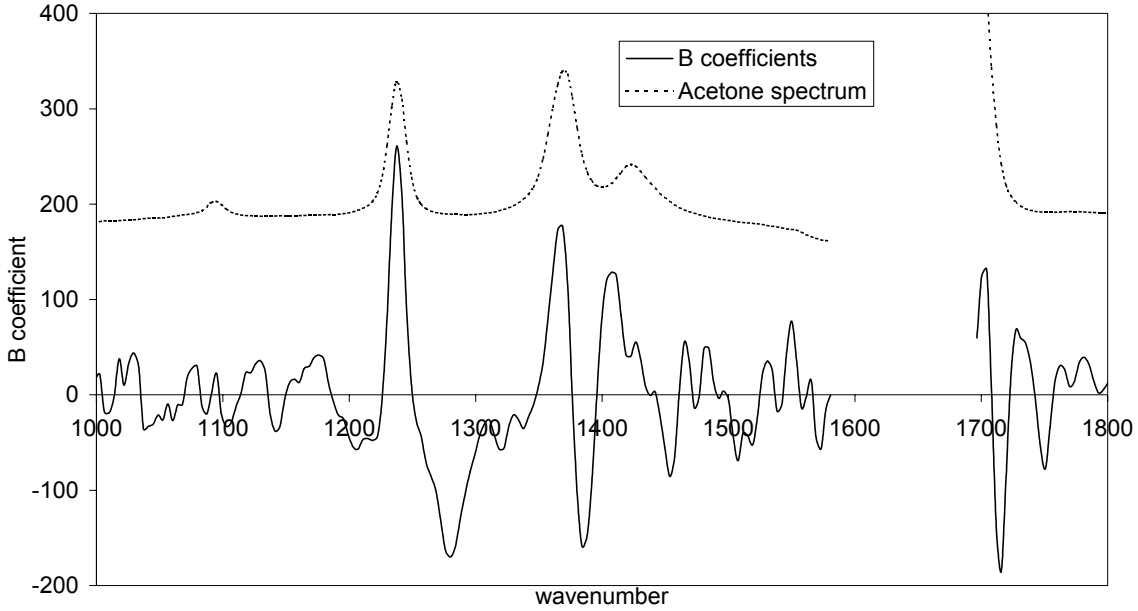


Figure 4.2 B coefficients from the 21-factor PLS model for acetone in milk based on natural samples. The spectrum of pure acetone in water is overlaid.

of samples originating from possibly ketotic cows.

4.3 INDEPENDENT INTERFERENCE REDUCTION (IIR)

There are two problems in relation to an FT-IR based acetone determination for milk:

- PLS models predicting acetone in milk use a high number of factors (*i.e.* they are very complex)
- natural acetone-containing samples are difficult to get hold of

These problems are incompatible, as more complex PLS models require more calibration samples.

Theoretically, it is possible to collect the number of samples necessary to build a 21-factor PLS model. But as acetone containing samples are rare, it seems a waste of expensive samples to use them for extracting 21 factors, of which maybe ten account for variations in the major milk components.

In the present case, special knowledge can be used for reducing the number of reference-analysed samples needed for calibration, *i.e.* the fact that samples known not to contain acetone are easy to obtain. This has led to the proposal of the IIR method, which can be useful in such cases. A detailed description of the method is given in Publication V (Hansen, 1998d). A brief summary will be given below.

4.3.1 Description of the method

The method relies on an assumption very similar to PCA data pre-treatment (Sun, 1997), DO (Andersson, 1998a), and Orthogonal Signal Correction, OSC (Wold *et al.*, 1998), *i.e.* that scatter effects and interferences can be modelled by chemometric means and removed from the data matrix prior to calibration. DO and OSC, however, require reference results in order to be able to perform this pre-processing. PCA data pre-treatment is only used for modelling of instrumental effects. IIR uses a set of samples in which the content of the component of interest is constant (or low) – and no reference results are required in the pre-processing step.

The algorithm is as follows:

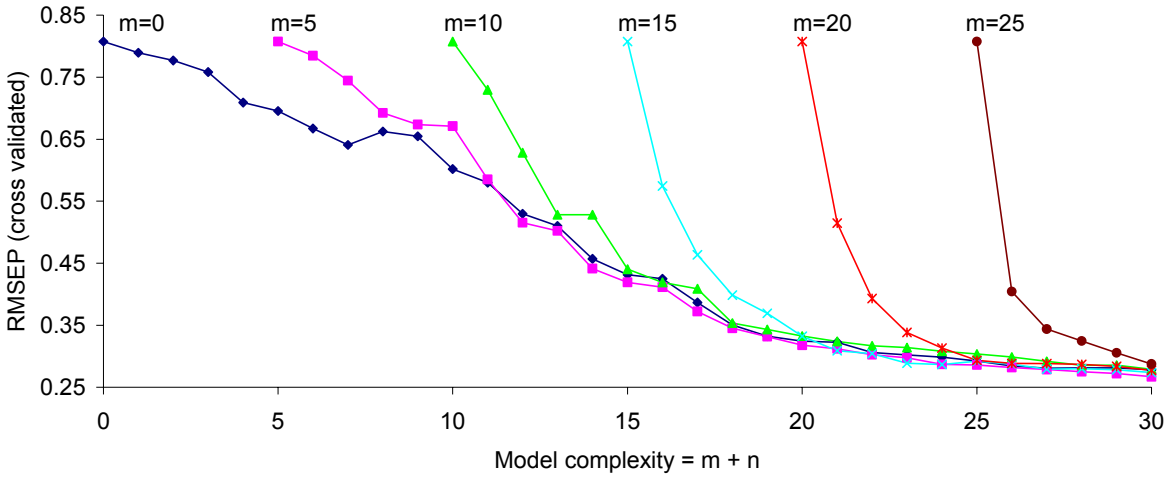


Figure 4.3 Cross validation error for acetone in milk (calculated using 10 cross validation segments) against model complexity for various dimensions (m) of the initial PCA in IIR.

1. Assume that a data matrix $\mathbf{X}_{simple}$ can be obtained, which contains typical sample variation in all parameters, apart from the component of interest. This matrix is modelled using PCA with m principal components. If $\mathbf{X}_{simple}$ is large enough this provides a number of well-defined loadings.

2. Another matrix, $\mathbf{X}_{special}$, containing a wide variation in the component of interest, is projected onto the loadings obtained in step 1. This results in m sample scores. This model (based on m scores and loadings) of the interference part of $\mathbf{X}_{special}$ is subtracted from it, yielding a new matrix $\mathbf{X}_{special\text{-}m}$.

3. This reduced matrix, $\mathbf{X}_{special\text{-}m}$, is used with its associated reference results, $\mathbf{Y}_{special}$, to generate a calibration model, *e.g.* using PLS with n factors.

The overall result is that the first loadings are extracted from a "cheap" data set with no associated reference results. The loadings describing the component of interest are extracted from an "expensive" data set, which may contain fewer samples than otherwise necessary.

4.3.2 Application to FT-IR

The method was tested using the calibration set based on 302 natural samples from individual cows described above (section 4.2.1). The samples were divided into two sets: one containing 198 samples with an acetone content below 0.1 mM ($\mathbf{X}_{simple}$) and another containing 104 samples with variation in the acetone content ($\mathbf{X}_{special}$). The independent test set (58 samples) was used without modifications.

IIR with the parameter m ranging from 0 to 25 was performed, followed by a PLS regression using n factors, where the sum of m and n varied from m+1 to 30. Cross validation on the calibration set was performed to find the optimal values of m and n. The change of the cross validation error (RMSEP at ten cross validation segments) as a function of m and n is shown in Figure 4.3.

The error stabilises at an overall model complexity (m+n) of approx. 25, irrespective of the value of m. Thus, use of IIR does not result in any reduction of the overall model complexity. The main advantage is therefore the reduced need for reference analyses and the fact that the model is divided into two parts describing the interferences and acetone, respectively.

When the final models were tested using the independent test set, another advantage emerged. This can be seen in Table 4.2: For

m	m+n	R^2	RMSEP	SEP	s_r
0	25	0.827	0.60	0.28	0.15
5	25	0.858	0.54	0.26	0.15
10	25	0.803	0.59	0.29	0.16
15	24	0.774	0.34	0.31	0.16
20	25	0.779	0.31	0.29	0.17
25	30	0.846	0.42	0.25	0.16

Table 4.2 IIR results for acetone on the independent test set (58 samples) using various combinations of the dimensions m and n.
m = dimension of the PCA used for pre-processing.
n = dimension of the PLS regression.

low values of m the RMSEP is high, but both R^2 and SEP are on the same level as during calibration. Thus, the predictions contain a bias when m is low. This suggests that without IIR pre-processing the models tend to be less robust to general changes in the samples, such as sample matrix or light scatter effects. When m equals 20 the RMSEP is close to the SEP, so this value of m would be a sensible choice for the optimal model. When m rises above 20, the overall model complexity increases accordingly, and the RMSEP (but not the SEP) starts to increase, supporting that m equal to 20 is the right choice.

In the present example, IIR pre-processing results in more robust and interpretable PLS models. This is because only relevant acetone information is present in the pre-processed matrix when PLS modelling is commenced. For example, with an appropriate choice of m, the first PLS loading contain features similar to the spectrum of pure acetone. IIR does not yield significant improvements in accuracy or repeatability.

An important perspective of the method is the fact that robustness is gained: Consider a large number of samples measured at various times at various locations on different instruments. If this set of samples is regarded as a general $\mathbf{X}_{simple}$, then its PCA loadings can be used for IIR pre-processing in future cases where a component present in low concentrations is to be measured. This would result in globally applicable calibrations based on relatively few new calibration samples.

4.4 CONCLUSIONS

The present studies show that an FT-IR based acetone determination is possible. Only natural samples could be used for calibration. The calibration based on Scandinavian samples could be transferred to New Zealand without any loss in accuracy. The accuracy (RMSEP) was determined to be 0.27 mM.

The obtained accuracy allowed for classification of the cows into the groups "well" and "possibly ill". The number of misclassifications depended on the classification limit used.

IIR pre-processing of the data results in more robust acetone predictions, as well as more interpretable calibration models. The method might be extended to other components present in low concentrations.

Chapter 5

Dairy Batch Process Analysis

5.1 BACKGROUND

Batch processes are carried out in all branches of the dairy industry. Examples are: production of cultured products and hydrolysis reactions, whereby chemical compounds are consumed and others are produced. If specific product properties are required it is a valuable tool to be able to estimate the concentrations of such compounds. FT-IR is an obvious choice for this task, as it detects almost any organic compound which can be produced. The use of FT-IR for batch process monitoring is therefore the subject of the present chapter. Lactose hydrolysis is used as an example, demonstrating the methods that can be used.

5.1.1 Lactose hydrolysis

When dairy products are ingested, lactose is cleaved into the monosaccharides galactose and glucose to make it digestible. Large racial groups suffer from an inability to perform this cleavage (the disease called lactose intolerance), which means that they cannot digest dairy products unless lactose is hydrolysed during production (Bayless and Rosensweig, 1966). In addition, cleavage of lactose results in improved properties of milk when used for production of ice creams and cultured products (Gyuricsek and Thompson, 1976).

The lactose hydrolysis process used in dairy production is the enzymatic cleavage of lactose performed by β-galactosidase (Figure 5.1). This process is carried out in batches. The course of the reaction is highly dependent on the initial conditions, such as temperature and enzyme concentration (Guy and Bingham, 1978; Forsman *et al.*, 1979). Therefore, it is of interest to follow the reaction in order to terminate the process as soon as the desired degree of hydrolysis is reached. As the process is completed within a few hours, a fast method is required. FT-IR spectroscopy is a good choice, since individual sugars can be determined using this method combined with *e.g.* PLS regression (Ridder and Kjær, 1995).

Depending on the conditions of the hydrolysis, oligosaccharides (*i.e.* tri- and tetrasaccharides) are produced during the process as a result of a side-reaction (Prenosil *et al.*, 1987a and 1987b). Their total concentration increases in the beginning of the process when the lactose concentration is high, and decreases in the later stages, accompanied by the production of galactose and glucose. It is of interest to identify these sugars and to predict their concentrations during the reaction, but they are difficult to estimate by a reference method. Thus, if FT-IR could provide such information, it would be of value to the dairy engineer.

5.1.2 Alternating Least Squares (ALS)

PLS regression is the usual tool when spectral data are correlated to reference analyses. PLS requires knowledge of the reference results, and in order to obtain a good calibration the reference data must be of high quality. In most cases this is no problem, but in the case of processes where an intermediate compound might only exist for a limited period of time (*i.e.* minutes), it can be difficult to terminate the process and to analyse a sample using the

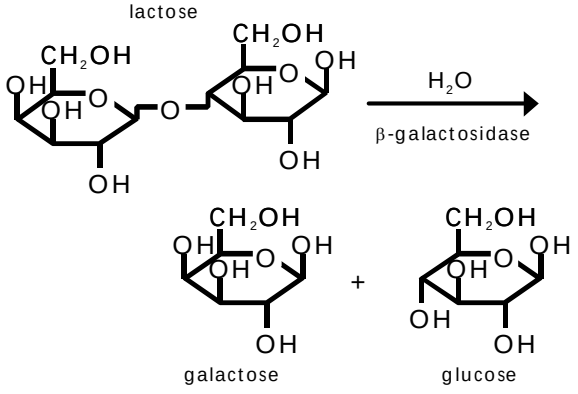


Figure 5.1 Lactose hydrolysis in milk as a result of the enzymatic action of β-galactosidase.

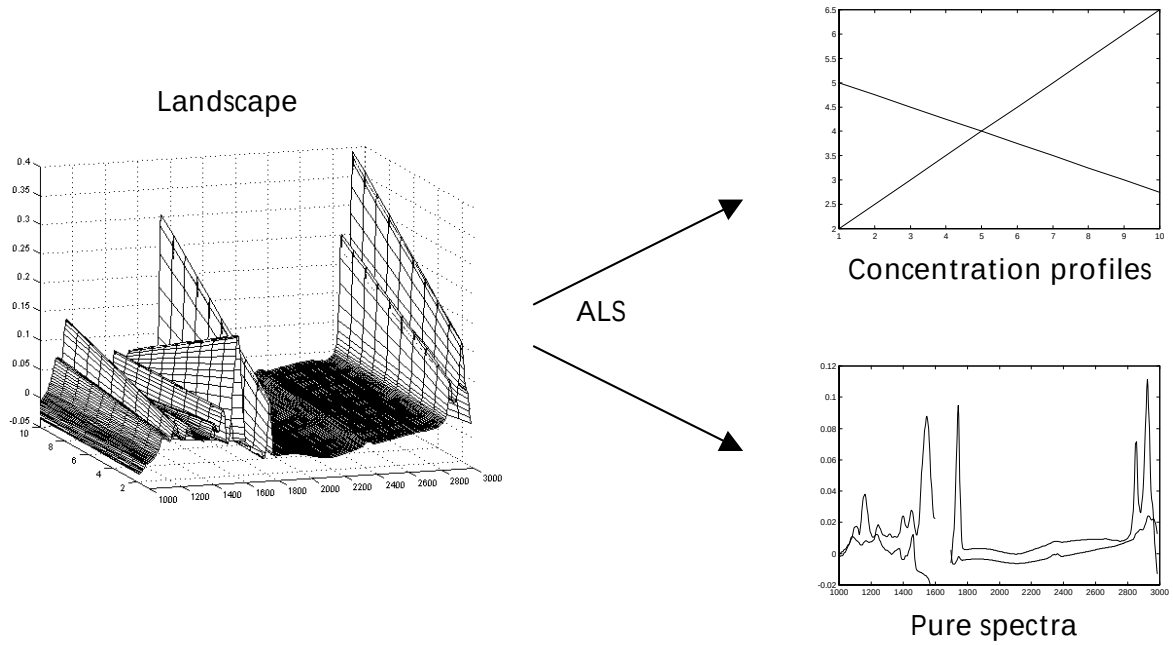


Figure 5.2 Illustration of the principle behind ALS: A landscape consisting of spectra measured *e.g.* at various times during a chemical process is resolved into the pure spectra and concentration profiles of which it consists. In the example, two chemical components (fat and protein) were changing during the process, and when ALS is used relative concentrations and pure spectra (in arbitrary units) of fat (concentration increasing) and protein (concentration decreasing) are obtained without the need for reference analysed samples.

reference method. Furthermore, it might only be of interest to know the relative concentration changes during the process, and an absolute reference result is therefore not necessary if the same information could be obtained by other means.

Lactose hydrolysis is an example of such a situation. It is usually the degree of conversion (not the actual lactose concentration), which is of interest. At the same time, the oligosaccharides are of interest, and, as mentioned above, their abundance is difficult to estimate using a reference method.

Alternating Least Squares (ALS), a special case of Alternating Regression (AR) (Karjalainen, 1989), provides a means for solving this problem. ALS is useful for handling second order data (see section 2.1.2), *e.g.* a landscape obtained as spectral measurements from a process collected at various times. The idea behind ALS is illustrated in Figure 5.2. If the Beer-Lambert law is obeyed perfectly, a landscape, in which the concentrations of the components change over time, will be generated by the pure spectra of the signal-giving components. In Figure 5.2, two signal-giving components (fat and protein) vary over time independently of each other. Therefore two pure spectra and two concentration profiles, both in arbitrary units, can be resolved iteratively by ALS, and without the need for any reference analyses.

The pure spectra obtained in this way can be used for prediction of fat and protein in future samples in a way similar to CLS (see section 1.1.2). A detailed mathematical description of ALS can be found in Publication VI (Hansen *et al.*, 1998).

5.2 ALS AND FT-IR DATA

ALS performed on FT-IR data has potentially useful applications: FT-IR spectra contain very detailed information on the chemical compounds in a mixture. They can therefore provide information on compounds for which no reference method exists. Even when a reference method is known it might be preferable not to use it, as it is time and resource demanding. Other attempts to use

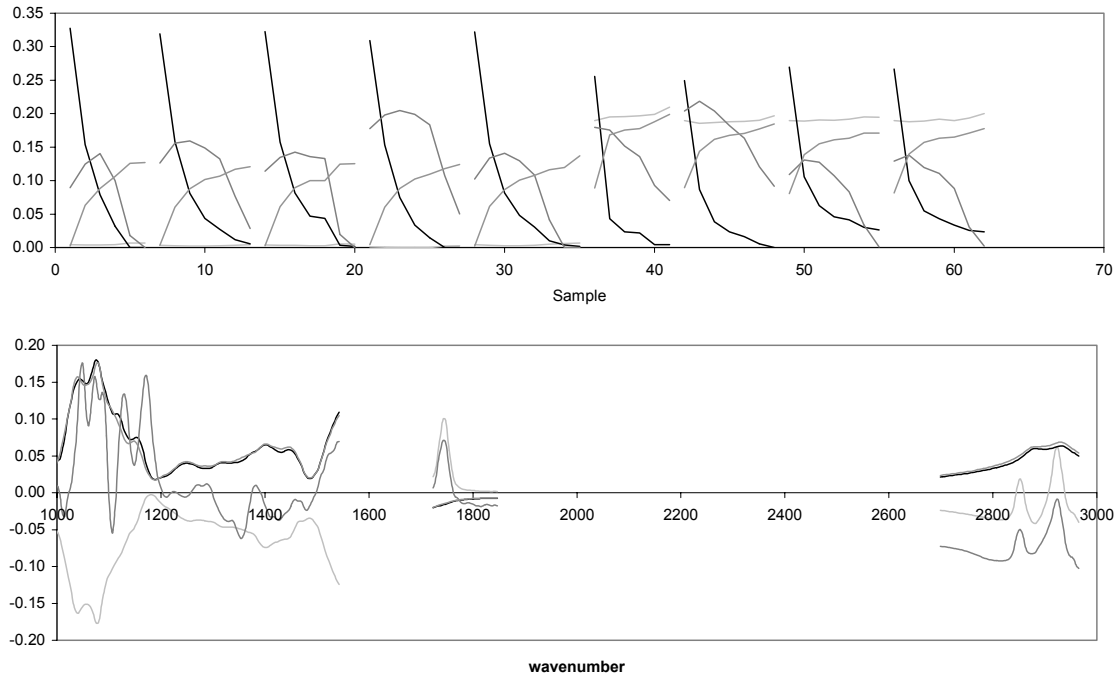


Figure 5.3 ALS resolved pure spectra and concentration profiles for the lactose hydrolysis process when four components are extracted from the spectral landscape of calibration set 2. The present solution was selected as the best (*i.e.* having the lowest SSE) out of 100 runs. The upper part shows the normalised concentration profiles for the nine process runs in sequence. The first five runs were based on skim milk, resulting in zero fat concentrations, while the last four runs were based on whole milk. The lower part shows the resolved pure spectra (normalised).

ALS on IR data have been very successful (Tauler *et al.*, 1993; Furusjö *et al.*, 1998).

For this reason, a series of experiments where ALS was applied to FT-IR spectra from an instrument monitoring a lactose hydrolysis process were carried out, in order to assess the possibilities for dairy batch process analysis. The results of this work have been reported in Publication VI (Hansen *et al.*, 1998) and patented (Hansen, 1998e (Publication VII)). A summary of the results will be given below.

5.2.1 Description of the data set

Two similar sets of calibration samples, each consisting of 62 samples, were obtained from nine process runs carried out using a laboratory set-up. The sets were similar, but not identical, as two samples were collected at every sampling with a short time delay. The two samples were handled independently in the subsequent measurements. 90 of these 124 samples had their lactose content determined by HPLC. The test set consisted of 23 samples collected from two process runs carried out six months later. HPLC references for lactose were determined on these samples, as well.

All samples were measured on an FT-IR instrument, and full spectra (apart from water bands and spectroscopically silent regions) were used in the calculations.

5.2.2 Result summary

ALS regressions were carried out on the two calibration sets assuming from three to six chemical components in the spectral landscapes. The concentration profiles were restricted to be *unimodal* (*i.e.* having only one maximum) and *non-negative*. In this way the possible solutions were limited to physically meaningful ones. The optimal models were selected by using the model error (in terms of the sum of squared errors (SSE)) or the correlation (R^2) of the lactose profiles to the lactose results as a selection criterion. Four

	Calibration	Test		
	R^2	R^2	[3]SEC	s_r
ALS, set 1 4 components	[1]0.942	[1]0.959	2.16	0.16
ALS, set 2 4 components	[1]0.971	[1]0.980	1.51	0.27
PLS 5 factors	[2]0.996	0.987	[4]1.55	0.23

Table 5.1 ALS and PLS results for dry base lactose (range 0 to 45 %) on the calibration set (set 1 or set 2, each containing 62 samples) and the test set (23 samples from two process runs) when predictions corresponding to lactose are correlated to reference data. The optimal ALS models were selected by use of the SSE (*i.e.* independently of the reference results).
[1]The correlation was obtained when comparing the profile corresponding to the pure lactose spectrum to the reference results.
[2]The correlation was obtained by using cross validation with ten cross validation segments on the calibration samples from set 1.
[3]SEC calculated following a univariate linear regression between the lactose profile and the reference results.
[4]For PLS the result stated is in terms of an SEP, as a substantial bias was present. A univariate linear regression was not performed in this case.

assumed chemical components gave the optimal models, both when SSE and R^2 were used. Thus, selection by use of the SSE is possible, which means that reference results are not needed for selecting the optimal model. The resolved pure spectra and concentration profiles for calibration set 2 are shown in Figure 5.3.

A PLS regression of FT-IR spectra against lactose reference data was carried out for comparative purposes. Five PLS factors were found to be optimal.

The optimal models, both from ALS and PLS, were used on the test set. The results for lactose in these tests, as well as the results during calibration are shown in Table 5.1. The result, when the lactose profiles resolved by ALS are correlated to the actual reference measurements, is very similar to the PLS result. Thus, in the present case, ALS is able to extract the same information as PLS from the data, but, very importantly, without using any reference results.

The normalised predictions from ALS and PLS in one of the two test runs are shown in Figure 5.4. Of the four components, one could be assigned to fat, while another was assigned to lactose. As fat does not vary within a single process run it has been omitted from Figure 5.4 for simplicity. The ALS and PLS lactose profiles have the same shapes as the profile based on the reference data, but the decrease of the latter is less smooth than what should be expected from a chemical reaction. Thus, FT-IR seems to be sensing the "true" decay of lactose better than the reference method.

The two remaining components in Figure 5.4 (denoted Component 3 and 4) are not known, but their spectra are similar to sugars (Figure 5.3). Component 4 shows a monotonic increase similar to what should be expected as a result of the build-up of the products: galactose and glucose (Greenberg and Mahoney, 1983). They cannot be resolved, probably because their profiles are strongly correlated. The profile for Component 3 peaks after approx. 40 minutes, followed by a monotonic decrease. This is the shape that would be expected if oligosaccharides are produced and consumed (Greenberg and Mahoney, 1983). Therefore the resolved profiles from ALS seem to detect this group of compounds which are difficult to determine by a reference method. However, more experiments, focusing on the oligosaccharides, are needed in order to confirm this.

Protein, known to be present in the samples, was not resolved from the spectra, as it did not vary between calibration samples. If the pure spectra are to be used for future prediction, a protein variation should be included in the data set in order to obtain a robust predictive model.

5.3 CONCLUSIONS

When samples from dairy processes are analysed over time using spectroscopic methods, new possibilities are emerging.

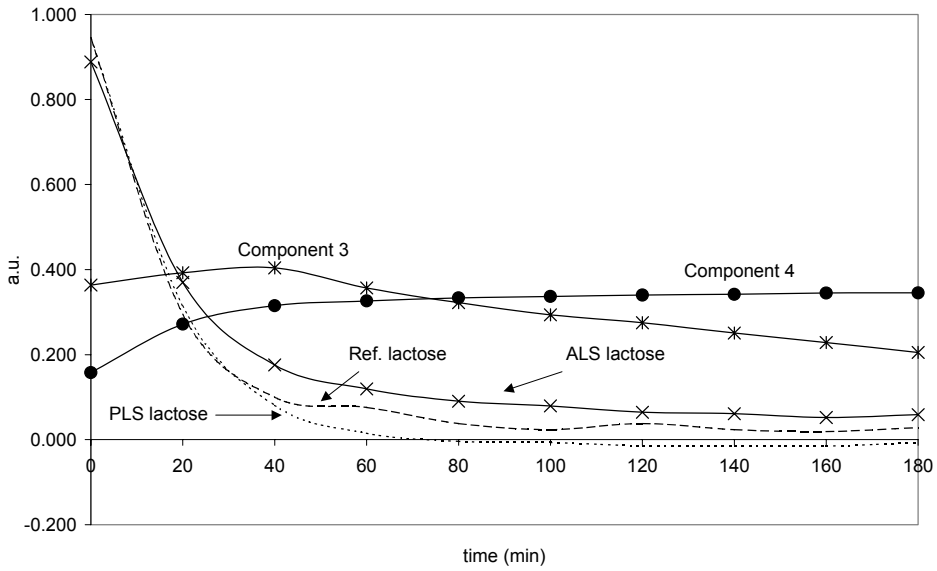


Figure 5.4 Concentration profiles for one of the two test runs. The fat profile is omitted, as it was constant during the process run. For lactose the corresponding profiles based on PLS predictions as well as the reference results have been added for comparison. All profiles have been normalised in order to be presented on the same scale.

In the present work, ALS made it possible for FT-IR to predict the constituents of a sample without using reference analysed samples for calibration. In the lactose hydrolysis process, four components could be extracted from the spectral data: fat, lactose, and two other components likely to be the sum of galactose and glucose, and a group of oligosaccharides. The prediction of lactose obtained by ALS was almost as good as the corresponding result from an ordinary PLS calibration.

ALS performed on FT-IR data is not limited to the present application. It is applicable wherever chemical reactions are used in the industry. Furthermore, second (or higher) order data might be generated in the laboratory, *e.g.* by changing the pH or temperature of a sample. By using ALS or another multivariate curve resolution method, such data can be used for detecting chemical compounds affected by the change, *i.e.* also intermediates – without using resource-demanding reference methods.

Chapter 6
Conclusions

The conclusions fall within two categories: the spectroscopic methods, and those related to the chemometric part.

6.1 SPECTROSCOPY

6.1.1 Spectroscopic methods in general

From the screening experiments involving NIR, Raman, UV/VIS, fluorescence and NMR spectroscopy, it was concluded that for applications involving liquid milk, mid-IR spectroscopy (using the FT-IR technique) is superior. This is not surprising, as mid-IR has been optimised for such use during the past two decades.

However, one method is promising with regard to its combination of easy sample handling and useful spectroscopic information. This method is low resolution ^{1}H-NMR spectroscopy. It is able to handle relatively voluminous samples with centimetre-scale diameters. This opens up for measuring dairy samples (powders, pastes, and other viscous products), considered "difficult" in mid-IR terms, with minimal sample preparation. Furthermore, the possibilities of "tuning" the NMR experiment towards measuring specific constituents are promising, as it allows for focusing on one (or a few) component(s) in the sample.

6.1.2 Mid-IR spectroscopy

Returning to mid-IR spectroscopy, this method is relatively sensitive when sample handling is no problem, *i.e.* when liquid milk is measured. When the FT-IR technique is used, the present work shows that it can be applied for measuring urea and acetone in milk from individual cows. The accuracies are on levels which make the FT-IR method useful as a screening tool. Both compounds occur in milk at very low levels (urea at approx. 30 mg/dl, acetone at approx. 6 mg/dl (1 mM)), thus showing the sensivity of the method.

In an attempt to ease sample handling for FT-IR, the ATR technique was tested. The result was poor when compared to the traditional method using a transmission cuvette. There were, however, some good indications showing that a decreased pathlength (*e.g.* 10 to 20 μm) through the sample allows for use of the range under the water band situated around 1650 cm^{-1}. This might improve *e.g.* the urea, acetone and protein determinations. Further experiments will be required to confirm this.

A general trend in all experiments where FT-IR was used, was that when a sufficient variation is present in the calibration samples there was no improvement in terms of accuracy when the spectra were pre-treated prior to calibration. The variation was obtained by measuring many samples (*i.e.* hundreds). It is therefore concluded that pre-treatment not necessarily leads to improved predictive models when used on FT-IR data. However, it cannot be denied that pre-processing of FT-IR data can be useful in cases where only few samples can be obtained.

Some of these experimental findings have been included into the basis for FT-IR instruments presently being developed by Foss A/S.

6.2 CHEMOMETRICS

In most cases, application of PLS regression for obtaining correlations between spectra and reference measurements yields satisfactory results. This is the case for urea and acetone in milk. In some cases, however, the data can contain properties allowing for higher order or other special treatments.

Fluorescence spectroscopy on undiluted milk is one such case. By use of PARAFAC it was possible to identify three components in the

spectral landscapes obtained by scanning over both the excitation and the emission mode.

Mid-IR spectroscopy does not contain the same fundamental multi-dimensionality as fluorescence, but a sensible experimental design can result in data which can be treated by multi-dimensional methods. In the present case, a batch process (lactose hydrolysis) was studied by use of ALS, and constituents for which no reference results were known could be resolved. This is a most promising result, since many different experiments of this type could be carried out, allowing for relative concentration predictions without the need for reference analyses in the calibration step.

For more information on the state-of-the-art of higher order methods the reader is referred to the PhD thesis by Bro (1998).

Another approach reducing the need for reference analyses is to pre-treat the calibration data in a way which removes interferences (not only scatter effects) before the actual calibration is performed. As a result of the acetone work performed here, the IIR method was proposed. Using IIR, it was possible to build robust predictive models based on a reduced set of reference analysed samples.

6.3 FUTURE TRENDS

Following the discussion given above, the author sees the following future trends for chemometrics and spectroscopy on milk and dairy products:

- FT-IR is, for the nearest future, the best spectroscopic method for applications involving liquid milk

Although mid-IR is a relatively old method, it still has potentials when full spectra generated by FT-IR are used. It is mainly in the area of data treatments that new accomplishments should be expected.

- Low resolution ^{1}H-NMR is a promising method for other sample types

NMR might not allow for measurements of components present in low concentrations, but it is likely that it can be adjusted to focus on specific proteins or fat properties.

- The generation of higher order data combined with the use of higher order chemometric tools will become of increasing importance

Especially in combination with FT-IR data that contains a wealth of information, multivariate curve resolution methods could be useful. This might include process monitoring or the development of hyphenated methods, such as HPLC or FIA combined with FT-IR.

References

Adar, F., Geiger, R., and Noonan, J. 1997. Raman spectroscopy for process/quality control. Appl. Spectrosc. Rev. 32:45-101.

Andersen, H. V., Kjær, L., Hansen, P. W., and Ridder, C., inventors. 1996. A method for standardizing a spectrometer. Foss Electric A/S, assignee. Int. Pat. Appl. No. WO 96/24832.

Andersson, C. A. 1998a. Direct orthogonalization. Chemomet. Intell. Lab. Sys. (in press).

Andersson, L. 1984. Concentrations of blood and milk ketone bodies, blood isopropanol and plasma glucose in dairy cows in relation to the degree of hyperketonaemia and clinical signs. Zentralbl. Veterinaermed. Reihe A. 31:683-693.

Andersson, L., Gustafsson, A. H., and Emanuelson, U. 1991. Effect of hyperketonaemia and feeding in dairy cows. Theriogenology 36:521-536.

Andersson, M. 1998b. Personal communication. Technical Analytical Chemistry, Lund University, Sweden.

Arnvidarson, B., Nygaard, L., and Hansen, P. W., inventors. 1998. Determination of extraneous water in milk samples, or the freezing point depression of milk samples. Foss Electric A/S, assignee. US Pat. No. 5739034.

Asselain, M., Manifacier, D., and Agnet, Y., inventors. 1996. Method and apparatus for the spectrophotometric assay of aqueous liquids. Foss Electric A/S, assignee. European Pat. No. 588892.

Bayless, T. M., and Rosensweig, N. S. 1966. A racial difference in incidence of lactase deficiency. J. Am. Med. Assoc. 197:968-972.

Bro, R. 1996. Multiway calibration. Multi-linear PLS. J. Chemomet. 10:47-61.

Bro, R. 1997. PARAFAC. Tutorial and applications. Chemomet. Intell. Lab. Sys. 38:149-171.

Bro, R. 1998. Multi-way analysis in the food industry. Models, algorithms and applications. PhD Thesis, University of Amsterdam.

Brosio, E., and Barbieri, R. 1996. Nuclear magnetic resonance in the analysis of dairy products. Rev. Anal. Chem. 15:273-291.

The DuraSamplIR™: User's guide to effective sample handling. 1997. ASI SensIR Technologies, Connecticut.

Ellis, G., Hendra, P. J., Hodges, C. M., Jawhari, T., Jones, C. H., Barazer, P. L., Passingham, C., Royaud, I. A. M., Sánchez-Blázquez, A., and Warnes, G. M. 1989. Routine analytical Fourier transform Raman spectroscopy. Analyst 114: 1061-1066

Fehrmann, A., Franz, M., Hoffmann, A., Rudzik, L., and Wüst, E. 1995. Dairy product analysis: Identification of microorganisms by mid-infrared spectroscopy and determination of constituents by Raman spectroscopy. J. AOAC Int. 78:1537-1542.

Forsman, E., Heikonen, M., Kiviniemi, L., Kreula, M., and Linko, P. 1979. Kinetic investigations of the hydrolysis of milk lactose with soluble Kluyveromyces lactis β-galactosidase. Milchwissenschaft 34:618-621.

Franz, M., Hoffmann, A., Rudzik, L., and Wüst, E. 1996. Determination of milk fat hardness by Raman spectroscopy [in German]. Deutsche Milchwirtschaft 47(1):5-6.

Furusjö, E., Danielsson, L.-G., Könberg, E., Rentsch-Jonas, M., and Skagerberg, B. 1998. Evaluation techniques for two-way data from in situ Fourier transform mid-infrared reaction monitoring in aqueous solution. Anal. Chem. 70: 1726-1734.

Geladi, P., MacDougall, D., and Martens, H. 1985. Linearization and scatter-correction for near-infrared reflectance spectra of meat. Appl. Spectrosc. 39:491-500.

Goulden, J. D. S., inventor. 1964. Method and apparatus for the analysis of emulsions and suspensions. National Research Development Corporation, assignee. US Pat. No. 3161768.

Greenberg, N. A., and Mahoney, R. R. 1983. Formation of oligosaccharides by β-galactosidase from Streptococcus thermophilus. Food Chem. 10:195-204.

Gustafsson, A. H., Andersson, L., and Emanuelson, U. 1993. Effect of hyperketonaemia, feeding frequency and intake of concentrate and energy on milk yield in dairy cows. Anim. Prod. 56:51-60.

Gustafsson, A. H., and Emanuelson, U. 1996. Milk acetone concentration as an indicator of hyperketonaemia in dairy cows: the critical value revised. Anim. Sci. 63:183-188.

Guy, E., and Bingham, E. 1978. Properties of β-galactosidase of Saccharomyces lactis in milk and milk products. J. Dairy Sci. 61:147-151.

Gyuricsek, D. M., and Thompson, M. P. 1976. Hydrolyzed lactose cultured dairy products. II. Manufacture of yoghurt, buttermilk and cottage cheese. Cult. Dairy Prod. J. 11:12-13.

Hansen, P. W. 1998a. Urea determination in milk using Fourier transform infrared spectroscopy and multivariate calibration. Milchwissenschaft 53: 251-255. Enclosed as Publication I.

Hansen, P. W., inventor. 1998b. Measurement of acetone in milk using IR spectroscopy. Foss Electric A/S, assignee. Int. Pat. Appl. No. WO 98/43070. Enclosed as Publication IV.

Hansen, P. W. 1998c. Screening of dairy cows for ketosis by use of infrared spectroscopy and multivariate calibration. J. Dairy Sci. (submitted). Enclosed as Publication III.

Hansen, P. W. 1998d. Pre-processing method minimising the need for reference analyses. To be published. Enclosed as Publication V.

Hansen, P. W., inventor. 1998e. Evaluation of spectroscopic data. Foss Electric A/S, assignee. Danish Pat. Appl. No. PA 1998 01177. Enclosed as Publication VII.

Hansen, P. W., van Brakel, A. S., Garman, J., and Nørgaard, L. 1998. Detection of specific sugars in dairy process samples using multivariate curve resolution. J. Dairy Sci. (submitted). Enclosed as Publication VI.

Hendra, P. J., Wilson, H. M. M., Wallen, P. J., Wesley, I. J., Bentley, P. A., Arruebarrena-Baez, M., Haigh, J. A., Evans, P. A., Dyer, C. D., Lehnert, R., and Pellow-Jarman, M. V. 1995. Routine analytical Fourier transform Raman spectroscopy. Part 2. An updated review. Analyst 120:985-991.

Höskuldsson, A. 1992. The H-principle in modelling with application to chemometrics. Chemomet. Intell. Lab. Sys. 14:139-153.

Karjalainen, E. J. 1989. The spectrum reconstitution problem. Use of alternating regression for unexpected spectral components in two-dimensional spectroscopies. Chemomet. Intell. Lab. Sys. 7:31-38.

Kemsley, E. K., Appleton, G. P., and Wilson, R. H. 1994. Quantitative analysis of emulsions using attenuated total reflectance (ATR). Spectrochim. Acta, Part A 50:1235-1242.

Lakowicz, J. R. 1983. Principles of fluorescence spectroscopy. Plenum Press, New York.

Leardi, R., Boggia, R., and Terrile, M. 1992. Genetic algorithms as a strategy for feature selection. J. Chemomet. 6:267-281.

Lefier, D. 1996. Analytical methods for the determination of the urea content in milk. Bull. IDF 315:35-38.

Lefier, D., Grappin, R., and Pochet, S. 1996. Determination of fat, protein, and lactose in raw milk by Fourier transform infrared spectroscopy and by analysis with a conventional filter-based milk analyzer. J. AOAC Int. 79:711-717.

Luinge, H. J., Hop, E., Lutz, E. T. G., van Hemert, J. A., and de Jong, E. A. M. 1993. Determination of the fat, protein and lactose content of milk using Fourier transform infrared spectrometry. Anal. Chim. Acta 284:419-433.

Marstorp, P., Anfält, T., and Andersson, L. 1983. Determination of oxidized ketone bodies in milk by flow injection analysis. Anal. Chim. Acta 149:281-289.

Martens, H., and Næs, T. 1989. Multivariate Calibration. Wiley, Chichester.

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

Morales, F. J., Romero, C., and Jiménez-Pérez, S. 1996. Fluorescence associated with Maillard reaction in milk and milk-resembling systems. Food Chem. 57:423-428.

Nathier-Dufour, N., Sedman, J., and van de Voort, F. R. 1995. A rapid ATR/FTIR quality control method for the determination of fat and solids in sweetened condensed milk. Milchwissenschaft 50:462-466.

Nikolajsen, R. P. H. 1998. ATR/FTIR spectrometric analysis of milk products [in Danish]. M.Sc. thesis, Technical University of Denmark.

Nikolajsen, R. P. H., and Hansen, P. W. 1998. ATR FT-IR measurements of low concentration components in milk. Internal report. Foss Electric A/S, Denmark. Enclosed as Publication II.

Nygaard, L., Lapp, T., and Arnvidarson, B., inventors. 1993. Method of determining urea in milk. Foss Electric A/S, assignee. US Pat. No. 5252829.

Official methods and recommended practices of the American Oil Chemists' Society. 1996. Solid fat content (SFC) by low-resolution nuclear magnetic resonance. AOCS official method Cd 16b-93. American Oil Chemists' Society, Illinois.

Oltner, R., and Wiktorsson, H. 1983. Urea concentrations in milk and blood as influenced by feeding varying amounts of protein and energy to dairy cows. Livest. Prod. Sci. 10:457-467.

Prenosil, J. E., Stuker, E., and Bourne, J. R. 1987a. Formation of oligosaccharides during enzymatic lactose hydrolysis and their importance in a whey hydrolysis process: Part I: State of art. Biotechnol. Bioeng. 30:1019-1025.

Prenosil, J. E., Stuker, E., and Bourne, J. R. 1987b. Formation of oligosaccharides during enzymatic lactose hydrolysis and their importance in a whey hydrolysis process: Part II: Experimental. Biotechnol. Bioeng. 30:1026-1031.

Reichardt, W., Krauss, M., Gernand, U., Klauke, G., and Grau, W. 1995. The quantitative estimation of the protein content in milk by means of ultraviolet spectral photometry. Part 7. The milk analyzer MAG 100 as possibility for simultaneous measuring of the protein and fat content in raw milk on the dairy farm [in German]. Die Nahrung 39:139-148.

Ridder, C., and Kjær, L. 1995. Sweet dreams come true. Applied FTIR technology in ice cream mix analysis. Scand. Dairy Inf. 9/95:34-36.

Rodriguez-Otero, J. L., Hermida, M., and Centeno, J. 1997. Analysis of dairy products by near-infrared spectroscopy: A review. J. Agric. Food Chem. 45:2815-2819.

Rutledge, D. N. 1992. Low resolution pulse nuclear magnetic resonance (LRP-NMR). Analusis 20(3):M58-M62.

Silverstein, R. M., Bassler, G. C., and Morrill, T. C. 1991. Spectrometric identification of organic compounds. 5th ed. Wiley, New York.

Stapelfeldt, H., and Skibsted, L. H. 1994. Modification of β-lactoglobulin by aliphatic aldehydes in aqueous solution. J. Dairy Res. 61:209-219.

Steger, H., Girschewski, H., Piatkowski, B., and Voigt, J. 1972. Evaluation of the state of ketosis in lactating cows by estimating the concentration of ketones in blood and that of acetone in milk [in German]. Arch. Tierernaehr. 22:157-162.

Steiner, G., Kaiser, S., and Raeuber, H.-J. 1992. Refractometry and ATR-spectroscopy for the control of yoghurt fermentation [in German]. ZFL, Int. Z. Lebensm. Tech. Verfahrenstech. Verpack. Anal. 43(9):EFS32-EFS35.

Sun, J. 1997. Statistical analysis of NIR data: Data pretreatment. J. Chemomet. 11:525-532.

Tauler, R., Kowalski, B., and Fleming, S. 1993. Multivariate curve resolution applied to spectral data from multiple runs of an industrial process. Anal. Chem. 65:2040-2047.

van de Voort, F. R., Sedman, J., Emo, G., and Ismail, A. A. 1992a. A rapid FTIR quality control method for fat and moisture determination in butter. Food Res. Int. 25:193-198.

van de Voort, F. R., Sedman, J., Emo, G., and Ismail, A. A. 1992b. Assessment of Fourier transform infrared analysis of milk. J. AOAC Int. 75:780-785.

Williams, P. C., and Norris, K. H. 1987. Qualitative applications of near-infrared reflectance spectroscopy *in* Near-infrared technology in the agricultural and food industries. Williams, P. C., and Norris, K. H., ed. American Association of Cereal Chemists, Inc., Minnesota.

Wold, S., Antti, H., Lindgren, F., and Öhman, J. 1998. Orthogonal signal correction of near-infrared spectra. Chemomet. Intell. Lab. Sys. (in press).

Ægidius, P. E. 1996. From water in grain to bacteria in milk [in Danish]. Foss Electric A/S, Denmark.

List of Abbreviations and Symbols

ALS	alternating least squares
AR	alternating regression
ATR	attenuated total reflectance
BM	B filter method
CCD	charge-coupled device
CIP	clean in place
CLS	classical least squares
CPMG	Carr-Purcell-Meiboom-Gill
CVD	chemical vaporisation diamond
DO	direct orthogonalisation
FIA	flow injection analysis
FID	free induction decay
FPD	freezing point depression
FT	Fourier transform
HPLC	high pressure liquid chromatography
IIR	independent interference reduction
IR	infrared
IRE	internal reflection element
MSC	multiplicative scatter (or signal) correction
NIR	near infrared
NMR	nuclear magnetic resonance
OSC	orthogonal signal correction
PARAFAC	parallel factor analysis
PCA	principal component analysis
PCR	principal component regression
PLS	partial least squares
PV	principal variables
R^2	correlation
RMSEP	root mean square error of prediction
SEC	standard error of calibration
SEP	standard error of prediction
S/N	signal-to-noise ratio
SNV	standard normal variate
s_r	repeatability
SSE	sum of squared errors
UV/VIS	ultraviolet/visible

Mathematical Expressions

Repeatability

Repeatability is expressed as a mean standard deviation (s_r) of multiple determinations performed under identical conditions and is calculated as:

$$s_r = \sqrt{\frac{1}{q(n-1)} \sum_{j=1}^{q} \sum_{i=1}^{n} (x_{j,i} - \bar{x}_j)^2}$$

where q is the number of samples, n is the number of replicates, $x_{j,i}$ is the result of the i'th replicate of the j'th sample and $\bar{x}_j$ is the average result of the j'th sample.

Accuracy

Accuracy is expressed as the Root Mean Square Error of Prediction (RMSEP) and is calculated as:

$$RMSEP = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{i,reference} - x_{i,predicted})^2}$$

where N is the number of determinations (number of samples (q) times number of replicates (n) from above) and $x_{i,reference}$ and $x_{i,predicted}$ are the reference and predicted values corresponding to the i'th determination, respectively.

When a bias (mean difference between reference results and predictions) is observed, the Standard Error of Prediction (SEP) is used. It is calculated as:

$$SEP = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_{i,reference} - x_{i,predicted} - bias)^2}$$

If two variables are related by performing a univariate linear regression (slope a and intercept b), between instrumental responses, $x_{i,instrumental}$, and reference results, $x_{i,reference}$, the accuracy of future predictions can be estimated by the use of the Standard Error of Calibration (SEC), calculated as:

$$SEC = \sqrt{\frac{1}{N-2} \sum_{i=1}^{N} (x_{i,reference} - (a x_{i,instrumental} + b))^2}$$

Correlation

Correlation is expressed as R^2, which is calculated as:

$$R^2 = \left[\frac{\frac{1}{N} \sum_{i=1}^{N} (x_{i,reference} - \bar{x}_{reference})(x_{i,predicted} - \bar{x}_{predicted})}{s_{reference} s_{predicted}} \right]^2$$

where N, $x_{i,reference}$ and $x_{i,predicted}$ are defined above and $\bar{x}_{reference}$, $s_{reference}$, $\bar{x}_{predicted}$ and $s_{predicted}$ are the means and standard deviations of the reference and predicted results, respectively.

Model fit

The fit of the model to **X** (in case of an ALS model) is expressed as the Sum of Squared Errors (SSE):

$$SSE = \sum_{i=1}^{N} \sum_{j=1}^{M} (x_{i,j} - \mathbf{c}_i \mathbf{a}_j)^2$$

where $x_{i,j}$ is an element in **X**, $\mathbf{c}_i$ is as row vector containing the concentrations of the i'th sample and $\mathbf{a}_j$ is a column vector containing the absorbencies of the j'th wavelength. M is the number of wavelengths in the spectra.

Equipment List

The following spectroscopic equipment was used for the experiments performed during the studies:

Method	Equipment	Locations
Fluorescence	Aminco-Bowman Series 2	Dept. of Dairy and Food Science, Food Chemistry, The Royal Veterinary and Agricultural University (Denmark)
Mid-IR	Bomem MB Series (ATR)	Foss Electric A/S (Denmark)
	Foss MilkoScan 4000	Foss Electric A/S (Denmark)
	Foss MilkoScan FT 120	Foss Electric A/S (Denmark)
		New Zealand Dairy Research Institute (New Zealand)
NIR	Bomem MB Series	Foss Electric A/S (Denmark)
NMR	Resonance Instruments MARAN	Dept. of Dairy and Food Science, Food Technology, The Royal Veterinary and Agricultural University (Denmark)
Raman	Dilor LABRAM II (CCD-Raman)	Dilor S.A. (France)
	Perkin-Elmer System 2000 (FT-Raman)	Dept. of Dairy and Food Science, Food Technology, The Royal Veterinary and Agricultural University (Denmark)
UV/VIS	Perkin-Elmer Lambda 2	Foss Electric A/S (Denmark)

Publication I

Urea Determination in Milk Using Fourier Transform Infrared Spectroscopy and Multivariate Calibration

Reprint from *Milchwissenschaft*, vol. 53, pp. 251-255 (1998)

Urea determination in milk using Fourier transform infrared spectroscopy and multivariate calibration

By P.W. HANSEN

The Royal Veterinary and Agricultural University, Department of Dairy and Food Science, Food Technology, Chemometrics Group, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark

1. Introduction

Determination of urea in milk is important in relation to the optimization of the feeding of dairy cows. Urea is synthesized in the liver on the basis of ammonia which comes from the breakdown of excess amino acids in the small intestine (1). A simple interpretation suggests that a low urea content of the milk shows the cow to be underfed with protein. Similarly, a high urea content shows the cow overfed with protein and thus is in need of energy to eliminate the excess amino acids. Both cases are undesirable and can be alleviated by optimized feeding of the cows.

Urea is found in single cow samples in the range from 10 to 80 mg/dl with a mean of 30 to 40 mg/dl and a standard deviation of approximately 10 mg/dl. These values depend on the geographical region from which the samples are taken.

Routine analysis of urea in milk is carried out either directly by enzymatic analysis (*e.g.* Flow Injection Analysis (FIA) or Autoanalyzer systems) or indirectly by infrared transmission spectroscopy (MilkoScan instruments) (2). The infrared method is the fastest (7.2 s for 1 analysis), but is less accurate and requires many calibration samples. In routine work it has a minimal requirement for chemical reagents which also makes it less complex, and thus less sensitive to operator induced errors.

Traditionally, infrared determination of urea is based on an absorption at approximately 1465 cm^{-1}, since the major absorptions of urea in the region of about 1700 to 1550 cm^{-1} are obscured by a strong water absorption in the same area of the spectrum. Urea also absorbs at approx. 1160 cm^{-1}, but this band is not normally used because of an overlap from a lactose absorption. The 1465 cm^{-1} band of urea, in addition to absorptions originating from fat, protein, lactose and citric acid, is used for urea determination.

The infrared instruments presently being used for urea analysis are based on filter technology, and the number of filters sets a physical limit on the amount of spectral information that can be obtained. Fourier transform infrared (FT-IR) spectroscopy is able to offer analyses that are of the same high accuracy and repeatability on fat, protein and lactose as traditional filter instruments (3, 4, 5). In addition to the larger amount of information that FT-IR spectroscopy provides, it is very easy to standardize different instruments to give the same responses to identical samples and thus provides transferability of calibrations (6).

Thus, FT-IR could prove beneficial for infrared urea determination when combined with multivariate data analysis such as Principal Component Regression (PCR) or Partial Least Squares (PLS). These methods are described by MARTENS and NÆS (7). The scope of this paper will be a study of multivariate calibration to obtain predictions for urea in milk using FT-IR spectroscopy.

2. Materials and methods

2.1 Milk samples

900 samples from individual cows (pooled from 1 evening and 1 morning milking) were collected over 1 month from different Danish herds with cows of different breeds. The samples were preserved using Broad Spectrum Microtabs (D&F Control Systems Inc., CA, USA) in the concentration prescribed by the manufacturer. To ensure sufficient variability in the data set, a part of the samples (5%) were fortified with approx. 30 mg urea per dl milk, and another part of the samples (5%) were given the double concentration of preservative.

2.2 Reference measurements

All samples were reference analysed for urea using CL-10 Micro (EuroChem, Italy). One third of the samples were tested in duplicate to check the repeatability of the instrument. This was satisfactory with a value of 1.4 mg/dl (calculated according to the formula below) compared to an expected accuracy of about 3.0 mg/dl.

2.3 Infrared measurements

FT-IR spectra of all samples were recorded in duplicate on a MilkoScan FT120 (Foss Electric A/S, Denmark). The instrument was modified to use a cuvette with a pathlength of 37 µm (standard is 50 µm) to ensure good repeatability. It records the infrared spectra from 5000 to 925 cm^{-1}. The instrument was standardized before each day's work.

For comparative purposes, the samples were also measured on a filter based instrument, MilkoScan 4000 (Foss Electric A/S, Denmark). It is equipped with 11 filters of which 10 are used for the urea analysis.

All measurements were ratioed against water and log-transformed to give absorbencies.

2.4 Calculations

The data analysis and calibration was performed on a PC with Matlab 4.2c software (The MathWorks Inc., MA, USA). The calibration routines were either programmed or taken from the PLS_Toolbox Version 1.5 (Eigenvector Technologies, WA, USA).

Repeatability is expressed as a mean standard deviation (s_r) of multiple determinations performed under identical conditions and is calculated as:

$$s_r = \sqrt{\frac{1}{q(n-1)} \sum_{j=1}^{q} \sum_{i=1}^{n} (x_{j,i} - \langle x_j \rangle)^2}$$

where q is the number of samples, n is the number of replicates, $x_{j,i}$ is the result of the i'th replicate of the j'th sample and $\langle x_j \rangle$ is the average result of the j'th sample.

Accuracy is expressed as the Root Mean Square Error of Prediction (RMSEP) and calculated as:

$$RMSEP = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{i,reference} - x_{i,predicted})^2}$$

where N is the number of determinations (number of samples (q) times number of replicates (n) from above) and $x_{i,reference}$ and $x_{i,predicted}$ are the reference and predicted values corresponding to the i'th determination, respectively.

Correlation is expressed as R^2, which is calculated as:

$$R^2 = \left[\frac{\frac{1}{N} \sum_{i=1}^{N} (x_{i,reference} - \langle x_{reference} \rangle)(x_{i,predicted} - \langle x_{predicted} \rangle)}{s_{reference}\, s_{predicted}} \right]^2$$

where N, $x_{i,reference}$ and $x_{i,predicted}$ are defined above and $\langle x_{reference} \rangle$, $s_{reference}$, $\langle x_{predicted} \rangle$ and $s_{predicted}$ are the mean and standard deviations of the reference and predicted results, respectively.

3. Results and discussion

The data set was divided into 3 sets: 400 samples were used for calibration, 300 samples were used for test set validation (to select the optimal number of factors) and the remaining 200 samples were used for the final independent test set validation. The 3 sets were found by sorting the data according to the reference results, followed by a systematic selection of samples for the 3 sets. Selected in this way, all 3 sets will span the urea variation in the best possible way.

The infrared spectra of 2 different samples are shown in Fig. 1. The following areas were removed from the spectra: 5000 to 2973 cm^{-1} (NIR area and water absorption), 2695 to 1816 cm^{-1} (area with no absorptions from milk), 1712 to 1585 cm^{-1} (water absorption) and 960 to 925 cm^{-1} (CaF_2 – cuvette material – absorption).

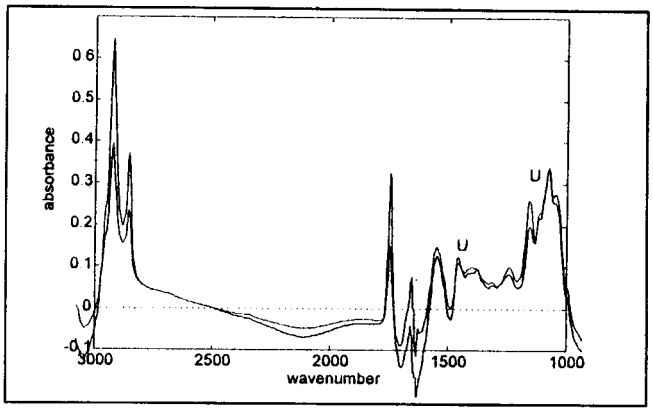


Fig. 1: Two typical infrared milk spectra ratioed against water. The 2 positions marked with 'U' are the positions of the 2 major urea absorptions in the spectrum

Before any attempt to perform a calibration on the sample set, a Principal Component Analysis (PCA) was performed on the absorbance data (scores plot, Fig. 2) and 7 samples were judged as spectral outliers. All of these samples had a somatic cell count from 1.6 to 11.0 million (Fossomatic 400 measurements) which suggests that they should be left out of a calibration for natural milk samples. In addition to these 7 samples, 10 were left out of the calibration because of missing or bad repeating reference results. The score plot with the remaining 883 samples is shown in Fig. 3, which shows that no other major outliers are present.

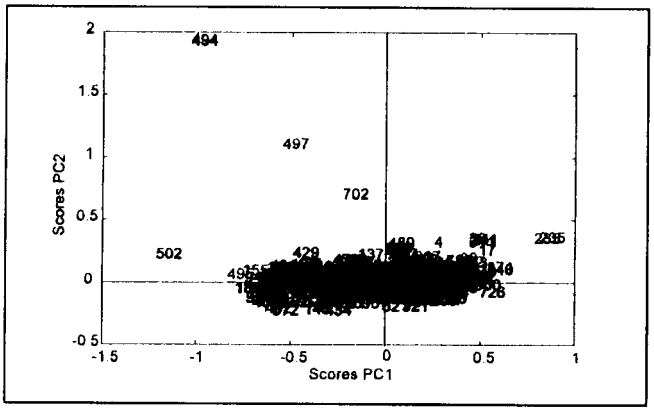


Fig. 2: PCA score plot, PC2 vs. PC1 on all samples (raw absorbance matrix) except 10 samples with bad repeating or missing reference results. Variance explained: PC1 79.48% and PC2 12.55%

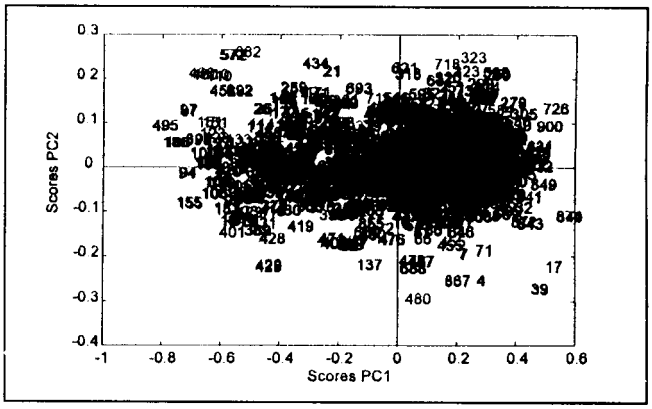


Fig. 3: PCA score plot, PC2 vs. PC1 as in Fig. 1, but with 7 outliers removed. Variance explained: PC1 87.02% and PC2 7.47%. Note how well the 2 replicates of each sample coincide

In the initial PLS calibration the scores on X (the raw absorbance matrix) were plotted against the residual variance in y (the urea reference results) for each PLS factor in order to find reference outliers. Two samples were removed because of their deviations from the others in the 6th factor and higher. The data set was thus reduced to consist of 395 calibration samples, 292 validation samples and 196 independent test samples.

The full spectrum data was pretreated in the following way:

1) No treatment
2) 1st derivative
3) 2nd derivative
4) Multiplicative Scatter Correction (MSC), only additive effect based on the area 1492 to 1481 cm^{-1}, commonly known as the "protein reference" area
5) MSC, both additive and multiplicative effect. The base was selected as all parts of the spectrum used for calibration, since no area without chemical information would span the whole absorbance range.
6) Division of the spectra by their vector norm, calculated by using all parts of the spectrum used for calibration

In addition, a calibration based on data from the filter instrument was performed.

In order to calculate the derivatives using a Savitzsky-Golay smooth, the number of points in each window must be chosen. This can be difficult, and consequently a test of different Savitzsky-Golay derivatives was carried out with the prediction error on the 300 test samples and the repeatability as measures of performance. The polynomial degree was fixed to 2. The results are shown in Table 1. These results show that a broad window was best, with the largest effect on the repeatability. Using these results, a 2nd order, 11-point Savitzsky-Golay was chosen in the calculation of the derivatives.

Table 1: Test of the sensivity of accuracy and repeatability to the number of points used in the calculation of the Savitzsky-Golay derivatives with a polynomial order of 2. RMSEP and s_r values were calculated using the 292 test samples

No. of points	1st derivative Factors	RMSEP	s_r	2nd derivative Factors	RMSEP	s_r
3	17	3.9	1.7	16	4.8	3.2
5	17	4.0	1.4	16	4.5	2.5
7	17	4.0	1.3	17	4.2	2.1
8	18	4.0	1.2	16	4.3	1.8
11	18	4.0	1.2	16	4.3	1.6

The results from the test set as well as the independent test set, in terms of correlation, accuracy and repeatability, are shown in Table 2 and 3. A plot of reference *vs.* measured with the best model (in terms of accuracy on the independent test set) using no pretreatment is given in Fig. 4. PLS loadings for factor 16 and B-coefficients from the same model are shown in Figs. 5 and 6, respectively.

The best model based on the available data was obtained without any data pretreatment, whereas most of the different treatments gave similar results. The results from the independent test set (Table 3), are very similar to the results from the test set (Table 2), so the model has not been overfitted. The loading spectra for the last

Table 2: Results from 292 test samples in terms of correlation coefficient, accuracy and repeatability. Results from the filter instrument (Milko Scan 4000) are shown for comparison. Repeatability is not estimated on this instrument, as only single measurements were performed

Treatment	Factors	R^2	RMSEP	s_r
None	16	0.89	3.7	1.2
1st derivative	19	0.86	4.0	1.2
2nd derivative	16	0.84	4.3	1.6
Additive MSC	17	0.88	3.7	1.1
Full MSC	14	0.88	3.7	1.2
Normalization	17	0.89	3.6	1.1
Filter instrument	8	0.76	5.3	–

Table 3: Results from the predictions of 196 independent test samples in terms of correlation coefficient, accuracy and repeatability

Treatment	R^2	RMSEP	s_r
None	0.87	3.9	1.3
1st derivative	0.86	4.1	1.3
2nd derivative	0.84	4.3	1.7
Additive MSC	0.86	4.0	1.2
Full MSC	0.86	4.0	1.3
Normalization	0.86	4.0	1.2
Filter instrument	0,74	5,5	–

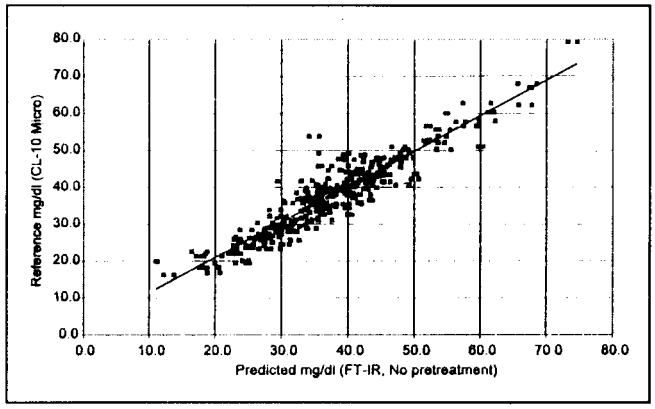


Fig. 4: Reference vs. predicted on the 196 independent test samples with the 16-factor model based on the raw spectra without any pretreatment. All replicates of all samples are shown

factor (Fig. 5) as well as the B-coefficient spectrum (Fig. 6) both contain structure to a large extent without much noise. This guarantees the stability of the model. These facts indicate that the data set is large enough to account for different interferences in the data, such as varying preservation, time drift, breed etc. The results further show that using derivative spectra results in poorer repeatability, which suggests that MSC of some type will be the best choice when data pretreatment is necessary.

That the FT-IR instrument actually uses the urea information in the spectrum is supported by the B-coefficient spectrum (Fig. 6). The plot shows that some of the most positive contributions to the urea determination are situated at 1469 and 1161 cm^{-1} which are consistent with the 2 absorptions from urea mentioned above. The primarily used absorption is still the 1465 cm^{-1} band, but the 1161 cm^{-1} band could be the one account

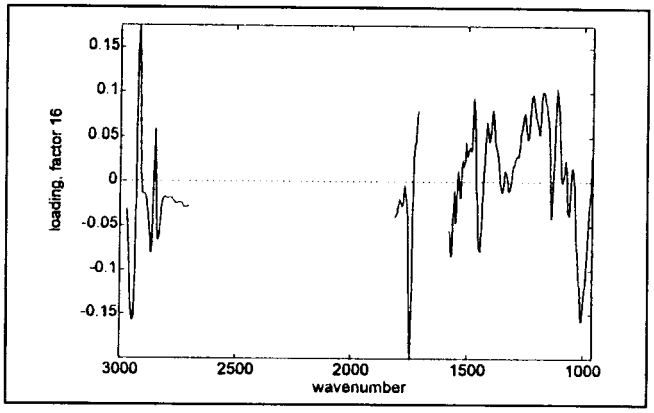


Fig. 5: Loading plot, PLS factor 16, for a model using data without any pretreatment

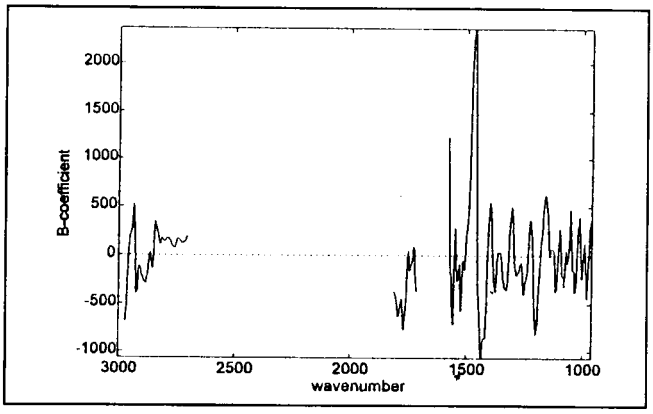


Fig. 6: Plot of B-coefficients vs. wave number for a model with 16 factors using data without any pretreatment

ing for the improved accuracy compared to the filter instrument. Negative B-coefficients are obtained at the primary absorptions from fat: 2853 (C-H stretch) and 1770 cm^{-1} (ester C=O stretch). This is reasonable, since fat has an absorption (1460 cm^{-1}, C-H bend) coinciding with the 1465 cm^{-1} band of urea.

The filter instrument shows a prediction error of 5.5 mg/dl on the independent test set which is typical for an instrument of this type. The accuracy of the FT-IR measurement compared to the filter-IR measurement is improved with 29% which provides better predictions for optimizing the feeding of the cows. The large number of PLS factors used in the models explains the poorer accuracy of the filter instrument: Ten filters cannot possibly account for 14 to 17 basic variations in the data. The number of factors could probably be lowered by different treatments of the spectra, *e.g.* a selection of the most significant variables, but to get as low as 10 would be very unlikely.

The repeatability of the FT-IR instrument is about 1.2 mg/dl, while it is 1.4 mg/dl on the reference results.

4. Conclusions

The present study shows that it is possible to obtain an improved urea prediction using the FT-IR technique compared to the traditional infrared filter measurement. The performance expressed as accuracy on identical samples is improved with 26%. It is clear that a comparable accuracy cannot be obtained using existing instrumentation based on filter technology, since more spectral information is needed to provide the better performance. Furthermore, FT-IR instruments can be standardized, and the developed urea calibration can thus be transferred between instruments.

Acknowledgements

The author would like to thank The Academy of Technical Sciences, Denmark, and Foss Electric A/S, Denmark, for supplying the funds for this research project and Inger Tholl and Jeannette Steinmeier for technical assistance. Foss Electric A/S also provided access to the instruments used in the experiments.

5. References

(1) OLTNER, R., WIKTORSSON, H.: Livestock Prod. Sci. **10** 457–467 (1983)
(2) NYGAARD, L., LAPP, T., ARNVIDARSON, B.: Internat. Publ. No. WO 93/19364
(3) VAN DE VOORT, F.R., SEDMAN, J., EMO, G., ISMAIL, A.A.: J. AOAC Internat. **75** 780–785 (1992)
(4) LUINGE, H.J., HOP, E., LUTZ, E.T.G., van HEMERT, J.A., DE JONG, E.A.M.: Analytica Chim. Acta **284** 419–433 (1993)
(5) LEFIER, D., GRAPPIN, R., POCHET, S.: J. AOAC Internat. **79** 711–717 (1996)
(6) ANDERSEN, H.V., KJÆR, L., HANSEN, P.W., RIDDER, C.: Internat. Publ. No. WO 96/24832
(7) MARTENS, H., NÆS, T.: Multivariate Calibration, Wiley (1989)

6. Summary

HANSEN, P.W.: **Urea determination in milk using Fourier transform infrared spectroscopy and multivariate calibration.** Milchwissenschaft **53** (5) 251–255 (1998).

24 Urea determination (Fourier transform infrared spectroscopy

Multivariate calibration of urea in milk, in the range from 10 to 80 mg/dl, with Fourier transform infrared spectroscopy is investigated on 900 single cow samples. Different data pretreatments, such as derivatives, multiplicative scatter correction and normalization, are used. The accuracy on an independent test set (expressed as the Root Mean Square Error of Prediction) is down to 3.9 mg/dl depending on the pretreatment. The existing fast method for urea analysis based on infrared filter instruments is also tested. The accuracy of this method is 5.5 mg/dl, which suggests that more information provided by the Fourier transform infrared spectra is needed to establish more accurate calibrations.

HANSEN, P.W.: **Harnstoffbestimmung in Milch mit Hilfe der Fourier-Transform-Infrarot-Spektroskopie und der multivariaten Kalibrierung.** Milchwissenschaft **53** (5) 251–255 (1998).

24 Harnstoffbestimmung (Fourier-Transform-Infrarot-Spektroskopie

An 900 Kuhmilchproben wurde die multivariate Kalibrierung von Harnstoff in Milch im Bereich 10-80 mg/dl mit der Fourier Transform-Infrarot-Spektroskopie untersucht. Unterschiedliche Daten-Vorbehandlungen wie Derivate,

multiplikative Streuungskorrektur und -normalisierung wurden eingesetzt. Die Genauigkeit bei einem unabhängigen Testset (ausgedrückt als Root Mean Square Error of Prediction) ist 3,9 mg/dl in Abhängigkeit von der Vorbehandlung. Die vorhandene Schnellmethode für die Harnstoffanalyse, basierend auf Infrarotfilterinstrumenten wird ebenfalls getestet. Die Genauigkeit dieses Verfahrens beträgt 5,5 mg/dl. Dies läßt vermuten, daß mehr Informationen von den Fourier-Transform-Infrarotspektren erforderlich sind, um genauere Kalibrierungen zu erreichen.

Publication II

ATR FT-IR Measurements of Low Concentration Components in Milk

Foss Electric A/S, Internal Report. Not previously published

ATR FT-IR Measurements of Low Concentration Components in Milk

Rikke P. H. Nikolajsen and Per Waaben Hansen
Foss Electric A/S,
Slangerupgade 69, P.O. Box 260, DK-3400 Hillerød, Denmark

ABSTRACT

ATR FT-IR measurements for the determination of urea and acetone in milk is compared to ordinary transmission IR spectroscopy. 105 individual cow samples fortified with urea and acetone were measured. It was possible to measure urea with an accuracy (RMSEP) of 6.9 mg/dl, while acetone was impossible to detect, using ATR. The effect of pre-processing and variable selection, as well as the benefits from using ATR, are discussed.

(**Key words:** ATR, FT-IR, FT 120, milk, urea, acetone, variable selection, pre-processing)

1. INTRODUCTION

Attenuated Total Reflectance (ATR) is a sampling method for infrared spectroscopy based on the fact that when a light ray inside a crystal with a high refractive index is reflected on the surface, a small amount of the light is travelling a short distance in the media next to the crystal, before it is reflected (Figure 1). This is used for measuring the infrared absorption spectra of strongly absorbing compounds which cannot be determined using an ordinary transmission cuvette. The method is described in detail elsewhere (Nikolajsen, 1997). Prior results on ATR measurements on milk products have been reported earlier (Nikolajsen, 1998).

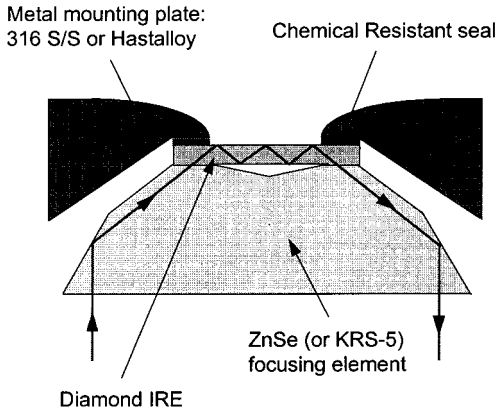


Figure 1 Scematic drawing showing the principle of ATR sampling. The sample is situated on top of the diamond. Reproduced from *The DurasamplIR: User's guide to effective sample handling.*

In infrared terms, milk is a strongly absorbing sample type, mainly due to its high water content (from 80 to 90 %). Thus, constituents absorbing under the water bands (obscuring the ranges from 3650 to 3030 and 1715 to 1580 cm^{-1}) cannot be observed using present infrared equipment (MilkoScan instruments) for quantitative determination of milk components. Today, the best results are obtained using Fourier transform infrared (FT-IR) instruments.

Urea and acetone are two components which could gain from inclusion of the water band. In milk, they occur in levels of approx. 0.03 % and 0.01 %, respectively. Although it is possible to make quantitative determinations of these two parameters, accuracy is low, and therefore they are presently regarded as being close to the "detection limit" of ordinary FT-IR equipment. The problem regarding determination of these components is that they are based on minor absorptions. The major absorption bands of both urea and acetone, stemming from the stretching of the carbonyl (-C=O) group, occur either very close to, or in the middle of, the water absorption centred at 1650 cm^{-1}.

In ATR the effective pathlength of the light through the sample is dependent on the wavenumber, the number of internal reflections in

the crystal, as well as the refractive index of both crystal and sample. It is typically shorter than the ordinary pathlength of a transmission cuvette, which is from 30 to 50 μm. This shorter pathlength opens up for measurements based on areas which are usually obscured by water. In relation to urea and acetone this could mean that the strong absorption of the carbonyl bond could be used to improve accuracy.

This is the scope of this study: to see whether the inclusion of the areas under the water bands (using ATR sampling) provides information that might improve the detection limit and accuracy of the urea and acetone determinations.

Other parameters with major absorptions under the water bands, such as specific proteins (*e.g.* caseins) could be studied. The choice of urea and acetone is based on the fact that Foss Electric A/S already had good in-house reference methods for these two parameters, and that it is possible to fortify the samples with these two compounds.

Ordinary transmission infrared spectroscopy will be compared to the ATR results.

2. SAMPLES

105 milk samples from individual cows from one morning milking were collected over a period of three weeks. Thirty samples were from Jersey herds. The samples were preserved using Broad Spectrum Microtabs.

In order to ensure a wide variation in both urea and acetone, the samples were fortified with the two compounds at random. 29 samples were fortified with urea – of these, 19 were added approx. 20 mg/dl urea, and 10 were added approx. 40 mg/dl urea. Likewise, 19 samples were added approx. 1 mM acetone, and 10 were added approx. 2 mM acetone. The samples for urea and acetone fortification were selected independently.

3. MEASUREMENTS

The 105 samples were measured using ATR sampling as well as an ordinary transmission instrument. Urea and acetone were determined using the reference methods.

3.1. ATR FT-IR measurements

The ATR FT-IR measurements were performed on a Bomem MB series interferometer, on which the ATR accessory was mounted. The accessory is built into an airtight box on top of the interferometer which was dried with silica gel, to prevent water interference. The interferometer is placed in a wooden box with a coarse temperature control, to avoid draft and other temperature changes in the surroundings to have an effect.

The ATR equipment is borrowed from ASI SensIR Technologies and consists of three units: the accessory with the internal reflection element with a special built steel cup for cleaning liquid on top of the crystal, a temperature controller and an ultrasonic cleaning system.

The internal reflection element (IRE) is an 18 reflection chemical vaporisation diamond (CVD), which is totally absorbing in the region: 1800 – 2300 cm^{-1} (pin 465-600), where no chemical information is present anyway. The IRE is attached to the temperature controller whose accuracy is ±0.1 °C . The experimental work was carried out at 40.0 °C. All milk samples were in 30 ml beakers and preheated in the water bath at 40 – 41 °C for approx. 20 minutes. All samples were measured in three replicates, and a water background was measured before each replicate. The resolution was 16 cm^{-1} and 60 scans were co-added. It was possible to measure 12 to 15 samples per day with the procedure described below. A preserved sucrose (5 % in water) standard was measured five times in the beginning and at the end of every working day. The intention was to use it for correction of the spectral data if day to day instabilities were observed. However, this

was not the case, but the spectra have been saved for later use.

The measurement procedure was as follows:

1. A cleaning procedure is carried out twice (see below for description of the cleaning procedure)
2. A water background is measured
3. A preheated milk sample is homogenised using the Foss Electric H-pump, and a few drops are placed on the ATR crystal and measured. The rest of the homogenised sample is placed in the water bath again (the milk samples were in the water bath approximately 30-35 minutes in total).
4. A cleaning procedure (see below) is performed
5. Point 2 – 4 are repeated until all measurements have been completed
6. A CIP procedure (see below) is performed

The cleaning procedure:

1. Remove milk sample
2. Apply a mixture of 0.01 % Brij, 0.5 % NaOH and 0.5 % Na_2EDTA in water to the crystal, and use the ultrasonic cleaner for 15 seconds (it is important that the ultrasonic cleaner head is in contact with the cleaning mixture).
3. Rinse with water
4. Remove water and dry with a paper tissue (Kimwipes)

Both the cleaning mixture and water is heated in the water bath.

CIP procedure:

CIP means Clean In Place (used in dairy plants to clean pipes).

1. Place an aqueous solution of 0.1 M NaOH on the crystal and leave it for 20 minutes
2. Rinse with water
3. Place an aqueous solution of 0.1 M HNO_3 on the crystal and leave it for 20 minutes
4. Rinse thoroughly with water

3.2. Transmission FT-IR measurements

The transmission measurements were carried out using a MilkoScan FT 120. The instrument was standardised before use. All measurements were carried out in duplicate. The samples were heated to a temperature of 40 °C before the determination.

3.3. Reference measurements

The urea reference results were obtained using the CL-10 Micro. All samples were measured in duplicate. The repeatability (s_r) was excellent: 0.34 mg/dl (compared to a usual FT-IR urea accuracy of 3 to 4 mg/dl). The range of the results was 8.7 to 67.2 mg/dl.

Acetone reference results were obtained by a modified vanillin method (Hansen, 1998). All samples were measured in duplicate. The repeatability (s_r) was also excellent: 0.02 mM (usual FT-IR acetone accuracy from 0.3 to 0.4 mM). The range of the results was 0.0 to 2.0 mM.

4. RESULTS

The average spectra from the two spectroscopic methods are shown in Figure 2. The ATR absorbances are generally lower than their FT 120 counterparts. This is due to the shorter pathlength in ATR sampling. The effective ATR pathlength is decreasing with the wavenumber of the infrared light. It is 2-3 times shorter in the left hand (sugar area) part of the spectrum (approx. 950 – 1150 cm^{-1}) and 7-8 times shorter in the Fat B area (2700 – 2900 cm^{-1}).

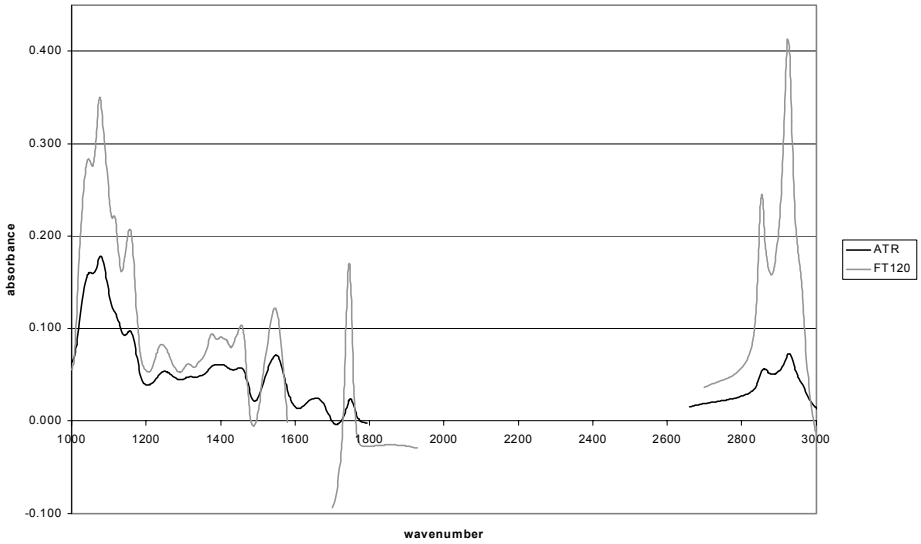


Figure 2 Mean ATR and FT 120 spectra (averaged over all samples), respectively.

The results fall within two groups – one for each of the two parameters: urea and acetone. Spectral data from both spectroscopic methods have been treated in exactly the same way in order to make the results comparable. In addition, spectral pre-processing using *e.g.* derivatives, as well variable selection has been tried – to obtain the optimal models.

Before any attempt to optimise the models, it is very important to remove outliers from the data set. The FT 120 data did not contain any obvious outliers, while eight outliers (sample No. 16, 38, 48 68, 79, 86, 95 and 99 – all replicates) were found in the ATR data set. They were found by use of the t vs. u plot during PLS calibration, as well as a combination of various measures introduced by Carsten Ridder.

After removing the outliers, the remaining samples were split into a calibration set (65 samples) and a test set (32 samples). The samples were sorted according to their reference results prior to the splitting, followed by a systematic selection, to obtain the same range in both sets. Therefore, the calibration and test sets used for urea and acetone were not the same.

It should be kept in mind that when comparing the results from the ATR method and the FT 120, the comparison is not totally fair – the FT 120 is an optimised instrument, while the ATR is an experimental set-up where improvements are possible. Therefore, even badly looking results may be promising.

4.1. Pre-processing methodologies

In pre-processing, the following methods were tried:

- Direct orthogonalisation (DO): A method developed at The Royal Veterinary and Agricultural University (Andersson, 1998a) which removes (by orthogonalisation, followed by a PCA on the resulting matrix) all information not related to the component of interest before a PLS calibration is performed. The number of principal components to be extracted depends on the actual application, so in this case from 1 to 10 components were tried.

- Derivatives: The classical spectroscopic pre-treatment. The Savitzsky-Golay method was used with various combinations of the number of points and the degree of the polynomial.

- Standard Normal Variate (SNV): A method where the individual spectra are treated to have a mean of zero and a standard deviation of one. This is done by

Pre-processing	#fac	RMSEP	R^2	s_r
None	13	6.9	0.725	5.3
1 DO component	12	6.8	0.730	5.4
2 DO components	10	6.9	0.727	5.5
3 DO components	9	6.9	0.727	5.5
4 DO components	8	6.9	0.727	5.5
5 DO components	7	6.9	0.727	5.5
6 DO components	6	6.9	0.730	5.6
7 DO components	5	6.8	0.734	5.6
8 DO components	5	6.8	0.731	5.5
9 DO components	4	6.8	0.735	5.5
10 DO components	3	7.1	0.713	5.3
1. derivative, 5-pt, 2. deg.	12	8.1	0.629	6.5
1. derivative, 5-pt, 3. deg.	13	8.6	0.584	6.9
1. derivative, 7-pt, 2. deg.	13	8.0	0.633	6.3
1. derivative, 7-pt, 3. deg.	12	8.5	0.594	6.9
1. derivative, 9-pt, 2. deg.	11	7.8	0.654	5.9
1. derivative, 9-pt, 3. deg.	12	8.3	0.615	6.5
2. derivative, 5-pt, 2. deg.	13	12.2	0.256	7.9
2. derivative, 5-pt, 3. deg.	13	12.2	0.256	7.9
2. derivative, 7-pt, 2. deg.	11	11.2	0.331	7.6
2. derivative, 7-pt, 3. deg.	11	11.2	0.331	7.6
2. derivative, 9-pt, 2. deg.	12	10.1	0.424	6.5
2. derivative, 9-pt, 3. deg.	12	10.1	0.424	6.5
SNV	11	7.7	0.659	5.8
Reference filter	11	7.8	0.655	5.6

Table 1 Result on the independent test set for various pre-processing methods applied to the ATR data for urea.
#fac = optimal number of PLS factors

subtracting the mean and dividing by the standard deviation.

- Reference filter: This is the classical pre-treatment in milk applications. A spectral area is selected, and its mean is subtracted from all points in the spectrum. Traditionally, the so-called Protein Reference (1492-1481 cm^{-1}) area is used.

To find the optimal number of PLS factors for each model multiple cross validations (with 4, 6, 8 and 10 cross validation segments) were carried out using the calibration set only. The average number of factors was used in the final model which was tested on the independent test set.

4.2. Variable selection methodologies

The variable selection was performed using both Principal Variables (PV) (Höskuldsson, 1992) and the B filter Method. (BM) (Andersson, 1998b). Genetic algorithms (Leardi *et al.*, 1992) (implemented by Carsten Ridder) with 100 generations were run following both PV and BM selection, and the optimal model was chosen for each of them.

- In PV variables are selected using the correlation between **X** and **Y** (*i.e.* the **w** vector). **X** is orthogonalised with respect to the variable showing the best correlation (*i.e.* the numerically largest element of **w**). Then a new model is calculated using the remaining data and a new "best variable" is found. In this way variables best suited for describing the relationship between **X** and **Y** are selected.

- In the B filter method the **X** matrix is scaled using the B coefficient spectrum (the regression coefficients of the calibration model). Then a new model is calculated, and **X** is scaled with the new B coefficients. In this way the most significant variables are encircled, and when repeating this procedure a number of times, filters which contain one or more variables are obtained.

4.3. Urea results

4.3.1. ATR FT-IR

The results of the various pre-processing methods on the independent test set are shown in Table 1.

Nothing is obtained by applying pre-processing to the spectra. In most cases the results are worse than the result obtained from the original spectra. Derivatives increase noise seriously, which is detrimental in this case where the noise (measured in terms of s_r) is the main contributor to the error.

The results of the variable selection methods, both for independent test set and cross validation are presented in Table 2.

Selection method	CVS	#fac	variables	Cross validation RMSEP	Cross validation RMSEP (mean)	Test set RMSEP	Test set R^2	Test set s_r
BM	8	13	29	5.5	4.1	7.2	0.698	5.6
PV	8	13	27	5.5	4.0	7.3	0.690	5.3

Table 2 Result of variable selection applied to the ATR data for urea.
RMSEP (mean) is the error when the three replicates from each sample are averaged before the calculations.
PV = Principal Variables, BM = B filter method, CVS = cross validation segments, #fac = optimal number of PLS factors.

The results of the variable selection are actually worse than what was obtained in the full spectrum calibration. The probable reason for this will be discussed later. A typical predicted vs. measured plot is presented in Figure 3.

4.3.2. Transmission FT-IR

The results for urea on the independent test set are shown in Table 3.

The result is much the same as what was seen from the ATR data, although the accuracy and repeatability is much better in this case. Pre-processing does not result in any significant improvements in this case, either.

The results of the variable selection methods, both for independent test set and cross validation are presented in Table 4.

The results of the variable selection are comparable to the full spectrum calibration results. A typical predicted vs. measured plot is shown in Figure 4.

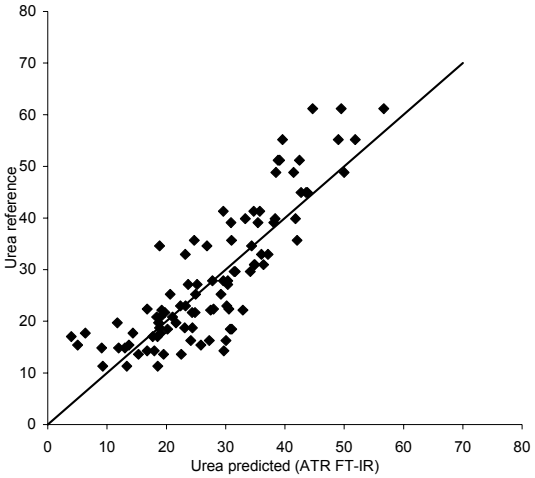


Figure 3 Reference vs. predicted from the ATR urea calibration (full spectrum, no pre-processing, 13 PLS factors). RMSEP 6.9 mg/dl, s_r 5.3 mg/dl.

4.3.3. Discussion

The ATR results are obviously worse than the results from the FT 120 (2 to 3 times larger in terms of RMSEP). This is mainly due to repeatability problems, which is demonstrated by the fact that the RMSEP calculated by

Pre-processing	#fac	RMSEP	R^2	s_r
None	16	2.9	0.953	1.1
1 DO component	15	2.9	0.953	1.1
2 DO components	14	2.9	0.953	1.1
3 DO components	13	2.9	0.954	1.1
4 DO components	12	2.9	0.954	1.1
5 DO components	11	2.9	0.954	1.1
6 DO components	10	2.9	0.953	1.1
7 DO components	9	2.9	0.953	1.1
8 DO components	8	2.9	0.954	1.1
9 DO components	7	2.9	0.953	1.1
10 DO components	6	2.9	0.952	1.1
1. derivative, 5-pt, 2. deg.	15	2.8	0.958	1.9
1. derivative, 5-pt, 3. deg.	15	3.1	0.950	1.8
1. derivative, 7-pt, 2. deg.	15	2.8	0.956	1.7
1. derivative, 7-pt, 3. deg.	15	2.9	0.955	1.7
1. derivative, 9-pt, 2. deg.	15	3.0	0.950	1.5
1. derivative, 9-pt, 3. deg.	15	2.8	0.957	1.8
2. derivative, 5-pt, 2. deg.	16	3.8	0.923	2.4
2. derivative, 5-pt, 3. deg.	16	3.8	0.923	2.4
2. derivative, 7-pt, 2. deg.	15	3.4	0.939	2.6
2. derivative, 7-pt, 3. deg.	15	3.4	0.939	2.6
2. derivative, 9-pt, 2. deg.	14	3.3	0.943	1.9
2. derivative, 9-pt, 3. deg.	14	3.3	0.943	1.9
SNV	13	3.3	0.941	1.2
Reference filter	13	3.0	0.952	1.4

Table 3 Result on the independent test set for various pre-processing methods applied to the FT 120 data for urea.
#fac = optimal number of PLS factors

Selection method	CVS	#fac	variables	Cross validation RMSEP	Cross validation RMSEP (mean)	Test set RMSEP	Test set R^2	Test set s_r
BM	8	11	22	2.7	2.6	3.0	0.951	1.2
PV	8	16	26	2.0	1.8	2.7	0.962	1.2

Table 4 Result of variable selection applied to the FT 120 data for urea.
RMSEP (mean) is the error when the three replicates from each sample are averaged before the calculations.
PV = Principal Variables, BM = B filter method, CVS = cross validation segments, #fac = optimal number of PLS factors.

cross validation for urea drops from 5.5 to 4.0 when averaging the replicates (Table 2), and that the accuracy and repeatability generally are very similar. The repeatability problem described earlier (Nikolajsen, 1998) prevails. This is seen from the fat, protein and lactose accuracies on the present test samples (calculated to be 10, 3 and 1 $\%C_V$, respectively) which are very similar to previous results.

In spite of this, there is no doubt that urea is detected by the ATR method. This is proved by the fact that the B coefficient spectra for the ATR and FT 120 full spectrum calibrations show the same characteristic peaks where urea has its major absorptions (Figure 5). It is also clear that the ATR calibration gathers some of its information from the peak under the water band in the ATR spectrum. This shows that additional urea information is obtained, but still the ordinary urea wavenumber (approx. 1465 cm^{-1}) has the largest peak. As the effective ATR pathlength in this range is approx. half of the FT 120 path (see section 4), the B coefficients would be expected to be doubled. This is not the case and is likely to be caused by the inclusion of the area under the water band.

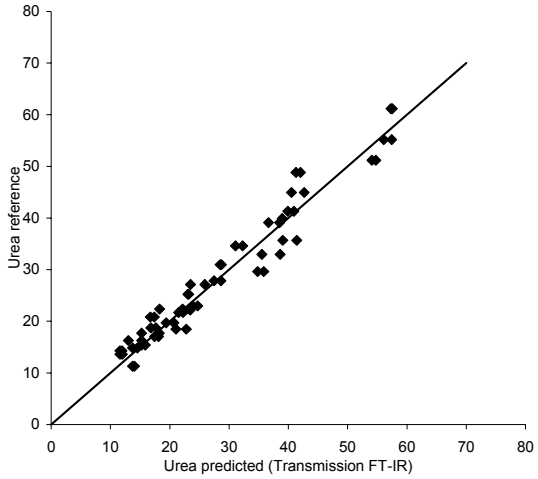


Figure 4 Reference vs. predicted from the FT 120 urea calibration (full spectrum, no pre-processing, 16 PLS factors). RMSEP 2.9 mg/dl, s_r 1.1 mg/dl.

This result is very promising, as it indicates that an improved ATR measurement or an FT 120 equipped with a 10-20 μm cuvette could give enhanced urea predictions.

Pre-processing of the data does not result in any significant improvements of the accuracy (Tables 1 and 3). In most cases both accuracy and repeatability becomes worse for both ATR and transmission results. The variable selection methods for the ATR data show much larger RMSEP values when predicting the test sets than for cross validation, which indicates an overfit (Table 2). Variable selection, when using FT 120 data, tends to yield accuracies comparable to full spectrum results (Table 4). This indicates that the rather irreproducible ATR spectra "disturb" the variable selection.

It should be stressed that if the reproducibility problems are related to the size of the particles in the milk, only better homogenisation or a greater penetration depth is expected to improve the method.

The overall result is therefore: Urea in milk can be measured using ATR – the area under the water band is used in the determination – and neither pre-processing nor variable selection improves the result.

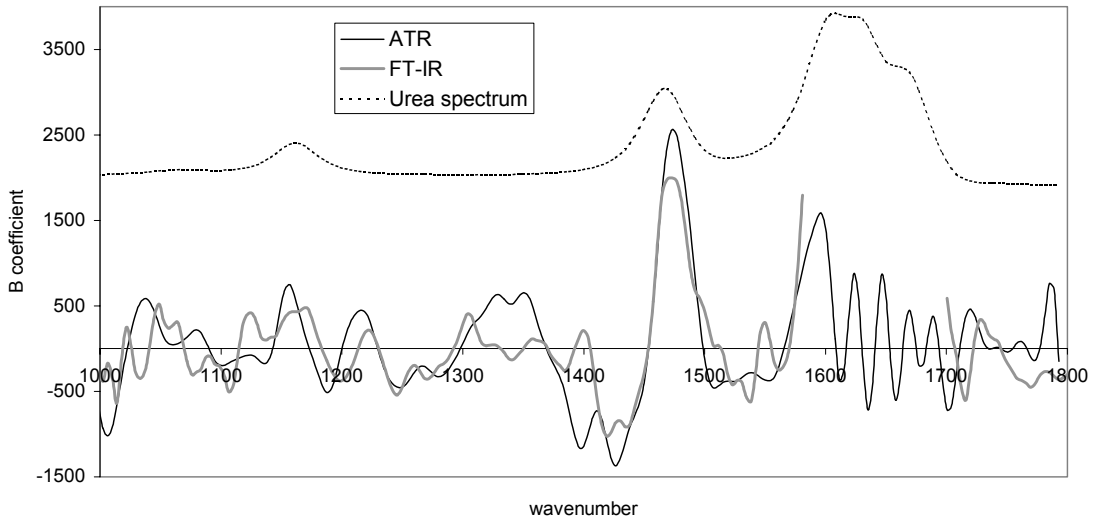


Figure 5 B coefficients from the two urea calibrations made using ATR and FT120 data, respectively. The ATR FT-IR spectrum of urea in water is overlaid.

4.4. Acetone

4.4.1. ATR FT-IR

It was not possible to establish any correlations between the ATR spectra and the acetone results. This is not surprising, as urea accuracy was already poor – and acetone occurs in even lower concentrations.

4.4.2. Transmission FT-IR

It was possible to establish an acetone calibration based on the FT 120 data. The results on the independent test set are shown in Table 5.

The results of the variable selection methods, both for the independent test set and cross validation are presented in Table 6.

Thus, acetone could be measured (as expected) using the FT 120. We will not go deeper into this result, as there are no ATR results to compare them to. We just note that the accuracy is similar to what has been observed on non-fortified samples in the past (Hansen, 1998). A typical predicted vs. measured plot is presented in Figure 6.

Pre-processing	#fac	RMSEP	R^2	s_r
None	21	0.22	0.820	0.11
1 DO component	20	0.22	0.820	0.11
2 DO components	19	0.23	0.819	0.11
3 DO components	18	0.22	0.819	0.11
4 DO components	17	0.22	0.820	0.11
5 DO components	16	0.22	0.820	0.11
6 DO components	15	0.22	0.822	0.11
7 DO components	14	0.22	0.822	0.11
8 DO components	13	0.22	0.822	0.11
9 DO components	12	0.22	0.823	0.11
10 DO components	11	0.22	0.824	0.11
1. derivative, 5-pt, 2. deg.	16	0.22	0.826	0.12
1. derivative, 5-pt, 3. deg.	16	0.22	0.833	0.15
1. derivative, 7-pt, 2. deg.	19	0.20	0.865	0.08
1. derivative, 7-pt, 3. deg.	16	0.21	0.840	0.14
1. derivative, 9-pt, 2. deg.	17	0.21	0.841	0.11
1. derivative, 9-pt, 3. deg.	16	0.21	0.836	0.11
2. derivative, 5-pt, 2. deg.	18	0.24	0.785	0.20
2. derivative, 5-pt, 3. deg.	18	0.24	0.785	0.20
2. derivative, 7-pt, 2. deg.	17	0.22	0.823	0.15
2. derivative, 7-pt, 3. deg.	17	0.22	0.823	0.15
2. derivative, 9-pt, 2. deg.	16	0.23	0.807	0.12
2. derivative, 9-pt, 3. deg.	16	0.23	0.807	0.12
SNV	18	0.27	0.736	0.13
Reference filter	16	0.30	0.681	0.11

Table 5 Result on the independent test set for various pre-processing methods applied to the FT 120 data for acetone.
#fac = optimal number of PLS factors

Selection method	CVS	#fac	variables	Cross validation RMSEP	Cross validation RMSEP (mean)	Test set RMSEP	Test set R^2	Test set s_r
BM	8	22	42	0.113	0.089	0.204	0.853	0.132
PV	8	19	32	0.147	0.125	0.211	0.853	0.133

Table 6 Result of variable selection applied to the FT 120 data for acetone.
RMSEP (mean) is the error when the three replicates from each sample are averaged before the calculations.
PV = Principal Variables, BM = B filter method, CVS = cross validation segments, #fac = optimal number of PLS factors.

5. CONCLUSIONS

From the present results it is concluded that:

- Urea can be determined by ATR FT-IR with an accuracy of 6.9 mg/dl (24 $\%C_V$, range: 8.7 – 67.2 mg/dl).
- The additional areas (water bands) provided by ATR are used in the urea determination.
- Acetone cannot be measured with the present ATR set-up.
- Neither pre-processing nor variable selection improves the calibration results.

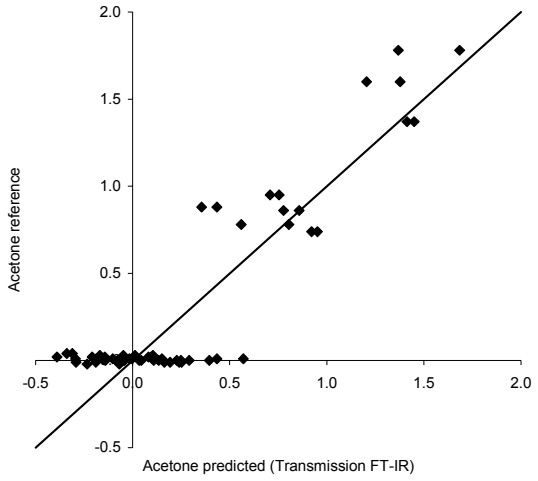


Figure 6 Reference vs. predicted from the FT120 acetone calibration (full spectrum, no pre-processing, 21 PLS factors). RMSEP 0.22 mM, s_r 0.11 mM.

6. REFERENCES

Andersson, C. A. 1998a. Direct orthogonalization. Chemomet. Intell. Lab. Sys. In press.

Andersson, M. 1998b. Personal communication. Technical Analytical Chemistry, Lund University, Sweden.

The DuraSamplIR™: User's guide to effective sample handling. 1997. ASI SensIR Technologies, Connecticut.

Hansen, P. W. 1998. Screening of dairy cows for ketosis by use of infrared spectroscopy and multivariate calibration. J. Dairy Sci. (submitted).

Höskuldsson, A. 1992. The H-principle in modelling with application to chemometrics. Chemomet. Intell. Lab. Sys. 14:139-153.

Leardi, R., Boggia, R., and Terrile, M. 1992. Genetic algorithms as a strategy for feature selection. J. Chemomet. 6:267-281.

Nikolajsen, R. P. H. 1997. Use of ATR/FT-IR spectrometry for food analysis [in Danish]. Pre-exam project, Technical University of Denmark.

Nikolajsen, R. P. H. 1998. ATR/FTIR spectrometric analysis of milk products [in Danish]. M.Sc. thesis, Technical University of Denmark.

Publication III

Screening of Dairy Cows for Ketosis by use of Infrared Spectroscopy and Multivariate Calibration

Submitted for publication in *Journal of Dairy Science*

Screening of Dairy Cows for Ketosis by use of Infrared Spectroscopy and Multivariate Calibration

Per Waaben Hansen
R&D Greenhouse, Foss Electric A/S,
Slangerupgade 69, P.O. Box 260, DK-3400 Hillerød, Denmark

ABSTRACT

A fast and inexpensive screening method for ketosis in cows is presented. It is based on the determination of acetone in the cow's milk by use of Fourier transform infrared spectroscopy combined with multivariate calibration. Only samples with a natural content of acetone could be used in the calibration step. On test samples ranging from 0.0 to 2.8 mM acetone, a correlation (R^2) of 0.81 and an accuracy (Root Mean Square Error of Prediction) of 0.27 mM was obtained. This accuracy was sufficient to allow for classification of the cows into two groups: one for healthy cows, and one for the cows to be kept under observation, with no false negatives on the present data set.

(**Key words:** acetone, Fourier transform infrared spectroscopy, ketosis, multivariate calibration)

Abbreviation Key: FIA = Flow Injection Analysis, **FT-IR** = Fourier Transform Infrared, **PLS** = Partial Least Squares.

INTRODUCTION

Ketosis is a metabolic disease observed in high-yielding lactating animals as well as diabetic humans. In cows, the disease means lower milk yield (5) as well as lower fertility (2). Especially in the time between the third and sixth week of the lactation, the cows are amenable to ketosis. Ketosis occurs less frequently if the cows are given sufficient feed, but in the case of pasture based farming this can be a problem. In such a case, additional feeding could solve the problem.

At least three chemical compounds are observed in the blood of a cow suffering from ketosis. They are: β-hydroxybutyrate, acetoacetate and acetone. These compounds – the so-called ketone bodies – occur in raised levels in the blood of a cow suffering from ketosis. It has been shown that there is a strong correlation between the amount of acetone in the milk and the ketone bodies in the blood from ketotic cows (11). The acetoacetate content is less than 15 % of the acetone content in the milk from a ketotic cow, and it is correlated to the acetone content (1). Therefore acetone is the best marker if only one compound in the milk is to be considered.

The acetone content of milk range typically from 0 to 2 mM. As a result of research in the area, the following limits for the milk acetone concentration have been proposed for its use as a ketosis marker (6):

less than 0.7 mM: The cow is healthy
between 0.7 and 1.4 mM: The cow might suffer from ketosis
above 1.4 mM: The cow suffers from ketosis

Classification of all cows according to these limits could help the dairy farmer in optimising the feeding of the herd, so there would be fewer incidences of ketosis in the future.

At present, ketosis testing is carried out using a colour strip test (4) or instruments based on Flow Injection Analysis (**FIA**) technology (9). FIA is capable of measuring up to 100 samples per hour with a high accuracy. Its main drawback is the low speed as it is not able to keep up with infrared instruments measuring *e.g.* fat, protein, lactose and urea at a rate of 500 samples per hour.

As Fourier transform infrared (**FT-IR**) spectroscopy has been shown capable of detecting

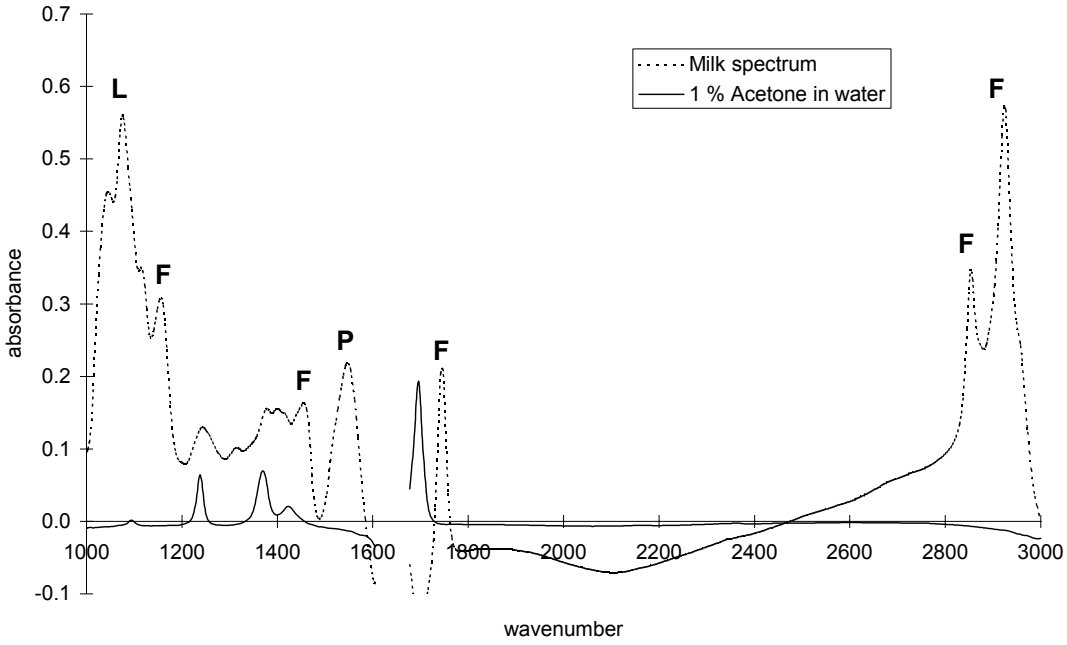


Figure 1 Infrared absorbance spectrum of acetone and milk. The major peaks of acetone are centred at 1696 (C=O stretch), 1370 (CH_3 symmetrical deformation) and 1238 cm^{-1} (C-C-C asymmetrical stretch). The major absorptions of fat (F), protein (P) and lactose (L) are indicated in the milk spectrum.

compounds like urea, occurring in levels of about 0.035 % (5.8 mM) in milk (7), the question rises whether other low concentration components, such as acetone, can be determined in this way. This would add this parameter to the existing dairy herd management systems without additional cost.

Acetone has three major absorptions in the mid-infrared so-called fingerprint area, as is evident from the spectrum of 1 % acetone in water shown in Figure 1. They are: 1696 (C=O stretch), 1370 (CH_3 symmetrical deformation) and 1238 cm^{-1} (C-C-C asymmetrical stretch). Other weaker absorptions are present as well. An average milk spectrum is also plotted in Figure 1. When comparing the milk absorptions to the acetone absorptions, it becomes clear that although appearing in very low concentrations, measurement of acetone might still be possible, as the absorptions are isolated from the major fat, protein and lactose peaks (indicated in the figure).

The usual method for establishing relationships between spectral data and wet chemistry (reference) results is by use of Partial Least Squares (PLS) regression (10). It is generally accepted that in order to obtain reliable calibrations for components occurring in minor concentrations, many samples (*i.e.* hundreds) are needed. In the present case, however, it is very difficult to obtain milk samples with a natural acetone content different from zero. Approximately 10 % (calculated from the data used in the present study) of the cows suffer from ketosis within a period of four weeks (third to sixth week) shortly after calving. This means that when using random sampling, only approximately 1 % of the samples would contain acetone on a ketotic level. Selective sampling or use of fortified samples is therefore necessary in order to cover the range of samples needed for calibration. In the present case fortification is preferable, as selective sampling is very resource demanding.

The scope of the present work is to answer the question whether acetone can be detected by the use of FT-IR and to determine the obtainable accuracy. It will be investigated whether fortified samples can be used for building calibrations when natural samples are to be tested. PLS regression will be used for calibration.

MATERIALS AND METHODS

Milk samples

Natural calibration samples. 310 milk samples from individual cows, pooled from one evening and one morning milking were collected from Norwegian (51 samples), Swedish (59 samples) and Danish (200 samples) cows between first and sixth week of the lactation. The samples were preserved using Broad Spectrum Microtabs (D&F Control Systems Inc., CA, USA) in the concentration prescribed by the manufacturer. The acetone content of these samples ranged from 0.0 to 4.7 mM. Eight outliers were detected and removed from the set.

Fortified calibration samples. 125 milk samples from individual cows in all stages of the lactation, were used. They were fortified with acetone on levels from 0 to 2 mM, in order to assure a wide range of the acetone results. A 1.0 M stock solution of acetone in demineralised water was used for fortification. The samples were preserved using Broad Spectrum Microtabs. The acetone content of these samples ranged from 0.0 to 2.3 mM.

Test samples. To make an evaluation on an independent test set, 58 samples collected by Dairying Research Corporation Ltd. (Hamilton, New Zealand) were measured. The samples were collected by the end of August where the frequency of ketosis is highest (due to the fact that most New Zealand cows calve at the same time). They were chosen among a total of 142 samples in the following way: 26 samples had infrared acetone predictions (using a preliminary acetone calibration based on natural samples) higher than 0.5 mM. Of the remaining 116 samples, 32 samples were chosen at random and added to the test set to check for false negatives. This set is thus expected to have a high occurrence of ketosis samples. All test samples were preserved using a bronopol solution. The acetone content of these samples ranged from 0.0 to 2.8 mM.

Reference measurements

Reference results on the acetone content of the samples were obtained using the vanillin test. The method described by Drewes (3) was modified in order to make it suitable for milk measurements. It is based on the fact that the reaction between acetone and vanillin under influence of strong base (crossed aldol condensation) produces the coloured vanillal-acetone and di-vanillalacetone which can be detected at 415 or 500 nm. The procedure was:

2 mL of 2 % vanillin (SIGMA V2375) in a 4 N KOH solution was added to a 30 mL plastic tube. A Teflon membrane (Cole-Parmer 2916-64, 0.5 μm pore size, diameter 47 mm) was placed on top of the tube, and 200 μL sample was pipetted on top of it. The tube was closed with a tight fitting lid so the membrane was trapped between the tube and the lid. The tubes were placed on a water bath at 55 °C for 1 hour. After cooling to room temperature, the tubes were opened, the sample and membrane was removed, and 3 mL of demineralised water was added to the reagent at the bottom of the tube. The absorbance at 500 nm was recorded. A 2.0 mM solution of acetone in demineralised water was used as the standard, and demineralised water was used as the blind.

A recovery of 90 to 95 % was obtained, and a repeatability (s_r, se below for a description) of 0.03 mM was observed. All samples were tested in duplicate. Acetoacetate gives a misreading, since approximately 10 % of the acetoacetate is measured as acetone. This is no major problem as the levels of acetoacetate in milk are very low (0.1 mM on average) (1).

Infrared measurements

The FT-IR spectra of the samples were recorded in triplicate (some fortified samples only in duplicate) on three different Milko-Scan FT 120 instruments (Foss A/S, Hillerød, Denmark). The infrared spectra from 5000 to 925 cm^{-1} were recorded. The instruments

were standardised before use. All measurements were ratioed against water and log-transformed to give absorbance spectra.

In the data analysis only the spectral ranges 964-1581, 1697-1812 and 2699-2969 cm^{-1} were used, as these are the areas containing useful chemical information. No spectral pre-processing was performed prior to calibration, as experience shows that only minor improvements, if any, are obtained when using full spectrum FT-IR data (7).

Calculations

The data analysis and calibration was performed on a PC using Matlab 5.2.1 software (The MathWorks Inc., MA, USA). The calibration routines were either programmed by the author or taken from the PLS_Toolbox Version 1.5 (Eigenvector Technologies, WA, USA).

Repeatability is expressed as a mean standard deviation (s_r) of multiple determinations performed under identical conditions and is calculated as:

$$s_r = \sqrt{\frac{1}{q(n-1)} \sum_{j=1}^{q} \sum_{i=1}^{n} (x_{j,i} - \bar{x}_j)^2}$$

where q is the number of samples, n is the number of replicates, $x_{j,i}$ is the result of the i'th replicate of the j'th sample and $\bar{x}_j$ is the average result of the j'th sample.

Accuracy is expressed as the Root Mean Square Error of Prediction (RMSEP) and is calculated as:

$$RMSEP = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{i,reference} - x_{i,predicted})^2}$$

where N is the number of determinations (number of samples (q) times number of replicates (n) from above) and $x_{i,reference}$ and $x_{i,predicted}$ are the reference and predicted values corresponding to the i'th determination, respectively.

Correlation is expressed as R^2, which is calculated as:

$$R^2 = \left[\frac{\frac{1}{N} \sum_{i=1}^{N} (x_{i,reference} - \bar{x}_{reference})(x_{i,predicted} - \bar{x}_{predicted})}{s_{reference} s_{predicted}} \right]^2$$

where N, $x_{i,reference}$ and $x_{i,predicted}$ are defined above and $\bar{x}_{reference}$, $s_{reference}$, $\bar{x}_{reference}$ and $s_{predicted}$ are the means and standard deviations of the reference and predicted results, respectively.

RESULTS AND DISCUSSION

PLS calibrations

Calibrations were performed using three different data sets:

1. Natural samples only
2. Fortified samples only
3. The combined set of all samples

Validation on the calibration sets was per-

	Natural samples[3]			Fortified samples[4]			All samples[5]		
CVS[1]	#factors[2]	R^2	RMSEP	#factors[2]	R^2	RMSEP	#factors[2]	R^2	RMSEP
2	22	0.78	0.25	16	0.72	0.34	12	0.53	0.40
4	20	0.78	0.25	17	0.83	0.26	23	0.82	0.24
6	23	0.79	0.24	23	0.89	0.21	23	0.83	0.23
8	21	0.79	0.24	23	0.90	0.20	23	0.83	0.24
10	23	0.80	0.24	22	0.89	0.22	23	0.84	0.23

Table 1 Results of PLS calibration for acetone (in mM units) using cross validation with various numbers of cross validation segments on the calibration set.
[1]CVS = Cross Validation Segments
[2]#factors = optimal number op PLS factors
[3]Results from a calibration based on 302 natural individual cow samples from Norway, Sweden and Denmark
[4]Results from a calibration based on 125 acetone-fortified individual cow samples from Denmark
[5]Results from a calibration based on the combined set of 427 samples

formed using cross validation with 2, 4, 6, 8 and 10 systematically selected cross validation segments. The results are shown in Table 1.

The cross validated results show that the natural samples give the most consistent results, as the number of PLS factors and the figures of merit are very similar, no matter how many cross validation segments that are used. The fortified samples show a tendency to give an increasing number of factors when more segments are used, which could suggest an overfit (*i.e.* the model fits the calibration samples very well, but not future samples). Even when combined with the natural samples, the fortified samples tend to deteriorate the model.

The poor results when the fortified samples are used could be due to the narrower range of acetone results in this set, or because of the lower number of samples when compared to the set of natural samples. But this is not supported by the results from the combined set of all samples. The results on the combined set are in some cases worse than what is obtained when using the natural *or* the fortified samples alone.

Based on the cross validation results in Table 1, the optimal number of PLS factors for each of the three calibration sets were chosen to be the average number for each set. The resulting PLS models were tested using the test set consisting of 58 selected samples from New Zealand. The results of these tests are presented in Table 2 and Figure 2. That fortified samples are not appropriate for calibration becomes even clearer here: The PLS model based on fortified samples is not able to predict acetone in the test samples. The model based on the combined set provides an average result. Only the model based on natural samples is able to predict the test samples with an error similar to the result obtained during calibration. Thus, the fortified samples do not contain the same acetone information as the natural ones, and they cannot be used for calibration. The most likely reason for this is that spectral information from acetoacetate and β-hydroxybutyrate may facilitate an acetone determination, and they are only present in the natural samples.

However, the models are still based on acetone information to a very large extent. This is evident from Figure 3 where the regression coefficients (the spectral weighting of the model, the so-called B coefficients) are plotted with the spectrum of pure acetone in water. The four most positive contributions among the regression coefficients correspond almost exactly to the four major peaks in the acetone spectrum. Thus, in spite of the very high complexity of the PLS models (seen as a high number of PLS factors), it is not some spurious correlation which has been found – the calibration *is* actually detecting acetone.

The use of FT-IR spectroscopy for determining acetone described here has been patented (8).

Samples	#factors[1]	R^2	RMSEP	s_r
Natural	21	0.81	0.27	0.12
Fortified	20	0.42	0.57	0.18
All	20	0.78	0.32	0.15

Table 2 Results of PLS calibration for acetone (in mM units) on the test set containing 58 samples from New Zealand.
[1]#factors = optimal number op PLS factors. They were selected as the average of the optimal numbers obtained through cross validation (Table 1).

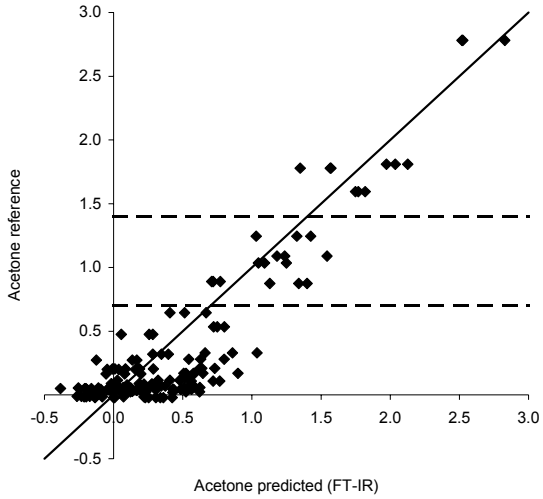


Figure 2 Result of the prediction of acetone in the test set (58 samples) using a 21-factor PLS model based on natural samples. The horizontal lines indicate the proposed limits at 0.7 and 1.4 mM (6). Figures of merit: correlation (R^2) 0.81, accuracy (RMSEP) 0.27 mM, and repeatability (s_r) 0.12 mM.

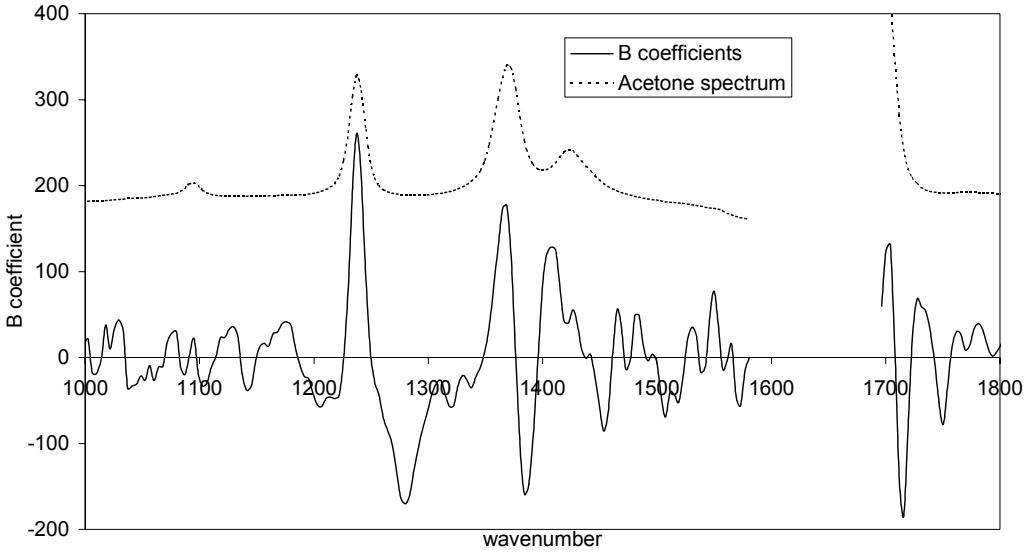


Figure 3 Regression (or B) coefficients from the 21-factor PLS model for acetone in milk based on natural samples plotted with the acetone spectrum in the range from 1000 to 1800 cm^{-1}.

Classification

The result of the prediction of acetone in the test set was an accuracy (RMSEP) of 0.27 mM. This corresponds to 95 % confidence limits of ±0.5 mM around a given true value. This means that the two limits of 0.7 mM ("possibly ill") and 1.4 mM ("ill") will be translated, when using FT-IR, into the intervals:

below 0.2 mM: The cow belongs to the group below 0.7 mM ("well")
between 0.2 and 0.9 mM: The cow might belong to one of the groups below 1.4 mM ("well" or "possibly ill")
above 0.9 mM: The cow might belong to the group above 1.4 mM ("ill").

It is clear that it will be difficult to distinguish between "possibly ill" and "ill" using FT-IR. The method will therefore only be useful for screening, *i.e.* to answer the question whether a given milk sample should be taken out for further analysis using the reference method or a FIA-based method. When using FT-IR for screening, a limit different from the ones mentioned above should be used. The actual value must be chosen according to how many false positives and false negatives that can be accepted.

An example of how the limit could be found is shown in Table 3, where the numbers of samples falling within the wrong group are shown, using limits from 0.2 to 0.7 mM. With these limits, false negatives are never encountered. The reason for this is the tendency to overestimation of low-acetone samples and underestimation of high-acetone samples, which is an undesired side-effect of the PLS method in cases where the major part of the error is on the indirect (infrared) method (10). This effect is clearly visible in Figure 2. In the present case it will therefore be sufficient to use an acetone limit of 0.7 mM for the FT-

Criterion	IR[1]>0.2	IR[1]>0.3	IR[1]>0.4	IR[1]>0.5	IR[1]>0.6	IR[1]>0.7
Samples selected (% of total)	93 (53 %)	78 (45 %)	68 (39 %)	58 (33 %)	45 (26 %)	37 (21 %)
False positives (%)	66 (71 %)	51 (65 %)	41 (60 %)	31 (53 %)	18 (40 %)	10 (27 %)
False negatives (%)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)

Table 3 Classification of the 58 test samples. The samples were split into two classes using the reference method: Above 0.7 mM acetone = positive. Below 0.7 mM acetone = negative. Every sample was tested in triplicate thus giving 174 determinations to distribute between the two groups.
[1]IR = infrared prediction of acetone using the 21-factor model based on natural samples

IR results. Using this limit, 27 % of the samples selected would be false positives. With the limited accuracy in mind this is a relatively low number, which suggests that the risk of overlooking real positive samples is high, although no false negatives were encountered in the present test. Thus, in some cases it might be preferred to use a limit lower than 0.7 mM.

CONCLUSIONS

The present studies show that it is possible to determine acetone in milk in the range from 0.0 to 2.8 mM using FT-IR spectroscopy combined with multivariate calibration. The obtained accuracy was 0.27 mM (a relative error of approximately 20 %) which is useful for screening purposes only, *i.e.* the results on acetone-containing samples should be confirmed by a more accurate method. As acetone is the primary marker for ketosis in the cow's milk, this can be used for finding cows that might be ill.

Only natural samples could be used for calibration. Calibrations based on samples fortified with acetone did not work when tested on samples from cows suffering from ketosis.

The classification power of the method was tested. The accuracy was too limited to allow for a classification into three groups ("well", "possibly ill" and "ill"), but it was possible to select samples from possibly ketotic cows with a relatively low number of false positives (27 %).

ACKNOWLEDGMENTS

The author would like to thank The Danish Academy of Technical Sciences and Foss A/S, Denmark, for providing the funding for this work. It was carried out in co-operation with Food Technology at the Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Denmark. Dairying Research Corporation Ltd. (Norman Thomson and Gwyneth Verkerk) and New Zealand Dairy Research Institute are thanked for providing samples and equipment for the New Zealand work. Pernille Bayer, Lilian J. Nicolaisen and Inger Tholl are thanked for technical assistance.

REFERENCES

1 Andersson, L. 1984. Concentrations of blood and milk ketone bodies, blood isopropanol and plasma glucose in dairy cows in relation to the degree of hyperketonaemia and clinical signs. Zentralbl. Veterinaermed. Reihe A 31:683-693.

2 Andersson, L., A. H. Gustafsson, and U. Emanuelson. 1991. Effect of hyperketonaemia and feeding on fertility in dairy cows. Theriogenology 36:521-536.

3 Drewes, P. A. 1974. Pages 1354-1364 *in* Clinical chemistry: Principles and techniques. 2nd ed. R. J. Henry, D. C. Cannon and J. W. Winkelman, ed. Harper & Row, Hagerstown.

4 Geishauser, T., K. Leslie, D. Kelton, and T. Duffield. 1998. Evaluation of five cowside tests for use with milk to detect subclinical ketosis in dairy cows. J. Dairy Sci. 81:438-443.

5 Gustafsson, A. H., L. Andersson, and U. Emanuelson. 1993. Effect of hyperketonaemia, feeding frequency and intake of concentrate and energy on milk yield in dairy cows. Anim. Prod. 56:51-60.

6 Gustafsson, A. H., and U. Emanuelson. 1996. Milk acetone concentration as an indicator of hyperketonaemia in dairy cows: the critical value revised. Anim. Sci. 63:183-188.

7 Hansen, P. W. 1998. Urea determination in milk using Fourier transform infrared spectroscopy and multivariate calibration. Milchwissenschaft 53:251-255.

8 Hansen, P. W., inventor. 1998. Measurement of acetone in milk using IR spectroscopy. Foss Electric A/S, assignee. Int. Pat. Appl. No. WO 98/43070.

9 Marstorp, P., T. Anfält, and L. Andersson. 1983. Determination of oxidized ketone bodies in milk by flow injection analysis. Anal. Chim. Acta 149:281-289.

10 Martens, H., and T. Næs. 1989. Multivariate calibration. Wiley, Chichester.

11 Steger, H., H. Girschewski, B. Piatkowski, and J. Voigt. 1972. Evaluation of the state of ketosis in lactating cows by estimating the concentration of ketones in blood and that of acetone in milk. Arch. Tierernaehr. 22:157-162.

Publication IV

Measurement of Acetone in Milk Using IR Spectroscopy

International Patent Application No. WO 98/43070

Measurement of Acetone in Milk Using IR Spectroscopy

Per Waaben Hansen
R&D Greenhouse, Foss Electric A/S,
Slangerupgade 69, P.O. Box 260, DK-3400 Hillerød, Denmark

ABSTRACT

A method for calibration of an IR spectrometry apparatus for providing and evaluating IR spectra in order to determine very low contents of specific components in a fluid, such as milk, and especially low contents of acetone, in a measuring range above 0, such as from 0.5 mM to 2.0 mM acetone in milk, using at least 50, such as from 50 - 300 known samples including at least 10 samples representing the fluid without any substantial content of the specific component for the calibration.

Preferably, a good calibration for acetone shall be based on spectral information including the spectral ranges 1712-1697, 1419-1396, 1378-1353 and 1249-1226 cm^{-1}, or at least a substantial portion/part of said ranges.

By use of the new calibration of an FT-IR-spectrometry apparatus it will be possible to determine the acetone content during the same IR measurement process used for determining other milk parameters such as fat and protein. In a similar way other small contents of a specified component in a fluid can be determined by use of a method according to the invention.

The present invention relates to measurement of small amounts of a specified component in a fluid, and especially acetone and/or acetoacetate in milk.

TECHNICAL BACKGROUND

It has for some time been desirable to be able to measure the content of acetone in milk in order to have a tool for an early detection of ketosis (a metabolic disease) in dairy cows. Methods for the determination are in fact available, but generally they are time-consuming. Besides the content of acetone there will typically be a related amount of acetoacetate. When ever in this specification the word acetone appears this means acetone and/or acetoacetate.

Recently IR spectrometry has become a preferred method for analysing milk, and accordingly it would be advantageous also to use IR spectrometry for finding the content of acetone and/or acetoacetate.

The application of IR spectral data for determination of concentrations of components in a composition is known *e.g.* from: WO 95/16201 (Foss Electric A/S), WO 96/24832 (Foss Electric A/S), US 5121337, (Brown), US 5252829, (Nygaard *et al.*), US 5606164 (Price et al) and EP 0751388 (Kyoto Dai-Ichi).

From WO 95/16201 it is known to determine added water and the related freezing point depression from IR spectral data. From this document it is also known to be advantageous to increase the leverage of the known calibration samples by adding extra water to natural samples. US 5121337 (Brown) discloses a method for correcting spectral data for data due to the spectral measurement process itself. Further it discloses how to estimate an unknown property and/or composition data of a sample by use of such method. US 5252829 (Nygaard), owned by the applicant, discloses a method of determining urea in milk. The content of urea in milk is generally above zero but fairly low, and the successful urea determination described in the patent is based on thorough compensation for the influence by other components on the urea

measurement, through use of contemporary determinations of the contributions from the other components, *i.e.* fat, lactose and protein. The samples used for calibration are generally samples having a considerable amount of urea, *i.e.* within the intended measuring range. US 5606164 (Price *et al*.) discloses a method and apparatus for biological fluid analyte concentration measurement using generalised distance outlier detection.

However, it has until now been considered impossible to use IR spectrometry for obtaining a reliable detection of the amounts of acetone appearing in cows suffering from ketosis, *cf*. Hendrik-Jan Luinge, B. Lutz, P. Dobbelaar and Y. H. Schukken: "Infrared spectrometry as a sensor for the early detection of ketosis in cows", 1996. Poster presentation, S.O.N. Analytische Chemie, Lunteren, Nov. 6 - 7, 1995.

The present invention provides a method for calibration of an IR spectrometry apparatus for providing and evaluating IR spectra in order to determine very low contents of specific components in a fluid, such as milk, and especially low contents of acetone, such as from about 0.5 mM to about 2.5 mM acetone in milk.

The art of extracting information on the chemical content of fluid compositions from measured spectra has for some time been based on a process of "learning" or "calibrating" the IR spectrometry apparatus to enable the data processor of the apparatus to recognise certain components in a fluid. Normally, the spectra of about 15 - 20 known samples are measured and used together with the known values of the content in the samples to derive a calibration for the IR spectrometry apparatus. It is general knowledge to people in the art that the set of samples used for calibration must be representative for the desired range of measurements. There are several methods of calculating such calibrations, and many methods (such as PCR, MLR or PLS regression) are well known to people in the art. Hitherto the methods and apparatus available have generally only been able to provide reliable measurements of contents which happen to appear in substantial amounts, *i.e.* that makes up a considerable fraction of the fluid, such as fat, protein and lactose in milk. In the case of acetone and/or acetoacetate the content in milk is generally zero or about zero.

SUMMARY OF THE INVENTION

The present invention provides a method for calibration of a spectrometry apparatus for providing and evaluating spectra for determination of very low concentrations of a specific component in a fluid in a specified measuring range above 0, (*e.g.* from 0.5 mM to 4.0 mM), by which calibration method a number of variables and corresponding coefficients (so-called B coefficients) are determined according to methods for multivariate calibration, such as PCR, MLR or PLS regression, comprising selecting and measuring a set of calibration samples including at least 30 - 50 known samples, *e.g.* from 50 - 300 known samples, for the calibration. According to the invention the set of calibration samples includes a number of samples representing the fluid without any substantial content of the specific component for the calibration, *i.e.* samples being below the specified measuring range.

The new method is based upon the use of a great number of known samples, which are measured by a spectrometry apparatus, providing a spectrum of each of the known samples (the content of the "known" samples either being known or determined by a reference method) and providing a calibration (*e.g.* by applying known calibration calculation methods such as PCR, MLR or PLS regression and by applying principal variables or genetic algorithms for variable selection).

The method was specifically developed for the determination of acetone in milk by use of IR spectroscopy. However, it is contemplated that the method is applicable to other types of spectroscopy for the determination of small amounts of other specific components.

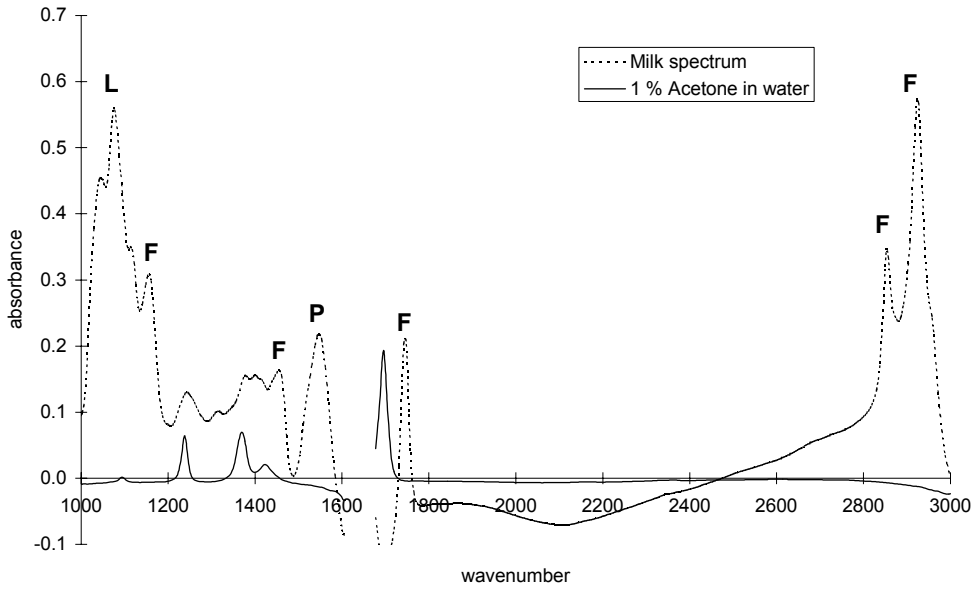


Figure 1 A measured IR spectrum of a milk sample with acetone and an IR spectrum of 1 % acetone in water.

Regarding the acetone and/or acetoacetate content, experience has indicated that a substantial number, *e.g.* about 20 - 30, and even better about 100 or 150, *i.e.* the vast majority of the calibration set (the known samples) may represent samples having none or almost no content of acetone.

The method according to the invention is advantageous because the vast majority of available samples are samples having no or almost no content of acetone. Only the few cows suffering from ketosis will provide samples which are representative for the range to be determined in order to be able to decide whether a cow suffer from ketosis or not. The most obvious solution to that problem would be to enrich a great number of natural samples with suitable amounts of acetone to provide a good calibration set.

According to the inventor's experience such extended enrichment is not necessary. A few enriched samples and/or natural samples from cows suffering from ketosis will do, and the vast majority of the calibration samples can be natural samples from healthy cows, *i.e.* samples without any significant amount of acetone, and samples outside the desired measuring range. In this way a fairly accurate determination is possible. Further the provision of the calibration set is fairly easy, as the vast majority of samples may be natural samples.

According to a further advantageous method some of the calibration samples may be enriched samples, *i.e.* samples having zero or almost zero content of the specified component, whereto a number of predetermined, known amounts of the specified component being added. Accordingly a preferred set of calibration samples comprises a great number of samples having almost no content of the specified component, and a small representative selection of samples covering the intended measuring range.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a measured IR spectrum of a milk sample with acetone and an IR spectrum of 1 % acetone in water.

Figure 2 shows measurement results using a full spectrum PLS model.

Figure 3 shows the spectral weighting of the calibration used in Figure 2.

Figure 4 shows measurement results using 100 samples and a full spectrum PLS model.

Figure 5 shows measurement results using 76 samples and a full spectrum PLS model.

Figure 6 shows measurement results using 39 samples and a full spectrum PLS model.

Figure 7 shows measurement results using 171 samples, and a reduced spectrum PLS model.

Figure 8 shows measurement results using 76 samples and a full spectrum PLS model.

DETAILED DESCRIPTION OF THE INVENTION

The method according to the invention will be explained in further details by use of examples based on measuring the content of acetone in milk. It should be emphasised that the method according to the invention in its broadest aspect also can be used for measuring other components appearing in very low concentrations in a fluid.

Acetone has a characteristic IR spectrum that appears from Figure 1 showing the IR spectrum of 1 % acetone in water. A few specific bands appear from the spectrum: at 1696, 1423, 1370, and 1238 cm^{-1}. For the sake of good order it is mentioned that the very high signals from about 1620 to 1670 cm^{-1} are caused by water.

The concentrations of acetone appearing in milk are very low. The normal level is about 0 mM, and in case of a cow suffering from ketosis the level might reach up to about 3 - 4 mM.

For diagnostic purposes the following acetone limits are recommended by Anders H. Gustafsson in report 222 from the Swedish University of Agricultural Sciences, Department of Animal Nutrition and Management: "Acetone and Urea Concentration in milk as indicators of the nutritional status and the composition of the diet of dairy cows":

< 0.7 mM: The cow does not suffer from ketosis

0.7-1.4 mM: The cow may be ill, *i.e.* milk yield may be lowered

> 1.4 mM: The cow is ill, *i.e.* milk yield is reduced by 10 - 20 %

It is mostly in the 3rd - 6th week of lactation (suckling) that the cow is liable to get ketosis, due to a very high yield in this period.

From these figures it appears that milk from a normal cow (healthy cow) has an acetone content of about 0 mM. The important measuring range extends from about 0.5 mM to about 3-4 mM, *i.e.* the range that will allow an accurate diagnosis of cows suffering from ketosis. Accordingly the term "without any substantial content" used in claim 1 shall by understood as a content below the intended measuring range. In the case of acetone this means that when given the relevant measuring range extending from 0.5 to 4 mM of acetone, the calibration samples may include a number of samples having less than about 0.1 mM or 0.2 mM.

The method according to the invention comprises:

1. collecting at least 30 - 50 single-cow milk samples, preferably about 50 - 300 samples;

2a. selecting at least 5 - 10 and preferably about 15 - 30 samples from cows being in the 1st - 6th week of lactation; and/or

2b. selecting at least 5 - 10 and preferably about 15 - 30 samples for enrichment with acetone, in order to provide a representative set of samples including samples having up to about 3 - 4 mM of acetone, (2b may be preferred as this will surely provide the desired representative set of samples);

3. measuring all samples by a reference method (*i.e.* all samples to be used for the calibration);

4. measuring all samples by use of the IR spectrometry apparatus to be calibrated;

5. entering the measurement results into a data processing unit arranged to (programmed to) calculate a calibration, *i.e.* to make a selection of wavebands and calculate the so-called B coefficients for the selected wavebands;

6. entering the calibration into the IR spectrometry apparatus to be calibrated.

A careful selection of single cow samples for the calibration is very important to the final result. The number of samples shall preferably be from 50 - 300 known samples for the calibration. A considerable part of the samples may have zero (0) or almost zero (0) content of acetone. A second important part of the samples shall represent the specified measuring range for the content of acetone. This part may include a selection of natural samples *i.e.* samples from cows having ketosis. According to the inventor's experience also a selection of enriched samples can be used, as well as a mixture of natural samples and enriched samples.

At least 10 - 20 samples may be enriched samples having a generally uniformly distributed variation of values covering the specified measuring range, and preferably covering more than the specified measuring range. The use of enriched samples can be an advantage in order to ensure that the calibration set includes a representative variation of the acetone content. A further advantage is that by using samples enriched by a known amount of acetone a corresponding reference measurement may be dispensed with.

In the following a number of examples will illustrate the measurement results that can be achieved by using the method according to the invention. 171 single-cow samples were measured. 20 samples were enriched by acetone. Each of the samples were measured by a FIA (flow injection analysis) reference method and three times by use of a Foss Electric MilkoScan FT 120, a FT-IR instrument using Fourier transform infrared technique, provided with a 37 µm cuvette. The signal to noise ratios for the FT 120 were determined at the most important wavelengths and the results are stated in Table 1 below.

The signal to noise ratios (S/N) shown in Table 1 were determined in the following way: Eight samples are measured three times each, with 20 seconds measuring time in a 37 µm sample cell. The transmittance is calculated on a water background. The resolution is 12 cm^{-1} measured as FWHH (Full Width Half Height). The RMS (Root Mean Square) noise on the three determinations is calculated as the standard deviation on the transmittance at the relevant wavenumbers. The total noise is then calculated as the RMS value for the eight samples. The signal to noise value is then calculated as the average transmittance for the 24 spectra at the relevant wavenumber divided by the total noise at the same wavenumber, *i.e.*:

$$S/N = \frac{\langle T_{wavenumber} \rangle}{\sigma_{wavenumber,sample}}$$

The signal to noise ratio (S/N) of the FT-IR apparatus in use will be important to the accuracy and reliability of the obtained measurements. A S/N of at least 500 and preferably at least 1000, and more preferably at least 1500 is considered to be an important parameter for the performance of the method. The rather low S/N in the last column is due to the fact that the wavenumber, 1700, is close to the water band, which gives rise to a high signal, but little information on the content of acetone.

As a first example the full spectrum – except for the water bands and a few other insignificant bands – is used. The result shown in Figure 2 is based on a full spectrum PLS model, using 17 factors. The validation is

cm^{-1}	1238	1365	1407	1700
S/N	1833	1820	1267	508

Table 1 The signal to noise ratios for the FT 120 determined at the most important wavelengths

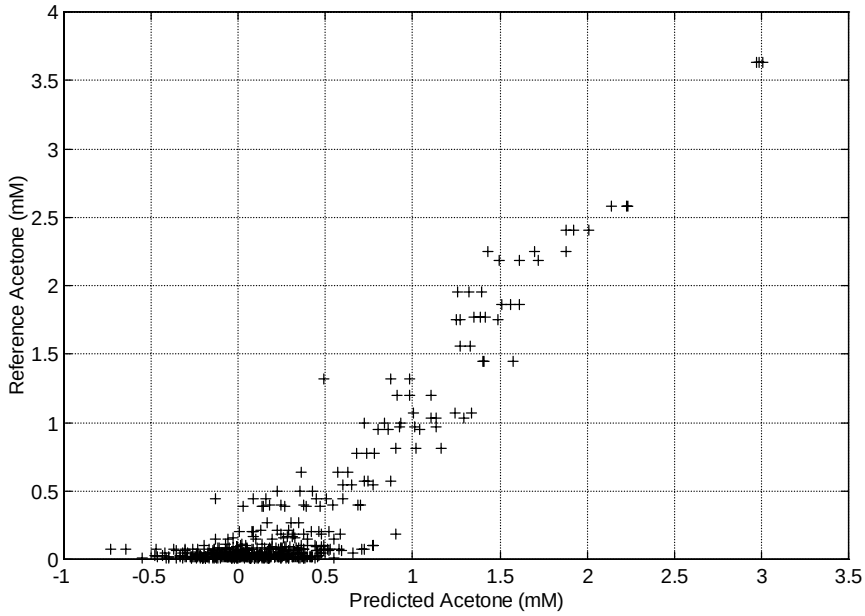


Figure 2 Measurement results using a full spectrum PLS model.

based on 6 cross validation segments. The "cross validation" is a normal validation procedure - used for testing the calibration. A fraction *e.g.* 5/6 of the samples are used for calibration, and the rest (the remaining segment, 1/6) is used for a validation to control whether the values obtained by use of the new calibration are comparable and preferably equal to or close to the measured reference values. The procedure is repeated 6 times - each time leaving a new segment of samples for the validation. The same type of 'cross validation' is used in all examples to follow later in this description.

Figure 2 and all the following Figures 4 – 7 and 8 show the acetone content measured by the IR-method versus the acetone content measured by the reference method. All IR-measurements are repeated three times, so for each sample three measurement points appear as '+'. As it appears from Figure 2 most measurement results are located in the vicinity of (0,0), in agreement with the fact that

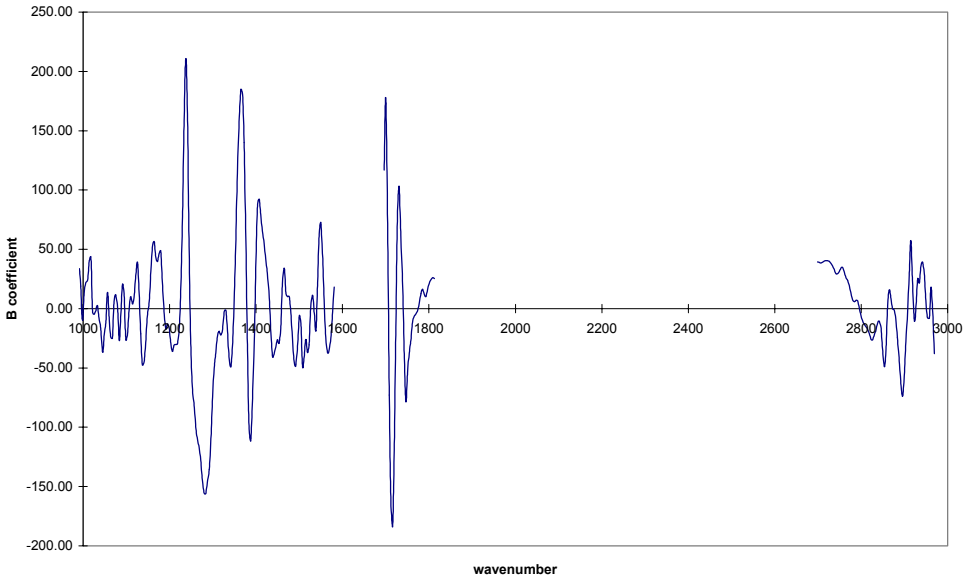


Figure 3 The spectral weighting of the calibration used in Figure 2.

No. of samples	No. of PLS factors	RMSEP	REP	Figure
171	17	0.26	0.13	2
100*	18	0.27	0.13	4
76*	19	0.28	0.13	5
76***	18	0.30	0.13	8
39**	15	0.46	0.14	6

Table 2: * The samples removed were selected randomly among all samples with an acetone content of less than 0.2 mM. ** The samples removed were all samples with an acetone content of less than 0.1 mM. *** The samples removed from the full dataset were selected ramdomly among all samples

most cows are healthy, not suffering from ketosis.

The calibration – which was used for obtaining the measurements shown in Figure 2 – is shown in Figure 3 as the spectral weighting of the calibration. The peaks appearing at 1700, 1407, 1365 and 1238 cm^{-1} correspond approximately, *i.e.* within the resolution of the MilkoScan FT 120, to the bands appearing in Figure 1.

From Figure 2 it appears that the few samples having a substantial content of acetone appear in close relation to a straight line from the origin, (0,0), to the point (3.0,3.6), indicating a close relation between the content measured through IR spectrometry and the content measured by the reference method. Accordingly, the acetone content in milk can be determined from IR spectrometry measurements using a calibration according to the present invention.

The performance of the method according to the invention is tested or evaluated by calculating repeatability and accuracy, defined as stated below:

Repeatability (REP) is stated as a mean standard deviation (s_r) of multiple determinations performed under identical conditions and is calculated as:

$$s_r = \sqrt{\frac{1}{q(n-1)} \sum_{j=1}^{q} \sum_{i=1}^{n} (x_{j,i} - \bar{x}_j)^2}$$

where q is the number of samples, n is the number of replicates, $x_{j,i}$ is the result of the i'th replicate of the j'th sample and $\bar{x}_j$ is the average result of the j'th sample.

Accuracy is stated as the Root Mean Square Error of Prediction (RMSEP) and calculated as:

$$RMSEP = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{i,reference} - x_{i,predicted})^2}$$

where N is the number of determinations (number of samples (q) times number of replicates (n) from above) and $x_{i,reference}$ and $x_{i,predicted}$ are the reference and predicted values corresponding to the i'th determination, respectively.

The Root Mean Square Error of Prediction (RMSEP) is found to be 0.26 mM, and the repeatability, REP is 0.13 mM (*i.e.* s_r = standard deviation on multiple determinations of the same sample; in the present case all samples are measured three times on the FT-IR instrument). This result indicates that the performance is sufficient in respect to the before-mentioned diagnostic threshold values at 0.7 mM and 1.4 mM.

Further, the results shown in Figure 2 indicate that even with RMSEP = 0.26 mM it is likely that all cows measured by the method according to the invention will be classified correctly as either healthy, perhaps ill, or ill, according to the criteria mentioned before.

The following examples illustrate the reliability of the measurements in respect to different conditions for the selection of calibration samples. Table 2 shows the data of five examples.

It is well known that a representative number of samples are needed for obtaining a reliable calibration model. The requested number will

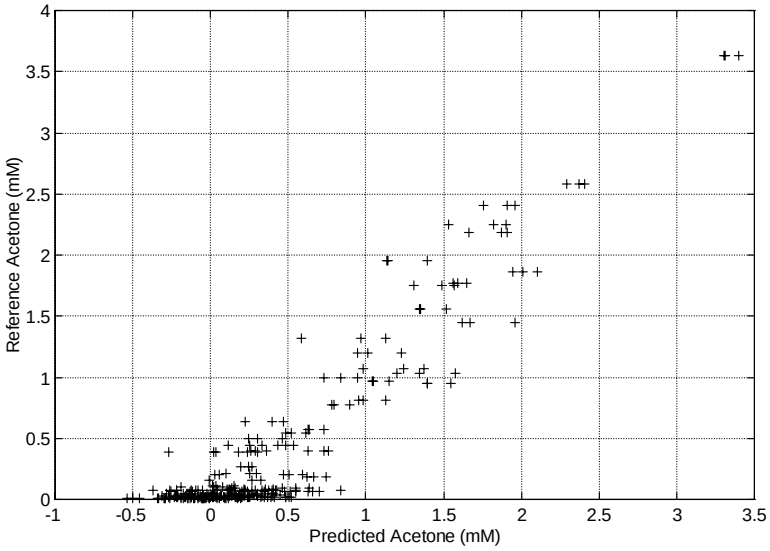


Figure 4 Measurement results using 100 samples and a full spectrum PLS model.

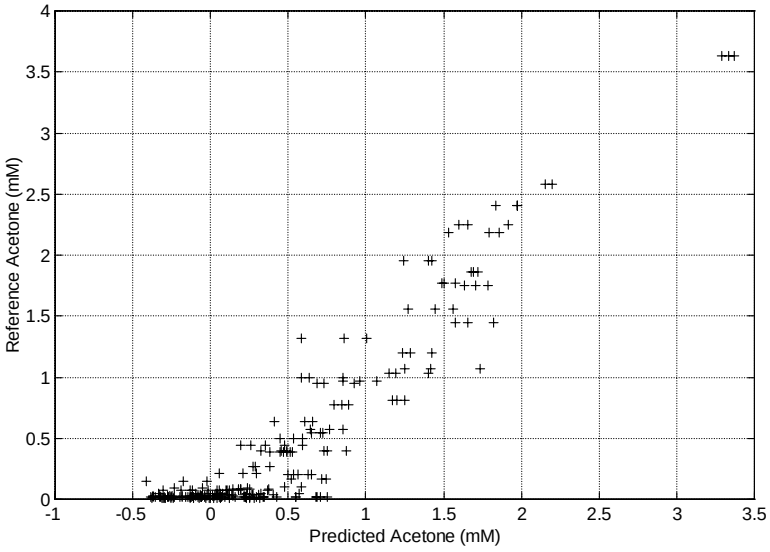


Figure 5 Measurement results using 76 samples and a full spectrum PLS model.

however depend on the type of measurement. A determination of very small amounts of acetone will require many samples. How many is investigated in the following examples referring to Table 2 and Figures 2, 4 - 6 and 8.

In the first example a total of 171 samples was applied for the calibration. 17 PLS factors were found providing measurement results having RMSEP = 0.26 and REP = 0.13. The diagram in Figure 2 illustrates this example. As it appears from Figure 2 the vast majority of the samples have almost zero content of acetone. About 20 samples represent the desired measuring range from about 0.5 to about 2.5 mM.

The next example (line 2 of Table 2) relates to the use of 100 samples. These were chosen by removing 71 randomly selected samples, having an acetone content of less than 0.2 mM. Figure 4 shows a full spectrum PLS model, using the 100 remaining samples leading to 18 PLS factors and measuring results having RMSEP: 0.27 and REP: 0.13. As it appears from Figure 4 the vast majority of the samples still have almost zero content of acetone. Again about 20 samples represent a

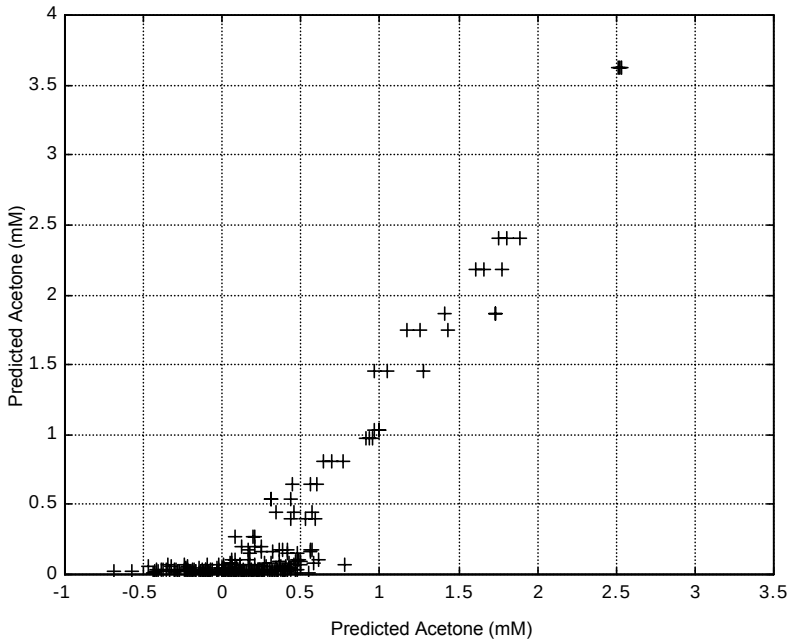


Figure 8 Measurement results using 76 samples and a full spectrum PLS model.

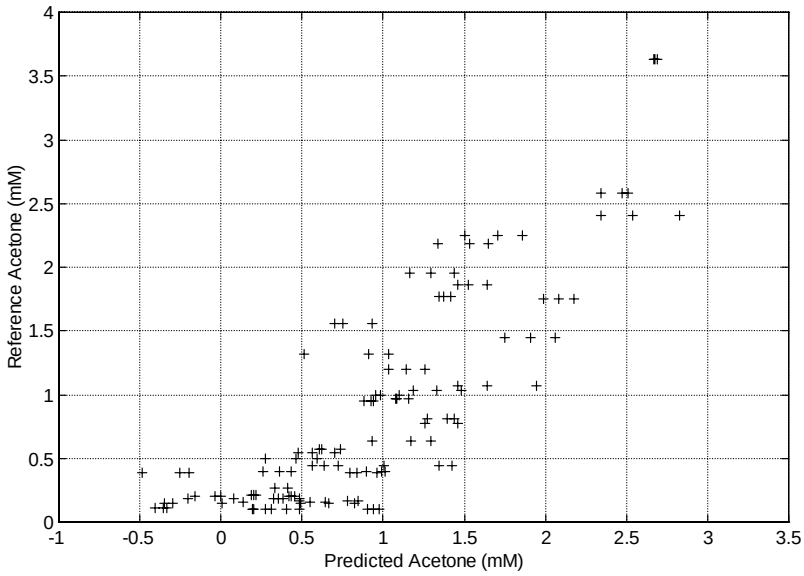


Figure 6 Measurement results using 39 samples and a full spectrum PLS model.

desired measuring range from about 0.5 to about 2.5 mM.

The third example (line 3 of Table 2) relates to the use of 76 samples. These were chosen by removing 95 samples randomly selected among the samples having an acetone content of less than 0.2 mM. Figure 5 shows a full spectrum PLS model using the remaining 76 samples leading to 19 PLS factors and measuring results having RMSEP: 0.28 and REP: 0.13. As it appears from Figure 5 a majority of the samples (about 44) have almost zero content of acetone. About 20 samples represent the desired measuring range from about 0.5 to about 2.5 mM.

The fourth example (line 4 of Table 2) also relates to the use of 76 samples. These samples were chosen by randomly removing 95 samples from the total set of 171 samples. Figure 8 shows a full spectrum PLS model using the remaining 76 samples leading to 18 PLS factors and measuring results having RMSEP: 0.30 and REP: 0.13. As it appears from Figure 8 the vast majority (about 56) of

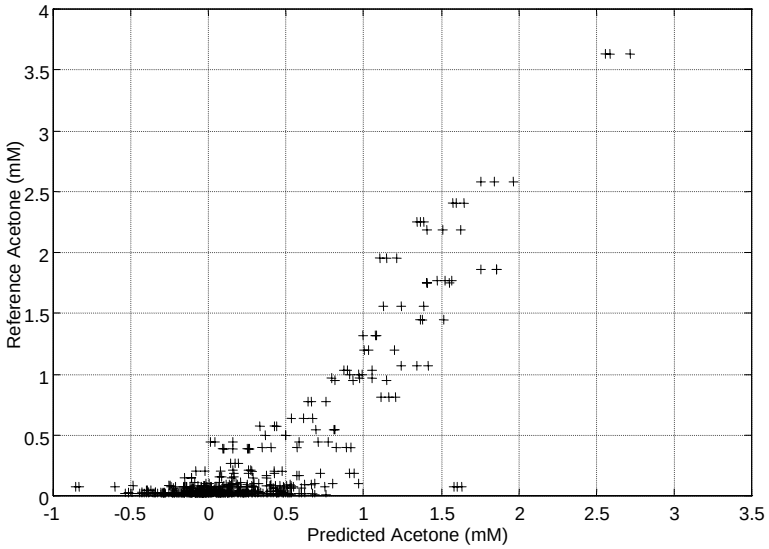


Figure 7 Measurement results using 171 samples, and a reduced spectrum PLS model.

the samples have almost zero content of acetone. Only about 10 samples represent the desired measuring range from about 0.5 to about 2.5 mM. Surprisingly, the accuracy seems to be satisfactory, almost as good as in example 3.

The fifth example (line 5 of Table 2) relates to the use of 39 samples. These were chosen among the 171 samples in example 1 by removing 132 samples randomly selected among samples having an acetone content of less than 0.1 mM. Figure 6 shows a full spectrum PLS model using the remaining 39 samples leading to 15 PLS factors and measuring results having RMSEP: 0.46 and REP: 0.14. As it appears from Figure 6 none of the samples has almost zero content of acetone. For about 18 samples the acetone content was from 0.1 to 0.5 mM. About 20 samples represent the desired measuring range from about 0.5 to about 2.5 mM. It is obvious from the Figure 6 that the results are more scattered. The calculated RMSEP indicates that the calibration set is too poor. Accordingly, a representative set of 20 samples covering the desired measuring range is not sufficient. A great number of samples without any or almost any content of acetone clearly contribute to improve the accuracy of the determination.

From Figures 4 - 6 and 8 it appears that 100 samples or even 76 samples including about 20 samples, representing the desired measuring range, lead to a calibration being almost as good as the first one using 171 samples shown in Figure 2 having RMSEP: 0.26 and REP: 0.13. Accordingly, the calibration set of samples may contain as few as about 10-20 representative samples and the remaining samples may have no or almost no content of acetone. Specifically, the fifth example indicates that the removal of the samples having the smallest amount of acetone results in very poor accuracy. Figure 2 clearly shows that the vast majority of calibration samples have an acetone concentration close to 0, and below 0.1 mM.

The Acetone spectrum of Figure 1 and the calibration shown in Figure 3 indicate that a good calibration for acetone shall preferably be based on spectral information including the spectral ranges 1712-1697, 1419-1396, 1378-1353 and 1249-1226 cm^{-1}, or at least a substantial part of said ranges. Also the range 1299-1276 cm^{-1} obviously add important information according to the large negative peak appearing among the calculated B coefficients shown in Figure 3.

Figure 7 shows an example by which two of the said ranges were removed: The waveband 1712-1697 and 1419-1396 cm^{-1}, leaving the waveband ranges: 1378-1353 and 1249-1226 cm^{-1}. The result was a RMSEP of 0.32, and REP of 0.11. It is the experience of the

inventor that at least two of the waveband ranges in question should be used in order to obtain a reliable measurement indicating whether the cow suffers from ketosis. It is preferred to use all the waveband ranges mentioned.

The method is specifically intended for measuring very small amounts of a component in a liquid. For acetone it is specifically interesting to know whether the acetone content in a milk sample is above or below 0.7 mM. Accordingly you may say that the really important measuring range is from about 0.5 to about 4 mM. Accordingly, you would expect from the knowledge of the prior art that a representative selection of known samples ought to be selected among samples in the range from about 0.5 to 4 mM. The inventor of the present method has realised the surprising fact that a great number of samples having less than 0.1 mM acetone apparently has a significant influence upon the quality of the calibration (when looking at the Figures 2, 4, 6, 7, 8). Here, it shall be kept in mind that the number of samples in the relevant measuring range from 0.5 mM to 4 mM are the same in the examples 1, 2, 3 and 5. Nevertheless, example 5 shows a significant decrease in accuracy compared to the examples 1, 2 and 3. The only difference is that a great number of samples having less than 0.1 and less than 0.2 mM were included in the examples 1, 2 and 3.

PATENT CLAIMS

1. A method for calibration of a spectrometry apparatus for providing and evaluating spectra for determination of very low concentrations of a specific component in a fluid in a measuring range above 0, (*e.g.* from 0.5 mM to 4.0 mM), by which calibration method a number of variables and corresponding coefficients (so-called B coefficients) are determined according to methods for multivariate calibration, such as PCR, MLR or PLS regression, the method comprising selecting and measuring a set of calibration samples including at least 30 - 50 known samples, *e.g.* from 50 - 300 known samples, for the calibration, *characterised in* that the set of calibration samples includes a number of samples representing the fluid without any substantial content of the specific component for the calibration, *i.e.* samples being below the measuring range.

2. A method according to claim 1, *characterised in* that the set of calibration samples includes at least 10 samples, and preferably at least 30, more preferably at least 50 known samples having 0 and/or almost 0 content of the specified component, such as acetone, for the calibration.

3. A method according to claim 1 or 2, *characterised in* that the set of calibration samples includes a majority of known samples having 0 and/or almost 0 content of the specified component, such as acetone, for the calibration, as well as at least 5 samples, preferably at least 10 samples and more preferably at least 20 samples representing the measuring range for the content of the specified component.

4. A method according to claim 3, *characterised by* at least some of the at least 10 samples representing the measuring range for the content of the specified component being enriched samples, *i.e.* samples, whereto an amount of the specified component being added.

5. A method according to claim 4, *characterised in* that the amounts of the specified component being added are predetermined known amounts.

6. A method according to any of the claims 1 - 5, *characterised by* using at least 17 specific wavebands for the determination of the content of the specified component, such as acetone.

7. A method according to any of the preceding claims, wherein the fluid is milk and the specified component is acetone, and the spectra are IR spectra, *characterised by* using at least two wavebands from a group of wavebands each comprising or being close to one of the

following wavenumbers: 1696, 1423, 1370 and/or 1238 cm^{-1}.

8. A method according to any of the claims 1-6, wherein the fluid is milk and the specified component is acetone, *characterised by* using waveband ranges substantially comprising at least two of the following waveband ranges: 1712 - 1697, 1419 - 1396, 1378 - 1353, 1249 - 1226 cm^{-1} and 1299 - 1276 cm^{-1} or parts thereof.

9. A method according to any of the claims 1-6, wherein the fluid is milk and the specified component is acetone, *characterised by* using waveband ranges substantially comprising at least two of the following waveband ranges: 1708 - 1697, 1415 - 1400, 1373 - 1357, 1245 - 1230 cm^{-1} and 1299 - 1276 cm^{-1} or parts thereof.

10. Method according to claim 8 or 9, *characterised by* using waveband ranges substantially comprising at least three of the said wavebands.

11. Method according to claim 8, *characterised by* using at least spectral information including the spectral ranges 1712 - 1697, 1419 - 1396, 1378 - 1353 and 1249 - 1226 cm^{-1}, or at least a substantial portion/part of said ranges.

12. Method according to claim 1, *characterised by* using any of the wavebands 1712 - 1697 cm^{-1}, 1419 - 1396 cm^{-1}, 1378 - 1353 cm^{-1}, 1299 - 1276 cm^{-1} or 1249-1226 cm^{-1}

13. Method for determining the content of acetone in milk, *characterised by* using an FT-IR apparatus calibrated by a method according to any of the claims 1 - 12.

14. Method for determining the content of a specified component in a fluid, *characterised by* using an FT-IR apparatus calibrated by a method according to any of the claims 1 - 6.

15. A method for determination of very low concentrations of a specific component in a fluid, said method including a calibration of a spectrometry apparatus according to any of the claims 1 – 14.

Publication V

Pre-processing Method Minimising the Need for Reference Analyses

To be published

Pre-processing Method Minimising the need for Reference Analyses

Per Waaben Hansen
R&D Greenhouse, Foss Electric A/S,
Slangerupgade 69, P.O. Box 260, DK-3400 Hillerød, Denmark

ABSTRACT

A new pre-processing method, called Independent Interference Reduction (IIR), is proposed. It is based on modelling of the interferences by use of Principal Component Analysis (PCA) using samples containing no variation in the parameter of interest. This is followed by a subtraction of the modelled effects from the calibration matrix. It is useful in cases where the parameter of interest gives only minor contributions to the calibration matrix, and where, at the same time, samples not containing variation in this parameter are easy to obtain. In such a case the method reduces the number of reference analyses required for establishing a robust calibration. When applied to acetone determination in milk samples by use of Fourier transform infrared (FT-IR) spectroscopy, the method provides a treatment of the data, which yields more robust models in the subsequent calibration step.

INTRODUCTION

When performing a calibration, the case where the data set has a high signal-to-noise ratio (S/N) is sometimes encountered. This is actually a good case, as it makes extraction of minor features from the data set possible. In spectroscopy, such examples are seen occasionally in mid-infrared (mid-IR) spectroscopy, especially when the spectra are generated by Fourier transform infrared (FT-IR) spectroscopy. Dimension-reducing methods, such as Principal Component Regression (PCR) or Partial Least Squares (PLS) regression (Martens and Næs, 1989) are necessary for making the information useful if the component of interest gives only rise to minor spectral features, *e.g.* because it is present in very low concentrations. In such cases, models of very high complexity result. FT-IR examples include urea (Hansen, 1998a) and acetone (Hansen, 1998b) in milk, where PLS models based on 15 to 20 factors are common. Such models require a high number of calibration samples (*i.e.* hundreds) to ensure a robust model.

In case of a component occurring in very low concentrations this may seem as a waste of good data, as the first PLS factors are usually describing the major constituents of the samples. This is because the weak spectral features from a low-concentration component might not be "visible" to the algorithm before the major interferences have been removed. Thus, expensive reference analysed data is wasted on the determination of scores and loadings for interfering compounds.

Therefore, a method removing all (or a major part) of the interferences *before* commencing regression on the reference analysed samples, would be useful. Usually, spectral data from a sample is easy and inexpensive to obtain – the reference analysis is the resource-demanding step. This is the basis for the method proposed in this paper: Independent Interference Reduction (IIR). It can be useful in cases where a low-concentration component is determined in a sample matrix with many interferences at high concentrations. The method is applied to an example from FT-IR spectroscopy, where the above mentioned requirements are met.

DESCRIPTION OF THE METHOD

IIR is a pre-processing method with similarities to two methods proposed recently:

Orthogonal Signal Correction (OSC) (Wold *et al.*, 1998) and Direct Orthogonalisation (DO) (Andersson, 1998). They both model the interferences and scatter effects by making a model of the data matrix with the information from the component of interest removed. This requires that: 1) all samples should be reference analysed, and that 2) the reference results are of a certain quality. Acting as pre-processing methods they remove scatter effects more efficiently than other schemes, *e.g.* Multivariate Signal Correction (MSC) (Martens and Næs, 1989), as they make a detailed model of the background. Furthermore, they provide models which are more interpretable, as interferences are removed *before* the calibration step. This is because the loadings are divided into those relating to the interferences, and those related to the useful part of the data.

IIR also removes interferences and scatter effects, but it does so by using *two* data sets: one set (matrix $\mathbf{X}_{simple}$) containing many samples showing variation in all parameters but the one of interest, and another set (matrices $\mathbf{X}_{special}$ and $\mathbf{Y}_{special}$) containing reference analysed samples with variation in all parameters, the one of interest in particular. The major requirement of the method is therefore that samples without variation in the parameter of interest are easily obtained. Parameters encountered in the dairy industry or in dairy farm management, such as compounds added during milk processing or components occurring as a result of a disease, may fulfil this requirement.

The basic steps for IIR are:

1. A data matrix, $\mathbf{X}_{simple}$, showing typical sample variation in all parameters, apart from the component of interest, is obtained. This matrix is modelled using Principal Component Analysis (PCA) using m principal components. If $\mathbf{X}_{simple}$ is large enough this results in a number of well-defined and accurate loadings.

2. Another matrix, $\mathbf{X}_{special}$, containing a wide variation in the component of interest is projected onto the loadings obtained in step 1, giving m sample scores. This model (based on m scores and loadings) of the interference part of $\mathbf{X}_{special}$ is subtracted from it, yielding a new matrix $\mathbf{X}_{special\text{-}m}$.

3. This "interference reduced" $\mathbf{X}_{special\text{-}m}$ is used with its associated reference matrix $\mathbf{Y}_{special}$ to generate a calibration model for predictive purposes, *e.g.* using PLS or MLR.

In step 1, the interferences contained in $\mathbf{X}_{simple}$ are modelled. It is therefore important that the component of interest is not present (or is constant) in $\mathbf{X}_{simple}$. However, deviations from this constraint will not be serious if the component is present in low concentrations (or shows weak spectral features), as the first principal components will not describe it anyway. Finally, if the error (*e.g.* the Root Mean Square Error of Prediction, RMSEP) is used for finding the optimal model in step 3, then problems originating from inclusion of spectral information stemming from the component of interest, will be revealed, as this will result in an increased prediction error. The data contained in $\mathbf{X}_{special}$ and $\mathbf{Y}_{special}$ should fulfil the usual requirements which apply to calibration sets (*e.g.* that the calibration data must represent future samples).

The steps are similar to the ones performed in PCA data pre-treatment (PCA-DP) (Sun, 1997), but where PCA-DP only models instrumental (baseline) effects, IIR is extended to model interferences as well. A detailed description of IIR is given below.

IIR during calibration

The matrices $\mathbf{X}_{simple}$, $\mathbf{X}_{special}$ and $\mathbf{Y}_{special}$ are required to perform the calibration steps. For the method to be useful, $\mathbf{X}_{simple}$ should contain more samples than $\mathbf{X}_{special}$.

1. $\mathbf{X}_{simple}$ is centred to give $\mathbf{X}^{c}_{simple}$, the mean spectrum being $\mathbf{x}_{m1}$

2. PCA is performed to give:

$$\mathbf{X}^{c}_{simple} = \mathbf{T}_{simple}\mathbf{P}^{t}_{simple}$$

using m principal components

3. $\mathbf{X}_{special}$ is scaled using $\mathbf{x}_{m1}$ to give $\mathbf{X}^{s}_{special}$

4. The scores of $\mathbf{X}^{s}_{special}$ on $\mathbf{P}_{simple}$ are calculated:

$$\mathbf{T}_{special} = \mathbf{X}^{s}_{special}\mathbf{P}_{simple}$$

5. $\mathbf{X}^{s}_{special}$ is reduced according to this model:

$$\mathbf{X}^{s}_{special\text{-}m} = \mathbf{X}^{s}_{special} - \mathbf{T}_{special}\mathbf{P}^{t}_{simple}$$

6. $\mathbf{X}^{s}_{special\text{-}m}$ is mean centred, and a regression against $\mathbf{Y}_{special}$ is performed. A PLS regression using n factors could be used in this step.

IIR during prediction

The vectors $\mathbf{x}_{pred}$ (a new spectrum), $\mathbf{x}_{m1}$ and $\mathbf{P}_{simple}$ (calculated during calibration) are required in the prediction step.

1. $\mathbf{x}_{pred}$ is scaled using $\mathbf{x}_{m1}$ to give $\mathbf{x}^{s}_{pred}$

2. The scores of $\mathbf{x}^{s}_{pred}$ on $\mathbf{P}_{simple}$ are calculated:

$$\mathbf{t}_{pred} = \mathbf{x}^{s}_{pred}\mathbf{P}_{simple}$$

3. $\mathbf{x}^{s}_{pred}$ is reduced according to this model:

$$\mathbf{x}^{s}_{pred\text{-}m} = \mathbf{x}^{s}_{pred} - \mathbf{t}_{pred}\mathbf{P}^{t}_{simple}$$

4. $\mathbf{x}^{s}_{pred\text{-}m}$ is scaled, and prediction is performed using the regression model calculated in calibration step 6

The dimensions of the initial PCA (m) and the final PLS calibration (n) have influence on the results in terms of the prediction error (RMSEP). Therefore validation should be performed calculating the prediction error using various values of m and n. From experience, a "closure" similar to the one observed with DO and PCA-DP, seems to exist: the sum of m and n (the overall dimensionality of the model) is constant. This could suggest that nothing is gained from use of IIR, but this is not the case: the first m loadings are generated from more samples than if only the samples from $\mathbf{X}_{special}$ had been used. Thus they are expected to be more well-defined and less noisy. Furthermore, the calibration is generated from a data set ($\mathbf{X}_{special}$) which is not "diluted" by samples not including variation from the component of interest. This may avoid the problem of underestimation of high samples and overestimation of low samples which is common in inverse calibration methods, such as PLS (Martens and Næs, 1989).

APPLICATION TO FT-IR DATA

The application presented here relates to acetone determination in milk from individual cows using FT-IR spectroscopy. This application fulfils the requirements of IIR, namely that: 1) acetone is found in milk low concentrations (*e.g.* 0.0 to 3.0 mM), and that 2) it is easy to obtain samples known not to contain acetone (acetone is only present when a cow suffers from ketosis, a metabolic disease). A detailed description of the data set can be found elsewhere (Hansen, 1998b). Only the information necessary for understanding the example will be given here.

Experimental

The calibration set consisted of 310 samples from individual cows collected in Norway, Sweden and Denmark. Eight outliers were removed from the set. The remaining samples were divided into two sets: 198 samples containing less than 0.1 mM acetone made up the matrix $\mathbf{X}_{simple}$, while the rest (104 samples, still containing some samples below 0.1 mM) made up $\mathbf{X}_{special}$. These samples were used for building a predictive model.

The test set consisted of 58 individual cow samples collected and measured in New Zealand. They were collected in a way which ensured a high proportion of acetone-containing samples. This set was used for testing the models.

All samples were measured on MilkoScan FT 120 FT-IR instruments (Foss A/S, Denmark),

either located in Denmark or New Zealand. The instruments were standardised before use. The samples were analysed for their acetone content using the reference method.

Calculations

The data analysis and calibration work was performed on a PC using Matlab 5.2.1 software (The MathWorks Inc., MA, USA). The calibration routines were either programmed by the author or taken from the PLS_Toolbox Version 1.5 (Eigenvector Technologies, WA, USA).

Repeatability is expressed as a mean standard deviation (s_r) of multiple determinations performed under identical conditions and is calculated as:

$$s_r = \sqrt{\frac{1}{q(n-1)} \sum_{j=1}^{q} \sum_{i=1}^{n} (x_{j,i} - \bar{x}_j)^2}$$

where q is the number of samples, n is the number of replicates, $x_{j,i}$ is the result of the i'th replicate of the j'th sample and $\bar{x}_j$ is the average result of the j'th sample.

Accuracy is expressed as the Root Mean Square Error of Prediction (RMSEP) and is calculated as:

$$RMSEP = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{i,reference} - x_{i,predicted})^2}$$

where N is the number of determinations (number of samples (q) times number of replicates (n) from above) and $x_{i,reference}$ and $x_{i,predicted}$ are the reference and predicted values corresponding to the i'th determination, respectively.

When a bias (mean difference between reference results and predictions) is observed, the Standard Error of Prediction (SEP) is used. It is calculated as:

$$SEP = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_{i,reference} - x_{i,predicted} - bias)^2}$$

Correlation is expressed as R^2, which is calculated as:

$$R^2 = \left[\frac{\frac{1}{N} \sum_{i=1}^{N} (x_{i,reference} - \bar{x}_{reference})(x_{i,predicted} - \bar{x}_{predicted})}{s_{reference} s_{predicted}} \right]^2$$

where N, $x_{i,reference}$ and $x_{i,predicted}$ are defined above and $\bar{x}_{reference}$, $s_{reference}$, $\bar{x}_{predicted}$ and $s_{predicted}$ are the means and standard deviations of the reference and predicted results, respectively.

RESULTS AND DISCUSSION

Calibrations were carried out for all values of m (PCA dimension) from 0 (ordinary PLS) to

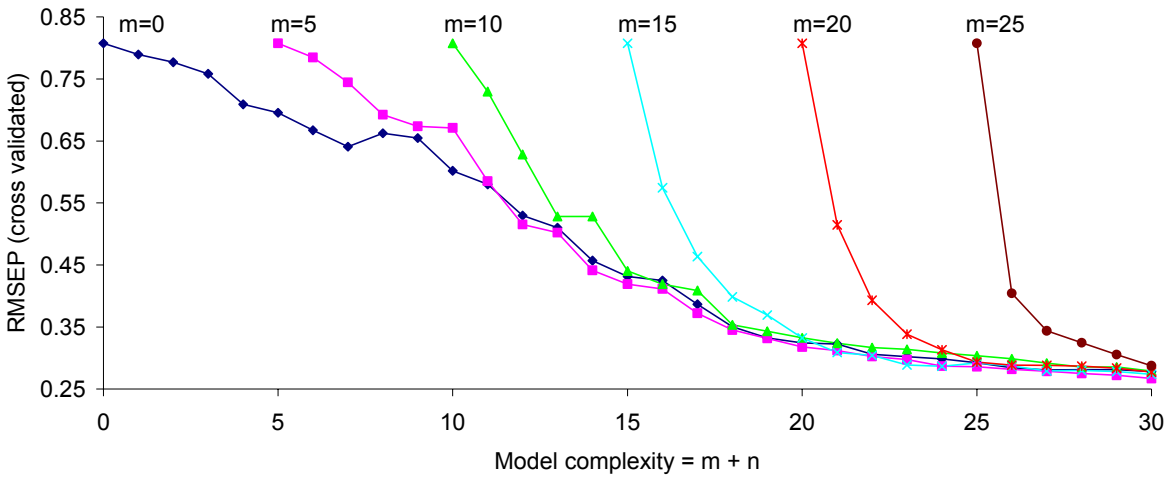


Figure 1 Cross validation error for acetone in milk (calculated using 10 cross validation segments) against model complexity for various dimensions (m) of the initial PCA in IIR.

25, and with all overall model complexities (*i.e.* m+n) from m+1 to 30. Cross validation was used for selecting the optimal model complexity, using 2, 4, 6, 8, and 10 cross validation segments. The average model complexity from these five trials was used when the final model was calculated.

The cross validated prediction errors using 10 cross validation segments for various values of m are plotted in Figure 1. From these plots it is clear that the overall model complexity stabilises at a value of approx. 25, irrespective of the value of m. This does not apply when m equals 25. In that case the same RMSEP is reached when m+n equals 30, so when m becomes high, the optimal model complexity is "pushed ahead". Thus, there seems to exist a lower limit for n, and this value must be 5 in the present case. Regarding m, using a too high value for this parameter does not seem to affect the predictive ability of the model.

The prediction error does not seem to improve when IIR is used, but it does not become poorer either. The main advantage of IIR is that it results in a smoother decay in the RMSEP, and in PLS loadings (not shown) with a closer relation to the component of interest.

When the final models are tested using the test set (with m+n determined by cross validation) there seems to be a benefit from using IIR. The results for various values of m are presented in Table 1. Both RMSEP and SEP are stated, as SEP is independent of a possible bias, while RMSEP is not.

In terms of the SEP the result is independent of the value of m, *i.e.* the use of IIR. The overall dimensionality of the models ends up at 24 or 25, apart from the case when m equals 25 itself. In this light, nothing is gained from IIR. When looking at the RMSEP, on the other hand, a significant improvement is seen when m is increased from 0 to 20. When m equals 20 the RMSEP is not significantly different from the SEP, and the bias obviously present at lower values of m, has disappeared. The bias could be a result of not all sample matrices being represented in $\mathbf{X}_{special}$, which is a common problem when the calibration set is too small. The application of IIR using the larger matrix $\mathbf{X}_{simple}$ removes this problem by adding a wider range of sample matrices. The repeatability, s_r, which is a measure of the noise in the experiment does not seem to be affected by the use of IIR: it is stable at a value of 0.15 to 0.17 mM.

m	m+n	R^2	RMSEP	SEP	s_r
0	25	0.827	0.60	0.28	0.15
5	25	0.858	0.54	0.26	0.15
10	25	0.803	0.59	0.29	0.16
15	24	0.774	0.34	0.31	0.16
20	25	0.779	0.31	0.29	0.17
25	30	0.846	0.42	0.25	0.16

Table 1 Results on the independent test set using various combinations of the dimensions m and n.

When m becomes greater than the apparent overall dimension, the RMSEP increases while the SEP remains the same. The overall result is that IIR tends to remove a bias, but it requires careful selection of the value of the parameter m, as the bias increases when m becomes too high. In the present example the optimal model seems to be one with m equal to 20 and n equal to 5. The predicted vs. measured plot for the test set in this case is shown in Figure 2.

In addition to the robustness gained, IIR pre-processing results in more interpretable PLS models. This is because only relevant acetone

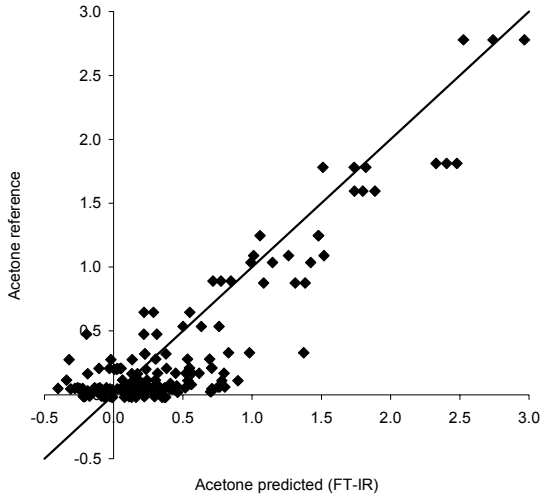


Figure 2 Reference vs. predicted acetone in milk for the test samples using a model with IIR pre-processing, where m = 20 and n = 5. R^2 is 0.779, RMSEP is 0.31 mM and s_r is 0.17 mM.

information is present in the pre-processed matrix when PLS modelling is commenced. For example, with an appropriate choice of m, the first PLS loading (not shown) contain features similar to the spectrum of pure acetone.

CONCLUSIONS

IIR is a new pre-processing method which can be useful for specific applications where the following requirements are met:

- The parameter of interest only gives minor contributions to the **X** matrix (*e.g.* a component in low concentrations)
- Samples in which the parameter of interest is constant, should be easy to obtain

The benefits of the method is that it reduces the need for resource-demanding experiments to obtain **Y**-data for calibration (*i.e.* reference analyses), and that the resulting calibration model is more interpretable. The main disadvantage is that two parameters, m and n, should be optimised, in case PLS regression is used for calibration.

In the specific case where acetone is determined by use of FT-IR, an improved predictive model could be obtained by use of IIR. On this data set the major improvement over the ordinary PLS calibration was that the calibration became more robust (having a lower bias when used on different instruments at different locations).

ACKNOWLEDGMENTS

The author would like to thank The Danish Academy of Technical Sciences and Foss A/S, Denmark, for providing the funding for this work. It was carried out in co-operation with Food Technology at the Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Denmark. Dairying Research Corporation Ltd. and New Zealand Dairy Research Institute are thanked for providing samples and equipment for the practical work. Claus A. Andersson, Rasmus Bro, Henrik V. Juhl, Lars Nørgaard, and Carsten Ridder are thanked for useful discussions.

REFERENCES

Andersson, C. A. 1998. Direct orthogonalization. Chemomet. Intell. Lab. Sys. (in press).

Hansen, P. W. 1998a. Urea determination in milk using Fourier transform infrared spectroscopy and multivariate calibration. Milchwissenschaft 53:251-255.

Hansen, P. W. 1998b. Screening of dairy cows for ketosis by use of infrared spectroscopy and multivariate calibration. J. Dairy Sci. (submitted).

Martens, H., and Næs, T. 1989. Multivariate Calibration. Wiley, Chichester.

Sun, J. 1997. Statistical analysis of NIR data: Data pretreatment. J. Chemomet. 11:525-532.

Wold, S., Antti, H., Lindgren, F., and Öhman, J. 1998. Orthogonal signal correction of near-infrared spectra. Chemomet. Intell. Lab. Sys. (in press).

Publication VI

Detection of Specific Sugars in Dairy Process Samples Using Multivariate Curve Resolution

Submitted for publication in *Journal of Dairy Science*

Detection of Specific Sugars in Dairy Process Samples Using Multivariate Curve Resolution

P. W. Hansen[*], A. S. van Brakel[†], J. Garman[†], and L. Nørgaard[‡]
[*]Foss Electric A/S, R&D Greenhouse,
Slangerupgade 69, P.O. Box 260, DK-3400 Hillerød, Denmark
[†]New Zealand Dairy Research Institute,
Private Bag 11 029, Palmerston North, New Zealand
[‡]The Royal Veterinary and Agricultural University, Department of Dairy and Food Science,
Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark

ABSTRACT

Dairy process monitoring by application of multivariate curve resolution using Alternating Least Squares is presented. Alternating Least Squares was used for resolving Fourier transform infrared spectral data from a dairy batch process in which lactose is enzymatically hydrolysed to glucose and galactose. It was possible to extract four compounds (fat, lactose and two other sugar components) from the spectral data obtained from nine process runs. Subsequently, the pure spectra obtained in this way were used to monitor the content of these compounds in two new process runs. In this way Alternating Least Squares made it possible to follow the hydrolysis process by Fourier transform infrared spectroscopy *without* the need for reference analyses. When the results were correlated to reference results for lactose the accuracy was similar to what was obtained when performing a Partial Least Squares regression on the same data: A lactose correlation of 0.980 was obtained when Alternating Least Squares was used and 0.987 when Partial Least Squares was used.

(**Key words:** lactose hydrolysis, multivariate curve resolution, process control, reference-independent estimation)

Abbreviation Key: **ALS** = Alternating Least Squares, **CVS** = Cross Validation Segments, **FT-IR** = Fourier Transform Infrared, **MIR** = Mid-Infrared, **NIR** = Near Infrared Reflection, **NIT** = Near Infrared Transmission, **PLS** = Partial Least Squares, **RMSEP** = Root Mean Square Error of Prediction, **SEC** = Standard Error of Calibration, **SEP** = Standard Error of Prediction, **SSE** = Sum of Squared Errors.

INTRODUCTION

Process control in industrial processes is increasing in importance as on-line analytical equipment providing fast and reliable results becomes available. Near infrared reflection/transmission (**NIR/NIT**) spectroscopy is the most frequently used method in many branches of industry, while mid-infrared (**MIR**) spectroscopy has proved very useful for process milk analysis. Milk analysis using MIR equipment is generally more accurate than the corresponding NIR/NIT method. This is because MIR contains more specific information (fundamental absorptions) and stronger signals than NIR/NIT, which detects derived information (overtones and combination bands). In addition, full spectrum instruments based on Fourier transform infrared (**FT-IR**) spectroscopy for dairy product analysis in the laboratory are showing promising results with regard to the number of components (*e.g.* specific sugars (13), casein (7, 9), urea (6)) that can be measured.

Hydrolysis of lactose in milk is of interest as large racial groups suffer from lactose intolerance, *i.e.* they are not able to cleave lactose into glucose and galactose. Therefore, low-lactose milk products produced by the action of the enzyme β-galactosidase are of commercial interest. The process is sensitive to the initial conditions such as temperature and β-

galactosidase concentration (2, 5). Therefore, the concentrations of the sugars need to be monitored during the course of the reaction, to control the process. The reaction is typically completed within a few hours, thus it requires a fast analytical method such as MIR.

The general approach when analysing spectral data with the intention of generating future predictions of milk constituents is to use one of several multivariate methods relating the data to wet chemistry results. These methods include Partial Least Squares (**PLS**) regression which is described elsewhere (10). The problem is that ordinary multivariate methods require an accurate and reproducible reference method to obtain a reliable calibration. This tends to be resource-demanding – especially when the typical number of calibration samples (fifteen to hundreds) is taken into account. In addition, during a process some intermediate species might only exist for a limited period of time, *i.e.* they are both produced and consumed during the reaction. In such a case it might be difficult to isolate the intermediates and to measure them using the reference methods.

Such problems can be solved using a regression method of higher order, such as PARAFAC (Parallel Factor Analysis) (1). MLR, PCR and PLS handle second order data, as the data can be arranged in a matrix. PARAFAC requires the data to be of an order higher than two, which is the case when *e.g.* a spectral *landscape* is obtained for each sample and the whole data set can be arranged in a cube. The landscape could be obtained by measuring a reacting sample with fixed time intervals during the process. The individual spectra constituting the landscape will be related and these relationships implicitly contain information on the concentrations of all infrared absorbing compounds in the sample.

PARAFAC is able to resolve these variations and to produce concentration profiles and pure spectra corresponding to the absorbing species present in the sample. The concentrations will be arbitrary, but proportional to the true concentrations. If correlating species are present, the concentration profiles will be the sums of such correlating compounds. PARAFAC is normally the most useful method for multivariate curve resolution, as it can handle more than one sample at a time, the solutions to the mathematical problem are unique (there is only one solution to each problem), and they might resemble real spectra and concentrations. Good results have been obtained on resolving absorption and emission profiles from fluorescence spectra of sugar samples using PARAFAC (1).

In the present case, PARAFAC will not work, as the actual shape of the concentration profiles will be strongly dependent on the initial conditions of the process. In such a case, it will only be able to analyse one landscape (sample) at a time *or* the unfolded data set. When a landscape is unfolded (*e.g.* when the spectra from the individual runs are appended to each other), one of the directions in the three-dimensional structure is lost.

Alternating Least Squares (**ALS**), sometimes referred to as Alternating Regression (8), is a two-way method which handles one landscape or unfolded data set at a time. Ref. (15) shows how such curve resolution methods are working in practice. ALS produces pure spectra and concentration profiles in a way similar to PARAFAC performed on the unfolded data set. The method has been used for resolution of infrared process data with excellent results (3, 14).

The aim of the present work is to investigate whether ALS is capable of resolving the changes occurring in the FT-IR spectra during the course of the lactose hydrolysis process and thus to obtain concentration profiles and pure spectra for the involved compounds. This will be done *without* the use of reference analyses, in order to show the resolving power of the ALS method. Having obtained the pure spectra it should be possible to monitor the concentrations of the components in new process runs. The results will be compared to results from an ordinary PLS regression performed on the same data set.

MATERIALS AND METHODS

Alternating Least Squares (ALS)

Alternating Least Squares relies on the assumption that the Beer-Lambert law is correct, *i.e.* that a spectrum (the row vector **x**) of a given sample can be seen as a linear combination of the pure constituent spectra (contained in the matrix **A**), thus

$$\mathbf{x} = \mathbf{cA} \quad [1]$$

where **c** is a row vector containing the concentrations of the constituents corresponding to the pure spectra in **A**. In the case where more than one spectrum is measured, the general expression becomes

$$\mathbf{X} = \mathbf{CA} \quad [2]$$

where **X** is a landscape containing the spectra in its rows and **C** is a matrix containing the concentrations corresponding to each spectrum. In this context one sample is named **X**, *i.e.* a collection of spectra from one process run.

A typical landscape from one lactose hydrolysis run with seven FT-IR spectral recordings is presented in Figure 1. Most of the variation in the spectra is in the range between 1000 and 1200 cm^{-1}. This is expected, as the only compounds affected by the reaction are sugars which show strong absorptions due to stretching of the sugar C-O bonds in this range. ALS calculates the pure spectra **A** from the input spectra in **X** (the landscape) and an estimate of **C** (*e.g.* random numbers) using a rearranged form of [2]:

$$\mathbf{A} = \mathbf{C}^{+}\mathbf{X} \quad [3]$$

where $\mathbf{C}^{+}$ is some pseudo-inverse of **C**, followed by a calculation of a better guess of **C** from this **A**:

$$\mathbf{C} = \mathbf{XA}^{+} \quad [4]$$

where $\mathbf{A}^{+}$ is the pseudo-inverse of **A**. The steps [3] and [4] are repeated until convergence (or a maximal number of iterations) has been reached.

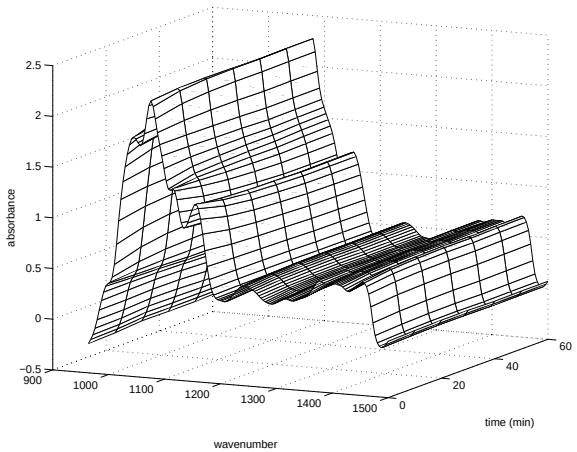


Figure 1 Infrared landscape obtained during 60 minutes (with sampling every 10 minutes) of a lactose hydrolysis process. Major changes are in the range from 1000 to 1200 cm^{-1} where sugars generally have strong absorptions.

Various constraints can be applied to the spectra and concentration profiles in **A** and **C** in order to avoid physically meaningless solutions (1). For example, the concentrations in **C** cannot possibly become negative, so a *non-negativity constraint* would be reasonable. In addition, when looking at compounds produced and consumed during a chemical process, the concentration profiles will be expected to have only *one* maximum during the course of the reaction. This leads us to applying the *unimodality constraint* which limits solutions to smooth concentration curves with only one maximum each. It could be argued that a non-negativity constraint would be appropriate for the pure spectra in **A** as well, but in this specific application it is not. As the FT-IR absorbance spectra are calculated using a water background slightly negative absorbances will result. Thus, constraining **A** to non-negativity will restrict the algorithm too much.

Non-negativity can be applied in various ways. The most straightforward approach is to force negative values to zero (*e.g.* in **C**) after each iteration. This is very simple and does not necessarily lead to the optimal description of **X** (*i.e.* the least squares solution). The approach employed here (adopted from *The N-Way Toolbox* by C. A. Andersson, Internet site: http://newton.foodsci.kvl.dk/matlab/nwaytoolbox) forces only one concentration profile (or spectrum) at a

time to zero, followed by a correction of the pure spectrum matrix, **A** (or concentration matrix, **C**). This modification leads to the optimal result.

In case of more process runs contained in the same data matrix **X** the unimodality constraint of **C** would not work. In such a case only the parts of **C** originating from the same process should be constrained. This approach was used in the present work.

After the concentration profiles and pure spectra have been obtained the same principle can be used for prediction of the constituents in an unknown sample by applying the vector form of [4] to the spectrum of the sample:

$$\mathbf{c} = \mathbf{x}\mathbf{A}^{+} \quad [5]$$

The concentration row vector **c** will be in arbitrary units, but linearly related to the actual concentrations.

Calculations

The data analysis and calibration was performed on a PC using Matlab 5.2.1 software (The MathWorks Inc., MA, USA). The pseudo-inverse in equations [3] and [4] were calculated using the built-in functions of Matlab. The calibration routines were either programmed by the authors or taken from the PLS_Toolbox Version 1.5 (Eigenvector Technologies, WA, USA).

Repeatability is expressed as a mean standard deviation (s_r) of multiple determinations performed under identical conditions and is calculated as:

$$s_r = \sqrt{\frac{1}{q(n-1)} \sum_{j=1}^{q} \sum_{i=1}^{n} (x_{j,i} - \bar{x}_j)^2}$$

where q is the number of samples, n is the number of replicates, $x_{j,i}$ is the result of the i'th replicate of the j'th sample and $\bar{x}_j$ is the average result of the j'th sample.

Accuracy is expressed as the Root Mean Square Error of Prediction (**RMSEP**) and is calculated as:

$$\mathrm{RMSEP} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{i,\mathrm{reference}} - x_{i,\mathrm{predicted}})^2}$$

where N is the number of determinations (number of samples (q) times number of replicates (n) from above) and $x_{i,\mathrm{reference}}$ and $x_{i,\mathrm{predicted}}$ are the reference and predicted values corresponding to the i'th determination, respectively.

When a bias (mean difference between reference results and predictions) is observed, the Standard Error of Prediction (**SEP**) is used. It is calculated as:

$$\mathrm{SEP} = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_{i,\mathrm{reference}} - x_{i,\mathrm{predicted}} - \mathrm{bias})^2}$$

If two variables are related by performing a univariate linear regression (slope a and intercept b), between instrumental responses, $x_{i,\mathrm{instrumental}}$, and reference results, $x_{i,\mathrm{reference}}$, the accuracy of future predictions can be estimated by the use of the Standard Error of Calibration (**SEC**), calculated as:

$$\mathrm{SEC} = \sqrt{\frac{1}{N-2} \sum_{i=1}^{N} (x_{i,\mathrm{reference}} - (a x_{i,\mathrm{instrumental}} + b))^2}$$

Correlation is expressed as R^2, which is calculated as:

$$R^2 = \left[\frac{\frac{1}{N} \sum_{i=1}^{N} (x_{i,\mathrm{reference}} - \bar{x}_{\mathrm{reference}})(x_{i,\mathrm{predicted}} - \bar{x}_{\mathrm{predicted}})}{s_{\mathrm{reference}} s_{\mathrm{predicted}}} \right]^2$$

where N, $x_{i,\mathrm{reference}}$ and $x_{i,\mathrm{predicted}}$ are defined above and $\bar{x}_{\mathrm{reference}}$, $s_{\mathrm{reference}}$, $\bar{x}_{\mathrm{predicted}}$ and $s_{\mathrm{predicted}}$ are the means and standard deviations of the reference and predicted results, respectively.

Finally, the ***fit of the model*** to **X** (equation [2]) is expressed as the Sum of Squared Errors (**SSE**):

$$SSE = \sum_{i=1}^{N}\sum_{j=1}^{M}(x_{i,j} - \mathbf{c}_i\mathbf{a}_j)^2$$

where $x_{i,j}$ is an element in **X**, $\mathbf{c}_i$ is as row vector containing the concentrations of the i'th sample and $\mathbf{a}_j$ is a column vector containing the absorbencies of the j'th wavelength. M is the number of wavelengths in the spectra. Note that the reference results have not been used in the calculation of the SSE.

Experimental

Sample sets. The samples obtained for this work are from New Zealand and fall within two groups:

Calibration samples: This set contains 124 samples. They were collected from nine process runs (five based on skim milk, four based on whole milk) carried out in May 1997 using an experimental set-up in the laboratory. Lactozym 3000 (Novo-Nordisk, Bagsværd, Denmark) was the enzyme used. Samples were taken from the reaction mixture at various time points over a three hour period, and they were immediately heated to 80 °C in a 750 W microwave oven in order to deactivate the enzyme. Duplicate samples were taken, and the following reference analyses and spectral measurements were carried out independently. Thus, the set of 124 samples comprises two very similar sets of 62 samples.

Test samples: This set contains 23 samples obtained from two process runs carried out in the laboratory in November and December 1997 using the same experimental set-up. Samples were taken at various intervals, and this time only the sub-sample used for reference analysis had the enzyme deactivated. The spectral measurement was carried out on the non-deactivated sample immediately (*i.e.* less than one minute) after sampling in order to make the FT-IR measurements as close to an on-line application as possible. The sub-samples for reference analysis were still subjected to a heat treatment.

Reference measurements. Lactose was determined on 90 of the calibration samples, as well as the 23 test samples, using the following HPLC set-up:

Equipment: Waters Maxima 820 Workstation (Millipore Corp., Milford, MA, USA) with Waters model 510 pump (Millipore Corp.) and a Waters WISP automatic injector (Millipore Corp.). Column heater (Jones Chromatography Ltd., Hengoed, UK). Shimadzu RID 6A detector (Shimadzu Corp., Kyoto, Japan). Alltima NH_2 column (250 × 4.6 mm, 5 micron particles) (Alltech Associates, Auckland, New Zealand). Adsorbosphere NH_2, 10 mm guard column cartridge (Alltech Associates). Silica saturation column between pump and injector.

Conditions: Column temperature: 28 °C. Mobile phase: 80 % acetonitrile at 2.0 mL/min. Injection volume: 100 μL.

Sample preparation: Liquid milk containing less than 1.0 g of total solids was diluted to 20 ml with water. The proteins were precipitated with barium hydroxide and zinc sulphate. The volume was made up to 40 mL before centrifugation for the unhydrolysed lactose samples and to 30 mL for the hydrolysed lactose samples. After centrifugation at 2000 rpm for 10 minutes, 2 mL of the clear supernatant was diluted to 10 mL with acetonitrile and filtered through a 0.45 micron syringe filter (Whatman, Singapore).

Spectral measurements. The FT-IR spectral measurements were carried out using a MilkoScan FT 120 (Foss, Hillerød, Denmark). The infrared spectrum from 925 to 5000 cm^{-1} was recorded. The calibration samples were measured in duplicate and the test samples were measured in triplicate.

In the data analysis only the ranges 964-1542, 1724-1847 and 2699-2965 cm^{-1} were used, as

	2 [1]CVS	4 [1]CVS	6 [1]CVS	8 [1]CVS	10 [1]CVS
[2]#factors	5	5	5	5	4
R^2	0.997	0.996	0.996	0.996	0.996
RMSEP	0.82	0.90	0.88	0.93	0.88
s_r	0.22	0.23	0.23	0.23	0.24

Table 1 PLS results from the cross-validated calibrations for the determination of lactose in the calibration samples. 44 samples ranging from 0 to 45 % dry base lactose were used.
[1]CVS = Cross-Validation Segments
[2]#factors = optimal number of PLS factors

these are the areas containing useful chemical information.

All measurements were ratioed against water and log-transformed to give absorbance spectra. A typical time-resolved landscape for the first hour of an experiment is shown in Figure 1. No spectral pre-processing was performed prior to data analysis, as experience shows that only minor improvements, if any, are obtained when using full spectrum data of the present type (6).

RESULTS AND DISCUSSION

PLS Calibration

For comparative purposes, a PLS calibration against the lactose reference results was performed.

As the calibration set consists of two dependent sample sets, only the first set (62 samples of which 44 had been reference-analysed) was used for finding the optimal model (regarding the number of PLS factors). To this end, a PLS calibration validated using cross-validation on the calibration samples was performed. The procedure was repeated using 2, 4, 6, 8 and 10 cross-validation segments to check the stability of the result. The results are shown in Table 1 and a reference vs. measured plot for lactose is shown in Figure 2. From these results, a model using 5 PLS factors is the best.

Then, all reference-analysed samples in the calibration set (a total of 90 samples) were

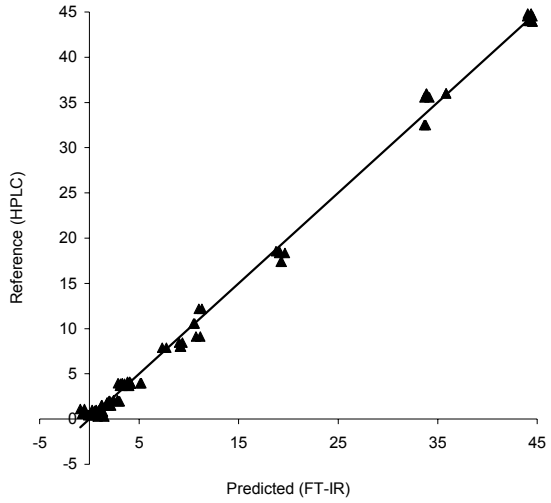


Figure 2 Reference versus predicted plot for lactose showing the 44 calibration samples (two replicates for each sample) predicted using cross validation (with 6 cross-validation segments) against the reference results. The model uses 5 PLS factors with R^2 0.996, RMSEP 0.88 and s_r 0.23. The reference results range from 0 to 45 % dry base lactose.

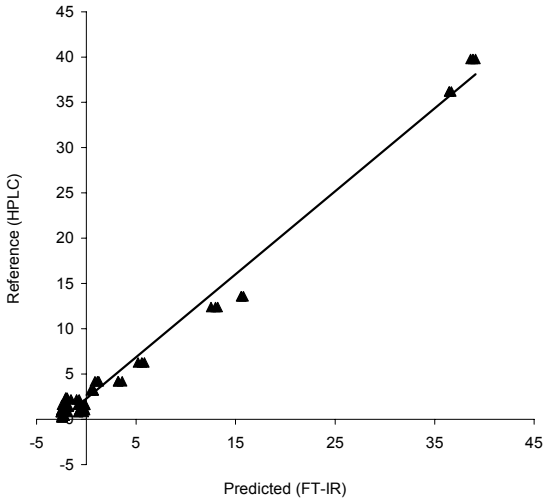


Figure 3 Reference vs. predicted plot for lactose showing the 23 test samples (three replicates for each sample) predicted using a PLS model with 5 factors. R^2 0.987, RMSEP 2.49, SEP 1.55, bias –1.96 and s_r 0.23. The reference results range from 0 to 40 % dry base lactose.

	3 components		4 components		5 components	
	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2
SSE	0.56	1.01	0.38	0.44	0.26	0.47
R^2	0.928	0.942	0.942	0.971	0.978	0.977
SEC	3.47	3.15	3.12	2.23	1.92	2.00
A: Results of the calibration sets (62 samples each) – the models selected on the basis of the lowest SSE.						
SSE	0.71	1.06	0.44	51.63	0.40	10.50
R^2	0.928	0.942	0.971	0.986	0.991	0.992
SEC	3.47	3.15	2.22	1.54	1.26	1.16
B: Results of the calibration sets (62 samples each) – the models selected on the basis of the highest correlation (R^2) to the lactose reference results.						
R^2	0.894	0.891	0.959	0.980	0.959	0.940
SEC	3.46	3.50	2.16	1.51	2.14	2.61
s_r	0.15	0.15	0.16	0.27	0.16	0.13
C: Results of the test set (23 samples) using the models (selected by use of the SSE) obtained from the individual calibration sets.						
R^2	0.894	0.891	0.959	0.953	0.978	0.973
SEC	3.46	3.50	2.15	2.29	1.59	1.78
s_r	0.15	0.15	0.16	0.14	0.19	0.19
D: Results of the test set (23 samples) using the models (selected by use of R^2) obtained from the individual calibration sets.						

Table 2 ALS results of the two calibration sets (part A and B) and the independent test set (part C and D) using 3, 4 and 5 components. The reference results range from 0 to 45 % dry base lactose.

used to build the final model using 5 PLS factors. This model was used for predicting lactose in the 23 test samples. The result is shown in Figure 3. It is clearly not as good as the result indicated by the cross-validation, but most of the error is due to a substantial bias. Thus, it seems as if the calibration set was not representative for the test set. The reason could be:

1. The time difference (more than 6 months) between the measurement of the calibration and test samples

2. Another chemical composition of the test samples due to seasonal variations in the milk. Seasonal variations are of great importance in New Zealand, where most dairy cattle calve simultaneously, as farming is pasture-based

3. From the spectral data (not shown) it is evident that the protein content of the calibration samples is virtually the same for all. Therefore, the calibration cannot account for protein variations in the test samples

Despite this lack of reproducibility, the reference results and lactose predictions correlate well which is the main issue in this context.

ALS Results

The PLS results suggest that five independent factors are necessary for predicting the lactose content of the samples. Thus, a similar number of components in the ALS is expected.

To obtain reasonable solutions, two constraints were applied during the regression:

1. Non-negativity of the concentration profiles

2. Unimodality of the concentration profiles (except for fat) in each process run

Although the lactose concentrations of most of the samples were already known, they were *not* used as start guesses, as the purpose of this study was to see whether it was possible to obtain good results having only a limited knowledge of the process. Therefore the following – very simple – start guesses were used:

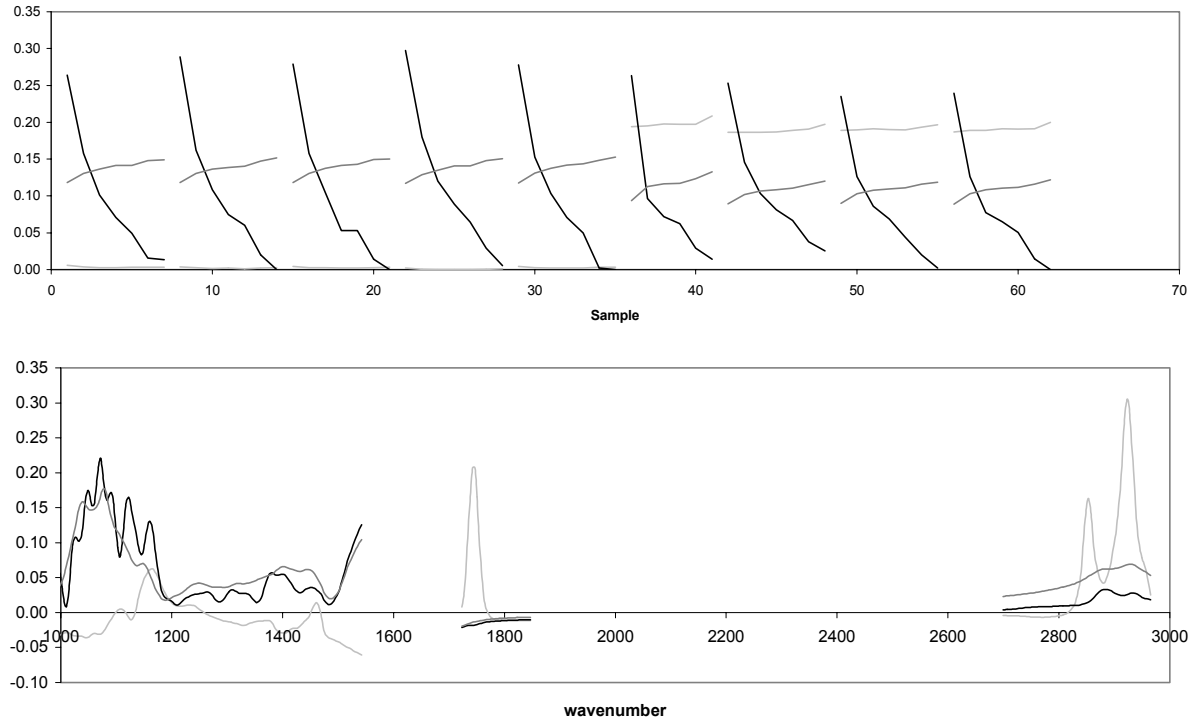


Figure 4 The 3-component ALS solution with lowest SSE (out of 100 runs) for set 1. Upper part shows the concentration profiles, lower part the pure spectra.

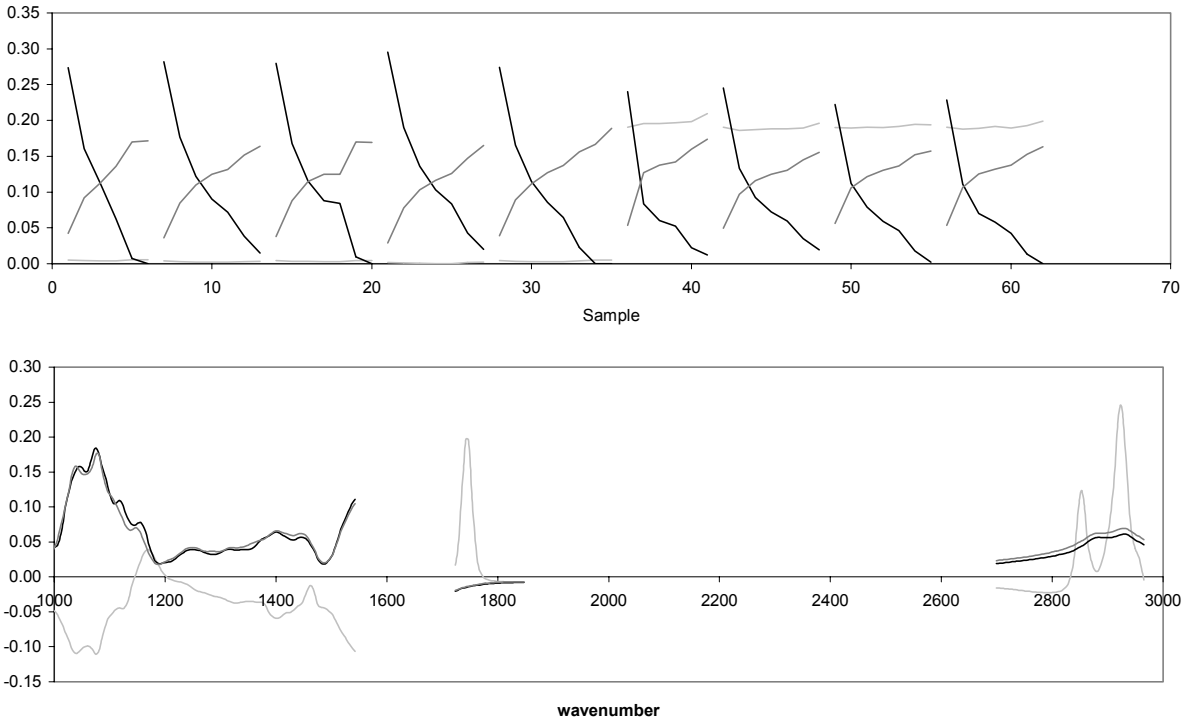


Figure 5 The 3-component ALS solution with lowest SSE (out of 100 runs) for set 2. Upper part shows the concentration profiles, lower part the pure spectra.

1. The skim and whole milk samples had a start guess for fat of 0 and 1, respectively
2. As lactose is known to be the only sugar present in the beginning of the process, the lactose concentration of the first sample of each run was set to 1, while all others were set to 0

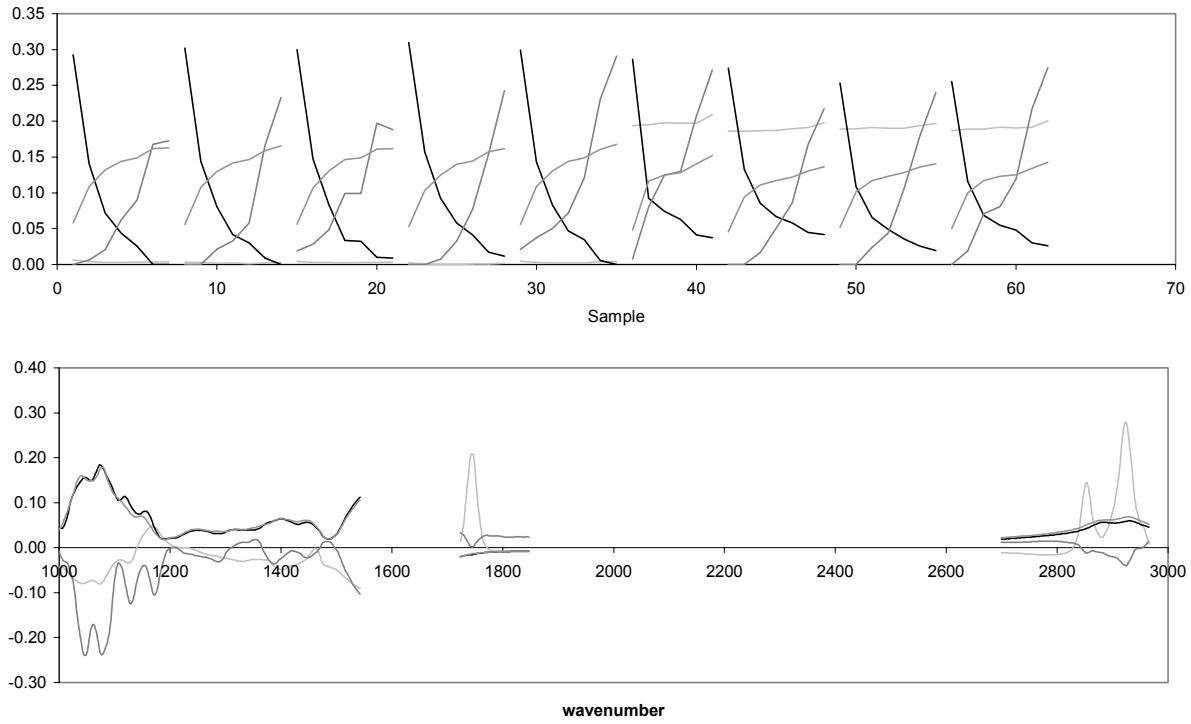


Figure 6 The 4-component ALS solution with lowest SSE for set 1. Upper part shows the concentration profiles, lower part the pure spectra.

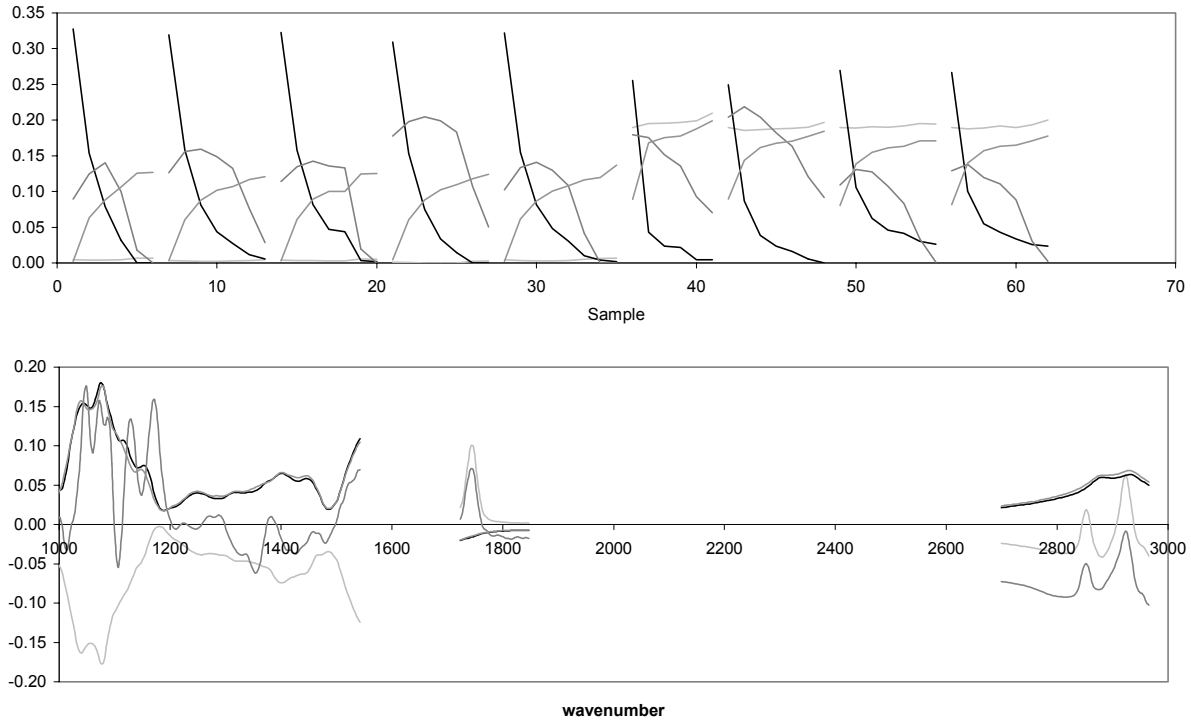


Figure 7 The 4-component ALS solution with lowest SSE for set 2. Upper part shows the concentration profiles, lower part the pure spectra.

3. The remaining components were given random numbers as start guesses (uniformly distributed numbers between 0 and 1)

As random numbers were used, many different solutions might result. For this reason the ALS run was carried out 100 times with new start guesses for each number of components. From three to six components (corresponding

to the number of pure spectra in **A**, equation [2]) were tried on both (almost identical) calibration sets of 62 samples. When six components were used, the concentration profiles and pure spectra became noisy and highly correlated, so they will not be discussed in the following text.

The best models, in terms of how well the **X** matrix is described (measured as SSE) and how well the "lactose" profile correlates with the reference results (measured as R^2), are shown in Table 2 (A and B parts). The results for three to five components were as follows:

With three components, the same pure spectra and concentration profiles were reached every time on both calibration sets – at least the differences were insignificant. The solutions with the lowest SSE are shown in Figures 4 and 5. Note that both pure spectra and concentration profiles have been normalised to make a presentation on the same scale possible. The pure spectra describing fat (having strong absorptions in the high end of the spectrum due to C-H stretching vibrations) generally have very negative contributions in the areas where the sugars absorb, because high-fat samples generally contain less lactose as a result of the displacement of the water phase by fat. This is unfortunate, as the ALS model will expect samples with a high fat content to contain less sugars, but it cannot be avoided with the present data set.

With four components many different solutions were reached. They belonged to a limited number of groups of solutions inside which the variations were small. Some results are shown in Figures 6 and 7. Of the four components at least two are sugars (having strong absorptions in the 1000 to 1200 cm^{-1} range), one is fat (strong absorptions between 2800 and 3000 cm^{-1}), while the last one is difficult to assign. Of the sugars, the component decreasing rapidly through each batch is lactose.

With five components the problem of finding the optimal solution becomes even more difficult. But, as is evident from Table 2, good lactose correlations were still obtained. As there is no major improvement when going from four to five components, the three- and four-component solutions were chosen for further examination.

An ALS with three components gives the most stable result and the most reasonable pure spectra, but the correlation to the actual lactose concentration is relatively poor. In Figure 8 the lactose concentration profiles are plotted against the reference results, and the resulting plot is clearly non-linear. The four-component model (Figure 9) gives a much more linear relationship to the lactose refe-

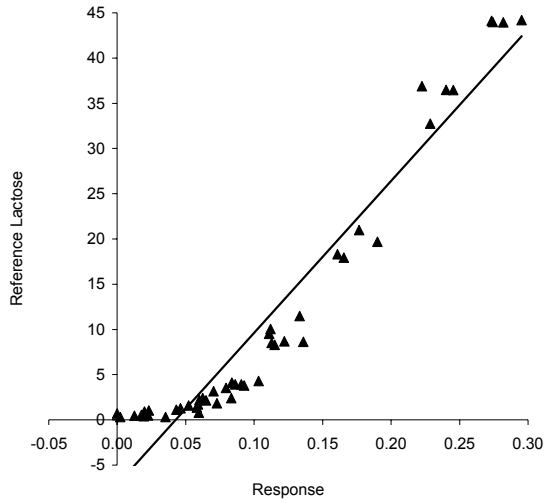


Figure 8 Lactose profile and reference lactose results plotted against each other. The profile originates from an ALS on calibration set 2 with 3 components. The relationship is strongly non-linear. R^2 is 0.942 and SEC is 3.15.

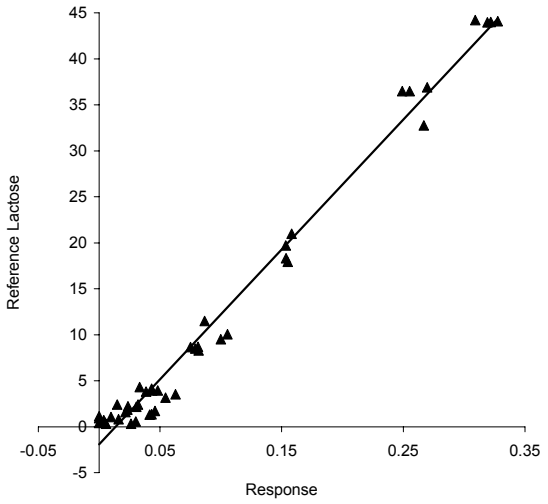


Figure 9 Lactose profile and reference lactose results plotted against each other. The profile originates from an ALS on calibration set 2 with 4 components. The linearity is much better than with only 3 components (Figure 8). R^2 is 0.971 and SEC is 2.23.

rence results. It cannot be due to overfitting (*i.e.* a too optimistic estimate of the error) as the reference results were not involved in the optimisation. In addition, the concentration profiles of the four-component model agree with previous observations that not only the monosaccharides (galactose and glucose), but also various other sugars (containing two or more monosaccharide units), generally known as the oligosaccharides, are formed during the process (4, 11, 12). The shapes of the concentration profiles are actually very similar to these previous observations. The way in which these sugars are distributed among the last two components (Figures 6 and 7) can vary between different ALS runs, which is the reason for the many different solutions seen when four (and five) components are tried.

Whether the SSE or the R^2 should be used as the selection criterion is not clear. When three components are used, both criteria give almost the same result, while the SSE criterion gives a somewhat higher prediction error in the case of four and five components. In both cases there is a significant improvement over the three-component result.

The final test of the hypothesis was done by using the ALS models on the test set obtained from two new process runs. The results are shown in Table 2 (C and D parts). There is a large improvement in SEC when going from three to four components. It is almost independent of the method (R^2 or SSE) used for selecting the optimal model. These results lead to the conclusion that there is nothing gained by selecting the optimal model by use of the lactose reference results (*i.e.* by looking at R^2). In fact, the best result is obtained by using the SSE. This is a very promising result, since there is no need for reference analyses when process data of the present type are analysed. ALS alone can be used for generating a model which can be used for future process monitoring. Note that the ALS concentrations are in arbitrary units, so only relative process changes can be detected.

The results in Table 2 should be compared to the PLS result shown in Figure 3. It is seen that the ALS predictions of lactose (using four components) are almost as good as the PLS predictions – in some cases the results are the same.

The normalised predictions of three of the four components (using the model with lowest SSE based on calibration set 2) for each of the two processes included in the test set are shown in Figure 10 together with the corresponding results from the PLS model and the reference method. Fat is omitted, as it is constant during each run. The ALS lactose predictions do not agree perfectly, neither with the reference nor the PLS results based on the same spectra, but they follow roughly the same curve. The main reason for the disagreement between the ALS and PLS results are the negative PLS lactose predictions which are due to the earlier discussed bias of the PLS calibration. Both PLS and ALS concentration profiles follow a smooth curve, which should be expected when dealing with a chemical reaction. Thus, the

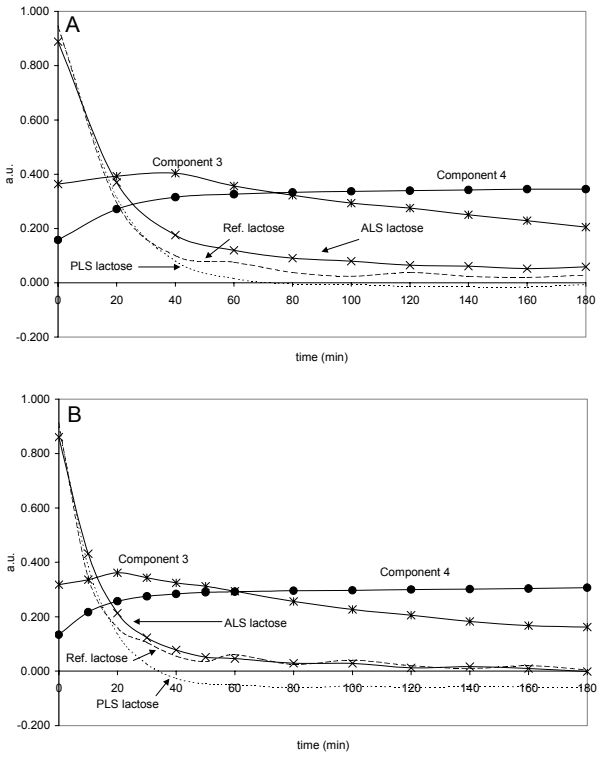


Figure 10 Concentration profiles from the reference and PLS results (lactose only) as well as the four components obtained through ALS for the first (A) and second (B) experiment present in the test set. All profiles are normalised to unity in order to make them comparable. Only the three sugar components are shown, as component 1 (fat) is constant throughout the process run.
a.u. = arbitrary units

fluctuations in the lactose reference results are likely to be caused by the lack of reproducibility of the reference method rather than real variations in the lactose content.

The two test runs (Figure 10) gave the same shapes of the concentration profiles of the third and fourth component as seen during calibration. Therefore Component 3 is assigned to the sum of galactose and glucose, and Component 4 is likely to be caused by oligosaccharides formed during the reaction. Another data set with reliable reference results on other sugars is required to confirm this.

The remaining problem allowing implementation of ALS for practical use in dairy process monitoring is how to select the optimal number of components in the ALS model. This corresponds to the problem of selecting factors in PLS, but in the ALS case there is no prediction error (*e.g.* RMSEP) to minimise. In the present case the obvious choice would have been three components, as this gave the most stable result. Only the comparison of the profiles to actual lactose results indicated that four components were optimal. Methods for determining the number of independently varying species present in the samples are therefore required.

Scores obtained through Principal Component Analysis (PCA) (10) could solve the problem. In the present case, when performing a PCA on all calibration samples, the scores (not shown) contain structure (originating from the batch structure of the data) revealing up to four or five components. Thus, four or five components would be expected to be optimal in ALS, which supports the actual findings shown above.

In this case we were successful in extracting concentrations from FT-IR data from a lactose hydrolysis process and to monitor new process runs using this knowledge. The present data set must be considered to be worst case, as the infrared spectra of reactants, intermediates and products are very similar, *i.e.* they are all sugars which give roughly the same absorption peaks. Resolving spectra from processes where the compounds are much less similar should therefore be easier, especially the problem of determining the number of components should be less difficult.

CONCLUSIONS

The present study has shown that ALS is a promising method for use in dairy process optimisation. Without the need for reference analyses it was possible to extract four components from lactose hydrolysis process data (fat, lactose and two other sugar components) and to obtain a lactose prediction error similar to the one obtained from an ordinary PLS regression. Such use of ALS for reference-independent prediction of process parameters is not limited to dairy products only, but is likely to be useful for process monitoring and identification of intermediates in all branches of the food and beverage industry.

By use of ALS combined with FT-IR it becomes possible to obtain quick information on compounds present during the process, but not necessarily by the end of it. A further advantage (in many cases the most important) is that the pure spectra obtained by ALS makes it possible to generate predictions of process parameters without the need for expensive and time consuming reference procedures.

ACKNOWLEDGMENTS

The authors would like to thank Foss A/S, Denmark, and The Danish Academy of Technical Sciences, for providing the funds for P. W. Hansen's work. E. Conaghan is thanked for his help in relation to the practical measurements. R. Bro and C. Ridder are thanked for useful discussions and help in the implementation of the mathematical routines.

REFERENCES

1 Bro, R. 1997. PARAFAC. Tutorial and applications. Chemomet. Intell. Lab. Sys. 38:149-171.

2 Forsman, E., M. Heikonen, L. Kiviniemi, M. Kreula, and P. Linko. 1979. Kinetic investigations of the hydrolysis of milk lactose with soluble Kluyveromyces lactis β-galactosidase. Milchwissenschaft 34:618-621.

3 Furusjö, E., L.-G. Danielsson, E. Könberg, M. Rentsch-Jonas, and B. Skagerberg. 1998. Evaluation Techniques for Two-Way Data from in Situ Fourier Transform Mid-Infrared Reaction Monitoring in Aqueous Solution. Anal. Chem. 70:1726-1734.

4 Greenberg, N. A., and R. R. Mahoney. 1983. Formation of Oligosaccharides by β-Galactosidase from Streptococcus thermophilus. Food Chem. 10:195-204.

5 Guy, E., and E. Bingham. 1978. Properties of β-galactosidase of Saccharomyces lactis in milk and milk products. J. Dairy Sci. 61:147-151.

6 Hansen, P. W. 1998. Urea determination in milk using Fourier transform infrared spectroscopy and multivariate calibration. Milchwissenschaft 53: 251-255.

7 Hewavitharana, A. K., and B. van Brakel. 1997. Fourier Transform Infrared Spectrometric Method for the Rapid Determination of Casein in Raw Milk. Analyst 122:701-704.

8 Karjalainen, E. J. 1989. The Spectrum Reconstitution Problem. Use of Alternating Regression for Unexpected Spectral Components in Two-Dimensional Spectroscopies. Chemomet. Intell. Lab. Sys. 7:31-38.

9 Kjær. L. 1997. Say cheese – and think of direct casein determination. Scand. Dairy Inform. 2/97: 28-30.

10 Martens, H., and T. Næs. 1989. Multivariate Calibration. Wiley, Chichester.

11 Prenosil, J. E., E. Stuker, and J. R. Bourne. 1987. Formation of oligosaccharides during enzymatic lactose hydrolysis and their importance in a whey hydrolysis process: Part I: State of the art. Biotechnol. Bioeng. 30:1019:1025.

12 Prenosil, J. E., E. Stuker, and J. R. Bourne. 1987. Formation of oligosaccharides during enzymatic lactose hydrolysis and their importance in a whey hydrolysis process: Part II. Biotechnol. Bioeng. 30:1026:1031.

13 Ridder, C., and L. Kjær. 1995. Sweet dreams come true. Applied FTIR technology in ice cream mix analysis. Scand. Dairy Inform. 9/95:34-36.

14 Tauler, R., B. Kowalski, and S. Fleming. 1993. Multivariate Curve Resolution Applied to Spectral Data from Multiple Runs of an Industrial Process. Anal. Chem. 65:2040-2047.

15 Tauler, R., A. Smilde, and B. Kowalski. 1995. Selectivity, Local Rank, Three-Way Data Analysis and Ambiguity in Multivariate Curve Resolution. J. Chemomet. 9:31-58.

Publication VII

Evaluation of Spectroscopic Data

Danish Patent Application No. PA 1998 01177

Evaluation of Spectroscopic Data

Per Waaben Hansen
R&D Greenhouse, Foss Electric A/S,
Slangerupgade 69, P.O. Box 260, DK-3400 Hillerød, Denmark

ABSTRACT

Dairy process monitoring by application of multivariate curve resolution using Alternating Least Squares is presented. Alternating Least Squares was used for resolving Fourier transform infrared spectral data from a dairy batch process in which lactose is enzymatically hydrolysed to glucose and galactose. It was possible to extract four compounds (fat, lactose and two other sugar components) from the spectral data obtained from nine process runs. Subsequently, the pure spectra obtained in this way were used to monitor the content of these compounds in two new process runs. In this way Alternating Least Squares made it possible to follow the hydrolysis process by Fourier transform infrared spectroscopy without the need for reference analyses. When the results were correlated to reference results for lactose the accuracy was similar to what was obtained when performing a Partial Least Squares regression on the same data: A lactose correlation of 0.980 was obtained when Alternating Least Squares was used and 0.987 when Partial Least Squares was used.

The present invention relates to evaluation of spectroscopic data, and more specifically the use of multivariate curve resolution, such as alternating least squares, ALS.

TECHNICAL BACKGROUND

Process control in industrial processes is increasingly important as on-line analytical equipment providing fast and reliable results becomes available. Near infrared reflection/transmission (NIR/NIT) spectroscopy is the most frequently used method in many branches of industry, while mid-infrared (MIR) spectroscopy has proved very useful for process milk analysis. Milk analysis using MIR equipment is generally more accurate than the corresponding NIR/NIT method. This is because MIR contains more specific information (fundamental absorptions) and stronger signals than NIR/NIT, which detects derived information (overtones and combination bands). In addition, full spectrum instruments based on Fourier transform infrared (FT-IR) spectroscopy for dairy product analysis in the laboratory are showing promising results with regard to the number of components (*e.g.* specific sugars (13), casein (7, 9), urea (6)) that can be measured.

Hydrolysis of lactose in milk is of interest as large racial groups suffer from lactose intolerance, *i.e.* they are not able to cleave lactose into glucose and galactose. Therefore, low-lactose milk products produced by the action of the enzyme β-galactosidase are of commercial interest. The process is sensitive to the initial conditions such as temperature and β-galactosidase concentration (2, 5). Therefore, the concentrations of the sugars need to be monitored during the course of the reaction, to control the process. The reaction is typically completed within a few hours, thus it requires a fast analytical method such as MIR.

The general approach when analysing spectral data with the intention of generating future predictions of milk constituents is to use one of several multivariate methods relating the data to wet chemistry results. These methods include Partial Least Squares (PLS) regression which is described elsewhere (10).

The problem is that ordinary multivariate methods require an accurate and reproducible reference method to obtain a reliable calibration. This tends to be resource-demanding – especially when the typical number of calibration samples (fifteen to hundreds) is taken into account. In addition, during a process some intermediate species might only exist for a limited period of time, *i.e.* they are both produced and consumed during the reaction. In such a case it might be difficult to isolate the intermediates and to measure them using the reference methods.

Such problems can be solved using a regression method of higher order, such as PARAFAC (Parallel Factor Analysis) (1). MLR, PCR and PLS handle second order data, as the data can be arranged in a matrix. PARAFAC requires the data to be of an order higher than two, which is the case when *e.g.* a spectral landscape is obtained for each sample and the whole data set can be arranged in a cube. The landscape could be obtained by measuring a reacting sample with fixed time intervals during the process. The individual spectra constituting the landscape will be related and these relationships implicitly contain information on the concentrations of all infrared absorbing compounds in the sample.

PARAFAC is able to resolve these variations and to produce concentration profiles and pure spectra corresponding to the absorbing species present in the sample. The concentrations will be arbitrary, but proportional to the true concentrations. If correlating species are present, the concentration profiles will be the sums of such correlating compounds. PARAFAC is normally the most useful method for multivariate curve resolution, as it can handle more than one sample at a time, the solutions to the mathematical problem are unique (there is only one solution to each problem), and they might resemble real spectra and concentrations. Good results have been obtained on resolving absorption and emission profiles from fluorescence spectra of sugar samples using PARAFAC (1).

In the present case, PARAFAC will not work, as the actual shape of the concentration profiles will be strongly dependent on the initial conditions of the process. In such a case, it will only be able to analyse one landscape (sample) at a time, or the unfolded data set. When a landscape is unfolded (*e.g.* when the spectra from the individual runs are appended to each other), one of the directions in the three-dimensional structure is lost.

Alternating Least Squares (ALS), sometimes referred to as Alternating Regression (8), is a two-way method which handles one landscape or unfolded data set at a time. Ref. (15) shows how such curve resolution methods are working in practice. ALS produces pure spectra and concentration profiles in a way similar to PARAFAC performed on the unfolded data set. The method has been used for resolution of infrared process data with excellent results (3, 14).

REFERENCES

1. Bro, R. 1997. PARAFAC. Tutorial and applications. Chemomet. Intell. Lab. Sys. 38:149-171.

2. Forsman, E., M. Heikonen, L. Kiviniemi, M. Kreula, and P. Linko. 1979. Kinetic investigations of the hydrolysis of milk lactose with soluble Kluyveromyces lactis β-galactosidase. Milchwissenschaft 34: 618-621.

3. Furusjö, E., L.-G. Danielsson, E. Könberg, M. Rentsch-Jonas, and B. Skagerberg. 1998. Evaluation Techniques for Two-Way Data from in Situ Fourier Transform Mid-Infrared Reaction Monitoring in Aqueous Solution. Anal. Chem. 70:1726-1734.

4. Greenberg, N. A., and R. R. Mahoney. 1983. Formation of Oligosaccharides by β-Galactosidase from Streptococcus thermophilus. Food Chem. 10:195-204.

5. Guy, E., and E. Bingham. 1978. Properties of β-galactosidase of Saccharomyces lactis in milk and milk products. J. Dairy Sci. 61:147-151.

6. Hansen, P. W. 1998. Urea determination in milk using Fourier transform infrared spectroscopy and multivariate calibration. Milchwissenschaft 53:251-255.

7. Hewavitharana, A. K., and B. van Brakel. 1997. Fourier Transform Infrared Spectrometric Method for the Rapid Determination of Casein in Raw Milk. Analyst 122:701-704.

8. Karjalainen, E. J. 1989. The Spectrum Reconstitution Problem. Use of Alternating Regression for Unexpected Spectral Components in Two-Dimensional Spectroscopies. Chemomet. Intell. Lab. Sys. 7:31-38.

9. Kjær. L. 1997. Say cheese – and think of direct casein determination. Scand. Dairy Inform. 2/97:28-30.

10. Martens, H., and T. Næs. 1989. Multivariate Calibration. Wiley, Chichester.

11. Prenosil, J. E., E. Stuker, and J. R. Bourne. 1987. Formation of oligosaccharides during enzymatic lactose hydrolysis and their importance in a whey hydrolysis process: Part I: State of art. Biotechnol. Bioeng. 30:1019:1025.

12. Prenosil, J. E., E. Stuker, and J. R. Bourne. 1987. Formation of oligosaccharides during enzymatic lactose hydrolysis and their importance in a whey hydrolysis process: Part II: Experimental. Biotechnol. Bioeng. 30:1026:1031.

13. Ridder, C., and L. Kjær. 1995. Sweet dreams come true. Applied FTIR technology in ice cream mix analysis. Scand. Dairy Inform. 9/95:34-36.

14. Tauler, R., B. Kowalski, and S. Fleming. 1993. Multivariate Curve Resolution Applied to Spectral Data from Multiple Runs of an Industrial Process. Anal. Chem. 65:2040-2047.

15. Tauler, R., A. Smilde, and B. Kowalski. 1995. Selectivity, Local Rank, Three-Way Data Analysis and Ambiguity in Multivariate Curve Resolution. J. Chemomet. 9:31-58.

In the following description the below stated abbreviations will be used:

ALS = Alternating Least Squares, AR = Alternating Regression, CVS = Cross Validation Segments, FT-IR = Fourier Transform Infrared, MIR = Mid-Infrared, NIR = Near Infrared Reflection, NIT = Near Infrared Transmission, PLS = Partial Least Squares, RMSEP = Root Mean Square Error of Prediction, SEC = Standard Error of Calibration, SEP = Standard Error of Prediction, SSE = Sum of Squared Errors, R^2 = Correlation, s_r = Repeatability.

The aim of the present work is to provide a spectroscopic method for monitoring processes in the food industry, and specifically dairy processes, *e.g.* the lactose hydrolysis process. Further it is the aim to provide a method for generating calibrations without using reference analysed samples. Specifically on-line monitoring is contemplated.

SUMMARY OF THE INVENTION

The present invention relates to a method for evaluation of spectroscopic data, comprising the use of multivariate curve resolution. According to the invention said method is used for determining the properties or constituents in food products, such as dairy products and/or monitoring a food process, *e.g.* a dairy process. Further the method may be used for generating a calibration to be used for future determining the properties of, or constituents in, food products, such as dairy products and/or monitoring a dairy process.

The method used according to the invention is specifically advantageous in that no reference samples has to be measured by a reference method; the method can be used for monitoring processes. Further the method can be used to recognise (or identify) unknown con-

stituents in the sample and predict their relative concentrations.

Specifically the method can be used to provide calibrations for a spectroscopic instrument, such as an FTIR instrument arranged to monitor a process such as the lactose hydrolysis process. The calibrations are derived on basis of the multivariate curve resolution and the calibrations can then be applied for future monitoring of the process.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows an infrared landscape from a chemical process.

Figure 2 is a reference versus predicted plot for lactose showing 44 calibration samples.

Figure 3 shows a reference versus predicted plot for lactose showing 23 test samples predicted using a PLS model with 5 factors.

Figure 4 shows the best three component ALS solution with lowest SSE (out of 100 runs) for set 1. Upper part shows the concentration profiles, lower part the pure spectra.

Figure 5 shows the best three component ALS solution with lowest SSE (out of 100 runs) for set 2. Upper part shows the concentration profiles, lower part the pure spectra.

Figure 6 shows the four component ALS solution with lowest SSE for set 1. Upper part shows the concentration profiles, lower part the pure spectra.

Figure 7 shows the four component ALS solution with lowest SSE for set 2. Upper part shows the concentration profiles, lower part the pure spectra.

Figure 8 shows a lactose profile and reference lactose results plotted against each other. The profile relates to calibration set 2 with 3 components.

Figure 9 shows a lactose profile and reference lactose results plotted against each other. The profile relates to calibration set 2 with 4 components.

Figure 10 A and B: Concentration profiles for 2 test runs (or batches).

DETAILED DESCRIPTION

In the following description a detailed explanation of a preferred method according to the invention will be given as a non limiting example, which is based on the use of Alternating Least Squares, ALS. Other methods according to the invention include: Curve Resolution, such as Multivariate Curve Resolution, and more specifically *e.g.* Alternating Regression (AR), Simplisma, Generalised Rank Annihilation Method (GRAM), Evolving Factor Analysis (EFA), Rank Annihilation Factor Analysis (RAFA), and Trilinear Decomposition (TLD).

Alternating Least Squares (ALS) relies on the assumption that the Beer-Lambert law is correct, *i.e.* that a spectrum (the row vector x) of a given sample can be seen as a linear combination of the pure constituent spectra (contained in the matrix A), thus

$$\mathbf{x} = \mathbf{cA} \quad [1]$$

where **c** is a row vector containing the concentrations of the constituents corresponding to the pure spectra in **A**. In the case where more than one spectrum is measured, the general expression becomes

$$\mathbf{X} = \mathbf{CA} \quad [2]$$

where **X** is a landscape containing the spectra in its rows and **C** is a matrix containing the concentrations corresponding to each spectrum. In this context one sample is named **X**, *i.e.* a collection of spectra from one process run.

A typical landscape obtained during 60 minutes (with sampling every 10 minutes) from one lactose hydrolysis run with seven FT-IR spectral recordings is presented in

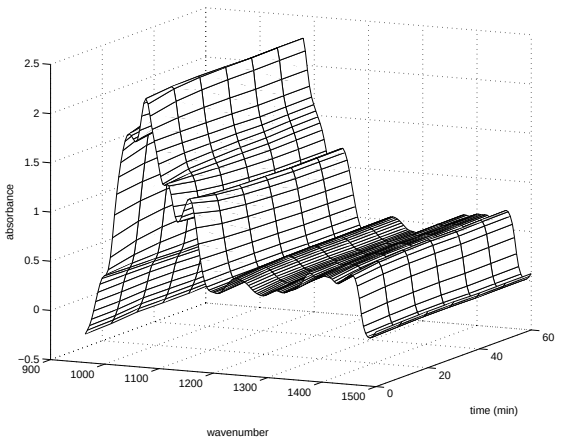


Figure 1 An infrared landscape from a chemical process.

Figure 1. Most of the variation in the spectra is in the range between 1000 and 1200 cm^{-1}. This is expected, as the only compounds affected by the reaction are sugars which show strong absorptions due to stretching of the sugar C-O bonds in this range. ALS calculates the pure spectra **A** from the input spectra in **X** (the landscape) and an estimate of **C** (*e.g.* random numbers) using a re-arranged form of [2]:

$$\mathbf{A} = \mathbf{C}^{+}\mathbf{X} \quad [3]$$

where $\mathbf{C}^{+}$ is some pseudo-inverse of **C**, followed by a calculation of a better guess of **C** from this **A**:

$$\mathbf{C} = \mathbf{X}\mathbf{A}^{+} \quad [4]$$

where $\mathbf{A}^{+}$ is the pseudo-inverse of **A**. The steps [3] and [4] are repeated until convergence (or a maximal number of iterations) has been reached.

Various constraints can be applied to the spectra and concentration profiles in **A** and **C** in order to avoid physically meaningless solutions (1). For example, the concentrations in **C** cannot possibly become negative, so a *non-negativity constraint* would be reasonable. In addition, when looking at compounds produced and consumed during a chemical process, the concentration profiles will be expected to have only one maximum during the course of the reaction. This leads us to applying the *unimodality constraint* which limits solutions to smooth concentration curves with only one maximum each. It can be argued that a non-negativity constraint will be appropriate for the pure spectra in **A** as well. In this specific application it is not. As the FT-IR absorbance spectra are calculated using a water background slightly negative absorbances will result. Thus, constraining **A** to non-negativity will restrict the algorithm too much.

Non-negativity can be applied in various ways. The most straightforward approach is to force negative values to zero (*e.g.* in **C**) after each iteration. This is very simple and does not necessarily lead to the optimal description of **X** (*i.e.* the least squares solution). The approach employed here (adopted from The N-Way Toolbox by C. A. Andersson, Internet site: http://newton.food-sci.kvl.dk/matlab/nwaytoolbox) forces only one concentration profile (or spectrum) at a time to zero, followed by a correction of the pure spectrum matrix, **A** (or concentration matrix, **C**). This modification leads to the optimal result.

In case of more process runs contained in the same data matrix **X** the unimodality constraint of **C** would not work. In such a case only the parts of **C** originating from the same process should be constrained. This approach was used in the present work.

After the concentration profiles and pure spectra have been obtained the same principle can be used for prediction of the constituents in an unknown sample by applying the vector form of [4] to the spectrum of the sample:

$$\mathbf{c} = \mathbf{x}\mathbf{A}^{+} \quad [5]$$

The concentration row vector **c** will be in arbitrary units, but linearly related to the actual concentrations.

Calculations

The data analysis and calibration may be performed on a computer, *e.g.* a PC preferably using Matlab 5.2.1 software (The MathWorks Inc., MA, USA). The pseudo-inverse in equations [3] and [4] can be calculated using the built-in functions of

Matlab. The calibration routines can be taken from the PLS_Toolbox Version 1.5 (Eigenvector Technologies, WA, USA).

Repeatability is expressed as a mean standard deviation (s_r) of multiple determinations performed under identical conditions and is calculated as:

$$s_r = \sqrt{\frac{1}{q(n-1)} \sum_{j=1}^{q} \sum_{i=1}^{n} (x_{j,i} - \bar{x}_j)^2}$$

where q is the number of samples, n is the number of replicates, $x_{j,i}$ is the result of the i'th replicate of the j'th sample and $\bar{x}_j$ is the average result of the j'th sample.

Accuracy is expressed as the Root Mean Square Error of Prediction (RMSEP) and is calculated as:

$$\mathrm{RMSEP} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{i,\mathrm{reference}} - x_{i,\mathrm{predicted}})^2}$$

where N is the number of determinations (number of samples (q) times number of replicates (n) from above) and $x_{i,\mathrm{reference}}$ and $x_{i,\mathrm{predicted}}$ are the reference and predicted values corresponding to the i'th determination, respectively.

When a bias (mean difference between reference results and predictions) is observed, the Standard Error of Prediction (SEP) is used. It is calculated as:

$$\mathrm{SEP} = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_{i,\mathrm{reference}} - x_{i,\mathrm{predicted}} - \mathrm{bias})^2}$$

If two variables are related by performing a univariate linear regression (slope a and intercept b), between instrumental responses, $x_{i,\mathrm{instrumental}}$, and reference results, $x_{i,\mathrm{reference}}$, the accuracy of future predictions can be estimated by the use of the Standard Error of Calibration (SEC), calculated as:

$$\mathrm{SEC} = \sqrt{\frac{1}{N-2} \sum_{i=1}^{N} (x_{i,\mathrm{reference}} - (a x_{i,\mathrm{instrumental}} + b))^2}$$

Correlation is expressed as R^2, which is calculated as:

$$R^2 = \left[\frac{\frac{1}{N} \sum_{i=1}^{N} (x_{i,\mathrm{reference}} - \bar{x}_{\mathrm{reference}})(x_{i,\mathrm{predicted}} - \bar{x}_{\mathrm{predicted}})}{s_{\mathrm{reference}} s_{\mathrm{predicted}}} \right]^2$$

where N, $x_{i,\mathrm{reference}}$ and $x_{i,\mathrm{predicted}}$ are defined above and $\bar{x}_{\mathrm{reference}}$, $s_{\mathrm{reference}}$, $\bar{x}_{\mathrm{predicted}}$ and $s_{\mathrm{predicted}}$ are the mean and standard deviations of the reference and predicted results, respectively.

Finally, the ***fit of the model*** to **X** (equation [2]) is expressed as the Sum of Squared Errors (SSE):

$$\mathrm{SSE} = \sum_{i=1}^{N} \sum_{j=1}^{M} (x_{i,j} - \mathbf{c}_i \mathbf{a}_j)^2$$

where $x_{i,j}$ is an element in **X**, c_i is a row vector containing the concentrations of the i'th sample and a_j is a column vector containing the absorbencies of the j'th wavelength. M is the number of wavelengths in the spectra. Note that the reference results have not been used in the calculation of the SSE.

Calibration samples

This set contains 124 samples. They were collected from nine process runs (five based on skim milk, four based on whole milk) carried out in May 1997 using an experimental set-up in the laboratory. Lactozym 3000 (Novo-Nordisk, Bagsværd, Denmark) was the enzyme used. Samples were taken from the reaction mixture at various time points over a three hour period, and they were immediately heated to 80 °C in a 750 W microwave oven in order to deactivate the enzyme. Duplicate samples were taken, and the following reference analyses and spectral measurements were carried out independently. Thus, the set of 124 samples comprises two very similar sets of 62 samples.

	2 [1]CVS	4 [1]CVS	6 [1]CVS	8 [1]CVS	10 [1]CVS
[2]#factors	5	5	5	5	4
R^2	0.997	0.996	0.996	0.996	0.996
RMSEP	0.82	0.90	0.88	0.93	0.88
s_r	0.22	0.23	0.23	0.23	0.24

Table 1 PLS results from the cross-validated calibrations for the determination of lactose in the calibration samples. 44 samples ranging from 0 to 45 % dry base lactose were used.
[1]CVS = Cross-Validation Segments
[2]#factors = optimal number of PLS factors

Test samples

This set contains 23 samples obtained from two process runs carried out in the laboratory in November and December 1997 using the same experimental set-up. Samples were taken at various intervals, and this time only the sub-sample used for reference analysis had the enzyme deactivated. The reference analysis described here is carried out only to check the applicability of the new method, it is not included in the method according to the invention. The spectral measurement was carried out on the non-deactivated sample immediately (*i.e.* less than one minute) after sampling in order to make the FT-IR measurements as close to an on-line application as possible. The sub-samples for reference analysis were still subjected to a heat treatment.

Reference measurements

These were carried out using refractive index detection. Lactose was determined on 90 of the calibration samples, as well as the 23 test samples, using a HPLC set-up.

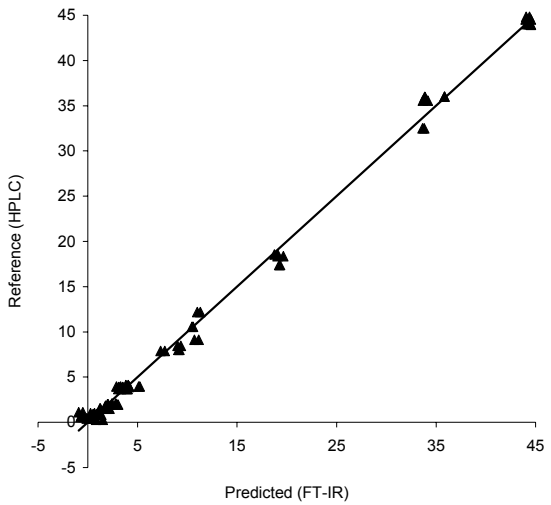


Figure 2 Reference versus predicted plot for lactose showing 44 calibration samples.

Spectral measurements

The FT-IR spectral measurements were carried out using a MilkoScan FT 120 (Foss, Hillerød, Denmark). The infrared spectrum from 925 to 5000 cm^{-1} was recorded. The calibration samples were measured in duplicate and the test samples were measured in triplicate.

In the data analysis only the ranges 964-1542, 1724-1847 and 2699-2965 cm^{-1} were used, as these are the areas containing useful chemical information.

All measurements were ratioed against water and log-transformed to give absorbance spectra. A typical time-resolved landscape for the first hour of an experiment is shown in Figure 1. No spectral pre-processing was performed prior to data analysis, as experience shows that only minor improvements, if any, are obtained when using full spectrum data of the present type (6).

RESULTS AND DISCUSSION

PLS Calibration

For comparative purposes, a PLS calibration against the lactose reference results was performed. This is not a part of the invention, it was done for comparative purposes only. The cross validated results are shown in Table 1 and a reference vs. measured plot for lactose is shown in Figure 2. From these

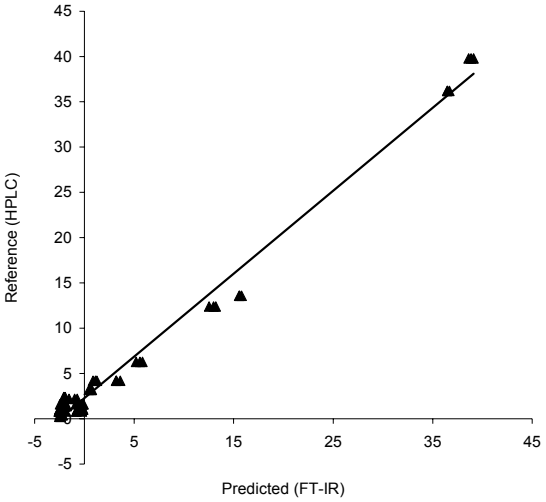


Figure 3 Reference versus predicted plot for lactose showing 23 test samples predicted using a PLS model with 5 factors.

results, a model using 5 PLS factors is the best. The corresponding result on the test set was an error (RMSEP) of 2.49. This result is shown in Figure 3. The reference results and lactose predictions correlate well which is the main issue in this context.

ALS Results

To obtain reasonable solutions, two constraints were applied during the regression:

- Non-negativity of the concentration profiles
- Unimodality of the concentration profiles (except for fat) in each process run
- Non-negativity of the spectra is a further possible constraint.

The calculations using regression are started by choosing initial values of the components. The following – very simple – start guesses were used:

- The skim and whole milk samples had a start guess for fat of 0 and 1, respectively
- As lactose is known to be the only sugar present in the beginning of the process, the lactose concentration of the first sample of each run was set to 1, while all others were set to 0
- The remaining components were given random numbers as start guesses (uniformly distributed numbers between 0 and 1)

As random numbers were used, many different solutions might result. For this reason the ALS run was carried out 100 times

	3 components		4 components		5 components	
	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2
SSE	0.56	1.01	0.38	0.44	0.26	0.47
R^2	0.928	0.942	0.942	0.971	0.978	0.977
SEC	3.47	3.15	3.12	2.23	1.92	2.00
A: Results of the calibration sets (62 samples each) – the models selected on the basis of the lowest SSE.						
SSE	0.71	1.06	0.44	51.63	0.40	10.50
R^2	0.928	0.942	0.971	0.986	0.991	0.992
SEC	3.47	3.15	2.22	1.54	1.26	1.16
B: Results of the calibration sets (62 samples each) – the models selected on the basis of the highest correlation (R^2) to the lactose reference results.						
R^2	0.894	0.891	0.959	0.980	0.959	0.940
SEC	3.46	3.50	2.16	1.51	2.14	2.61
s_r	0.15	0.15	0.16	0.27	0.16	0.13
C: Results of the test set (23 samples) using the models (selected by use of the SSE) obtained from the individual calibration sets.						
R^2	0.894	0.891	0.959	0.953	0.978	0.973
SEC	3.46	3.50	2.15	2.29	1.59	1.78
s_r	0.15	0.15	0.16	0.14	0.19	0.19
D: Results of the test set (23 samples) using the models (selected by use of R^2) obtained from the individual calibration sets.						

Table 2 ALS results of the two calibration sets (part A and B) and the independent test set (part C and D) using 3, 4 and 5 components. The reference results range from 0 to 45 % dry base lactose.

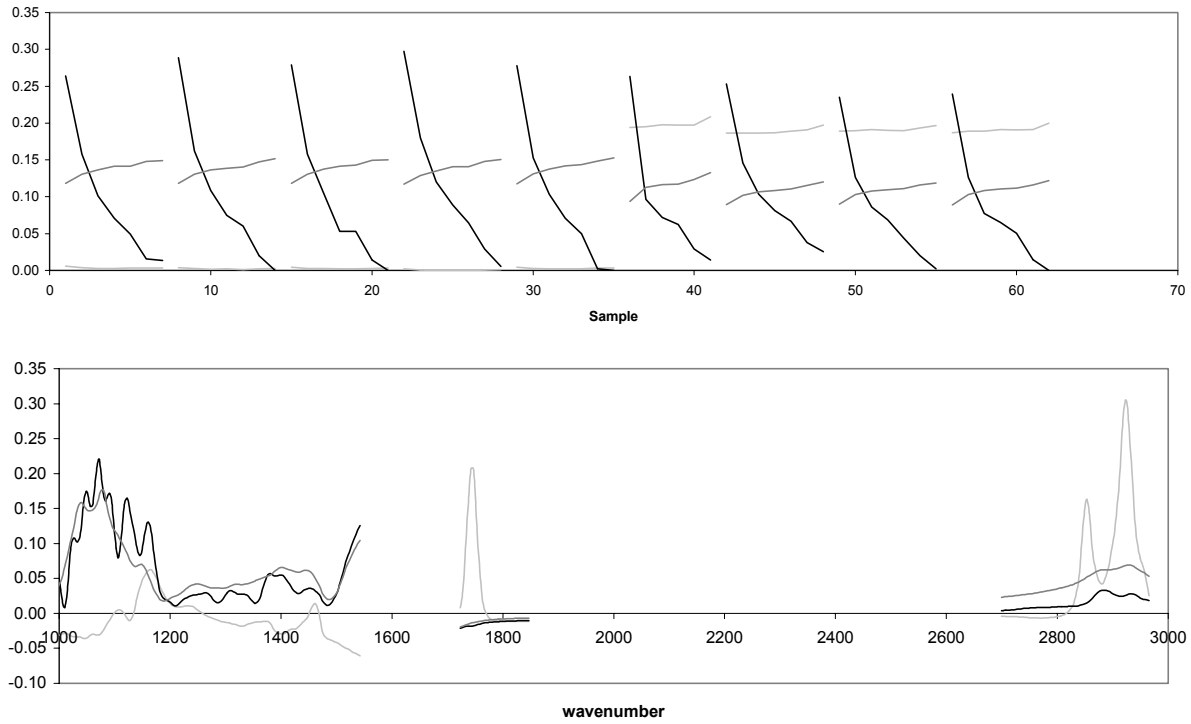


Figure 4 The best 3 component ALS solution with lowest SSE (out of 100 runs) for set 1. Upper part shows the concentration profiles, lower part the pure spectra.

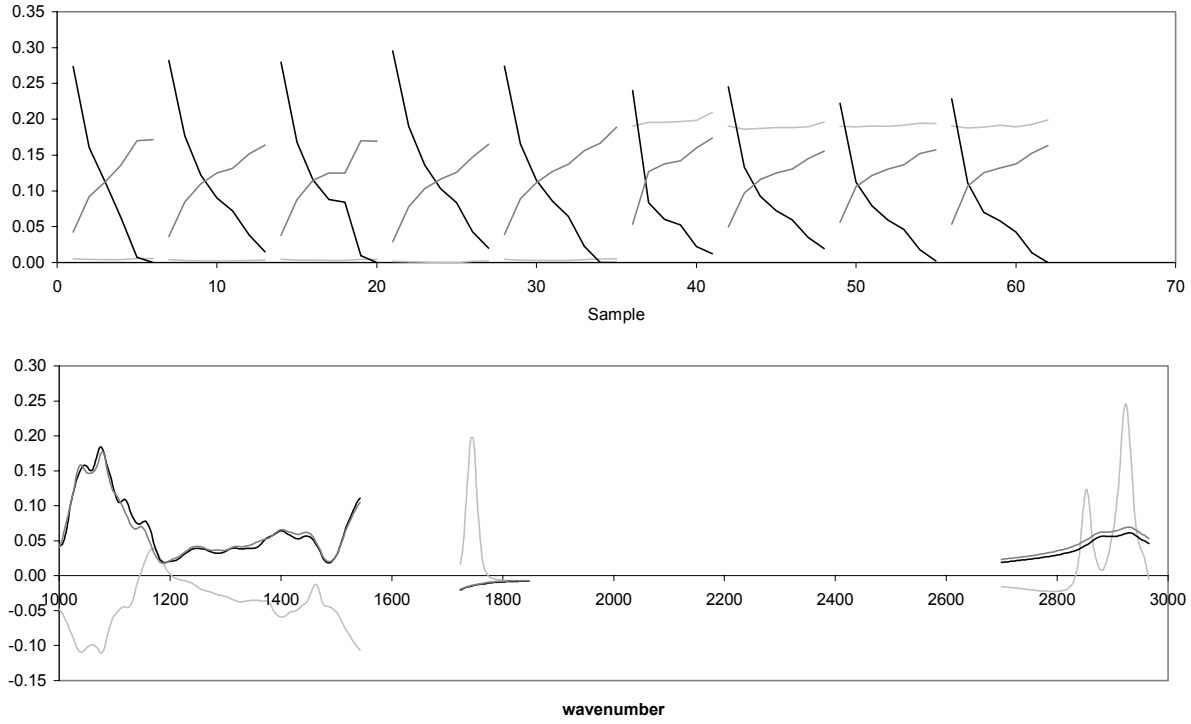


Figure 5 The best 3 component ALS solution with lowest SSE (out of 100 runs) for set 2. Upper part shows the concentration profiles, lower part the pure spectra.

with new start guesses for each number of components. From three to six components (corresponding to the number of pure spectra in **A**, equation [2]) were tried on both (almost identical) calibration sets of 62 samples. When six components were used, the concentration profiles and pure spectra became noisy and highly correlated, so they will not be discussed in the following text.

The best models, in terms of how well the **X** matrix is described (measured as SSE) and

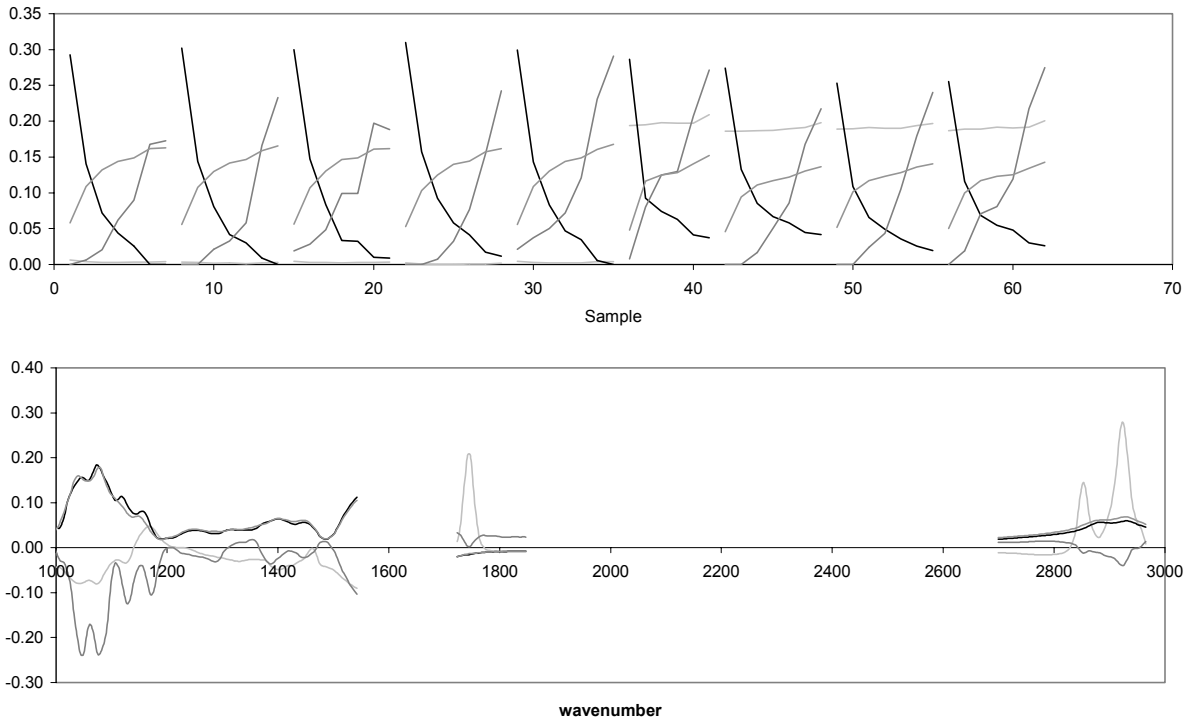


Figure 6 The four component ALS solution with lowest SSE for set 1. Upper part shows the concentration profiles, lower part the pure spectra.

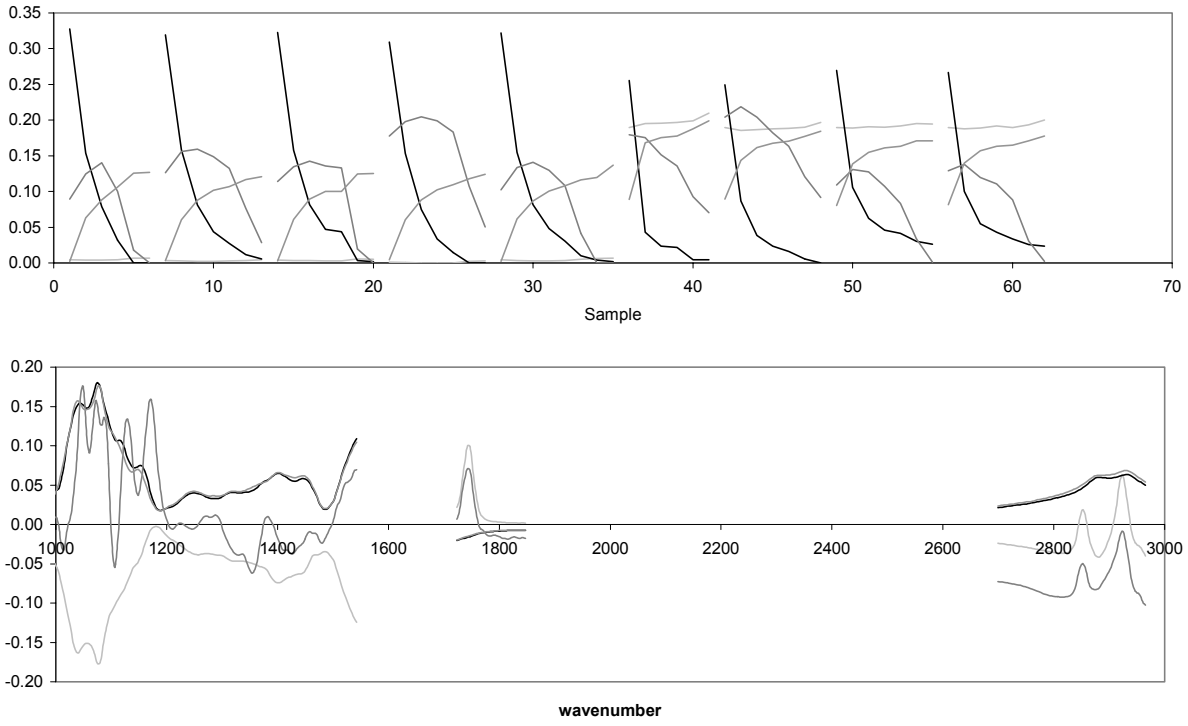


Figure 7 The four component ALS solution with lowest SSE for set 2. Upper part shows the concentration profiles, lower part the pure spectra.

how well the "lactose" profile correlates with the reference results (measured as R^2), are shown in Table 2 (A and B parts). The results for three to five components were as follows:

With three components, the same pure spectra and concentration profiles were reached every time on both calibration sets – at least the differences were insignificant. The solutions with the lowest SSE are shown in Figures 4 and 5. Note that both pure spectra and con-

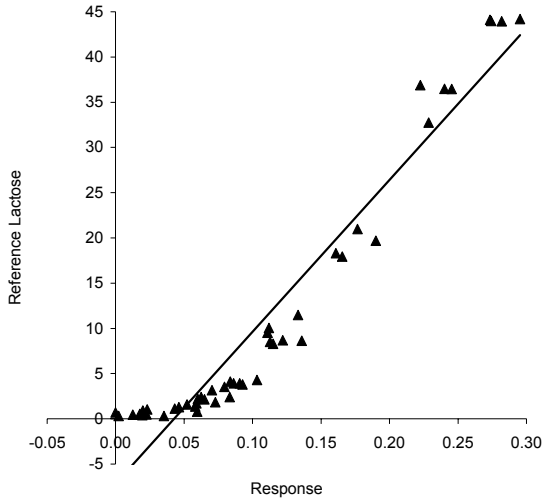


Figure 8 A lactose profile and reference lactose results plotted against each other. The profile relates to calibration set 2 with 3 components.

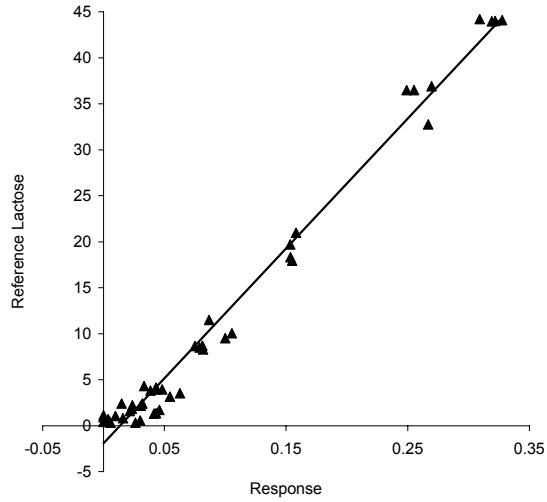


Figure 9 A lactose profile and reference lactose results plotted against each other. The profile relates to calibration set 2 with 4 components.

centration profiles have been normalised to make a presentation on the same scale possible. The pure spectra describing fat (having strong absorptions in the high end of the spectrum due to C-H stretching vibrations) have very negative contributions in the areas where the sugars absorb, reflecting that high-fat samples generally contain less lactose as a result of the displacement of the water phase by fat.

With four components many different solutions were reached. They belonged to a limited number of groups of solutions inside which the variations were small. Some results are shown in Figures 6 and 7. Of the four components at least two are sugars (having strong absorptions in the 1000 to 1200 cm^{-1} range), one is fat (strong absorptions between 2800 and 3000 cm^{-1}), while the last one is difficult to assign. Of the sugars, the component decreasing rapidly through each batch is lactose.

With five components the problem of finding the optimal solution becomes even more difficult. But, as is evident from Table 2, good lactose correlations were still obtained. As there is no major improvement when going from four to five components, the three- and four-component solutions were chosen for further examination.

An ALS with three components gives the most stable result and the most reasonable pure spectra, but the correlation to the actual lactose concentration is relatively poor. In Figure 8 the lactose concentration profiles are plotted against the reference results, and the resulting plot is clearly non-linear. The four-component model (Figure 9) gives a much more linear relationship to the lactose reference results. It cannot be due to overfitting (*i.e.* a too optimistic estimate of the error) as the reference results were not involved in the optimisation. In addition, the concentration profiles of the four-component model agree with previous observations that not only the monosaccharides (galactose and glucose), but also various other sugars (containing two or more monosaccharide units), generally known as the oligosaccharides, are formed during the process (4, 11, 12). The shapes of the concentration profiles are actually very similar to these previous observations. The way in which these sugars are distributed among the last two components (Figures 6 and 7) can vary between different ALS runs, which is the reason for the many different solutions seen when four (and five) components are tried.

The SSE and/or the R^2 may be used as the selection criterion. When three components are used, both criteria give almost the same result, while the SSE criterion gives a somewhat higher prediction error in the case of four and five components. In both cases there is a significant improvement over the three-component result. As use of SSE and R^2

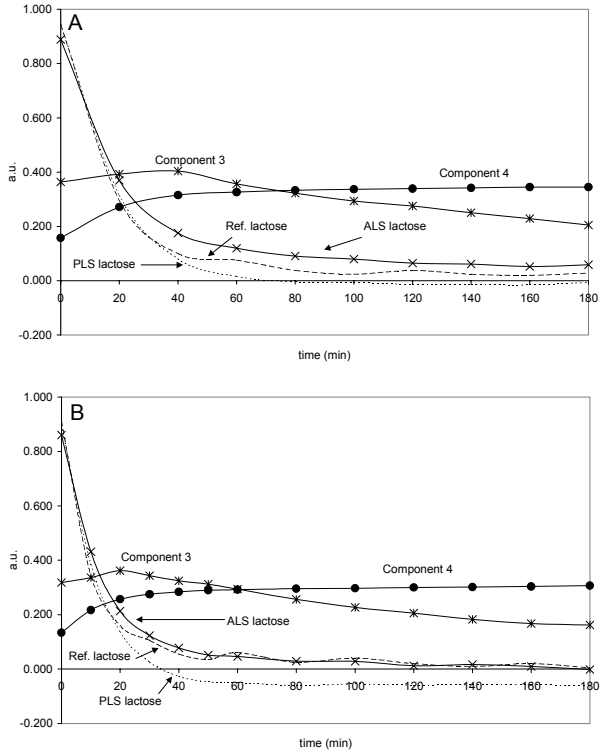


Figure 10 Concentration profiles for 2 test runs (or batches).

give similar results SSE is preferred as it does not involve reference data.

The final test of the new method was done by using the ALS models on the test set obtained from two new process runs. The results are shown in Table 2 (C and D parts). There is a large improvement in SEC when going from three to four components. It is almost independent of the method (R^2 or SSE) used for selecting the optimal model. These results lead to the conclusion that there is nothing gained by selecting the optimal model by use of the lactose reference results (*i.e.* by looking at R^2). In fact, the best result is obtained by using the SSE. This is a very promising result, since there is no need for reference analyses when process data of the present type are analysed. ALS alone can be used for generating a model which can be used for future process monitoring. Note that the ALS concentrations are in arbitrary units, so only relative process changes can be detected.

The results in Table 2 should be compared to the PLS result shown in Figure 3. It is seen that the ALS predictions of lactose (using four components) are almost as good as the PLS predictions – in some cases the results are the same.

The normalised predictions of three of the four components (using the model with lowest SSE based on calibration set 2) for each of the two processes included in the test set are shown in Figure 10 together with the corresponding results from the PLS model and the reference method. Fat is omitted, as it is constant during each run. The ALS lactose predictions do not agree perfectly, neither with the reference nor the PLS results based on the same spectra, but they follow roughly the same curve. The main reason for the disagreement between the ALS and PLS results are the negative PLS lactose predictions which are due to the earlier discussed bias of the PLS calibration. Both PLS and ALS concentration profiles follow a smooth curve, which should be expected when dealing with a chemical reaction. Thus, the fluctuations in the lactose reference results are likely to be caused by the lack of reproducibility of the reference method rather than real variations in the lactose content.

The two test runs (Figure 10) gave the same shapes of the concentration profiles of the third and fourth component as seen during calibration. Therefore Component 3 is assigned to the sum of galactose and glucose, and Component 4 is likely to be caused by oligosaccharides formed during the reaction. Another data set with reliable reference results on other sugars is required to confirm this.

The remaining problem allowing implementation of ALS for practical use in dairy process monitoring is how to select the optimal number of components in the ALS model. This corresponds to the problem of selecting factors in PLS, but in the ALS case there is no prediction error (*e.g.* RMSEP) to minimise. In the present case the obvious choice would have been three components, as this gave the most stable result. Only the comparison of the profiles to actual lactose results indicated that four components were optimal. Methods for determining the number of independently varying species present in the samples are therefore required.

Scores obtained through Principal Component Analysis (PCA) (10) could solve the problem.

In the present case, when performing a PCA on all calibration samples, the scores (not shown) contain structure (originating from the batch structure of the data) revealing up to four or five components. Thus, four or five components would be expected to be optimal in ALS, which supports the actual findings shown above.

In this case it was possible to extract concentrations from FT-IR data from a lactose hydrolysis process and to monitor new process runs using this knowledge. The present data set must be considered to be worst case, as the infrared spectra of reactants, intermediates and products are very similar, *i.e.* they are all sugars which give roughly the same absorption peaks. Resolving spectra from processes where the compounds are much less similar should therefore be easier, especially the problem of determining the number of components should be less difficult.

CONCLUSIONS

It appears from the foregoing that ALS is a promising method for use in dairy process optimisation. Without the need for reference analyses it is possible to extract *e.g.* four components from lactose hydrolysis process data (fat, lactose and two other sugar components) and to obtain a lactose prediction error similar to the one obtained from an ordinary PLS regression. Such use of ALS for reference-independent prediction of process parameters is not limited to dairy products only, but is likely to be useful for process monitoring and identification of intermediates in all branches of the food and beverage industry.

By use of ALS combined with FT-IR it becomes possible to obtain quick information on compounds present during the process, but not necessarily by the end of it. A further advantage (in many cases the most important) is that the pure spectra obtained by ALS makes it possible to generate predictions of process parameters without the need for expensive and time consuming reference procedures.

PATENT CLAIMS

1. Method for evaluation of spectroscopic data, comprising the use of curve resolution, such as multivariate curve resolution, such as Alternating Least Squares (ALS), and/or Alternating Regression, (AR), *characterised by* using the method for determination of the properties of and/or constituents in food and/or beverages, such as dairy products.

2. Method according to claim 1, *characterised by* using the method for monitoring a process involving food and/or beverages, such as dairy processes, and especially for on-line monitoring of the process.

3. Method according to claim 1, *characterised by* generating calibrations by use of curve resolution for determining the properties of and/or constituents in food and/or beverages, such as dairy products and/or monitoring a process involving food and/or beverages, such as a dairy process, and especially for on-line monitoring of the process.

4. Method for detection of constituents in fluids involved in dairy processes, in which method at least two samples are extracted from a process line and wherein the samples are transferred into apparatus comprising optical equipment allowing the registration of data identifying a (*e.g.* absorbance or reflectance) spectrum of the samples, *characterised in* that the data are processed using alternating regression, AR, or alternating least squares, ALS.

5. Method according to claim 1 or 4, *characterised by* using at least one of the following constraints:

   - Non-negativity of at least one of the concentration profiles,

- Non-negativity of at least one of the spectra, and

- Unimodality of at least one of the concentration profiles in each process run.

6. Method according to claim 1 or 4, *characterised by* starting the calculations by setting the initial concentration of at least one component such as lactose to 1 while all other concentrations are set to 0.

7. Method according to claim 1 or 4, *characterised by* using the SSE (Sum of Squared Errors) for selecting the optimal number of components.

8. Method according to claim 1 or 4, *characterised by* using three, four or five components.

9. Use of ALS or parameters generated by it for monitoring dairy processes.

10. Use of ALS or parameters generated by it for optimising dairy processes.

11. Method for determining the content of lactose in a milk sample, *characterised by* using ALS.

12. On-line monitoring of a dairy process, *characterised by* using multivariate curve resolution, such as alternating least squares, ALS.

13. On-line monitoring of a dairy process, *characterised by* using calibrations or parameters generated by use of curve resolution, or multivariate curve resolution, such as alternating least squares, ALS.

14. Method according to claim 1, *characterised by* using the method for determination of the properties of and/or constituents in dairy products and/or monitoring a dairy process.

FOSS ELECTRIC
69, Slangerupgade
DK 3400 Hillerød, Denmark
Fax: +45 70 10 33 71
Phone: +45 70 10 33 70
e-mail: info@foss-electric.dk

ATV
The Danish Academy of Technical Sciences
Lundtoftevej 266
DK-2800 Lyngby, Denmark

KVL
The Royal Veterinary and Agricultural University
Department of Dairy and Food Science
Rolighedsvej 30
DK-1958 Frederiksberg C, Denmark

FOSS
FIRST IN FOOD ANALYSIS