Norwegian University
of Life Sciences

Master's Thesis 2017 30 ECTS
Faculty of Chemistry, Biotechnology and Food Science

Fourier transform infrared spectroscopy (FTIR) for detection and quantification of milk components for cattle health monitoring

Felicia Judith Marie Stehr
Chemistry and biotechnology, organic analysis

Acknowledgements

This project was conducted at Nofima (Ås, Norway) in the period from September 2016 to December 2017. It was a part of the larger project "New approaches for management and breeding of dairy cows in automatic milking systems" (the AMS project). I am very thankful for this interesting period and for meeting all the nice people working at Nofima.

I am very grateful to my supervisor at Nofima, *Nils Kristian Afseth,* for guiding me through this project. This friendly, skilled and busy man helped me structure the thesis, explained the data analysis part and provided the data for the calibration models.

I would like to express my gratitude to my supervisor at NMBU, *Elling-Olav Rukke,* for excellent guidance and support during the work with this Master's thesis. He provided valuable advices regarding both writing skills and the content of the thesis.

I was very lucky to get one extra supervisor, *Amira Rachah,* during the work with the thesis. I am grateful to her for all valuable suggestions and inputs during the writing period and for working with me on the data using MATLAB.

I also would like to thank the fabulous *Bjørg Narum* for showing me how to do the milk analyses and for always being very helpful and friendly.

Thanks to
the *staff at SHF* for providing the milk samples,
TINE (Heimdal) for information about the chemical composition of the samples,
ProdMed for providing data for betahydroxybuturate content,
Vitas (Oslo) for analysing the reference samples,
Olav Reksen for showing me the AMS stable and for leading the AMS project,
Silje Aaser Basmoen for helping me with the analyses,
and *friends and family* for encouragement during my time as a student.

Finally, I would like to thank my dear *Terje Kvambe* for always supporting and motivating me. Thanks for showing me smart functions in Excel, for general IT support during my time as a student and for proofreading the thesis.

Ås, 2017-12-14

Felicia J. M. Stehr

Sammendrag

Antallet automatiske melkesystemer (AMS-er) har økt jevnt de siste årene på grunn av større besetninger og høyere krav til effektivitet. Med et AMS kan kua velge når og hvor ofte hun vil melkes. Ulike målinger av kua og melken kan utføres ved hver melking. Bruk av teknologier for måling av helseparametere for individuelle kyr kan sikre tidlig deteksjon av sykdommer og maksimere potensialet til den individuelle kua og flokken. I dette prosjektet ble tørrfilm-FTIR (Fourier transform infrarød spektroskopi) evaluert for det formål å detektere og kvantifisere melkekomponenter for overvåking av storfehelse.

Melkeprøver fra 64 NRF-kyr (Norsk rødt fe) ble tatt ut i løpet av deres første 100 dager i melk. Totalt antall prøver brukt i analysene var 2333. De resulterende FTIR-spektraene ble brukt til eksplorative analyser av melkekomposisjonen. Prinsipalkomponentanalyse (PCA) ble brukt til å søke etter systematiske endringer i melken under laktasjonen. De observerte trendene i FTIR-spektraene var i samsvar med trendene for den kjemiske sammensetningen levert av TINE. Dette betyr at spektraene inneholdt logisk informasjon om melkeprøvene.

FTIR-spektraene ble også brukt til prediksjon av proporsjonene av fettsyrene C10:0, C14:0, C16:0, C18:0, C18:1*cis*-9, konjugert linolsyre, summen av mettede fettsyrer, summen av enumettede fettsyrer og summen av flerumettede fettsyrer. Tidligere etablerte delvis minste kvadraters regresjonsmodeller (PLSR), med referanseprøver fra det nåværende datasettet inkludert, ble brukt til prediksjonene. Moderat gode til gode kalibreringsmodeller (med determinasjonskoeffisient, R^2, fra 0,76 til 0,97) ble oppnådd for fettsyrene. Gjennomsnittsverdiene av de predikerte proporsjonene var i samsvar med hva som ble funnet i litteraturen. Deretter ble den predikerte fettsyresammensetningen analysert for å finne systematiske endringer under laktasjonen. De observerte trendene for gjennomsnittlige fettsyreinnhold var også i samsvar med litteraturen. Disse trendene var imidlertid vage for den individuelle ku på grunn av store daglige variasjoner.

Denne studien viser at tørrfilm-FTIR er en potensiell metode for deteksjon og kvantifikasjon av melkekomponenter for overvåking av storfehelse. Imidlertid er det nødvendig med videre forskning og utvikling før denne teknologien kan inkorporeres i AMS-er for overvåking av individuelle kyr. Prosedyrene for prøvetaking og analyse bør forbedres og praktiske utfordringer vedrørende implementering av FTIR-instrumentet i et AMS må undersøkes. Utvikling av kalibreringsmodeller for prediksjon av energibalanse vil også være viktig.

Abstract

The number of automatic milking systems (AMSs) has increased steadily over the past years due to larger herd sizes and higher efficiency demands. With an AMS, a cow can choose when and how often she wants to be milked. Various measurements of the cow and the milk may be performed at each milking. The use of technologies for measurements of health parameters of individual cows may ensure early detection of diseases and maximization of individual cow and herd potential. In the present project, dry-film Fourier transform infrared spectroscopy (FTIR) was evaluated for the purpose of detecting and quantifying milk components for cattle health monitoring.

Milk samples from 64 Norwegian Red cows were collected during their first 100 days in milk. The total number of samples used in the analyses was 2333. The resulting FTIR spectra were used for explorative analyses of the milk composition. Principal component analysis (PCA) was used to search for systematic changes in the milk during the lactation. The trends observed in the FTIR spectra were consistent with the trends for the chemical composition provided by TINE. This means that the spectra contained logical information about the milk samples.

The FTIR spectra were also used for prediction of the proportions of the fatty acids C10:0, C14:0, C16:0, C18:0, C18:1*cis*-9, conjugated linoleic acid, the sum of saturated fatty acids, the sum of monounsaturated fatty acids and the sum of polyunsaturated fatty acids. Previously established partial least squares regression (PLSR) models, with reference samples from the current data set included, were used for the predictions. Moderately good to good calibration models (with coefficients of determination, R^2, ranging from 0.76 to 0.97) were obtained for the fatty acids. The mean values of the predicted proportions were in agreement with what was found in the literature. Subsequently, the predicted fatty acid composition was analysed in order to find systematic changes during the lactation. The observed trends for the average fatty acid contents were also consistent with the literature. However, these trends were vague on the individual cow level due to large day-to-day variations.

This study shows that dry-film FTIR is a potential method for detection and quantification of milk components for cattle health monitoring. However, before incorporating this technology in AMSs for monitoring of individual cows, further research and development is needed. The procedures for sampling and analysis should be improved and practical challenges regarding implementation of the FTIR instrument in AMSs must be investigated. Development of calibration models for prediction of energy balance from the spectra will also be important.

List of abbreviations

AMS	Automatic milking system
DIM	Days in milk
EB	Energy balance
EMSC	Extended multiplicative signal correction
FA	Fatty acid
– CLA	Conjugated linoleic acid
– FFA	Free fatty acid
– MUFA	Monounsaturated fatty acid
– PUFA	Polyunsaturated fatty acid
– SFA	Saturated fatty acid
FTIR	Fourier transform infrared spectroscopy
GC	Gas chromatography
OLA	Onset of luteal activity
PCA	Principal component analysis
– PC1	The first principal component
– PC2	The second principal component
PCR	Principal component regression
PLSR	Partial least squares regression
– R^2	Coefficient of determination
– RMSECV	Root mean square error of cross-validation
ProdMed	The Department of Production Animal Clinical Sciences
SD	Standard deviation
SHF	The Animal Production Experimental Centre

Table of contents

1. Introduction

1.1. Background and objectives

In Norway, the first automatic milking system (AMS) was installed in year 2000 and since then the number of AMSs has increased steadily due to larger herd sizes and higher efficiency demands (Ottinsen 2013). This is also true throughout the world, and in 2015 more than 25 000 farmers worldwide used AMS (Barkema et al. 2015). With an AMS, a cow can choose when, and how often, she wants to be milked. The AMS can perform different measurements of the cow and the milk at each milking (Ottinsen 2013).

Over the years, milk farmers have used experience and judgement to detect deviating conditions in cows. Cow diseases often start with changes in physiological parameters like heart rate, body temperature or milk composition, which are not easy to detect by the human eye. Thus, by use of real-time data for monitoring of these parameters, and identification of deviations, the farmer may intervene sooner than with examination by humans only. The use of technologies for measurements of health parameters of individual cows for improved farm performance is called precision dairy farming. In addition to early detection of disease, precision dairy farming may be used for maximization of individual cow potential. The use of AMS offers new possibilities for management of the individual cow and the herd (Bewley 2009). However, sensor technologies that provide chemical information of the milk and relate this information to the health status of a cow are still lacking (Neethirajan 2017). In the present project, dry-film Fourier transform infrared spectroscopy (FTIR) was evaluated for this purpose.

The main objective of this work was to analyse milk from 64 cows during their first 100 days in milk (DIM) using dry-film FTIR, and to investigate if the resulting FTIR spectra were suitable for detection and quantification of milk components for cattle health monitoring. The objective was divided into two sub-objectives:

1. To investigate if the FTIR spectra could be used to detect any systematic change in milk composition during the first 100 DIM. Principal component analysis (PCA) was used for the explorative analysis of the data.

2. To predict the fatty acid (FA) composition of all milk samples using existing partial least squares regression (PLSR) models and search for any systematic change in the

predicted FA composition during the first 100 DIM. One part of this objective was to investigate if there were any differences in the predicted FA composition between cows with subclinical ketosis and normal cows.

1.2. Dairy cows and cow milk

The most common cow breed for milk and meat production in Norway is the Norwegian Red cattle (Geno 2016) (Figure 1). A Norwegian Red cow produces on average about 7700 kg milk per year. The cow's lactation period (the period in which the cow produces milk) starts on the day of calving and lasts for about 310 days (Geno 2017). DIM is defined as the number of days from calving and is equivalent to days of lactation. Cow milk contains approximately 87 % water, 4.6 % lactose, 4.2 % fat, 3.4 % protein and 0.7 % minerals (Ottinsen 2013). The lactose content has been reported to be highest between 30 and 60 DIM (Ptak et al. 2012). It varies the least of the chemical components in the milk. Fat content varies the most of the component in milk. It varies with diet and lactation stage. It is often highest in the start of the lactation and decreases during the first two months. Then it slowly increases. The fat content is also affected by milking interval (Linn 1988). The concentration of protein is highest immediately after calving and decreases during the 30 first days of lactation (Linn 1988; Tsioulpas et al. 2007). The protein content is affected by milking interval and diet in a much lesser extent than what the fat content is (Linn 1988). The concentration of free fatty acids (FFAs) in milk may vary from below 0.9 mmol/L (milk of highest quality) to more than 1.7 mmol/L (milk of bad quality) (TINE SA 2017). The urea content is normally between 3.0 mmol/L and 6.0 mmol/L and is lowest in the beginning of the lactation (Geno 2014c).

Figure 1. A Norwegian Red cow in the cowshed (A) and in a pasture (B). Photo: Felicia Stehr.

The Norwegian Red cattle are known for good fertility and health, and they calve easily (Geno 2017). The cows are usually inseminated for the first time when they are from 14 to 16 months. The gestation length is approximately 280 days. Since the Norwegian dairy farmers often want their cows to give birth to one calf per year, the cows are inseminated again a few months after calving (Geno 2014a). Thus, the cows only have a few important weeks for recovery between the end of the first lactation period and the start of the next (Geno 2017). Since it is common to have a voluntary waiting period of about 45−60 days before inseminating the cow again, the cow must get pregnant within a relatively short period (about 35 days) to obtain the goal of one calf per year (Martin 2015). Figure 2 provides an overview of the different periods during the first year after calving. The period of milk sampling in this project is also marked in the figure.

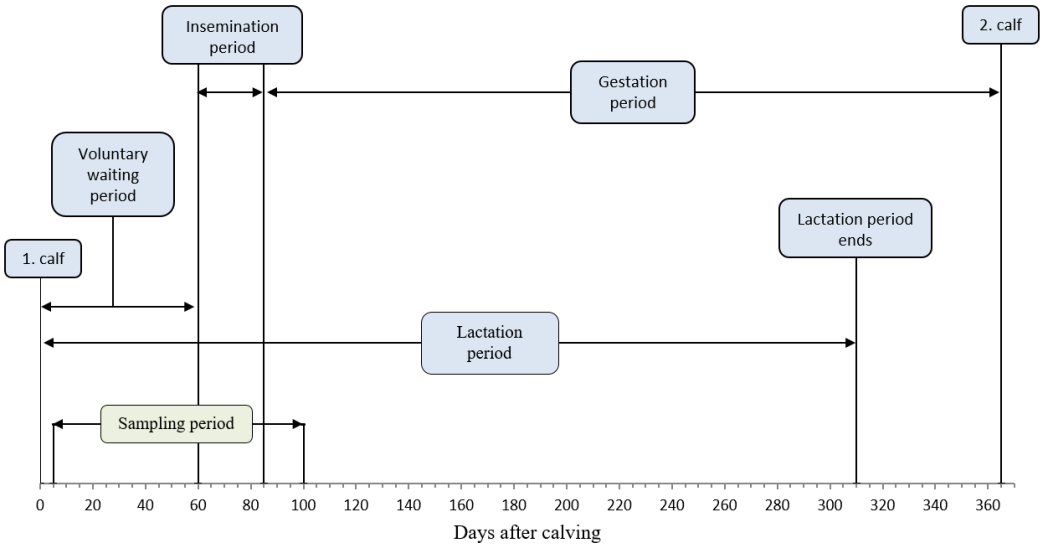


Figure 2. The different periods during the first year after calving. The sampling period in this project was from day 5 to day 100 in the lactation period.

A healthy cow is experiencing oestrus regularly every third week (Walsh et al. 2011). The oestrus typically last between 12 and 18 hours (Martin 2015). Oestrus durations between 0.2 and 36 hours have been observed. To get the cow pregnant, she must be inseminated before ovulation (which occurs in the late oestrus) (Martin 2015).

1.2.1. Cow health

Although the Norwegian Red cattle are known for good health and fertility, some fertility challenges associated with dairy cows exist. One of the most common problems is difficulty in getting the cow pregnant, and this is often due to problems related to the oestrus. Total absence of oestrus is often not the case. Instead, the problems usually are poor ability to show oestrus signs and/or a shorter period of oestrus than normal. The problems with weak oestrus signs and shorter oestrus periods are frequent in the first months after calving (Geno 2014b), and the cow often does not show oestrus signs at all during the first oestrus after calving (Martin 2015).

About four days after ovulation there will be a rise in progesterone levels (from the progesterone producing structure called corpus luteum). Onset of luteal activity (OLA) is defined as the first significant rise in progesterone after calving. The time from calving to OLA is known to be related to the reproductive performance in cows and it is a good measure of fertility. There will in general be an increased chance of conception in cows experiencing early OLA. These cows also have a better ability to show oestrus signs than cows experiencing late OLA. This is due to earlier resumption of ovarian cyclicity, which leads to increased number of ovulations before insemination. Very early OLA, on the other hand, may be associated with irregular ovarian cyclicity (Martin 2015). For cows experiencing late OLA it has been found that a waiting period of at least 70 days (instead of 45−60 days) is optimal (Inchaisri et al. 2011). Extending the waiting period for the cows that are predicted to experience late OLA could therefore result in reduced insemination cost and increased production efficiency (Martin 2015).

A cow requires a lot of energy in the beginning of the lactation due to the large milk production (which is at its maximum between four and eight weeks after calving). It is not possible for the cow to replace the lost energy just by eating, so she will experience a negative energy balance (EB). A negative EB could also occur later in the lactation due to insufficient amount, or bad quality, of the food (Gross et al. 2011). Negative EB may lead to ketosis, which is a disease that causes reduced appetite, reduced milk yield and weight loss. Subclinical ketosis is when the cow does not have visible symptoms of ketosis (TINE Rådgiving 2008). During negative EB, the cow will mobilize body fat (Walsh et al. 2011), and body condition score (BCS) loss will occur (Martin 2015). BCS is a commonly used measure for monitoring of a cow's body condition. Both a too low (1.5−2.5) and a too high

(>3.5) BCS at calving, and considerable BCS loss in the period after calving, will have a negative impact on cow fertility (Walsh et al. 2011).

Severe negative EB has a negative impact on the cow's reproductive performance (Walsh et al. 2011), and it is known that it will lead to delayed OLA (Martin 2015; Walsh et al. 2011). Also previously performed studies (Butler et al. 1981; de Vries & Veerkamp 2000) have found that there is a correlation between EB and number of days to OLA. Martin et al. (2015) found that cows that were experiencing early OLA on average entered a state of positive EB after calving three weeks earlier than the cows experiencing late OLA. They also found that the EB was more negative in the beginning of the lactation for the cows experiencing late OLA than for the cows experiencing early OLA. The difference in EB between cows experiencing early and late OLA was largest three weeks after calving.

For improved fertility, it is important to minimize the extent of negative EB and BCS loss early in the lactation. This could be done by optimal feeding of the cows before and after calving (Walsh et al. 2011). Early detection of delayed OLA, and other problems, makes it easier to give the individual cow optimal treatment and dietary supplements. For daily monitoring of EB and prediction of time to OLA, it is important to have good methods for prediction of these parameters. For this purpose, measurements of components in the milk may be used.

1.2.2. Cow milk as a diagnostic tool

Milk analyses may be used as a diagnostic tool for monitoring of a cow's health. As an example, detection of cows with mastitis (inflammation of the udder) is done by measurement of the cell count in milk. More than 100 000 cells/mL may indicate subclinical mastitis, while milk from cows with clinical mastitis might contain more than one million cells per mL (Ottinsen 2013). The urea content is another important parameter to monitor because a too low or a too high concentration is associated with reduced fertility (Geno 2014c).

Løvendahl et al. (2010) found that prediction of EB based on milk fat content, fat:protein ratio and milk yield was not reliable enough for daily cow management and decision support. They suggested that for that purpose one would need variables in milk that are more correlated to the EB. Different FAs in milk have different origins (diet, body fat, mammary gland or the rumen) (Martin et al. 2015). It is known that the FA composition in milk is affected by variables such as season, herds, feeding regime and stage in lactation (Gross et al. 2011; Palmquist et al. 1993). The negative EB immediately after calving results

in body fat mobilization, which results in a larger amount of long-chained FAs in the milk. The large amount of long-chained FAs in the milk inhibits the synthesis of short-chained FAs (Palmquist et al. 1993). Therefore, there will in general be a higher proportion of long-chained FAs early in the lactation than late in the lactation. The opposite is true for short-chained FAs (except for C4:0) (Gross et al. 2011; Palmquist et al. 1993). A negative EB will therefore have an influence on the FA composition in the milk and thus, the FA composition may be used for monitoring of a cow's EB (Palmquist et al. 1993). Similar conclusions were drawn by Stoop et al. (2009) who discovered that FA composition was correlated both to EB and stage in lactation. Results of Gross et al. (2011) showed that particularly changes in 18:1*cis*-9 and in groups of FAs would be of interest for monitoring of EB. Van Haelst et al. (2008) discovered that cows with subclinical ketosis had a higher proportion of 18:1*cis*-9 in the milk, compared to normal cows, already two weeks before diagnosis. Martin et al. (2015) identified relationships between the proportion of some FAs and OLA already during the first week after calving. The prediction of early and late OLA was best in the fourth week after calving. A significant relationship between FA and OLA was observed for C14:0, C16:0 and C18:1*cis*-9. A significant correlation between protein concentration and OLA, fat concentration and OLA, or fat:protein ratio and OLA was not detected. In contrast to detection of OLA by progesterone content, prediction of OLA by FA composition gives the farmer a possibility to optimize the treatment of individual cows before OLA occurs (Martin et al. 2015).

1.2.3. Fatty acids in cow milk

FAs are hydrocarbon chains with a carboxyl group in one end and a methyl group in the other end. An FA without double bonds is saturated with hydrogen atoms and is therefore called a saturated fatty acid (SFA). A monounsaturated fatty acid (MUFA) contains one double bond whilst a polyunsaturated fatty acid (PUFA) has more than one double bond (Rustan & Drevon 2005). The relative concentrations of SFAs, MUFAs and PUFAs in milk are approximately 70 %, 25 % and 5 %, respectively (Soyeurt et al. 2006).

The double bonds can be placed in various positions along the carbon chain and they can be of *cis* or *trans* configuration. A systematic notation for FAs includes the number of carbon atoms, the number of double bonds, and the configuration and position of the double bonds. The position of the double bond may either be indicated as the number of carbon atoms from the methyl end or from the carboxylic end. When counting from the methyl end,

the letter n or ω is included (Rustan & Drevon 2005). The notation for the FA in Figure 3 will be C18:1*cis*-9 in this thesis.

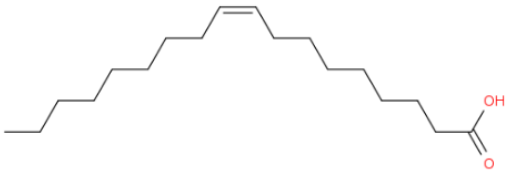


Figure 3. Structure of the fatty acid with notation C18:1cis-9 (PlantFAdb 2016a). This is the most common unsaturated fatty acid in cow milk.

About 98 % of the fat in milk is present as FAs bound in triglycerides (Soyeurt et al. 2006). The formation of triglycerides is illustrated in Figure 4. Other types of lipids present in small amounts in milk are FFAs, cholesterol and phospholipids (Christie 1995).

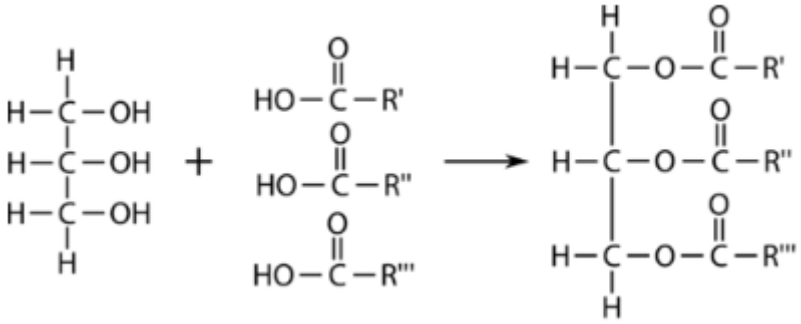


Figure 4. Esterification of three fatty acids (middle) with a glycerol (left) results in a triglyceride (right) (LibreTexts 2016).

In the mammary cells, the triglycerides are gathered in droplets enclosed by a membrane consisting of protein and polar lipids (Keenan & Dylewski 1995; Soyeurt et al. 2006). In the cytoplasm, these small droplets may fuse together and form larger droplets. During the secretion process of these droplets to the milk, they become enveloped by the cell membrane such that when they are detached they are completely covered with cell membrane (Keenan & Dylewski 1995). The membrane that now covers the fat globule (three layers), is called the milk fat globule membrane (MFGM) (Christie 1995). When the droplets have been secreted to the milk from the mammary cells, they are called milk fat globules (Keenan & Dylewski 1995). The diameter of a fat globules is on average 2−4 µm (Christie 1995). Because milk contains these fat globules, milk is an emulsion of fat and water (Soyeurt et al. 2006).

The FAs in milk are derived either from lipids in the blood plasma or from synthesis in the mammary gland. The lipids in the blood plasma originate from the diet or from lipolysis of body fat. Synthesis of FAs with 16 or fewer carbon atoms occurs in the mammary gland (Christie 1995). The synthesis is catalysed by the enzyme fatty acid synthase (FAS). The mammary gland is not capable of synthesising FAs with 18 or more carbon atoms, so C18 and longer FAs come from mobilized body fat or from the diet. However, desaturation of C18:0 to C18:1 occurs in the mammary gland (Hawke & Taylor 1995).

Over 400 different FAs have been detected in cow milk. However, only about 15 of these are present in a concentration of 1 % or higher (Månsson 2008). The composition of the abundant FAs in milk from Norwegian Red cattle is presented in Table 1. These data were obtained in a study performed by Afseth et al. (2010).

Table 1. Fatty acid composition in milk from Norwegian Red cattle expressed as percentage of total weight of fatty acids (mean) with the corresponding standard deviations (SD). The composition was determined by gas chromatography in the study by Afseth et al. (2010). The names are from Dewick (2009) and the database PlantFAdb (2016b).

Notation	Systematic name	Common name	Mean	SD
C4:0	Butanoic acid	Butyric acid	4.4	0.4
C6:0	Hexanoic acid	Caproic acid	2.6	0.2
C8:0	Octanoic acid	Caprylic acid	1.6	0.2
C10:0	Decanoic acid	Capric acid	3.4	0.6
C12:0	Dodecanoic acid	Lauric acid	3.8	0.7
C14:0	Tetradecanoic acid	Myristic acid	12.1	1.2
C14:1*cis*-9	*cis*-9-Tetradecenoic acid	Myristoleic acid	0.9	0.3
C15:0	Pentadecanoic acid	Pentadecylic acid	1.2	0.2
C16:0	Hexadecanoic acid	Palmitic acid	27.3	2.8
C16:1*cis*-9	*cis*-9-Hexadecenoic acid	Palmitoleic acid	1.1	0.3
C17:0	Heptadecanoic acid	Margaric acid	0.5	0.1
C18:0	Octadecanoic acid	Stearic acid	10.7	1.5
C18:1*trans*-9	*trans*-9-Octadecenoic acid	Elaidic acid	0.2	0.02
C18:1*trans*-10	*trans*-10-Octadecenoic acid	Isooleic acid	0.3	0.06
C18:1*trans*-11	*trans*-11-Octadecenoic acid	*trans*-Vaccenic acid	1.4	0.5
C18:1*cis*-9	*cis*-9-Octadecenoic acid	Oleic acid	18.2	3.1
C18:1*cis*-11	*cis*-11-Octadecenoic acid	*cis*-Vaccenic acid	0.7	0.1
C18:2*cis*-9, *cis*-12	*cis*-9, *cis*-12-Octadecadienoic acid	Linoleic acid	1.1	0.3
C18:2*cis*-9, *trans*-11	*cis*-9, *trans*-11-Octadecadienoic acid	Conjugated linoleic acid	0.6	0.2
C18:3*cis*-9, *cis*-12, *cis*-15	*cis*-9, *cis*-12, *cis*-15-Octadecatrienoic acid	α-linolenic acid	0.4	0.1
C20:0	Eicosanoic acid	Arachidic acid	0.2	0.03
C20:4*cis*-5, *cis*-8, *cis*-11, *cis*-14	*cis*-5, *cis*-8, *cis*-11, *cis*-14-Eicosatetraenoic acid	Arachidonic acid	0.07	0.02
C20:5*cis*-5, *cis*-8, *cis*-11, *cis*-14, *cis*-17	*cis*-5, *cis*-8, *cis*-11, *cis*-14, *cis*-17-Eicosapentaenoic acid	Timnodonic acid	0.06	0.01
C22:6*cis*-4, *cis*-7, *cis*-10, *cis*-13, *cis*-16, *cis*-19	*cis*-4, *cis*-7, *cis*-10, *cis*-13, *cis*-16, *cis*-19-Docosahexaenoic acid	Cervonic acid	0.04	0.01
Summed saturated fatty acids	-	-	67.8	3.7
Summed monounsaturated fatty acids	-	-	22.8	3.5
Summed polyunsaturated fatty acids	-	-	2.3	0.3
Iodine value	-	-	21.9	3.1
Summed C18:1 *trans* isomers	-	-	1.9	0.5

As presented in Table 1, the four major FAs in milk are C14:0, C16:0, C18:0 and C18:1*cis*-9. There is a high proportion of SFAs and a very low proportion of PUFAs and *trans* isomers. The concentration of PUFAs in milk is relatively low due to hydrogenation of

unsaturated FAs from the diet by microorganisms in the rumen (Christie 1995). FAs with 20 and 22 carbon atoms are present in very low concentrations. The FAs present at a concentration of 1 % or higher in Table 1 are presented in Figure 5.

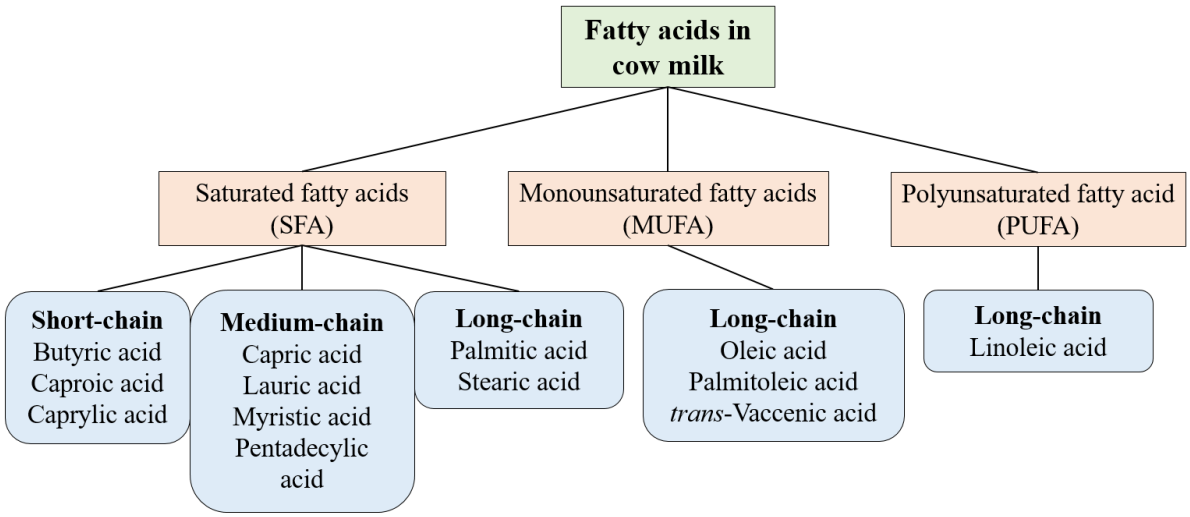


Figure 5. Fatty acids present in a concentration of 1 % (of total weight of fatty acids) or higher in cow milk (from Table 1).

Since there is a difference in carbon chain length between the FAs derived from the blood and from the mammary gland, the composition of FAs will be affected by factors such as diet and lactation stage. During the lactation, the amount of C4:0, C18:0 and C18:1 will decline, while there will be an increase of C16:0. The amount of C6:0 to C14:0 will first increase, and approximately 8 to 12 weeks after calving the amount will slowly decrease (Hawke & Taylor 1995).

Hydrolytic degradation of the triglycerides results in an increased concentration of FFAs. FFAs can give a bad taste to milk and milk products. This process is called hydrolytic rancidity. Another word for hydrolysis of milk fat is lipolysis. The lipolysis is catalysed by enzymes called lipases. There are two types of lipases involved in hydrolysis in milk: enzymes that originate from the mammary gland and enzymes from microorganisms. The lipoprotein lipase (LPL) is synthesised in the mammary gland and is used for degradation of triglycerides in the blood. This enzyme is inactivated by heat treatment. Lipolysis can be induced or spontaneous. Induced lipolysis may occur from vigorous stirring due to breakage of the MFGM, which leads to more exposed triglycerides. Lipolysis may also be induced by temperature changes, and freezing and thawing of the milk. Lipolysis that occurs in stored milk that has been cooled soon after milking is called spontaneous lipolysis. Milk in which

spontaneous lipolysis occur is called spontaneous milk. Cows in late lactation are more likely to produce spontaneous milk than cows in early lactation. Feeding regime has also been found to influence the probability of spontaneous milk. Factors that can increase the probability of hydrolytic rancidity are cows in late lactation, poor feeding conditions and agitation of raw milk (especially with air). Also, storage for long periods, storage at too high temperatures and poor hygiene are risk factors, since hydrolytic rancidity may be caused by microbial contamination (Deeth & Fitz-Gerald 1995).

1.3. Fourier transform infrared spectroscopy

1.3.1. Principle

Spectroscopy is the study of the interaction between electromagnetic radiation and matter (Helmenstine 2017). A spectrometer is used for measurements of incoming light and may for example be used for temperature determination of objects in space. A spectrophotometer contains a radiation source and can measure the difference in intensity of radiation, at different wavelengths, before and after light has passed through a sample. Its application areas include investigation of a sample's absorption properties (Finlay 2017).

Light (radiation) can be thought of both as traveling particles (photons) and as waves. The waves consist of magnetic and electric fields, which oscillate perpendicular to each other. A wave can be described by its wavelength and frequency. The wavelength is the distance between two crests (one complete oscillation), see Figure 6A. Light of different wavelengths can be categorized into regions of the electromagnetic spectrum (Figure 6B) (Harris 2010).

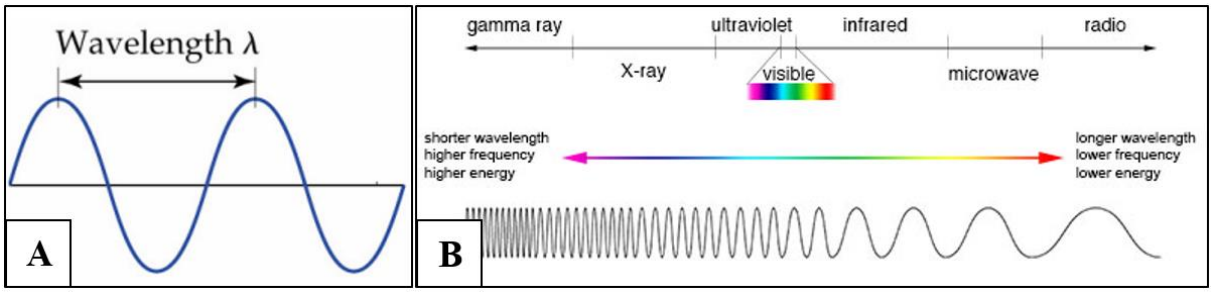


Figure 6. A wave with wavelength λ (A) (SCENIHR 2008) and the electromagnetic spectrum (B) (NASA 2013).

The frequency is defined as the number of finished oscillations per second. One hertz (Hz) is defined as one oscillation per second. The relationship between wavelength and frequency is given in Equation 1 (Harris 2010).

$$\nu = \frac{c}{\lambda} \qquad (1)$$

where ν is the frequency, c is the speed of light (2.998×10^8 m/s in vacuum), and λ is the wavelength. As Equation 1 states, the frequency increases with decreased wavelength. The number of oscillations per a defined distance is called the wavenumber and is often given in cm^{-1} (number of oscillations per cm) (Harris 2010). The relationship between wavenumber and wavelength is given in Equation 2.

$$\tilde{\nu} = \frac{1}{\lambda} \quad (2)$$

where $\tilde{\nu}$ is the wavenumber. The energy of a photon can be described by Equation 3 (Harris 2010).

$$E = h\nu \quad (3)$$

where h is Planck's constant (6.626×10^{-34} J·s). By combining Equation 1, 2 and 3, Equation 4 is obtained (Harris 2010).

$$E = h\frac{c}{\lambda} = hc\tilde{\nu} \quad (4)$$

Equation 3 and 4 illustrates that the energy of a wave will increase with decreased wavelength, increased frequency or increased wavenumber. Since ultraviolet light has shorter wavelength than infrared light, ultraviolet light has higher energy than infrared light (Harris 2010).

When radiation strikes a molecule, the molecule can absorb a photon and thereby get excited from the ground energy state to a higher energy level (Harris 2010). The absorbance (also called optical density) can be expressed by Equation 5.

$$A = log\left(\frac{I_0}{I}\right) \quad (5)$$

where I_0 is the intensity of the radiation that strikes the sample, and I is the transmitted light (the remaining part of the light that leaves the sample) (Williams & Fleming 2008). According to Equation 5, absorption of 90 % of the light will give an absorbance of 1. Absorption of 99 % will give an absorbance of 2. Regarding the precision of the spectrophotometer, it is preferable to achieve an absorbance approximately between 0.3 and 2. According to Beer's law (Equation 6), the absorbance is proportional to the concentration of the absorbing molecules in the sample (Harris 2010).

$$A = \varepsilon lc \quad (6)$$

where ε is the molar absorptivity (also known as the molar extinction coefficient) [$M^{-1}cm^{-1}$], l is the path length of the sample [cm] and c is the concentration [M]. The absorbance and the molar absorptivity are dependent on the wavelength. An absorbance spectrum can be constructed by plotting absorbance against wavelength (Harris 2010).

Different types of radiation affect the molecules in different ways. For example, when visible and ultraviolet radiations are absorbed, electrons will jump to a higher energy orbital, while absorption of microwave radiation will stimulate rotation of the molecule. Infrared light, which is used in this project, will stimulate vibrations in the molecule. The atoms naturally vibrate in different directions in the ground energy state of a molecule. The most interesting vibrations are symmetric and asymmetric stretching, and different types of bending movements (Harris 2010). These vibrations are presented in Figure 7.

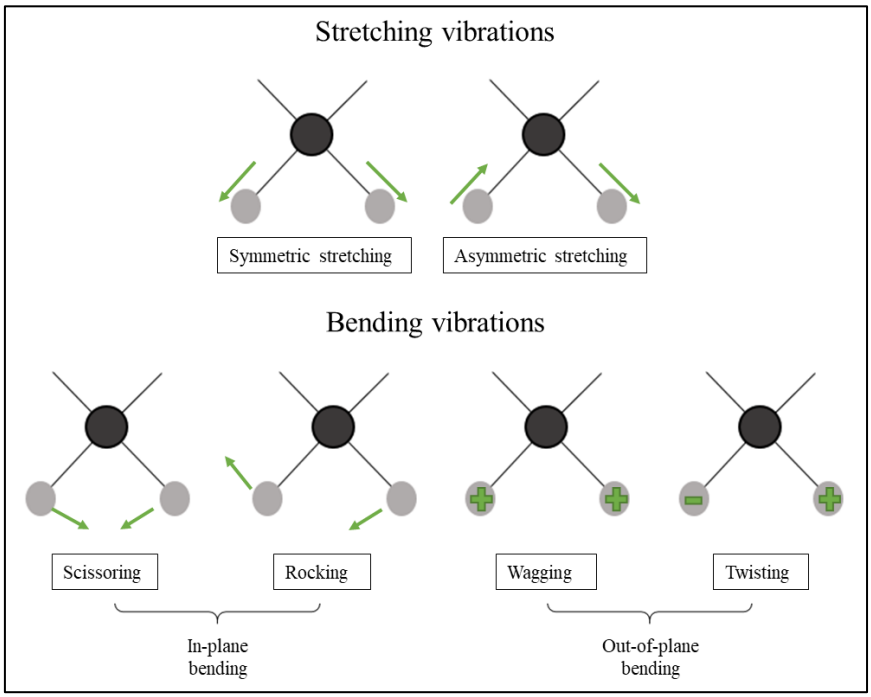


Figure 7. Different types of molecular vibrations. The plus sign means motion towards the reader and the minus sign means motion away from the reader.

The amplitude of the oscillations, and the energy of the molecule, will increase when infrared radiation is absorbed. When the molecule goes back to its ground state, it converts the energy from the photon to heat. The wavelength of the absorbed light depends on the type of vibration and the atoms involved. In other words, different amounts of energy are required to increase the amplitude of the oscillation for different types of chemical bonds and vibrations (Harris 2010). Functional groups will therefore absorb infrared light at characteristic frequency ranges. Due to this, it is possible to identify the functional groups in a molecule by use of infrared spectrophotometry. The most useful infrared region is from 4000 cm^{-1} to 625 cm^{-1} since many functional groups vibrate with frequencies within this range (Williams & Fleming 2008).

Modern infrared spectrophotometers analyse light of all chosen wavelengths at the same time. This is done by use of Fourier transformation, and this spectroscopy method is thus called FTIR. The key component of an FTIR instrument is the interferometer. The basic construction of an interferometer is illustrated in Figure 8.

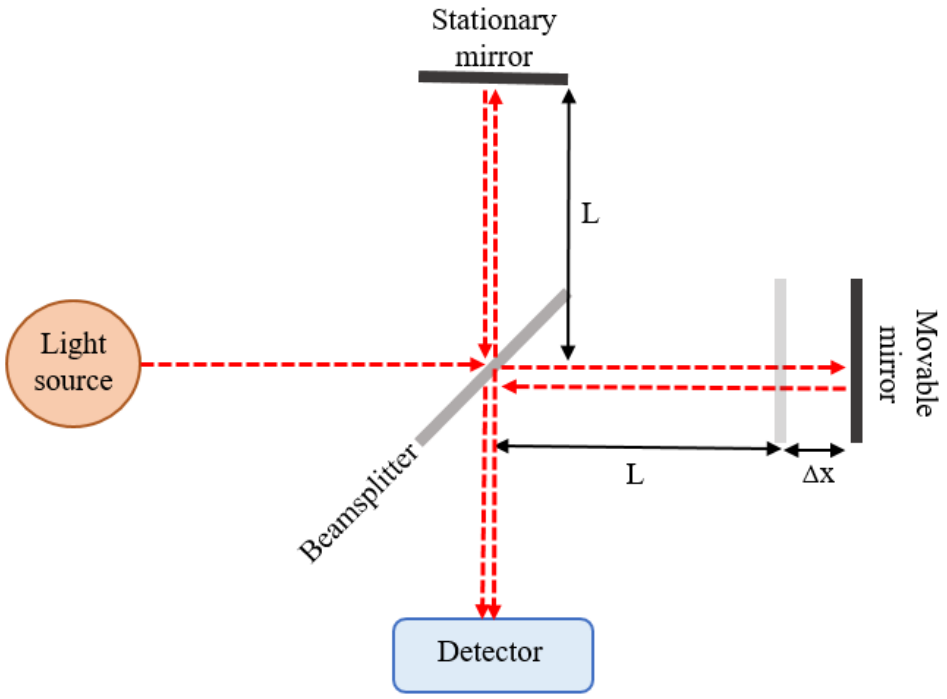


Figure 8. Basic construction of an interferometer.

In the interferometer, radiation from an infrared light source hits a beamsplitter. Approximately 50 % of the light is reflected towards a stationary mirror. When it hits the mirror, it will have travelled a length of L from the beamsplitter. The other 50 % of the light is transmitted towards a movable mirror. This beam will travel a distance of L+Δx. The difference in travelled path length between the two beams will be 2Δx when they meet in the beamsplitter. This difference is called the retardation. As the mirror moves back and forth, the retardation is changing. This will lead to a pattern of constructive and destructive interference. A plot of light intensity, measured by the detector, against the retardation is called an interferogram (Harris 2010). A computer transforms the interferogram into a plot of absorbance against wavenumber by Fourier transformation. Since light of all wavenumbers are investigated at once, the analysis only takes a few seconds (Williams & Fleming 2008).

1.3.2. Interpretation of spectra

The vibrational frequency of a bond depends on the strength of the bond and the mass of the atoms involved. Triple bonds have higher stretching frequencies than double and single bonds. Increased mass of the atoms involved results in decreased frequency. A C−C bond will therefore vibrate with a lower frequency than a C−H bond (Williams & Fleming 2008). In this project, absorbance spectra from the mid-infrared range are used. The mid-infrared spectrum ranges from 4000 cm^{-1} to 400 cm^{-1} (Li-Chan 2010) and can roughly be divided into four ranges. The first is the high-frequency area (4000−2400 cm^{-1}) where the stretching vibrations of single bonds O−H, C−H and N−H occur. The second area (2400−1900 cm^{-1}) is where triple bonds have stretching frequencies. The double bonds (stretching) and the single bond of N−H (bending) have vibration frequencies in the third area (1900−1500 cm^{-1}). The last area (1500−400 cm^{-1}), which is called the fingerprint area, contains frequencies of most of the bending and stretching vibrations of some single bonds. The fingerprint area contains complex absorption patterns and each molecule has its own characteristic fingerprint area (Williams & Fleming 2008). The absorption peaks in the region above 1500 cm^{-1} are distinct and may be used for identification of the functional groups in the molecule. (Williams & Fleming 2008).

The absorption peaks for the functional groups in the common FAs in milk are marked in Figure 9. The carbonyl group in an FA will be part of an ester linkage, and has been reported to absorb radiation around 1745 cm^{-1} (Afseth et al. 2010; Soyeurt et al. 2006). The stretching vibrations of C−H in a *cis* double bond will give an absorption band around 3010 cm^{-1}. For C−H in methyl and methylene groups, the stretching vibrations occur in the range between 3000 cm^{-1} and 2800 cm^{-1}. The bending vibrations of C−H in methyl and methylene groups occur in the range from 1440 cm^{-1} to 1300 cm^{-1}. The stretching vibrations of C−O bonds occur in the range between 1200 cm^{-1} and 1000 cm^{-1}. An absorption band around 966 cm^{-1} is related to the out-of-plane stretch of C−H in a double bond with *trans* configuration (Afseth et al. 2010). Extensive tables of absorption frequencies can be found in the literature, e.g. Williams and Fleming (2008).

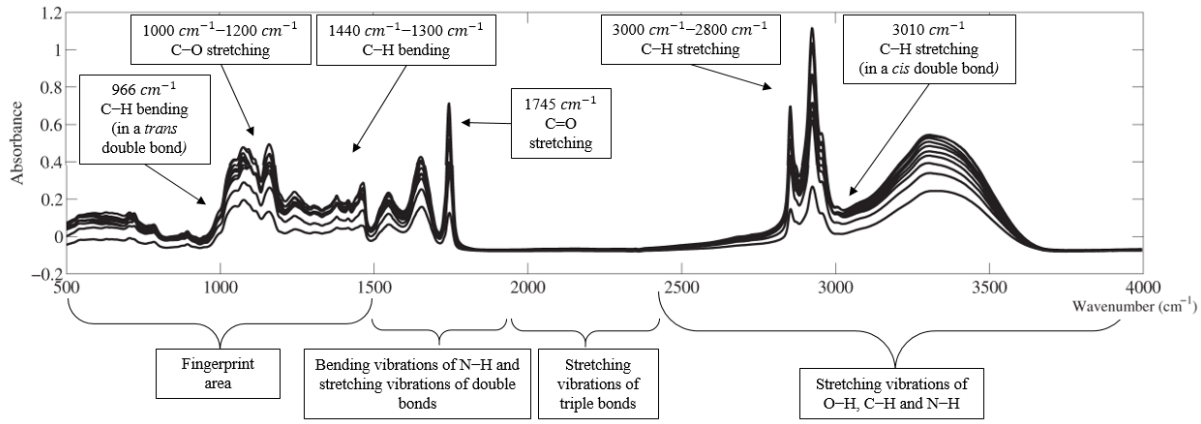


Figure 9. Vibration frequencies of the functional groups in common fatty acids in milk (upper part of the figure). The four main areas where vibrations of different bonds occur (lower part of the figure). The figure of the FTIR spectra is taken from Afseth et al. (2010) and is modified by Felicia Stehr.

Milk consists of several different components such as fat, protein and lactose. The FTIR spectrum of a milk sample will therefore be complex, as can be seen in Figure 9. It will consist of absorption peaks for all the bonds in the different molecules (Soyeurt et al. 2006). In general, it is difficult to interpret the peaks in the fingerprint area (Cozzolino 2013). However, by use of calibration equations the spectra may be used for prediction of various characteristics (Soyeurt et al. 2006).

1.3.3. Applications

FTIR is widely used for routine analysis of liquid milk samples for determination of its content of fat, protein and lactose (McParland & Berry 2016; Rutten et al. 2009). These quality parameters are often used for milk price calculations (Rutten et al. 2009). Research has also revealed that milk minerals, individual milk proteins and free amino acids can be predicted by FTIR analysis with low to moderate accuracy (McParland & Berry 2016). Other applications of FTIR include identification of microorganisms in milk, determination of urea content, and screening for subclinical ketosis. Near-infrared spectroscopy has been used for detection of cows with mastitis (Kohler et al. 2010).

Determination of FA composition in milk is not yet performed by FTIR on a routinely basis (Rutten et al. 2009). The determination of FA composition is often done by use of gas chromatography (GC), as this method provides accurate results (Afseth et al. 2010). Since GC is a time-consuming and expensive method, the possibility to use FTIR (which is both cheaper and easier to use) for prediction of FA composition has been investigated. Soyeurt et al. (2006) investigated the possibility of prediction of FA composition using FTIR. The experiment only included 49 milk samples, but it was reported that the spectra from FTIR

analysis of liquid milk samples could be used for prediction of most of the FAs (in g/dL) in milk (with square root of cross-validated coefficients of determination ranging from 0.71 to 0.94). The exceptions were for the FAs present in very low concentration (C10:1*cis*-9, C14:1*cis*-9 and C18:3*cis*-9, *cis*-12, *cis*-15). Their results showed that the prediction with FTIR became more efficient with increasing concentration of the FAs. A few years later, a similar experiment was conducted by another research group (Rutten et al. 2009). This experiment included 3622 milk samples and they concluded that all major FAs and combined groups of FAs could be predicted accurately with FTIR. These results were also consistent with the previous research regarding the inadequate prediction of the minor FAs in milk.

One way to increase the concentration of all FAs in the milk samples is to introduce a drying step, i.e. make dried thin films of the milk samples before FTIR analysis. A comparison of this dry-film procedure and liquid milk FTIR was performed by Afseth et al. (2010). Since liquid samples contain a considerable proportion of water, another aspect with FTIR analysis of liquid milk samples is that an IR spectrum will show a lot of absorption in the O−H bending and stretching regions (around 1600−1700 cm^{-1} and above 3020 cm^{-1}, respectively). This causes these regions to be more or less opaque to infrared light and the information in these regions are therefore lost (Afseth et al. 2010). In Figure 10, spectra of ten milk samples from the experiment, analysed with liquid FTIR (A) and dry-film FTIR (B) are presented. In Figure 10A, the O−H bending and stretching regions are very noisy compared to the same regions in Figure 10B.

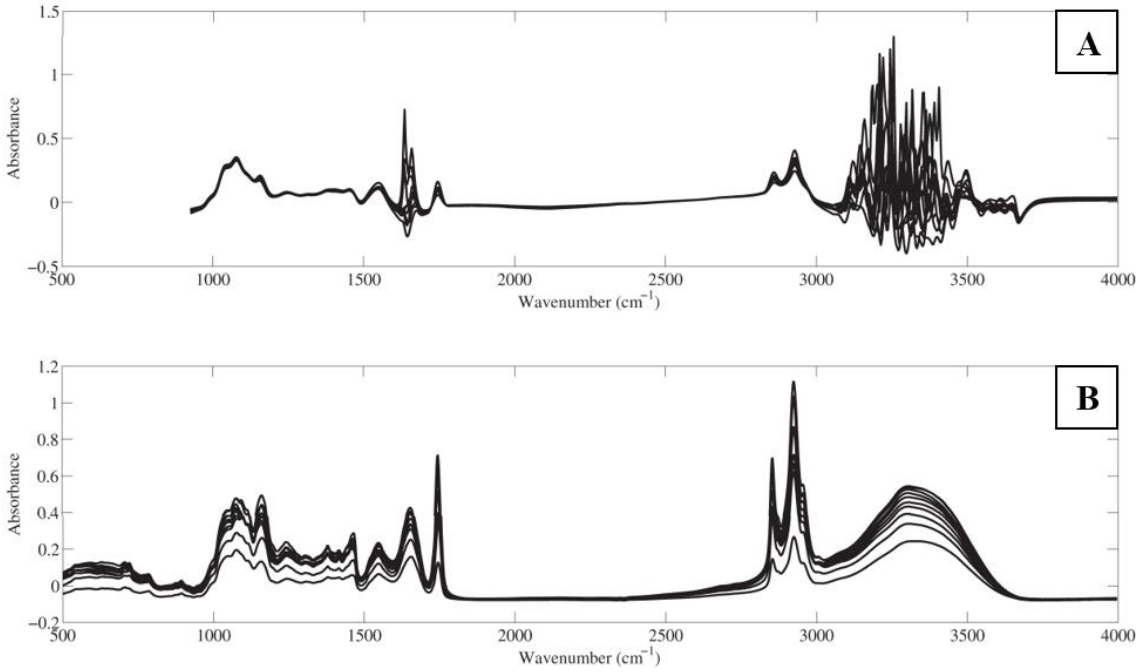


Figure 10. Ten raw FTIR spectra of liquid milk samples (A) and the corresponding spectra obtained with dry-film analysis (B) (Afseth et al. 2010).

In the experiment of Afseth et al. (2010), the estimation errors in the calibrations were overall lower for the dry-film method than for the liquid milk method. Only the dry-film method resulted in calibrations that were good enough for the major FAs C16:0, C18:0 and the minor FAs C4:0, C18:2*cis*-9, *trans*-11, summed PUFA and summed C18:1*trans* isomers. Calibrations obtained for C6:0−C14:0, C18:1*cis*-9, summed SFA, summed MUFA and the iodine value, were good for both approaches. The regions of 600 cm^{-1} to 1800 cm^{-1} and 2800 cm^{-1} to 3200 cm^{-1} were used in the dry-film FTIR analysis.

In the previously mentioned study by Martin et al. (2015), where the link between FA composition and OLA was investigated, they also explored the possibility of using dry-film FTIR for determination of FA composition in thawed milk. It was observed that the estimation errors were generally higher and the correlation coefficients were generally lower for the calibration developed for the previously frozen samples than for the fresh milk samples in the study by Afseth et al. (2010). They did however conclude that the calibrations still were feasible for determination of FA composition, and that this method could be used for analysis of historically frozen samples. Both these research groups believe that there is a potential for incorporation of dry-film FTIR analysis into AMSs and that this could provide the farmer with real-time information about the herd and the individual cow.

Since FA composition and EB are correlated (chapter 1.2.2), it could be expected that also EB may be predicted by FTIR spectra. This was investigated on liquid milk samples by McParland and Berry (2016), who obtained a calibration model with square root of cross-validated coefficients of determination of 0.78. However, when the calibration model was tested on a completely external validation data set, the prediction was of poor accuracy.

1.4. Multivariate data analysis

A multivariate data set is a data set that consists of more than one measured variable per sample (CAMO 2017), but the number of variables may be very high. Often these high-dimensional data sets consist of a larger number of variables $X_1, X_2, ..., X_p$ than the number of samples $(p > n)$ (James et al. 2013). As an example, absorbance is measured at several thousand wavenumbers for each sample in FTIR analysis. These spectra are collected in a matrix with the rows corresponding to the objects (samples) and the columns corresponding to the measured variables (wavenumbers).

A relationship between two or three variables (dimensions) is possible to represent graphically. However, for more than three variables it is no longer possible (Miller & Miller 2010). Multivariate data analysis may be used to reveal patterns in the high dimensional space, visualize the data (data exploration) and for predictions (Cozzolino 2013; Miller & Miller 2010). According to McParland and Berry (2016), chemometrics may be defined as "*the process of extracting information from chemical systems using data-mining techniques relating the spectral data to phenotypes*". The multivariate data analysis in this work includes pre-processing of the spectra, PCA and PLSR.

1.4.1. Pre-processing

FTIR spectra of milk are influenced by milk chemistry, milk physics and instrument physics (Kohler et al. 2010). Pre-processing can be used to enhance the chemical information in the data (Rajalahti & Kvalheim 2011). FTIR spectra may be pre-processed before calibration to remove scaling and offset variations. For example, varying thickness of the dry-films of milk can lead to scaling effects in the spectra (Afseth et al. 2010).

The physical effects may be reduced by taking the first or second derivative of the raw spectra. The derivative is often calculated using the Savitzky-Golay method. The minima in the spectrum of the second derivative will correspond to the peaks in the raw spectrum. The second-derivative spectrum may reveal information that is hidden in a spectrum containing many overlapping bands. In addition, the second derivative removes baseline effects (Kohler et al. 2010).

Normalisation is often performed to adjust for physical effects such as differences in thickness of samples and variations in the fraction of the transmitted light that hits the detector. Changes in sample thickness will change the optical path length, which cause multiplicative effects. Examples of normalisation techniques are standard normal variate,

multiplicative signal correction or extended multiplicative signal correction (EMSC) (Kohler et al. 2010).

In summary, the second derivative will correct for variations in the baseline of raw spectra while EMSC will remove physical variations such as differences in sample thickness (Kohler et al. 2010). In the previously established calibration models for prediction of FA composition by use of FTIR spectra (which are used in this project), the FTIR spectra were pre-processed by the Savitzky-Golay method followed by normalization by EMSC. The pre-processing was performed before calculating the mean of the three replicates (Afseth et al. 2010).

1.4.2. Principal component analysis

PCA is a technique that reduces the dimensions of a multivariate data matrix with n observations and p variables (X_1, X_2, ..., X_p). PCA does not involve any associated response (Y), and it is therefore often used for visualization of the data. PCA compresses the variables into principal components (James et al. 2013). The principal components are calculated from the covariance in the data set. The covariance is a measure of the joint variation of two variables. If the covariance of two variables is close to zero, there is no relationship between them. The principal components are linear combinations of all variables (X_1, X_2, ..., X_p) measured for the samples (Miller & Miller 2010). In Equation 7, this is illustrated for the first principal component (PC1):

$$Z_1 = \Phi_{11}X_1 + \Phi_{21}X_2 + \ldots + \Phi_{p1}X_p \qquad (7)$$

where Z_1 is the first principal component, X_1, ..., X_p are the original variables (absorbance at different wavelengths in FTIR) and Φ_{11}, ..., Φ_{p1}, are the coefficients of the principal component (James et al. 2013; Miller & Miller 2010). These are also called the loadings of the principal component (James et al. 2013). The loadings are chosen to result in non-correlating principal components and so that the PC1 describes most of the variation in the data set. The second principal component (PC2) will describe second most of the variation and will be orthogonal to PC1. This is illustrated in Figure 11.

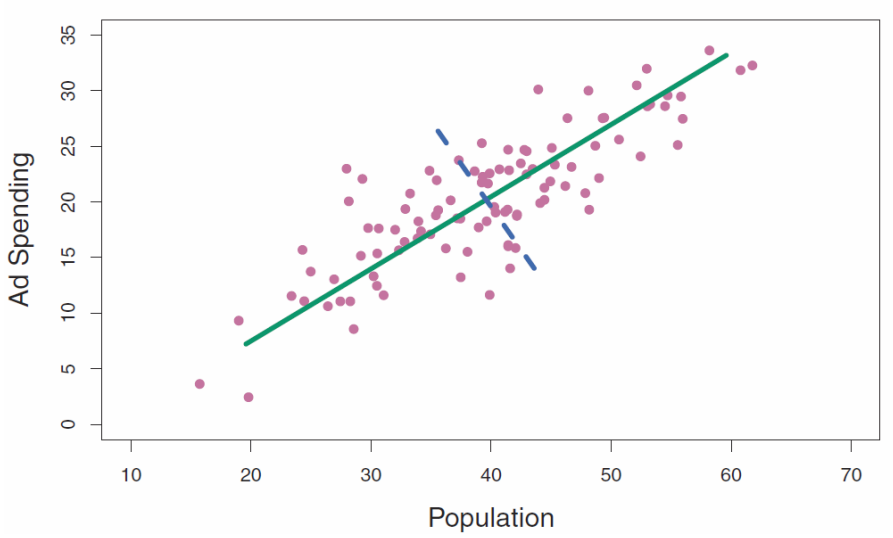


Figure 11. In this two-dimensional space, the green line represents the first principal component and the blue dashed line represents the second principal component. This example is taken from James et al. (2013). The ad spending is plotted against population size for 100 different cities (purple points).

PCA of a data set that consists of p variables will result in p principal components. However, when a significant correlation between the original variables exists, a reduction of the data set may be achieved. This is because only a few of all the principal components will contain useful information. Often, it may be possible to describe the major part of the information in a multivariate data set with only two dimensions (PC1 and PC2). It may also turn out that not all of the original variables in the dataset contribute significantly to the principal components. These may therefore be removed (Miller & Miller 2010).

Substitution of the loadings and the relevant absorbance values into Equation 7 for each of the samples, results in the scores of PC1. When plotting the scores for PC1 and PC2 against each other, a score plot is obtained. A score plot illustrates how the samples are related to each other. It is easier to detect any groupings in a score plot than by just looking at a raw data set (Miller & Miller 2010). Similar samples will be located close to each other in a score plot. Therefore, PCA is an important tool for visualization of data and can be used for identification of patterns or groupings (Cozzolino 2013). In addition, the variables may be visualized by plotting the loadings in a loading plot. Loading plots may be used for interpretation of the patterns in the score plots (Rajalahti & Kvalheim 2011). PCA may also detect outliers. A sample that is located far away from the other samples in the score plot may be an outlier. This sample may be analysed one more time or removed from the data set (Cozzolino 2013).

1.4.3. Partial least squares regression

In multivariate calibration the variables are divided into predictor variables and a response variable (X_1, X_2, ..., X_p and Y, respectively) (James et al. 2013; Miller & Miller 2010). A regression model that shows the quantitative relationship between the response and the predictor variables is calculated (Rajalahti & Kvalheim 2011). When determining the FA composition from FTIR spectra, the FA concentrations are response variables (Y-block) and the absorbance at the different wavenumbers are predictor variables (X-block). When spectra of several constituents in a sample are overlapping, multivariate calibration may be used to predict the concentration of each of the constituents. When a relationship between the concentration of an FA and the absorbance at different wavenumbers has been established, the concentration in new samples may be predicted. Common methods used in multivariate calibration are multiple linear regression, principal component regression (PCR) and PLSR (Miller & Miller 2010)**.**

The principal components in PCR are calculated in the same way as in PCA; they are chosen to describe as much of the variation in the original predictor variables as possible. The first principal components are then used as predictor variables in the regression. In PLSR, the variation in the original variables is described by a reduced number of components, like in PCA and PCR. The difference is that the original variables that are highly correlated to the response variable are given extra weight. Therefore, the components (the new predictor variables) in PLSR both describe the variation in the original variables and take into account the correlations between the original variables and the response variables (James et al. 2013; Miller & Miller 2010). It may turn out that a prediction model contains variables that are not significantly related to the response. These variables may be removed from the model without any significant loss of information. There are several methods for selection of the most significant variables (Rajalahti & Kvalheim 2011).

The regression coefficients in PLSR may describe which of the variables that are of importance (Minitab 2017). The regression coefficients may be used for evaluation of the robustness of a model and, as stated by Afseth et al. (2010), "*finding logical spectral features in the regression coefficients provides strong indications that the calibrations are based on reasonable chemical features and not on random or spurious correlations*".

1.4.3.1. Validation of calibration models

It is of great importance to investigate the predictive performance of an established calibration model, i.e. to validate the model. This is due to the risk of overfitting of a model. Overfitting may occur when the model is based on more variables than samples (Rajalahti & Kvalheim 2011) or when the same data is used both for calculation and validation of the model (Kohler et al. 2010). The most-used validation method is the cross-validation method (Rajalahti & Kvalheim 2011). In cross-validation, one or more samples are held out when building the calibration model. These samples are used for prediction. This procedure is repeated until all samples have been held out once. When using full cross-validation, only one sample is held out each time. In segmented cross-validation, the data set is divided into groups and one group is held out each time (Afseth 2007). Cross-validation is often used when determining the appropriate number of factors in a PLSR calibration (James et al. 2013).

To evaluate the precision and robustness of a calibration equation, several statistical parameters may be calculated (McParland & Berry 2016; Soyeurt et al. 2006). In the previously established PLSR calibration models (Afseth et al. 2010) for determination of FA composition from FTIR spectra of dry milk films used in this project, the calibration models were evaluated using the root mean square error of cross-validation (RMSECV) and the coefficient of determination (R^2) between the reference (y_i) and the predicted ($\hat{y}_i$) values. The optimal number of factors was determined using segmented cross-validation (20 segments) (Afseth et al. 2010). R^2 is a common measure of model fit. Calculation of R^2 always results in a value between 0 and 1. This number reflects the proportion of variance in the response explained by the predictors. If R^2 is close to 0, the regression model does not explain much of the variability in the response variable. Conversely, if R^2 is close to 1, a large proportion of the variability in the response variable has been explained by the predictor variables (James et al. 2013). The RMSECV is calculated with Equation 8:

$$RMSECV = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \qquad (8)$$

where i indicates the samples from 1 to N (total number of samples), y_i are the reference values and $\hat{y}_i$ are the predicted values (Afseth et al. 2010). The RMSECV is a measure of fit and is the average distance between a reference and a predicted value. One advantage of RMSECV is that it has the same unit as the y-values (Vernier Software & Technology 2001).

Optimally, prediction models should also be validated by use of an independent data set (a test set) (i.e. samples from a new population or new sampling dates). Validation using an independent data set often results in poorer prediction results than what cross-validation do. Thus, it is important to use the prediction models only on populations that are similar to the population from which the calibration data was collected (McParland & Berry 2016).

2. Materials and methods

2.1. Chemicals

Cow milk
Bronopol tablets, 2-bromo-2-nitro-1,3-propanediol
Distilled water
Ice
Isopropanol, 70 %
SDS solution, 0.05 %

2.2. Equipment

Pipettes:

Producer:	VWR
Model:	Ultra high performance
Volumes:	2−20 µL and 100−1000 µL
Producer:	Sartorius
Model:	Biohit mLINE M10
Volume:	0.5−10 µL
Producer:	Thermo Labsystems
Model:	Finnpipette Digital
Volume:	200−1000 µL
Producer:	Thermo Fisher Scientific
Volume:	0.5−10 µL

Vortex mixer:

Producer:	VWR collection
Model:	VV3

Sample well plates:

Supplier:	Bruker Optik GmbH
Model:	Silicon, 96 wells

FTIR instrument:

Producer:	Bruker Optik GmbH, Germany
Model:	Tensor 27
Year:	2004
Serial number:	T27.1472
Detector:	DLaTGS
Coupled unit:	HTS-XT (High-Throughput-Screening-Extension)
Software:	OPUS version 6.5

2.3. Project description

This study was performed on the research herd at the Animal Production Experimental Centre (SHF), which is a part of the Norwegian University of Life Sciences. The 64 cows in the project were divided into four feeding regimes. The first sub-objective was to investigate if PCA of the FTIR spectra could detect any systematic change in milk composition during the first 100 DIM. The chemical composition of the milk samples was provided by TINE (Heimdal, Norway). The second sub-objective was to predict FA composition of the milk samples using existing PLSR models and to search for any systematic change in the predicted FA composition during the first 100 DIM. It was also examined if there were any differences in the predicted FA composition between cows with subclinical ketosis and normal cows. Information about if the cows had subclinical ketosis or not was provided by the Department of Production Animal Clinical Sciences (ProdMed). A flow chart illustrating the structure of the project is given in Figure 12.

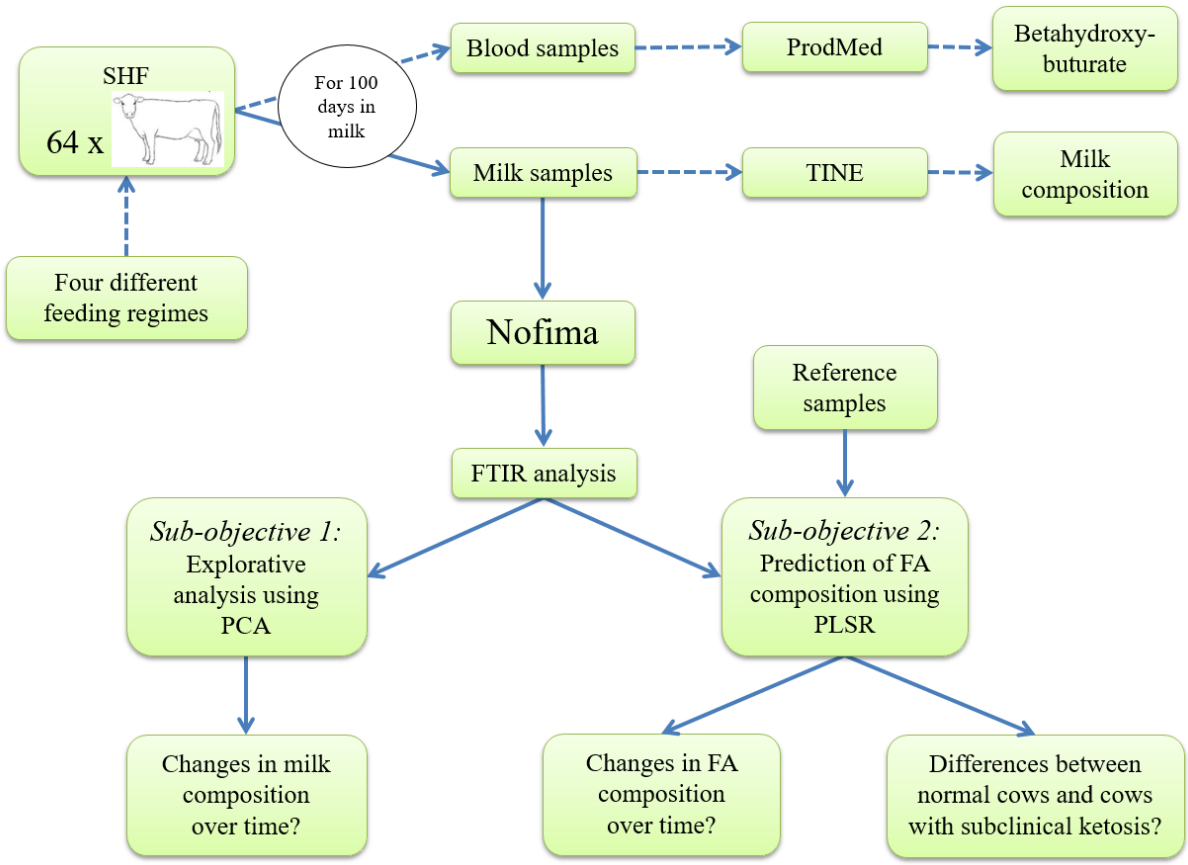


Figure 12. The structure of the present project. Abbreviations: the Animal Production Experimental Centre (SHF), the Department of Production Animal Clinical Sciences (ProdMed), principal component analysis (PCA), partial least squares regression (PLSR) and fatty acid (FA).

2.4. Sample preparation at The Animal Production Experimental Centre

Three times a week from September 2016 to February 2017, milk samples from the cows in the project were collected at SHF. OLA-beakers, with one tablet of bronopol added (for preservation), were filled with 70 mL milk. Bronopol was added to the milk to inhibit growth of microbes. A low concentration of bronopol does not influence the results significantly (Afseth et al. 2010; Butler & Stergiadis 2011). The beakers were marked with cow ID number and a period code representing the date of the sample outtake. The samples were heated to 39 °C in a water bath and shaken before they were split into three fractions. A volume of 40 mL was sent to Tine and about 10 mL were sent to ProdMed. The OLA-beakers with the remaining volume (approximately 20 mL) were kept in a refrigerator and later picked up by Nofima.

2.5. Sample preparation at Nofima

The OLA-beakers with the milk samples (Figure 13) were picked up at SHF, and placed in a refrigerator at Nofima. Before preparation of the samples, they were room tempered for about 15 to 30 minutes. The sample volume in the OLA-beakers varied a bit.

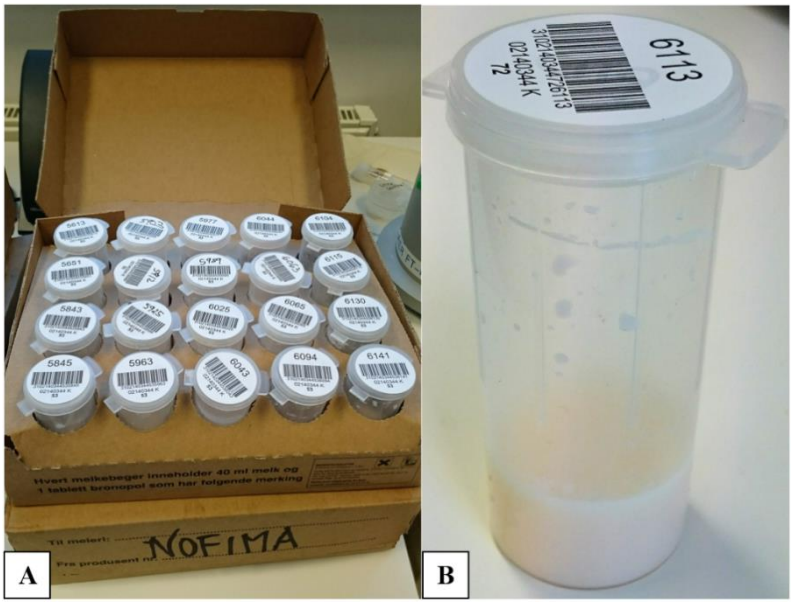


Figure 13. Milk samples in OLA-beakers from the Animal Production Experimental Centre (SHF) (A) and a milk sample from the cow with ID number 6113 from period 72 (B). Photo: Felicia Stehr.

For each milk sample, an Eppendorf tube (1.5 µL) was marked with the cow ID number and the period code of the milk sample (Figure 14). The empty tubes were placed on ice. The milk in each OLA-beaker were mixed for 10 seconds with a vortex mixer at speed 4 (out of

maximum 6), and 750 µL was pipetted into an Eppendorf tube (Figure 15A). Finally, 250 µL of distilled water was added to all the Eppendorf tubes and the lids were closed (Figure 15B).

Figure 14. Eppendorf tubes were marked with cow ID number and period code. Photo: Felicia Stehr.

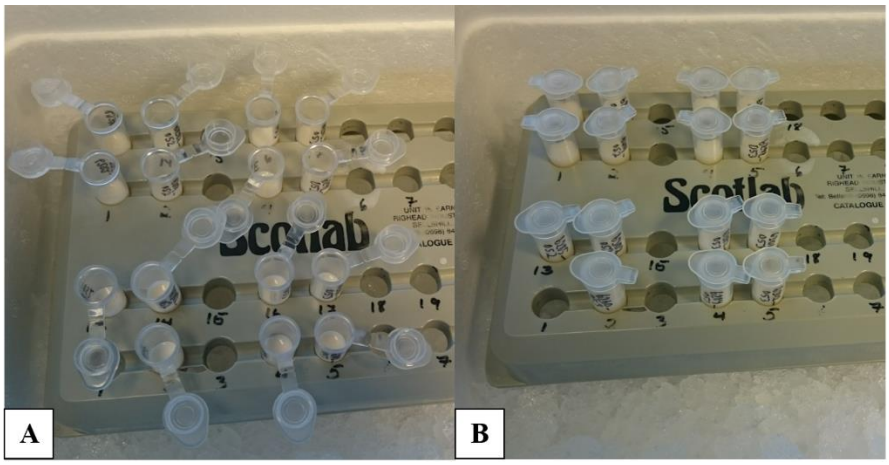


Figure 15. Eppendorf tubes on ice with milk (A) and water added (B). Photo: Felicia Stehr.

The diluted samples in the Eppendorf tubes were mixed for 5 seconds at maximum speed (6), before three replicates (2.5 µL) of each sample were transferred to a sample well plate. The sample ID number was written down in a paper scheme before each sample was transferred to the plate (Figure 16), to be sure of which well contained which sample. Then the droplets in the three wells were smeared out to thin films using the tip of the pipette. When all samples were transferred, the plate was left to air dry for at least one hour before FTIR analysis.

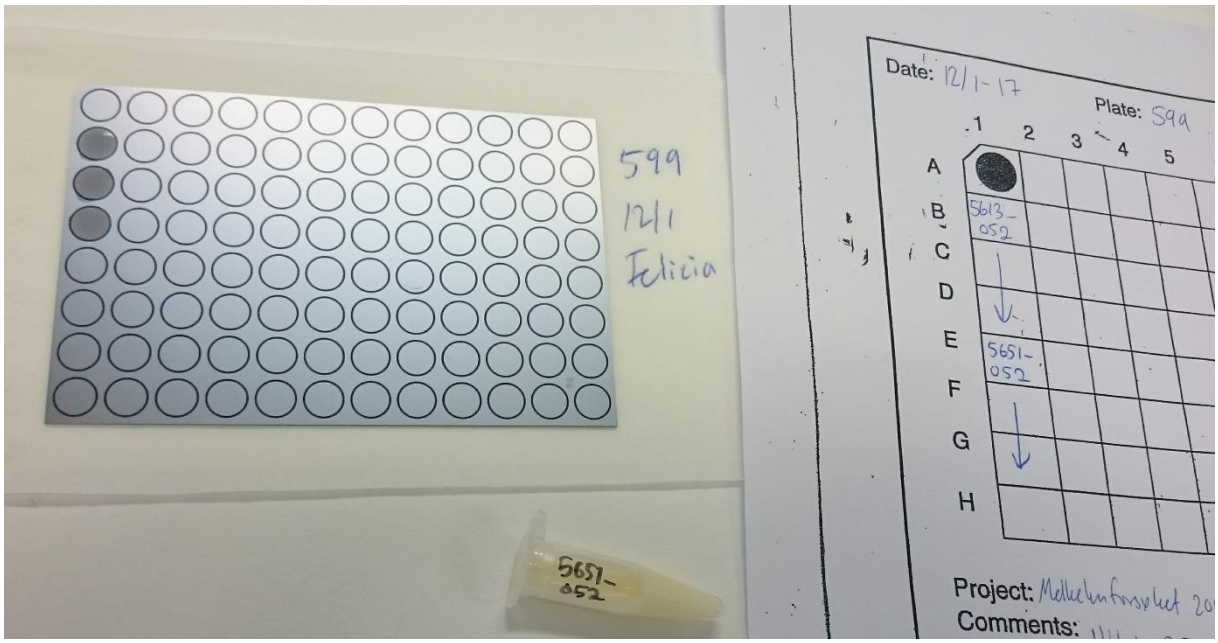


Figure 16. Three replicates (2.5 µL) of each milk sample were transferred to a sample well plate. The sample ID numbers were written down in a paper scheme before each sample was transferred to the plate. Photo: Felicia Stehr.

2.6. Fourier transform infrared spectroscopy analysis

Once a day, the FTIR instrument (Figure 17) was calibrated with the "position test" and a calibration plate, and the signal was checked. Performance qualification tests of the sample compartment and the external port were performed weekly.

Figure 17. The FTIR instrument with a HTS-XT coupled to it. Photo: Felicia Stehr.

Experiment file and storing options were selected in the OPUS software, and an identification table for the sample positions was created and saved on the computer. The instrument settings used for all analyses are listed in Table 2.

Table 2. The FTIR instrument settings used in the experiment.

Resolution (cm^{-1})	4
Sample scan time (number of scans)	40
Background scan time (number of scans)	40
Interval of saved data (cm^{-1})	4000 to 400
Result spectrum	Absorbance

Background spectra of the plate (the reference well) were collected before each sample measurement. Each spectrum was based on an average of 40 scans. More information about the instrument settings is provided in Appendix A. The plate with the dried milk samples was inserted into the FTIR instrument (Figure 18) and the analysis was started.

Figure 18. The port where the sample plate was inserted. Photo: Felicia Stehr.

When the analysis was finished, the spectra of all samples were controlled with respect to appropriate absorbance and other errors. The spectra were accepted if the absorbance peaks were approximately between 0.1 and 1.7. If any sample had a too high or too low absorbance spectrum, that sample was diluted with a more appropriate factor and analysed one more time.

When the FTIR analysis of the samples was completed and the spectra seemed to be fine, the silica plate was cleaned in several steps. First, the dry milk films were washed away with tap water and the plate was wiped with paper tissue. The plate was then cleaned with 0.05 % SDS and paper tissue, and rinsed with tap water. Thereafter the plate was coated with 70 % isopropanol for 30 seconds and then wiped with paper tissue. Finally, the plate was rinsed with distilled water and dried with paper tissue.

2.7. Data analysis

The data analysis in this project was divided into two parts. The first part consisted of explorative analysis of the milk composition in the milk samples. This was conducted using information about the chemical composition provided from TINE, and by performing PCA of the FTIR spectra. The second part consisted of establishment of PLSR models and prediction of FA composition in all milk samples. The data analyses (line plots, pre-processing, PCA and PLSR) of the FTIR spectra were performed with The Unscrambler X (version 10.3, CAMO Software AS). Microsoft Excel was used for various calculations, creation of tables and for keeping track of information about each sample. MATLAB (version R2017a, MathWorks, Inc.) was used to create additional plots.

2.7.1. Data collection

This study was performed on the research herd at SHF, which is a part of the Norwegian University of Life Sciences. The herd consisted of 64 Norwegian Red cows. The cows were followed for 100 DIM. DIM is defined as the number of days from calving. This expression is equivalent to days of lactation. Milk samples from the cows were collected three times a week from September 2016 to February 2017. The cows were milked with an AMS. The 64 cows in the project were divided into two groups: one group was fed with high quality food and the other group was fed with food of medium quality. Within these two groups, there were two subgroups: static feed allocation according to standard lactation curve and dynamic feed allocation. Three cows were excluded from the experiments due to clinical mastitis or other diseases. The remaining number of cows was 61. Samples that were collected before DIM 5 or after DIM 100 were removed. The number of remaining milk samples that were used in the data analyses was 2333. The number of samples, first and last sampling date and the corresponding DIM-range for each of the 61 cows is listed in Appendix B. The appendix also includes information about feeding regimes and whether the cow experienced subclinical ketosis.

2.7.2. Explorative analysis

2.7.2.1. Chemical composition of milk

The content of fat, protein, lactose, urea and FFAs for all milk samples was determined at TINE (Heimdal, Norway). These analyses were performed with a Bentley FCM 600 instrument (2015). Using Excel, the minimum, maximum, mean and standard deviation (SD) were calculated for these parameters. Samples with missing or unlikely values were removed. These statistics were therefore based on 2143 samples.

To explore how the content of the chemical components vary with lactation stage, the components were plotted for the first 100 DIM for three randomly selected cows. The cows had ID numbers 5977, 6238 and 6290. Due to large day-to-day variation, the data for fat, protein and lactose were smoothed, as described by Lindstrøm (2017). The smoothing model used was Legendre polynomials. The smoothed data was only used for visualization of trends, not for predictions of FAs. The average composition of the chemical components for all cows at DIM 5 to DIM 100 was calculated. The number of samples per DIM is provided in Appendix C. The average content was plotted for DIM 5 to DIM 100 to see if any trends became clearer compared to the plots for the individual cows.

2.7.2.2. Principal component analysis of spectra

To detect any visible outliers of the FTIR spectra, line plots (absorbance against wavenumber) were created. A PCA of all raw spectra from all sample replicates was performed to obtain an overview of the samples and to identify any outliers. This and further analyses with PCA were performed with segmented cross-validation (20 segments). The raw spectra were then pre-processed to correct for variations in baselines and sample thicknesses. First, the second derivative of all spectra was calculated using the Savitzky-Golay algorithm with a polynomial degree of 2, and 13 smoothing points. Thereafter all spectra were normalized by EMSC. A PCA was performed of all pre-processed replicates. Outliers were detected and evaluated for removal. Finally, the average spectrum of the three replicates for each sample was calculated.

Since the milk composition is known to vary through the lactation (Linn 1988; Ptak et al. 2012), PCA was used to investigate if any trends were detectable in the FTIR spectra. The average spectra of all samples from all cows at DIM 5 to DIM 100, respectively, were used. The number of samples per DIM is provided in Appendix C. It was also investigated if there were any trends for the three individual cows (with ID numbers 5977, 6238 and 6290). Milk

composition is also influenced by diet (Linn 1988), and the cows in this experiment were grouped in four different feeding regimes (Appendix B). A PCA was performed to investigate if any differences between cows with different feeding regimes could be revealed. This PCA shows the four feeding regimes as four different colours. Spectra of all samples at DIM 70−71 were used. These DIM were chosen because the results of Lindstrøm (2017) showed differences in EB between the groups at this point in time.

2.7.3. Prediction of fatty acid composition

2.7.3.1. Partial least squares regression models

Multivariate calibration models based on PLSR were established for prediction of FA composition in milk using FTIR spectra. The PLSR models were calibrated using the 219 reference samples from Afseth et al. (2010) and 25 reference samples from the current data set (244 samples in total). The FA composition of the 25 new reference samples was determined at Vitas AS in Oslo. The analyses of the reference samples were performed using GC, as described by Afseth et al. (2010). The fat fraction (separated from the water phase by centrifugation) was transesterified before analysis. The FA peaks were identified using a FAME (fatty acid methyl esters) mix. The proportions of the individual FAs were expressed in percentage of the total weight of FAs. The sums of SFAs, MUFAs and PUFAs were calculated directly from the GC results.

The data used in the calibration models were pre-processed in a different order than the data used for the explorative analysis. The pre-processing of the calibration samples also included the 2333 prediction samples. Since the calibration data from Afseth et al. (2010) was already averaged, the average spectrum of the three replicates of all samples was calculated first. Thereafter, the second derivative was calculated using the Savitzky-Golay algorithm with a polynomial degree of 2, and 13 smoothing points. The two wavenumber regions (from 600 cm^{-1} to 1800 cm^{-1} and from 2800 cm^{-1} to 3200 cm^{-1}) used in the prediction models were then normalized by EMSC.

PLSR models for prediction of the proportions of C10:0, C14:0, C16:0, C18:0, C18:1*cis*-9, CLA (conjugated linoleic acid) and the sums of SFAs, MUFAs and PUFAs were established using the 244 calibration samples. The FAs included in the sums of SFAs, MUFAs and PUFAs are listed in Table 1, Chapter 1.2.3. These FAs were chosen because they are interesting from a biological perspective (Chapter 1.2.2). The optimal number of PLSR

factors was determined using segmented cross-validation. Samples for the 20 segments were randomly chosen.

2.7.3.2. Fatty acid composition of milk

Prediction of the proportions of the FAs in the 2333 samples was performed using the optimal number of factors found for each model. The minimum, maximum, mean and SD was calculated for the FA contents. To explore how the content of the FAs vary with lactation stage, the contents were plotted for the first 100 DIM for the three randomly selected cows (with ID numbers 5977, 6238 and 6290). The average content of FAs for all cows was plotted for DIM 5 to DIM 100 to see if any trends became clearer compared to the plots for the individual cows. The number of samples per DIM used when calculating the average content for all cows is provided in Appendix C. A PCA of the average FA composition from DIM 5 to DIM 100 was performed to visualize how the content of different FAs changed during the lactation.

The level of betahydroxybuturate in blood samples from the cows was measured at ProdMed. This was done in order to detect cows with subclinical ketosis. Clinical ketosis is a disease that causes reduced appetite, reduced milk yield and weight loss. Ketosis may come from negative EB (TINE Rådgiving 2008) and EB is related to FA composition (Gross et al. 2011). Therefore, it was investigated whether the FTIR spectra could reveal any differences between cows with subclinical ketosis and normal cows. This was done by plotting the predicted FA contents over time for the two groups. The average content at DIM 5 to DIM 100 of samples from the two groups was used. The number of samples per DIM used when calculating the average content for the two groups is provided in Appendix C. There were 41 normal cows and 20 cows with subclinical ketosis (Appendix B). In this project, a cow was defined to experience subclinical ketosis if the concentration of betahydroxybuturate was more than 1 ng/mL in week three of lactation.

3. Results

The main objective of this work was to analyse milk from cows during their first 100 days in milk (DIM) using dry-film FTIR, and to investigate if the resulting FTIR spectra were suitable for detection and quantification of milk components. The first sub-objective was to investigate if PCA of the FTIR spectra could detect any systematic change in milk composition during the first 100 DIM. The second sub-objective was to predict fatty acid (FA) composition of all milk samples using existing PLSR models and to search for any systematic change in the predicted FA composition during the first 100 DIM. It was also examined if there were any differences in the predicted FA composition between cows with subclinical ketosis and normal cows.

3.1. Explorative analysis

The explorative analyses were performed in order to visualize the chemical composition of milk during the first 100 DIM. The information about the chemical composition provided from TINE was first used to describe the chemical composition for all samples. Then it was used to make plots of the chemical content during the lactation. This information was later used for explanation of trends in the FTIR spectra. The trends in the spectra were evaluated using PCA.

3.1.1. Chemical composition of milk

The minimum, maximum, mean and SD of the content of fat, protein, lactose, urea and FFAs are provided in Table 3. The statistics are based on 2143 samples from 61 cows.

Table 3. Minimum, maximum, mean and standard deviation (SD) of parameters measured for 2143 milk samples from 61 cows.

Chemical component	Min	Max	Mean	SD
Fat (%)[1]	2.0	8.0	4.0	1.0
Protein (%)[1]	2.5	4.7	3.4	0.3
Lactose (%)[1]	4.1	5.5	4.9	0.2
Urea (mmol/L)	2.4	8.3	5.1	0.8
FFAs[2] (mmol/L)	0.1	5.4	0.5	0.4

[1] Percentage of milk weight. [2] Free fatty acids (FFAs).

3.1.2. Chemical composition during the first 100 days in milk

To explore trends for the chemical components, the contents of fat, protein, lactose, urea and FFAs during the lactation were plotted for the three randomly selected cows (with ID numbers 5977, 6238 and 6290). The average content for all cows were also investigated. The number of samples per DIM is provided in Appendix C. Due to large day-to-day variation in the data for fat, protein and lactose, these data were smoothed with Legendre polynomials as described by Lindstrøm (2017). The smoothed data made it easier to visualize trends.

Plots of the fat content over time for the three cows, based on unsmoothed data, are provided in Figure 19A. Plots based on smoothed data are provided in Figure 19B.

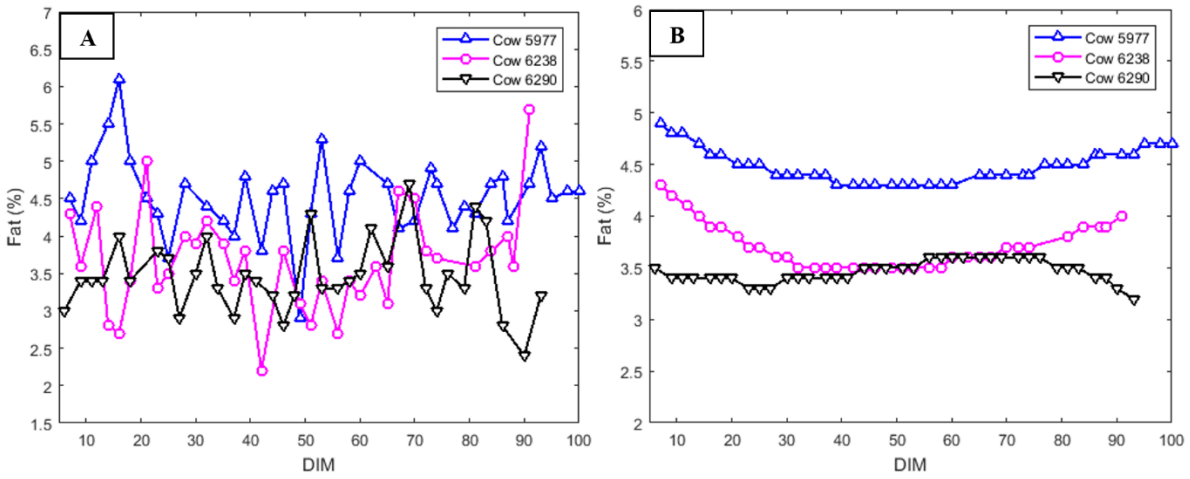


Figure 19. Fat content (in percentage of milk weight) during the first 100 days in milk (DIM) for three randomly selected cows. Unsmoothed data is used in Figure A and smoothed data is used in Figure B.

Figure 19A illustrates that there were large day-to-day variations in fat content for each cow. The smoothed data in Figure 19B made it easier to explore the trends. For two of the cows, the fat content decreased in the period from DIM 5 to DIM 30. Between DIM 30 and DIM 50, the fat content seemed to be stable and thereafter it increased.

The average fat content for all 61 cows during the lactation is plotted in Figure 20, both for unsmoothed and smoothed data.

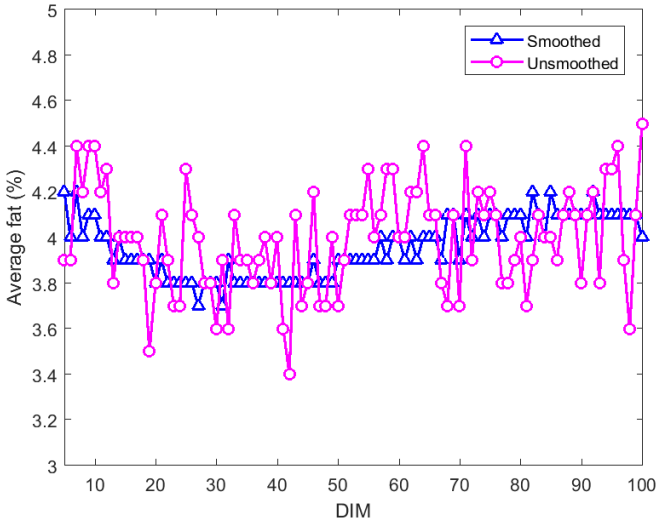


Figure 20. Average fat content (in percentage of milk weight) of milk samples from 61 cows during the first 100 days in milk (DIM).

The graph of unsmoothed data in Figure 20 shows large variation between the DIM. When looking at the smoothed data, there is a clear trend for fat content during the lactation. The fat content decreased in the period from DIM 5 to approximately DIM 30. Between DIM 30 and DIM 50, it was stable and thereafter it increased. This trend was hard to detect for the unsmoothed data due to the large variations.

The protein content over time for the three cows is provided in Figure 21, based on unsmoothed and smoothed data in Figure 21A and B, respectively.

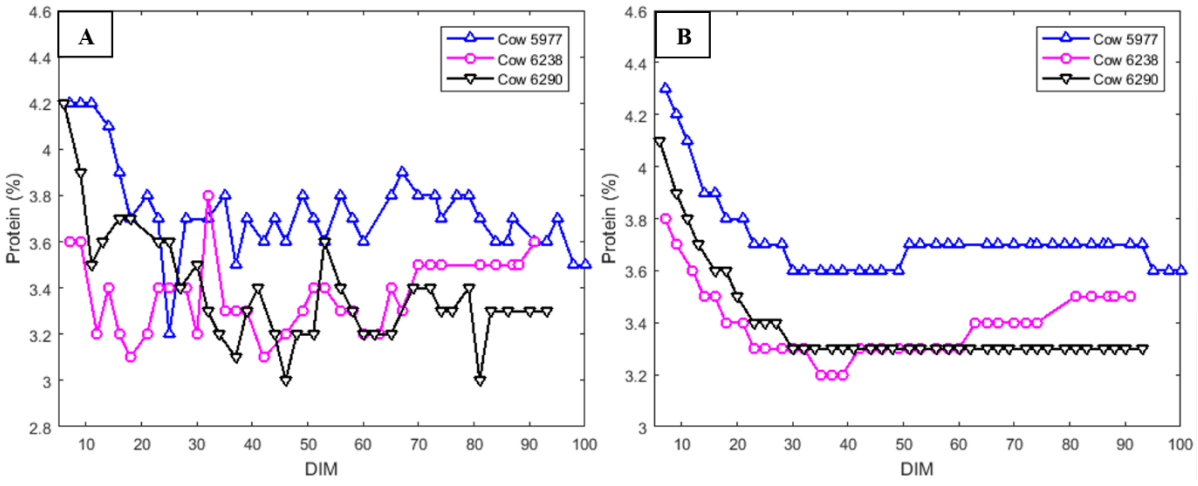


Figure 21. Protein content (in percentage of milk weight) during the first 100 days in milk (DIM) for three randomly selected cows. Unsmoothed data is used in Figure A and smoothed data is used in Figure B.

Both the unsmoothed and smoothed data in Figure 21 indicates that the protein content was highest in the beginning of the lactation. The smoothed data shows that it decreased in

the period from DIM 5 to about DIM 30. After this point, it was stable for two cows, and increased a bit for one cow.

The average protein content during the lactation is plotted in Figure 22. This figure contains graphs for both unsmoothed and smoothed data for all 61 cows.

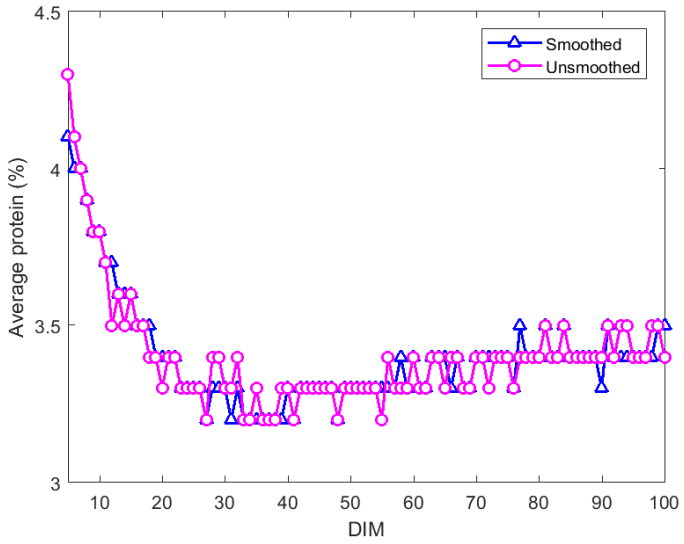


Figure 22. Average protein content (in percentage of milk weight) of milk samples from 61 cows during the first 100 days in milk (DIM).

In Figure 22, there is a distinct trend for the protein content during the first 100 DIM. From DIM 5 to DIM 30, the content decreased and then it increased slightly during the next 70 DIM.

The lactose content over time for the three cows is provided in Figure 23. The graphs are based on unsmoothed and smoothed data in Figure 23A and B, respectively.

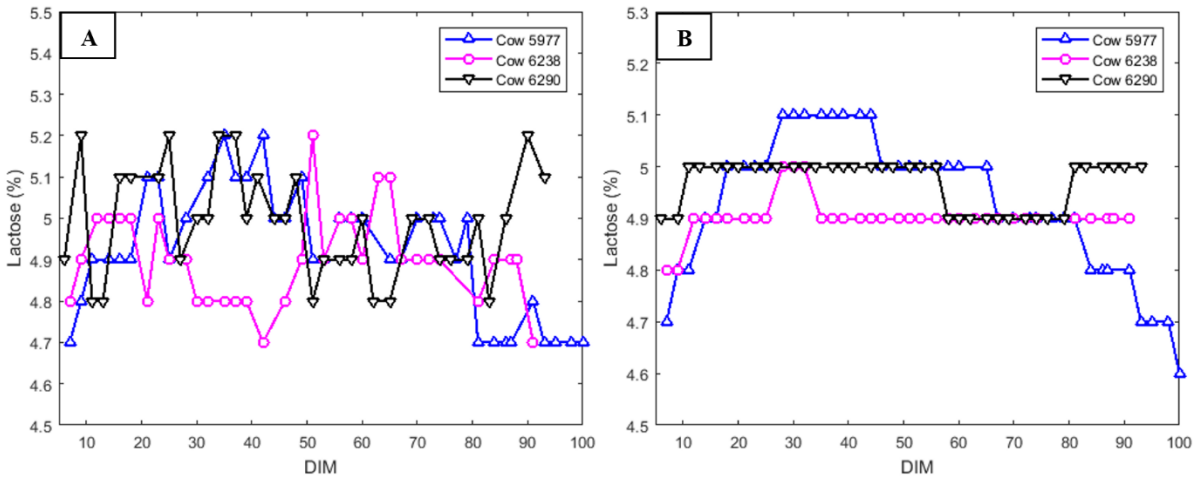


Figure 23. Lactose content (in percentage of milk weight) during the first 100 days in milk (DIM) for three randomly selected cows. Unsmoothed data is used in Figure A and smoothed data is used in Figure B.

The range of lactose variation is narrow compared to, for instance, the fat range. The lactose content for the cow with ID number 5977 (blue line in Figure 23A) only varied between 4.7 % and 5.2 % during the 100 DIM. There are different trends for the three cows. Inspection of Figure 23B reveals that the content seemed to increase during the first 30 DIM for two cows, while it seemed to be stable for one cow. Thereafter it seemed to decrease before it increased for one cow, remained stable for the second cow and decreased for the third cow. The graphs of unsmoothed and smoothed data for the cow with ID number 6238 differ quite much from each other.

The average lactose content during the lactation is plotted in Figure 24. This figure contains graphs for both unsmoothed and smoothed data for all 61 cows.

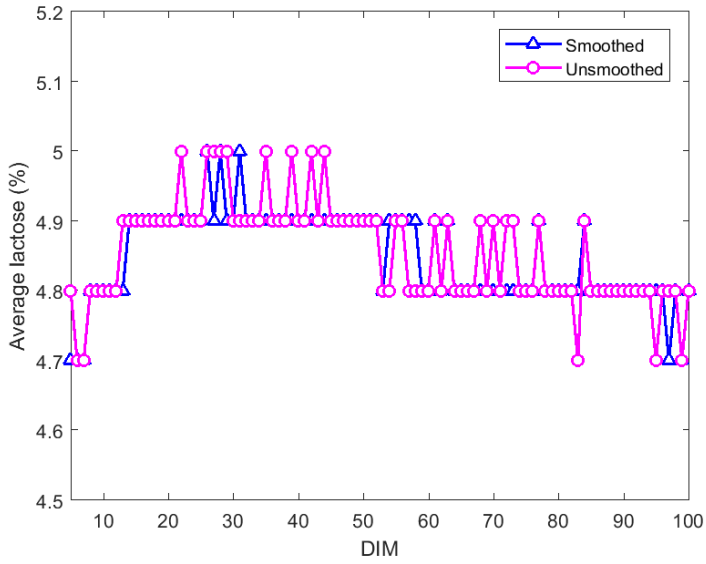


Figure 24. Average lactose content (in percentage of milk weight) of milk samples from 61 cows during the first 100 days in milk (DIM).

The average lactose content for all 61 cows resulted in graphs that seem to follow a trend (Figure 24). The lactose content increased during the first 25 DIM. The following 15 days it seemed to be stable and thereafter it decreased. The content during the lactation only varied between 4.7 and 5.0 %.

The content of FFAs during the first 100 DIM for the three cows is plotted in Figure 25A. The average concentration for all cows is plotted in Figure 25B.

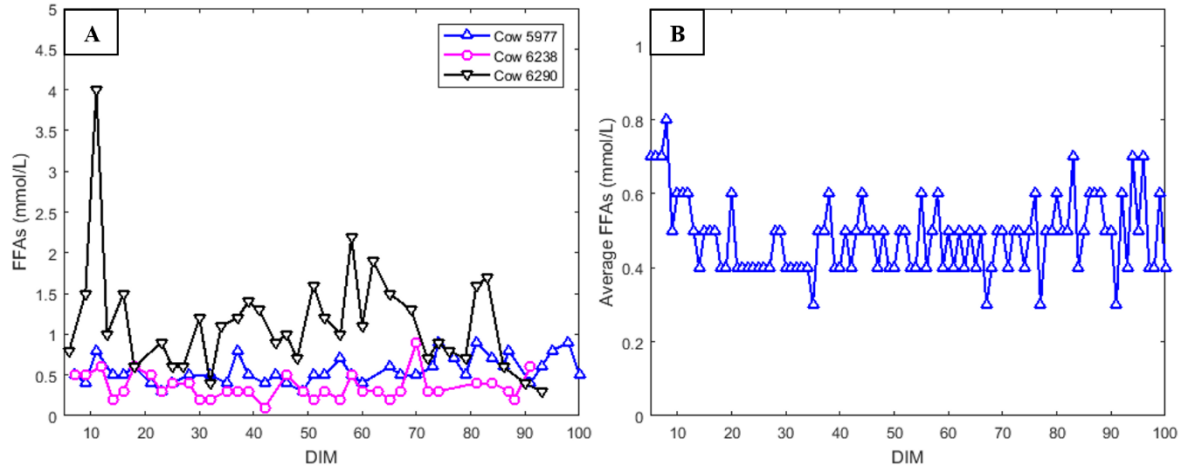


Figure 25. Content of free fatty acids (FFAs) (mmol/L) during the first 100 days in milk (DIM) for three randomly selected cows (A) and for the average of all 61 cows (B).

Figure 25A illustrates that the content of FFAs in the samples from the cow with ID number 6290 was higher than in the samples from the other two. There are no visible trends for the individual cows. However, for the average in Figure 25B the content of FFAs seemed to be a bit higher in the start of the lactation.

The concentration of urea during the first 100 DIM for the three cows is plotted in Figure 26A. The average concentration for all cows is plotted in Figure 26B.

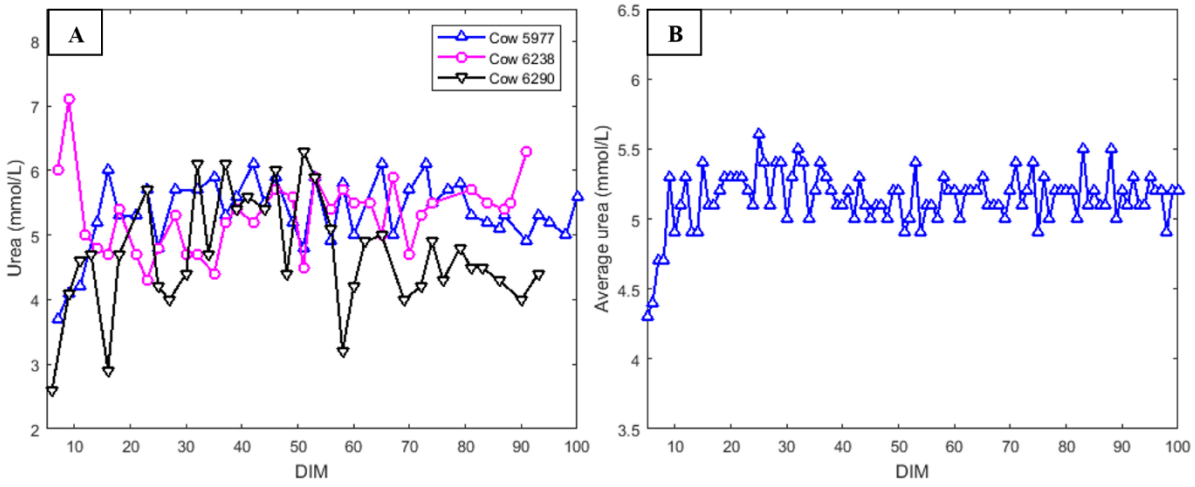


Figure 26. Urea content (mmol/L) during the first 100 days in milk (DIM) for three randomly selected cows (A) and for the average of all 61 cows (B).

Figure 26B indicates that the concentration of urea was lowest in the beginning of the lactation. The urea content for two of the individual cows (Figure 26A) seemed to follow this trend, while it seemed to be the opposite for the third cow.

3.1.3. The Fourier transform infrared spectra

3.1.3.1. Raw spectra and outlier detection

To illustrate the raw FTIR absorbance spectra, the spectra of the three replicates of one sample from each of the three selected cows (with ID numbers 5977, 6238 and 6290) are provided in Figure 27. These spectra were from the third week of lactation.

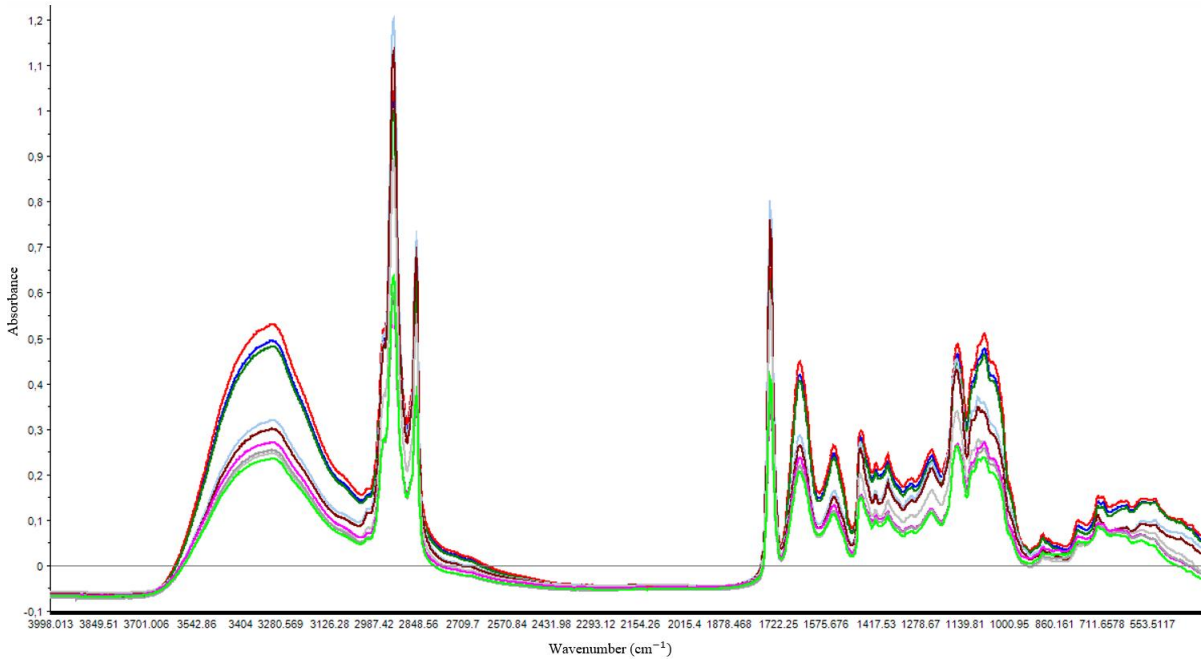


Figure 27. Raw absorbance FTIR spectra of three replicates of milk samples from each of the three cows. These samples are from the third week of lactation.

The absorbance spectra in Figure 27 show main absorbance peaks at wavenumbers 3300 cm^{-1}, 2923 cm^{-1}, 2854 cm^{-1}, 1745 cm^{-1}, 1654 cm^{-1}, 1548 cm^{-1} and a range of different peaks in the fingerprint area. Information about how to interpret FTIR spectra are provided in Chapter 1.3.2 (illustrated in Figure 9). The absorbance peak at 3300 cm^{-1} comes from stretching vibration of O−H and N−H in lactose and protein, respectively. The peaks at 2923 cm^{-1} and 2852 cm^{-1} are associated with symmetric and asymmetric stretching vibration of C−H in methyl and methylene (in fat). A peak at wavenumber 1745 cm^{-1} is characteristic for stretching vibration of C=O in an ester (in fat). The peaks at 1656 cm^{-1} and 1548 cm^{-1} are characteristic for the two stretching vibrations in a secondary amide (in protein) (Afseth et al. 2010; Williams & Fleming 2008). From 6999 replicate spectra, only five spectra were removed. This removal was due to strong water vapour signals in the FTIR spectra.

A PCA of all remaining sample replicates was performed to explore how the samples related to each other and to identify any outliers. The score plot and the loading plot are provided in Figure 28A and B, respectively.

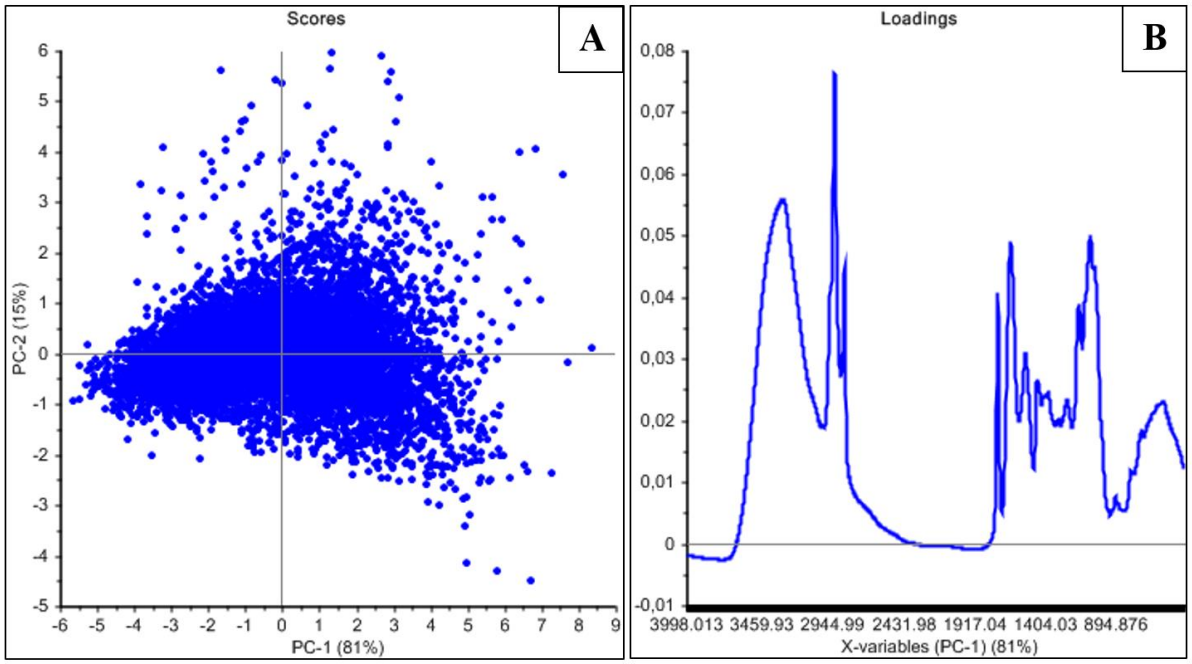


Figure 28. Score (A) and loading plot (B) from a PCA of raw spectra measured for the sample replicates of raw milk.

The score plot in Figure 28A shows how the samples relate to each other. No obvious outliers were detected in the score plot. PC1 and PC2 explained 81 % and 15 % of the variance, respectively. The loading plot in Figure 28B reveals the wavenumbers that are of importance in PC1. This plot is similar to the raw spectra, which is expected due to intensity variations related to film thickness.

3.1.3.2. Pre-processed spectra and outlier detection

All FTIR spectra were pre-processed with the Savitzky-Golay method and EMSC to correct for variations in baseline and sample thickness. The pre-processed spectra (for the same samples as the raw spectra in Figure 27) are shown in Figure 29A. To illustrate the difference between spectra pre-processed only with the Savitzky-Golay algorithm and spectra pre-processed both with the Savitzky-Golay and the EMSC methods, smaller parts of the spectra are shown in Figure 29A and B. The range from 1450 cm^{-1} to 1300 cm^{-1} was chosen. This is the area where the bending vibrations of C−H occur. As illustrated in Figure 29B and C, the multiplicative effects were reduced by EMSC.

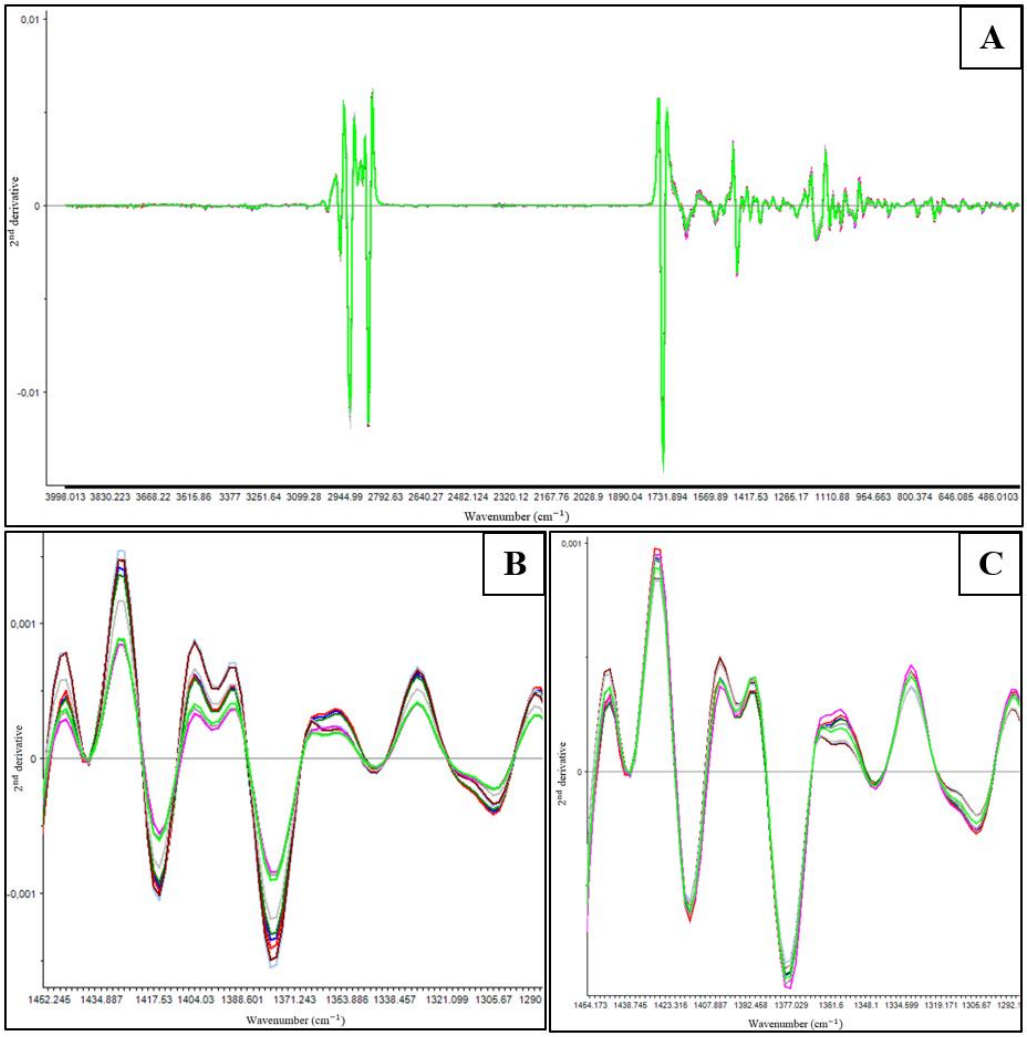


Figure 29. Savitzky-Golay + EMSC spectra of three sample replicates of three milk samples (A). Savitzky-Golay pre-processed spectra (B) and Savitzky-Golay + EMSC pre-processed spectra (C) in the wavenumber range of 1450−1300 cm^{-1}.

The minima in the pre-processed spectra in Figure 29A correspond to the absorbance peaks in the raw spectra (Figure 27). The main minima are located at wavenumbers 3006 cm^{-1}, 2960 cm^{-1}, 2923 cm^{-1}, 2852 cm^{-1}, 1745 cm^{-1}, 1656 cm^{-1}, 1463 cm^{-1}. Absorbance at 3006 cm^{-1} is characteristic for C−H stretch in a *cis* double bond (in fat). The peaks at 2960 cm^{-1}, 2923 cm^{-1} and 2852 cm^{-1} are associated with symmetric and asymmetric stretching vibration of C−H in methyl and methylene (in fat). The peak at wavenumber 1745 cm^{-1} is characteristic for stretching vibration of C=O in an ester (in fat). The peak at 1656 cm^{-1} is characteristic for one of the two stretching vibrations in a secondary amide (in protein). The peak at wavenumber 1463 cm^{-1} is probably due to bending vibration of C−H (in fat) (Afseth et al. 2010; Williams & Fleming 2008).

A PCA of the pre-processed spectra of all replicates was performed to investigate if there were any outliers after pre-processing. The score plot and loading plot from a PCA of the pre-processed spectra of all replicates are provided in Figure 30A and B, respectively.

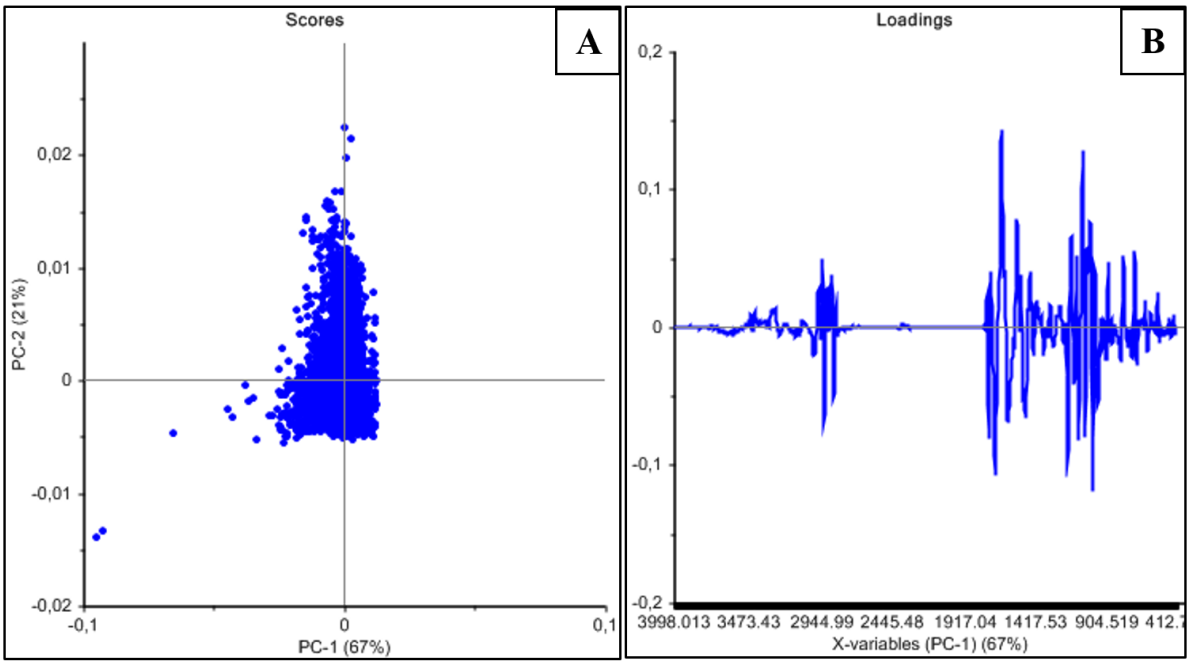


Figure 30. Score (A) and loading plot (B) from a PCA of the pre-processed spectra measured for the sample replicates of raw milk.

In Figure 30, the three replicates of one sample from the cow with ID number 5989 at period code 40 are located far away from the other samples in the score plot. However, all the three replicates lie very close to each other (the two points in the lower left corner are actually three points). Since they were close to each other and the absorbance spectra seemed fine, they were not considered outliers that should be removed. PC1 and PC2 explained 67 % and 21 % of the variance in the pre-processed data set, respectively. The loading plot illustrates that the wavenumbers around 2800 cm^{-1} and in the range from about 1800 cm^{-1} to 900 cm^{-1} are important for explanation of the variance.

Finally, the average spectrum of the pre-processed replicates was calculated for all samples.

3.1.3.3. Spectra during the first 100 days in milk

PCA of the FTIR spectra was performed to investigate if the spectra could reveal trends in chemical composition during the lactation. The PCA was based on the average spectra of all the samples from all cows at DIM 5 to DIM 100. The result is provided in Figure 31. The number of samples per DIM is provided in Appendix C. Blue, red and green scores represent the DIM ranges 5−36, 37−68 and 69−100, respectively.

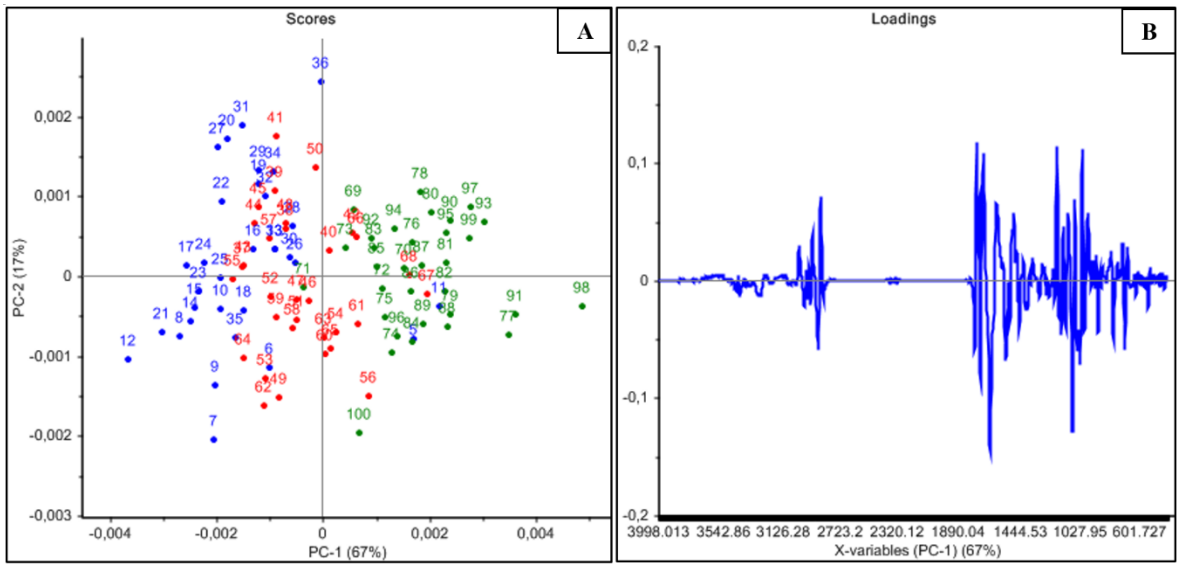


Figure 31. Score (A) and loading plot (B) from a PCA of the average spectra of all samples from all cows at different days in milk (DIM). This PCA includes milk samples from DIM 5 to DIM 100.

Figure 31A indicates a trend in chemical composition during the first 100 DIM. The samples from early and late DIM are located on opposite sides in the plot. The loading plot in Figure 31B shows that the most important wavenumbers in this PCA are between 1760 cm^{-1} and 1450 cm^{-1} and between 1200 cm^{-1} and 1000 cm^{-1}. The most important peaks in the loading plot were found at 1751 cm^{-1}, 1656 cm^{-1} and 1074 cm^{-1}. There was also a small band at 2852 cm^{-1}. Information about how to interpret FTIR spectra are provided in Chapter 1.3.2 (illustrated in Figure 9).

The sampling period was then divided into two ranges to explore if there were trends during shorter time intervals. The results of the PCAs based on DIM 5−50 and DIM 51−100 are provided in Figure 32 and Figure 33, respectively. Blue scores represent DIM 5−25 (in Figure 32) and DIM 51−75 (in Figure 33). Red scores represent DIM 26−50 (in Figure 32) and DIM 76−100 (in Figure 33).

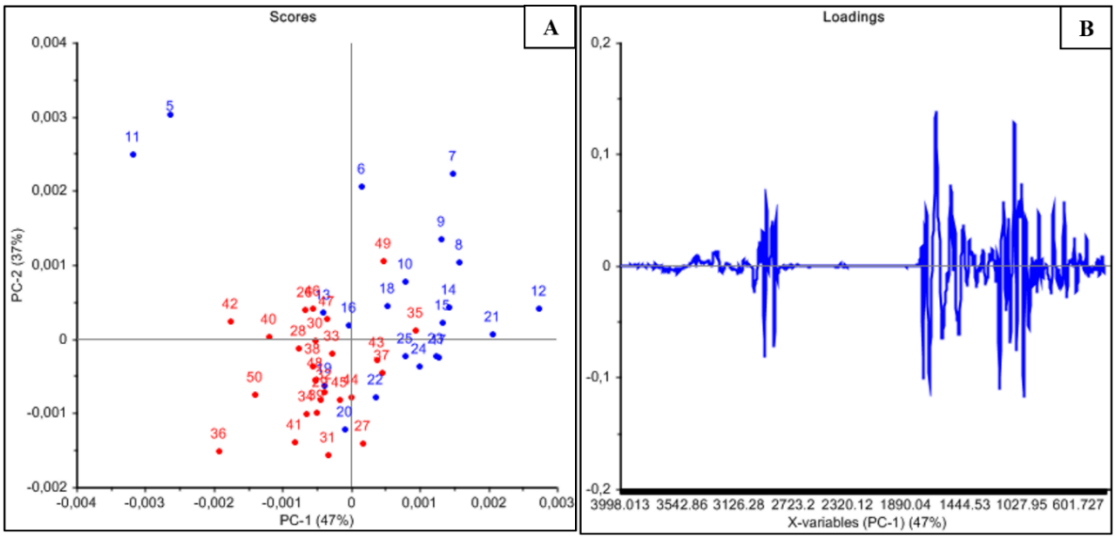


Figure 32. Score (A) and loading plot (B) from a PCA of the average spectra of all samples from all cows at different days in milk (DIM). This PCA includes milk samples from DIM 5 to DIM 50.

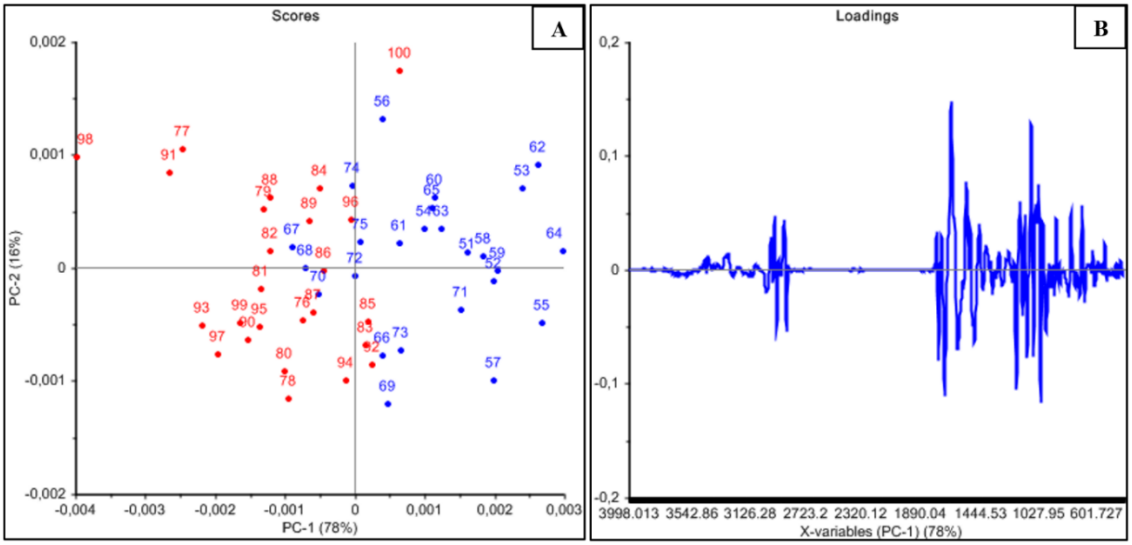


Figure 33. Score (A) and loading plot (B) from a PCA of the average spectra of all samples from all cows at different days in milk (DIM). This PCA includes milk samples from DIM 51 to DIM 100.

In Figure 32 and Figure 33, samples from different stages in the lactation are separated to some extent in the score plots. The loadings are similar to the ones in Figure 31B. The most important peaks in the loading plot were found at 1656 cm^{-1} and 1074 cm^{-1}.

Then it was investigated if there were any trends for the three individual cows (with ID number 5977, 6238 and 6290) during the sampling period. The results of the PCAs of the spectra from samples from the three individual cows at DIM 5−100 are provided in Figure 34, Figure 35 and Figure 36.

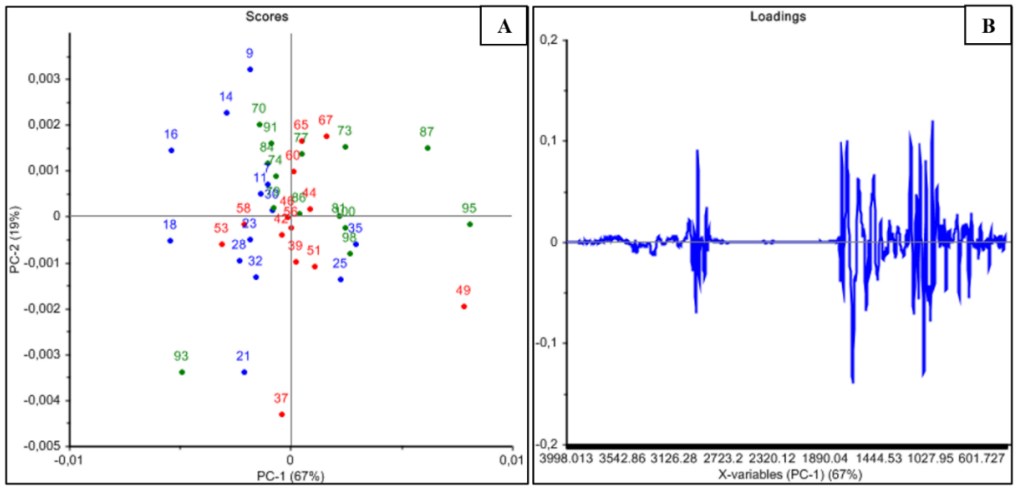


Figure 34. Score (A) and loading plot (B) from a PCA of the spectra measured for the cow with ID number 5977 at the different days in milk (DIM). This PCA includes milk samples from DIM 5 to DIM 100.

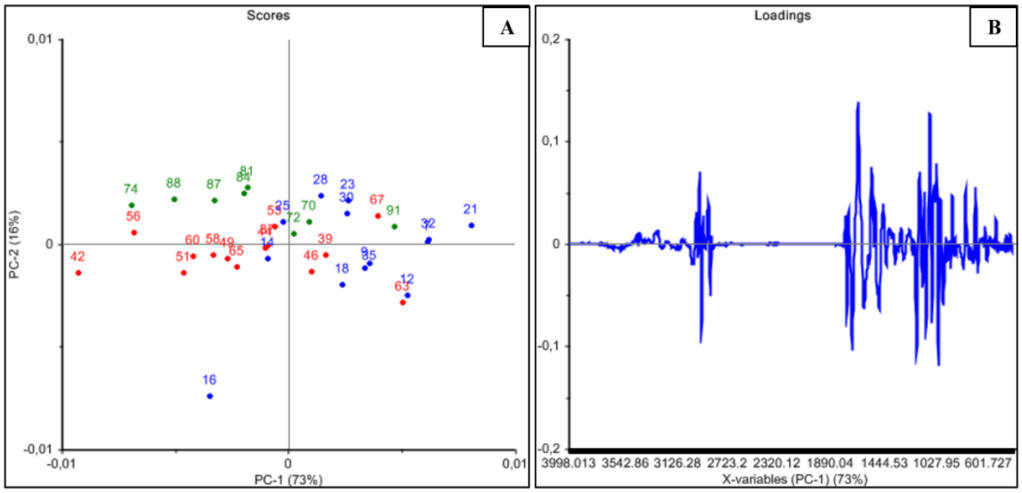


Figure 35. Score (A) and loading plot (B) from a PCA of the spectra measured for the cow with ID number 6238 at the different days in milk (DIM). This PCA includes milk samples from DIM 5 to DIM 100.

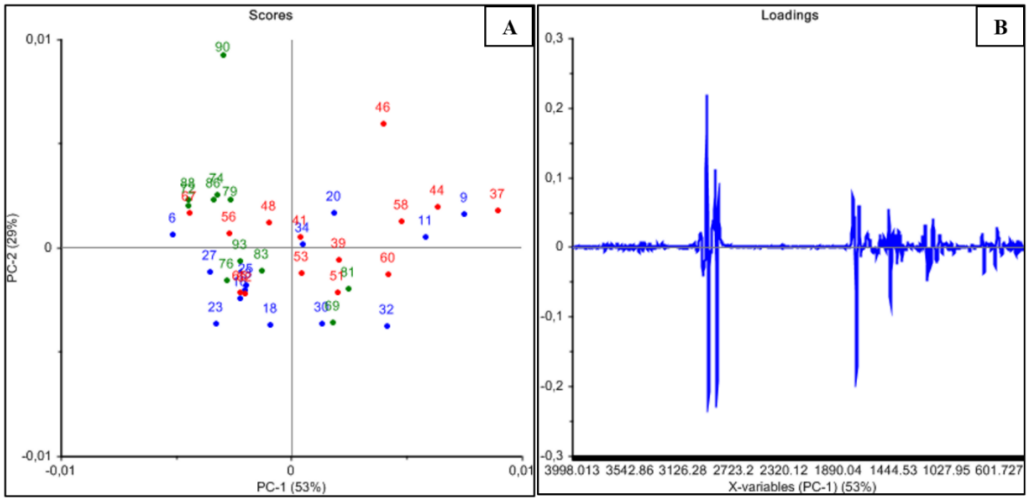


Figure 36. Score (A) and loading plot (B) from a PCA of the spectra measured for the cow with ID number 6290 at the different days in milk (DIM). This PCA includes milk samples from DIM 5 to DIM 100.

There were no distinct trends in the spectra from DIM 5 to DIM 100 for two of the individual cows (Figure 34A and Figure 36A). For the cow with ID number 6238 (Figure 35), the early DIM seemed to be located to the right, while the middle and late DIM seemed to be located to the left. However, this was not as obvious as for the average spectra for all cows. Any trends did not become clearer by plotting other principal components than PC1 and PC2. The loadings in Figure 34B and Figure 35B indicate that important wavenumbers are 1745 cm^{-1}, 1656 cm^{-1}, 1074 cm^{-1} and 2923 cm^{-1}. The loading plots for these two cows are similar to the loading plots for the average spectra. The loading plot for the cow with ID number 6290 (Figure 36B) differed from the others. The most important peaks in the loading plot were found at wavenumber 2923 cm^{-1} and 1737 cm^{-1}.

3.1.3.4. Feeding regime

Milk composition is influenced by factors such as breed, health and diet (Linn 1988). The cows in this experiment were grouped in four different feeding regimes (Appendix B). One part of the project was to examine if a PCA of the spectra could visualize any differences between milk samples from cows with different feeding regimes. The PCA was based on samples from cows at DIM 70 and DIM 71. These DIM were chosen because the results of Lindstrøm (2017) showed differences in EB between the groups at this time point. The score and loading plot from this PCA are provided in Figure 37A and B, respectively. The four different feeding regimes are shown as four colours. The score plot did not show any distinct groupings of samples from cows with the same feeding regime.

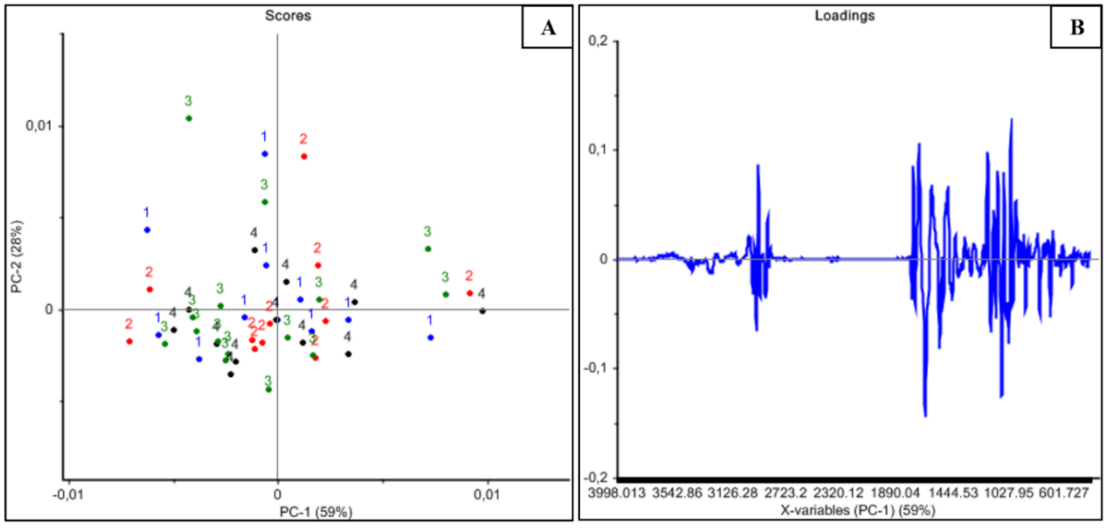


Figure 37. Score (A) and loading plot (B) from a PCA of the pre-processed spectra of milk samples from 50 cows at DIM 70 and DIM 71. The four different feeding regimes (numbered from 1 to 4) are shown in different colours.

3.2. Prediction of fatty acid composition

Prediction of the FA composition in the 2333 milk samples from 61 cows was conducted using modified PLSR models based on previously established calibrations (Afseth et al. 2010). The predicted proportions of the FAs were plotted for the first 100 DIM to explore any trends in the predicted FA composition.

3.2.1. Partial least squares regression models

PLSR models for prediction of FA composition in milk from dry-film FTIR spectra were established. The 219 calibration samples from Afseth et al. (2010) and 25 samples from the current experiment were included in the models (244 samples in total). It was also tested to calibrate the PLSR models only with the 219 old samples and to use the 25 new samples for test set validation. Results from this validation indicated that the FA content in the 25 new reference samples could be predicted well.

Calibration models were obtained for C10:0, C14:0, C16:0, C18:0, C18:1*cis*-9, CLA, and sums of SFAs, MUFAs and PUFAs. From a biological perspective, these were the most interesting FAs (Chapter 1.2.2). Except for CLA and sum of PUFA, these are the most abundant FAs in milk and good calibration models for these were established by Afseth et al. (2010). The FAs included in the sums of SFAs, MUFAs and PUFAs are listed in Table 1, Chapter 1.2.3. The optimal number of PLSR factors was determined using segmented cross-validation. A PCA of the 244 calibration samples was performed to explore how the old and new samples related to each other (Figure 38A and B).

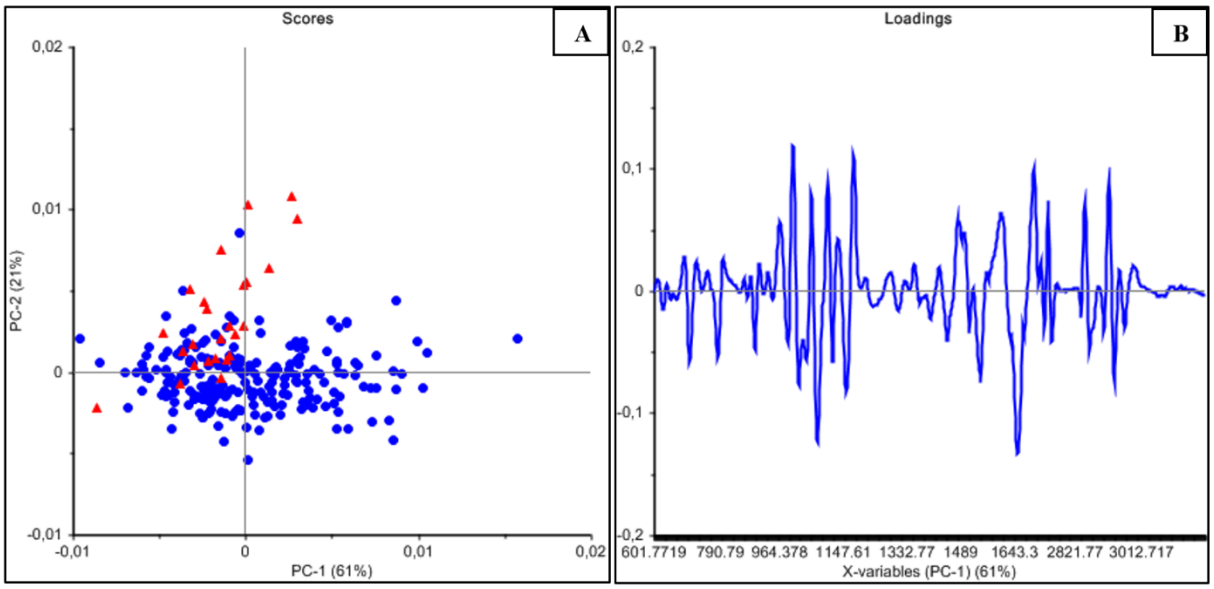


Figure 38. Score (A) and loading plot (B) from a PCA of the pre-processed spectra of the 244 calibration samples. The blue points in the score plot represent the 219 samples from Afseth et al. (2010). The red triangles represent the 25 reference samples from the current data set.

There are no obvious outliers in the score plot in Figure 38A. Most of the new samples (red triangles) are located near the old samples (blue points). The loading plot in Figure 38B indicates that many wavenumbers over the analysed ranges (from 600 cm^{-1} to 1800 cm^{-1} and from 2800 cm^{-1} to 3200 cm^{-1}) are important for describing the variations in the data set. The calibration results for prediction of the FA composition are provided in Table 4.

Table 4. Calibration results obtained for the PLSR models for prediction of fatty acid proportions in milk samples from dry-film FTIR spectra.[1]

Fatty acid	RMSECV	R^2	Factors
C10:0	0.15	0.93	10
C14:0	0.43	0.85	9
C16:0	0.80	0.92	11
C18:0	0.50	0.88	14
C18:1*cis*-9	0.55	0.97	10
CLA	0.08	0.82	15
SFA	0.75	0.96	10
MUFA	0.56	0.97	10
PUFA	0.16	0.76	18

[1] Abbreviations: root mean square error of cross-validation (RMSECV), coefficient of determination (R^2), optimal number of PLSR factors determined using segmented cross-validation (factors), conjugated linoleic acid (CLA), sum of saturated fatty acids (SFA), sum of monounsaturated fatty acids (MUFA) and sum of polyunsaturated fatty acids (PUFA).

The results in Table 4 show that calibration models with the best R^2 were obtained for C10:0, C16:0, C18:1*cis*-9, and sums of SFAs and MUFAs. The model with the lowest R^2 was obtained for the sum of PUFAs. The plots of predicted vs. reference *y*-values, and the regression coefficients for the calibration models are provided below. Information about the vibrations and bonds corresponding to the different wavenumbers is provided in Chapter 1.3.2 (illustrated in Figure 9). The 25 new reference samples are indicated with blue colour in the following figures.

The plot of predicted and reference response values obtained for the C10:0 prediction model is provided in Figure 39A. The regression coefficients are provided in Figure 39B.

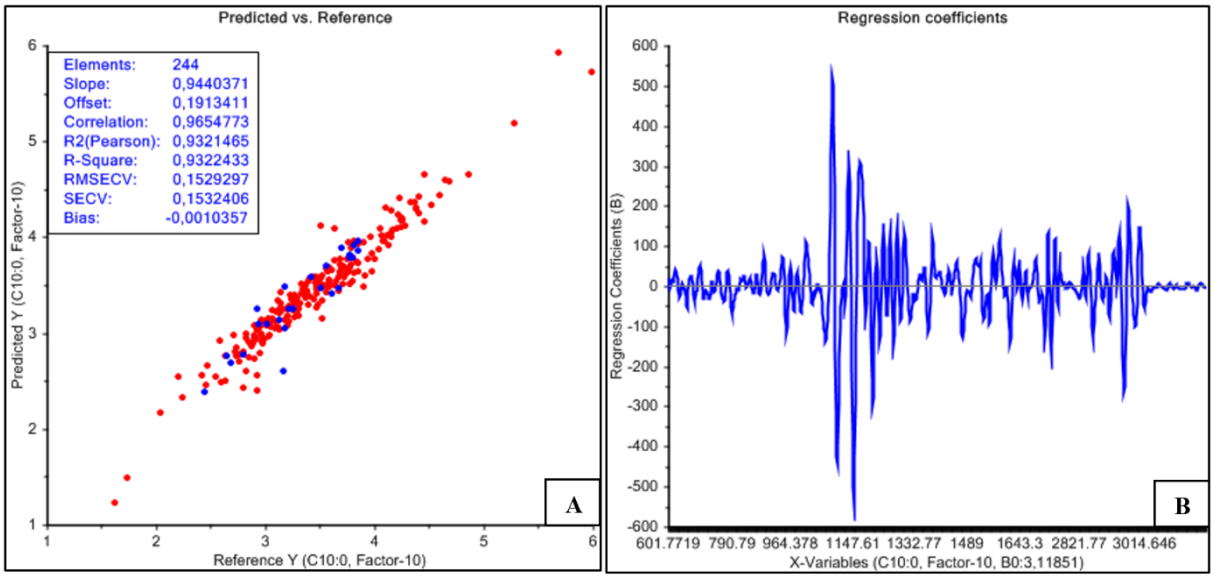


Figure 39. Predicted and reference response values (A) and regression coefficients (B) of the PLSR model obtained for C10:0. The model was based on 219 samples from Afseth et al. (2010) (red dots) and 25 samples from the current data set (blue dots).

The plot in Figure 39A reveals a good correlation between reference and predicted values. The blue points, which represent the new reference samples, are close to the old samples. The most important peaks in the plot of the regression coefficients in the prediction model for C10:0 (Figure 39B) were found at wavenumbers 1089 cm^{-1}, 1103 cm^{-1} and 1153 cm^{-1}.

The plot of predicted and reference response values obtained for the C14:0 prediction model is provided in Figure 40A. The regression coefficients are provided in Figure 40B.

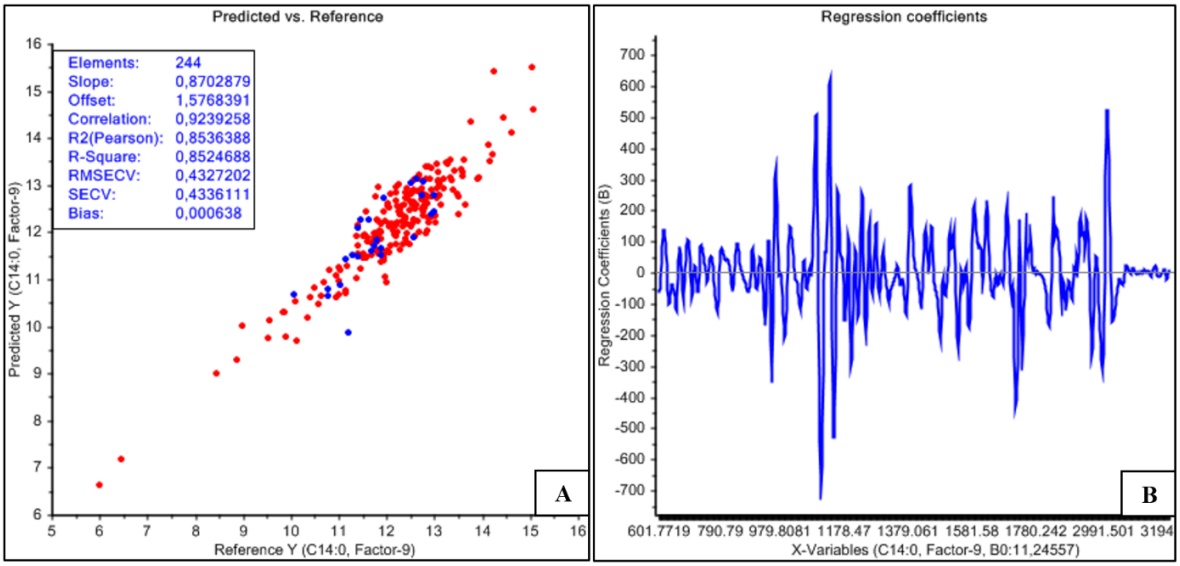


Figure 40. Predicted and reference response values (A) and regression coefficients (B) of the PLSR model obtained for C14:0. The model was based on 219 samples from Afseth et al. (2010) (red dots) and 25 samples from the current data set (blue dots).

The plot in Figure 40A reveals a good correlation between reference and predicted values. The blue points, which represent the new reference samples, are close to the old samples. The most important peaks in the plot of the regression coefficients in the prediction model for C14:0 (Figure 40B) were found at wavenumbers 966 cm^{-1}, 1095 cm^{-1}, 1163 cm^{-1}, 1720 cm^{-1} and 3006 cm^{-1}.

The plot of predicted and reference response values obtained for the C16:0 prediction model is provided in Figure 41A. The regression coefficients are provided in Figure 41B.

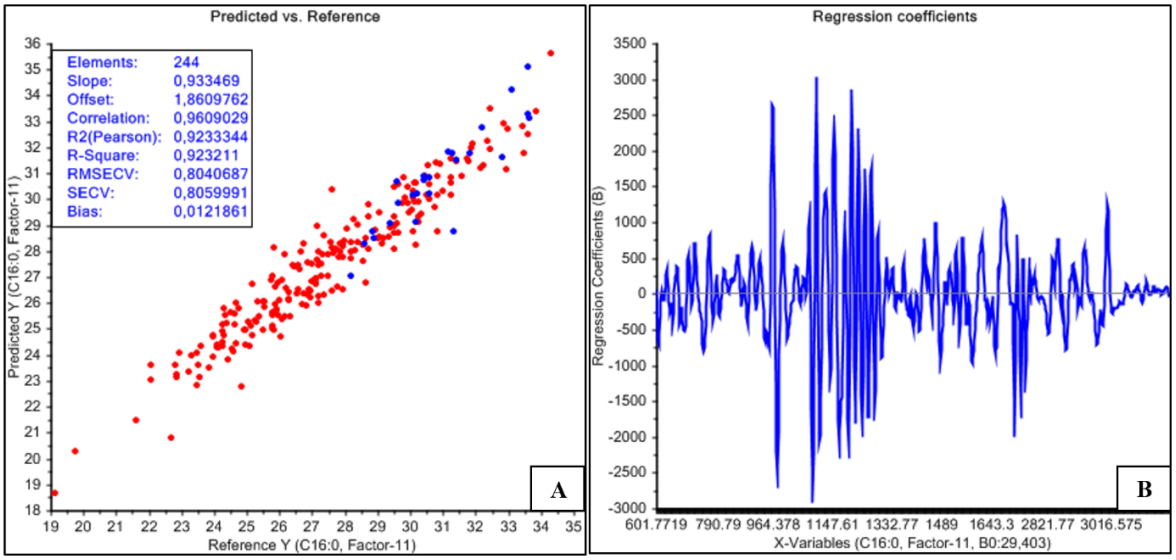


Figure 41. Predicted and reference response values (A) and regression coefficients (B) of the PLSR model obtained for C16:0. The model was based on 219 samples from Afseth et al. (2010) (red dots) and 25 samples from the current data set (blue dots).

The plot in Figure 41A reveals a good correlation between reference and predicted values. The blue points, which represent the new reference samples, are close to the old samples. The most important peaks in the plot of the regression coefficients in the prediction model for C16:0 (Figure 41B) were found at wavenumbers 966 cm^{-1}, 1172 cm^{-1} and 1261 cm^{-1}.

The plot of predicted and reference response values obtained for the C18:0 prediction model is provided in Figure 42A. The regression coefficients are provided in Figure 42B.

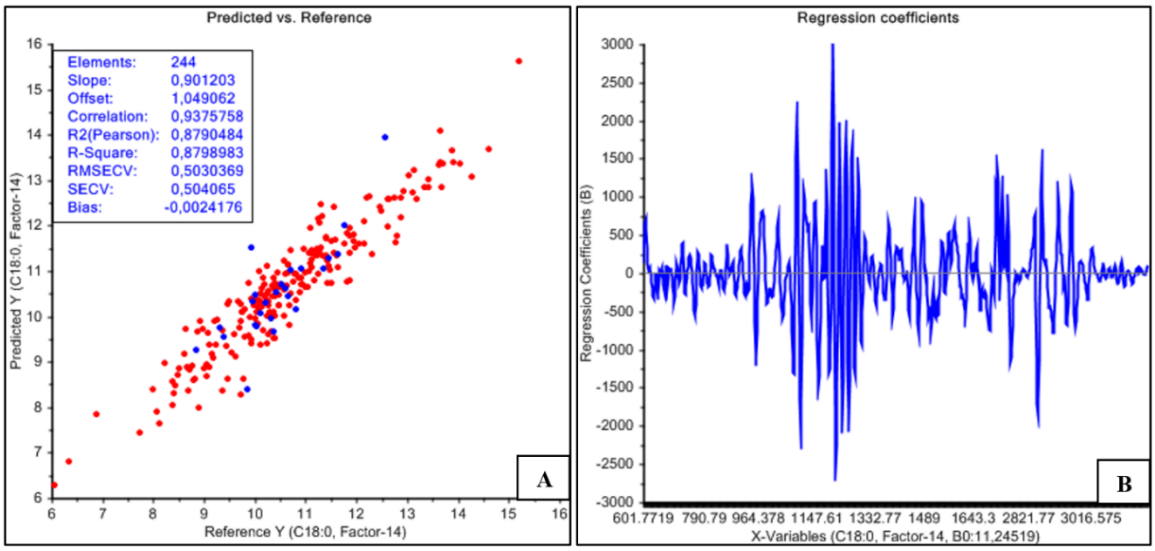


Figure 42. Predicted and reference response values (A) and regression coefficients (B) of the PLSR model obtained for C18:0. The model was based on 219 samples from Afseth et al. (2010) (red dots) and 25 samples from the current data set (blue dots).

The plot in Figure 42A reveals a good correlation between reference and predicted values. The blue points, which represent the new reference samples, are close to the old samples. The most important peaks in the plot of the regression coefficients in the prediction model for C18:0 (Figure 42B) were found at wavenumbers 947 cm^{-1}, 1074 cm^{-1}, 1242 cm^{-1}, 1280 cm^{-1}, and 2852 cm^{-1}.

The plot of predicted and reference response values obtained for the C18:1*cis*-9 prediction model is provided in Figure 43A. The regression coefficients are provided in Figure 43B.

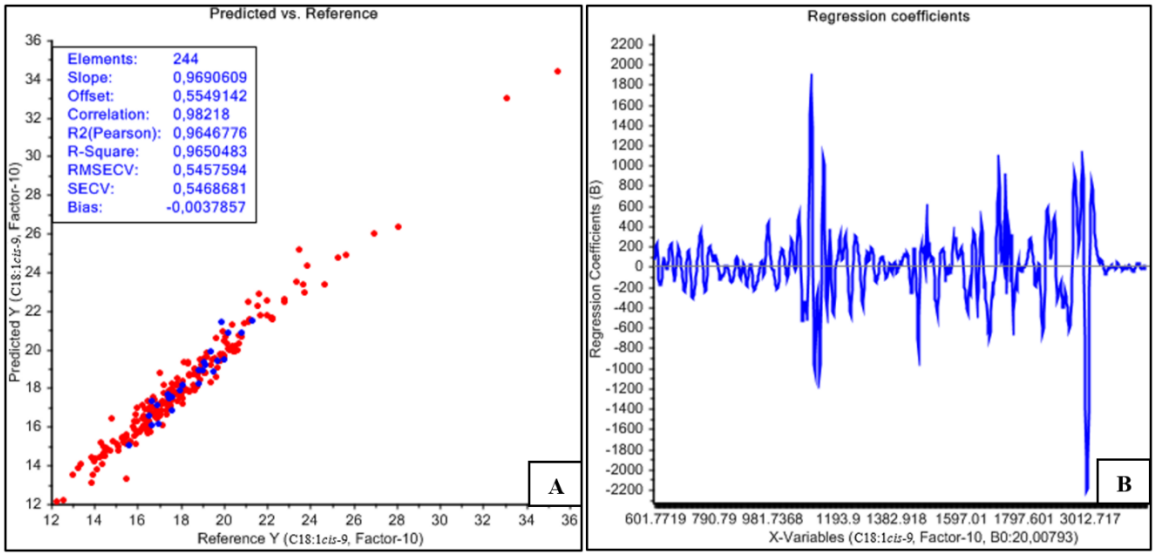


Figure 43. Predicted and reference response values (A) and regression coefficients (B) of the PLSR model obtained for C18:1cis-9. The model was based on 219 samples from Afseth et al. (2010) (red dots) and 25 samples from the current data set (blue dots).

The plot in Figure 43A reveals a good correlation between reference and predicted values. The blue points, which represent the new reference samples, are close to the old samples. The most important peak in the plot of the regression coefficients in the prediction model for C18:1*cis*-9 (Figure 43B) was found at wavenumber 3006 cm^{-1}.

The plot of predicted and reference response values obtained for the CLA prediction model is provided in Figure 44A. The regression coefficients are provided in Figure 44B.

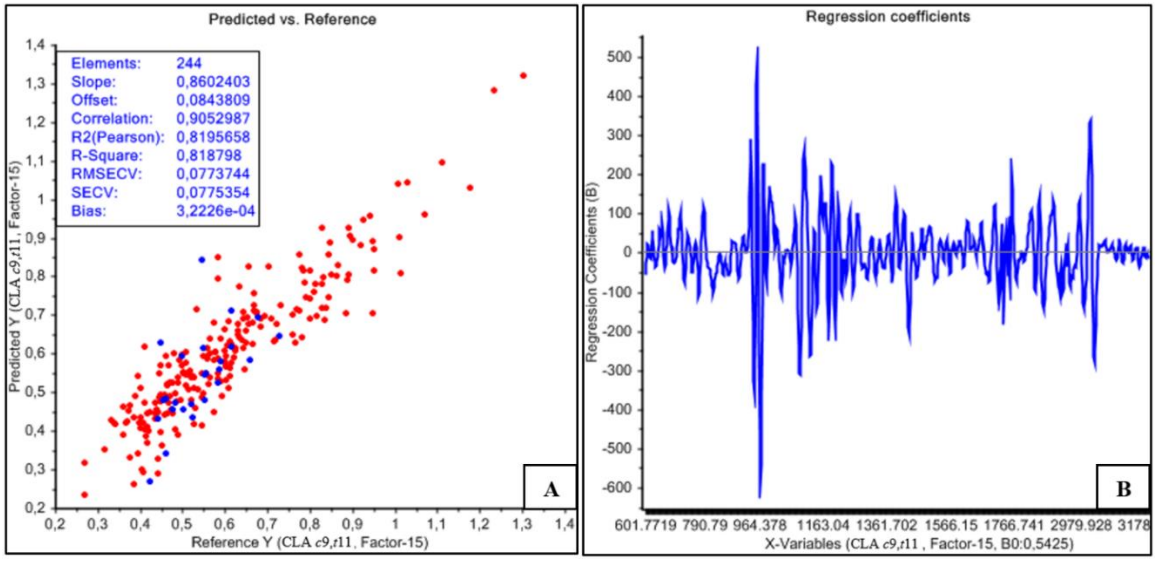


Figure 44. Predicted and reference response values (A) and regression coefficients (B) of the PLSR model obtained for conjugated linoleic acid (CLA). The model was based on 219 samples from Afseth et al. (2010) (red dots) and 25 samples from the current data set (blue dots).

The plot in Figure 44A reveals a moderate correlation between reference and predicted values. The blue points, which represent the new reference samples, are close to the old samples. The most important peaks in the plot of the regression coefficients in the prediction model for CLA (Figure 44B) were found at wavenumbers 947 cm^{-1}, 966 cm^{-1} and 3010 cm^{-1}.

The plot of predicted and reference response values obtained for the SFA prediction model is provided in Figure 45A. The regression coefficients are provided in Figure 45B.

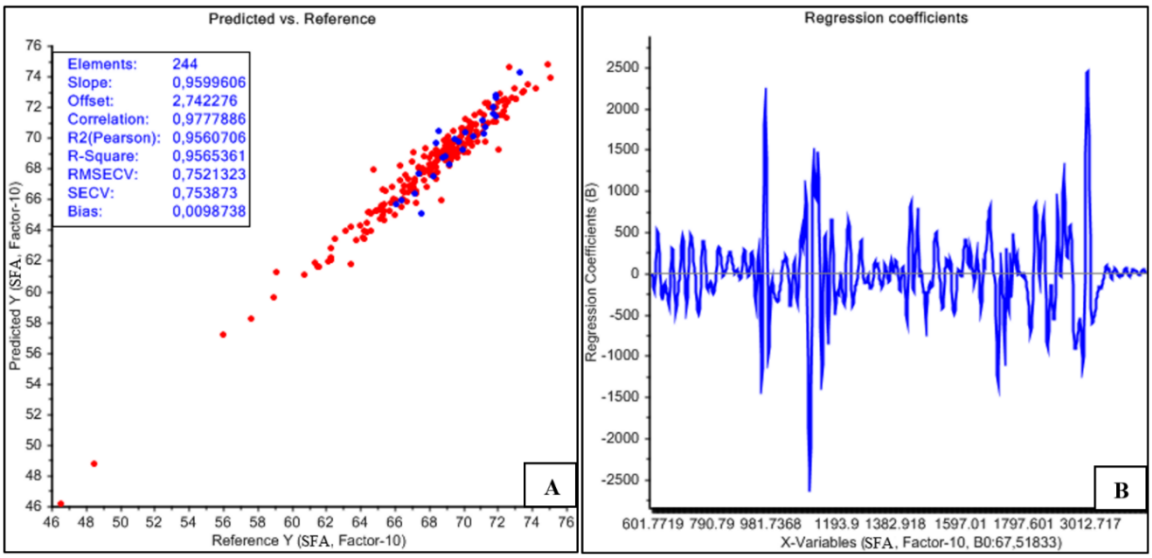


Figure 45. Predicted and reference response values (A) and regression coefficients (B) of the PLSR model obtained for the sum of saturated fatty acids (SFA). The model was based on 219 samples from Afseth et al. (2010) (red dots) and 25 samples from the current data set (blue dots).

The plot in Figure 45A reveals a good correlation between reference and predicted values. The blue points, which represent the new reference samples, are close to the old samples. The most important peaks in the plot of the regression coefficients in the prediction model for the sum of SFAs (Figure 45B) were found at wavenumbers 966 cm^{-1}, 1095 cm^{-1} and 3006 cm^{-1}.

The plot of predicted and reference response values obtained for the MUFA prediction model is provided in Figure 46A. The regression coefficients are provided in Figure 46B.

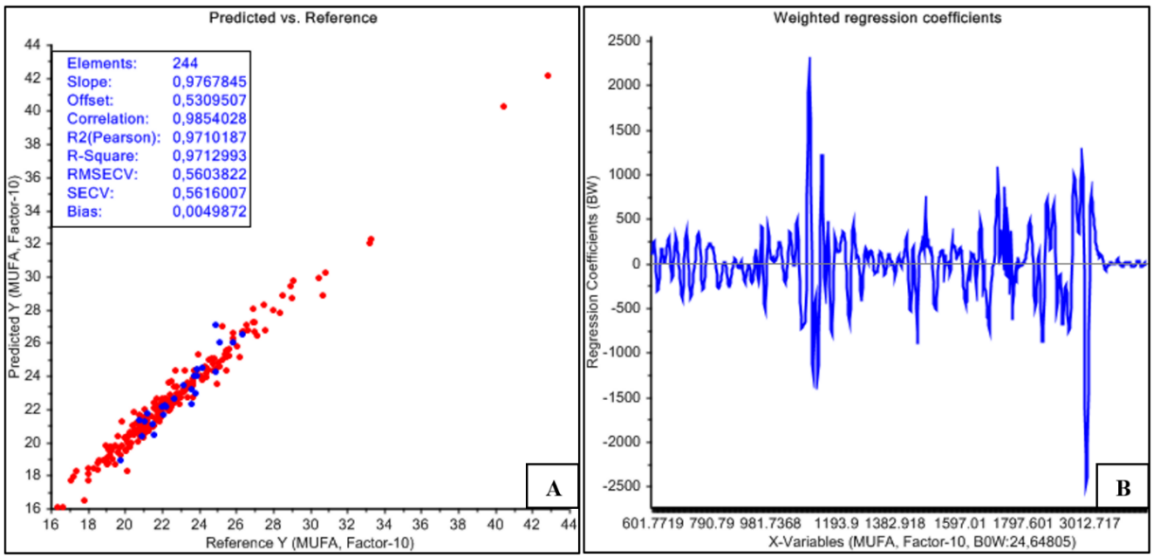


Figure 46. Predicted and reference response values (A) and regression coefficients (B) of the PLSR model obtained for the sum of monounsaturated fatty acids (MUFA). The model was based on 219 samples from Afseth et al. (2010) (red dots) and 25 samples from the current data set (blue dots).

The plot in Figure 46A reveals a good correlation between reference and predicted values. The blue points, which represent the new reference samples, are close to the old samples. The most important peak in the plot of the regression coefficients in the prediction model for the sum of MUFAs (Figure 46B) was found at wavenumber 3006 cm^{-1}.

The plot of predicted and reference response values obtained for the PUFA prediction model is provided in Figure 47A. The regression coefficients are provided in Figure 47B.

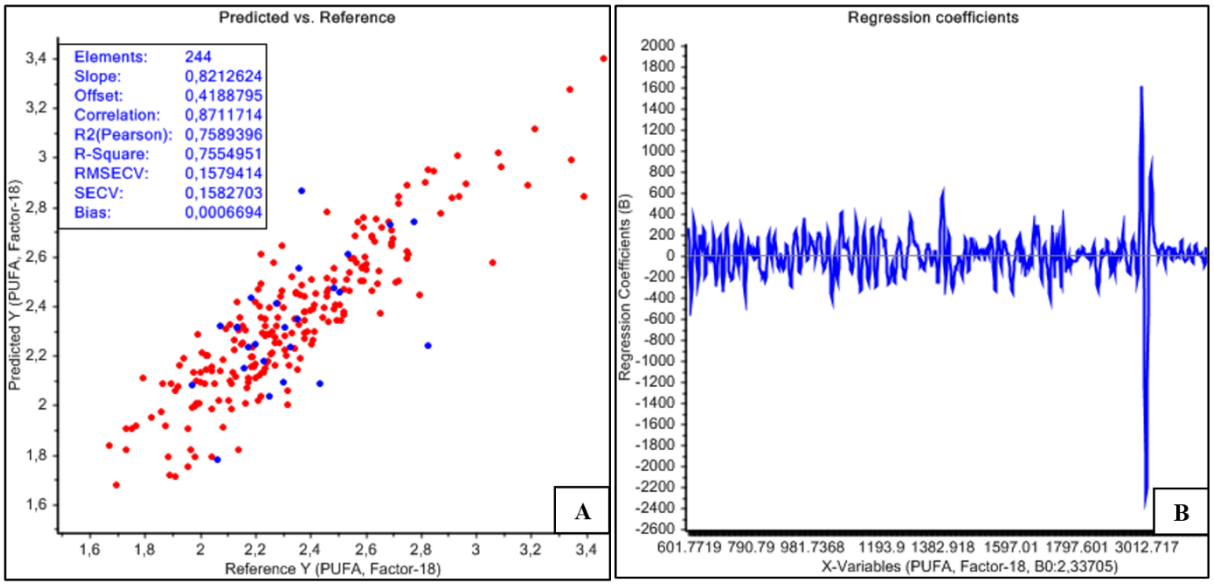


Figure 47. Predicted and reference response values (A) and regression coefficients (B) of the PLSR model obtained for the sum of polyunsaturated fatty acids (PUFA). The model was based on 219 samples from Afseth et al. (2010) (red dots) and 25 samples from the current data set (blue dots).

The plot in Figure 47A reveals a moderate correlation between reference and predicted values. The blue points, which represent the new reference samples, are close to the old samples. The most important peak in the plot of the regression coefficients in the prediction model for the sum of PUFAs (Figure 47B) was found at wavenumber 3012 cm^{-1}.

3.2.2. Fatty acid composition of milk

Since the results of the calibration models were moderately good to good (Chapter 3.2.1), the models were used for prediction of the FA composition of the 2333 milk samples. Prediction of the proportion of C10:0, C14:0, C16:0, C18:0, C18:1*cis*-9, CLA, and sums of SFAs, MUFAs and PUFAs in the 2333 milk samples was performed using the optimal number of factors found for each PLSR model (Table 4, Chapter 3.2.1). The minimum, maximum, mean and SD of the predicted FA content is provided in Table 5. The statistics are based on 2333 samples from 61 cows.

Table 5. Minimum, maximum, mean and standard deviation (SD) of the predicted fatty acid content in 2333 milk samples from 61 cows. All values are expressed as percentage of total weight of fatty acids.[1]

Fatty acid	Min	Max	Mean	SD
C10:0	0.5	6.1	3.1	0.7
C14:0	4.1	23.0	11.1	1.6
C16:0	16.0	43.5	29.7	3.1
C18:0	3.7	22.1	11.5	1.9
C18:1*cis*-9	7.1	39.4	20.3	3.6
CLA	-0.1	1.1	0.5	0.1
SFA	46.4	76.8	67.5	4.0
MUFA	14.9	46.2	24.9	3.9
PUFA	1.4	4.9	2.3	0.3

[1] Abbreviations: Conjugated linoleic acid (CLA), sum of saturated fatty acids (SFA), sum of monounsaturated fatty acids (MUFA) and sum of polyunsaturated fatty acids (PUFA).

3.2.3. Fatty acid composition during the 100 first days in milk

The predicted FA content was plotted against DIM to explore if there were any trends during the lactation. Plots were created both for the three individual cows and for the average content for all cows. The individual cows were the same as in previous chapters (with ID numbers 5977, 6238 and 6290). The number of samples per DIM used when calculating the average content for all cows is provided in Appendix C.

The predicted C10:0 content during the first 100 DIM for the three individual cows is plotted in Figure 48A. The average content for all cows is plotted in Figure 48B.

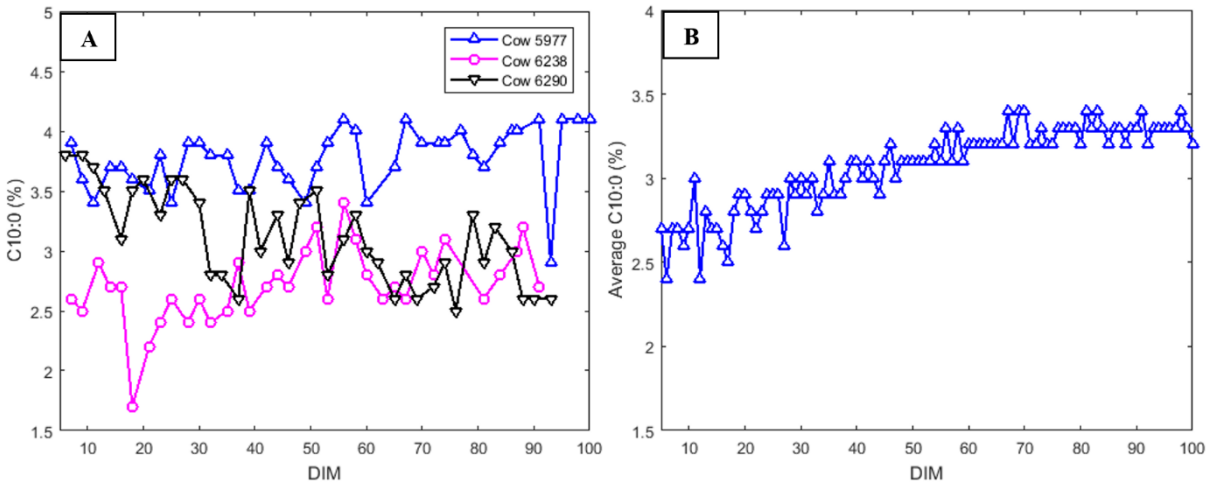


Figure 48. C10:0 content (percentage of total weight of fatty acids) during the first 100 days in milk (DIM) for three randomly selected cows (A) and the average content of samples from 61 cows (B).

The proportion of C10:0 was lowest in the beginning of the lactation and increased during the first 70 DIM (Figure 48B). There was larger day-to-day variation for the individual cows compared with the average content. In Figure 48A, the content of C10:0 seems to increase a bit for two of the cows and decrease for the cow with ID number 6290.

The predicted C14:0 content during the first 100 DIM for the three individual cows is plotted in Figure 49A. The average content for all cows is plotted in Figure 49B.

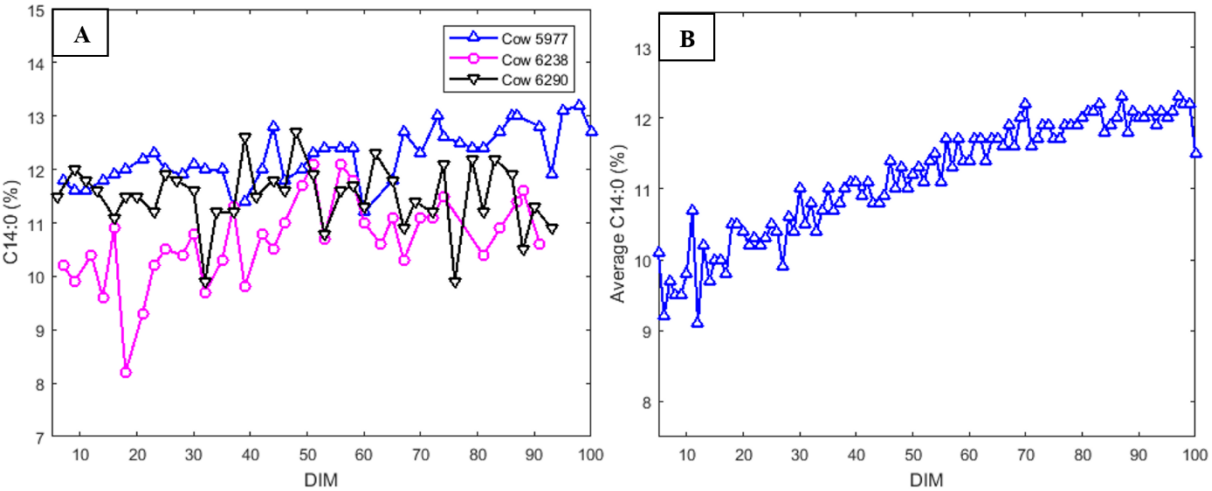


Figure 49. C14:0 content (percentage of total weight of fatty acids) during the first 100 days in milk (DIM) for three randomly selected cows (A) and the average content of samples from 61 cows (B).

The trend for C14:0 was similar to the observed trend for C10:0. It increased during the first 80 DIM (Figure 49B). For two of the individual cows, the content seemed to increase (Figure 49A). For the cow with ID number 6290 it seemed to be stable, except for the large day-to-day variation.

The predicted C16:0 content during the first 100 DIM for the three individual cows is plotted in Figure 50A. The average content for all cows is plotted in Figure 50B.

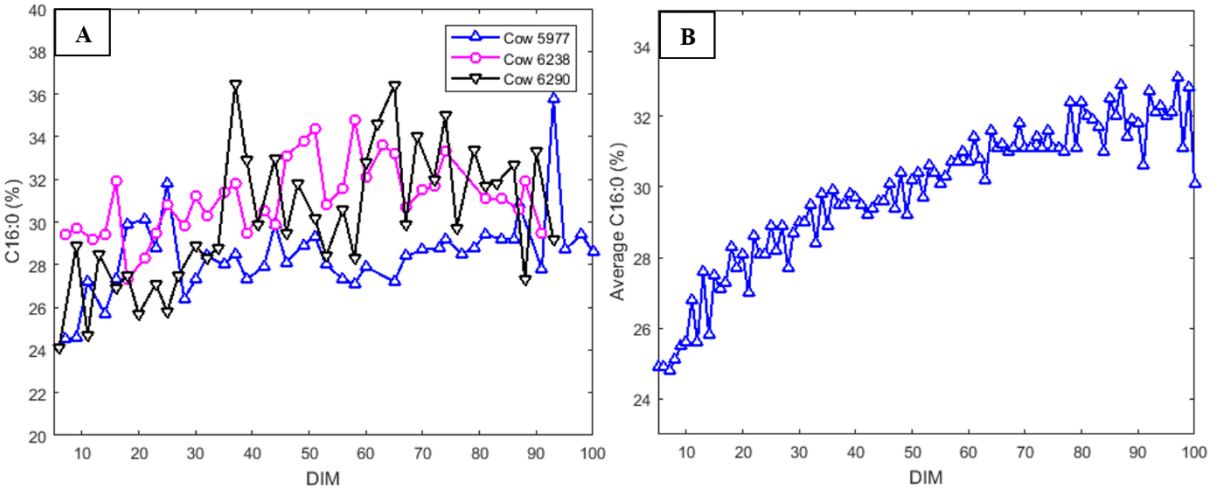


Figure 50. C16:0 content (percentage of total weight of fatty acids) during the first 100 days in milk (DIM) for three randomly selected cows (A) and the average content of samples from 61 cows (B).

Figure 50B illustrates that the average content of C16:0 increased from 25 % to 32 % during the first 80 DIM. The content also seemed to increase for the individual cows. However, the trends for the three cows are vague due to large variations from DIM to DIM.

The predicted C18:0 content during the first 100 DIM for the three individual cows is plotted in Figure 51A. The average content for all cows is plotted in Figure 51B.

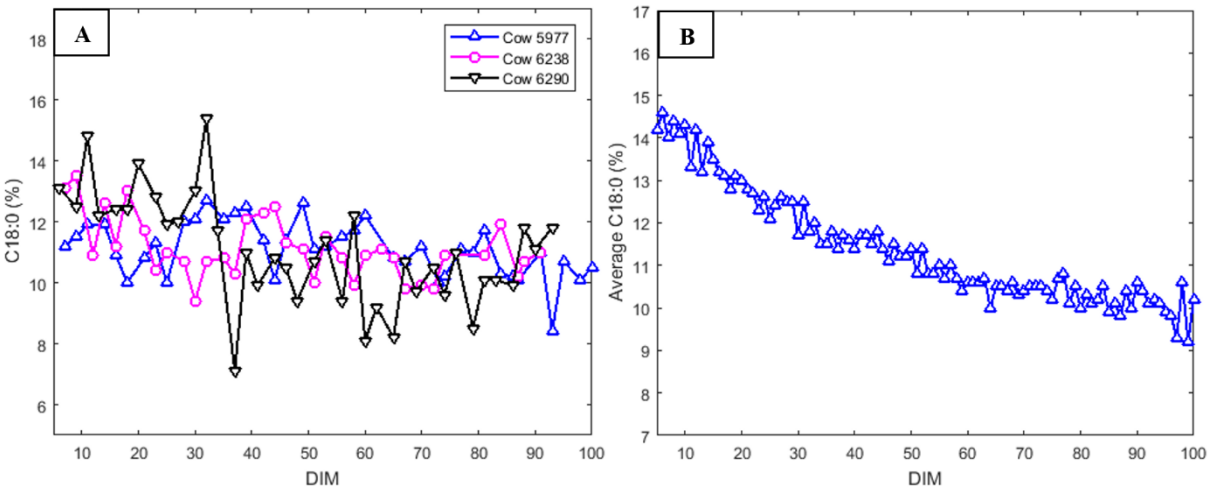


Figure 51. C18:0 content (percentage of total weight of fatty acids) during the first 100 days in milk (DIM) for three randomly selected cows (A) and the average content of samples from 61 cows (B).

The proportion of C18:0 decreased during the lactation (Figure 51). The trend is more distinct for the average content than for the individual cows.

The predicted C18:1*cis*-9 content during the first 100 DIM for the three individual cows is plotted in Figure 52A. The average content for all cows is plotted in Figure 52B.

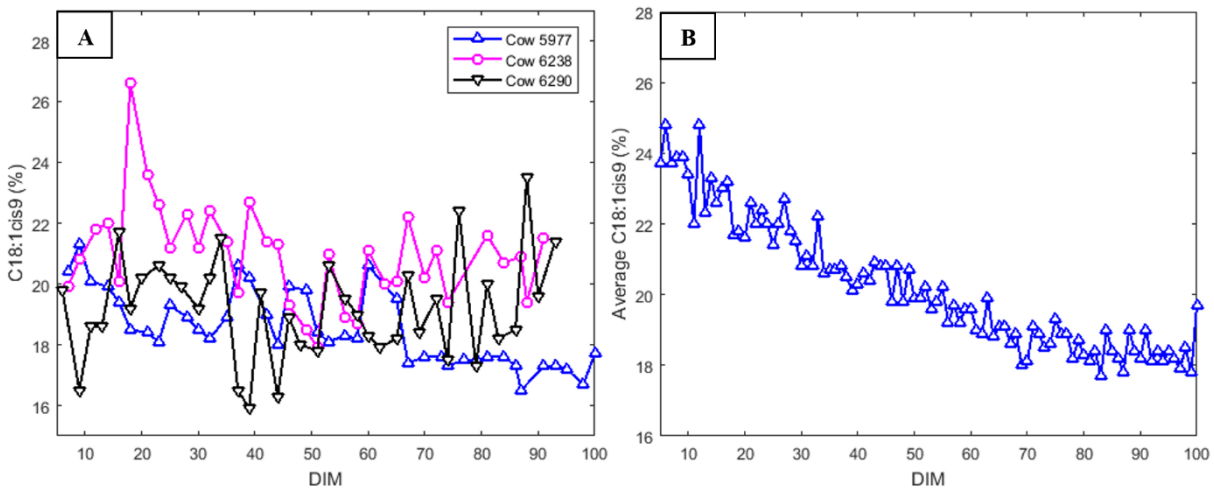


Figure 52. C18:1cis-9 content (percentage of total weight of fatty acids) during the first 100 days in milk (DIM) for three randomly selected cows (A) and the average content of samples from 61 cows (B).

In Figure 52B, the proportion of C18:1*cis*-9 decreased during the first 80 DIM. This trend was only visible for the cow with ID number 5977. There were no trends for the two other cows.

The predicted CLA content during the first 100 DIM for the three individual cows is plotted in Figure 53A. The average content for all cows is plotted in Figure 53B.

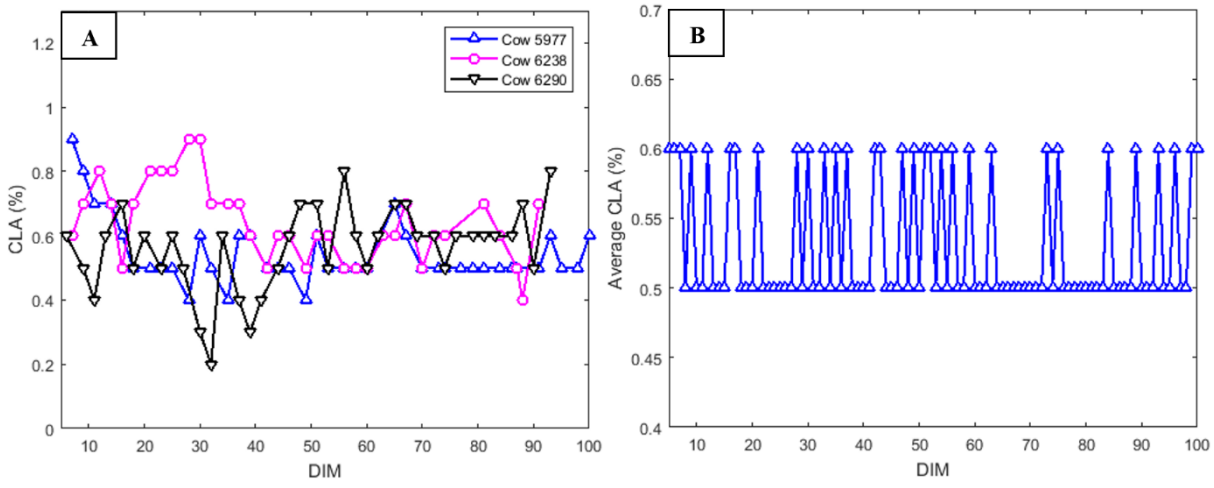


Figure 53. CLA (conjugated linoleic acid) content (percentage of total weight of fatty acids) during the first 100 days in milk (DIM) for three randomly selected cows (A) and the average content of samples from 61 cows (B).

There were no clear trends for the content of CLA during the lactation for the individual cows (Figure 53A). The average content of CLA was stable during the first 100 DIM and only varied between 0.5 % and 0.6 % (Figure 53B).

The predicted content of sum of SFAs during the first 100 DIM for the three individual cows is plotted in Figure 54A. The average content for all cows is plotted in Figure 54B. This group contain C10:0, C14:0, C16:0 and C18:0 (and the other SFAs in Table 1, Chapter 1.2.3).

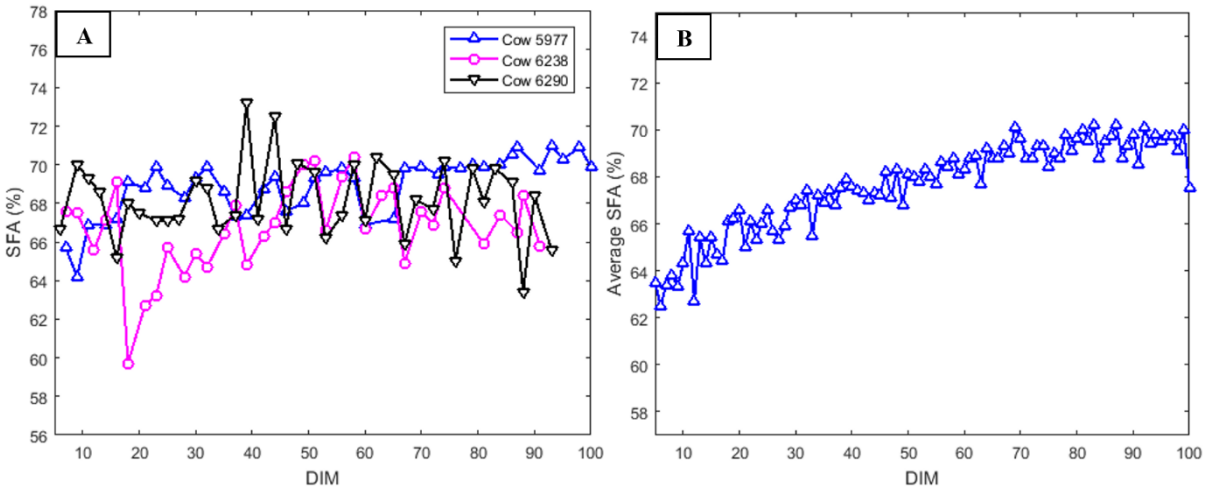


Figure 54. Sum of saturated fatty acids (SFA) content (percentage of total weight of fatty acids) during the first 100 days in milk (DIM) for three randomly selected cows (A) and the average content of samples from 61 cows (B).

The average proportion of the sum of SFAs increased during the lactation (Figure 54B). There were no distinct trends to observe for two of the individual cows. The content seemed to increase a bit for the cow with ID number 5977 (Figure 54A).

The predicted content of the sum of MUFAs during the first 100 DIM for the three individual cows is plotted in Figure 55A. The average content for all cows is plotted in Figure 55B. This group contain C18:1*cis*-9 (and the other MUFAs in Table 1, Chapter 1.2.3).

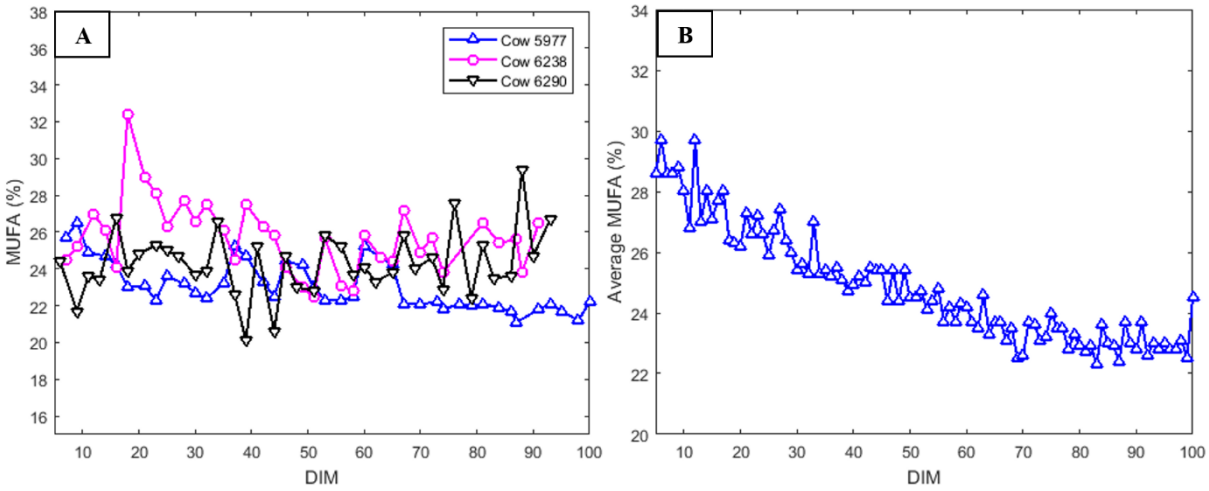


Figure 55. Sum of monounsaturated fatty acids (MUFA) content (percentage of total weight of fatty acids) during the first 100 days in milk (DIM) for three randomly selected cows (A) and the average content of samples from 61 cows (B).

From Figure 55A, there were no clear trends in the content of the sum of MUFAs during the lactation. In Figure 55B, a distinct trend can be observed. The proportion decreases during the first 80 DIM.

The predicted content of the sum of PUFAs during the first 100 DIM for the three individual cows is plotted in Figure 56A. The average content for all cows is plotted in Figure 56B. This group contain CLA (and the other PUFAs in Table 1, Chapter 1.2.3).

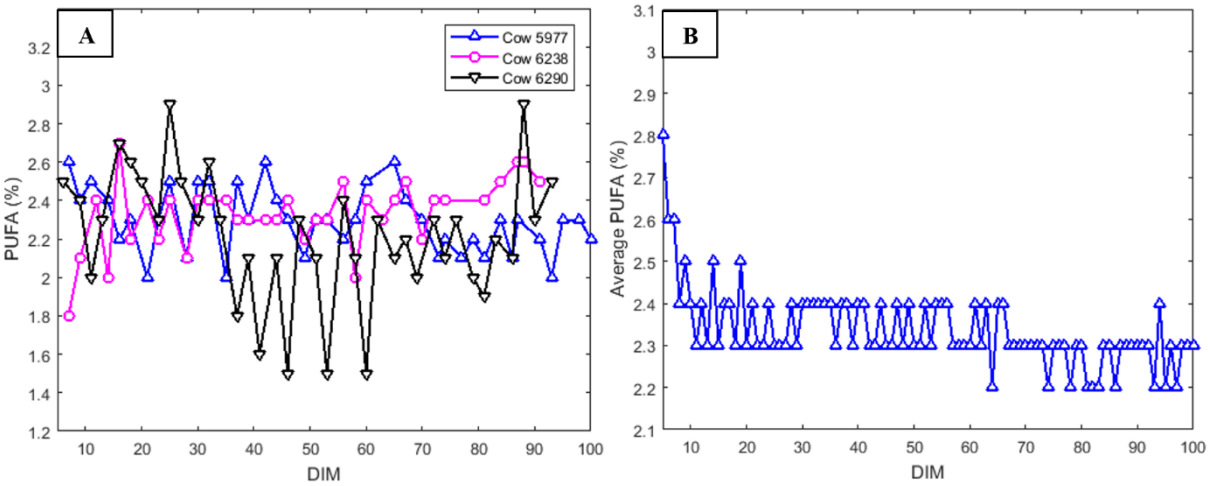


Figure 56. Sum of polyunsaturated fatty acids (PUFA) content (percentage of total weight of fatty acids) during the first 100 days in milk (DIM) for three randomly selected cows (A) and the average content of samples from 61 cows (B).

The average content of the sum of PUFAs was highest in the beginning of the lactation (Figure 56B). From DIM 5 to DIM 10, the content decreased from 2.8 % to 2.4 %. From DIM

10 to DIM 100, it seemed to decrease slightly. There were no clear trends for the individual cows (Figure 56A).

3.2.3.1. Principal component analysis

To sum up how the predicted FA composition changed during the lactation, a PCA of the proportions of the FAs for all samples was performed. The average composition of samples from all cows at DIM 5 to DIM 100, respectively, was used. The number of samples per DIM used when calculating the average content for all cows is provided in Appendix C. The result of the PCA is presented in Figure 57.

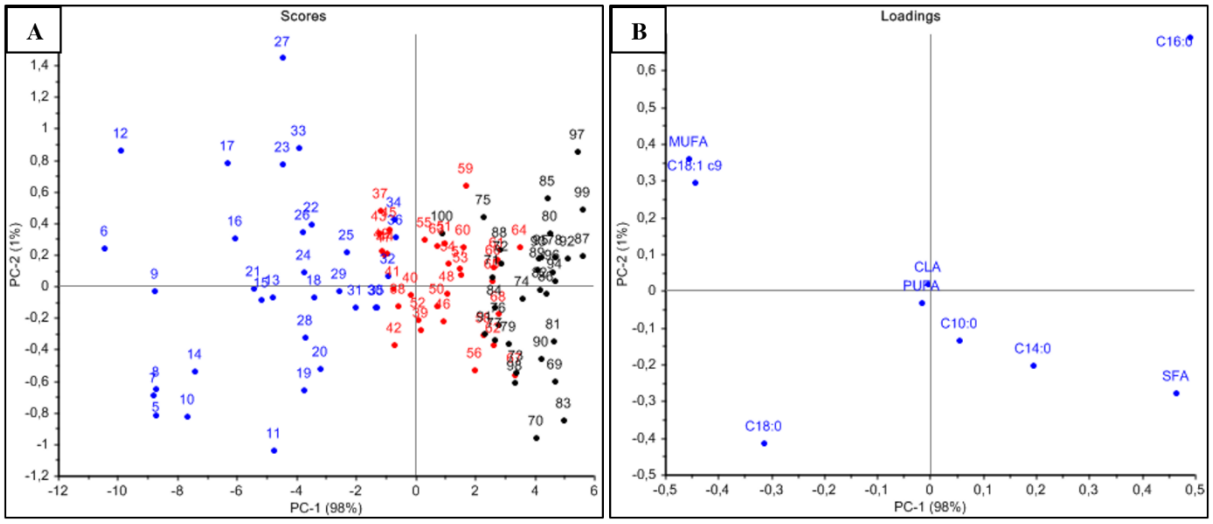


Figure 57. Score (A) and loading plot (B) from a PCA of the average proportion of fatty acids predicted for all milk samples at different days in milk (DIM). This PCA includes milk samples from DIM 5 to DIM 100.

The score plot in Figure 57A illustrates a distinct trend for the samples. The samples from DIM 5 to DIM 29 are located to the left (negative PC1). Most of the samples from DIM 30 to 60 are located in the middle and most of the samples from about DIM 61 to DIM 100 are located to the right. PC1 explains 98 % of the variation in this data set. PC2 only explains 1 %. The loadings in Figure 57B shows how the FAs relate to the each other and the samples. Sum of MUFAs and C18:1*cis*-9 are located to the left and have the largest negative PC1 values. This means that samples with a negative PC1 contains a high proportion of these two FAs. C16:0 and the sum of SFAs are located to the right in the loading plot. Therefore, samples located to the right in the score plot will contain a higher proportion of these FAs than the samples located to the left. This is consistent with what was found previously in this chapter. Figure 50B, for instance, illustrated that the average content of C16:0 was higher at DIM 100 than at DIM 5.

3.2.4. Fatty acid composition of milk from cows with subclinical ketosis

The level of betahydroxybuturate in blood samples from the cows was measured at ProdMed in order to detect cows with subclinical ketosis. To investigate if the FTIR spectra could reveal any differences between cows with subclinical ketosis and normal cows, the predicted FAs were plotted over time for the two groups. The average content at DIM 5 to 100 of samples from the two groups was used. The number of samples per DIM used when calculating the average content for the two groups is provided in Appendix C. There were 41 normal cows and 20 cows with subclinical ketosis (Appendix B). The average contents of C14:0, C16:0, C18:0 and C18:1*cis*-9 for the two groups are plotted in Figure 58A, B, C and D, respectively.

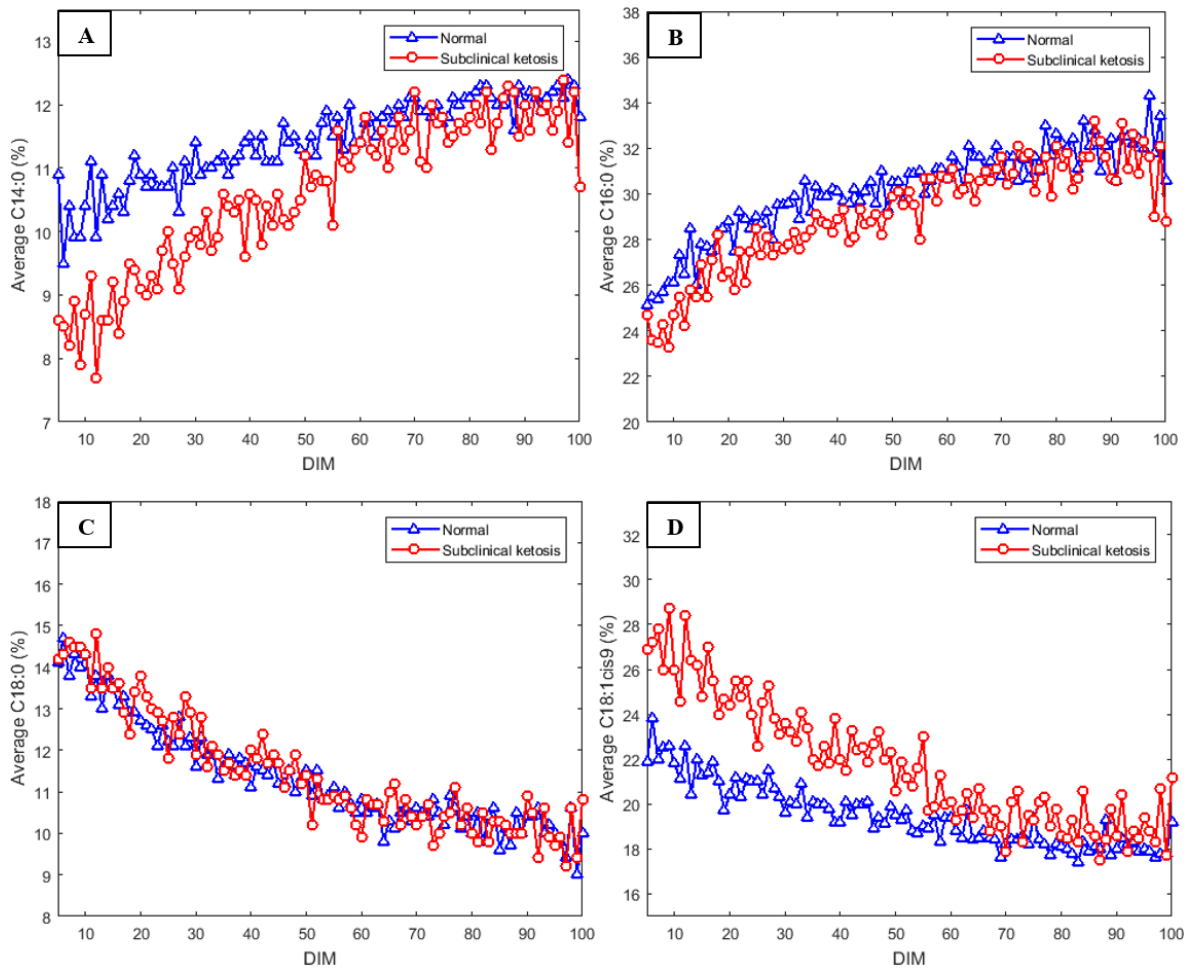


Figure 58. Average content of C14:0 (A), C16:0 (B), C18:0 (C), C18:1cis-9 (D) in samples from cows with subclinical ketosis and normal cows. The percentage of total weight of fatty acids is plotted during the 100 first days in milk (DIM).

There were clear differences between the two groups in the average content of C14:0 and C18:1*cis*-9 during the lactation (Figure 58A and D). The milk from cows with subclinical

ketosis had on average a lower proportion of C16:0 than the milk from normal cows (Figure 58B). The proportion of C18:0 was similar for both groups (Figure 58C).

The plots for C10:0 and the sums of SFAs, MUFAs and PUFAs are not shown. The plots of the average content of C10:0 and the sum of SFAs for the two groups were similar to the plot for C14:0 (Figure 58A). The plot for the sum of MUFAs was similar to the plot for C18:1*cis*-9 (Figure 58D). The plot for the sum of PUFAs did not show any differences between the two groups.

The plot for CLA (Figure 59) did not show any distinct differences between the two groups. The content only varied between 0.5 % and 0.6 % (except at DIM 20). However, it seems like the normal cows on average had a bit lower proportion of CLA between DIM 8 and 27.

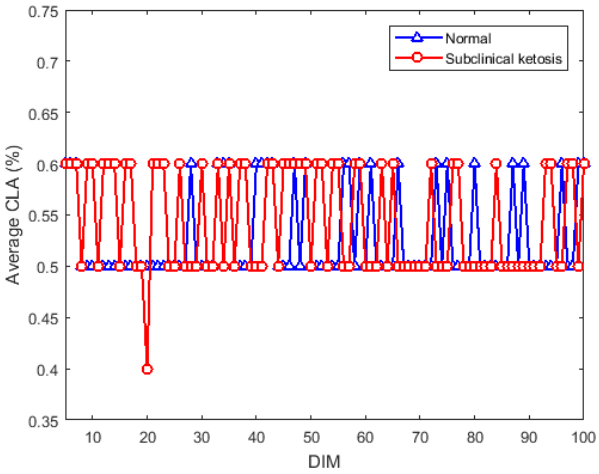


Figure 59. Average content of CLA in samples from cows with subclinical ketosis and normal cows. The percentage by total weight of fatty acids is plotted during the 100 first days in milk (DIM).

4. Discussion

A cheap, rapid and accurate method for determination of fatty acid (FA) composition of milk could be advantageous for monitoring of cows' health and fertility. The FA composition is, for instance, associated with energy balance (EB), subclinical ketosis and onset of luteal activity (OLA) (Gross et al. 2011; Martin et al. 2015; Van Haelst et al. 2008). The main objective of this work was to analyse milk from cows during their first 100 days in milk (DIM) using dry-film FTIR, and to investigate if the resulting FTIR spectra were suitable for detection and quantification of milk components by using PCA and PLSR.

4.1. Explorative analysis

4.1.1. Chemical composition of milk

Since milk consists of many different components such as fat, protein, lactose, minerals, vitamins etc., a dry-film FTIR spectrum of a milk sample will be complex. Therefore, it was very relevant to explore and describe the chemical composition (determined at TINE) of the milk samples before analysing the spectra. The minimum, maximum, mean and standard deviation (SD) of the chemical components was calculated. These statistics were based on 2143 samples because samples with missing or unlikely values were removed. The mean values of fat, protein, lactose, urea and FFAs were 4.0 %, 3.4 %, 4.9 %, 5.1 mmol/L and 0.5 mmol/L, respectively. These results were in agreement with the composition of milk described in the literature (Geno 2014c; Ottinsen 2013; TINE SA 2017). There were large variations between the samples in the data set. There were largest variations in fat content, which ranged from 2.0 % to 8.0% after removing values below or above these limits. This was expected, as fat content is highly affected by diet, stage of lactation and milking interval (Linn 1988). The fat content may also have been over- or underestimated due to insufficient mixing of samples prior to analysis.

4.1.2. Chemical composition during the first 100 days in milk

To explore how the content of the chemical components varied, the content of each component was plotted from DIM 5 to DIM 100. This was done both for the three randomly selected cows and for the average of all cows. Due to the large day-to-day variation, the data for fat, protein and lactose, were smoothed with Legendre polynomials. The smoothed data were only used for visualization of trends, not for prediction of FAs. Regarding all chemical components, the large variations from day to day and between the cows made it difficult to

distinguish trends in the unsmoothed data for the individual cows. The smoothed data, and the average data for all cows, made it easier to explore trends.

From the smoothed data for the individual cows and the average of all cows, the fat content decreased during the first 30 DIM. Between DIM 30 and 50, it seemed to be stable and thereafter it increased. This trend is partly consistent with what has been described in the literature. According to Linn (1988), the content decreases during the first 60 DIM and then it increases. Gross et al. (2011) observed that the fat content decreased in the interval from calving to about 40 DIM. Thereafter it increased slightly. It was difficult to recognise this trend for the unsmoothed data, both for the individual cows and for the average of all cows. The large variations in fat content between the DIM could be due to sampling challenges. Insufficient mixing of the samples before splitting or analysing a sample may over or underestimate the fat content (due to separated fat and water phases). Other explanations are natural variations in the cow during the day and in the feed intake (Linn 1988), and varying number of milking times per day in an AMS. This leads to varying milking intervals, which may result in variations in fat content (Lindstrøm 2017).

The trend for protein content during the first 100 DIM was easier to detect than the trend for fat content. For the averaged unsmoothed data and for the smoothed data (both for the individual cows and the average), the protein content was highest immediately after calving. It seemed to reach a minimum at about DIM 30−40 and thereafter it increased slightly. This is in agreement with what has been described in previous studies (Gross et al. 2011; Linn 1988; Tsioulpas et al. 2007). This trend was observed to some extent for the unsmoothed data for two of the individual cows. There were variations from DIM to DIM for all individual cows. However, the fluctuations were not as large as for the fat content. The protein content may be affected by diet and milking interval, but not as much as the fat content (Linn 1988).

The maximum lactose content during the 100 DIM seemed to be between DIM 25 and DIM 40. This is consistent with the results of Ptak et al. (2012). Gross et al. (2011) observed that the lactose content increased a bit during the first 40 DIM and that it thereafter was relatively stable. The variations of lactose content in the unsmoothed data for the individual cows were smaller than the variations for fat and protein. This was expected as lactose content has been described in the literature to vary the least of the chemical components (Linn 1988).

For FFAs and urea, only the unsmoothed data was available. The average content of FFAs for all cows decreased during the first 10 DIM. This trend was not recognised in the plots for the individual cows. The observed trend is not consistent with what was found in the literature. It is known that cows in late lactation are more likely to produce milk with increased levels of FFAs than cows in early lactation (Deeth & Fitz-Gerald 1995). The high average value in the beginning of the lactation may be due to a very high concentration in a few samples. It could also be due to uncertainties in the measurements due to the naturally narrow range of variation in FFA content. Reasons for increased levels of FFA content may be poor feeding regime, microbial contamination or agitation of the milk sample (Deeth & Fitz-Gerald 1995). The urea content was lowest in the beginning of the lactation for two of the individual cows and for the average of all cows. This is in agreement with the literature (Geno 2014c).

Since distinct trends were observed (for the average of all cows) for fat, protein, lactose, FFAs and urea during the lactation, it was expected to find a trend in the averaged FTIR spectra. For the individual cows, it was expected that no clear trends in the FTIR spectra could be revealed due to the large day-to-day variation.

4.1.3. The Fourier transform infrared spectra

4.1.3.1. Spectra and outlier detection

From the raw spectra, only five of 6999 replicate spectra were removed. This removal was due to strong water vapour signals in the FTIR spectra. The small number of removed spectra is promising for future analyses of milk by dry-film FTIR. To correct for baseline variations in the raw dry-film FTIR absorbance spectra, the second derivative was calculated with the Savitzky-Golay algorithm. FTIR spectra of milk will, as already mentioned, be very complex and contain several overlapping absorption bands. When calculating the second derivative of the spectra, overlapping peaks were resolved and more details were revealed (Kohler et al. 2010). It is not possible to create dried sample films with equal thicknesses. Therefore, analysis of films from three sample replicates may results in spectra with different absorbance. Normalization by EMSC was performed to correct for such multiplicative effects. As expected, the raw and pre-processed spectra showed peaks at wavenumbers characteristic for the functional groups found in fat, protein and lactose.

PCA of the pre-processed spectra resulted in a score plot that revealed that three replicates were different from the other samples. Since all three replicates were located very

close to each other, they were not considered outliers that should be removed. They probably represented a somewhat different chemical composition than the other samples, and they are therefore important to include when describing the composition of milk. PC1 and PC2 explained 67 % and 21 %, respectively, of the variance in the pre-processed data set. The loading plot illustrated that the wavenumbers around 2800 cm^{-1} and in the range from about 1800 cm^{-1} to 900 cm^{-1} were important for explanation of the variance. This was expected as the pre-processed spectra contained many peaks in these areas.

4.1.3.2. Spectra during the first 100 days in milk

It was investigated whether the observed trends in chemical composition during the lactation (Chapter 0) could be detected by performing a PCA of the FTIR spectra. A PCA of the average spectra from all cows in the periods DIM 5−100, DIM 5−50 and DIM 51−100 was performed. The results of the three analyses were similar. A clear trend for the spectra over time was observed in the score plots. With exception of some overlaps, the samples from early and late DIM were located on opposite sides in the score plots. The loading plots showed that the most important wavenumbers were between 1760 cm^{-1} and 1450 cm^{-1} and between 1200 cm^{-1} and 1000 cm^{-1}. The most important peaks were found at 1751 cm^{-1}, 1656 cm^{-1} and 1074 cm^{-1}. There was also a small band at 2852 cm^{-1}. As already mentioned, wavenumber 1656 cm^{-1} is characteristic for vibrations in proteins. This indicates that the variation in protein content or composition is one reason for the observed trend. This is consistent with what was found in Chapter 0. There was a distinct trend for average protein content, especially during the first 30 DIM. Absorption at wavenumber 1074 cm^{-1} probably corresponds to the stretching vibration of C−O. This might be the C−O bond in lactose or in the ester linkage in triglycerides. According to Lei et al. (2010), the C−O bond in fat has characteristic absorption at 1161 cm^{-1} while the C−O bonds in lactose absorb in the range from 1150 cm^{-1} to 1030 cm^{-1}. The band at 1751 cm^{-1} corresponds to C=O stretching vibration in fat. The small band found at 2852 cm^{-1} is associated with C−H bending in fat. From this, it seems like the contents of fat, protein and lactose are the reasons for the observed trends in the spectra. This is consistent with the results in Chapter 0. This was expected, as these are the three main components in milk. When plotting the averaged unsmoothed data, distinct trends were observed for protein and lactose. The unsmoothed data for fat followed a trend as well, but it was unclear do to large day-to-day variation.

It was also investigated whether there were any trends for the three individual cows (with ID numbers 5977, 6238 and 6290). No trends became visible in the score plots for two

of the cows. This was consistent with the observations for the chemical components in Chapter 0. This was probably due to large day-to-day variations. For the cow with ID number 6238, the samples from early DIM seemed to be located to the right, while most of the samples from middle and late DIM seemed to be located to the left. However, the trend was not as obvious as the trend for the average spectra for all cows. It was difficult to distinguish trends for this cow in Chapter 0. The loading plots for the cows with ID number 5977 and 6238 were similar to the loading plots for the average spectra. Important wavenumbers were 1745 cm^{-1}, 1656 cm^{-1}, 1074 cm^{-1} and 2923 cm^{-1}. This means that fat, protein and lactose were important for describing the variations in the milk samples from these cows. The loading plot for the cow with ID number 6290 differed from the others. The most important peak in the loading plot was found at wavenumber 2923 cm^{-1}. There was also an important peak found at 1737 cm^{-1}, which probably corresponds to C=O stretching in fat. This means that the fat content was most important for describing the variations in the milk samples from this cow. This was difficult to understand from the plots of fat, protein and lactose (Chapter 0). However, the content and variations of FFAs were largest for this cow.

Since the trends observed in the FTIR spectra were consistent with the trends for the chemical composition in Chapter 0, the spectra contained logical information about the milk samples. This was expected as these components were determined by use of liquid FTIR at TINE.

4.1.3.3. Feeding regime

Milk composition is influenced by factors such as breed, health and diet (Linn 1988). The cows in this experiment were grouped in four different feeding regimes (Appendix B). One part of the project was to examine if a PCA of the spectra could visualize any groupings of milk samples from cows with the same feeding regime. Spectra of all samples at DIM 70−71 were used in the PCA. These DIM were chosen because the results of Lindstrøm (2017) showed differences in EB between the groups at this time point. The score plot did not show any distinct groupings of samples from cows with the same feeding regime. It is possible that groupings had become visible if the average spectra for the four groups had been used in the PCA instead of spectra from the individual cows. Since the FTIR spectra contain information about the chemical composition and not the EB directly, it is possible that groupings could have been visible at other time points. However, from this limited investigation, it seems like other parameters affect the milk composition more than what

feeding regime does. This means that the observed trends are probably not due to the feeding regime only.

4.2. Prediction of fatty acid composition

4.2.1. Partial least squares regression models

PLSR models for prediction of FA composition in milk from dry-film FTIR spectra were established using the 219 calibration samples from Afseth et al. (2010) with 25 samples from the current data set included (244 samples in total). The optimal number of PLSR factors was determined using segmented cross-validation (20 segments).

Calibration models for prediction of the content of C10:0, C14:0, C16:0, C18:0, C18:1*cis*-9, CLA, and sums of SFAs, MUFAs and PUFAs in milk were established. The FAs included in the sums of SFAs, MUFAs and PUFAs are listed in Table 1, Chapter 1.2.3. The chosen FAs are interesting from a biological perspective. C14:0, C16:0, C18:1*cis*-9 have been reported to have a significant relationship with OLA (Martin et al. 2015). C18:1*cis*-9 and groups of FAs have been suggested as interesting parameters for monitoring of EB (Gross et al. 2011). Cows with subclinical ketosis have been reported to produce milk with a higher proportion of C18:1*cis*-9 than normal cows (Van Haelst et al. 2008). It was interesting to predict the proportion of C10:0 and C18:0 as they have different origins in the cow, and the proportion of these should follow different trends during the lactation. Prediction models for C6:0 and C8:0 were not established since these are present in a lower proportion than C10:0 in milk. In addition, these two FAs have the same origin as C10:0 and should therefore follow the same trend as C10:0. CLA is present in a very low concentration in milk. It was therefore expected that a poorer model would be obtained for CLA than for FAs present in higher concentrations.

In the plots of predicted vs. reference values for all established calibration models, the 25 samples from the current data set were located close to the 219 old samples. This means that the old and new samples are similar regarding FA composition. Consequently, results for all prediction models in this project were similar to the results obtained with only the 219 old samples (Afseth et al. 2010). The PCA of the 244 calibration samples did not show any outliers and most of the new samples were located near the old samples in the score plot.

RMSECV and R^2, together with visual inspection of the plots, indicated that moderately good to good models were obtained for all FAs. If R^2 was close to 1, a large proportion of the

variability in FA composition was explained by the FTIR spectra (James et al. 2013). The RMSECV expresses the average distance between a reference and a predicted value (Vernier Software & Technology 2001). The best prediction models were obtained for C10:0, C16:0, C18:1*cis*-9, sums of SFAs and MUFAs. The R^2 for these models were 0.93, 0.92, 0.97, 0.96 and 0.97, respectively. The optimal numbers of factors were 10, 11, 10, 10 and 10, respectively. Models for C14:0 and C18:0 resulted in R^2 of 0.85 and 0.88, respectively. The optimal numbers of factors were 9 and 14, respectively. The model with lowest R^2 (0.76) was obtained for the sum of PUFAs. This model and the model for CLA (R^2 of 0.82) had 18 and 15 as the optimal number of factors, respectively. The optimal number of factors indicates how complex a model is (Afseth et al. 2010).

It was expected to obtain calibration models with lower R^2 for the FAs present in very low concentration than for the FAs present in high concentration in milk. This was true for summed PUFA and CLA, which are present at about 2.3 % and 0.6 %, respectively, of total FAs in milk. However, a better model was obtained for C10:0 (R^2 of 0.93), which is only present at about 3.4 % of total FAs in milk.

If concentration was the only parameter affecting the prediction model, prediction of CLA should have resulted in the worst model since CLA is present in the lowest concentration (of the nine FAs predicted in the current experiment). This was not the case. The prediction model for the sum of PUFAs resulted in a lower R^2 than the model for CLA. This might be explained by the regression coefficients. Important wavenumbers observed for CLA and PUFAs were 3010 cm^{-1} and 3012 cm^{-1}, respectively, which correspond to the vibration of C−H in a *cis* double bond. The difference between these two wavenumbers may be due to the degree of unsaturation (Afseth et al. 2010). The plot of the regression coefficients for CLA revealed that also the wavenumber 966 cm^{-1} was important for this model. This wavenumber corresponds to the bending vibration of C−H in a double bond with *trans* configuration (Afseth et al. 2010). The peak at this wavenumber was expected since CLA contains a *trans* double bond. Regarding the sum of PUFAs, this group only included one FA with a *trans* double bond (CLA). It contains five FAs with between two and six *cis* double bonds (Table 1, Chapter 1.2.3). This might explain why the regression coefficient peak at 966 cm^{-1} is very small compared to the one at 3012 cm^{-1} in the model for the sum of PUFAs. The fact that the prediction model for CLA contained two important peaks in the plot of the regression coefficients could explain why a better prediction model was obtained for

CLA than for sum of PUFAs. This shows that both concentration and type of bonds in a component affect a prediction model.

The most important peak in the plot of the regression coefficients in the prediction model for the sum of MUFAs and in the model for C18:1*cis*-9 was at wavenumber 3006 cm^{-1}. This is the characteristic frequency of the stretching vibration of C−H in a *cis* double bond (Afseth et al. 2010). The sum of MUFAs included both FAs containing a *cis* or *trans* double bond. However, the FAs with *trans* double bond are present in a very low concentration in milk compared to C18:1*cis*-9. This might explain why there was not an important peak located at 966 cm^{-1} in the plot of the regression coefficients. The most important peaks in the plot of the regression coefficients in the prediction model for C18:0 were found at wavenumbers 1074 cm^{-1}, 1242 cm^{-1}, 1280 cm^{-1} and 2852 cm^{-1}. The first three listed wavenumbers are probably related to C−O stretching (Williams & Fleming 2008), while 2852 cm^{-1} is related to C−H stretching. Wavenumber 2852 cm^{-1} could be related to the length of the carbon chain. The most important peaks in the plot of the regression coefficients in the prediction model for C16:0 were found at wavenumber 966 cm^{-1}, 1172 cm^{-1} and 1261 cm^{-1}. Wavenumbers 1172 cm^{-1} and 1261 cm^{-1} are probably related to vibration of the C−O bonds. C16:0 does not contain a *trans* double bond, so the peak at 966 cm^{-1} was not expected. However, when determining the proportion of one particular FA, information about the proportions of the other FAs are important as well. Since the peak was positive, wavenumber 966 cm^{-1} might be important for describing the relationship between the proportions of C16:0 and FAs with *trans* double bonds. In a similar manner, one of the most important peaks in the prediction model for C14:0 was at wavenumber 3006 cm^{-1} (positive coefficients). Since C14:0 does not contain any *cis* double bonds, wavenumber 3006 cm^{-1} might be important for describing the relationship between the proportion of C14:0 and FAs with *cis* double bonds. This model also had a large peak at wavenumber 1163 cm^{-1}. This wavenumber is probably related to the stretching vibration of C−O. The model for the sum of SFAs had the most important peaks at wavenumbers 966 cm^{-1}, 1095 cm^{-1} and 3006 cm^{-1} in the plot of the regression coefficients. The coefficients at 966 cm^{-1} and 3006 cm^{-1} were positive, and can be explained in the same manner as for C14:0 and C16:0. This suggests that a sample containing a large proportion of C14:0, C16:0 or SFAs will contain a smaller proportion of FAs with *cis* and *trans* double bonds (i.e. MUFAs, PUFAs or C18:1*cis*-9). The important peaks in the plot of the regression coefficients for C10:0 were found at 1089 cm^{-1}, 1103 cm^{-1} and 1153 cm^{-1}. These wavenumbers probably correspond to C−O stretching vibrations.

The plots of the regression coefficients in all models contained peaks at wavenumbers that represent vibrations of functional groups found in the respective FAs. This indicates that the models are based on actual chemical properties of the molecules and not random correlations in the data set (Afseth et al. 2010).

4.2.2. Fatty acid composition of milk

Since the results of the calibration models were moderately good to good (Chapter 3.2.1), the models were used for prediction of the FA composition of the 2333 milk samples. Prediction of the FA composition (percentage of total weight of FAs) was performed using the optimal number of factors found for each PLSR model. The predicted mean values for C10:0, C14:0, C16:0, C18:0, C18:1*cis*-9, CLA, and sums of SFAs, MUFAs and PUFAs were 3.1 %, 11.1 %, 29.7 %, 11.5 %, 20.3 %, 0.5 %, 67.5 %, 24.9 % and 2.3 %, respectively. These values were similar to mean values found in the literature (Afseth et al. 2010; Månsson 2008). Some differences were expected since the mean values of the reference samples are based on 219 samples while the mean values of the predicted contents are based on 2333 samples. The minimum and maximum values of the predicted values were also a bit different from the ranges for the 219 samples. This is probably due to more natural variations among 2333 samples than 219 samples. The most unlikely value from the predictions was the minimum value of −0.1 % for CLA content. The fact that the PLSR model for CLA was only moderately good, could be the explanation of this obviously incorrect value.

4.2.3. Fatty acid composition during the 100 first days in milk

FA composition is known to be influenced by diet, EB and lactation stage (Gross et al. 2011). To explore how the predicted FA composition in milk varied with lactation stage, the FA contents were plotted for the first 100 DIM for the three randomly selected cows (with ID numbers 5977, 6238 and 6290). Only vague trends were observed for the individual cows due to day-to-day variations and variation between the cows. These variations could be due to feeding and variations in the EB of the cow. According to Wiking et al. (2006), the FA composition is not altered by varying milking frequencies. Varying temperature of the samples, when diluting and transferring them to the plate, is probably not a reason for the day-to-day variations. Varying sample temperatures may cause variations in total fat content, but not in the FA composition. Lipolysis (Chapter 1.2.3) of the milk fat may have occurred in samples with long drying times of the sample films or long time from milking to analysis. However, lipolysis might only have had an impact on absorption at wavenumber 1745 cm^{-1}. Vibration of C=O in an ester (in a triglyceride) will absorb at wavenumber 1745 cm^{-1} while

vibration of C=O in a carboxylic acid (in an FFA) will absorb at about wavenumber 1720 cm^{-1}. Therefore, lipolysis probably did not have large influence on the results.

Trends became clearer when plotting the average FA contents for all cows. The proportion of C10:0 was lowest in the beginning of the lactation and increased from about 2.5 % to 3.3 % during the first 70 DIM. The proportion of C14:0 and C16:0 increased from 9.5 % to 12 % and from 25 % to 32 %, respectively, during the first 80 DIM. The proportion of C18:0 and C18:1*cis*-9 decreased from 14.5 % to 9.5 % and from 25 % to 18 %, respectively, during first 80 DIM. These results are consistent with previously described trends for FAs during the lactation. Immediately after calving, body fat mobilization will occur due to negative EB. The fat mobilization results in a larger amount of long-chained FAs in the milk. The large amount of long-chained FAs in the milk inhibits the synthesis of short- and medium-chained FAs in the mammary gland. Thus, there will be a higher proportion of long-chained FAs at early DIM than at a late DIM. Conversely, there will be a lower proportion of short- and medium-chained FAs at an early DIM than at a late DIM (Palmquist et al. 1993). The observed increase of C10:0, C14:0 and C16:0, and the decrease of C18:0 and C18:1*cis*-9, are therefore in agreement with what have been reported in the literature (Gross et al. 2011; Hawke & Taylor 1995). The CLA content only varied between 0.5 % and 0.6 % during the first 100 DIM. This is consistent with what has been reported by Gross et al. (2011). However, in that study the content only varied between 0.3 % and 0.4 %. The average proportion of the sum of SFAs increased from 63 % to 70% during the first 80 days. Since the major part of the sum of SFAs consists of the short- and medium-chained FAs, the trend for the sum of SFAs was expected to follow the same trend as observed for C10:0, C14:0 and C16:0. Likewise, the trend for the sum of MUFAs should be similar to the trend for C18:1*cis*-9. The sum of MUFAs decreases (as expected) from 29 % to 23 % during the first 80 DIM. The content of the sum of PUFAs decreased from 2.8 % to 2.4 % in the interval from DIM 5 to DIM 10. From DIM 10 to DIM 100, it seemed to decrease from approximately 2.4 % to 2.2%. The proportion of the sum of PUFAs has been observed to be relatively stable during the first 150 DIM (Gross et al. 2011). On the other hand, the sum of PUFAs in that study did not contain the same PUFAs as in this experiment. The observed relationships between the FAs are in accordance with what was expected from the regression coefficients discussed in Chapter 4.2.1. At early DIM, there were observed larger proportions of 18:1*cis*-9, sums of MUFAs and PUFAs, while there were smaller proportions of C14:0, C16:0 and sum of SFAs.

To sum up the results of the trends in FA composition during the lactation, a PCA of the FAs was performed. The average composition of samples from all cows at DIM 5 to DIM 100, respectively, was used. The score plot revealed a distinct trend for the samples. Samples from DIM 5 to DIM 29 were located to the left. Most of the samples from DIM 30 to DIM 60 were located in the middle, while most of the samples from about DIM 61 to DIM 100 were located to the right. PC1 explained 98 % of the variation in the data set. PC2 only explained 1 %. The sum of MUFAs had the most negative loading in PC1 (located to the left in the loading plot). This indicated that samples located to the left in the score plot contain a high proportion of MUFAs. Likewise, the samples located to the right in the score plot should contain a lower proportion of MUFAs. This is consistent with what was discussed previously in this chapter. There was a higher proportion of sum of MUFAs and C18:1*cis*-9 in samples from early DIM than late DIM. C18:0 are located to the left in the loading plot as well. This is also in agreement with the observed trend for C18:0. The variables CLA, sum of PUFAs and C10:0 are located in the middle of the loading plot, which reflects that these variables do not contribute much to the variation from DIM 5 to DIM 100. This is in accordance with what was observed in the plots for these variables during the first 100 DIM. The smallest variations were observed for these three variables. Of these three, CLA is located closest to 0, which is in agreement with the fact that CLA was the variable that varied the least. Sum of SFAs and C16:0 are located to the right in the loading plot (positive loading). Samples with a positive PC1 in the score plot will therefore contain a larger proportion of these two variables than samples with a negative PC1. This result is also in agreement with the observed trends for C16:0 and the sum of SFAs. The high explained variation in PC1 (98 %) might be explained by the clear trends in FA composition shown previously in this chapter. By using the predicted FA composition in the PCA instead of the whole FTIR spectra, the amount of redundant information is reduced.

4.2.4. Fatty acid composition of milk from cows with subclinical ketosis

The level of betahydroxybuturate in blood samples from the cows was measured in order to detect cows with subclinical ketosis. In this project, a cow was defined to experience subclinical ketosis if the concentration of betahydroxybuturate in blood samples was more than 1 ng/mL in week three. To investigate if the FTIR spectra could reveal any differences between cows with subclinical ketosis and normal cows, the predicted FAs were plotted over time for the two groups. The average content at DIM 5 to 100 of samples from the two groups was used. Clear visual differences in the content of C14:0 and C18:1cis-9 during the lactation

were observed between the two groups. Cows diagnosed with subclinical ketosis produced on average milk with a higher proportion of C18:1*cis*-9 and a lower proportion of C14:0 than normal cows. The milk from cows with subclinical ketosis had on average a bit lower proportion of C16:0 than the milk from normal cows. The proportion of C18:0 was similar for both groups. The result for C10:0 and sum of SFAs were similar to the results for C14:0, while the result for sum of MUFAs was similar to the result for C18:1*cis*-9. The plots of the content of CLA and the sum of PUFAs did not show any obvious differences between the two groups. These similarities were expected, due to the same argumentation as in the previous discussion (Chapter 4.2.3). Van Haelst et al. (2008) discovered that cows with subclinical ketosis had a higher proportion of C18:1*cis*-9 in the milk, compared to normal cows, already two weeks before diagnosis. Since negative EB may lead to ketosis (TINE Rådgiving 2008), it also was expected to observe differences in FAs related to EB between the two groups. As discussed earlier, higher proportions of long-chained FAs will be present in the beginning of the lactation (during negative EB) than later in the lactation (when the EB is positive). The opposite is the case for short- and medium-chained FAs. Gross et al. (2011) observed that milk from cows exposed to feed restriction contained a higher proportion of 18:1*cis*-9 than the milk from cows on normal diet during the first 120 DIM. During the same time interval, the milk from the feeding restricted cows contained a lower proportion of C16:0 and shorter FAs. From this, it was expected to find higher amounts of long-chained FAs and lower amounts of medium-chained FAs in milk from cows with subclinical ketosis. The observed logical and expected trends indicate that the dry-film FTIR spectra are suitable for prediction of these FAs and groups of FAs. In this experiment, a cow was considered to have subclinical ketosis if the level of betahydroxybuturate exceeded 1 ng/mL in blood samples from the third week in milk. Possibly, the difference between the two groups may become more distinct by increasing the threshold of betahydroxybuturate.

Statistical tests to check if there were significant differences between the two groups were not performed. However, the standard deviations (SDs) for the content of C14:0 and C18:*1cis*-9 for the two groups were calculated at each DIM to get an indication of how large the variations were. The SDs are included in the plots for the average contents of C14:0 (Figure 60) and C18:*1cis*-9 (Figure 61).

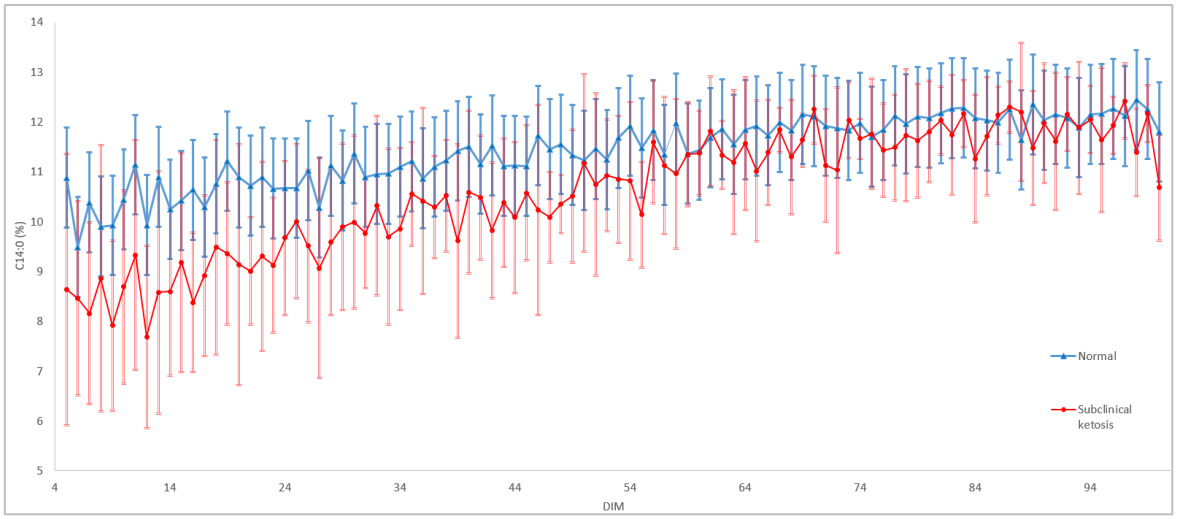


Figure 60. Average content of C14:0 in samples from cows with subclinical ketosis and normal cows. The percentage of total weight of fatty acids is plotted during the 100 first days in milk (DIM). Corresponding standard deviations are included as error bars.

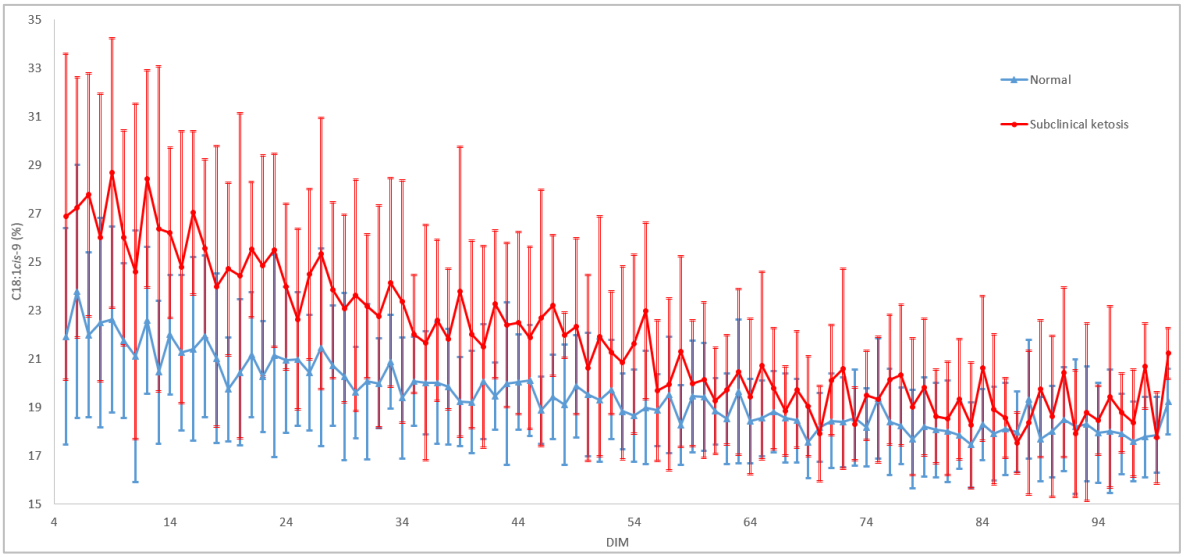


Figure 61. Average content of C18:1cis-9 in samples from cows with subclinical ketosis and normal cows. The percentage of total weight of fatty acids is plotted during the 100 first days in milk (DIM). Corresponding standard deviations are included as error bars.

The SDs reflect large variations between the individual cows. Despite the large variations, there are still clear trends in the data. The large SDs are due to the natural variations in fat composition among cows. A cow that naturally produces milk with a low proportion of C18:1*cis*-9 might, if experiencing subclinical ketosis, produce milk with a proportion of C18:1*cis*-9 in the lower part of the range for the subclinical ketosis group. Even though the elevated proportion for this cow may be very low compared to the average cow, this does not necessarily mean that the elevation is not significant for that specific cow. In

order to learn more about this, more exploration of how the predicted FA composition varies for the individual cows will be important. To minimize the variations that originate from the analysis, improvement of sampling and analysis procedures will be necessary.

It is obvious from Figure 60 and Figure 61 that the SDs are largest in the subclinical group. This might indicate that a large part of the variations is due to biological variations and not due to variations caused by analysis procedures. In general, the range (mean value ± 1 SD) of C14:0 content for the normal cows did not drop below the mean values for the cows with subclinical ketosis during the first 30 DIM. Conversely, the range of C18:1*cis*-9 for the normal cows did not exceed the mean value for the cows with subclinical ketosis during the first 50 DIM. This might indicate that, even though there are uncertainties associated with the analyses, the FTIR analysis can be used for screening of cows with subclinical ketosis. For instance, if a measurement of C18:1*cis*-9 in the beginning of the lactation results in a value below or close to the normal group's mean value, the cow can either be normal or at risk of developing subclinical ketosis. However, if the measurement results in a value above the subclinical group's mean value, the cow is likely to experience subclinical ketosis. Frequent measurements of individual cows will result in more knowledge about the normal values for each cow. Therefore, the cow could be examined further if she produces milk with abnormal values of, for instance, C18:1*cis*-9.

For improved fertility, it is important to minimize the extent of negative EB and this could be done by optimal feeding of the cows (Walsh et al. 2011). By detection of cows with subclinical ketosis, the farmer may intervene early to avoid reduced cow health and financial loss (Van Haelst et al. 2008). Early detection of cows that are likely to experience late OLA makes it easier to give the individual cow optimal treatment and dietary supplements (Martin et al. 2015). Both Afseth et al. (2010) and Martin et al. (2015) suggest that there is a potential for incorporation of dry-film FTIR analysis into AMSs. The results in the present project confirmed that prediction of FA composition from dry-film FTIR is possible. By incorporation in an AMS and by determination of the FA composition in the milk at every milking, important information about the cow can be acquired. However, practical challenges regarding implementation of the FTIR instrument in an AMS must first be investigated. For instance, the dry-film method includes a drying step that needs to be automatized before the method can be used in the barn.

This study has shown that the predicted contents of FAs from FTIR spectra are suitable for describing the average FA composition in a whole herd. Regarding the individual cows, there are larger uncertainties associated with the results. This means that there is future work to be done before using the predicted FA composition for monitoring of individual cows. Trends in FA composition for more cows than the three selected cows in this experiment should be explored to see if the large variations were present for all 61 cows. For instance, the day-to-day variation for the cow with ID number 5977 was very small for content of C14:0 in the period from DIM 5 to DIM 25, and for the sum of MUFAs in the period from DIM 70 to DIM 100. Another aim for future work is to improve the sampling and analysis procedures to ensure that all samples are treated equally prior to analysis. In an AMS, the sampling procedure is automated and the samples are fresh. Automated sampling may reduce errors introduced by humans, such as pipetting errors or mixing the samples. The milk in this experiment was analysed only three days a week. With incorporation in an AMS, measurement of the FA composition could be performed several times per day. This could result in smoother trends for the individual cows. Another possibility is to use the average predicted FA composition of all measurements per day, or the average of several days. Frequent measurements also would make it possible to follow a specific trend for one individual cow, and to detect abnormal values early.

This and other studies have shown that the EB is related to FA composition. Gross et al. (2011) showed that particularly changes in 18:1*cis*-9 and in groups of FAs (such as SFAs and MUFAs) could be of interest for monitoring of EB. Martin et al. (2015) identified significant relationships between C14:0, C16:0 and C18:1*cis*-9 and OLA. These relationships were explained by the correlation between EB and OLA. Prediction of EB from FTIR spectra of liquid milk samples was tested by McParland and Berry (2016). These prediction models resulted in R^2 of 0.78, but the prediction of EB on an external validation data set was poor. The logical trends observed between normal cows and cows with subclinical ketosis in the current experiment indicate that the EB could be predicted from dry-film FTIR spectra. Based on this, further work should include establishment of PLSR calibration models for prediction of EB, either directly from the pre-processed FTIR spectra or from the predicted FA composition.

The results of the prediction models for FA composition in this work, were in general better than previously established models calibrated with FTIR spectra of liquid samples (MilkoScan) (Afseth et al. 2010; Rutten et al. 2009; Soyeurt et al. 2006). The reasons for this

are the increased concentrations achieved by drying the samples. PCA of the dry-film spectra revealed patterns of the chemical composition during the lactation. Therefore, it is possible to obtain good calibration models for prediction of fat, protein, lactose and other parameters in milk from dry-film FTIR spectra. If dry-film FTIR was used in an AMS, information about contents of fat, FA composition, protein and lactose could be collected almost simultaneously. These parameters could be used in combination with each other or with other parameters for cow health monitoring. An additionally interesting feature in this respect is protein composition, which today is complicated and tedious to measure using standard chemical analysis tools.

5. Conclusion and perspectives

In the present project, dry-film FTIR was evaluated for the purpose of detecting and quantifying milk components for cattle health monitoring. Overall, the study showed that dry-film FTIR spectra could be used for this purpose.

Systematic changes in chemical composition of cow milk during the first 100 DIM were detected by using PCA of the dry-film FTIR spectra. No groupings of samples from cows with the same feeding regime were detected by PCA of the spectra. From the investigation of spectra and feeding regimes, it seems like other parameters affect the milk composition more than what the feeding regime does.

Previously established PLSR models, with reference samples from the current data set included, were used for prediction of the proportions of C10:0, C14:0, C16:0, C18:0, C18:1*cis*-9, CLA, and the sums of SFAs, MUFAs and PUFAs. Moderately good to good calibration models (with coefficients of determination, R^2, ranging from 0.76 to 0.97) were obtained. Systematic changes during the first 100 DIM were observed for the predicted FA composition. However, these trends were vague for the individual cows due to large day-to-day variations. Differences in the predicted FA composition between cows with subclinical ketosis and normal cows were observed (especially for C14:0 and C18:1*cis*-9). However, it was not examined if the differences were statistically significant.

Further research and development is needed before incorporation of this technology in AMSs for cow health monitoring. This work should include:

- Exploration of the predicted FA composition in the milk from all the individual cows.
- Testing whether the differences between the two groups (normal cows and cows with subclinical ketosis) were significant in any periods of the lactation.
- Establishment of PLSR models for prediction of other components (e.g. protein composition) in milk from dry-film FTIR spectra.
- Establishment of PLSR models for prediction of EB, either directly from the pre-processed FTIR spectra or from the predicted FA compositions.
- Improvement of procedures for sampling and analysis to minimize variations between samples.
- Investigation of practical challenges regarding implementation of dry-film FTIR in AMSs.

6. References

Afseth, N. K. (2007). *Vibrational spectroscopy and chemometrics for rapid and reliable characterisation of foods – A few aspects concerning Raman spectroscopy*. Philosophiae doctor thesis. Ås: Norwegian University of Life Sciences.

Afseth, N. K., Martens, H., Randby, Å., Gidskehaug, L., Narum, B., Jørgensen, K., Lien, S. & Kohler, A. (2010). Predicting the Fatty Acid Composition of Milk: A Comparison of Two Fourier Transform Infrared Sampling Techniques. *Applied Spectroscopy*, 64 (7): 700-707.

Barkema, H. W., von Keyserlingk, M. A. G., Kastelic, J. P., Lam, T. J. G. M., Luby, C., Roy, J. P., LeBlanc, S. J., Keefe, G. P. & Kelton, D. F. (2015). Invited review: Changes in the dairy industry affecting dairy cattle health and welfare. *Journal of Dairy Science*, 98 (11): 7426-7445.

Bewley, J. (2009). *Precision Dairy Farming*: University of Kentucky. Available at: http://www.uky.edu/Ag/AnimalSciences/dairy/dairysystems/jb0209.pdf (accessed: 2017-08-12).

Butler, G. & Stergiadis, S. (2011). Suitability of bronopol preservative treated milk for fatty acid determination. *Journal of Dairy Research*, 78 (2): 220-225.

Butler, W. R., Everett, R. W. & Coppock, C. E. (1981). The Relationships between Energy Balance, Milk Production and Ovulation in Postpartum Holstein Cows1. *Journal of Animal Science*, 53 (3): 742-748.

CAMO. (2017). *Multivariate Data Analysis*. Available at: http://www.camo.com/multivariate_analysis.html (accessed: 2017-11-03).

Christie, W. W. (1995). Composition and structure of milk lipids. In Fox, P. F. (ed.) Lipids, vol. 2 *Advanced dairy chemistry*, pp. 1-36. London: Chapman & Hall.

Cozzolino, D. (2013). Overview of infrared spectroscopy: hardware, chemometrics and software. In Cozzolino, D. (ed.) *Infrared Spectroscopy: Theory, Developments and Applications*, pp. 1-14. Hauppauge, New York: Nova Science Publishers, Inc.

de Vries, M. J. & Veerkamp, R. F. (2000). Energy Balance of Dairy Cattle in Relation to Milk Production Variables and Fertility. *Journal of Dairy Science*, 83 (1): 62-69.

Deeth, H. C. & Fitz-Gerald, C. H. (1995). Lipolytic enzymes and hydrolytic rancidity in milk and milk products. In Fox, P. F. (ed.) Lipids, vol. 2 *Advanced dairy chemistry*, pp. 247-308. London: Chapman & Hall.

Dewick, P. M. (2009). *Medicinal natural products: a biosynthetic approach*. 3rd ed. West Sussex: John Wiley & Sons Ltd. 539 pp.

Finlay, C. (2017). *Difference Between Spectrometer and Spectrophotometer*: Sciencing. Available at: http://sciencing.com/difference-between-spectrometer-spectrophotometer-8577067.html (accessed: 2017-09-08).

Geno. (2014a). *Fertility in the NRF cow (Fruktbarhet hos NRF-kua)*. Available at: https://www.geno.no/Start/Brunst/Fagstoff/KUFRUKTBARHET/Fruktbarhet-i-norge/ (accessed: 2017-08-03).

Geno. (2014b). *Shortage of oestrus (Brunstmangel)*. Available at: https://www.geno.no/Start/Brunst/Fagstoff/Fruktbarhetsproblemer/Brunstmangel/ (accessed: 2017-08-04).

Geno. (2014c). *Urea and protein*. Available at: https://www.geno.no/Start/Brunst/FAGSTOFF/Foring-og-fruktbarhet/Urea-og-protein/ (accessed: 2017-11-13).

Geno. (2016). *NRF in Norway and the world (NRF i Norge og verden)*. Available at: https://www.geno.no/Start/Geno-Avler-for-bedre-liv/om-nrf-kua/NRF-i-Norge-og-verden/?parent=282 (accessed: 2017-08-02).

Geno. (2017). *Characteristics of NRF (Karakteristikk hos NRF)*. Available at: https://www.geno.no/Start/Geno-Avler-for-bedre-liv/om-nrf-kua/Karakteristikk-hos-NRF/ (accessed: 2017-08-02).

Gross, J., van Dorland, H. A., Bruckmaier, R. M. & Schwarz, F. J. (2011). Milk fatty acid profile related to energy balance in dairy cows. *Journal of Dairy Research*, 78 (4): 479-488.

Harris, D. C. (2010). *Quantitative chemical analysis*. 8th ed. New York: W. H. Freeman and Company. 719 pp.

Hawke, J. C. & Taylor, M. W. (1995). Influence of nutritional factors on the yield, composition and physical properties of milk fat. In Fox, P. F. (ed.) Lipids, vol. 2 *Advanced dairy chemistry*, pp. 37-88. London: Chapman & Hall.

Helmenstine, A. M. (2017). *Spectroscopy definition and difference from spectrometry*: ThoughtCo. Available at: https://www.thoughtco.com/definition-of-spectroscopy-605676 (accessed: 2017-09-08).

Inchaisri, C., Jorritsma, R., Vos, P. L. A. M., van der Weijden, G. C. & Hogeveen, H. (2011). Analysis of the economically optimal voluntary waiting period for first insemination. *Journal of Dairy Science*, 94 (8): 3811-3823.

James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013). *An Introduction to Statistical Learning*. with Applications in R. New York: Springer. Available at: www.StatLearning.com.

Keenan, T. W. & Dylewski, D. P. (1995). Intracellular origin of milk lipid globules and the nature and structure of the milk lipid globule membrane. In Fox, P. F. (ed.) Lipids, vol. 2 *Advanced dairy chemistry*, pp. 89-130. London: Chapman & Hall.

Kohler, A., Afseth, N. K., Jørgensen, K., Randby, Å. & Martens, H. (2010). Quality analysis of milk by vibrational spectroscopy. In Li-Chan, E. C. Y., Griffiths, P. R. & Chalmers, J. M. (eds) *Applications of vibrational spectroscopy in food science*, pp. 483–499: John Wiley & Sons, Inc.

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

Li-Chan, E. C. Y. (2010). Introduction to vibrational spectroscopy in food science. In Li-Chan, E. C. Y., Chalmers, J. M. & Griffiths, P. R. (eds) *Applications of vibrational spectroscopy in food science*, pp. 3-29: John Wiley & Sons, Inc.

LibreTexts. (2016). *14.2: Lipids and Triglycerides*. Available at: https://chem.libretexts.org/LibreTexts/University_of_Kentucky/UK%3A_CHE_103_-_Chemistry_for_Allied_Health_(Soult)/Chapters/Chapter_14%3A_Biological_Molecules/14.2%3A_Lipids_and_Triglycerides (accessed: 2017-10-04).

Lindstrøm, K. S. (2017). *Feeding dairy cows after a standard lactation curve compared with a dynamic approach of the lactation curve (Fôring av melkekyr etter en standard laktasjonskurve sammenlignet med en dynamisk tilnærming av laktasjonskurven)*. Master's thesis. Ås: Norwegian University of Life Sciences, Faculty of Biosciences. 77 pp.

Linn, J. G. (1988). *Factors Affecting the Composition of Milk from Dairy Cows*. Available at: https://www.ncbi.nlm.nih.gov/books/NBK218193/ (accessed: 2017-11-30).

Løvendahl, P., Ridder, C. & Friggens, N. C. (2010). Limits to prediction of energy balance from milk composition measures at individual cow level. *Journal of Dairy Science*, 93 (5): 1998-2006.

Martin, A. D. (2015). *Surveillance of reproductive performance in Norwegian cattle farming*. Philosophiae doctor thesis. Ås: Norwegian University of Life Sciences.

Martin, A. D., Afseth, N. K., Kohler, A., Randby, Å., Eknæs, M., Waldmann, A., Dørum, G., Måge, I. & Reksen, O. (2015). The relationship between fatty acid profiles in milk identified by Fourier transform infrared spectroscopy and onset of luteal activity in Norwegian dairy cattle. *Journal of Dairy Science*, 98 (8): 5374-5384.

McParland, S. & Berry, D. P. (2016). The potential of Fourier transform infrared spectroscopy of milk samples to predict energy intake and efficiency in dairy cows1. *Journal of Dairy Science*, 99 (5): 4056-4070.

Miller, J. N. & Miller, J. C. (2010). *Statistics and chemometrics for analytical chemistry*. 6th ed. Harlow, England: Pearson Education Limited. 278 pp.

Minitab. (2017). *Coefficients and standardized coefficients for Partial Least Squares Regression*. Available at: https://support.minitab.com/en-us/minitab/18/help-and-how-to/modeling-statistics/regression/how-to/partial-least-squares/interpret-the-results/all-statistics-and-graphs/coefficients/ (accessed: 2017-10-14).

Månsson, H. L. (2008). Fatty acids in bovine milk fat. *Food & Nutrition Research*, 52 (1): 1821-1823.

NASA. (2013). *The Electromagnetic Spectrum*. Available at: https://imagine.gsfc.nasa.gov/science/toolbox/emspectrum1.html (accessed: 2017-09-08).

Neethirajan, S. (2017). Recent advances in wearable sensors for animal health management. *Sensing and Bio-Sensing Research*, 12: 15-29.

Ottinsen, H. G. (2013). *How the milking system affects the mastitis frequency in Norwegian dairy herds (Hvordan mjølkesystem påvirker mastittfrekvensen i norske besetninger)*. Master's thesis. Ås: Norwegian University of Life Sciences, Department of Animal and Aquacultural Sciences. 63 pp.

Palmquist, D. L., Beaulieu, A. D. & Barbano, D. M. (1993). Feed and animal factors influencing milk fat composition. *Journal of Dairy Science*, 76 (6): 1753-1771.

PlantFAdb. (2016a). *9-Octadecenoic acid*. Available at: https://plantfadb.bch.msu.edu/fatty_acids/10074 (accessed: 2017-09-15).

PlantFAdb. (2016b). *Fatty acids*. Available at: https://plantfadb.bch.msu.edu/fatty_acids?utf8=%E2%9C%93&query=&commit=Go (accessed: 2017-09-18).

Ptak, E., Brzozowski, P. & Bieniek, J. (2012). Genetic parameters for lactose percentage in the milk of Polish Holstein-Friesians. *Journal of Animal and Feed Sciences*, 21 (2): 251-262.

Rajalahti, T. & Kvalheim, O. M. (2011). Multivariate data analysis in pharmaceutics: A tutorial review. *International Journal of Pharmaceutics*, 417 (1): 280-290.

Rustan, A. C. & Drevon, C. A. (2005). *Fatty Acids: Structures and Properties*: John Wiley & Sons, Ltd. Available at: https://www.uio.no/studier/emner/matnat/farmasi/FRM2041/v06/undervisningsmateriale/fatty_acids.pdf (accessed: 2017-09-15).

Rutten, M. J. M., Bovenhuis, H., Hettinga, K. A., van Valenberg, H. J. F. & van Arendonk, J. A. M. (2009). Predicting bovine milk fat composition using infrared spectroscopy based on milk samples collected in winter and summer. *Journal of Dairy Science*, 92 (12): 6202-6209.

SCENIHR. (2008). *Scientific opinion on light sensitivity*. Available at: http://ec.europa.eu/health/ph_risk/committees/04_scenihr/docs/scenihr_o_019.pdf (accessed: 2017-09-08).

Soyeurt, H., Dardenne, P., Dehareng, F., Lognay, G., Veselko, D., Marlier, M., Bertozzi, C., Mayeres, P. & Gengler, N. (2006). Estimating Fatty Acid Content in Cow Milk Using Mid-Infrared Spectrometry. *Journal of Dairy Science*, 89 (9): 3690-3695.

Stoop, W. M., Bovenhuis, H., Heck, J. M. L. & van Arendonk, J. A. M. (2009). Effect of lactation stage and energy status on milk fat composition of Holstein-Friesian cows. *Journal of Dairy Science*, 92 (4): 1469-1478.

TINE Rådgiving. (2008). *Ketosis (Ketose)*. Available at: http://storfehelse.no/dyrehelse/produksjonssjukdommer/ketose (accessed: 2017-11-13).

TINE SA. (2017). *TINE's regulatory framework for assessment and payment of milk by quality when delivered to TINE Raw material (TINEs regelverk om bedømmelse og betaling av melk etter kvalitet ved levering til TINE Råvare)*. Available at: https://medlem.tine.no/praktisk-informasjon/tines-regelverk/_attachment/397755?_ts=159e00e1600 (accessed: 2017-11-13).

Tsioulpas, A., Grandison, A. S. & Lewis, M. J. (2007). Changes in Physical Properties of Bovine Milk from the Colostrum Period to Early Lactation. *Journal of Dairy Science*, 90 (11): 5012-5017.

Van Haelst, Y. N. T., Beeckman, A., Van Knegsel, A. T. M. & Fievez, V. (2008). Short Communication: Elevated Concentrations of Oleic Acid and Long-Chain Fatty Acids in Milk Fat of Multiparous Subclinical Ketotic Cows. *Journal of Dairy Science*, 91 (12): 4683-4686.

Vernier Software & Technology. (2001). *What are Mean Squared Error and Root Mean Squared Error?* Available at: https://www.vernier.com/til/1014/ (accessed: 2017-10-14).

Walsh, S. W., Williams, E. J. & Evans, A. C. O. (2011). A review of the causes of poor fertility in high milk producing dairy cows. *Animal Reproduction Science*, 123 (3): 127-138.

Wiking, L., Nielsen, J. H., Båvius, A. K., Edvardsson, A. & Svennersten-Sjaunja, K. (2006). Impact of Milking Frequencies on the Level of Free Fatty Acids in Milk, Fat Globule Size, and Fatty Acid Composition. *Journal of Dairy Science*, 89 (3): 1004-1009.

Williams, D. H. & Fleming, I. (2008). *Spectroscopic methods in organic chemistry* 6th ed. Maidenhead: McGraw-Hill Education. 291 pp.

Appendix A. The instrument settings

The optic, acquisition and Fourier transform (FT) settings used in OPUS 6.5 are provided in Figures A1, A2 and A3, respectively.

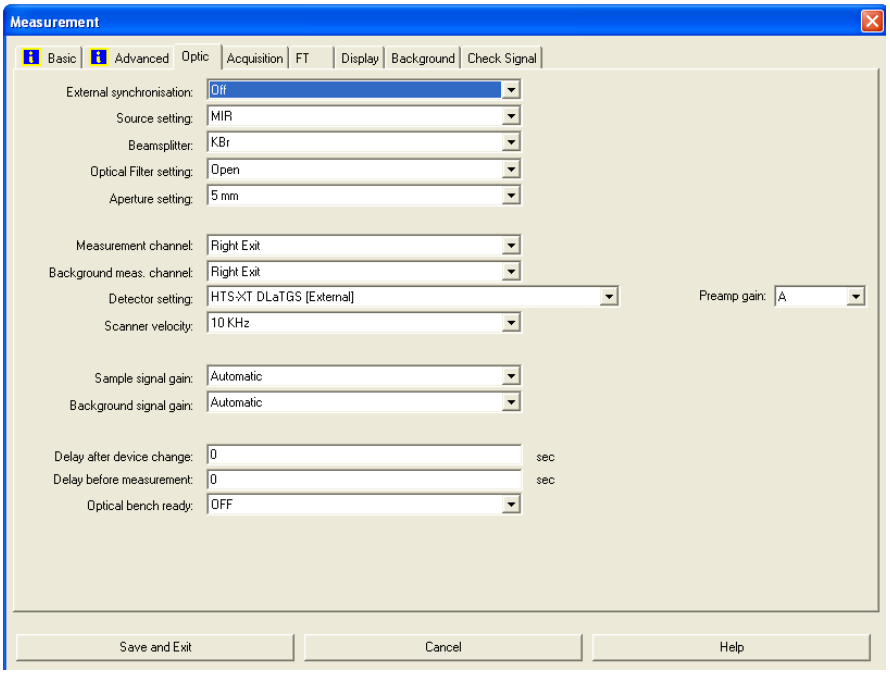


Figure A1. The optic settings in OPUS 6.5.

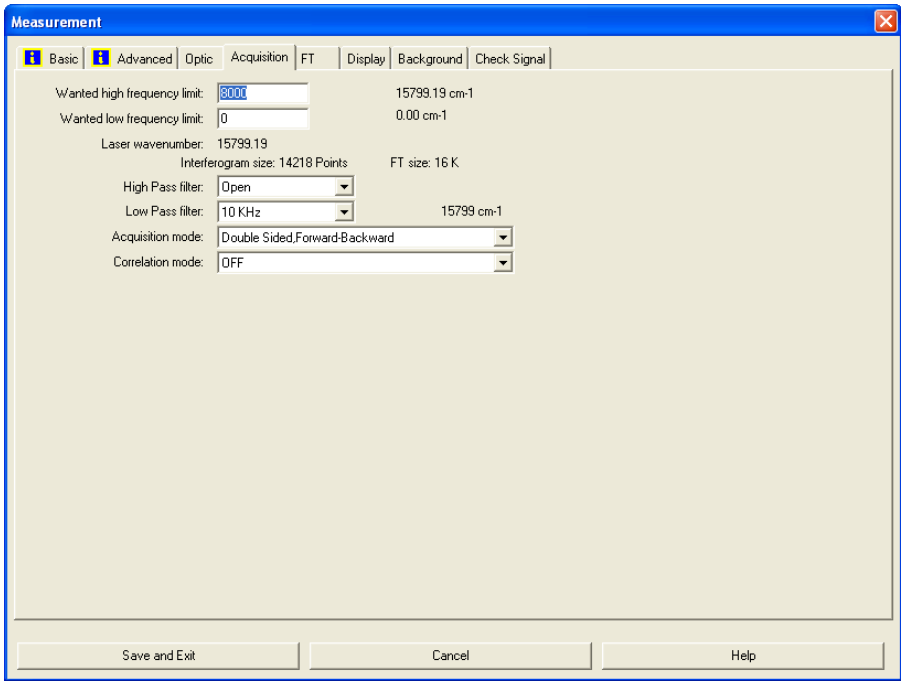


Figure A2. The acquisition settings in OPUS 6.5.

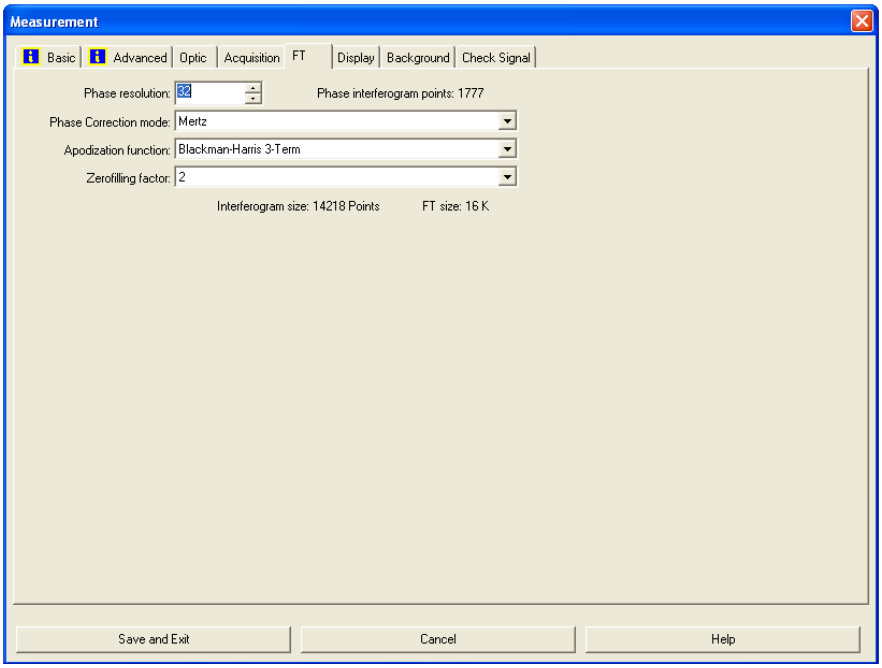


Figure A3. The Fourier transform (FT) settings in OPUS 6.5.

Appendix B. Information about the cows

Cow ID number	Feeding regime[1]	Subclinical ketosis[2]	Number of samples	First sampling date	DIM (min)	Last sampling date	DIM (max)
5613	3	0	39	2016-11-02	5	2017-02-01	96
5651	1	0	37	2016-11-07	5	2017-02-01	91
5725	1	1	40	2016-09-14	5	2016-12-14	96
5756	4	1	41	2016-09-09	6	2016-12-12	100
5830	3	1	40	2016-09-14	6	2016-12-14	97
5840	3	1	37	2016-09-05	6	2016-11-28	90
5843	1	1	40	2016-10-24	6	2017-01-25	99
5845	2	0	37	2016-11-04	6	2017-02-01	95
5879	4	1	35	2016-09-05	5	2016-11-28	89
5903	1	0	27	2016-12-02	6	2017-02-06	72
5912	2	0	40	2016-11-16	5	2017-02-17	98
5920	3	1	37	2016-09-19	7	2016-12-14	93
5925	2	1	34	2016-12-09	5	2017-02-27	85
5946	3	1	36	2016-09-05	6	2016-11-28	90
5963	3	1	36	2016-10-17	5	2017-01-09	89
5965	2	1	39	2016-09-16	5	2016-12-14	94
5977	4	0	40	2016-10-31	7	2017-02-01	100
5989	4	0	35	2016-12-02	6	2017-02-27	93
6025	2	0	40	2016-10-07	5	2017-01-09	99
6042	4	0	40	2016-09-14	5	2016-12-14	96
6043	2	0	39	2016-11-25	5	2017-02-27	99
6044	1	0	39	2016-09-23	5	2016-12-23	96
6065	1	0	37	2016-11-28	6	2017-02-27	97
6088	2	0	41	2016-09-12	5	2016-12-14	98
6094	2	1	41	2016-09-21	6	2016-12-23	99
6104	3	1	39	2016-09-19	10	2016-12-16	98
6105	1	0	39	2016-09-16	6	2016-12-14	95
6113	4	0	41	2016-09-09	5	2016-12-12	99
6115	4	0	36	2016-12-05	7	2017-02-27	91
6116	1	1	38	2016-09-19	7	2016-12-14	93
6130	3	1	39	2016-11-18	5	2017-02-20	99
6138	3	0	36	2016-11-07	6	2017-02-01	92
6141	4	0	36	2016-11-09	6	2017-02-01	90
6143	3	0	40	2016-10-17	7	2017-01-18	100
6153	2	0	36	2016-11-04	6	2017-02-01	95
6156	1	1	38	2016-11-16	6	2017-02-17	99
6160	4	0	36	2016-11-09	5	2017-02-01	89
6169	2	0	35	2016-10-19	6	2017-01-09	88
6196	3	0	41	2016-09-12	7	2016-12-14	100
6222	3	0	39	2016-08-31	5	2016-11-28	94
6228	1	0	41	2016-09-26	5	2016-12-30	100
6234	2	0	39	2016-09-16	5	2016-12-14	94
6238	2	0	35	2016-10-17	7	2017-01-09	91
6242	4	0	41	2016-09-12	7	2016-12-14	100
6245	3	0	38	2016-09-19	7	2016-12-14	93
6249	4	1	41	2016-09-26	5	2016-12-30	100
6253	1	0	36	2016-09-23	6	2016-12-23	97
6258	4	1	39	2016-10-31	5	2017-02-01	98
6260	3	1	40	2016-10-28	5	2017-01-30	99
6264	1	0	37	2016-09-02	5	2016-11-28	92
6266	1	0	39	2016-10-24	6	2017-01-23	97
6268	1	0	40	2016-09-14	6	2016-12-14	97
6274	1	1	40	2016-10-28	5	2017-01-30	99
6275	4	0	41	2016-09-12	6	2016-12-14	99
6282	3	0	41	2016-10-22	5	2017-01-25	100
6284	3	0	38	2016-10-28	5	2017-01-30	99
6285	2	0	41	2016-09-21	5	2016-12-23	98
6290	2	0	38	2016-09-02	6	2016-11-28	93
6292	2	0	39	2016-10-31	6	2017-02-01	99
6301	4	0	35	2016-11-09	6	2017-02-01	90
6302	2	0	38	2016-11-25	5	2017-02-27	99

[1] The types of feeding regimes were (1) high quality food and static feed allocation, (2) high quality food and dynamic feed allocation, (3) medium quality food and static feed allocation and (4) medium quality food and dynamic feed allocation.

[2] Normal cows (0) and cows with subclinical ketosis (1).

Appendix C. Number of samples per days in milk

The number of samples used when calculating the average content of chemical components per DIM (days in milk) is provided in Table C1. The number of samples used when calculating the average FTIR spectra and fatty acid composition is provided in Table C2.

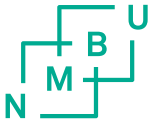
Table C1. Number of samples per DIM (days in milk). This information was used when calculating the average content of chemical components of all samples from all cows at each DIM. Total number of samples was 2143.

DIM	Samples	DIM	Samples	DIM	Samples	DIM	Samples	DIM	Samples	DIM	Samples
5	25	**21**	24	**37**	22	**53**	25	**69**	19	**85**	25
6	22	**22**	24	**38**	27	**54**	25	**70**	23	**86**	21
7	24	**23**	20	**39**	18	**55**	20	**71**	26	**87**	28
8	24	**24**	28	**40**	28	**56**	25	**72**	20	**88**	23
9	23	**25**	22	**41**	20	**57**	27	**73**	21	**89**	22
10	25	**26**	25	**42**	21	**58**	24	**74**	24	**90**	20
11	21	**27**	21	**43**	23	**59**	27	**75**	20	**91**	18
12	28	**28**	20	**44**	22	**60**	27	**76**	20	**92**	21
13	19	**29**	21	**45**	24	**61**	20	**77**	21	**93**	20
14	24	**30**	20	**46**	24	**62**	19	**78**	24	**94**	20
15	26	**31**	26	**47**	27	**63**	22	**79**	21	**95**	14
16	22	**32**	25	**48**	20	**64**	24	**80**	28	**96**	15
17	28	**33**	26	**49**	26	**65**	26	**81**	23	**97**	9
18	23	**34**	19	**50**	27	**66**	26	**82**	29	**98**	12
19	26	**35**	24	**51**	21	**67**	19	**83**	19	**99**	7
20	19	**36**	20	**52**	25	**68**	25	**84**	22	**100**	8

Table C2. Number of samples per DIM (days in milk). This information was used when calculating the average spectrum of all samples from all cows at each DIM. It was also used when calculating the average predicted FA composition. Total number of samples was 2333.

DIM	Samples[1]	DIM	Samples[1]	DIM	Samples[1]	DIM	Samples[1]	DIM	Samples[1]	DIM	Samples[1]
5	28 (10)	**21**	25 (8)	**37**	25 (8)	**53**	25 (9)	**69**	22 (6)	**85**	29 (13)
6	23 (7)	**22**	27 (10)	**38**	30 (10)	**54**	26 (10)	**70**	23 (8)	**86**	22 (6)
7	24 (7)	**23**	24 (7)	**39**	20 (4)	**55**	20 (6)	**71**	27 (11)	**87**	30 (9)
8	27 (11)	**24**	29 (10)	**40**	31 (12)	**56**	26 (9)	**72**	24 (6)	**88**	26 (8)
9	24 (5)	**25**	25 (6)	**41**	23 (8)	**57**	28 (11)	**73**	26 (8)	**89**	24 (8)
10	29 (11)	**26**	28 (11)	**42**	23 (6)	**58**	24 (7)	**74**	26 (8)	**90**	21 (7)
11	23 (6)	**27**	22 (7)	**43**	26 (10)	**59**	27 (10)	**75**	26 (10)	**91**	22 (6)
12	29 (11)	**28**	24 (8)	**44**	26 (8)	**60**	27 (6)	**76**	23 (7)	**92**	21 (8)
13	22 (7)	**29**	26 (11)	**45**	26 (10)	**61**	24 (10)	**77**	24 (8)	**93**	21 (6)
14	25 (8)	**30**	21 (6)	**46**	25 (6)	**62**	21 (7)	**78**	25 (10)	**94**	20 (6)
15	27 (10)	**31**	30 (10)	**47**	28 (10)	**63**	25 (8)	**79**	24 (8)	**95**	14 (4)
16	25 (7)	**32**	25 (7)	**48**	21 (5)	**64**	25 (9)	**80**	29 (10)	**96**	17 (6)
17	28 (10)	**33**	27 (11)	**49**	26 (9)	**65**	26 (7)	**81**	24 (6)	**97**	11 (5)
18	25 (6)	**34**	22 (7)	**50**	28 (11)	**66**	29 (10)	**82**	30 (11)	**98**	12 (3)
19	27 (11)	**35**	25 (8)	**51**	22 (5)	**67**	23 (6)	**83**	20 (7)	**99**	13 (6)
20	23 (7)	**36**	26 (10)	**52**	27 (8)	**68**	27 (10)	**84**	24 (7)	**100**	8 (2)

[1] The number of samples from cows with subclinical ketosis is shown in brackets.



Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway