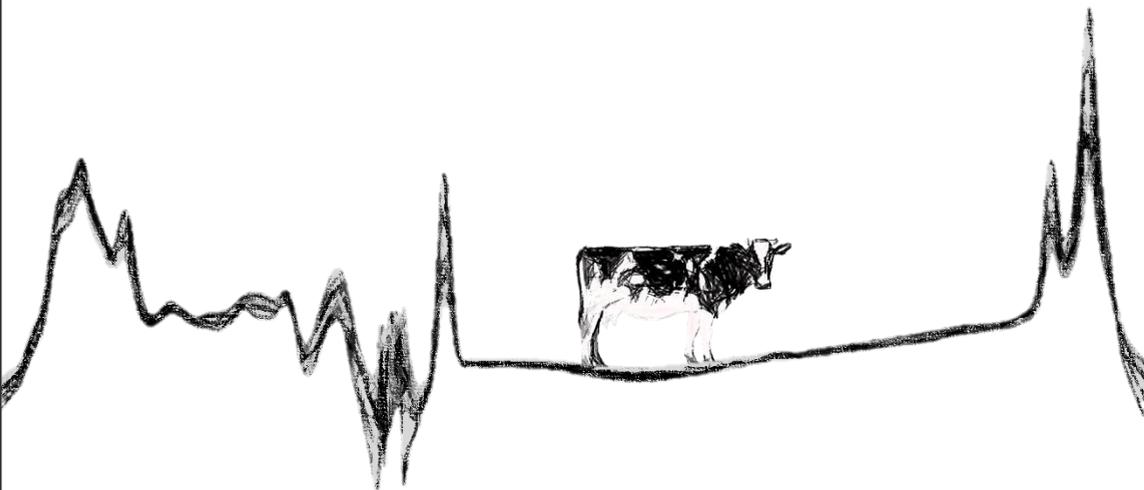


Promoting international prediction models
through standardization of milk
mid-infrared spectra

Clément Grelet



COMMUNAUTÉ FRANÇAISE DE BELGIQUE ACADÉMIE
UNIVERSITAIRE WALLONIE-EUROPE UNIVERSITÉ DE LIÈGE –
GEMBLoux AGRO-BIO TECH

PROMOTING INTERNATIONAL PREDICTION MODELS THROUGH
STANDARDIZATION OF MILK MID-INFRARED SPECTRA

Clément GRELET

Dissertation originale présentée en vue de l'obtention du grade de docteur en
sciences agronomiques et ingénierie biologique

Promoteurs: Prof. Hélène Soyeurt & Dr. Pierre Dardenne

2019

Copyright. Aux termes de la loi belge du 30 juin 1994, sur le droit d'auteur et les droits voisins, seul l'auteur a le droit de reproduire partiellement ou complètement cet ouvrage de quelque façon et forme que ce soit ou d'en autoriser la reproduction partielle ou complète de quelque manière et sous quelque forme que ce soit. Toute photocopie ou reproduction sous une autre forme est donc faite en violation de la dite loi et des modifications ultérieures.

Abstract

Analysis of milk by Fourier transform mid-infrared (**FT-MIR**) spectrometry provides a large amount of information on the physico-chemical composition of individual milk samples. Hence, it has been used for decades to predict fat, protein and lactose contents, and more recently fine milk composition, milk processing qualities and status of cows. This fast and cost-effective technology is a perfect candidate to provide new information for the management of individual cows. However, its concrete use by field organizations is still suboptimal given the difficulty of sharing data and models among spectrometers. The aim of this research was to optimize the use of FT-MIR analysis of milk with the final purpose of enabling the development of new management tools for dairy farmers.

In order to harmonize spectral responses among instruments and allow sharing of data and models, the first objective was to test a standardization method, well known from the NIR sector, in the frame of FT-MIR spectrometers dedicated to milk analysis. The possibility of standardizing such instruments was assessed by using the Piecewise Direct Standardization (**PDS**) method and common raw milk samples constituted from the IDF norm (ISO 9622:2013 | IDF 141:2013). The performances of spectral harmonization were assessed by the transfer of a robust fat model from a master instrument into 21 slave instruments. Regressions were performed between master and each slave fat predictions, before and after PDS. The biases and the root mean square errors between the predictions decreased after PDS from 0.378 to 0.000 and from 0.461 to 0.016 (g of fat/100 mL of milk), respectively. These preliminary results showed that the PDS method permits a reduction of the inherent spectral variability between instruments and the use of common robust models by all the spectrometers included in the constituted network.

The second objective was to ensure that models of interest with low precision could also be transferred from instrument to instrument. The effect of standardization on network spectral reproducibility was assessed on 66 instruments from 3 different brands by comparing the spectral variability of the slave instruments with and without standardization. With standardization, the standardized Mahalanobis distance (GH) between the slaves and master spectra was reduced on average from 2,656 to 14. The transfer of models from instrument to instrument was then tested using 3 FT-MIR models predicting the quantity of daily methane emitted by dairy cows, the concentration of polyunsaturated fatty acids in milk, and the fresh cheese yield. The differences, in terms of root mean squared error, between master and slaves predictions were reduced after standardization on average from 103 to 17 g/d for methane, from 0.032 to 0.005 g/100 mL of milk for polyunsaturated fatty acids, and from 2.55 to 0.49 g of curd/100 g of milk for fresh cheese yield. For all models, an improvement of prediction reproducibility within the network has also been observed. Concretely, the spectral standardization allows the transfer and use of multiple models on all instruments as well as the improvement of spectral and prediction reproducibility within the network. The method offers opportunities for data exchange

as well as the creation and use of common database and models, at an international level, to provide more information for the management of dairy herds.

After ensuring the possibility of using spectral data under optimal conditions, the third objective was to concretize the development of models providing information on cow status to be used as management tools by dairy farmers. This work aimed to develop models to predict milk citrate, reflecting early energy imbalance, and milk acetone and β -hydroxybutyrate (BHB) as indicators of (sub)clinical ketosis. Milk samples were collected in commercial and experimental farms in Luxembourg, France, and Germany. Milk mid-infrared spectra were recorded locally and standardized. Prediction equations were developed using partial least square regression. The coefficient of determination (R^2) of cross-validation was 0.73 for acetone, 0.71 for BHB, and 0.90 for citrate with root mean square error of 0.248, 0.109, and 0.70 mmol/L, respectively. Finally, an external validation was performed and R^2 obtained were 0.67 for acetone, 0.63 for BHB, and 0.86 for citrate, with a root mean square error of validation of 0.196, 0.083, and 0.76 mmol/L, respectively. The results demonstrated the potential of FT-MIR spectrometry to predict citrate content with good accuracy and to supply indicative contents of BHB and acetone in milk, thereby providing rapid and cost-effective tools to manage ketosis and negative energy balance in dairy farms.

This research highlights new knowledge and possibilities regarding the harmonization of spectral format from different instruments in order to create, share and use FT-MIR models providing information for the management of dairy cows. More concretely, it contributes outputs as procedures to standardize instruments in routine and models to predict indicators of negative energy balance and ketosis to help farmers in the management of early lactation period.

Résumé

L'analyse du lait par spectroscopie moyen infrarouge à transformée de Fourier (FT-MIR) est une source d'information importante sur la composition physico-chimique des échantillons de lait. Pour cette raison, elle est utilisée depuis des dizaines d'années afin de prédire les teneurs en matières grasses, protéines et lactose du lait, et plus récemment la composition fine du lait, les propriétés technologiques ou encore l'état de la vache. Cette technologie rapide et économique a donc un potentiel élevé dans le but de générer de nouvelles informations permettant d'améliorer la gestion individuelle des vaches laitières au sein des exploitations. Cependant, son utilisation concrète pour le conseil en élevage reste sous-optimal en raison de la difficulté de partager les données et modèles entre spectromètres. Le but de cette recherche a donc été d'optimiser l'utilisation de l'analyse FT-MIR du lait afin de permettre le développement d'outils de gestion pour les éleveurs laitiers.

Afin d'harmoniser les réponses spectrales entre les instruments et de permettre le partage des données et des modèles, le premier objectif a été de tester une méthode de standardisation, utilisée dans le secteur du proche infrarouge, dans le cadre de spectromètres dédiés à l'analyse moyen infrarouge du lait. La méthode Piece-wise Direct Standardization (PDS) a été testée en combinaison avec des échantillons de lait cru générés selon la norme IDF (ISO 9622:2013 | IDF 141:2013). Les performances de cette harmonisation spectrale ont été évaluées en transférant un modèle robuste de prédiction de la matière grasse d'un instrument de référence vers 21 instruments secondaires. Le biais et l'erreur standard ont été réduits après PDS, respectivement de 0.378 à 0.000 et de 0.461 à 0.016 g/100 mL de lait. Ces résultats préliminaires ont montré que l'utilisation de la PDS permet de réduire la variabilité spectrale entre les instruments et d'utiliser des modèles communs sur toutes les machines intégrant le réseau de standardisation.

Le deuxième objectif visait à tester la possibilité de transférer sur différents instruments des modèles d'intérêt ayant une faible précision. L'effet de la standardisation sur la reproductibilité spectrale du réseau a été évalué sur 66 machines provenant de 3 constructeurs, en comparant la variabilité spectrale des instruments secondaires avec ou sans standardisation. Après standardisation, la distance standardisée de Mahalanobis (GH) entre les spectres des appareils référence et secondaires a été réduite de 2 656 à 14. Le transfert de modèles sur différents instruments a été testé en utilisant 3 modèles prédisant la quantité de méthane émis par les vaches laitières, la teneur en acides gras polyinsaturés et le rendement fromager frais du lait. Les différences, en terme d'erreur standard, entre les prédictions des instruments secondaires et de la référence ont été réduites en moyenne après standardisation de 103 à 17 g/j pour le méthane, de 0.032 à 0.005 g/100mL de lait pour les acides gras polyinsaturés et de 2.55 à 0.49 g de caillé/100 g de lait pour le rendement fromager frais. Pour les 3 modèles, une amélioration de la reproductibilité des prédictions au sein du réseau a aussi été observée. Concrètement, la standardisation spectrale permet le transfert et l'utilisation de modèles multiples sur différents instruments, ainsi que l'amélioration de la reproductibilité des spectres et

des prédictions au sein du réseau. Cette méthode ouvre donc des perspectives en terme d'échange de données, de création et d'utilisation de modèles à un niveau international afin de générer plus d'informations pour la gestion des troupeaux laitiers.

Après avoir permis l'utilisation de données spectrales dans des conditions optimisées, le troisième objectif de ce travail a été de concrétiser le développement de modèles générant des informations sur le statut des vaches laitières afin d'être utilisés comme des outils de gestion de troupeaux. Le but concret de ce travail a été de développer des modèles de prédiction du citrate dans le lait, comme indicateur précoce de balance énergétique négative, ainsi que du BHB et de l'acétone dans le lait, comme indicateurs d'acétonémie (sub)clinique. Pour cela, des échantillons individuels de lait ont été collectés dans des fermes expérimentales et commerciales en France, Allemagne et Luxembourg. Les spectres FT-MIR de lait ont été collectés localement et standardisés dans un format commun. Les équations de prédictions ont été développées en utilisant la Partial Least Square (PLS) régression. Les coefficients de régression obtenus en cross-validation (R^2_{cv}) sont de 0.73 pour l'acétone, 0.71 pour le BHB et 0.90 pour le citrate avec une erreur standard (RMSE) respective de 0.248, 0.109 et 0.70 mmol/L. En validation externe, les R^2 obtenus sont de 0.67 pour l'acétone, 0.63 pour le BHB et 0.86 pour le citrate, avec des erreurs respectives de 0.196, 0.083 et 0.76 mmol/L. Ces résultats montrent le potentiel de la spectroscopie FT-MIR afin de prédire la teneur en citrate avec une bonne précision et de générer des teneurs indicatives en BHB et acétone dans le lait, fournissant ainsi des outils rapides et économiques de gestion de l'acétonémie et de la balance énergétique des vaches en ferme laitière.

Cette recherche apporte de nouvelles connaissances et possibilités en ce qui concerne l'harmonisation du format des spectres provenant de différents instruments, afin de créer, diffuser et utiliser des outils d'aide à la décision pour la gestion des vaches laitières. Plus concrètement, ce travail aura permis de mettre au point une méthode et des procédures afin de standardiser des instruments en routine et de créer des modèles de prédiction de la balance énergétique et de l'acétonémie afin d'aider les éleveurs à mieux gérer le début de lactation des vaches laitières.

Comité de thèse

Promoteurs

Hélène Soyeurt, Université de Liège, Gembloux Agro-Bio Tech

Pierre Dardenne, Centre Wallon de Recherches Agronomiques

Membres

Frédéric Dehareng, Centre Wallon de Recherches Agronomiques

Nicolas Gengler, Université de Liège, Gembloux Agro-Bio Tech

Yves Beckers, Université de Liège, Gembloux Agro-Bio Tech

Juan Antonio Fernandez Pierna, Centre Wallon de Recherches Agronomiques

Frédéric Colinet, Université de Liège, Gembloux Agro-Bio Tech

Vincent Baeten, Centre Wallon de Recherches Agronomiques

Remerciements

En tout premier lieu à mes Promoteurs, Hélène Soyeurt pour m'avoir incité à réaliser ce travail, l'avoir suivi et l'avoir tiré vers le haut, et à Pierre Dardenne pour m'avoir transmis une partie de son savoir tellement précieux en terme d'infrarouge et de modélisation.

A Frédéric Dehareng, pour m'avoir confié ce boulot, pour la bienveillance, la confiance, pour faire vivre l'équipe et lui permettre de travailler dans une ambiance humaine et décontractée.

Aux membres du comité de thèse qui m'ont guidé au cours de ce travail, Juan Antonio Fernandez pour m'avoir appris Matlab, pour les scripts et pour les précieux conseils touristiques, Vincent Baeten pour l'expertise en spectroscopie et pour l'optimisme sans faille, Nicolas Gengler pour les idées toujours plus innovantes, Catherine Bastin et Frédéric Colinet pour toutes les chouettes collaborations, pour les karaokés absurdes, discussions tardives et autres défis d'EAAP et à Yves Beckers pour avoir accepté de présider le jury.

To Dr. P. Berzaghi, and Pr. J. Bindelle who accepted to provide their experience and time to evaluate this work with an external point of view.

Au CRA-W, à Mr Sinnaeve, Mr Berben et Mr Poismans pour avoir permis de faire ce doctorat. A EMR, aux contrôles laitiers, laboratoires, universités et centres de recherches du projet OptiMIR, pour cette collaboration fructueuse et pour les données utilisées dans ce travail. A Interreg IV-B et à la Région Wallonne pour le financement de ce projet.

Evidemment à l'équipe du labo lait, Amélie, ma colloque de bureau, pour les heures (années ?) de discussion aussi bien scientifiques que sans queue ni tête, Olivier pour sa précieuse expérience dans le domaine laitier et pour refaire le monde de temps à autre, Petimat pour la bonne humeur, pour être la maman du labo, pour travailler sans relâche tout en nous apprenant le russe, Claire pour sa gentillesse et pour la qualité de son travail, Mohammed et Matthieu pour leur amabilité et le renfort bienvenu pour la création des gammes, et Octave qui vient nous enlever une vilaine épine du pied.

A tous les collègues du Henseval, anciens et actuels, pour l'entraide et la bonne ambiance quotidienne, mais aussi pour les désormais incontournables barbecs au ravel, descentes de la less(ff)e et pour me ramener quand je loupe le train ;)

Finalement, et le plus important, à tous les proches, pour m'avoir encouragé à faire un doctorat, ou tout simplement parce qu'ils ont une valeur infinie à mes yeux : Lorianne grâce à qui chaque journée est nouvelle et un bonheur, Hugo, petit loulou/rayon de soleil arrivé en cours de route, et tous ceux qui sont là depuis le début (ou très longtemps ;) qui sont indispensable à mon monde : Papa, Maman, Charlotte, Joseph, Loulou, la petite Anna et tout le reste de la famille, tous les zouaves Vendéens, Nalou, Piou, Caro, Anouck, Margaux et Léane, les copains Belges, Katia, Yvo et Paul, les french tocards et tous les autres !

Table of content

Chapter 1	: General introduction	1
Chapter 2	: Standardization of milk mid-infrared spectra from a European dairy network	14
Chapter 3	: Standardization of milk mid-infrared spectrometers for the transfer and use of multiple models	32
Chapter 4	: Development of Fourier transform mid-infrared calibrations to predict acetone, β -hydroxybutyrate, and citrate contents in bovine milk through a European dairy network	55
Chapter 5	: General discussion	72
Chapter 6	: Perspectives and conclusion	84
Annex		91

Figures

Figure 1-1. Spectrum of a milk sample	3
Figure 2-1. Graphical view of the piecewise direct standardization technique. m_j = the response measured at wavenumber j on the master instrument; s_j = the response measured in a small window of size n around j (neighboring) on the slave instrument; T = transmittance.....	17
Figure 2-2. Raw milk spectrum from the master instrument. T = transmittance...18	
Figure 2-3. Master milk spectrum with assignment of the main spectral bands, after removal of noise areas from water absorption. T = transmittance.	19
Figure 2-4. Root mean square error (RMSE) between predictions of each slave instruments and the master, before and after standardization ($n = 20$). Before STD = before standardization; After STD = after standardization; Lab = laboratory. Bentley (Chaska, MN), Delta (Drachten, the Netherlands), and Foss (Hillerød, Denmark). .21	
Figure 2-5. Effect of standardization of a Delta (Drachten, the Netherlands) slave spectra on differences with master spectra and on fat predictions. (A) Spectra of the master and a Delta slave instrument before (left) and after standardization (right), from 926 to 3,008 cm^{-1} . (B) Differences between spectra of the master and a Delta slave instrument before (left) and after standardization (right), from 926 to 3,008 cm^{-1} . (C) Comparison of fat predictions by the master and a Delta slave before (left) and after (right) standardization. RMSE = root mean square error. T = transmittance.	23
Figure 2-6. Effect of standardization of a Bentley (Chaska, MN) slave spectra on differences with master spectra and on fat predictions. (A) Spectra of the master and a Bentley slave instrument before (left) and after standardization (right), from 926 to 3,008 cm^{-1} . (B) Differences between spectra of the master and a Bentley slave instrument before (left) and after standardization (right), from 926 to 3,008 cm^{-1} . (C) Comparison of fat predictions by the master and a Bentley slave before (left) and after (right) standardization. RMSE = root mean square error. T = transmittance.	24
Figure 2-7. Effect of standardization of a Foss (Hillerød, Denmark) slave spectra on differences with master spectra and on fat predictions. (A) Spectra of the master and a Foss slave instrument before (left) and after standardization (right), from 926 to 3,008 cm^{-1} . (B) Differences between spectra of the master and a Foss slave instrument before (left) and after standardization (right), from 926 to 3,008 cm^{-1} . (C) Comparison of fat predictions by the master and by a Foss slave before (left) and after (right) standardization. RMSE = root mean square error. T = transmittance.	25
Figure 3-1. Principal component analysis (PCA) of the spectra of all instruments before standardization ($n = 66$). The PCA is based on the common analysis of 5 standardization samples in triplicate after selection of 212 informative wave numbers and a first derivative. Plot of principal components (PC) 5 and 6 summarizes the spectral variability between instruments. Each color–symbol association represents an individual instrument.....	38

Figure 3-2. Principal component analysis (PCA) of the spectra of all instruments, including the master, before and after standardization (n = 66). The PCA is based on the common analysis of 5 standardization samples in triplicate after selection of 212 informative wave numbers and a first derivative. Plot of principal components (PC) 5 and 6 summarizes the spectral variability between instruments..... 39

Figure 3-3. Comparison of methane predictions (g/d) by the master and by 3 slaves from the 3 brands before (left) and after (right) standardization. Dashed line is $y = x$; continuous line is the regression line between slave and master predictions. RMSE = root mean squared error..... 40

Figure 3-4. Box plot representation of PUFA predictions for a sample of the interlaboratory study analyzed on 66 instruments for master spectra, nonstandardized spectra, and spectra after standardization after removing aberrant values. 42

Figure 3-5. Plots representing the link between the quality of the model (coefficient of determination of cross-validation; R^2_{cv}) and the performance of the transfer by piecewise direct standardization [relative root mean squared error (RMSE) between slaves and master predictions after standardization] from the 5 interlaboratory study samples analyzed on 66 instruments. CH_4 = methane emitted by dairy cows; PUFA = PUFA in milk; FCY = fresh cheese yield; Fat = fat model used in Grelet et al. (2015). 45

Figure 4-1. Effect of editing the acetone data set on the distribution of values: (A) Original data set after preliminary editing (n = 548), (B) data set after logarithmic transformation and removing randomly low values (n = 224). 60

Figure 4-2. Plot of BHB reference values from flow injection analysis and BHB values predicted from Fourier transform mid-infrared analysis (A) for the cross-validation (n = 325) and (B) for the validation (n = 108). 62

Figure 4-3. Plot of acetone reference values from flow injection analysis and acetone values predicted from Fourier transform mid-infrared analysis (A) for the cross-validation (n = 168) and (B) for the validation (n = 56). 64

Figure 4-4. Plot of citrate reference values from flow injection analysis and citrate values predicted from Fourier transform mid-infrared analysis (A) for the cross-validation (n = 380) and (B) for the validation (n = 126). 66

Figure 5-1. Scheme of the organization of monthly standardization..... 73

Figure 5-2. Principal component analysis of the spectra of the 5 interlaboratory study samples analyzed on 66 instruments, before and after standardization, and projected on the fatty acid dataset containing spectra from 1,822 individual milk samples. All the spectral data were plotted after selection of 212 informative wave numbers. Principal components (PC) 6 and 7 reported on the picture. 76

Figure 5-3. Plots representing the link between the quality of the model (coefficient of determination of cross-validation; R^2_{cv}) and the performance of the transfer by piecewise direct standardization [relative root mean squared error (RMSE) between slaves and master predictions after standardization] from the 5 interlaboratory study

samples analyzed on 66 instruments. The different models originate from Soyeurt et al. (2011), Soyeurt et al. (2009), Vanlierde et al. (2016), Grelet et al. (2016), Colinet et al. (2015) and Soyeurt et al. (2007).78

Figure 5-4. Plots representing the link between the number of latent variables within the PLS models and the performance of the transfer by piecewise direct standardization [relative root mean squared error (RMSE) between slaves and master predictions after standardization] from the 5 interlaboratory study samples analyzed on 66 instruments. The different models originate from Soyeurt et al. (2011), Soyeurt et al. (2009), Vanlierde et al. (2016), Grelet et al. (2016), Colinet et al. (2015) and Soyeurt et al. (2007).79

Figure 5-5. Plot of fat predictions over time originating from a fat model applied on non-standardized spectra from an instrument analysing a common batch of UHT milk during a period of 3 months.....80

Figure 6-1. Example of valorization of raw predictions (milk BHB, acetone, oleic acid and fat/protein ratio) into usable information by the « Cétolait » tool developed by the Walloon Breeding Association (AWE).86

Figure A-1. Principal component analysis of the spectra of the 5 interlaboratory study samples analyzed on 66 instruments, before and after standardization, and projected on the fatty acids dataset containing spectra from 1,822 individual milk samples. All the spectral data were plotted after selection of 212 informative wave numbers. Principal components (PC) 1 and 2 reported on the picture.....92

Figure A-2. Principal component analysis of the spectra of the 5 interlaboratory study samples analyzed on 66 instruments, before and after standardization, and projected on the fatty acids dataset containing spectra from 1,822 individual milk samples. All the spectral data were plotted after selection of 212 informative wave numbers. Principal components (PC) 3 and 4 reported on the picture.92

Figure A-3. Principal component analysis of the spectra of the 5 interlaboratory study samples analyzed on 66 instruments, before and after standardization, and projected on the fatty acids dataset containing spectra from 1,822 individual milk samples. All the spectral data were plotted after selection of 212 informative wave numbers. Principal components (PC) 4 and 5 reported on the picture.93

Tables

Table 2-1. Description of the Fourier transform mid-infrared instruments included in the study	18
Table 2-2. Means of statistical results from regressions between fat predictions by the Master and the slaves, before and after standardization of interlaboratory study 1 (n=20) ¹	21
Table 2-3. Root mean square error (RMSE) between fat predictions of master and slaves, averaged by brand, before and after PDS (n = 20) ¹	22
Table 2-4. Mean statistic results of comparison of fat prediction by the master and the slaves, before and after standardization of the second interlaboratory study spectra with coefficients from the first interlaboratory study (n = 20) ¹	26
Table 3-1. Calibration and cross-validation statistics of equations used	36
Table 3-2. Root mean squared error (RMSE) between master and slaves predictions, averaged by brand, before and after piecewise direct standardization (PDS; n = 66 instruments)	41
Table 3-3. Maximum root mean squared error (RMSE) between master and slaves predictions, sorted by brand, before and after piecewise direct standardization (PDS; n = 66 instruments)	41
Table 3-4. Accuracy and reproducibility of predictions within the network; comparison of predictions from master, nonstandardized, and standardized spectra, from samples of the interlaboratory study analyzed on 66 instruments using 3 different Fourier transform mid-infrared calibrations	43
Table 3-5. Reproducibility of predictions among the different brands after standardization; comparison of predictions from master and standardized spectra, from samples of the interlaboratory study analyzed on 66 instruments using 3 different Fourier transform mid-infrared calibrations	44
Table 4-1. Overview of calibration and validation statistics from various studies aiming at predicting acetone content in milk by Fourier transform mid-infrared spectrometry; reference method for quantifying acetone, number of samples used, root mean square error (RMSE), standard error of cross-validation (SECV), and R ² are presented ¹	57
Table 4-2. Characteristics of the 4 sample source	58
Table 4-3. Characteristics of the Fourier transform mid-infrared instruments used in the study ¹	59
Table 4-4. Descriptive statistics of the results from the reference analysis for BHB, acetone, and citrate	61

Table 4-5. Cross-validation and validation statistics for BHB, acetone, and citrate contents in milk ¹	63
Table 4-6. Results of classification of BHB predictions into 2 classes (% of samples; using a threshold of 0.200 mmol/L) for cross-validation and validation data sets....	64
Table 4-7. Classification of the acetone predictions into 2 classes (% of samples; using a threshold of 0.150 mmol/L) for cross-validation and validation data sets....	65
Table 5-1. Overview of various studies aiming to standardize spectrometers using PDS method.....	74

Abbreviations

A: absorbance
BHB: β -hydroxybutyrate
CH₄: methane
cm⁻¹: 1/centimetre, wavenumber unit
d: day
DIM: days in milk
FCY: fresh cheese yield
FT-MIR: Fourier-Transform Mid Infrared
g: gramme
GH: standardized Mahalanobis distance
IDF: International Dairy Federation
l: litre
LV: latent variables
Max: maximum
Min: minimum
MIR: Mid Infrared
ml: millilitre
mmol: millimole
MRO: milk recording organization
N: number of samples
NADH: nicotinamide dinucleotide
NEB: negative energy balance
NIR: near Infrared
nm: nanometer
PC: principal component
PCA: principal component analysis
PDS: piece-wise direct standardization
PLF : precision livestock farming
PLS: partial least square
PLS-DA: partial least square discriminant analysis
PUFA: poly unsaturated fatty acids
R² : determination coefficient
R²_{cv} : determination coefficient of cross-validation
R²_v : determination coefficient of validation

RMSE: root mean square error
RMSE_{cv}: root mean square error of cross-validation
RMSE_v: root mean square error of validation
RPD: ratio of SD to SEC
RPD_{cv}: ratio of SD to SECV
SD: standard deviation
SEC: standard error of calibration
SECV: standard error of cross-validation
SEL: standard error of laboratory
STD: standardization
SVM: Support Vector Machine
T: transmittance
UHT: ultra high temperature processing

Chapter 1

General introduction

Chapter 1 : General introduction

Current context involving the need for new phenotypes

This study has been conducted with the objective of optimizing the use of FT-MIR analysis of cow milk under routine conditions, within a network of labs in order to enable the development of new indicators for the dairy sector, and particularly for dairy farmers. It takes place in a context where the dairy production has been facing profound changes in recent decades prompting the need for new herd management tools.

Firstly, a reduction in the number of dairy farm as well as an increase of herd size and milk production per cow have been reported in the main dairy regions across the world (Gargiulo et al., 2018). Monitoring and management of cows are becoming more challenging and complex, requiring enhanced management ability from dairy farmers (Bewley, 2016). At the same time administrative tasks have also increased, with the combined result of making the adequate monitoring of individual animals extremely difficult for farmers (Berckmans, 2006). Additionally, due to demographic evolution, today's farmer families have fewer children and lower involvement of adult children in the family farm. Consequently, more farms in developed countries depend on nonfamily labour (Barkema et al., 2015). However, the necessary skills and competences to monitor and manage dairy herds differ considerably among farm workers, and availability of competent laborers is often problematic (Winsten et al., 2010). Insufficient training of farm workers could lead to failure in the detection of health problems (Barkema et al., 2015). The limited availability of trained workers, or the cost linked to hiring labor with the necessary management skills, combined with the increase in herd size, are changes that make the management of herds more complex and time consuming, potentially affecting the profitability of dairy farms (Gargiulo et al., 2018).

Secondly, the European Union has abandoned the quota system, and due to market pressures farmers have often found fluctuating, typically declining, milk price to be aligned with global price (Barkema et al., 2015). The profitability of dairy farms is also endangered by this volatile (and often too low) price. In this context, optimal management of cows to prevent or detect diseases, improve animal health and welfare, reproduction and production, is considered a viable strategy for increasing economic outcome (Moyes et al., 2013).

Thirdly, society is becoming increasingly concerned with animal welfare, healthy animals and the environmental impact of livestock productions (Berckmans, 2006). For Belgian consumers, animal welfare comes second, after food safety, when deciding what food to purchase (Vanhonacker et al., 2007). To preserve their 'social license to operate', farmers need to provide objective evidence of the welfare and health of their animals (Barkema et al., 2015).

Consequently, there is a growing need for monitoring and management tools to satisfy these different necessities. In modern agriculture, digitalized monitoring techniques are becoming more and more important to support farmers in managing their production processes (Berckmans, D., 2006). In an animal production context,

these techniques have been grouped under the term Precision Livestock Farming (**PLF**). For Wathes et al. (2008), the background for the development of PLF is an environment where profits are marginal and the availability of labor unit with the necessary management skills is limited or too expensive. First PLF applications were devoted to monitoring the growth of housed pigs and poultry, but PLF can be applied to any species, including animals grown in extensive systems (Frost, 2001). In dairy production, potential field of applications are the detection of diseases and sub-optimal health status, improvement of health and welfare, reproduction, feed efficiency and environmental impact, among others. The objectives of PLF are to provide information on these topics, offer time saving opportunities to farmers, objectify the measures in favour of less emotional decisions, obtain more precise information or information previously unavailable, and benefit from additional phenotypes of interest without inducing animal stress due to invasive techniques. To cope with these objectives and farmers' needs, the technology being used must be available under routine condition, economically affordable, applicable on a large scale and provide quick results to farmers.

FT-MIR analysis of milk

Among all the potential technologies, analysis of milk by Fourier-Transform Mid Infrared (**FT-MIR**) spectrometry is already being used in main dairy areas of the world for the purposes of quantifying fat, protein and lactose contents, milk payment and milk recording (Soyeurt et al., 2006; Dehareng et al., 2012). FT-MIR allows a fast, cost-effective and non-destructive quantification of milk chemical properties in order to avoid reference methods, which are usually tedious, expensive and time consuming.

Spectroscopy is defined as the study of a matrix through the electromagnetic radiation with which it interacts or that it produces. The history of spectroscopy is linked to the analysis of scattered light through a prism as experienced by Isaac Newton. The discovery of infrared radiation is attributed to Herschel who analysed the temperature of the diffracted sunlight just beyond the red end of visible spectrum (Herschel, 1800). Spectrometry is the measurement of electromagnetic radiation as a means of obtaining information about the systems and their components (IUPAC, 2007).

In a spectrometer, milk samples are crossed by infrared radiations at different wavelengths (nm). In mid infrared (**MIR**), these infrared radiations are more often characterized by a frequency, expressed in wavenumbers (cm^{-1}), which are an inverted function of wavelengths. The analysis is based on the interactions between light and chemical bonds within the molecules constituting milk. Two atoms constituting a molecule vibrate at a precise frequency and can interact with infrared rays having the same frequency and absorb energy of the light (Rouessac et al., 2016). Covalent bonds can absorb light at different wavenumbers because different vibrations as stretching and bending take place. Consequently, matter absorbs light at precise frequencies according to its chemical composition. This is physically expressed by the Beer-Lambert law stating that absorbance is linearly linked to the coefficient of molar

absorption, which is dependent of the wavenumber, pathlength and concentration of the analyte of interest.

Spectrometers usually contain a light emitting source, a system designed to select wavenumbers of interest, a cell containing the sample, and a light intensity measuring detector. Interactions between light and chemical bonds within milk molecules are measured on the basis of their transmittance (**T**) which is the ratio of light emitted by the source (**I₀**) to the light received by the detector after passing through the sample (**I**), or considering the absorbance (**A**) of the sample which is a logarithmic function of the transmittance.

$$T = I/I_0 \quad A = -\log(T)$$

Milk is analysed in the near-IR region, from 800 to 2,500 nm (or 4,000 to 12,500 cm^{-1}) but more often in the MIR region, from 2,500 to 25,000 nm (or 400 to 4,000 cm^{-1}) as the MIR region contains the fundamental vibrations and thus provides better results.

The first MIR instruments dedicated to milk analysis measured the absorption of infrared energy at specific frequencies using a monochromator. Fat, protein and lactose contents were estimated by measuring absorbance of carbonyl groups in the ester linkages, the peptide linkages between amino acids and the O-H bonds, respectively. A second generation of infrared instrumentation used optical filters for wavenumber selection (Grappin and Jenet, 1976).

With the combined use of interferometers and Fourier transform algorithms, it has been possible to analyse milk samples in a large range of the infrared region. Measurements in the MIR region are taken at thousands of different wavenumbers. These thousands of absorbance (or transmittance) values obtained within this range of wavenumbers constitute a MIR spectrum (Figure 1-1).

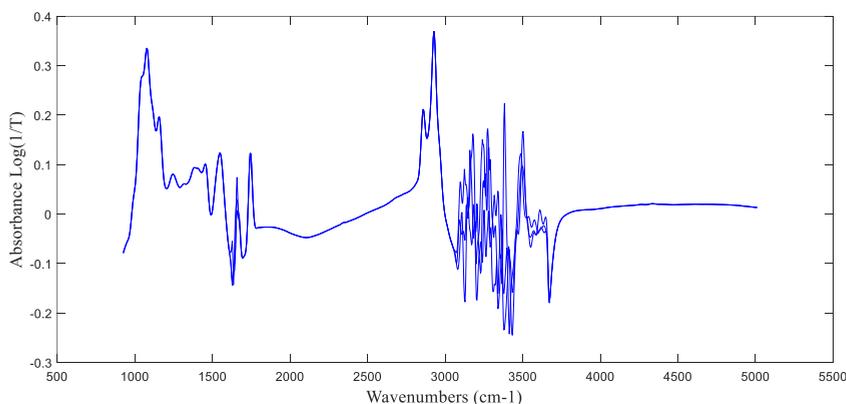


Figure 1-1. Spectrum of a milk sample

Absorbance values at these different wavenumbers are used in equations to predict the content of milk components of interest (Williams and Norris, 1987). This is an

extension of the Beer–Lambert law linking absorbance to concentration. The use of the entire spectra, and consequently a higher amount of information, has increased the potential of this technology to provide accurate predictions while in most milk labs FT-MIR spectrometers have replaced filter instruments.

As previous technologies, FT-MIR instruments have been mainly used to produce fat, protein, and lactose predictions. Fat and protein contents are currently used for milk payment in many countries and have been therefore used as phenotypes for genetic improvement. FT-MIR analysis of milk has also been used in herd management. The ratio fat/protein is still used as the main indicator for ketosis and acidosis detection (Koeck et al., 2014). In milk processing, protein composition is a good indicator of milk ability to be processed into cheese (Amenu & Deeth, 2007). Models predicting urea in milk and free fatty acids are also implemented on manufacturer's instruments. Urea in milk is useful for diet management and especially to estimate nitrogen efficiency in rumen fermentation (Hof et al., 1997). Free fatty acids are considered as an indicator of lipolysis in milk and associated organoleptic issues (Deeth H.C., 2006).

More recently, the use of the entire spectra in combination with advance of informatics and multivariate statistics have allowed the development of models that delve deeper into the composition of fine milk. Models have been built for the determination of fatty acids profiles (Soyeurt et al., 2006), lactoferrin (Soyeurt et al., 2007), minerals (Soyeurt et al., 2009), protein composition (Bonfatti et al., 2011), ketone bodies (Van Knegsel et al., 2010). Other studies have used FT-MIR spectra to build calibrations predicting technological properties of milk such as milk acidity (De Marchi et al., 2009), ability to coagulate, and firmness of curd or cheese yield (Dal Zotto et al., 2008; Colinet et al., 2015). Recent researches also focused on the prediction of blood composition through milk spectra, particularly for β -hydroxybutyrate (**BHB**) (Broutin, 2014). FT-MIR spectrum of milk has even been considered as a reflection of cows' 'status', and models were developed to predict methane emissions of dairy cows (Dehareng et al., 2012), likelihood of conception (Hempstalk et al., 2015), body energy status (McParland et al., 2011) or energy intake and efficiency (McParland et al., 2014).

All these FT-MIR models aim to provide more information for the dairy sector, in order to optimize and rationalize the management of herds by providing valuable indications on the status of individual lactating cows. Additionally, these models are of great interest for genetic studies, because large amount of phenotypes of interest are needed. Finally, in the framework of milk processing, models predicting the technological properties of milk before the beginning of the process would be beneficial to production management.

Standardization

Most of these models have been developed by scientific institutions in the framework of research projects, and very few of them have been really used by field and commercial organizations devoted to providing management tools to farmers. Actually, there is a suboptimal use of technology designed to provide useful information to farmers due to the weak transfer mechanism between research and field

organizations (De Marchi et al., 2014). The main reasons for this are the difficulty to obtain robust models providing correct predictions on real field conditions (Gengler et al., 2016), and the difficulty to share and transfer models between instruments (Wang et al., 1991).

The development of robust models is indeed very expensive, time consuming and practically difficult as it implies to constitute calibration datasets based on a large number of samples, that require high variability in terms of breeds, diets, management systems, geographic origins, reference and spectral data. When applying the models, the potential variability that could be met under field conditions must be covered. Combinations of different datasets containing complementary information is therefore essential (De Marchi et al., 2014). In order to save resources and obtain reliable tools there is an urgent need to merge datasets and develop and share common models.

However, this is not directly possible due to the specific instrumental response produced by each instrument (Feudale et al., 2002). These differences originate from the specific physical characteristic of each brand and models, using different technologies, materials, and optical paths... The main manufacturers of FT-MIR spectrometers dedicated to milk analysis are Foss (Hillerød, Denmark), Bentley (Chaska, MN, USA) and Delta Instruments (Drachten, the Netherlands). These three brands provide spectra with different ranges, from 925.66 to 5,010.15 cm^{-1} for Foss instruments, from 649.03 to 3,998.59 cm^{-1} for Bentley instruments and from 397.31 to 4,000 cm^{-1} for Delta instruments. Additionally, instruments originating from the same models are also individually characterized by specific spectral responses due to different uses, deterioration of pieces, environmental factors, maintenance operations or piece replacements. Finally, these factors added to electronic drifts and detector instability lead to the instability of each individual spectrometer response over time (Bonfatti et al., 2017).

This heterogeneity prevents merging datasets from different instruments: a model developed on instrument A provides biased predictions if transferred to instrument B (Rodriguez et al., 2011). This is a practical limiting factor to the development and use of FT-MIR models to provide management tools for farmers under routine conditions.

In terms of fat, protein and lactose content delivery, this spectral heterogeneity has traditionally been handled by the dairy sector by performing slope and bias adjustments on the predictions after analysing standard milks with reference values obtained from wet chemistry (ISO 9622:2013 | IDF 141:2013). In the context of providing new phenotypes for the management of dairy herds, as energy or health status, it is too expensive or even impossible to constitute standard samples with known values of the variable of interest (Bonfatti et al., 2017). Given that predictions of developed models cannot be corrected through slope and bias adjustment, the creation of robust models predicting new phenotypes of interest is suboptimal and the transfer of such models among different spectrometers often leads to incorrect predictions.

Standardization of the FT-MIR spectra in order to harmonize directly the spectral responses of instruments appears as a potential solution (Rodriguez et al., 2011). This could allow the merging of datasets and the transfer of models through the dairy sector. Harmonization of spectral responses would serve to optimize and enhance

practically the use of FT-MIR spectrometry as an efficient technology designed to provide management information for dairy farms.

Such standardization methods have been historically developed in the context of near infrared (NIR) analysis of agricultural products (Wang et al., 1991). These methods aimed to solve two situations where the models were rendered invalid due to the 'format' of spectral responses (Feudale et al., 2002). First, as mentioned earlier, the transfer of a model developed on a specific instrument and then transferred to a second instrument was problematic given the differences in spectral responses. The second issue was caused by the change in time of the spectral responses of individual instruments, resulting in models not adapted to this changed spectral response. Although standardization methods have mostly been developed for NIR data, it has also been applied to transfer UV-Visible, fluorescence and Raman spectral data (Feudale et al., 2002). However, these methods have not been tested for the harmonization of FT-MIR spectra of milk in order to transfer models predicting new phenotypes of interest for management of dairy herds.

Research objectives

In this context, this study aims as a first step to evaluate the performance of harmonizing milk FT-MIR spectral data, using Piece-wise Direct Standardization method, in order to transfer a robust model predicting fat content in milk, from instrument to instrument (**Chapter 2**). In a second step, the effect of this method on data harmonization is evaluated further, to assess the possibilities of using in practice innovative models of interest for the dairy sector in a network of instruments (**Chapter 3**). This is done by analysing, on 66 spectrometers, the impact on spectral reproducibility, the transfer of models with low robustness and the accuracy and reproducibility of predictions within the network. After insuring the possibility of using spectral data in optimal conditions, the final step aims to concretize the development of models that provide information on cow status so they may be used by dairy farmers as management tools (**Chapter 4**). Emphasis is placed on the biomarkers of energy status and ketosis associated risks. Indeed, in early lactation negative energy balance has been identified as one of the major factors influencing fertility and health of dairy cows (Collard et al., 2000; Butler, 2003). This is done with the objective of detecting suboptimal status before becoming clinical cases. Concretely, this third part aims to develop models to predict citrate, reflecting early energy imbalance, and acetone and BHB as indicators of (sub)clinical ketosis.

References

- Amenu B., & Deeth H. C., 2007. The impact of milk composition on cheddar cheese manufacture. *Aust. J. Dairy Technol.*, 62:3, 171.
- Barkema H. W., Von Keyserlingk M. A. G., Kastelic J. P., Lam T. J. G. M., Luby C., Roy J. P., & Kelton D. F., 2015. Invited review: Changes in the dairy industry affecting dairy cattle health and welfare. *J. Dairy Sci.*, 98:11, 7426-7445.
- Belay T.K., Dagnachew B.S., Kowalski Z.M., Ådnøy T., 2017. An attempt at predicting blood β -hydroxybutyrate from Fourier-transform mid-infrared spectra of

milk using multivariate mixed models in Polish dairy cattle. *J. Dairy Sci.* 100, 6312-6326.

Berckmans D., 2006. Automatic on-line monitoring of animals by precision livestock farming. *Livestock production and society*, 287.

Bewley J., 2016. Opportunities for monitoring and improving animal welfare using precision dairy monitoring technologies. *J. Anim. Sci.* 94:2:11.

Bonfatti V., Di Martino G., Carnier P., 2011, Effectiveness of mid-infrared spectroscopy for the prediction of detailed protein composition and contents of protein genetic variants of individual milk of Simmental cows. *J. Dairy Sci.* 94:12:5776-5785.

Bonfatti V., Fleming A., Koeck A., and Miglior F., 2017. Standardization of milk infrared spectra for the retroactive application of calibration models. *J. Dairy Sci.* 100:2032–2041.

Broutin P., 2015. Determination of the concentration of a component in one fluid of an animal by spectroscopic analysis of another fluid. Patent. Pub. No: WO/2015/055966. International Application No: PCT/FR2014/052650. Publication Date: 23.04.2015. Retrieved on 6 May 2018 from <https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2015055966>.

Butler W.R., 2000. Nutritional interactions with reproductive performance in dairy cattle. *Anim. Reprod. Sci.* 60, 449-457.

Colinet F., Troch T., Baeten V., Dehareng F., Dardenne P., Sindic M., Gengler N., 2015. Genetic variability of MIR predicted milk technological properties in Walloon dairy cattle. In Book of Abstracts of the 66th Annual Meeting of the European Federation of Animal Science (p. 400). Wageningen Academic Publishers.

Collard B.L., Boettcher P.J., Dekkers J.C., Petitclerc D., and Schaeffer L.R., 2000. Relationships between energy balance and health traits of dairy cattle in early lactation. *J. Dairy Sci.*, 83, 2683–2690.

Dal Zotto R., De Marchi M., Cecchinato A., Penasa M., Cassandro M., Carnier P., Bittante G., 2008. Reproducibility and repeatability of measures of milk coagulation properties and predictive ability of mid-infrared reflectance spectroscopy. *J. Dairy Sci.* 91(10): 4103-4112.

De Marchi M., Fagan C., O'donnell C., Cecchinato A., Dal Zotto R., Cassandro M., Bittante G., 2009. Prediction of coagulation properties, titratable acidity, and pH of bovine milk using mid-infrared spectroscopy. *J. Dairy Sci.*, 92:1 :423-432.

De Marchi M., Toffanin V., Cassandro M., and Penasa M., 2014. Invited review: Mid-infrared spectroscopy as phenotyping tool for milk traits. *J. Dairy Sci.* 97:1171–1186.

Deeth H. C., 2006. Lipoprotein lipase and lipolysis in milk. *Int. Dairy J.*, 16:6, 555-562.

Dehareng F., Delfosse C., Froidmont E., Soyeurt H., Martin C., Gengler N., Vanlierde A. and Dardenne P., 2012. Potential use of milk mid-infrared spectra to predict individual methane emission of dairy cows, *Animal*, 10:1694-701.

Feudale R. N., Woody N. A., Tan H., Myles A. J., Brown S. D., & Ferré J., 2002. Transfer of multivariate calibration models: a review. *Chemometr. Intell. Lab. ,* 64:2, 181-192.

Frost A. R., 2001. An overview of integrated management systems for sustainable livestock production. *BSAP Occasional Publication*, 28, 45-50.

Gargiulo J. I., Eastwood C. R., Garcia S. C., & Lyons N. A., 2018. Dairy farmers with larger herd sizes adopt more precision dairy technologies. *J. Dairy Sci.* 101:6, 5466-5473.

Gengler N., Soyeurt H., Dehareng F., Bastin C., Colinet F., Hammami H., Vanrobays M.L., Lainé A., Vanderick S., Grelet C., Vanlierde A., Froidmont E., Dardenne P., 2016. Capitalizing on fine milk composition for breeding and management of dairy cows. *J. Dairy Sci.* 99:5, 4071-4079.

Grappin R., Jeunet R., 1976. Essais de l'appareil Milko-Scan 300 utilisé pour le dosage en série de la matière grasse et des protéines du lait, *Lait*, 56:558,498-520.

Hempstalk K., McParland S., Berry D.P., 2015. Machine learning algorithms for the prediction of conception success to a given insemination in lactating dairy cows. *J. Dairy Sci.*, 98:8, 5262-5273.

Herschel W., 1800. XIV. Experiments on the refrangibility of the invisible rays of the sun. *Philosophical Transactions of the Royal Society of London*, 90, 284-292.

Hof G., Vervoorn M. D., Lenaers P. J., & Tamminga S., 1997. Milk urea nitrogen as a tool to monitor the protein nutrition of dairy cows. *J. Dairy Sci.* 80:12, 3333-3340.

ISO 9622:2013 | IDF 141:2013 – Whole milk – Determination of milk fat, protein and lactose content – Guidance on the operation of mid-infrared instruments. ISO, Geneva, Switzerland.

IUPAC. *Compendium of Chemical Terminology*, 2nd ed. (the "Gold Book"). Compiled by McNaught A.D., and Wilkinson A., Blackwell Scientific Publications, Oxford. 1997. XML on-line corrected version: <http://goldbook.iupac.org> (2006-) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. ISBN 0-9678550-9-8. <https://doi.org/10.1351/goldbook>.

Koeck A., Jamrozik J., Schenkel F. S., Moore R. K., Lefebvre D. M., Kelton D. F., & Miglior F., 2014. Genetic analysis of milk β -hydroxybutyrate and its association with fat-to-protein ratio, body condition score, clinical ketosis, and displaced abomasum in early first lactation of Canadian Holsteins. *J. Dairy Sci.* , 97:11, 7286-7292.

McParland S., Banos G., Wall E., Coffey M.P., Soyeurt H., Veerkamp R.F., Berry D.P., 2011., The use of mid-infrared spectrometry to predict body energy status of Holstein cows, *J. Dairy Sci.* 95:3,1222-1239.

McParland S., Lewis E., Kennedy E., Moore S.G., McCarthy B., O'Donovan M., Butler S.T., Pryce J.E., Berry D.P., 2014. Mid-infrared spectrometry of milk as a predictor of energy intake and efficiency in lactating dairy cows, *J. Dairy Sci.* 97:9,5863-5871.

Moyes K. M., Larsen T., & Ingvarsten K. L., 2013. Generation of an index for physiological imbalance and its use as a predictor of primary disease in dairy cows during early lactation. *J. Dairy Sci.*, 96:4, 2161-2170.

Pralle R. S., Weigel K. W., & White H. M., 2018. Predicting blood β -hydroxybutyrate using milk Fourier transform infrared spectrum, milk composition, and producer-reported variables with multiple linear regression, partial least squares regression, and artificial neural network. *J. Dairy Sci.* 101:5,4378-4387.

Rodriguez J. D., Westenberger B. J., Buhse L. F., and Kauffman J. F., 2011. Standardization of Raman spectra for transfer of spectral libraries across different instruments. *Analyst* 136:4232–4240.

Rouessac F., Rouessac A., Cruché D., Duverger-Arfulso C., Martel A., 2016. *Analyse chimique-8e éd.: Méthodes et techniques instrumentales modernes*. Dunod. Paris, ISBN 978-2-10-074688-0.

Soyeurt H., Dardenne P., Dehareng F., Lognay G., Veselko D., Marlier M., Gengler N., 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. *J. Dairy Sci.*, 89 :9, 3690-3695.

Soyeurt H., Colinet F., Arnould V., Dardenne P., Bertozzi C., Renaville R., Portetelle D., Gengler N., 2007. Genetic Variability of Lactoferrin Content Estimated by Mid-Infrared Spectrometry in Bovine Milk, *J. Dairy Sci.* 90:9,4443-4450.

Soyeurt H., Bruwier D., Romnee J.-M., Gengler N., Bertozzi C., Veselko D., Dardenne P., 2009. Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry, *J. Dairy Sci.* 92:6,2444-2454.

Van Kneysel A.T.M., van der Drift S.G.A, Horneman M., de Roos A.P.W., Kemp B., Graat E.A.M., 2010. Short communication: Ketone body concentration in milk determined by Fourier transform infrared spectroscopy: Value for the detection of hyperketonemia in dairy cows, *J. Dairy Sci.* 93:7,3065-3069.

Vanhonacker F., Verbeke W., Van Poucke E., and Tuytens F. A., 2007. Segmentation based on consumers' perceived importance and attitude toward farm animal welfare. *Int. J. Soc. Food Agric.* 15:91–101.

Vanlierde A., Vanrobays M.L., Gengler N., Dardenne P., Froidmont E., Soyeurt H., McParland S., Lewis E., Deighton M.H., Mathot M., and Dehareng F., 2016. Milk mid-infrared spectra enable prediction of lactation-stage dependent methane emissions of dairy cattle within routine population-scale milk recording schemes. *Animal Production Science*, 2016, 56, 258–264.

Wang Y., Veltkamp D.J, Kowalsky B.R., 1991. Multivariate instrument standardisation, *Anal. Chem.*, 63:23,2750–27.

Wathes C. M., Kristensen H. H., Aerts J. M., & Berckmans D., 2008. Is precision livestock farming an engineer's daydream or nightmare, an animal's friend or foe, and a farmer's panacea or pitfall? *Comput. electron. agr.*, 64:1, 2-10.

Williams P., Norris K., 1987. *Near-Infrared technology in the agricultural and food industries*. American association of cereal chemists, Inc. St. Paul, Minnesota 55121 USA.

Winsten J. R., Kerchner C. D., Richardson A., Lichau A., and Hyman J. M., 2010. Trends in the Northeast dairy industry: Large-scale modern confinement feeding and management-intensive grazing. *J. Dairy Sci.* 93:1759–1769.

Chapter 2

**Standardization of milk mid-infrared
spectra from a European dairy network**

Introduction to chapter 2

Numerous standardization methods have been reported in scientific literature. The common points of all these methods are the analysis of common samples performed through various instruments, to be used as a basis for the transfer procedure, followed by the adjustment of the responses of interest into a response considered as the reference.

A first type of standardization requires considering the adjustment of the final predictions provided by each model for a given instrument. In this scheme, the values predicted by secondary instruments are corrected to fit the reference values, which could be chemical values from wet chemistry or the prediction provided by an instrument considered as a reference. Predictions of the models are adjusted a posteriori using a univariate slope and bias correction obtained from a linear regression. This is the method classically used in the dairy sector to ensure the quality of fat, protein and lactose predictions (ISO 9622:2013 | IDF 141:2013). However, this method involves the use of samples with known reference values, which is difficult, expensive or even impossible to obtain in the case of some phenotypes predicting cow status (Bonfatti et al., 2017).

A second type of method is based on the standardization of the calibration datasets, and consequently of the models, in order to transfer them to a secondary instrument. Some methods, such as the classical or inverse model standardization, have been described by Wang et al. (1991). However, both methods also need the use of samples with known reference values (Feudale et al., 2002).

A third type of method focuses on standardizing the spectral response of instruments. Thus, models remain unique and unchanged, and each spectral response is corrected to match the spectral response of a reference instrument on which the models are developed. A first method called univariate standardization performs a regression at each wavelength between absorbance values of master and secondary instruments (Shenk and Westerhaus, 1993). Univariate standardization only allows correcting for intensity differences and not wavenumber axes. Hence, there is a need to combine with a wavenumber adjustment (Feudale et al., 2002). Another method called direct standardization uses the entire spectrum of the secondary instrument to match the spectral response of the reference instrument at each wavenumber (Wang et al., 1991). The main problem with this method is that a significant amount of chemical information can be modelled, and if variation happens within the sample analysed by different instruments it can be integrated into the transfer model and disrupt the standardization performances (Feudale et al., 2002). To cope with this, the Piece-wise Direct Standardization method (**PDS**), presented by Wang et al. (1991) takes into account that in real spectroscopic data the spectral response at one wavenumber is more likely to be related to the response at nearby spectral points. The response of the reference instrument at a given wavenumber is reconstructed based on the measurement on a small window around this wavenumber on the secondary instrument. Among all the tested methods, PDS provides the best results, and can reduce the differences between instruments even when a small number of samples were used for the transfer (Wang et al., 1991). The better results obtained through

PDS can be attributed to its local and multivariate nature, allowing to correct for wavenumber shifts, intensity differences and peak broadening (Feudale et al., 2002).

Consequently, the PDS seems the more appropriate method in this context to standardize mid-infrared spectrometers dedicated to milk analysis. This technique has been mostly evaluated with NIR data based on standard samples constituted of cereals, forages or oil. Accordingly, there is a need to validate the use of this method when applied on milk spectra acquired in mid-infrared range. Moreover, the nature and the number of common standard samples used for transfer procedures has a high impact on the performance of the method (Wang et al., 1991). There is also a need to test whether the samples classically used to perform slope and bias correction and produced following the **IDF** (International Dairy Federation) norm 141, ISO/DIS 9622, can be used as a physical common reference to apply the method. To achieve this practically, the following chapter aims to evaluate the performance of harmonizing milk FT-MIR spectral data, using Piece-wise Direct Standardization method and a set of standard raw milk samples, with the objective of transferring a robust fat model from instruments to instruments.

References

Bonfatti, V., Fleming A., Koeck A., and Miglior F., 2017. Standardization of milk infrared spectra for the retroactive application of calibration models. *J. Dairy Sci.* 100:2032–2041.

Feudale R. N., Woody N. A., Tan H., Myles A. J., Brown S. D., & Ferré J. 2002. Transfer of multivariate calibration models: a review. *Chemometr. Intell. Lab. ,* 64(2), 181-192.

ISO 9622:2013 | IDF 141:2013 – Whole milk – Determination of milk fat, protein and lactose content – Guidance on the operation of mid-infrared instruments. ISO, Geneva, Switzerland.

Shenk J., & Westerhaus M. 1993. Comments on standardisation: Part 2. *NIR news*, 4(5), 13-15.

Wang Y., Veltkamp D.J, Kowalsky B.R., 1991. Multivariate instrument standardisation, *Anal. Chem.*, 63(23)2750–27.

Chapter 2: Standardization of milk mid-infrared spectra from a European dairy network

C. Grelet¹, J. A. Fernández Pierna¹, P. Dardenne, V. Baeten, and F. Dehareng

Walloon Agricultural Research Center, Valorisation of Agricultural Products Department,
24 Chaussée de Namur, 5030 Gembloux, Belgium

J. Dairy Sci. 98 :2150–2160 <http://dx.doi.org/10.3168/jds.2014-8764>

© American Dairy Science Association®, 2015

¹ These authors contributed equally

Abstract

The goal of this study was to find a procedure to standardize dairy milk mid-infrared spectra from different Fourier transform mid-infrared spectrophotometers (different brands or models) inside a European dairy network to create new farm-management indicators (e.g., fertility, health, feed, environmental impact) based on milk infrared spectra. This step is necessary to create common spectral databases, allowing the building of statistical tools, to be used by all instruments of the network. The method used was piecewise direct standardization (PDS), which matches slave instrument spectra on master-instrument spectra. To evaluate the possibility of using common equations on different instruments, the PDS method was tested on a set of milk samples measured on each machine, and an equation predicting fat content of milk is applied on all. Regressions were performed between master and slaves fat predictions, before and after PDS. Bias and root mean square error between predictions were decreased after PDS, respectively, from 0.3781 to 0.0000 and from 0.4609 to 0.0156 (g of fat/100 mL of milk). The stability over time of these results was confirmed by an application of the coefficients created by PDS 1 mo later on the slave spectra. These preliminary results showed that the PDS method permits a reduction of the inherent spectral variability between instruments, allowing the merging of Fourier transform mid-infrared milk spectra from different instruments into a common database, the creation of new types of dairy farm management indicators, and the use of these common calibrations for all Fourier transform mid-infrared instruments of the European dairy network.

Key words: Fourier transform mid-infrared spectrometry, standardization, dairy milk, piecewise direct standardization

Introduction

This work is the first step of a project aiming to develop innovative farm-management web applications based on the use of Fourier transform mid-infrared (FT-MIR) spectrometry analysis of milk to enable a sustainable and profitable management of the milk production. Fourier transform mid-infrared spectrometry is the worldwide method of choice for composition and quality controls during routine liquid milk testing. It allows a fast, nondestructive quantification of milk chemical properties to avoid reference methods, which are usually tedious, expensive, and time consuming. In 1961, a patent application for a FT-MIR method determining fat, protein, and lactose in milk was introduced (Goulden, 1964). The first apparatus, an

IRMA (Infrared Milk Analyzer, Grubb Parsons, Newcastle upon Tyne, UK) using a monochromator, was based on the principle of measuring direct absorption of the infrared energy at specific frequencies by carbonyl groups in the ester linkages of the fat molecules, by peptide linkages between amino acids of protein molecules, and by the O-H groups in lactose molecules. A second generation of infrared instrumentation has adopted the change from wavenumber selection by diffraction grating to optical filters (Grappin and Jeunet, 1976) and was largely used by Central milk laboratory testing, where both tank milk and individual-cow samples were tested. Fourier transform mid-infrared supplies complementary chemical information and allows a high throughput with high sensitivity in a short response time from a very small quantity of sample (Ghosh and Jayas, 2009). In 1993, the first purpose-built FT-MIR instrument based on the Fourier transform infrared (FT-MIR) technology was marketed (Anadis MI-200; Asselain et al., 1996). With the introduction of the FT-MIR, new applications have been developed because of the use of the full spectrum of the sample. In this way, FT-MIR has been applied for the determination of more and more milk components such as proteins composition (Bonfatti et al., 2011), minerals (Soyeurt et al., 2009), ketone bodies (van Kneysel et al., 2010), lactoferrin (Soyeurt et al., 2007), and fatty acid profile (Rutten et al., 2009; Soyeurt et al., 2011). Then recent studies were performed using these milk components predicted by FT-MIR to predict physiological indicators of the animal (Friggens et al., 2007; Mohammed et al., 2011). In the context of this research project the FT-MIR spectrum is directly considered as a reflection of the state of the cows, avoiding the step of milk composition, to obtain indicators concerning fertility, health, environment, and feeding among others. Until now, only a few studies have been performed to show the potential of the entire FT-MIR spectra as an indicator of those parameters. Only recent studies have shown that predictions based on direct spectra are much more global, sensitive, and accurate than those based on milk components when they are predicted from FT-MIR. Dehareng et al. (2012) have shown that enteric methane was better predicted when directly working with FT-MIR spectra than the results based on fatty acid predictions. Also recently, the FT-MIR spectrum of milk was shown to be a good indicator of body energy status (McParland et al., 2011), energy intake and efficiency (McParland et al., 2014), and fertility diagnosis (Laine et al., 2013) in dairy cattle. This innovative approach of using FT-MIR spectroscopy needs the support of important spectral databases associated with reference values for each of the properties to be studied. For this reason, the OptiMIR project was built; it is a European Interreg project involving 6 countries and focuses on the development of prediction tools directly based on FT-MIR spectra. In this work, a large number of commercially available mid-infrared spectrometers (21) from different manufacturers (3) installed in different laboratories (10) located in different countries (3) were used. Because of differences of the instrumental responses between different FT-MIR spectrometers, spectra obtained on one instrument cannot readily be compared with a library acquired on a different instrument. Moreover, the use of calibration models developed on an instrument with FT-MIR spectra obtained on another instrument will usually lead to an increased uncertainty of the prediction model. This is a drawback when recalibrating an instrument or using a historical database. Therefore, spectral corrections adapted to each instrument (standardization procedures) are needed

(Rodriguez et al., 2011). One of the most common techniques for instrument standardization is the piecewise direct standardization (PDS) proposed by Wang et al. (1991). However, in previous studies it was mainly used with near-infrared spectra (Bouveresse and Massart, 1996) and was not tested for milk spectra. Then, the objective of this work was to demonstrate and validate the use of the PDS to standardize spectra from different models and manufacturers of instruments, by reducing the inherent instrument-to-instrument variability, within the dairy network, such that the milk spectra from all spectrometers (the slaves) can be compared with the milk spectra of a standard instrument (the master). In such a way, it should be possible to create and maintain international databases containing data collected by all the FT-MIR instruments and to relate them to chemical characteristics of the milk (e.g., protein, fat, fatty acids content among others) and to animal physiology (fertility, nutrition, health, and environment).

Materials and methods

Standardization

Standardization samples are measured on a master instrument and on a slave instrument, leading to response matrices M and S . The PDS method is based on the fact that the variation of spectroscopic data is limited to small spectral regions. In PDS, the response m_j measured at wavenumber j on the master instrument is related to the wavenumbers located in a small window (s_j) of size n around j (neighboring) measured on the slave instrument (Figure 1-1). The window (s_j) was composed of 5 wavenumbers and was the same for all instruments:

$$s_j = [S_{(j-n)}, \dots, S_{(j)}, \dots, S_{(j+n)}].$$

A regression using the principal component regression method is calculated between each spectral response on the master at wavenumber j and the corresponding window s_j on the slave. Vector b_j is the vector of transformation coefficients for the j th wavenumber, and b_{0j} is the offset term:

$$m_j = s_j b_j + b_{0j}.$$

The F matrix contains the b_j coefficient transformation vectors for all wavenumbers. This way of calculating the b_j parameter using a moving spectral window leads to a banded diagonal matrix. The b_0 vector contains the offset terms for all wavenumbers. Each time a new sample is measured on the slave instrument, the obtained spectra S_{new} can be standardize into $(S_{\text{new}})_{\text{std}}$ using F and b_0 :

$$(S_{\text{new}})_{\text{std}} = S_{\text{new}} F + b_0.$$

The standardization model for every master–slave combination needs to be designed, describing the shift between each slave instrument and the master instrument.

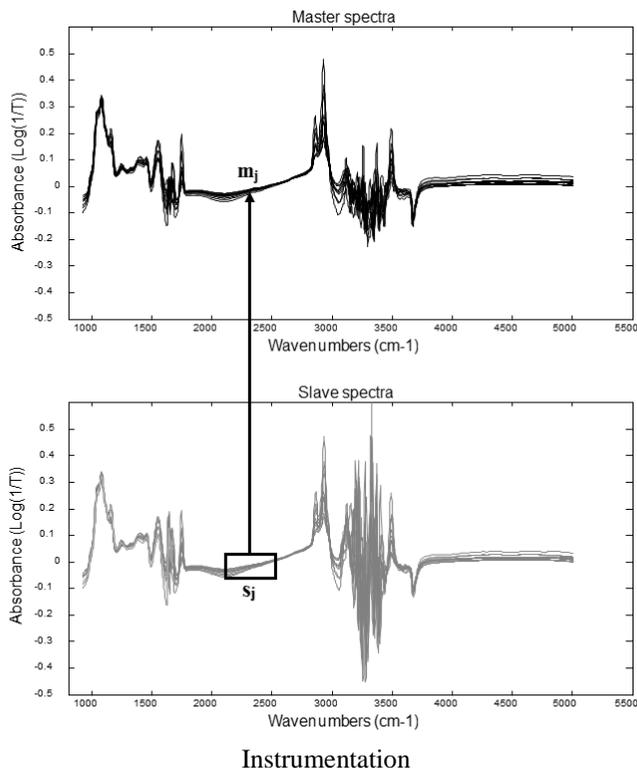


Figure 2-1. Graphical view of the piecewise direct standardization technique. m_j = the response measured at wavenumber j on the master instrument; s_j = the response measured in a small window of size n around j (neighboring) on the slave instrument; T = transmittance

Table 2-1 shows a summary of the 21 different instruments, located in 10 laboratories, and their characteristics used in this study. The different machine types are FT 6000 and FT+ (Foss, Hillerød, Denmark), FTS (Bentley, Chaska, MN) and Standard Lactoscope FT-MIR automatic (Delta Instruments, Drachten, the Netherlands). All the instruments of this study are located in Germany (11), France (7), and Belgium (3). The wavenumber ranges of the different brand are 925.66 to $5,010.15$ cm^{-1} for Foss FT 6000 and FT+ instruments, 649.03 to $3,998.59$ cm^{-1} for Bentley instruments and 397.31 to $4,000$ cm^{-1} for Delta instruments. The resolution used was 8 cm^{-1} for Delta and Bentley Instruments and unknown for Foss instruments.

Milk Samples

Analysis of identical samples is needed to standardize each machine. To achieve this, several interlaboratory studies were organized; 21 sets of identical samples were distributed to the different laboratories to standardize the spectra. These sets of samples were produced according to the IDF (International Dairy Federation) norm 141, ISO/DIS 9622 IDF 141 (IDF, 2012). All the sets created consisted of 10 samples of raw milk with large variations in fat (between 1 and 5% mass/ vol) and protein

(between 2.9 and 5% mass/vol). In all cases, once received, milk samples were homogenized and analyzed at $40 \pm 2^\circ\text{C}$ in triplicate in each laboratory following a predefined protocol.

Table 2-1. Description of the Fourier transform mid-infrared instruments included in the study

Brand	Type	Number of instruments	Frequency reported by constructors (cm^{-1})	Number of Wavenumbers
Foss Electric A/S (Hillerød/Denmark)	FT 6000	8	926 - 5010	1060
Foss Electric A/S	FT+	5	926 - 5010	1060
Bentley (Chaska, MN)	FTS	7	649 - 3999	899
Delta Instruments (Drachten, the Netherlands)	Lactoscope	1	397 - 4000	935

Methodology

An FT 6000 instrument (Foss) was defined as the master instrument. Figure 2-2 shows a typical milk spectrum measured using the master; as for all milk spectra, the water spectrum was subtracted. Different characteristic peaks can be clearly characterized, as well as a noisy area induced by H_2O absorption.

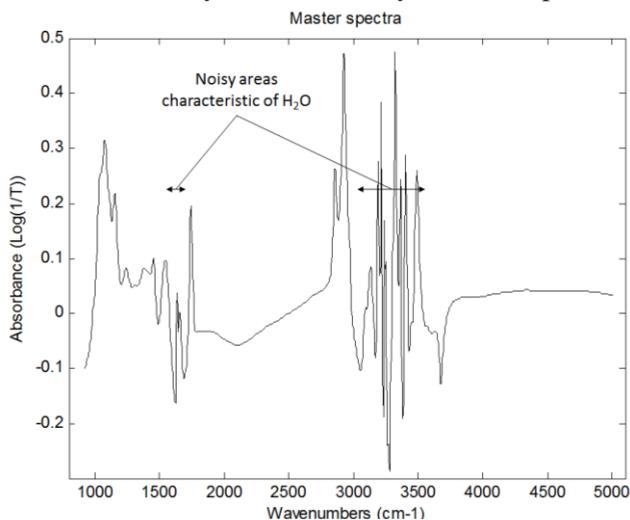


Figure 2-2. Raw milk spectrum from the master instrument. T = transmittance

Figure 2-3 shows the same master spectrum after removing the noisy areas; that is, $1,600$ to $1,689 \text{ cm}^{-1}$ and $3,008$ to $5,010 \text{ cm}^{-1}$.

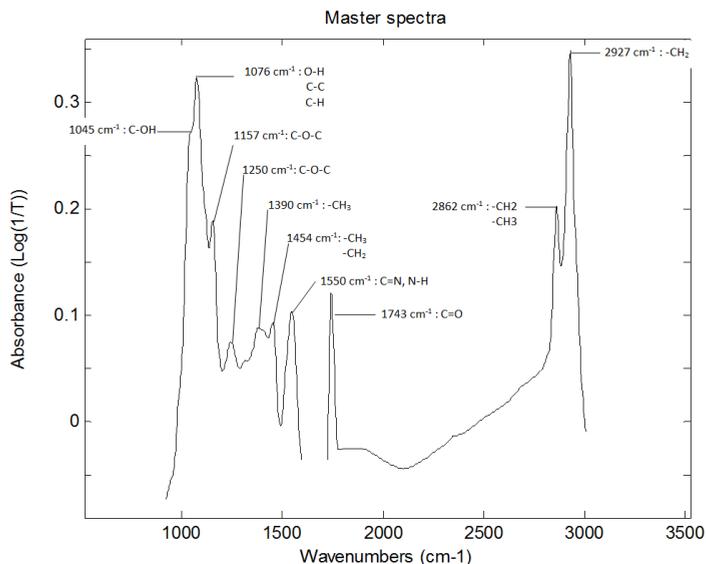


Figure 2-3. Master milk spectrum with assignment of the main spectral bands, after removal of noise areas from water absorption. T = transmittance.

The main components of dairy milk can be linked with characteristic bands on milk FT-MIR spectra because of their chemical composition and chemical bonds absorbing light at specific wavenumbers. Lactose induces a response around $1,045\text{ cm}^{-1}$ with C–O stretching vibration of alcohols functions, $1,076\text{ cm}^{-1}$ with C–O, C–C, and C–H stretching vibration, and $1,157$ and $1,250\text{ cm}^{-1}$ with C–O–C ether stretching. Proteins appear around $1,550\text{ cm}^{-1}$ with peaks of C–N and N–N stretching. Fat chains appear around $1,390$ and $1,454\text{ cm}^{-1}$ with C–H bending of $-\text{CH}_3$ and $-\text{CH}_2$, and around $2,862$ and $2,927\text{ cm}^{-1}$ with C–H stretching of $-\text{CH}_3$ and $-\text{CH}_2$. Fat also appears around $1,743\text{ cm}^{-1}$ because of the C=O ester stretching (Socrates, 1980). In a first step, an interlaboratory study was organized with all the laboratories and apparatus, and the standardization coefficients F and b0 were created as previously explained, using the milk set measured in all the instruments. To validate the standardization method, a fat-prediction model developed on the master instrument was applied on all slave instruments before and after the standardization procedure. The fat model was developed by partial least squares method and based on the whole milk spectrum. Then all the slave predictions were compared with the prediction obtained by the master instrument. In a second step and to really validate the method, the coefficients obtained during the first interlaboratory study were applied to spectra obtained during a second interlaboratory study, realized 1 mo later. All the results were expressed in terms of R^2 (determination coefficient), root mean square error (RMSE), slope (deviation from 1), and bias between the slave and the master predictions. Effects of instrument brands, PDS, and interlaboratory studies on these results were assessed by an ANOVA. All computations, chemometric analysis, and graphics were carried out with programs developed in Matlab v7.5.0 (The Mathworks Inc., Natick, MA) and the

PLS toolbox v. 4.11 (Eigenvector Research Inc., Wenatchee, WA). For ANOVA the Minitab Statistical Software (Minitab Inc., State College, PA) was used.

Results and discussion

Standardization

To be able to perform the standardization, a first step to harmonize the number of wavenumbers was applied. A linear interpolation was performed on all slave spectra. In the case of the Bentley instruments, spectra were interpolated from 649 to 3,999 cm^{-1} to 926 to 3,999 cm^{-1} and in the case of Delta from 397 to 4,000 cm^{-1} to 926 to 4,000 cm^{-1} . Spectra in transmittance values are transformed into absorbance by a \log_{10} . Then the PDS procedure was applied to each of the slave instruments. Figures 2-4, 2-5, and 2-6 illustrate the effect of standardization on spectra from instruments of the 3 brands, on the differences between the absorbance values of master and these slave spectra, and on fat predictions from an equation created on the master. For the 3 brands, slave spectra perfectly matched master spectra after PDS, and differences were strongly reduced after standardization of the slaves except in noisy areas that had to be discarded because they were induced by water response. The fat predictions from the slaves were compared with fat predictions from the master before and after standardization. For the 3 brands, the PDS method allowed reduction in the differences between the slaves and the master predictions. Bias for Delta, Bentley, and Foss instruments, respectively, decreased from -0.5287 to 0.0000 , from -0.8654 to 0.0000 , and from 0.1557 to 0.0000 . The RMSE also decreased for the 3 instruments, respectively, from 0.6114 to 0.0271 , from 0.9979 to 0.0127 , and from 0.1693 to 0.0056 g of fat/100 mL of milk, showing that the fat prediction developed on the master can be used on instruments from these 3 brands with limited error. Figure 2-7 illustrates the RMSE between fat predictions of the master and each slave instrument, before and after standardization.

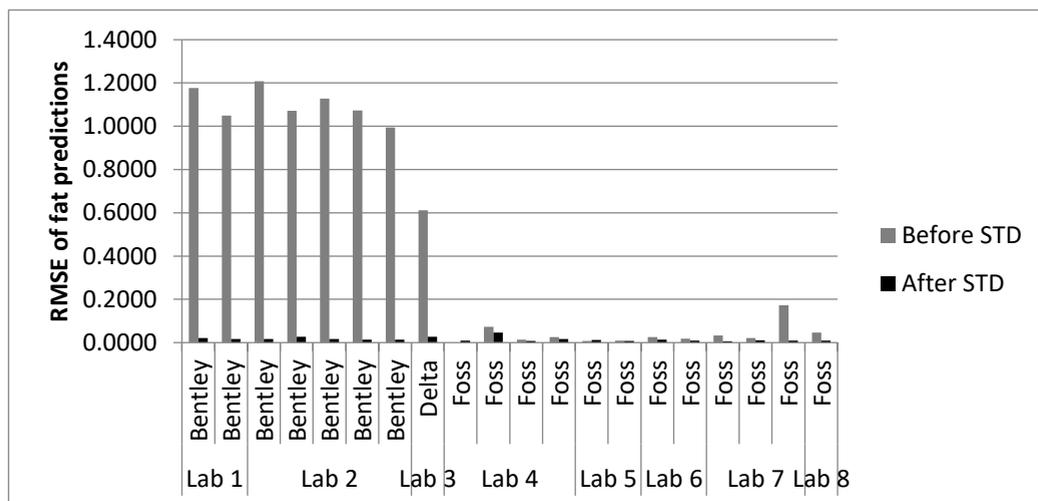


Figure 2-4. Root mean square error (RMSE) between predictions of each slave instruments and the master, before and after standardization ($n = 20$). Before STD = before standardization; After STD = after standardization; Lab = laboratory. Bentley (Chaska, MN), Delta (Drachten, the Netherlands), and Foss (Hillerød, Denmark).

As expected, when comparing fat predictions before standardization of the slaves with the master, larger differences were found with Delta and Bentley slave instruments than with Foss slaves. For all instruments before standardization, RMSE ranged between 0.0083 and 1.2074. However, after PDS, these values decreased and ranged between 0.0066 and 0.0466 for all slave instruments, allowing a clear use of the fat prediction developed on the master in all the slave instruments included in the interlaboratory study. The R^2 , slope, bias, and RMSE between master and slave fat predictions before and after PDS were calculated for each instrument and averaged (Table 2-2) to get an overview of the PDS effect.

Table 2-2. Means of statistical results from regressions between fat predictions by the Master and the slaves, before and after standardization of interlaboratory study 1 ($n=20$)¹

	Global means		
	Before PDS	After PDS	p value
R^2	0.9999	0.9999	0.512
Slope	0.1379	0.0035	0.000
Bias	0.3781	0.0000	0.000
RMSE	0.4609	0.0156	0.000

¹PDS: Piece-wise direct standardization; Slope: deviation according to 1; RMSE: Root mean square error

The R^2 between master and slaves predictions did not change after standardization ($P = 0.512$) and were higher than 0.999. Slope deviation between predictions was greatly reduced, on average from 0.1379 to 0.0035 ($P = 0.000$). Bias and RMSE between predictions decreased after standardization, from 0.3781 to 0.0000 for bias ($P = 0.000$) and from 0.4609 to 0.0156 on average for RMSE ($P = 0.000$). Table 2-3 shows the average of RMSE by brands, before and after PDS. Before standardization the RMSE mean values were different for each brand ($P = 0.001$), i.e., 0.0404, 0.6114, and 1.1001 for Foss, Delta, and Bentley, respectively.

Table 2-3. Root mean square error (RMSE) between fat predictions of master and slaves, averaged by brand, before and after PDS ($n = 20$)¹

	Averaged RMSE ²			P-value
	Bentley (n=7)	Delta (n=1)	Foss (n=12)	
Before PDS	1.1001	0.6114	0.0404	0.001
After PDS	0.0179	0.0271	0.0133	0.279

¹PDS = piecewise direct standardization. ²Bentley (Chaska, MN), Delta (Drachten, the Netherlands), and Foss (Hillerød, Denmark).

After PDS, the averaged RMSE decreased to 0.0133 for Foss, 0.0271 for Delta, and 0.0179 for Bentley. Analysis of variance indicated that these values were not significantly different regarding the 3 brands ($P = 0.279$). For all brands, the PDS method allowed great reduction in the predictions errors with the master predictions.

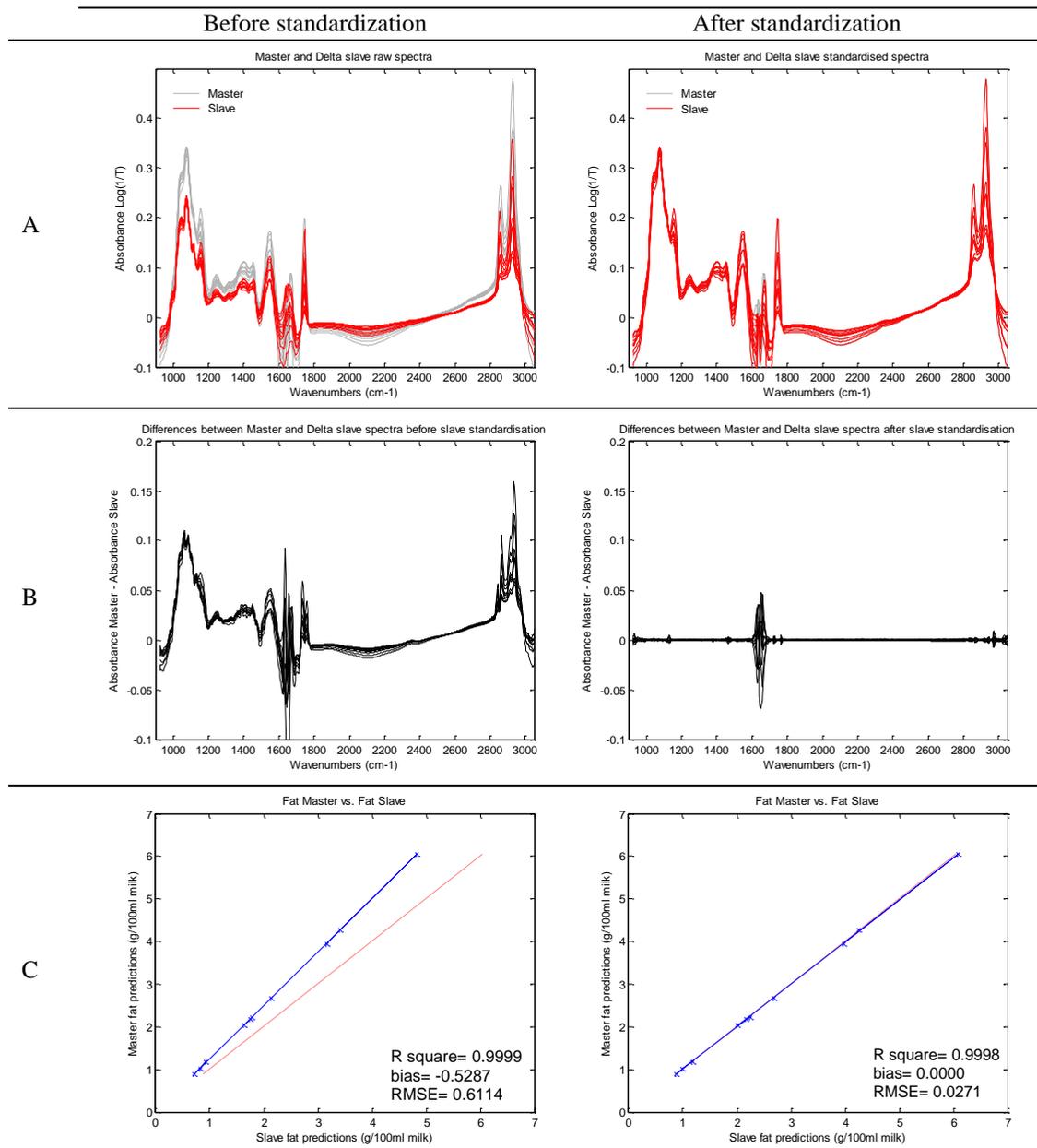


Figure 2-5. Effect of standardization of a Delta (Drachten, the Netherlands) slave spectra on differences with master spectra and on fat predictions. (A) Spectra of the master and a Delta slave instrument before (left) and after standardization (right), from 926 to 3,008 cm^{-1} . (B) Differences between spectra of the master and a Delta slave instrument before (left) and after standardization (right), from 926 to 3,008 cm^{-1} . (C) Comparison of fat predictions by the master and a Delta slave before (left) and after (right) standardization. RMSE = root mean square error. T = transmittance.

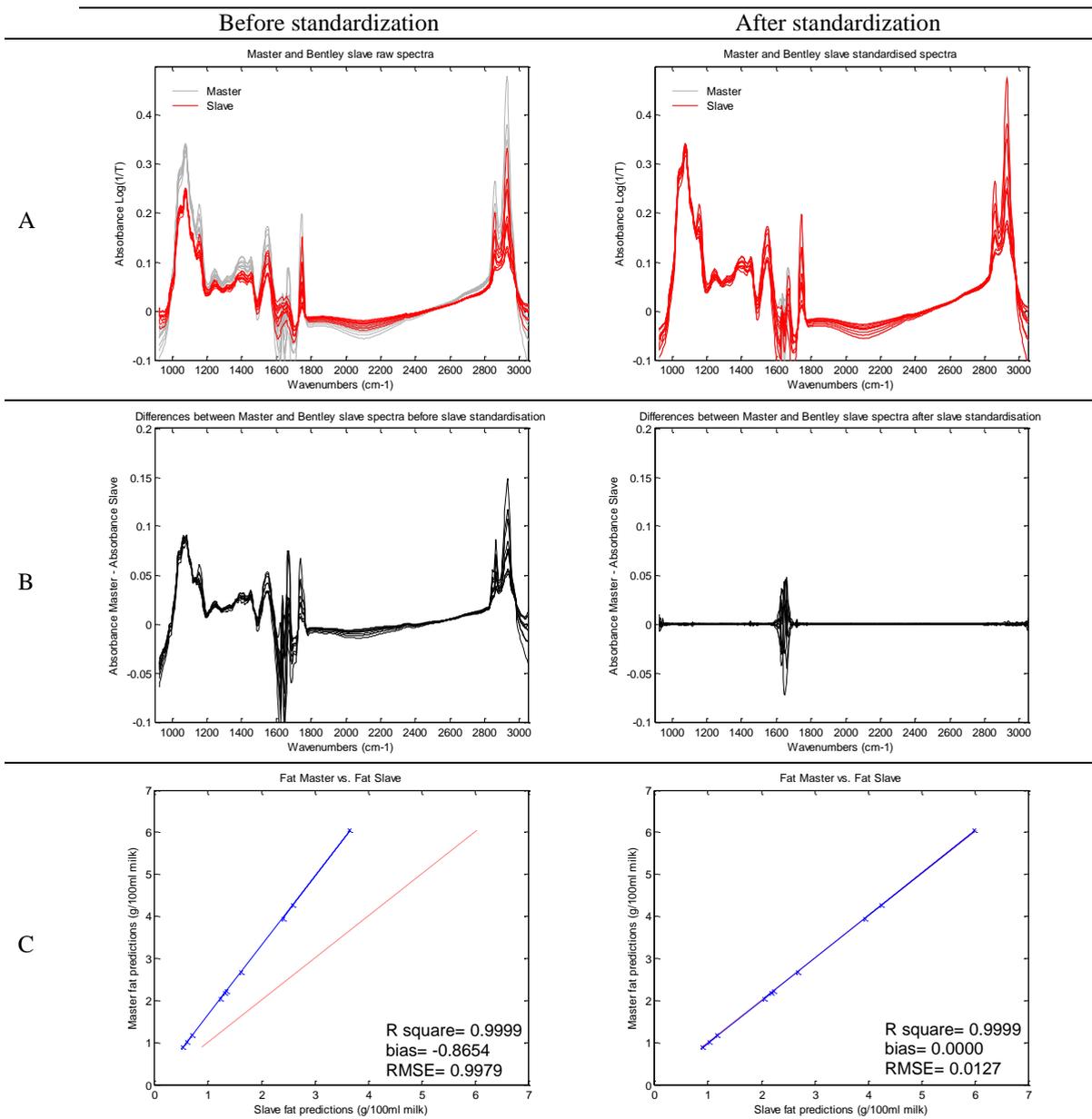


Figure 2-6. Effect of standardization of a Bentley (Chaska, MN) slave spectra on differences with master spectra and on fat predictions. (A) Spectra of the master and a Bentley slave instrument before (left) and after standardization (right), from 926 to 3,008 cm⁻¹. (B) Differences between spectra of the master and a Bentley slave instrument before (left) and after standardization (right), from 926 to 3,008 cm⁻¹. (C) Comparison of fat predictions by the master and a Bentley slave before (left) and after (right) standardization. RMSE = root mean square error. T = transmittance.

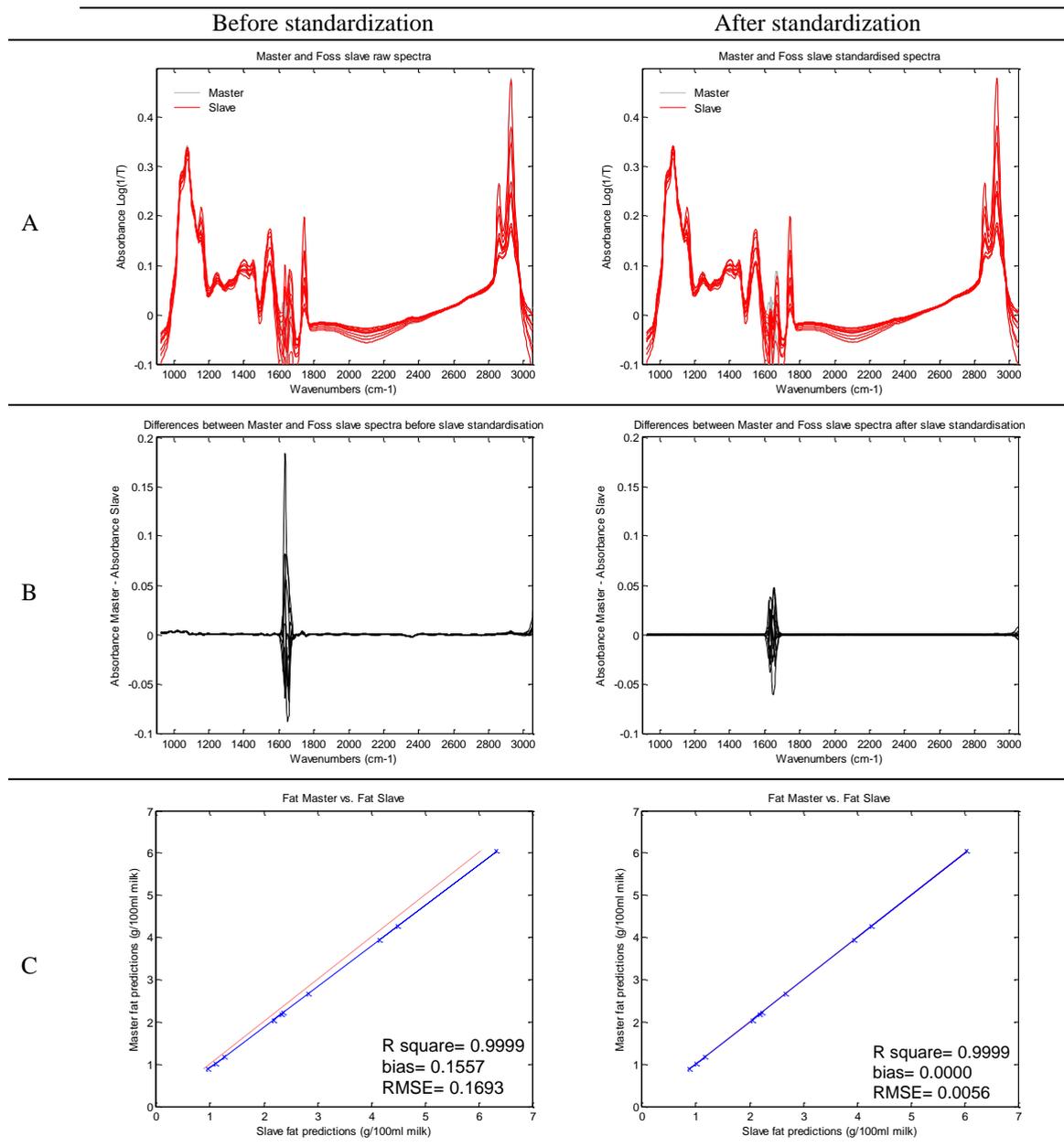


Figure 2-7. Effect of standardization of a Foss (Hillerød, Denmark) slave spectra on differences with master spectra and on fat predictions. (A) Spectra of the master and a Foss slave instrument before (left) and after standardization (right), from 926 to 3,008 cm^{-1} . (B) Differences between spectra of the master and a Foss slave instrument before (left) and after standardization (right), from 926 to 3,008 cm^{-1} . (C) Comparison of fat predictions by the master and by a Foss slave before (left) and after (right) standardization. RMSE = root mean square error. T = transmittance.

Validation of the Standardization Coefficients

The coefficients obtained from the first interlaboratory study were applied to the spectra of a second interlaboratory study 1 mo later. Table 2-4 shows the results of these standardizations with the first interlaboratory study coefficients. The R^2 was still higher than 0.999, and no significant differences were observed ($P = 0.283$) with or without PDS. Slope, bias, and RMSE still decreased greatly after applications of first interlaboratory study coefficients on spectra from the second interlaboratory study. This clearly indicates the stability of the model as well as the PDS procedure between instruments along time. In average, slope, bias, and RMSE were, respectively, reduced from 0.1304 to 0.0093 ($P = 0.000$), from 0.4118 to 0.0350 ($P = 0.000$), and from 0.4458 to 0.0393 ($P = 0.000$). Slope deviation to one, bias, and RMSE were all significantly higher in the second interlaboratory study than in the first interlaboratory study, respectively, 0.0093 versus 0.0035 ($P = 0.010$), 0.035 versus 0.000 ($P = 0.000$), and 0.0393 versus 0.0156 ($P = 0.000$). These results show that standardization coefficients seem to be less adapted to reduce differences between master and slaves spectra 1 mo later after their creation. This study, through several interlaboratory studies, has also shown that the standardization coefficients can be used during time, even if the error is slightly increasing, which can be explained by physical wear and perturbation on apparatus resulting in spectral deviations in time. Further monthly interlaboratory study should bring more robustness to the standardization coefficients, making the standardization step stable for a long time and allowing a harmonization of the predictions of the network for important farm management indicators (health, fertility, feeding, environmental impact). The methodology proposed here will allow a complete correction of physical variation within and between different instruments and brands, allowing the construction of a FT-MIR international database and models. In this work this database has been successfully related to an important chemical characteristic of the milk (fat content). Further work is needed to validate the extension of these results to common calibrations related to animal characteristics such as fertility, nutrition, health, and environment.

Table 2-4. Mean statistic results of comparison of fat prediction by the master and the slaves, before and after standardization of the second interlaboratory study spectra with coefficients from the first interlaboratory study ($n = 20$)¹

	Global means		
	Before PDS	After PDS	p value
R^2	0.9998	0.9991	0.283
Slope	0.1304	0.0093	0.000
Bias	0.4118	0.0350	0.000
RMSE	0.4458	0.0393	0.000

¹PDS: Piece-wise direct standardization; Slope: deviation according to 1; RMSE: Root mean square error

Conclusions

The study described here illustrates that a network of FT-MIR instruments can be standardized; this is illustrated by the reduction of differences between fat predictions of different instruments and brands. The methodology used consisted of a simple interpolation followed by the chemometric PDS tool as a standardization technique. This is an important first step to build a common transnational database with spectra coming from different FT-MIR instruments, including different brands, which should allow the creation of new common indicators for farm management. Robust models built on historical data sets collected by a single instrument or master over several years can then be applied or transferred to other instruments from the same or different brands that are working in the same or different wavenumber ranges.

Acknowledgments

The authors thank INTERREG and the Walloon Region for their financial support in this project; the laboratories of the OptiMIR project for their collaboration and participation in interlaboratory studies; and Comité du lait (Belgium), AWE (Walloon Breeding Association, Belgium), CRA-W (Walloon Agricultural Research Center, Belgium), LKV-BW (Germany), LKV-NRW (Germany), LACOLAIT (France), CLASEL (France), ADECL62 (France), and GxABT (Belgium) for their help on organization, spectra collection and transfer. The authors acknowledge Ouissam Abbas (CRA-W) for spectral bands interpretation and Claire Darimont for her valuable work on interlaboratory studies organization.

References

- Asselain M., Manificier D., and Agnet Y., 1996. Method and Apparatus for the Spectrophotometric Assay of Aqueous Liquids. Foss Electric A/S, assignee. European Pat. No. 588892.
- Bonfatti V., Di Martino G., and Carnier P., 2011. Effectiveness of mid-infrared spectroscopy for the prediction of detailed protein composition and contents of protein genetic variants of individual milk of Simmental cows. *J. Dairy Sci.* 94:5776–5785.
- Bouveresse E., and Massart D. L., 1996. Improvement of the piecewise direct standardisation procedure for the transfer of NIR spectra for multivariate calibration. *Chemom. Intell. Lab. Syst.* 32:201–213.
- Dehareng F., Delfosse C., Froidmont E., Soyeurt H., Martin C., Gengler N., Vanlierde A., and Dardenne P., 2012. Potential use of milk mid-infrared spectra to predict individual methane emission of dairy cows. *Animal* 6:1694–1701.
- Friggens N.C., Ridder C., and Løvendahl P., 2007. On the use of milk composition measures to predict the energy balance of dairy cows. *J. Dairy Sci.* 90:5453–5467.
- Ghosh P.K., and Jayas D. S., 2009. Use of spectroscopic data for automation in food processing industry. *Sens. Instrumen. Food Qual.* 3:3–11.
- Goulden J.D.S., 1964. Analysis of milk by infrared absorption. *J. Dairy Res.* 31:273–284.
- Grappin R., and Jeunet R., 1976. Essais de l'appareil Milko-Scan 300 utilisé pour le dosage en série de la matière grasse et des protéines du lait. *Lait* 56:498–520.

International Dairy Federation (IDF), 2012. Milk and liquid milk products—Guidelines for the application of mid-infrared spectrometry. IDF norm 141. ISO/DIS 9622:2012. IDF, Brussels, Belgium.

Laine A., Goubau A., Dale L. M., Bel Mabrouk H., Hammami H., and Gengler N., 2013. Potential use of mid-infrared milk spectrum in pregnancy diagnosis of dairy cows, Page 523 in Proc. EAAP – 64th Annual Meeting, Nantes. EAAP, Rome, Italy.

McParland S., Banos G., Wall E., Coffey M. P., Soyeurt H., Veerkamp R. F., and Berry D. P., 2011. The use of mid-infrared spectrometry to predict body energy status of Holstein cows. *J. Dairy Sci.* 94:3651–3661.

McParland S., Lewis E., Kennedy E., Moore S. G., McCarthy B., O'Donovan M., Butler S. T., Pryce J. E., and Berry D. P., 2014. Midinfrared spectrometry of milk as a predictor of energy intake and efficiency in lactating dairy cows. *J. Dairy Sci.* 97:5863–5871.

Mohammed R., McGinn S. M., and Beauchemin K. A., 2011. Prediction of enteric methane output from milk fatty acid concentrations and rumen fermentation parameters in dairy cows fed sunflower, flax, or canola seeds. *J. Dairy Sci.* 94:6057–6068.

Rodriguez J.D., Westenberger B.J., Buhse L.F., and Kauffman J.F., 2011. Standardization of Raman spectra for transfer of spectral libraries across different instruments. *Analyst* 136:4232–4240.

Rutten M.J.M., Bovenhuis H., Hettinga K. A., van Valenberg H. J. F., and van Arendonk J. A. M., 2009. Predicting bovine milk fat composition using infrared spectroscopy based on milk samples collected in winter and summer. *J. Dairy Sci.* 92:6202–6209.

Socrates G., 1980. *Infrared Characteristic Group Frequencies*. 2nd ed. Wiley & Sons Ltd., Chichester, UK.

Soyeurt H., Bruwier D., Romnee J.-M., Gengler N., Bertozzi C., Veselko D., and Dardenne P., 2009. Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 92:2444–2454.

Soyeurt H., Colinet F., Arnould V., Dardenne P., Bertozzi C., Renaville R., Portetelle D., and Gengler N., 2007. Genetic variability of lactoferrin content estimated by mid-infrared spectrometry in bovine milk. *J. Dairy Sci.* 90:4443–4450.

Soyeurt H., Dehareng F., Gengler N., McParland S., Wall E., Berry D. P., Coffey M., and Dardenne P., 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *J. Dairy Sci.* 94:1657–1667.

van Knegsel A.T.M., van der Drift S.G.A., Horneman M., de Roos A.P.W., Kemp B., and Graat E.A.M., 2010. Short communication: Ketone body concentration in milk determined by Fourier transform infrared spectroscopy: Value for the detection of hyperketonemia in dairy cows. *J. Dairy Sci.* 93:3065–3069.

Wang Y., Veltkamp D.J., and Kowalsky B.R., 1991. Multivariate instrument standardisation. *Anal. Chem.* 63:2750–2756.

Chapter 3

**Standardization of milk mid-infrared
spectrometers for the transfer and use of
multiple models**

Introduction to chapter 3

The following chapter aims to go deeper into the analysis of the standardization performances, especially to validate the possibility of using models to predict phenotypes of interest for the dairy sector under practical conditions. Indeed, some aspects seem very important to take in consideration in connection with the routine use of such tools, since they have not been studied in previous chapters and no information can be found in the literature with reference to these points.

The first aspect to be considered is the possibility of using models with low robustness and accuracy under realistic conditions. In current literature, studies have mainly focused on evaluating the ability of methods to transfer models for quality control of products in chemistry laboratories (Bouveresse et al., 1994; Bouveresse & Massart, 1996; Rodriguez et al., 2011). These are well-used and robust models predicting major analytes of interest, such as protein, cellulose or fat in forage and cereals (Bouveresse et al., 1994). In the dairy sector, models predicting main components such as fat, protein and lactose are already well controlled for decades. However, recent research has focused mostly on developing models predicting fine milk composition or cow status but these models are characterized by a relatively low accuracy when compared to main components. Such new models are nonetheless considered very useful in a context of animal management or genetic studies (cf Chapter 3, General discussion). An objective of the following chapter is to validate the performance of the standardization method to transfer, from instrument to instrument, such low accuracy models.

A second aspect is dealing with the homogeneity of the spectral responses among instruments, especially because it constitutes a prerequisite for the use of qualitative methods. Qualitative approaches are often used to explore the variability of a dataset. Some studies have attempted to classify spectra of milk samples in classes following feeding systems (Valenti et al., 2013), metabolic status of cows (Grelet et al., 2018) or lameness status of cows (Mineur et al., 2017). Others aim to detect outliers and abnormal samples (Fernandez Pierna et al., 2016). Although the qualitative models based on milk FT-MIR spectra is a great opportunity for the dairy sector, most of the previous studies on standardization have focused on the transfer of quantitative models. In these studies, standardization performances have been evaluated through standard errors (e.g., RMSE) between quantitative predictions of master and slave instruments. The homogeneity of spectral responses was not directly considered, although it is a necessary requirement for qualitative analysis. Indeed, the most common techniques such as Principal Component Analysis (PCA), Partial Least Square Discriminant Analysis (PLS-DA) or Support Vector Machine (SVM) are based principally on differences or distances between spectra. The spectra generated by different instruments must be in the same format to avoid individual instrument specificities interfering with the information of interest. Therefore, a second objective of the following chapter is to evaluate the impact of standardization on spectral homogeneity among instruments.

The last aspect to be considered is the accuracy and reproducibility of predictions among all apparatus. Testing the impact of standardization on accuracy, in order to ensure that all instruments will provide a true response, especially when results are

passed on to farmers, is particularly important. It is also essential to assess the reproducibility of predictions, to prevent heterogeneous predictions among the instruments, particularly within the different spectrometers of the same lab.

References

Bouveresse E., Massart D. L., & Dardenne P., 1994. Calibration transfer across near-infrared spectrometric instruments using Shenk's algorithm: effects of different standardisation samples. *Anal. chim. acta*, 297:3, 405-416.

Bouveresse E., & Massart D.L., 1996. Improvement of the piecewise direct standardisation procedure for the transfer of NIR spectra for multivariate calibration. *Chemometr. Intell. Lab. , 32:2*, 201-213.

Fernandez Pierna J.F., Vincke D., Baeten V., Grelet C., Dehareng F., & Dardenne P., 2016. Use of a multivariate moving window PCA for the untargeted detection of contaminants in agro-food products, as exemplified by the detection of melamine levels in milk using vibrational spectroscopy. *Chemometr. Intell. Lab. , 152*, 157-162.

Grelet C., Vanlierde A., Hostens M., Foldager L., Salavati M., Ingvarlsen K., Crowe M., Sorensen M.T, Froidmont E., Ferris C.P., Marchitelli C., Becker F., Larsen T., Carter F., GplusE Consortium, & Dehareng F., 2018. Potential of milk mid-IR spectra to predict metabolic status of cows through blood components and an innovative clustering approach. *Animal*, 1-10.

Mineur A., Köck A., Grelet C., Gengler N., Egger-Danner C., & Sölkner J., 2017. First results in the use of milk mid-infrared spectra in the detection of lameness in Austrian dairy cows. *Agriculturae Conspectus Scientificus*, 82:2, 163-166.

Rodriguez J.D., Westenberger B.J., Buhse L.F., & Kauffman J.F., 2011. Standardization of Raman spectra for transfer of spectral libraries across different instruments. *Analyst*, 136:20, 4232-4240.

Valenti B., Martin B., Andueza D., Leroux C., Labonne C., Lahalle F., Larroque H., Brunschwig P., Lecomte C., Brochard M., & Ferlay A., 2013. Infrared spectroscopic methods for the discrimination of cows' milk according to the feeding system, cow breed and altitude of the dairy farm. *Int. Dairy J.*, 32:1, 26-32.

Chapter 3 : Standardization of milk mid-infrared spectrometers for the transfer and use of multiple models

C. Grelet,* J. A. Fernández Pierna,* P. Dardenne,* H. Soyeurt,† A. Vanlierde,* F. Colinet,† C. Bastin,‡ N. Gengler,† V. Baeten,* and F. Dehareng*1

*Valorization of Agricultural Products Department, Walloon Agricultural Research Center, 5030 Gembloux, Belgium

†Agriculture, Bio-Engineering, and Chemistry Department, University of Liège, Gembloux Agro-Bio Tech, 5030 Gembloux, Belgium

‡Walloon Breeding Association, B-5590 Ciney, Belgium

J. Dairy Sci. 100:7910–7921 <https://doi.org/10.3168/jds.2017-12720>

© American Dairy Science Association®, 2017.

Abstract

An increasing number of models are being developed to provide information from milk Fourier transform mid-infrared (FT-MIR) spectra on fine milk composition, technological properties of milk, or even cows' physiological status. In this context, and to take advantage of these existing models, the purpose of this work was to evaluate whether a spectral standardization method can enable the use of multiple equations within a network of different FT-MIR spectrometers. The piecewise direct standardization method was used, matching “slave” instruments to a common reference, the “master.” The effect of standardization on network reproducibility was assessed on 66 instruments from 3 different brands by comparing the spectral variability of the slaves and the master with and without standardization. With standardization, the global Mahalanobis distance from the slave spectra to the master spectra was reduced on average from 2,655.9 to 14.3, representing a significant reduction of noninformative spectral variability. The transfer of models from instrument to instrument was tested using 3 FT-MIR models predicting (1) the quantity of daily methane emitted by dairy cows, (2) the concentration of polyunsaturated fatty acids in milk, and (3) the fresh cheese yield. The differences, in terms of root mean squared error, between master predictions and slave predictions were reduced after standardization on average from 103 to 17 g/d, from 0.0315 to 0.0045 g/100 mL of milk, and from 2.55 to 0.49 g of curd/100 g of milk, respectively. For all the models, standard deviations of predictions among all the instruments were also reduced by 5.11 times for methane, 5.01 times for polyunsaturated fatty acids, and 7.05 times for fresh cheese yield, showing an improvement of prediction reproducibility within the network. Regarding the results obtained, spectral standardization allows the transfer and use of multiple models on all instruments as well as the improvement of spectral and prediction reproducibility within the network. The method makes the models universal, thereby offering opportunities for data exchange and the creation and use of common robust models at an international level to provide more information to the dairy sector from direct analysis of milk.

Key words: Fourier transform mid-infrared spectra, standardization, milk, model transfer

Introduction

Over the past decade, the number of research studies seeking to extract more quantitative information from the Fourier transform mid-infrared (FT-MIR) spectra has increased constantly (De Marchi et al., 2014). Equations based on the full spectrum have been developed for the determination of fine milk components such as fatty acid profiles (Soyeurt et al., 2006; Rutten et al., 2009), protein composition (Bonfatti et al., 2011), minerals (Soyeurt et al., 2009), ketone bodies (van Kneysel et al., 2010), citrate (Grelet et al., 2016), and lactoferrin (Soyeurt et al., 2007). Other studies have focused on FT-MIR spectra to build equations predicting technological properties of milk such as milk acidity (De Marchi et al., 2009), ability to coagulate, firmness of curd, or cheese yield (Dal Zotto et al., 2008; Colinet et al., 2015). Recent work has directly considered the FT-MIR spectrum of milk as a reflection of cows' status, with FT-MIR equations being developed to predict methane emissions of dairy cows (Dehareng et al., 2012; Vanlierde et al., 2016), likelihood of conception (Hempstalk et al., 2015), body energy status (McParland et al., 2011), energy intake and efficiency (McParland et al., 2014). In the work of Lainé et al. (2017), the spectrum is even considered as a response for which the effect of pregnancy is evaluated. Hence, the FT-MIR analysis of milk allows the measurement of multiple variables to be used for fine milk quality control in industry, management of herds, or the generation of new phenotypes for genetic studies. Even if some models could be statistically considered as low quality, they are of major interest for the dairy sector because they provide the opportunity to predict key variables that were not available before on a large scale and in a cost-effective way. However, developing such models is time consuming and expensive given that they require the analysis of a large number of samples to cover the whole distribution of the studied trait as well as a large spectral variability. Therefore, there is a clear interest in sharing predictive models among milk laboratories and milk recording organizations. However, a major issue with FT-MIR data is related to the specific instrumental response produced by each spectrometer. These differences between spectral responses of instruments originate from the physical characteristics and acquisition modes specific to each model of machine and from the different uses, piece replacements, and maintenance operations specific to each spectrometer. Differences in spectral response cause difficulties in combining spectra as well as bias in predictions when transferring a calibration model built on one instrument to another instrument. Consequently, exchanges of data and models are limited. To cope with this issue, classical models predicting the main milk components by FT-MIR (e.g., fat and proteins) are monitored and adjusted over instruments and over time using slope and intercept correction. The method is based on the adjustment of the models according to interlaboratory study samples, in which the content of the relevant components is known. However, for most of the new predicted variables (e.g., cows' physiological status or hard-to-measure fine milk components), it is expensive or almost impossible to produce interlaboratory study samples with a known content of the variable of interest. This makes it difficult or impossible to adjust a model after transfer to another spectrometer. Consequently, a model developed on one instrument theoretically can be used only by that instrument because of its specific format. In the context of increasing interest in using new models, the impossibility of transferring them leads to a suboptimal situation, as the

creation of robust models is difficult and expensive. For this reason, it is necessary to implement a preliminary step of spectral standardization permitting the sharing of models. In the context of projects involving international networking, since December 2011 a large instrument standardization network has been developed to harmonize the format of FT-MIR milk spectral response. The objective is to clear the way for potential collaborations between organizations using FT-MIR spectrometers for milk analysis. The possibility of creating common data sets and common models that can be transferred from laboratory to laboratory and used by all instruments allows financial and technical resources to be pooled. Moreover, the possibility of merging spectral data, as far as the reference methods are comparable, allows the inclusion of different feeding systems, breeds, and management, thus increasing the robustness of the developed common models. Over the years the network size has increased, and as many as 127 instruments of 3 different brands coming from 14 countries on 4 continents (North America, Asia, Europe, and Oceania) have been standardized. Recently, it has been shown that using the piecewise direct standardization (PDS) method it is possible to transfer a high-quality fat model from one instrument to more than 20 different instruments in the network (Grelet et al., 2015). However, there is no information about the possibility of transferring models with lower accuracy or predicting fine milk composition or indirect variables, which are not milk components and consequently are predicted indirectly, despite the fact that these models are of great interest to the dairy sector. Furthermore, the effects of standardization on spectral and prediction reproducibility over the network have never been assessed even though it is essential for management or breeding purposes. Therefore, the objectives of this study were to evaluate the effect of the PDS standardization method (1) on spectral reproducibility over spectrometers in a network, (2) on transferring multiple and varied FT-MIR models from one instrument to another, and (3) on the accuracy and reproducibility of predictions among all apparatus. The global perspective is to make all spectrometers speak the same language, thereby allowing the transfer and exchange of developed models predicting classical and new parameters throughout the network.

Materials and methods

Instrumentation

The different instruments available through the network are FT 6000, FT+, FT2, and FT120 (Foss, Hillerød, Denmark); FTS (Bentley, Chaska, MN); and Standard Lactoscope FT-MIR automatic (Delta Instruments, Drachten, the Netherlands). The wave number ranges of the different brands were 925.66 to 5,010.15 cm^{-1} for Foss instruments, 649.03 to 3,998.59 cm^{-1} for Bentley instruments, and 397.31 to 4,000 cm^{-1} for Delta instruments. The resolution used was 8 cm^{-1} for Delta and Bentley instruments and unknown for Foss instruments. As the goal of this work was to validate the standardization method rather than to compare the results of the different brands, the brands were anonymized as brand A, brand B, and brand C. In this study we used data coming from a December 2015 interlaboratory study involving 66 instruments of the 3 brands located in 26 laboratories in Austria, Belgium, Canada, France, Germany, Luxembourg, Switzerland, and the United Kingdom.

Standardization

The standardization procedure, based on the PDS method, is described in Grelet et al. (2015). A set of standardization samples were measured on a reference instrument (the “master” instrument) and on each instrument that needs to be aligned (the “slave” instruments), leading to different response matrices. As reported by Grelet et al. (2015) and Wang et al. (2016), individual FT-MIR spectrometers suffer from instability over time. To cope with this instability, the master cannot be a single instrument. To bring stability to the reference, the master was therefore a fictitious machine that is an average of 18 instruments selected for their stability over time. In this configuration, the reference was linked to and dependent on the network but had a proper and stable spectral response to which all the slaves were matched. The response measured at a precise wave number on the master instrument was related to the response located in a small window around the same wave number measured on each slave instrument. A linear regression was then performed between the spectral response of the master at each wave number and the corresponding windows on the slaves. The coefficients generated for all wave numbers are called standardization coefficients. Whenever a new sample was measured on the slave instrument, the obtained spectra could be standardized to the master response format using these standardization coefficients. A standardization model needed to be designed for every master–slave combination, correcting the shift between each slave instrument and the master instrument. To match each slave to the master, a set of standardization samples needed to be analyzed by all instruments following the same procedure. To achieve this, interlaboratory studies have been organized in the network every month since December 2011, distributing sets of identical samples to the different participating laboratories. The various partial least squares (PLS) models can make use of different spectral areas, so all the spectral regions containing information need to be standardized independently to the model used. The samples were created to cover sufficient variability in the absorbance values at each wave number to allow a regression between slave and master absorbance values at each region of the spectra. All the sets generated consisted of 5 samples of raw milk with large and orthogonal variations in fat (between 2 and 6% wt/vol) and protein (between 2 and 5% wt/vol). The samples were created by blending skim milk, cream, ultrafiltration retentate, and permeate. These sets of samples were produced according to the method described in ISO (1999). Samples were preserved with bronopol (0.02%) and sent at 4°C in isothermal packages containing ice packs and using express delivery (within 24 h). The day of receipt, the milk samples were homogenized and analyzed at $40 \pm 2^\circ\text{C}$ in triplicate by each laboratory following a common protocol. The creation of samples and the generation of standardization coefficients were done centrally at Walloon Agricultural Research Center (Gembloux, Belgium). The instrument-specific standardization coefficients were transferred to the respective labs to be applied on the raw spectra of the corresponding slave spectrometers to obtain standardized spectra.

Spectral Reproducibility Within the Network

The spectral variability between instruments of the network was assessed by performing a principal component analysis (PCA) with the spectra of the master and

all 66 instruments before standardization based on the analysis of the 5 common interlaboratory study samples in triplicate. The effect of standardization on the spectral reproducibility of the network was assessed by performing a second PCA with spectra of the master and all 66 instruments before and after standardization based on the same samples. All the spectra were transformed in absorbance and interpolated to match the wave number range of the master to observe differences from spectral response only. A PCA was performed on spectra after a first derivative with a gap of 5 and using 212 selected wave numbers, from 968.1 to 1,577.5 cm^{-1} , 1,731.8 to 1,762.6 cm^{-1} , 1,781.9 to 1,808.9 cm^{-1} , and 2,831.0 to 2,966.0 cm^{-1} (Gretet et al., 2016). Based on the second PCA, the improvement of spectral reproducibility was quantitatively assessed by comparing the global Mahalanobis distances (GH) of the slaves from the master before and after standardization.

Transfer of Individual Calibration Models

To cover the wide diversity of predicted variables that can be found in the milk sector, the effect of standardization on the transfer of models from instrument to instrument was tested for 3 varied models relating to (1) cows' status (daily CH_4 emitted by dairy cows), (2) fine milk composition (PUFA), and (3) technological properties of milk (fresh individual laboratory cheese yield of milk; FCY). The CH_4 model was developed by Vanlierde et al. (2016) and contains samples from Belgium and Ireland. In this study the Legendre polynomial transformation was removed because the interlaboratory study samples, which are not natural samples, cannot be associated with DIM information. The PUFA model comes from the work performed by Soyeyrt et al. (2011) and contains samples from 7 different countries in the European Union. The FCY equation was built within the framework of research by Colinet et al. (2015) based on Belgian samples. All the models were developed using PLS regression with a first derivative and a gap of 5 and using the 212 wave numbers mentioned previously. Calibration and cross-validation statistics of the calibrations are shown in Table 3-1.

Table 3-1. Calibration and cross-validation statistics of equations used

Predicted variable	Terms	Samples	Mean	SD	SEC	R ² c	CV groups	SECV	R ² cv	RPDcv
Methane emitted (g/d)	12	532	430	129	66	0.74	5	72	0.69	1.79
Total of PUFA (g/100mL of milk)	11	1799	0.159	0.045	0.021	0.78	4	0.021	0.77	2.10
Fresh cheese yield (g of curd/100g of milk)	8	337	26.51	7.11	3.44	0.77	4	3.62	0.74	1.96

¹Terms = number of terms used in the regressions; Samples = number of samples in the calibration data sets; SEC = SE of calibration; R²c = coefficient of determination of calibration; CV groups = number of subsets used in cross-validation; SECV = SE of cross-validation; R²cv = coefficient of determination of cross-validation; RPDcv = ratio of SD to SECV.

Based on the analysis of the standardization samples, the calibration models were applied to the master and slave instruments before and after the standardization procedure. All the slave predictions were then compared with the predictions obtained by the master instrument. The results of the comparison between slave predictions and master predictions, before and after standardization, are expressed by the root mean squared error (RMSE). This reflects the differences between predictions of the master and predictions of the slaves due to specific spectral responses, highlighted by model transfer. However, as RMSE is related to the level and the unit of the variable predicted, a relative error was also calculated. The relative error due to model transfer was calculated by looking at the ratio of RMSE between slaves and master predictions divided by the average of the reference values from the calibration data sets.

Accuracy and Reproducibility of Predictions Over the Network

For the 3 models, the accuracy of the predictions within the network was assessed by comparing the global averages of the master predictions and of the predictions of all slave instruments before and after standardization. Comparisons were done using the Tukey test. The reproducibility of the predictions within the network was approached by calculating the standard deviation of the predictions of all instruments before and after standardization for the 5 samples and with the 3 models. Reproducibility within the network was also compared with the 10-d repeatability of predictions of individual instruments calculated, for each model, with the analysis of a common UHT milk set during 10 d. All computations, chemometric analyses, and graphics were carried out with programs developed in Matlab version 7.5.0. (The MathWorks Inc., Natick, MA) and PLS toolbox version 4.11 (Eigenvector Research Inc., Wenatchee, WA).

Results and discussion

Spectral Reproducibility Within the Network

From the first PCA done with spectra of all instruments before standardization, the principal components from 1 to 4 discriminate the 5 samples of the interlaboratory study. Therefore, these principal components comprise spectral information on milk composition. Principal components 5 and 6 allow discriminating the instruments and thus report information on the spectral variability among instruments (Figure 3-1). On this PCA figure, the spectra of the 66 instruments are represented by a color-symbol association. The heterogeneous distribution of the instruments highlights the considerable variability of spectral response between the different spectrometers within the network. Principal components 5 and 6 of the second PCA performed on spectra of the master (red stars) and all 66 instruments before (blue triangles) and after (green squares) standardization are reported in Figure 3-2. On this PCA figure, the master showed reduced variability, with spectra concentrated into a small space showing good homogeneity of the reference spectral response. By contrast with the considerable spectral heterogeneity observed without standardization, the spectral variability of the slaves was relatively limited after PDS. Spectral reproducibility was considerably improved after standardization, and standardized slaves' spectra were concentrated around the master spectra. This indicates that slaves' spectral responses

were much closer to the master’s spectral response than before and that the spectral homogeneity within the network increased.

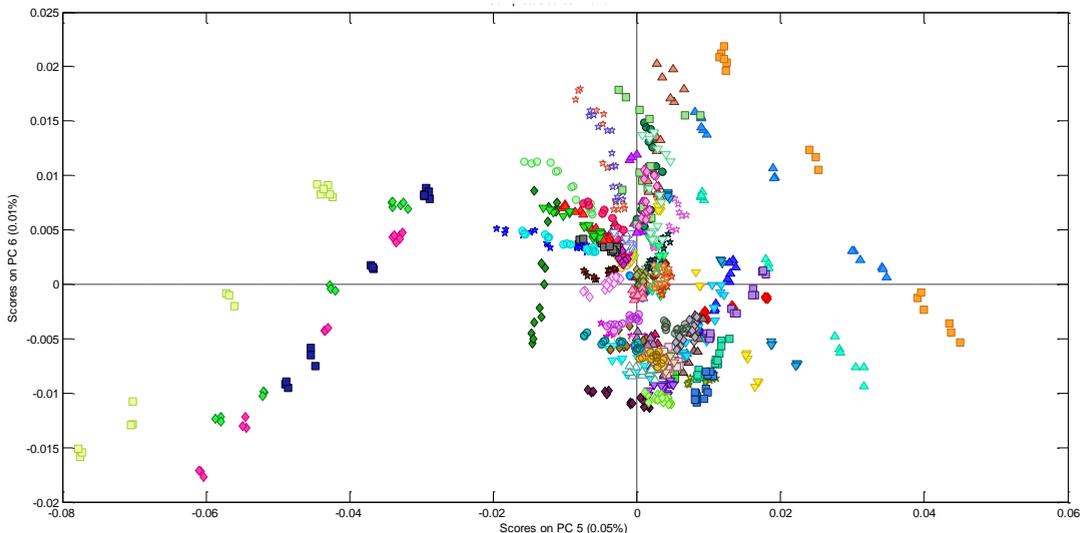


Figure 3-1. Principal component analysis (PCA) of the spectra of all instruments before standardization (n = 66). The PCA is based on the common analysis of 5 standardization samples in triplicate after selection of 212 informative wave numbers and a first derivative. Plot of principal components (PC) 5 and 6 summarizes the spectral variability between instruments. Each color–symbol association represents an individual instrument

The GH from the slaves to the master were calculated to evaluate quantitatively the spectral homogeneity within the network before and after standardization. Before PDS, the GH ranged from 6.17 to 86,759.36, with an average value of 2,655.92. These very high GH can be explained by the fact that GH is the ratio of the Mahalanobis distance of a spectrum to the average of Mahalanobis distances of a reference data set. Classically, the GH is used to compare a sample with a database in a calibration step or as quality control when using an equation to predict new samples. These databases are built to contain as much variability as possible, which makes the denominator -the global Mahalanobis distance of the database- high. The threshold of 3 is then frequently used to define samples as outliers. In this study the GH was used in a different way, by comparing spectra with a reference containing very low variability, making the denominator really low. This reflects the fact that slaves’ spectra contain an important variability compared with the master spectra, which constitute a homogeneous reference with limited spectral variations. Therefore, the threshold of 3 is not adapted in this case because the GH is used to compare another type of data, with another variability. After standardization the GH ranged from 0.50 to 350.24, with an average of 14.26. These quantitative results, with an average GH from the slaves to the master that is 186 times smaller, confirm the first conclusions obtained by visual observation of the PCA. The standardization process strongly reduces the

spectral variability within the network and makes the spectra closer to the master response, which is expected to have a positive effect on the transfer of models.

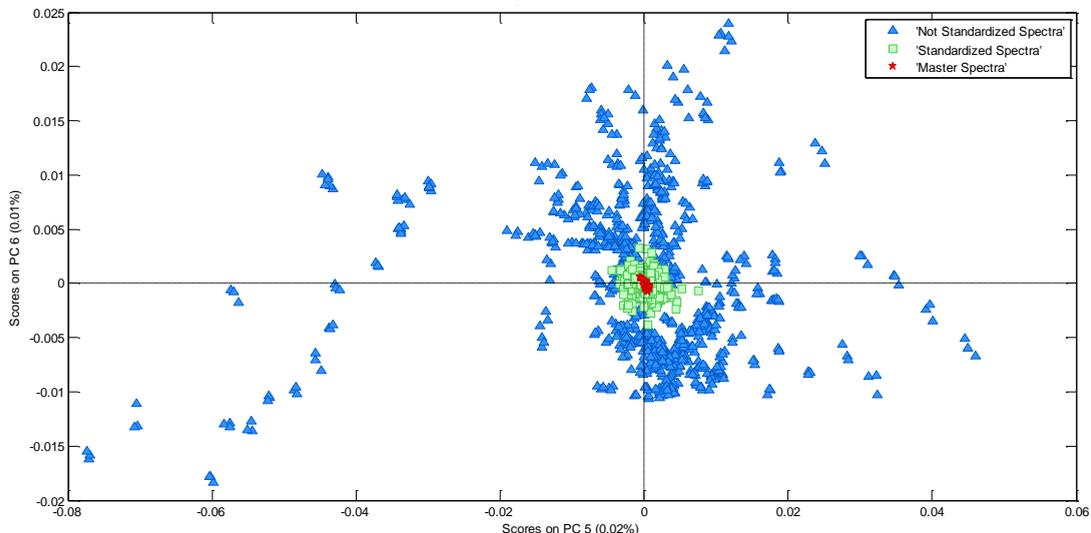


Figure 3-2. Principal component analysis (PCA) of the spectra of all instruments, including the master, before and after standardization ($n = 66$). The PCA is based on the common analysis of 5 standardization samples in triplicate after selection of 212 informative wave numbers and a first derivative. Plot of principal components (PC) 5 and 6 summarizes the spectral variability between instruments.

Transfer of Individual Calibration Models

Figure 3-3 shows the transfer of the CH_4 model to an instrument of each brand before and after standardization. It illustrates the bias potentially generated in predictions by the transfer of a model without a preliminary step of spectral standardization. The figure also shows the reduction of the differences between slave and master predictions induced by the standardization. For all instruments, the differences between master and slave predictions before standardization were substantial, with RMSE ranging from 6 to 422 g/d for CH_4 , from 0.0017 to 0.1333 g/100 mL for the PUFA model, and from 0.1110 to 39.57 g of curd/100 g of milk for FCY. Average RMSE for the 66 instruments without standardization was 103 g/d, 0.0315 g/100 mL, and 2.55 g of curd/100 g of milk, respectively (Table 3-2).

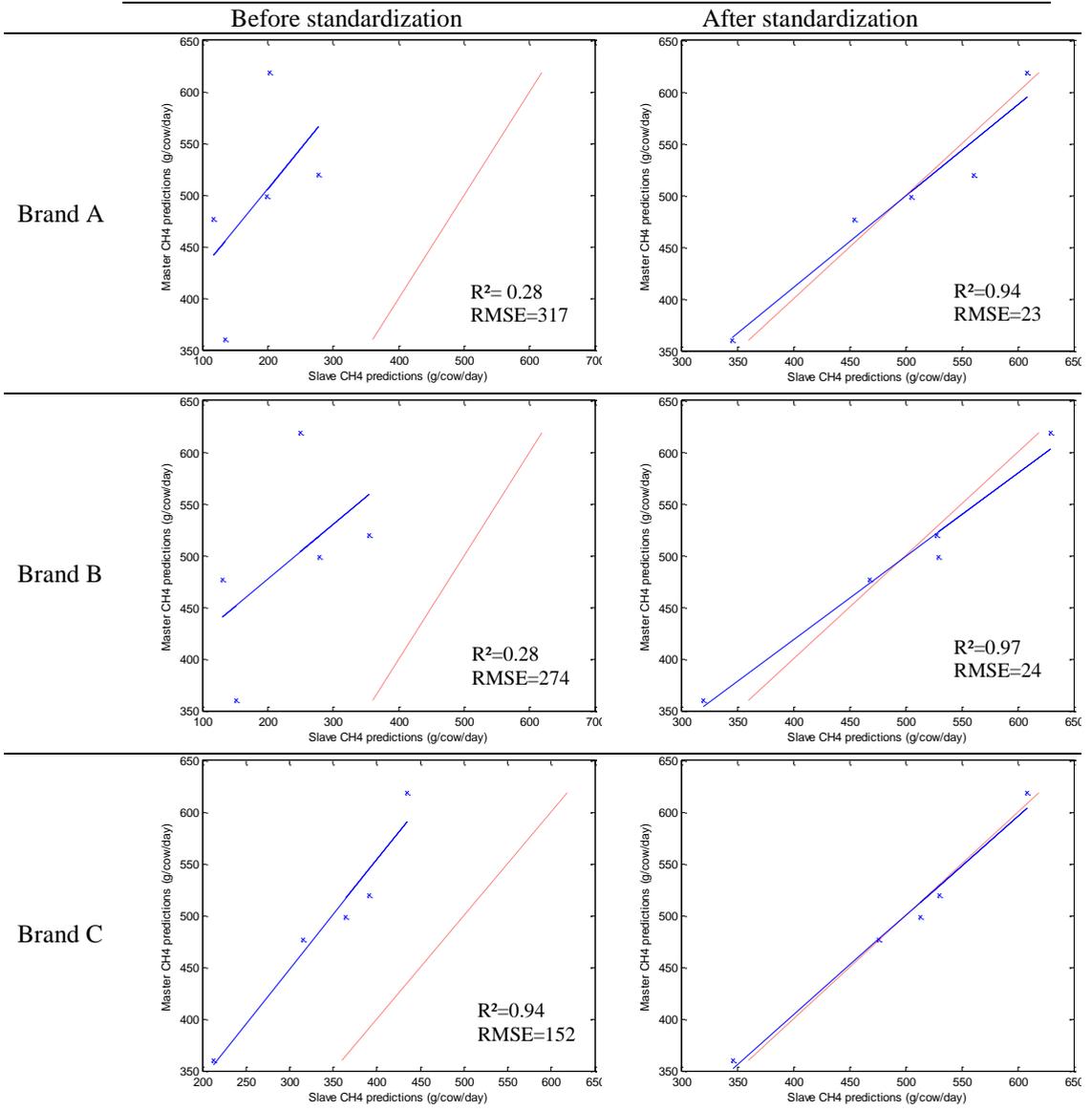


Figure 3-3. Comparison of methane predictions (g/d) by the master and by 3 slaves from the 3 brands before (left) and after (right) standardization. Dashed line is $y = x$; continuous line is the regression line between slave and master predictions. RMSE = root mean squared error.

Table 3-2. Root mean squared error (RMSE) between master and slaves predictions, averaged by brand, before and after piecewise direct standardization (PDS; n = 66 instruments)

	Methane emitted (g/d)		PUFA (g/100mL of milk)		Cheese fresh yield (g of curd/100g of milk)	
	Before PDS	After PDS	Before PDS	After PDS	Before PDS	After PDS
Brand A	231	33	0.0733	0.0089	6.73	1.47
Brand B	348	25	0.1281	0.0027	23.94	1.15
Brand C	73	14	0.0211	0.0039	1.09	0.31
Global average	103	17	0.0315	0.0045	2.55	0.49

These errors due to model transfer were not negligible, with relative RMSE of 23.9, 19.8, and 9.6%. Without standardization, RMSE fluctuated among brands, with different levels of average (Table 3-2) and maximum (Table 3-3) RMSE. However, these levels are relatively high compared with the standard error of cross-validation of the respective equations (Table 3-1), meaning that the transfer of models will add a fairly considerable error to the predictions compared with the inherent error of the models. For the CH₄ and PUFA models, for all brands the transfer from the master to other instruments led in the majority of cases to strong bias in the predictions, making the transfer of models inconceivable. In the case of the FCY model, for brands A and B the transfer from the master to other instruments also led to important bias in the predictions. Concerning brand C, the average difference from the master prediction before PDS was limited, suggesting that the model could be transferred without the standardization step. However, some instruments show elevated RMSE (up to 3.95 g of curd/100 g of milk), inducing substantial errors in predictions when transferring the model without standardization. After standardization, the RMSE between slave and master predictions was considerably decreased and ranged from 2 to 61 g/d for CH₄, 0.0013 to 0.0152 g/100 mL for PUFA, and 0.09 to 2.10 g of curd/100 g of milk for FCY. The average RMSE for the 66 instruments was reduced to 17 g/d, 0.0045 g/100 mL, and 0.49 g of curd/100 g of milk, respectively (Table 3-2).

Table 3-3. Maximum root mean squared error (RMSE) between master and slaves predictions, sorted by brand, before and after piecewise direct standardization (PDS; n = 66 instruments)

	Methane emitted (g/d)		PUFA (g/100mL of milk)		Cheese fresh yield (g of curd/100g of milk)	
	Before PDS	After PDS	Before PDS	After PDS	Before PDS	After PDS
Brand A	420	61	0.1009	0.0152	12.14	2.10
Brand B	422	26	0.1333	0.0041	39.58	1.35
Brand C	247	27	0.0529	0.0112	3.95	1.07

Consequently, the relative RMSE decreased after standardization from 23.9 to 4.0% for CH₄, from 19.8 to 2.8% for PUFA, and from 9.6 to 1.8% for FCY, meaning that the relative error induced by model transfer was reduced to a more acceptable level for the routine use of the predictions. Indeed, after PDS, the average RMSE was relatively limited compared with the inherent standard error of cross-validation of each model, meaning that the transfer did not add a significant error to the final predictions. The averaged and maximum RMSE still varied among the 3 brands but were considerably decreased for all of them after PDS (Tables 3-2 and 3-3). The average differences between master and slave predictions were 6.8, 8.7, and 4.5 times less, respectively. For the 3 models and the 3 brands, the effect of the standardization was a considerable reduction of the differences from master predictions, allowing the transfer of the models to all slave instruments.

Accuracy, Reproducibility, and Use of Predictions Over the Network

Figure 3-4 illustrates, for one sample of the interlaboratory study, the improvement of the reproducibility of PUFA predictions after standardization compared with predictions from raw spectra. After standardization the distribution of the PUFA predictions was much tighter, meaning that the predictions were more precise. The mean was also closer than the master prediction mean, showing an increase in the accuracy of the predictions.

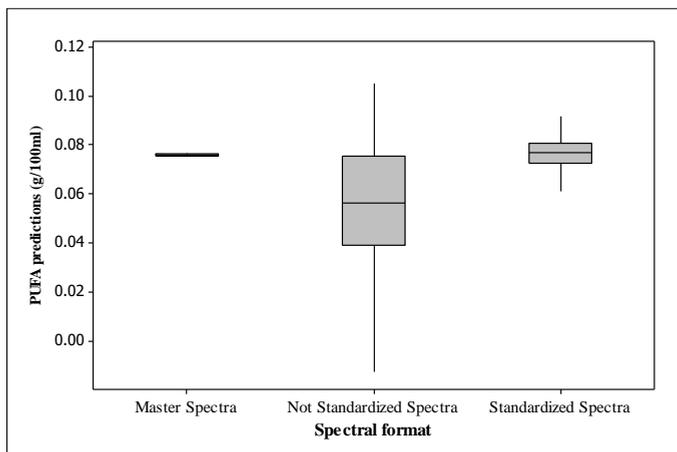


Figure 3-4. Box plot representation of PUFA predictions for a sample of the interlaboratory study analyzed on 66 instruments for master spectra, nonstandardized spectra, and spectra after standardization after removing aberrant values.

To evaluate the effect of standardization on network accuracy for the 66 instruments, the slave prediction means were compared with the master prediction means (Table 3-4). For CH₄ and PUFA, the prediction means for all instruments without standardization were significantly different from the master prediction means, with 431 versus 496 g/d for CH₄ and 0.093 versus 0.117 g/100 mL for PUFA. After standardization, the prediction means were 495 g/d and 0.117 g/100 mL, respectively, and could not be significantly differentiated from the master prediction mean. For

FCY, the prediction means without (25.92 g of curd/100 g of milk) and with (24.46 g of curd/100 g of milk) standardization could not be significantly differentiated from the master prediction mean (24.47 g of curd/100 g of milk), although the prediction mean after PDS seemed closer. Network accuracy was significantly improved after standardization for CH₄ and PUFA models. This also seemed to be the case for FCY, although this could not be demonstrated statistically.

Table 3-4. Accuracy and reproducibility of predictions within the network; comparison of predictions from master, nonstandardized, and standardized spectra, from samples of the interlaboratory study analyzed on 66 instruments using 3 different Fourier transform mid-infrared calibrations

Item	Mean			SD	
	Master	Slaves before PDS ¹	Slaves after PDS	Before PDS	After PDS
Methane emitted (g/d)	496 ^a	431 ^b	495 ^a	126	25
PUFA (g/100mL of milk)	0.117 ^a	0.093 ^b	0.117 ^a	0.0346	0.0069
Fresh cheese yield (g of curd/100g of milk)	24.47 ^{ab}	25.92 ^b	24.46 ^a	4.893	0.694

^{a,b} Means within a row with different superscripts are significantly different by the Tukey test ($P < 0.05$).

¹PDS = piecewise direct standardization.

After PDS, for the 3 brands taken separately and for each model, there was no difference between the predictions of the master and of the slaves of each brand of instrument (Table 3-5). Network reproducibility was improved, with SD between instruments' predictions reduced after standardization from 126 to 25 g/d for CH₄, from 0.0346 to 0.0069 g/100 mL for PUFA, and from 4.89 to 0.69 g of curd/100 g of milk for FCY (Table 3-4). For the 3 models and for brands A, B, and C, the average 10-d repeatability of the instruments was 45, 22, and 50 g/d per cow; 0.0108, 0.0095, and 0.0063 g/100 mL of milk; and 1.06, 0.72, and 0.57 g of curd/100 g of milk, respectively. The reproducibility levels obtained were therefore in the same order of magnitude as the inherent repeatability of individual spectrometers, meaning that predictions throughout the network were as precise as for an individual instrument. These results show that at the network level the method improved the accuracy and reproducibility of predictions by matching all spectrometers to a common reference response format. As shown in Figure 3-3, improved network reproducibility harmonized the regression lines between spectrometers. In breeding studies this increases the usefulness of prediction. Indeed, if intra-herd differences can be adjusted through the herd means, heterogeneity of variances due to instruments would be a major challenge in genetic evaluations and could only be approximately adjusted post prediction with very complex methods (e.g., Gengler et al., 2004) if sources of variation, which are numerous and unforeseeable, are correctly identified for each instrument, which is not realistic on a network-wide scale.

Table 3-5. Reproducibility of predictions among the different brands after standardization; comparison of predictions from master and standardized spectra, from samples of the interlaboratory study analyzed on 66 instruments using 3 different Fourier transform mid-infrared calibrations

	Master	Brand A	Brand B	Brand C	<i>P-value</i>
Methane emitted (g/d)	496.1 ^a	496.0 ^a	487.5 ^a	494.93 ^a	0.967
PUFA (g/100mL of milk)	0.117 ^a	0.117 ^a	0.118 ^a	0.117 ^a	0.999
Fresh cheese yield (g of curd/100g of milk)	24.47 ^a	24.50 ^a	24.35 ^a	24.45 ^a	1.000

^aMeans within a row with different superscripts are significantly different by the Tukey test ($P < 0.05$)

General Discussion

The goal of this work was to evaluate the effect of a standardization method on the transfer of multiple models with low accuracy from instrument to instrument. A previous study (Grelet et al., 2015) demonstrated that until now it was possible to transfer only high-quality models (fat). The results confirm the first conclusions obtained with the fat model: without the use of a standardization step, the transfer of models leads to considerable errors in the predictions, reducing the value of information from FT-MIR milk spectra for the dairy sector. This study shows that standardization greatly reduced the spectral variability between spectrometers of the network by bringing the spectra closer to a common reference response. Moreover, the use of PDS strongly increased the reproducibility and accuracy of predictions across all instruments. With the 3 models used (CH₄ emitted by dairy cows, PUFA in milk, and FCY), the developed method substantially reduced the relative error due to transfer of equations. The differences between master and slave predictions were 6.8, 8.7, and 4.5 times less, respectively. This demonstrates the possibility of transferring different models relating to cow status, fine milk composition, or technological properties of milk within the network. However, the levels of reduction obtained were less impressive than those obtained for the fat model, where the differences between slave and master predictions were reduced on average by a factor of 29.5, whereas the spectral correction was the same. Furthermore, the relative RMSE after standardization were 4.0, 2.8, and 1.8% for CH₄, PUFA, and FCY, respectively, whereas it was only 0.4% for the fat model. Compared with the models used in the study, the main characteristics of the fat model were greater accuracy and a direct and strong signature of fat molecules in the spectra. Figure 3-5 illustrates the link between the coefficient of determination of cross-validation (R^2_{cv}) of the models and the performances of the transfer.

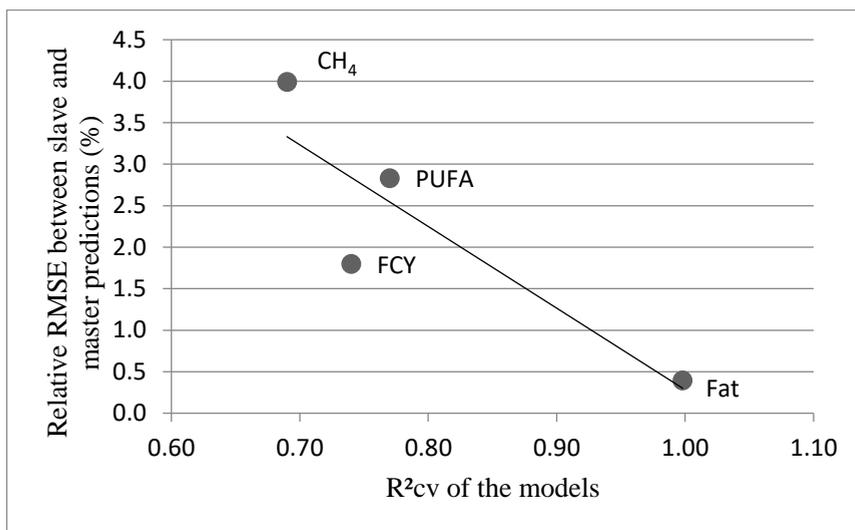


Figure 3-5. Plots representing the link between the quality of the model (coefficient of determination of cross-validation; R^2_{cv}) and the performance of the transfer by piecewise direct standardization [relative root mean squared error (RMSE) between slaves and master predictions after standardization] from the 5 interlaboratory study samples analyzed on 66 instruments. CH₄ = methane emitted by dairy cows; PUFA = PUFA in milk; FCY = fresh cheese yield; Fat = fat model used in Grelet et al. (2015).

This clearly shows that the error does not depend only on the standardization method, which does not increase the error in the final predictions, but mainly on the quality of models used. As the standardization did not interfere with the error of the final results, the decision to put effort into developing and sharing a model relies only on the quality of the model regarding the accuracy needed. Williams (2014) proposed a scale regarding the quality of models using the ratio of performance to deviation (RPD), which is the standard deviation of the reference values divided by the standard error of prediction. From this scale, the models with an RPD of <2.3 are very poor and not recommended. Hence, one can think that developing and transferring such models is useless. However, in Williams (2014), the aim was related to quality control of products, and the scaling of the RPD was done in line with this objective, which is quite demanding. Furthermore, the paper also mentioned that due to some complications (e.g., difficulties obtaining high variance in the sample set) high RPD can be difficult to obtain, whereas the models can still be of interest for industry or research. In the present study, the models used were not dedicated to quality control, and in another context (e.g., animal management advisory and especially breeding) they can be of interest despite their low accuracy. There are several reasons for this. First, even if these models are phenotypically imprecise, they provide useful information that was not available before (e.g., predictors for direct traits that were very difficult to obtain). Second, if prediction errors are random, having multiple records reduces the predictive noise globally. Third, in genetic studies, they are repeated throughout a family, and a common genetic background (being heritable) can

be genetically correlated with other traits of interest (e.g., direct health traits). Usefulness of models with lower phenotypic predictive power is linked to their genetic correlation to other traits of interest. Concretely, there are various examples of how such models with low statistics ($RPD < 2.3$, which is equivalent to $R^2 < 0.81$) can be of interest for the dairy sector. For example, Leclercq et al. (2013) studied the genetic variability of lactoferrin based on a model with $R^2_{cv} = 0.71$, Cecchinato et al. (2009) used coagulation property models with R^2_{cv} between 0.46 and 0.69 to estimate heritabilities and genetic correlations, and the models developed by de Roos et al. (2007) with R^2 of 0.72 for acetone and 0.64 for BHB are currently routinely used for ketosis screening. In a genetic study, Bonfatti et al. (2017b) concluded that genetic progress will be faster with good models and that less accurate equations might be successfully used for breeding purposes. Finally, McParland et al. (2015) showed that the ultimate issue for the use of such models for breeding is the existence of genetic correlations. Based on models predicting energy intake and energy balance with R^2_{cv} of 0.56 and 0.53, respectively, they reported genetic correlations between measured and MIR-predicted traits of 0.84 for energy intake and 0.54 for energy balance, indicating that selection based on MIR-predicted variables would improve true energy intake and energy balance. Consequently, the RPD needed should be defined by the users, and this level will be different following their own purposes and applications. Furthermore, the presented standardization method will improve the usefulness of models with low predictive power. These models provide useful information that was not available before; however, this type of data has to be accumulated across large populations, therefore involving many instruments. Having multiple records reduces the predictive noise globally, but only if these records are comparable across time and instruments. For breeding studies, usefulness of models is linked to their genetic correlation to other traits of interest. However, these studies have to rely on many comparable records. Nonstandardized data would inflate residual -not modeled- variance and therefore reduce heritability and genetic progress. It would also affect genetic correlations between FT-MIR models and direct traits. McParland et al. (2015) reported good genetic correlations between measured and MIR-predicted traits, but this study was based on a single spectrometer. One can hypothesize that genetic correlation would have been lower with nonstandardized FT-MIR data from several instruments. In addition to the standardization methods, some important parameters need to be considered to ensure reproducible and accurate predictions for routine use. First, complementary to the homogeneity of spectral response, the robustness of models is an essential element. This capacity of the models to be all terrain is affected not only by the statistical performances of the models (e.g., the RPD), which are well known, but mainly by others factors not frequently mentioned. Models can be transferred into the network, but it is necessary to ensure that calibration data sets cover the spectral variability of the different geographical regions, breeds, and diets to obtain valid predictions. Robustness is also affected by the number of latent variables used in the PLS models (which should be reasonable) as well as the precision of the reference method, the use of a repeatability file, the integration of several brands of FT-MIR spectrometers into the data set, and the reproducibility of wave number areas selected within the models. Second, the models have to be developed with standardized spectra to be compatible with the reference spectral response and to be

used by all the spectrometers. Third, there is a need for a thorough investigation of the spectral stability of individual instruments over time, as this could potentially affect the predictions within 2 interlaboratory studies. The developed method makes it possible to harmonize a network precisely and hence to constitute a standardized historical database usable for multiple purposes. A new model predicting an interesting phenotype can be applied to past standardized spectra to take advantage of a depth of data (e.g., to realize a genetic study). However, the developed method is valid only once a slave instrument has been integrated in the network and has analyzed the standardization samples. Recently, a study aimed to standardize spectra over instruments and over time using historical data sets as a basis (Bonfatti et al., 2017a). This method allows the harmonizing of historical databases and the use a posteriori of models when instruments have not been standardized. However, it is concluded in this work that to guarantee the correct application of the calibration models on a running spectrometer the instrument should be standardized using the traditional standardization methods, which make use of spectra acquired on common reference samples. The potential risk induced by using historical data sets is that it may correct not only differences attributable to instruments but also those attributable to other factors that may affect the different data sets, such as feed diets, breeds, or seasons. The methods are therefore complementary to retroactively standardize an instrument and to precisely harmonize a running network.

Conclusions

The results obtained in this work show that spectral standardization allows the transfer and use of multiple models on all instruments and the improvement of the spectral and prediction reproducibility within the network. The transfer does not add significant error to the final predictions, which are largely affected by the quality of the models used. The method makes the equations universal, thereby offering opportunities for data exchange and the creation and use of common robust models at an international level to provide more information to the dairy sector from milk analysis.

Acknowledgments

The authors thank Interreg (European Regional Development Fund) and the Walloon region for their financial support of this project. The authors are grateful to European Milk Recording (Ciney, Belgium) and particularly Xavier Massart for the management of the network as well as to the laboratories participating in the interlaboratory studies: Qualitätslabor NÖ (Gmünd, Austria), Milchprüfing Oberösterreich (Ried, Austria), QLM–Qualitätslabor St. Michael (Sankt Michael, Austria), and Prüfstelle Labor Rotholz (Jenbach, Austria); Comité du lait (Battice, Belgium) and Centre Wallon de Recherches Agronomiques (Gembloux, Belgium); Central Milk Testing Lab of Edmonton (Edmonton, Canada); Adecl62 (Maroeuil, France), Analis 53 (St Berthevin, France), Cialso (Auch, France), Galilait (Clermont-Ferrand, France), Labilait (Aumale, France), Lacolait (Roulans, France), LDA39 (Poligny, France), Lial (Rioz, France), Lial-Mc (Aurillac, France), Lilano (Saint Lo, France), Lilco (Surgère, France), Urcil (Carhaix, France), and Uriane (La Capelle, France); MPR–Kirchheim (Kirchheim, Germany) and LKV-NRW (Krefeld,

Germany); Teagasc (Moorepark, Ireland); Convis (Ettelbrück, Luxembourg); NMR–Wolverhampton (Wolverhampton, UK), NMR–Glasgow (Glasgow, UK), and IML (Bailieborough, United Kingdom); and Suisselab (Zollikofen, Switzerland). The authors are especially indebted to Claire Darimont, Petimat Kitaeva, Matthieu Dubuisson, Mohamed El Morabit, and Olivier Genard for their valuable work on the organization of the interlaboratory study (CRA-W, Gembloux, Belgium).

References

Bonfatti V., Di Martino G., and Carnier P., 2011. Effectiveness of mid-infrared spectroscopy for the prediction of detailed protein composition and contents of protein genetic variants of individual milk of Simmental cows. *J. Dairy Sci.* 94:5776–5785.

Bonfatti V., Fleming A., Koeck A., and Miglior F., 2017a. Standardization of milk infrared spectra for the retroactive application of calibration models. *J. Dairy Sci.* 100:2032–2041.

Bonfatti V., Vicario D., Lugo A., and Carnier P., 2017b. Genetic parameters of measures and population-wide infrared predictions of 92 traits describing the fine composition and technological properties of milk in Italian Simmental cattle. *J. Dairy Sci.* 100:5526–5540.

Cecchinato A., De Marchi M., Gallo L., Bittante G., and Carnier P., 2009. Mid-infrared spectroscopy predictions as indicator traits in breeding programs for enhanced coagulation properties of milk. *J. Dairy Sci.* 92:5304–5313.

Colinet F., Troch T., Baeten V., Dehareng F., Dardenne P., Sindic M., and Gengler N., 2015. Genetic variability of MIR predicted milk technological properties in Walloon dairy cattle. Page 400 in *Book of Abstracts of the 66th Annual Meeting of the European Federation of Animal Science*. Wageningen Academic Publishers, Wageningen, the Netherlands.

Dal Zotto R., De Marchi M., Cecchinato A., Penasa M., Cassandro M., Carnier P., and Bittante G., 2008. Reproducibility and repeatability of measures of milk coagulation properties and predictive ability of mid-infrared reflectance spectroscopy. *J. Dairy Sci.* 91:4103–4112.

De Marchi M., Fagan C., O'Donnell C., Cecchinato A., Dal Zotto R., Cassandro M., and Bittante G., 2009. Prediction of coagulation properties, titratable acidity, and pH of bovine milk using midinfrared spectroscopy. *J. Dairy Sci.* 92:423–432.

De Marchi M., Toffanin V., Cassandro M., and Penasa M., 2014. Invited review: Mid-infrared spectroscopy as phenotyping tool for milk traits. *J. Dairy Sci.* 97:1171–1186.

de Roos A.P.W., van den Bijgaart H.J.C.M., Hørlyk J., and de Jong G., 2007. Screening for subclinical ketosis in dairy cattle by Fourier transform infrared spectrometry. *J. Dairy Sci.* 90:1761–1766.

Dehareng F., Delfosse C., Froidmont E., Soyeurt H., Martin C., Gengler N., Vanlierde A., and Dardenne P., 2012. Potential use of milk mid-infrared spectra to predict individual methane emission of dairy cows. *Animal* 6:1694–1701.

Gengler N., Wiggans G.R., and Gillon A., 2004. Estimated heterogeneity of phenotypic variance of test-day yield with a structural variance model. *J. Dairy Sci.* 87:1908–1916.

Grelet C., Fernández Pierna J.A., Dardenne P., Baeten V., and Dehareng F., 2015. Standardization of milk mid-infrared spectra from a European dairy network. *J. Dairy Sci.* 98:2150–2160.

Grelet C., Fernández Pierna J.A., Dardenne P., Soyeurt H., Vanlierde A., Colinet F., Gengler N., Baeten V., and Dehareng F., 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, β -hydroxybutyrate and citrate contents in bovine milk through a European dairy network. *J. Dairy Sci.* 99:4816–4825.

Hempstalk K., McParland S., and Berry D. P., 2015. Machine learning algorithms for the prediction of conception success to a given insemination in lactating dairy cows. *J. Dairy Sci.* 98:5262–5273.

ISO (International Organization for Standardization. 1999. Whole milk—Determination of milk fat, protein and lactose content—Guidance on the operation of mid-infrared instruments. Standard number 9622:1999. ISO, Geneva, Switzerland.

Lainé A., Bastin C., Grelet C., Hammami H., Colinet F. G., Dale L. M., Gillon A., Vandenplas J., Dehareng F., and Gengler N., 2017. Assessing the effect of pregnancy stage on milk composition of dairy cows using mid-infrared spectra. *J. Dairy Sci.* 100:2863–2876.

Leclercq G., Gengler N., Soyeurt H., and Bastin C., 2013. Genetic variability of the mid-infrared prediction of lactoferrin content in milk for Walloon Holstein first-parity cows. *Livest. Sci.* 151:158–162.

McParland S., Banos G., Wall E., Coffey M.P., Soyeurt H., Veerkamp R.F., and Berry D. P., 2011. The use of mid-infrared spectrometry to predict body energy status of Holstein cows. *J. Dairy Sci.* 94:1222–1239.

McParland S., Kennedy E., Lewis E., Moore S. G., McCarthy B., O'Donovan M., and Berry D. P., 2015. Genetic parameters of dairy cow energy intake and body energy status predicted using midinfrared spectrometry of milk. *J. Dairy Sci.* 98:3651–3661.

McParland S., Lewis E., Kennedy E., Moore S. G., McCarthy B., O'Donovan M., Butler S. T., Pryce J. E., and Berry D. P., 2014. Mid-infrared spectrometry of milk as a predictor of energy intake and efficiency in lactating dairy cows. *J. Dairy Sci.* 97:5863–5871.

Rutten M.J.M., Bovenhuis H., Hettinga K. A., van Valenberg H. J. F., and van Arendonk J. A. M., 2009. Predicting bovine milk fat composition using infrared spectroscopy based on milk samples collected in winter and summer. *J. Dairy Sci.* 92:6202–6209.

Soyeurt H., Bruwier D., Romnee J.-M., Gengler N., Bertozzi C., Veselko D., and Dardenne P., 2009. Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 92:2444–2454.

Soyeurt H., Colinet F., Arnould V., Dardenne P., Bertozzi C., Renaville R., Portetelle D., and Gengler N., 2007. Genetic variability of lactoferrin content estimated by mid-infrared spectrometry in bovine milk. *J. Dairy Sci.* 90:4443–4450.

Soyeurt H., Dehareng F., Gengler N., McParland S., Wall E., Berry D. P., Coffey M., and Dardenne P., 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *J. Dairy Sci.* 94:1657–1667.

Soyeurt H., Dardenne P., Dehareng F., Lognay G., Veselko D., Marlier M., and Gengler N., 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 89:3690–3695.

van Knegsel A.T.M., van der Drift S.G.A., Horneman M., de Roos A.P.W., Kemp B., and Graat E.A.M., 2010. Short communication: Ketone body concentration in milk determined by Fourier transform infrared spectroscopy: Value for the detection of hyperketonemia in dairy cows. *J. Dairy Sci.* 93:3065–3069.

Vanlierende A., Vanrobays M.L., Gengler N., Dardenne P., Froidmont E., Soyeurt H., McParland S., Lewis E., Deighton M. H., Mathot M., and Dehareng F., 2016. Milk mid-infrared spectra enable prediction of lactation-stage dependent methane emissions of dairy cattle within routine population-scale milk recording schemes. *Anim. Prod. Sci.* 56:258–264.

Wang Q., Hulzebosch A., and Bovenhuis H., 2016. Genetic and environmental variation in bovine milk infrared spectra. *J. Dairy Sci.* 99:6793–6803. <https://doi.org/10.3168/jds.2015-10488>. Williams, P. 2014. Tutorial: The RPD statistic: A tutorial note. *NIR News* 25:22–26.

Chapter 4

Development of Fourier transform mid-infrared calibrations to predict acetone, β -hydroxybutyrate, and citrate contents in bovine milk through a European dairy network

Introduction to chapter 4

While the two previous chapter focused on the practical conditions to use in routine models developed from milk MIR spectra, the following chapter aims to concretize this work by describing the development of models predicting phenotypes of interest for the dairy sector.

Among all the phenotypes of interest, difficulties arising in early lactation due to negative energy balance (NEB) are among the most important issues. Collard et al. (2000) and Butler (2003) identified the length and intensity of postpartum NEB as major parameters influencing health and fertility in dairy cows. Following Suthar et al. (2013), 75% of dairy cows' diseases occur during the first month of lactation. The NEB is a consequence of an increased milk production per cow, partly due to a genetic selection for milk production, while dry matter intake failed to increase enough to compensate for the higher energy expenditure (Barkema et al., 2015). Physiologically, altered metabolism due to NEB in early lactation causes liver damage and dysfunctions (Turk et al., 2004), inflammation (Wathes et al., 2009), alteration of hormone regulation (Esposito et al., 2014) and immune response (Hammon et al., 2006; Moyes et al., 2010). Consequently, this increases the risk of ketosis, lameness, displaced abomasum, milk fever, metritis, retained placenta and mastitis (Collard et al., 2000; LeBlanc, 2010; Esposito et al., 2014). Among these disorders, subclinical ketosis is the metabolic disease with the highest prevalence and represents a gateway condition for other metabolic and infectious diseases (Suthar et al., 2013). In the US dairy herds, it has been the most important metabolic disease, surpassing ruminal acidosis and milk fever in clinical significance since 1990 (Oetzel, 2007). Ketosis is characterized by increased levels of circulating ketone bodies and occurs when physiologic mechanisms for the adaptation to NEB fail (Herdt, 2000). In addition to metabolic and infectious diseases, NEB can also reduce reproductive performances (Esposito et al., 2014). Issues in dairy farms associated with NEB can represent an important source of economic losses. As an example, total cost per case of hyperketonemia are estimated at 289\$ in average (McArt et al., 2015). Considering the associated costs, as well as the high incidence of problems linked to NEB and ketosis, there is a clear interest to have information on these phenotypes in the early lactation period.

Among the biomarkers of these statuses, blood Non-Esterified Fatty acids (NEFA) and β -hydroxybutyrate (BHB) have been identified as gold standards to reflect body reserve mobilization and ketosis, respectively (Leblanc, 2010; Suthar et al., 2013). Alternatively, some metabolites measured in milk are also known to be good indicators for NEB and ketosis and readily accessible at large scale in dairy herds via milk recording organizations (MROs). The high frequency of milk sampling by the MROs could allow detecting an animal in suboptimal condition before turning into observable clinical cases. In the study by BjerreHarpøth et al. (2012), it was concluded that citrate content had the greatest response during NEB between various metabolites measured in milk. Baticz et al. (2002) stated that in order to evaluate the energy status of cows, content of citrate should be measured by simple and automated methods such as FT-MIR technology. In addition, ketone bodies in milk, such as BHB and acetone, have been identified as the most common indicators of (sub)clinical ketosis in milk

(Enjalbert et al., 2001). Consequently, the objective of the following chapter is to build calibration models to predict milk BHB and acetone using the milk MIR spectra, and assess the possibility of predicting milk citrate through this technology. This work could be the basis for the creation of useful tools for dairy herd management based on a cost-effective technology easily accessible through the classical milk recording system. The following chapter is based on the results described in Chapters 2 & 3, considering that datasets from several instruments have been combined into a common calibration database using the standardization method. Finally, due to the universal standardized spectral format, such tools could be used routinely on all network instruments.

References

- Baticz O., Tömösközi S., Vida L., Gaál T., 2002. Relationship between concentration of citrate and ketone bodies in cow's milk. *Acta Vet. Hung.* 50:253–61.
- Bjerre-Harpøth V., Friggens N.C., Thorup V.M., Larsen T., Damgaard B.M., Ingvarsen K.L., Moyes K.M., 2012. Metabolic and production profiles of dairy cows in response to decreased nutrient density to increase physiological imbalance at different stages of lactation. *J. Dairy Sci.* 95:2362–80.
- Butler W.R., 2000. Nutritional interactions with reproductive performance in dairy cattle. *Anim. Reprod. Sci.* 60, 449-457.
- Collard B.L., Boettcher P.J., Dekkers J.C., Petitclerc D., Schaeffer L.R., 2000. Relationships between energy balance and health traits of dairy cattle in early lactation. *J. Dairy Sci.* 83,2683–2690.
- Enjalbert F., Nicot M.C., Bayourthe C., Moncoulon R., 2001. Ketone bodies in milk and blood of dairy cows: relationship between concentrations and utilization for detection of subclinical ketosis. *J. Dairy Sci.* 84:583–9.
- Esposito G., Irons P.C., Webb E.C., Chapwanya A., 2014. Interactions between negative energy balance, metabolic diseases, uterine health and immune response in transition dairy cows. *Anim. Reprod. Sci.* 144:3,60-71.
- Hammon D.S., Evjen I.M., Dhiman T.R., Goff J.P., Walters J.L., 2006. Neutrophil function and energy status in Holstein cows with uterine health disorders. *Vet. Immunol. Immunop.* 113:1,21-29.
- Herdt H., 2000, Ruminant Adaptation to Negative Energy Balance. *Vet Clin. N. AM: Food animal Practice.* 16:2,215-230.
- Koeck A., Jamrozik J., Schenkel F. S., Moore R. K., Lefebvre D. M., Kelton D. F., Miglior F., 2014. Genetic analysis of milk β -hydroxybutyrate and its association with fat-to-protein ratio, body condition score, clinical ketosis, and displaced abomasum in early first lactation of Canadian Holsteins. *J. Dairy Sci.* 97:11,7286-7292.
- LeBlanc S., 2010. Monitoring metabolic health of dairy cattle in the transition period. *J. Reprod. Develop.* 56:S29-S35.
- McArt J.A.A., Nydam D.V., Overton M.W., 2015. Hyperketonemia in early lactation dairy cattle: A deterministic estimate of component and total cost per case. *J. Dairy Sci.* 98:3, 2043-2054.

Moyes K.M., Drackley J.K., Morin D.E., Rodriguez-Zas S.L., Everts R.E., Lewin H.A., Looor J.J., 2010. Mammary gene expression profiles during an intramammary challenge reveal potential mechanisms linking negative energy balance with impaired immune response. *Physiol. Genomics*, 41:2, 161-170.

Oetzel G.R., 2007, September. Herd-level ketosis–diagnosis and risk factors. In *Proceedings of the 40th annual conference of bovine practitioners*, Vancouver, Canada.

Suthar V.S., Canelas-Raposo J., Deniz A., Heuwieser W., 2013. Prevalence of subclinical ketosis and relationships with postpartum diseases in European dairy cows. *J. Dairy Sci*, 96:5,2925-2938.

Turk R., Juretic D., Geres D., Turk N., Rekić B., Simeon-Rudolf V., Svetina A., 2004. Serum paraoxonase activity and lipid parameters in the early postpartum period of dairy cows. *Res. Vet. Sci.*, 76:1, 57-61.

Wathes D.C., Cheng Z., Chowdhury W., Fenwick M.A., Fitzpatrick R., Morris D.G., Patton J., Murphy J.J., 2009. Negative energy balance alters global gene expression and immune responses in the uterus of postpartum dairy cows. *Physiol. Genomics.*, 39:1,1-13.

Chapter 4 : Development of Fourier transform mid-infrared calibrations to predict acetone, β -hydroxybutyrate, and citrate contents in bovine milk through a European dairy network

C. Grelet,*¹ C. Bastin,†¹ M. Gelé, J.B. Davière,‡ M. Johan,‡ A.Werner,# R.Reding, J.A. Fernandez Pierna,* F. G. Colinet,† P. Dardenne,* N. Gengler,† H. Soyeurt,† and F. Dehareng*²

*Walloon Agricultural Research Center (CRA-W), Valorization of Agricultural Products Department, 24 Chaussée de Namur, 5030 Gembloux, Belgium

†University of Liège, Gembloux Agro-Bio Tech, Agriculture, Bio-engineering and Chemistry Department, 2 Passage des Déportés, 5030 Gembloux, Belgium

‡French Livestock Institute (IDELE), 9 rue André Brouard, CS 70510, 49105 Angers cedex 02, France §Clasel, 141 Boulevard des Loges, 53942 Saint Berthevin, France

#Landeskontrollverband (LKV) Baden Württemberg, Heinrich-Baumann Str. 1-3, 70190 Stuttgart, Germany

lConvis S.C., 4 Zone Artisanale et Commerciale, 9085 Ettelbruck, Luxembourg

¹ These authors contributed equally

J. Dairy Sci. 99:4816–4825 <http://dx.doi.org/10.3168/jds.2015-10477>

© American Dairy Science Association®, 2016

Abstract

To manage negative energy balance and ketosis in dairy farms, rapid and cost-effective detection is needed. Among the milk biomarkers that could be useful for this purpose, acetone and β -hydroxybutyrate (BHB) have been proved as molecules of interest regarding ketosis and citrate was recently identified as an early indicator of negative energy balance. Because Fourier transform mid-infrared spectrometry can provide rapid and cost-effective predictions of milk composition, the objective of this study was to evaluate the ability of this technology to predict these biomarkers in milk. Milk samples were collected in commercial and experimental farms in Luxembourg, France, and Germany. Acetone, BHB, and citrate contents were determined by flow injection analysis. Milk mid-infrared spectra were recorded and standardized for all samples. After edits, a total of 548 samples were used in the calibration and validation data sets for acetone, 558 for BHB, and 506 for citrate. Acetone content ranged from 0.020 to 3.355 mmol/L with an average of 0.103 mmol/L; BHB content ranged from 0.045 to 1.596 mmol/L with an average of 0.215 mmol/L; and citrate content ranged from 3.88 to 16.12 mmol/L with an average of 9.04 mmol/L. Acetone and BHB contents were log-transformed and a part of the samples with low values was randomly excluded to approach a normal distribution. The 3 edited data sets were then randomly divided into a calibration data set (3/4 of the samples) and a validation data set (1/4 of the samples). Prediction equations were developed using partial least square regression. The coefficient of determination (R^2) of cross-validation was 0.73 for acetone, 0.71 for BHB, and 0.90 for citrate with root mean square error of 0.248, 0.109, and 0.70 mmol/L, respectively. Finally, the external validation was performed and R^2 obtained were 0.67 for acetone, 0.63 for BHB, and 0.86 for citrate, with

respective root mean square error of validation of 0.196, 0.083, and 0.76 mmol/L. Although the practical usefulness of the equations developed should be further verified with other field data, results from this study demonstrated the potential of Fourier transform mid-infrared spectrometry to predict citrate content with good accuracy and to supply indicative contents of BHB and acetone in milk, thereby providing rapid and cost-effective tools to manage ketosis and negative energy balance in dairy farms.

Key words: Fourier transform mid-infrared spectrometry, milk, acetone, β -hydroxybutyrate, citrate

Introduction

Fourier transform mid-infrared (FT-MIR) spectrometry is a method of choice to perform composition and quality controls during routine liquid milk testing. It allows a fast and nondestructive quantification of milk chemical properties to avoid reference methods, which are usually tedious, expensive, and time consuming. Today, FT-MIR spectrometry is used worldwide to predict contents of fat, protein, urea, and lactose in official milk-recording schemes and milk payment systems. In addition, several studies undertaken over the last decade demonstrated the potential of FT-MIR to predict detailed milk composition (De Marchi et al., 2014), for example, fatty acid profile (Rutten et al., 2009; Soyeurt et al., 2011), protein composition (Bonfatti et al., 2011), lactoferrin (Soyeurt et al., 2012), minerals (Soyeurt et al., 2009), technological properties of milk (e.g., coagulation properties, curd firmness, and cheese yield; De Marchi et al., 2014), and the physiological state of the cow (e.g., methane emissions; Vanlierde et al., 2015), pregnancy status (Lainé et al., 2014), body energy status (McParland et al., 2011), energy intake, and efficiency (McParland et al., 2014). Hence the analysis of milk by FT-MIR spectrometry offers the opportunity to record a whole range of new phenotypes to develop tools enabling profitability and sustainability of the dairy sector (Gengler et al., 2015) as well as genetic and genomic evaluations (Gengler et al., 2016). Over the range of traits potentially predictable by FT-MIR spectrometry, biomarkers of negative energy balance state and ketosis are of primary importance for optimized fertility, health, and welfare of high-yielding dairy cows. The extent and the duration of the postpartum negative energy balance have been identified as one of the major factor influencing fertility and health of dairy cows (Collard et al., 2000; Butler, 2003). BjerreHarpøth et al. (2012) mentioned that citrate content, among various metabolites measured in milk, had the greatest response during a period of negative energy balance. Baticz et al. (2002) concluded that content of citrate should be measured by easy and automated method such as FT-MIR technology to evaluate the energy status of cows. Furthermore, (sub)clinical ketosis, caused by excessive body fat mobilization due to severe negative energy balance, is one of the most frequent production diseases, with a prevalence ranging from 7 to 43% in the first 2 mo of lactation (Suthar et al., 2013). Ketosis negatively affects milk yield and reproductive performances; it also increases the risk of subsequent diseases such as displaced abomasum (Duffield, 2000). The major ketone bodies in milk (i.e., BHB, acetone, and acetoacetate) are the most common indicators of ketosis in milk (Enjalbert et al., 2001).

Previous studies attempted to predict acetone content in milk by using FT-MIR spectrometry. As shown in Table 4-1, these studies differ in the reference method used to quantify acetone content but also in the number of samples used. Coefficients of determination of calibration and of cross-validation ranged from 0.39 to 0.80 (Table 4-1). Except Hansen (1999), these authors did not perform external validation. de Roos et al. (2007) investigated the prediction of BHB by FT-MIR spectrometry. Using 1,069 samples, they obtained a R^2 of cross-validation of 0.63 and a standard error of crossvalidation of 0.065 mmol/L. Using FT-MIR predictions of acetone and BHB contents in milk, van Knegsel et al. (2010) and van der Drift et al. (2012) investigated the opportunity of such predictions for the detection of ketosis in dairy cows. To our knowledge, no study had reported calibration statistics for the prediction of citrate contents in milk by FT-MIR spectrometry although FT-MIR predictions of citrate have been used in the work of Bjerre-Harpøth et al. (2012). The objectives of this study were (1) to build calibrations predicting acetone and BHB contents in milk and to evaluate their usefulness for use on field with an external validation data set and (2) to assess the potential of FT-MIR spectrometry to predict citrate content in milk. The first novelty of this work lies in the combination of data from cows of different breeds collected in different countries and production systems as well as in the combination of spectral data from several apparatus of different brands. The merging of spectral data relies on the FT-MIR standardization procedure developed by Grelet et al. (2015). This method brings possible the collation of the data set, thereby increasing the robustness of calibrations and the use of the developed calibrations by all standardized instruments. The second novelty of this work is to provide a detailed external validation procedure to assess the robustness of the calibrations developed.

Table 4-1. Overview of calibration and validation statistics from various studies aiming at predicting acetone content in milk by Fourier transform mid-infrared spectrometry; reference method for quantifying acetone, number of samples used, root mean square error (RMSE), standard error of cross-validation (SECV), and R^2 are presented¹

Reference	Reference method	Calibration			Cross validation		External validation			
		N	RMSE	R^2	SECV	RMSE	R^2	N	RMSE	R^2
Hansen et al., 1999	Vanilin test	302	-	-	-	0.240	0.80	58	0.270	0.81
Heuer et al., 2001	Gas chromatography	180	-	-	0.210	-	-	-	-	-
De Roos et al., 2007	Continuous flow analyser	1063	-	-	0.184	-	0.72	-	-	-
Hanus et al., 2011	Microdiffusion photometric	14	-	0.65	-	-	-	-	-	-
Hanus et al., 2014	Microdiffusion photometric	89	-	0.39	-	-	-	-	-	-

¹RMSE and SECV are expressed in mmol/L.

Materials and methods

Sampling

The sampling was undertaken from August 2013 to June 2014. Milk samples were collected by 4 organizations from 3 countries in both research and commercial farms (Table 4-2). After removing subsets of samples thawed during transport or poorly preserved and subsets of samples showing issues during spectra acquisition or during reference analysis, the initial data set included 566 samples. A total of 256 milk samples originated from 2 experimental farms from cows at 7 to 56 d postcalving to focus on the postpartum period when cows are at the greatest risk of metabolic disorders due to negative energy balance. Eighty-two samples were collected in Neumühle experimental farm (Germany) from Holstein cows fed mainly with maize silage. A total of 174 samples were recorded in the Poisy experimental farm (France) from Abondance and Montbéliarde cows fed with fresh grass during summer or with hay and maize during winter season. In addition, 310 samples were collected in commercial farms by Milk Recording Organizations. Several 200 samples were collected by Clasel in the west of France, from Holstein and Normande cows fed mainly with maize silage during winter and grass during summer. Samples were selected at the laboratory based on milk parameters known to be related to ketosis status such as fat to protein ratio. In Luxembourg, 110 samples were collected by Convis s.c. from Holstein cows fed mainly with maize silage supplemented by grazing during summer. For all cows, milk samples were collected by using sampling systems approved by ICAR. Morning and evening samples were pooled to obtain daily milk and 2 identical samples were generated to be analyzed by FT-MIR and reference analysis.

Table 4-2. Characteristics of the 4 sample source

	N	Region	Breed	Feed	DIM
Neumühle	82	Western Germany	Holstein	Maize silage	7-56
Poisy	174	Eastern France	Abondance and Montbéliarde	Fresh grass or hay and maize silage	7-56
Clasel	200	Western France	Holstein and Normande	Maize silage or fresh grass	7-305
Convis	110	Luxembourg	Holstein	Maize silage supplemented by grazing during summer	5-60

Fourier Transform Mid-Infrared Analysis

All samples were analyzed fresh by local milk recording organizations. A total of 10 different instruments, located in 5 laboratories, were used in the study. The instruments were located in Germany (6), France (2), and Belgium (2). Table 4-3 provides the technical information related to the instruments used in this study. To combine them into a common database, the spectra recorded from all these instruments were standardized into a common format using the piece-wise direct standardization method and the protocol developed within the OptiMIR project

(Grelet et al., 2015). For each spectrum, the standardized Mahalanobis distance (GH) was calculated and no spectrum was considered as an outlier.

Table 4-3. Characteristics of the Fourier transform mid-infrared instruments used in the study¹

Brand	Type	Number of instruments	Frequency reported by constructors (cm ⁻¹)	Number of Wavenumbers	Resolution
Foss Electric A/S (Denmark)	FT 6000	3	926 – 5,010	1060	Unknown
Foss Electric A/S (Denmark)	FT+	2	926 – 5,010	1060	Unknown
Bentley (United States)	FTS	5	649 – 3,999	899	8cm ⁻¹

¹Number of wavenumbers is the total data points numbers of the raw spectra before interpolation and area selection.

Reference Analysis

The samples dedicated to the reference analysis were sent to the Walloon Agricultural Research Center (CRAW, Gembloux, Belgium) to be analyzed. Samples from Germany and France were frozen before being sent, and the samples from Luxembourg were preserved with bronopol at 0.01% and analyzed fresh. The BHB and acetone content in milk were analyzed by a continuous flow analyzer (Scan ++, Skalar, Breda, the Netherlands) using the procedure described by de Roos et al. (2007). Citrate was also analyzed with continuous flow analyzer, based on enzyme-catalyzed reaction. Citrate is converted into oxaloacetate and acetate, catalyzed by citrate lyase, and oxaloacetate decarboxylases into pyruvate. Oxaloacetate and pyruvate are then reduced by nicotinamide dinucleotide (NADH) into malate and lactate, which are catalyzed by malate and lactase dehydrogenase. Decrease of NADH is stoichiometric with citrate content, and the remaining NADH is measured by optical density at 340 nm. Ranges of analysis of the continuous flow analyzer provided by the supplier are 0.02 to 1 mmol/L for acetone, 0.04 to 2 mmol/L for BHB, and 0.03 to 24 mmol/L for citrate; otherwise values are estimated by extrapolation. All samples were analyzed twice, and samples with variation higher than 5% were re-analyzed. Standard error of laboratory, which is the repeatability of the reference method, was calculated. From the 566 samples, for each component the samples with missing value or with value under the lower limit of detection of the respective analysis were removed, leading to 558, 548, and 506 samples, respectively, for BHB, acetone, and citrate data set.

Calibration and Validation

Preliminary statistics indicated that acetone and BHB values were not normally distributed, with a higher proportion of low values. When performing calibration, this type of distribution gives too much weight to the low values, leading to a low accuracy in predicting high values. Therefore, editing of data was needed to use a more

balanced data set between low and high values. Several samples in the cloud of low values were randomly removed to obtain a reduced data set covering the same range of values, but giving less weight on low values. Visual inspection of the data indicated that this large number of low values was situated between 0.025 and 0.085 mmol/L for acetone. Hence, 66% of these data was randomly removed. After this edit, the number of samples for acetone was 224. The same process was used for BHB values for which the large number of samples with low values was situated between 0.100 and 0.250 mmol/L, thereby bringing the number of data to 434. A logarithmic (10) transformation was then applied on both acetone and BHB reference values to approach a normal distribution. Figure 4-1 shows the distribution of the acetone data set before and after the logarithmic transformation and after removing randomly a part of the low values.

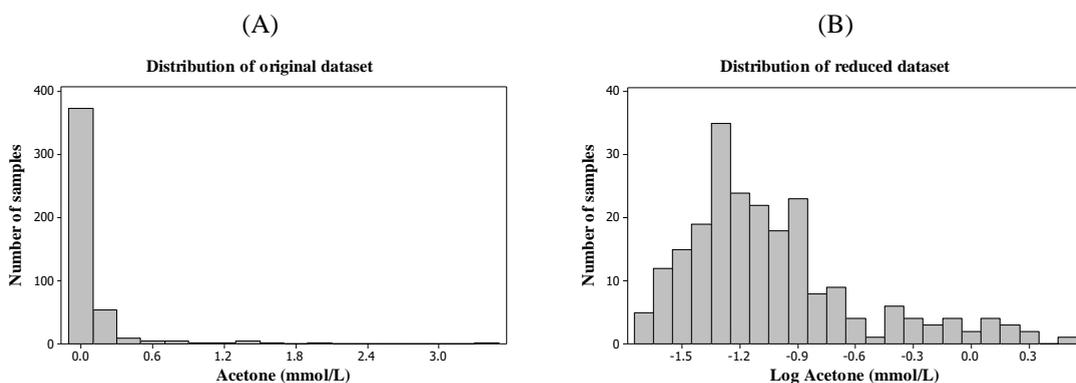


Figure 4-1. Effect of editing the acetone data set on the distribution of values: (A) Original data set after preliminary editing ($n = 548$), (B) data set after logarithmic transformation and removing randomly low values ($n = 224$).

The citrate reference data set was normally distributed, and it was therefore not further edited. For all data sets, a quarter of the data was randomly excluded in the calibration process to be used as an external validation. As pretreatment of FT-MIR spectra, a first derivative was used with a gap of 5 wavenumbers, associated with an autoscale preprocess only for acetone and BHB. The spectral area selected were 968.1 to 1,577.5 cm^{-1} , 1,731.8 to 1,762.6 cm^{-1} , 1,781.9 to 1,808.9 cm^{-1} , and 2,831.0 to 2,966.0 cm^{-1} . Detailed evaluation of the spectra based on previous knowledge has permitted the selection of those wavenumber bands. Noisy areas induced by water were removed. And to bring the models as robust as possible, only the wavenumbers with a spectral response highly correlated between different instruments when analyzing common samples were used. Calibrations were done using partial least square (PLS) regression. Cross-validations were performed on the calibration data sets using 10 subsets randomly constituted, and samples with residuals higher than 2.5 times the SD of the global residuals were considered as outliers (Rousseeuw et al., 2006). Models were then applied on the external validation data sets. Both cross-validation and validation results were expressed in terms of R^2 , root mean square error (RMSE), and ratio performance/deviation (RPD). The goal of the RPD criterion is to

show simultaneously the accuracy of predictions and the global variability of the reference values (Williams and Sobering, 1993). The RPD is defined as the ratio SD/RMSE as RMSE can be the one calculated in cross-validation or the one of a validation set. When the RPD is between 1.5 and 2, the model can discriminate low from high values; a RPD between 2 and 2.5 indicates that rough quantitative predictions are possible, and a RPD between 2.5 and 3 or above corresponds to good and excellent prediction accuracy (Nicolai et al., 2007). Then, to illustrate the usefulness of the developed equations for the detection of subclinical ketosis, the percentage of data well-classified into 2 classes (low vs. high values) was calculated based on a threshold 0.15 mmol/L for acetone (de Roos et al., 2007), and the same classification was done with a threshold of 0.2 mmol/L for BHB (Denis-Robichaud et al., 2014). All computations and chemometric analysis were carried out with programs developed in Matlab v7.5.0. (The Mathworks Inc., Natick, MA) and the PLS toolbox v. 4.11 (Eigenvector Research Inc., Wenatchee, WA).

Results and discussion

Reference Analysis

Descriptive statistics of the initial data set are provided in Table 4-4. Content of BHB in the milk samples analyzed ranged from 0.045 to 1.595 mmol/L with an average of 0.215 mmol/L and a standard deviation of 0.174 mmol/L. On average, BHB content was higher in the present data set than in the data set of de Roos et al. (2007; average of 0.146 mmol/L with value ranging from -0.021 to 3.960) and of Denis-Robichaud et al. (2014; average of 0.18 mmol/L with value ranging from -0.03 to 1.09 mmol/L). Acetone content in the milk samples analyzed ranged from 0.020 to 3.355 mmol/L with an average of 0.103 mmol/L and a standard deviation of 0.260 mmol/L. These values were in the same range than values presented by de Roos et al. (2007; from -0.021 to 3.960 mmol/L with an average of 0.146 mmol/L) and by Denis-Robichaud et al. (2014; from -0.03 to 2.63 with an average of 0.100 mmol/L).

Table 4-4. Descriptive statistics of the results from the reference analysis for BHB, acetone, and citrate

Component	Unit	#N	Min	Max	Mean	SD	SEL ¹
BHB	mmol/L	558	0.045	1.596	0.215	0.174	0.005
Acetone	mmol/L	548	0.020	3.355	0.103	0.260	0.006
Citrate	mmol/L	506	3.88	16.12	9.04	2.21	0.216

¹SEL = standard error of laboratory.

Citrate content in milk samples analyzed ranged from 3.88 to 16.12 mmol/L with an average of 9.04 mmol/L and a standard deviation of 2.21 mmol/L. Calculated standard errors of the laboratory were respectively 0.005, 0.006, and 0.216 mmol/L for BHB, acetone, and citrate, meaning that the reference method was precise and did not affect the statistics obtained in the calibration step. Considering their respective molar mass of 104.11, 58.08, and 192.12 g/mol, these 3 molecules are present in milk on average at a level of 21.7 ppm for BHB, 5.8 ppm for acetone, and 1,684.31 ppm for citrate. Dardenne et al. (2015) mentioned that using FT-MIR technology,

constituents cannot be detected below 100 ppm. Therefore, it is worth noting that calibration of BHB and acetone contents in milk cannot be done by the specific spectral response of these molecules in milk but by indirect links with global milk composition.

BHB Cross-Validation and Validation Results

After edits to obtain a more balanced data set between low and high values, the BHB data set contained 433 samples of which 325 samples were randomly included into the calibration data set. Because the best results were obtained using a log-transformation of the reference data, only these results are presented here. A PLS model was done using 8 latent variables, and 7 samples were considered as outliers and discarded. In the calibration data set, after removing the outliers, BHB content ranged from 0.045 to 1.595 mmol/L with an average of 0.235 mmol/L and a standard deviation of 0.193 mmol/L. The average of predicted values is 0.219 mmol/L, which is slightly smaller than the average of reference values; this is due to the slope between reference and predicted values of 1.1842 that can be observed in Figure 4-2.A.

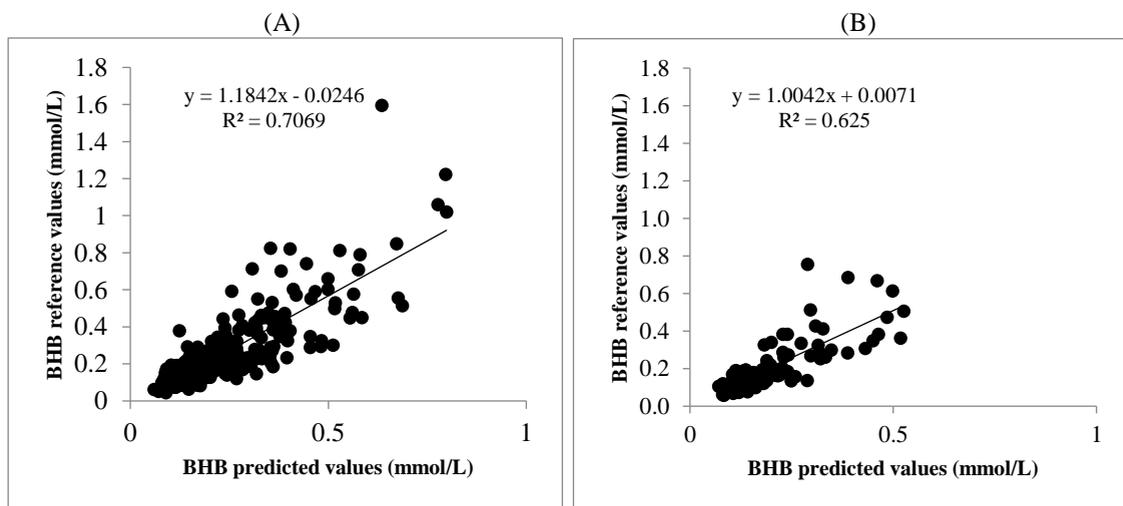


Figure 4-2. Plot of BHB reference values from flow injection analysis and BHB values predicted from Fourier transform mid-infrared analysis (A) for the cross-validation ($n = 325$) and (B) for the validation ($n = 108$).

Table 4-5 shows the cross-validation statistics. The RMSE of cross-validation obtained is 0.109 mmol/L, with an R^2 of cross-validation of 0.71 and a RPD of 1.77. The error is relatively high due to the lack of precision of the model on high values of the data set (Figure 4-2.A), combined with the artificial removal of a series of samples with low values. The bias on high values might be induced by the logarithmic transformation of the reference data, giving more weight to low values in the model. However, when using a threshold of 0.200 mmol/L, 86.5% of the samples were well classified, with 87.4% of samples with low content of BHB and 85.0% of samples with high content of BHB well classified (Table 4-6), thereby demonstrating the usefulness of this equation to detect cows with an abnormally high content of BHB.

A data set of 108 samples was then used to perform an external validation. The BHB content ranged from 0.058 to 0.755 mmol/L with an average of 0.204 mmol/L and a standard deviation of 0.136 mmol/L (Table 4-5). The average of predicted values, which is 0.198 mmol/L, is comparable to the mean of reference values, showing a slope close to 1 between reference and predicted values in the validation step (Figure 4-2.B).

Table 4-5. Cross-validation and validation statistics for BHB, acetone, and citrate contents in milk¹

Item	N	No. of LV	No. of Outliers	Min	Max	Mean R	SD	Mean P	RMSE	R ²	RPD
BHB (mmol/L)											
Cross-validation	325	8	7	0.045	1.596	0.235	0.193	0.219	0.109	0.71	1.77
Validation	108	-	-	0.058	0.755	0.204	0.136	0.198	0.083	0.63	2.36
Acetone (mmol/L)											
Cross-validation	168	7	2	0.020	3.355	0.190	0.397	0.146	0.248	0.73	1.60
Validation	56	-	-	0.021	1.968	0.179	0.306	0.145	0.196	0.67	2.03
Citrate (mmol/L)											
Cross-validation	380	9	2	3.88	16.12	9.03	2.26	9.02	0.70	0.90	3.21
Validation	126	-	-	4.44	15.16	9.08	2.03	9.10	0.76	0.86	2.96

¹Number of samples used, number of latent variables (LV), descriptive statistics of the reference values [minimum, maximum, mean (mean R), and SD], mean of the predicted values (mean P), root mean square error (RMSE), R², and RPD (ratio SD of calibration/RMSE) are presented.

The RMSE of validation was 0.083 mmol/L, with an R² of 0.63 and a RPD of 2.33. The distribution of the data in the validation data set was slightly different than in the calibration data set, mainly due to reference values in average lower. Considering the lack of precision of the model for high values, this difference between the distribution of the calibration and validation data sets probably explained the difference in the cross-validation and validation statistics. The R² highly depends on the distribution of the data and especially on the range of data (Davies and Fearn, 2006). The R² of validation is lower than the R² of cross-validation probably because of a reduced range of values (Figure 4-2.B). However, the RMSE of validation is better, probably due to a higher proportion of samples with a relatively low content of BHB. Nevertheless, the accuracy shown by the model is satisfying and brings the RPD of validation higher than 2, which is the considered limit to start screening. Therefore, the equation developed in this study can provide an indicative value of the BHB content. When classifying the data of the validation data set into 2 classes by using a threshold of 0.200 mmol/L, 90.8% of the samples were well classified, with 90.9% of samples with low BHB content and 90.6% of samples with high BHB content properly classified (Table 4-6). Therefore, one can conclude that the model is not good enough to provide precise quantitative values of milk BHB, especially when BHB content is elevated, but it can allow discrimination between high and low values of BHB with an acceptable rate of good classification.

Table 4-6. Results of classification of BHB predictions into 2 classes (% of samples; using a threshold of 0.200 mmol/L) for cross-validation and validation data sets

Item (%)	Low BHB content (<0.200 mmol/mL)	High BHB content (>0.200 mmol/L)	Global good classification
Cross-Validation	n=198	n=120	
Predicted low	87.4%	15.0%	86.5%
Predicted high	12.6%	85.0%	
Validation	n=77	n=32	
Predicted low	90.9%	9.4%	90.8%
Predicted high	9.1%	90.6%	

Acetone Cross-Validation and Validation Results

After edits to obtain a data set balanced between low and high values, the acetone data set contained 224 samples of which 168 samples that were randomly selected and included in the calibration data set. The remaining 56 samples were included in the validation data set. In the calibration data set, acetone content ranged from 0.020 to 3.355 mmol/L with an average of 0.190 mmol/L and a standard deviation of 0.397 mmol/L. A PLS model was done using 7 latent variables, and 2 samples were considered as outliers. The average of predicted values is 0.146 mmol/L; this highlights the slope effect of 1.6295 that can be observed in Figure 4-3.A between reference and predicted values.

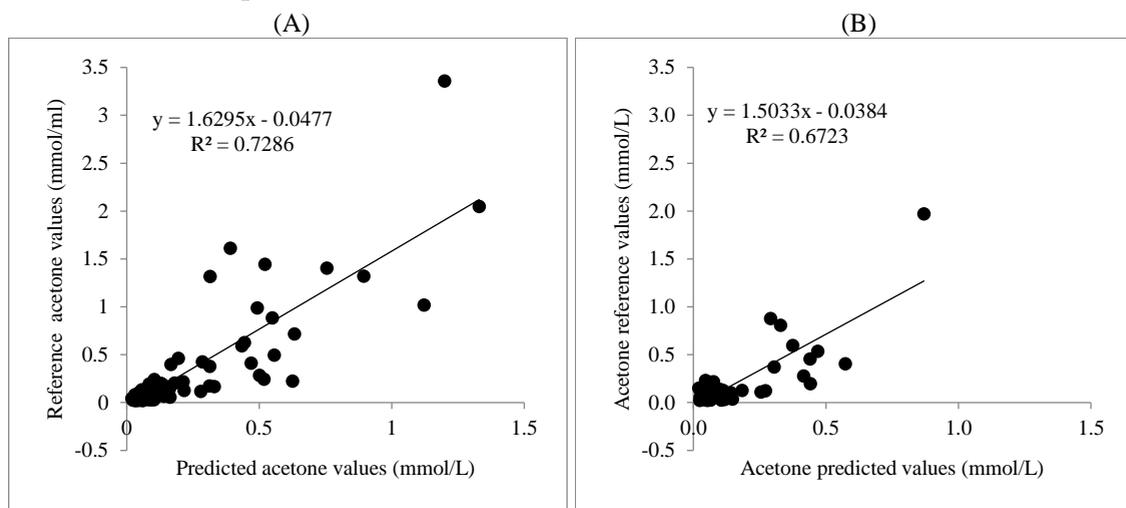


Figure 4-3. Plot of acetone reference values from flow injection analysis and acetone values predicted from Fourier transform mid-infrared analysis (A) for the cross-validation (n = 168) and (B) for the validation (n = 56).

The RMSE of cross-validation was 0.248 mmol/L, with an R^2 of 0.73 and a RPD of 1.60 (Table 4-5). This RMSE was in a similar range to the one obtained by Hansen (1999) in cross-validation (0.240 mmol/L), by Heuer et al. (2001; 0.210 mmol/L), and

de Roos et al. (2007; 0.184 mmol/L). Similarly to the BHB model, the acetone model was relatively imprecise when acetone values were high (Figure 4-3), and this lack of precision on high values affected dramatically the RMSE. However when classifying the data of the validation data set into 2 classes by using a threshold of 0.150 mmol/L, 93.4% of the samples were properly classified, with 95.5% of samples with low acetone content and 84.4% of samples with high acetone content well classified (Table 4-7). In the validation data set, acetone content ranged from 0.021 to 1.968 mmol/L with an average of 0.179 mmol/L and a standard deviation of 0.306 mmol/L. The slope between reference and predicted values is confirmed in validation step with an average of predicted values of 0.145 mmol/L and a slope of 1.5033 (Figure 4-3.B). The RMSE of validation was 0.196 mmol/L, with an R^2 of 0.67 and a RPD of 2.03 (Table 4-5). Even if the R^2 was slightly lower in validation than in crossvalidation (0.67 instead of 0.73), the RMSE was also lower (0.196 instead of 0.248), meaning that the error was lower in the validation data set. Similarly than for BHB, these results can be explained by the distribution of both validation and calibration data sets. The RMSE of validation is lower than the one obtained by Hansen (1999). In the latter study, the validation data set was constituted by samples from New Zealand, whereas the calibration data set included with samples collected in Norway, Sweden, and Denmark. When classifying the data of the validation data set into 2 classes by using a threshold of 0.150 mmol/L, 89.3% of the samples were properly classified, with 93.0% of samples with low acetone content and 76.9% of samples with high acetone content well classified (Table 4-7). Therefore, similarly to the BHB model, the acetone model seems to be not appropriate to provide precise quantitative values, especially when acetone content is elevated, but it can allow discriminating high from low acetone values.

Table 4-7. Classification of the acetone predictions into 2 classes (% of samples; using a threshold of 0.150 mmol/L) for cross-validation and validation data sets

Item (%)	Low acetone content (<0.150 mmol/mL)	High acetone content (>0.150 mmol/mL)	Global good classification
Cross-Validation	n=134	n=32	
Predicted low	95.5%	15.6%	93.4%
Predicted high	4.5%	84.4%	
Validation	n=43	n=13	
Predicted low	93.0%	23.1%	89.3%
Predicted high	7.0%	76.9%	

Citrate Cross-Validation and Validation Results

The calibration data set for citrate contained 380 samples and the validation data set included 126 samples. In the calibration data set, citrate content ranged from 3.88 to 16.12 mmol/L, with an average of 9.03 mmol/L and a standard deviation of 2.26 mmol/L. The PLS model was done with 9 latent variables and 2 samples were considered as outliers. The slope between predicted and reference values is close to 1 (Figure 4-4) and the average of predicted values is comparable to the average of

reference values, with respectively 9.03 and 9.02 mmol/L. The RMSE of cross-validation obtained was 0.70 mmol/L, which is very low compared with mean or standard deviation, leading to an RPD of 3.21 (Table 4-5), indicating a fair estimation of citrate content and the use of the model for screening (Williams, 2004). The R^2 of cross-validation of the model was 0.90 (Figure 4-4.A). In the validation data set, the citrate content ranged from 4.44 to 15.16 mmol/L, with an average of 9.08 mmol/L and a standard deviation of 2.06 mmol/L (Table 4-5). The RMSE was 0.76 mmol/L, and the R^2 and the RPD were respectively 0.86 (Figure 4-4.B) and 2.96. Hence, these results were satisfactory, thereby allowing the use of citrate as a potential novel biomarker in milk potentially useful for management and breeding of dairy cows.

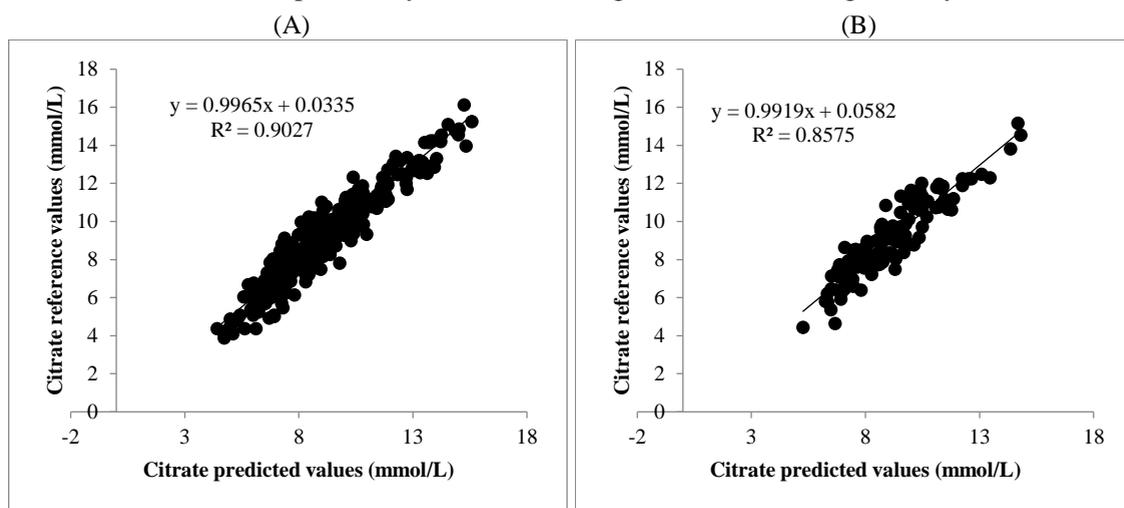


Figure 4-4. Plot of citrate reference values from flow injection analysis and citrate values predicted from Fourier transform mid-infrared analysis (A) for the cross-validation ($n = 380$) and (B) for the validation ($n = 126$).

Implications

Because standardized spectra were used in the development of the prediction equations for BHB, acetone, and citrate content in milk, these equations can be used on all standardized instruments from the OptiMIR network (Grelet et al., 2015), thereby allowing the utilization of these new biomarkers in the development of breeding and management tools for dairy cows. Hence, the equations developed in the frame of this project have been disseminated throughout the OptiMIR network. In laboratories, the spectra are extracted from the instruments, stored into external database, and standardized before application of the equations. With automation and IT development, these steps can be run daily, producing a quick feedback on field. Taking into account the precision of the different calibrations, the predictions could be used for herd management, or at individual level by using thresholds or relative values to cope with low accuracy. In 2015, advisory tools for the detection of cows potentially suffering from (sub-)clinical ketosis or from energy deficit have been deployed in Alsace region in France (Pezon, 2015) and in Luxembourg and are still

under development in Germany, United Kingdom, and the Walloon Region of Belgium (Baugnies, 2015). Further studies will investigate the opportunity of using these traits as indicators of health traits and fertility in breeding programs.

Conclusions

This work confirmed the usefulness of the FT-MIR spectrometry to predict milk biomarkers such as the content in milk of acetone, BHB, and citrate. Cross-validation statistics of the developed equations for acetone and BHB were similar or better than previous work and highlighted the opportunity to use these predictions for the detection of cows with high or low levels of ketone bodies in milk rather than for the determination of their exact content. Such results are expected given that the low concentration of BHB and acetone in milk implies only an indirect calibration of these components. Additionally, external validation statistics were provided in this work and confirmed the cross-validation results. Although further research is warranted to demonstrate the interest of mid-infrared predicted citrate as a useful biomarker for the dairy industry, this study showed that the development of FT-MIR calibration for citrate content in milk is promising. Finally, this work emphasized the usefulness of the standardization of mid-infrared spectra from different FT-MIR spectrophotometers to merge data sets and create more robust calibrations that can be used through a large network.

Acknowledgments

The authors gratefully acknowledge the laboratories and experimental farms providing data for this work: Comité du lait de Battice (Belgium), MPR-Kirchheim (Germany), ANALIS 53 (France), Trinottières experimental farm (France), Marcenat experimental farm (France), and Neumühle experimental farm (Germany). The National Fund for Scientific Research (FNRS, Brussels, Belgium) is acknowledged for grant agreement project 2.4604.11 (COMPOMILK project). The authors also thank the INTERREG IVB NWE program and the Ministry of Agriculture of Walloon Region of Belgium (Service Public de Wallonie, Direction générale opérationnelle “Agriculture, Ressources naturelles et Environnement”; DGARNE) for their financial support to the OptiMIR project (www.optimir.eu). Nicolas Gengler, as a former senior research associate, also acknowledges the support of the FNRS. The authors acknowledge especially Claire Darimont (CRA-W, Gembloux, Belgium) for her valuable work on chemical analysis.

References

- Baticz O., Tömösközi S., Vida L., and Gaál T., 2002. Relationship between concentration of citrate and ketone bodies in cow's milk. *Acta Vet. Hung.* 50:253–261.
- Baugnies S., 2015. Mise au point d'un indicateur de défaut énergétique des vaches à partir de nouvelles molécules disponibles via l'analyse spectrale du Contrôle Laitier dans le cadre du projet européen OptiMIR. Ms Thesis. Haute Ecole Condorcet, Prov. Hainaut, Belgium.

Bjerre-Harpøth V., Friggens N.C., Thorup V.M., Larsen T., Damgaard B.M., Ingvarsen K.L., and Moyes K.M., 2012. Metabolic and production profiles of dairy cows in response to decreased nutrient density to increase physiological imbalance at different stages of lactation. *J. Dairy Sci.* 95:2362–2380.

Bonfatti V., Di Martino G., and Carnier P., 2011. Effectiveness of mid-infrared spectroscopy for the prediction of detailed protein composition and contents of protein genetic variants of individual milk of Simmental cows. *J. Dairy Sci.* 94:5776–5785.

Butler W.R., 2003. Energy balance relationships with follicular development, ovulation and fertility in postpartum dairy cows. *Livest. Prod. Sci.* 83:211–218.

Collard B.L., Boettcher P.J., Dekkers J.C., Petitclerc D., and Schaeffer L.R., 2000. Relationships between energy balance and health traits of dairy cattle in early lactation. *J. Dairy Sci.* 83:2683–2690.

Dardenne P., Fernandez Pierna J.A., Abbas O., Baeten V., and Lecler B., 2015. Untargeted multivariate characterization of contaminants in powdered milk. *Proc. Final Optimir Scientific and Expert Meeting: From Milk Analysis to Advisory Tools. Namur (Belgium), 16–17 April 2015, BASE, vol. 19(2).*

Davies A., and Fearn T., 2006. Back to basics: Calibration statistics. *Spectroscopy Europe* 18:31–32.

De Marchi M., Toffanin V., Cassandro M., and Penasa M., 2014. Invited review: Mid-infrared spectroscopy as phenotyping tool for milk traits. *J. Dairy Sci.* 97:1171–1186.

de Roos A.P.W., van den Bijgaart H.J.C.M., Hørlyk J., and de Jong G., 2007. Screening for subclinical ketosis in dairy cattle by Fourier transform infrared spectrometry. *J. Dairy Sci.* 90:1761–1766.

Denis-Robichaud J., Dubuc J., Lefebvre D., and DesCôteaux L., 2014. Accuracy of milk ketone bodies from flow-injection analysis for the diagnosis of hyperketonemia in dairy cows. *J. Dairy Sci.* 97:3364–3370.

Duffield T., 2000. Subclinical ketosis in lactating dairy cattle. *Vet. Clin. North Am. Food Anim. Pract.* 16:231–253.

Enjalbert F., Nicot M. C., Bayourthe C., and Moncoulon R., 2001. Ketone bodies in milk and blood of dairy cows: Relationship between concentrations and utilization for detection of subclinical ketosis. *J. Dairy Sci.* 84:583–589.

Gengler N., Soyeurt H., Dehareng F., Bastin C., Colinet F.G., Hammami H., and Dardenne P., 2015. Overview of possibilities and challenges of the use of infrared spectrometry in cattle breeding, Page 353 in *Proc. EAAP 66th Annual Meeting, Warsaw. Wageningen Academic Publishers, Wageningen, the Netherlands.*

Gengler N., Soyeurt H., Dehareng F., Bastin C., Colinet F.G., Hammami H., Vanrobays M.L., Lainé A., Vanderick S., Grelet C., Vanlierde A., Froidmont E., and Dardenne P., 2016. Capitalizing on fine milk composition for breeding and management of dairy cows. *J. Dairy Sci.* 99:1–9.

Grelet C., Fernández Pierna J.A., Dardenne P., Baeten V., and Dehareng F., 2015. Standardization of milk mid-infrared spectra from a European dairy network. *J. Dairy Sci.* 98:2150–2160.

Hansen P.W., 1999. Screening of dairy cows for ketosis by use of infrared spectroscopy and multivariate calibration. *J. Dairy Sci.* 82:2005–2010.

Hanuš O., Genčurová V., Zhang Y., Hering P., Kopecký J., Jedelská R., Dolínková A., and Motyčka Z., 2011. Milk acetone determination by the photometrical method after microdiffusion and via FT infra-red spectroscopy. *J. Agrobiol.* 28:33–48.

Hanuš O., Roubal P., ěhal J., Klimešová M.Y., Samková E., Jedelská R., and Kopecký J., 2014. Development in indirect infra-red determination of milk acetone. *Acta Univ. Agric. Silvic. Mendel. Brun.* 62:919–927.

Heuer C., Luinge H.J., Lutz E.T., Schukken Y.H., van der Maas J.H., Wilmink H., and Noordhuizen J.P., 2001. Determination of acetone in cow milk by Fourier transform infrared spectroscopy for the detection of subclinical ketosis. *J. Dairy Sci.* 84:575–582.

Lainé A., Bel Mabrouk H., Dale L. M., Bastin C., and Gengler N., 2014. How to use mid-infrared spectral information from milk recording system to detect the pregnancy status of dairy cows. *Commun. Agric. Appl. Biol. Sci.* 79:33–38.

McParland S., Banos G., Wall E., Coffey M.P., Soyeurt H., Veerkamp R.F., and Berry D. P., 2011. The use of mid-infrared spectrometry to predict body energy status of Holstein cows. *J. Dairy Sci.* 94:3651–3661.

McParland S., Lewis E., Kennedy E., Moore S. G., McCarthy B., O'Donovan M., Butler S.T., Pryce J.E., and Berry D.P., 2014. Mid-infrared spectrometry of milk as a predictor of energy intake and efficiency in lactating dairy cows. *J. Dairy Sci.* 97:5863–5871.

Nicolai B.M., Beullensa K., Bobelyna E., Peirs A., Saeys W., Theronb K.I., and Lammertyn J., 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol. Technol.* 46:99–118.

Pezon J., 2015. La cétose subclinique détectée par le spectre MIR. *L'éleveur Laitier* 234:48–51.

Rousseuw P.J., Debruyne M., Engelen S., and Hubert M., 2006. Robustness and outlier detection in chemometrics. *Crit. Rev. Anal. Chem.* 36:221–242.

Rutten M.J.M., Bovenhuis H., Hettinga K.A., van Valenberg H.J.F., and van Arendonk J.A.M., 2009. Predicting bovine milk fat composition using infrared spectroscopy based on milk samples collected in winter and summer. *J. Dairy Sci.* 92:6202–6209.

Soyeurt H., Bastin C., Colinet F.G., Arnould V.M.-R., Berry D.P., Wall E., Dehareng F., Nguyen H.N., Dardenne P., Schefers J., Vandenplas J., Weigel K., Coffey M., Théron L., Detilleux J., Reding E., Gengler N., and McParland S., 2012. Mid-infrared prediction of lactoferrin content in bovine milk: potential indicator of mastitis. *Animal* 6:1830–1838.

Soyeurt H., Bruwier D., Romnee J.-M., Gengler N., Bertozzi C., Veselko D., and Dardenne P., 2009. Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 92:2444–2454.

Soyeurt H., Dehareng F., Gengler N., McParland S., Wall E., Berry D.P., Coffey M., and Dardenne P., 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *J. Dairy Sci.* 94:1657–1667.

Suthar V.S., Canelas-Raposo J., Deniz A., and Heuwieser W., 2013. Prevalence of subclinical ketosis and relationships with postpartum diseases in European dairy cows. *J. Dairy Sci.* 96:2925–2938.

van der Drift S.G., van Hulzen K.J.E., Teweldemedhn T.G., Jorritsma R., Nielen M., and Heuven H.C.M., 2012. Genetic and nongenetic variation in plasma and milk β -hydroxybutyrate and milk acetone concentrations of early-lactation dairy cows. *J. Dairy Sci.* 95:6781–6787.

van Knegsel A.T., van der Drift S.G.A., Horneman M., de Roos A.P.W., Kemp B., and Graat E.A.M., 2010. Short communication: Ketone body concentration in milk determined by Fourier transform infrared spectroscopy: Value for the detection of hyperketonemia in dairy cows. *J. Dairy Sci.* 93:3065–3069.

Vanlierde A., Vanrobays M.-L., Dehareng F., Froidmont E., Soyeurt H., McParland S., Lewis E., Deighton M.H., Grandl F., Kreuzer M., Gredler B., Dardenne P., and Gengler N., 2015. Hot topic: Innovative lactation-stage-dependent prediction of methane emissions from milk mid-infrared spectra. *J. Dairy Sci.* 98:5740–5747.

Williams P.C., 2004. Variables affecting near-infrared reflectance spectroscopic analysis. Pages 145–169 in *Near Infrared Technology in the Agriculture and Food Industries*. 2nd ed. P. Williams and K. Norris, ed. Am. Cereal Assoc. Cereal Chem., St. Paul, MN.

Williams P.C., and Sobering D.C., 1993. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *J. Near Infrared Spectrosc.* 1:25–32.

Chapter 5

General discussion

Chapter 5: General discussion

In a context where dairy farmers face difficulties to properly manage animals due to an increasing herd size, lack of relevant workforce, and growing administrative work, tools providing new indicators reflecting cow status represent a potential solution to help farmers in the daily management of farm. Such tools are also of interest in the objective of optimizing management and maintaining profitability in a period where price volatility is endangering farm viability. Additionally, a generation of new phenotypes could allow the dairy sector to provide objective measurements of animal health, welfare and environmental impact, to meet society's requirements. These new management indicators must be provided by a cost-effective technology available on a large scale. It is evident that analysis of milk through FT-MIR spectrometry is a potential solution to generate new information of interest for farmers. However, as mentioned earlier (cf. Introduction. Standardization), the use of this technology is currently suboptimal, and standardization of spectral data is needed to enable a global and efficient use of this resource. Consequently, the first objective of this work was to evaluate the possibility of applying a standardization method -widely recognised in the NIR sector- on MIR spectra of milk in order to create, transfer and use models of interest developed on different instruments and moments. Then, the study emphasises the creation of tools to help farmers manage negative energy balance and ketosis, which are two major issues threatening the sustainability of dairy farming.

Based on the results obtained, the study provided new knowledge on the possibility of using PDS method to standardize MIR spectra of milk as well as the feasibility of transferring different types of models, with high or low robustness, and predicting milk composition, processing qualities or even cow's status. The second part of this work validated the possibility of discriminating samples with high or low milk BHB and acetone and demonstrated the possibility of accurately predicting citrate in milk, offering dairy farms a good opportunity to detect early negative energy balance.

In this chapter, concrete outputs arising from the thesis will be described and results will be compared to those obtained in the literature, principally in the frame of studies using NIR spectrometers. Two essential elements regarding the standardization, reduction of spectral variability and interaction with quality of models, will be evaluated in more detail. Finally, the need for quality insurance systems will be raised.

Concrete and practical implementation of thesis outputs

This work has contributed to the development of a standardization methodology to be applicable in routine for the use of models developed in a research context. The first concrete output concerns the global methodology of standardization. This is concretely comprised of protocols defining the constitution of raw milk samples -to serve as the common basis for the standardization procedure-, the sending of samples among laboratories, the procedure to analyse standardization samples, the constitution of the master and application of PDS, the criteria for analysing the results, the distribution of standardization coefficients, and the procedure to use them to

routinely standardize milk spectra. The organization of these steps is shown in Figure 5-1.

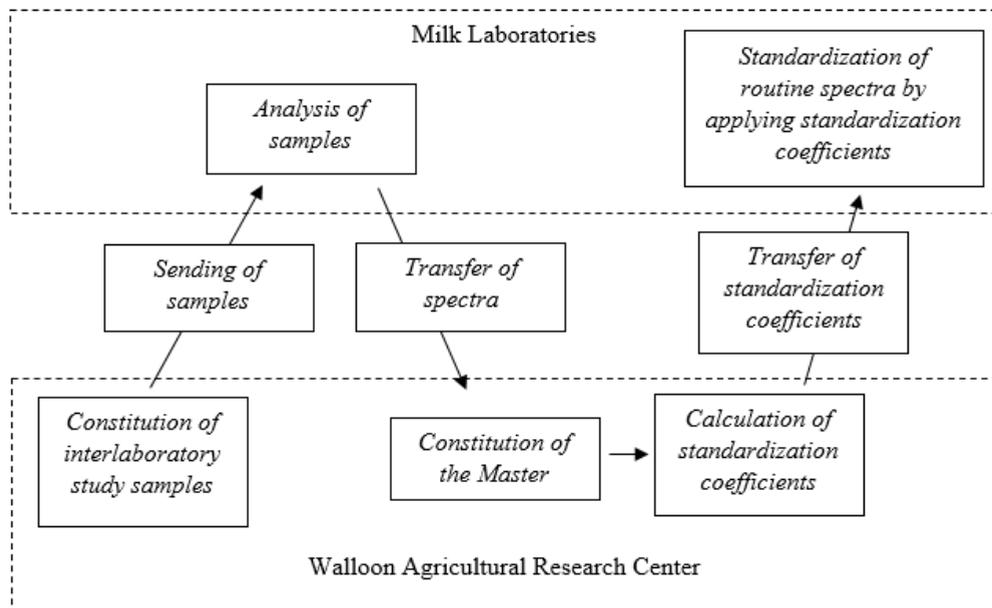


Figure 5-1. Scheme of the organization of monthly standardization

The second significant output of this thesis concerns the creation of a standardization network. Indeed, the first interlaboratory study was conducted in December 2011 in the framework of the Interreg OptiMIR project (www.optimir.eu). It included 26 spectrometers from 12 laboratories in 3 countries. Since then, the number of spectrometers and laboratories has been constantly increasing. The Interreg OptiMIR project was completed in 2015 but the standardization network remains operational. In 2018, approximately 103 spectrometers from 41 labs in 12 countries have been contributing monthly to standardization efforts. This network, consisting mainly of milk recording organizations and research institutions, represents a concrete exchange platform. Indeed, this platform constitutes an opportunity to share resources, exchange data, work jointly to create common models and use them within the constituted working group.

Thus, the models developed within the Interreg OptiMir project to predict milk BHB, acetone and citrate have been circulated among all the partners. Practically, after milk recording collection, individual milk samples are analysed on spectrometers and extracted FT-MIR spectra are standardized by the application of corresponding standardization coefficients. Spectra are usually stored in an external database, treated by automated homemade softwares developed within each institution and allowing to perform these steps daily. Standardized spectra are pre-treated with the same pre-treatment as described in Chapter 4 and coefficients of each equations are applied to generate predictions for each individual sample. Full automation of these systems allow providing feedback to farmers a few days after milk sampling. Based on BHB,

acetone and citrate predictions from MIR spectra of milk, tools helping farmers to detect negative energy balance or ketosis have been deployed by milk recording organizations since 2015 in France, Germany and Luxembourg are still under development in United Kingdom and Belgium.

Comparison with NIR standardization networks

In this study, the PDS method has been used in order to standardize mid-infrared spectrometers dedicated to milk analysis. Indeed, from the literature, this seems the more appropriate method to harmonize the spectral format of the instruments. However, this literature is based almost exclusively on NIR instruments used for the analysis of fuels or feed products (Table 5-1). In addition, from Wang *et al.* (1991), transfer performance can be impacted by the number and nature of the standardization samples. Currently, the samples used are real milk samples produced according to IDF norm 141. It is interesting to verify whether the use of PDS method to standardize FT-MIR instruments used for analysing milk and based on reconstituted milk samples, performs equally well than in the framework of NIR instruments.

The current study shows that the relative RMSE between master and slaves predictions, assessing the quality of model transfer, range from 0.4% to 4% for the fat and methane models, respectively (Chapter 3). In the literature, the first studies testing the PDS methods have focused on the analysis of gasoline and jet fuel using NIR instruments (Wang *et al.*, 1991; 1992). In these studies, relative RMSE ranged from 2.4 to 2.9 % when transferring models predicting different analytes of gasoline, and from 0.5 to 12.8% when transferring models predicting saturates and aromatic compounds in jet fuel, respectively. Other studies have focused on gasoline or diesel oil and reported relative RMSE from 0.3 to 5.9 % (Bouveresse & Massart, 1995) and from 0.6 to 19% (Lima & Borges, 2002). Relative RMSE in studies focusing on feed yielded higher values, namely, 10.5% in the case of a model predicting crude protein in silage (Liu *et al.*, 2011) and 7% in the case of a model predicting fibre content of flax stem (Sohn *et al.*, 2007). The main characteristics of these studies are summarized in table 5-1.

Table 5-1. Overview of various studies aiming to standardize spectrometers using PDS method

Study	Instrument	Matrice	Models	Relative RMSE (%)
Grelet <i>et al.</i> , 2017	FT-MIR	Milk	Fat, PUFA, FCY, CH ₄ ¹	0.4 - 4
Wang <i>et al.</i> , 1991	FT-NIR	Gasoline	Unknown analytes	2.4 - 2.9
Wang <i>et al.</i> , 1992	FT-NIR	Jet fuel	Aromatics, saturates	0.5 - 12.8
Bouveresse & Massart, 1995	FT-NIR	Gasoline	Unknown analytes	0.3 - 5.9
Liu <i>et al.</i> , 2011	FT-NIR	Silage	Crude protein	10.5
Sohn <i>et al.</i> , 2007	FT-NIR	Flax stem	Fiber content	7
Lima & borges, 2002	FT-NIR	Diesel oil	Density, viscosity, cetane number, cetane content, aromatics, distillation points	0.6 - 19

¹ CH₄ = methane emitted by dairy cows; PUFA = PUFA in milk; FCY = fresh cheese yield;

Compared against these studies, the results obtained in the framework of MIR analysis of milk show a similar range. In these studies, transfer performances are considered satisfactory for real world applications. Hence, this confirms the possibility of using PDS to standardize the MIR spectra of milk. This validates especially that the 5 milk interlaboratory samples are appropriate for this use. This also implies that the recombined milk samples contain enough spectral variability at each wavenumber to be used as standardization samples, which was expected due to the milk composition with large and orthogonal variations in fat and protein contents.

In the current study, the averaged RMSE between master and slaves predictions after standardization were relatively limited compared with the inherent standard cross-validation error of each model, meaning that transfers did not add a significant error to the final predictions (Chapter 3). Indeed, the averaged RMSE were 3.7, 4.7, 7.4 and 4.2 times smaller than the standard error of the models for fat, PUFA, FCY and methane, respectively. In previous studies, a transfer error slightly higher than the model-inherent error was observed. This ranged from 1.1 to 1.3 times for Wang et al. (1992), and from 1.2 to 6 times for Lima & Borges (2002) while a third study (Sohn et al., 2007) observed a slightly smaller transfer error, 1.1 time, than the model error. The results obtained in the present study can be considered satisfactory since the transfer procedure adds fewer error to the final predictions -in comparison with the inherent error of models- than previous studies employing PDS and NIR spectrometers. This could be partially explained by the use of low performance models in the framework of the current work. Indeed, as discussed in the Chapter 3, low accuracy models are also of interest for farm management or genetic studies as they provide useful information not previously available on a large scale.

Further evaluation regarding the reduction of spectral variability through standardization

Chapter 3 shows succinctly the impact of standardization on the reduction of spectral variability among instruments. This reduction has been quantified in this chapter by comparing the GH between the slave instruments and the centroid of the master spectra, based on the 5 samples of the interlaboratory study, and calculated with 10 principal components. This showed on average a reduction factor of 186 after standardization (GH values decreased from 2656 to 14). This was illustrated visually by performing a PCA plotting the spectra of the 5 interlaboratory study samples analysed by master and slaves instruments before and after standardization (Figure 3-2). This methodology allows to compare those samples to themselves, whether or not standardized. However, the small number of samples used only covers a limited milk spectral variability when compared to a real dataset obtained, for instance, during routine milk recording. This evaluation was consequently not representative of the effect of the standardization on the reduction of spectral variability under real conditions. In order to estimate this effect on real data, it is interesting to project the interlaboratory study samples, before and after standardization, into a real and highly variable dataset. This is done by projecting the data from Chapter 3, the interlaboratory study samples analyzed on the master and the slaves before and after standardization, on the fatty acid dataset obtained by Soyeurt et al. (2011). The original dataset has been updated with additional data for a total of 1,822 samples

from different instruments, breeds, diets and 7 countries. The test is performed with this dataset as it represents an important variability of real data coming from different systems and geographical regions. The spectra from all the slave instruments are interpolated on the master range. All the spectra are reduced to the 212 informative wave numbers and the PCA is done without pre-treatment of the spectra. As in Chapter 3, the first principal components, from 1 to 5, discriminate the interlaboratory studies samples and report information on milk composition. Figures reporting the projection of the data on those principal components, from PC 1 to PC 5, are shown in annex A. Figure 5-2 shows the projection of Chapter 3 standardized and raw interlaboratory data on PC 6 and PC 7 of the fatty acids dataset. On these principal components, the visual discrimination of interlaboratory study samples composition is limited and the effect of standardization on harmonization of spectral format can be observed. The interlaboratory study samples analysed by master instrument are represented by red points in the picture. The spectra of the interlaboratory study samples analysed by the different slave instruments without standardization are represented as green points. The spectra of the same samples after standardization are represented in blue. The fatty acids dataset samples, used only to evaluate the impact of standardization on the interlaboratory study samples under real conditions, are represented by grey points in the picture.

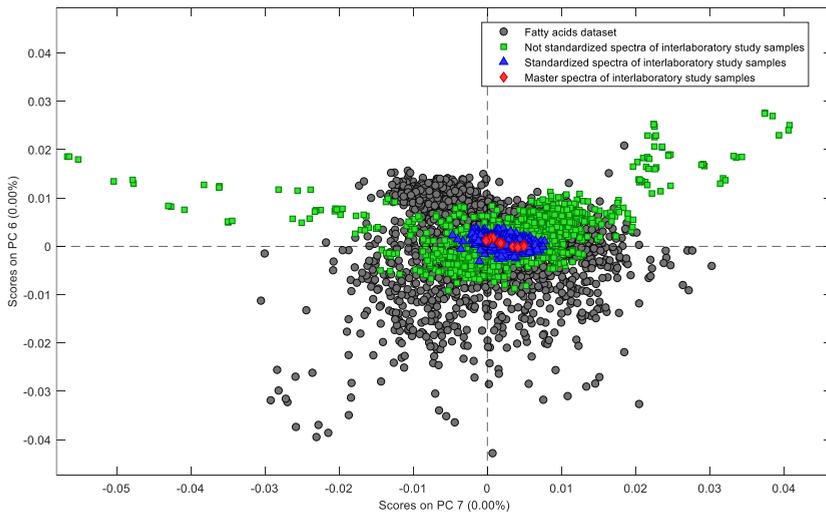


Figure 5-2. Principal component analysis of the spectra of the 5 interlaboratory study samples analyzed on 66 instruments, before and after standardization, and projected on the fatty acid dataset containing spectra from 1,822 individual milk samples. All the spectral data were plotted after selection of 212 informative wave numbers. Principal components (PC) 6 and 7 reported on the picture.

It can be observed that some non-standardized spectra show a large deviation from the master and even from the fatty acid dataset. This reflects an important spectral difference between the response of those instruments and the response of the master.

Following standardization, the spectral variability is relatively limited, and spectra are situated in the areas surrounding the master spectra. Compared to non-standardized spectra, the variability is considerably reduced, with no spectra deviating from either the master samples or fatty acid variability areas. Variability of non-standardized spectra is visually less important when projected on fatty acids dataset than when projected alone (Figure 3-2). This was expected given that only 5 identical samples were analysed by the instruments whereas the fatty acid dataset contain numerous and significantly different individual samples, which comparatively reduce the variability due to spectral differences between instruments as if this variability was visually zoomed out. This application under real conditions confirms the conclusions derived from Chapter 3: it is visually possible to observe that spectral homogeneity within the network increases around master spectral response following standardization.

Interaction between standardization and quality of models

As described in Chapter 3, the final precision when evaluating the transfer of models was not only due to the standardization process itself but was mainly impacted by the quality of the model being transferred. Indeed, Figure 3-5, indicates that a direct relationship seems to exist between the performance of the transfer (relative RMSE between slave and master) and the R^2_{cv} of the model. However, this observation has been made by comparing only 4 models. In order to confirm this hypothesis, it is necessary to study this relationship with a higher number of models. This has been done by calculating the relative RMSE between slaves and master predictions after standardization for models predicting fatty acids (Soyeurt et al., (2011), minerals (Soyeurt et al., 2009), methane (Vanlierde et al., 2016), casein (Not published), citrates (Grelet et al., 2016), FCY (Colinet et al., 2015) and lactoferrin (Soyeurt et al., 2007). RMSE were calculated for the same interlaboratory study samples analysed on 66 instruments than in Chapter 3. Figure 5-3 plots the relationship between the quality of transfer (relative RMSE) and the quality of models (R^2_{cv}). Relative RMSE range from 0.4 to 9.6%, which is in the same range than in the studies using PDS with NIR instruments described above (Table 5-1). From Figure 5-3, the conclusions are not as evident as in Chapter 3. There is a very weak tendency of relative RMSE to decrease when model R^2_{cv} increases. The relationship is characterized by an R^2 of 0.057 compared to 0.77 in Chapter 3. This suggests a potential impact of model quality on transfer performance, but particularly that the quality of model transfer should be influenced by additional parameters, such as robustness. The robustness of models is defined in Chapter 3 as the capacity of the models to function in all terrains and provide good results under various conditions. An important point influencing the robustness of models is the variability included in calibration datasets, both in terms of reference and spectral data. Indeed, to provide good results under real conditions, the calibration dataset must include a maximum variability of the different geographical regions, breeds, and diets encountered on the field. Robustness is also affected, among other factors, by the number of latent variables used in the PLS models, the precision of the reference method, the use of a repeatability file, the integration of several brands of FT-MIR spectrometers into the data set, and the reproducibility of wave number areas selected within the models.

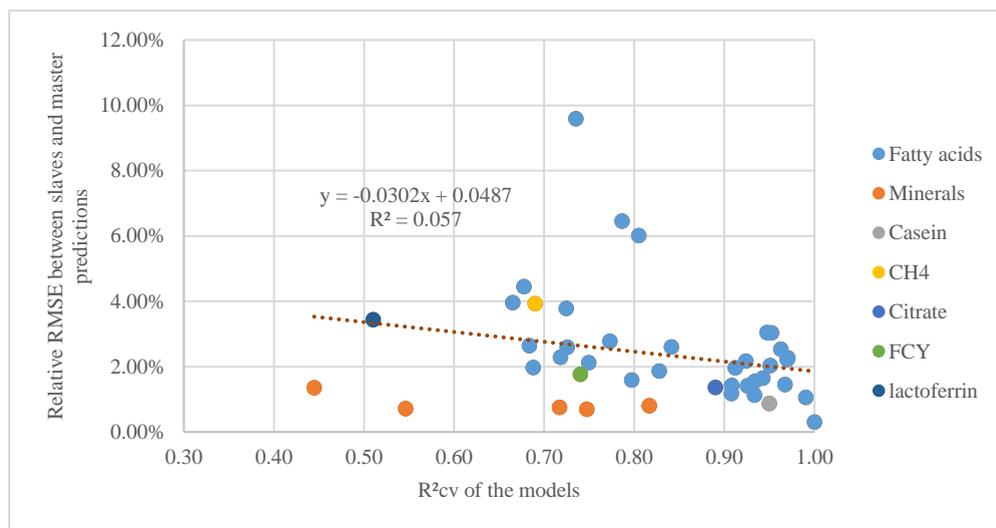


Figure 5-3. Plots representing the link between the quality of the model (coefficient of determination of cross-validation; R^2_{cv}) and the performance of the transfer by piecewise direct standardization [relative root mean squared error (RMSE) between slaves and master predictions after standardization] from the 5 interlaboratory study samples analyzed on 66 instruments. The different models originate from Soyeurt et al. (2011), Soyeurt et al. (2009), Vanlierde et al. (2016), Grelet et al. (2016), Colinet et al. (2015) and Soyeurt et al. (2007).

In the current example, based on Figure 5-3, the relationship among fatty acids or minerals models seems different, and differs in terms of slope and level of relative RMSE. In particular, the relative RMSE between slaves and master is low for mineral models, while the R^2_{cv} is not higher than for fatty acids. In connection with robustness, the coverage of interlaboratory study samples variability could explain this low relative RMSE. Indeed, the minerals calibration dataset contains skimmed milk samples that could play a role in the coverage of the extremely low fat interlaboratory study samples (around 2%) and explain the better predictions using this model. Additionally, the number of latent variables used within the PLS models has been mentioned as a parameter influencing the robustness of models. In Figure 5-4, the relative RMSE are plotted against the number of latent variables used in the PLS models, for the same models shown in Figure 5-3. According to Figure 5-4, a relationship seems to exist between those 2 parameters. Even if the R^2 between those 2 variables is low, it also argues in favour of an influence of model robustness on the quality of model transfer. As reported in previous chapters, the selection of 212 wavenumbers resulted from the analysis of the repeatability of the spectral response at the different wavenumbers when analysing identical milks on different instruments. However, the spectral response is not equally repeatable within the selected 212 wavenumbers as well. The relative importance of regression B coefficients along the spectra, in combination with the different repeatability of those spectral areas, are also likely to influence the quality of model transfer.

This additional study moderates the impact of model performances (R^2_{cv}) on model transfer performances. It suggests that other important model parameters could influence the results but also that part of the error could derive from the standardization process itself. Indeed, since this is simply a mathematical transfer, it cannot correct for “artefact” resulting from differences in the analysis, as good vs. insufficient homogenization of the samples, creating fundamental differences in the physico-chemical information contained within spectra from identical samples.

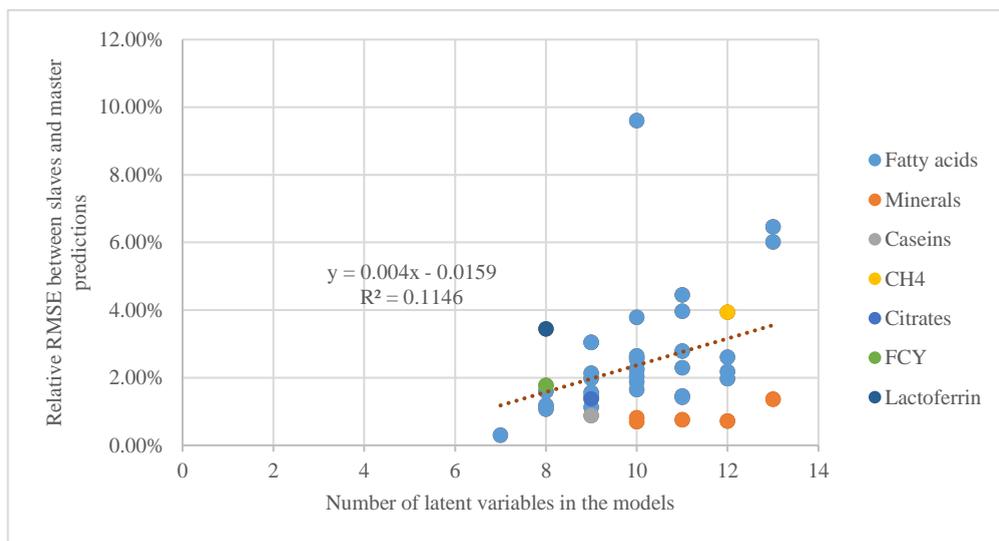


Figure 5-4. Plots representing the link between the number of latent variables within the PLS models and the performance of the transfer by piecewise direct standardization [relative root mean squared error (RMSE) between slaves and master predictions after standardization] from the 5 interlaboratory study samples analyzed on 66 instruments. The different models originate from Soyeurt et al. (2011), Soyeurt et al. (2009), Vanlierde et al. (2016), Grelet et al. (2016), Colinet et al. (2015) and Soyeurt et al. (2007).

Need for quality assurance systems to use models in routine

As described in the Introduction chapter, deterioration of pieces, changes of temperature and humidity, maintenance operations, piece replacements, electronic drifts and detector instability lead to the instability of each individual spectrometer response over time. The sending of reference samples is currently limited to a monthly frequency, due to technical and economic reasons. However, the spectrometers can potentially be affected by drifts or perturbations between two interlaboratory studies leading to biased final predictions until the estimation of new standardization coefficients. Consequently, there is a need to monitor the individual stability in time of the instruments at least, and ideally design a mechanism to stabilize the observed deviations in order to insure correct predictions. Complementary to the monthly standardization, a method has been developed to monitor the instrument stability daily. It is based on the analysis of an identical UHT milk, from a unique production

batch, for 3 months, minimally at a daily frequency. A fat prediction model is then applied to the raw spectra of this UHT milk acquired daily on a specific instrument, and raw predictions are plotted in time (Figure 5-5). Since the composition of the milk samples and the model applied on the spectra remain the same in time, the potential drifts observed for predictions originate consequently from the raw spectra themselves and reveal perturbations occurring on the spectrometer. A web tool has been developed by EMR (European Milk Recording; <https://www.milkrecording.eu>) to automatically produce this graph following the update of daily spectra, and allow laboratory managers to monitor the daily stability of their instruments. Figure 5-5 shows an example of perturbation detected on an instrument. The detection of a perturbation could allow to run a new analysis of interlaboratory study samples in order to generate new standardization coefficients that take this drift into account. However, this is not an ideal solution as several days could elapse between the detection of a drift and the generation of new coefficients considering the time needed for dispatching the necessary samples. Consequently, further research is required in order to stabilize these perturbations on a daily basis.

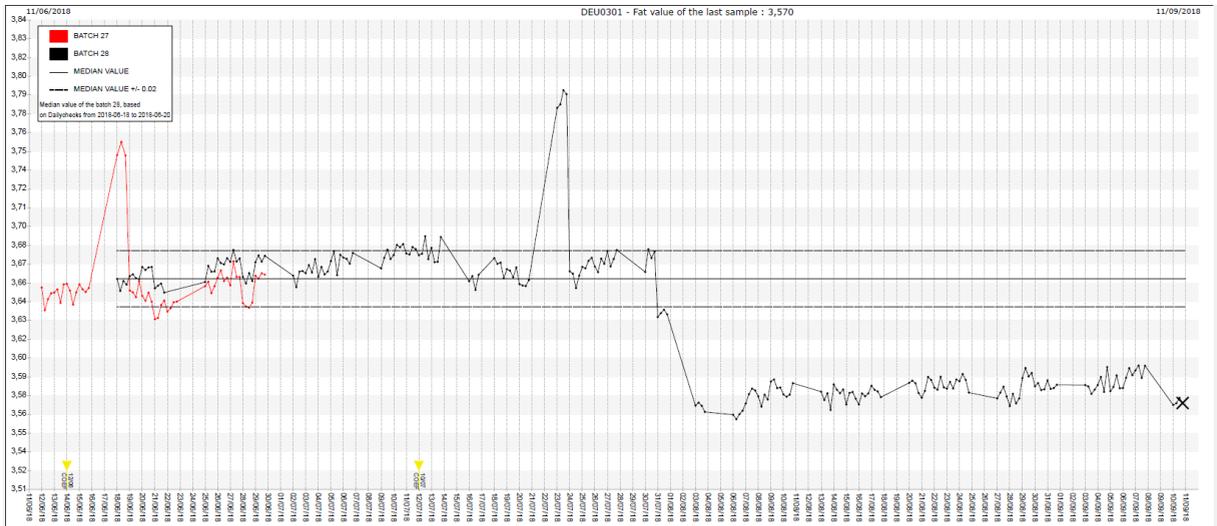


Figure 5-5. Plot of fat predictions over time originating from a fat model applied on non-standardized spectra from an instrument analysing a common batch of UHT milk during a period of 3 months.

An additional quality assurance issue is the representativeness of a calibration dataset regarding a new sample to predict. Indeed, the use of extrapolation in infrared predictions is dangerous, as mentioned by Dardenne (2010). Applying a model for a sample not covered by the calibration dataset is likely to lead to erroneous predictions. It is consequently important to develop models covering the existing variability of milk MIR spectra and studied trait(s) as completely as possible, by including samples from different geographical regions, breeds, diets, and by analysing milks on different types of instruments in order to maximize the probability of covering the variability of the samples predicted under routine conditions. Additionally, it becomes essential to validate the suitability of the model for every new sample to be analysed by

calculating the distance between the sample to be predicted and the calibration dataset used to build the predictive model.

References

Bouveresse E., Massart D.L., Dardenne P., 1995. Modified algorithm for standardization of near-infrared spectrometric instruments. *Anal. Chem.* 67:8,1381-1389.

Colinet F., Troch T., Baeten V., Dehareng F., Dardenne P., Sindic M., Gengler N., 2015. Genetic variability of MIR predicted milk technological properties in Walloon dairy cattle. Page 400 in *Book of Abstracts of the 66th Annual Meeting of the European Federation of Animal Science*. Wageningen Academic Publishers, Wageningen, the Netherlands.

Dardenne P., 2010. Some considerations about NIR spectroscopy, closing speech at NIR-2009. *NIR news*, 21:1,8-14.

Grelet C., Fernández Pierna J.A., Dardenne P., Soyeurt H., Vanlierde A., Colinet F., Gengler N., Baeten V., Dehareng F., 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, β -hydroxybutyrate and citrate contents in bovine milk through a European dairy network. *J. Dairy Sci.* 99:4816–4825.

ISO 9622:2013 | IDF 141:2013 – Whole milk – Determination of milk fat, protein and lactose content – Guidance on the operation of mid-infrared instruments. ISO, Geneva, Switzerland.

Lima F.S.G., Borges L.E.P., 2002. Evaluation of standardisation methods of near infrared calibration models. *J. Near Infrared Spec.* 10:4,269-278.

Liu X., Han L.J., Yang Z.L., 2011. Transfer of near infrared spectrometric models for silage crude protein detection between different instruments. *J. Dairy. Sci.* 94 :11, 5599-5610.

Sohn M., Barton F.E., Himmelsbach D.S., 2007. Transfer of near-infrared calibration model for determining fiber content in flax: effects of transfer samples and standardization procedure. *App. Spectrosc.* 61:4,414-418.

Soyeurt H., Colinet F., Arnould V., Dardenne P., Bertozzi C., Renaville R., Portetelle D., and Gengler N., 2007. Genetic variability of lactoferrin content estimated by mid-infrared spectrometry in bovine milk. *J. Dairy Sci.* 90:4443–4450.

Soyeurt H., Bruwier D., Romnee J.-M., Gengler N., Bertozzi C., Veselko D., Dardenne P., 2009. Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 92:2444–2454.

Soyeurt H., Dehareng F., Gengler N., McParland S., Wall E., Berry D. P., Coffey M., Dardenne P., 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *J. Dairy Sci.* 94:1657–1667.

Vanlierde A., Vanrobays M.L., Gengler N., Dardenne P., Froidmont E., Soyeurt H., McParland S., Lewis E., Deighton M. H., Mathot M., Dehareng F., 2016. Milk mid-infrared spectra enable prediction of lactation-stage dependent methane emissions of dairy cattle within routine population-scale milk recording schemes. *Anim. Prod. Sci.* 56:258–264.

Wang Y., Veltkamp D.J, Kowalsky B.R., 1991. Multivariate instrument standardisation, *Anal. Chem.*, 63:23,2750–27.

Wang Y., Lysaght M.J., Kowalski B.R., 1992. Improvement of multivariate calibration through instrument standardization. *Anal. Chem.*, 64:5,562-564.

Chapter 6

Perspectives and conclusion

Chapter 6: Perspectives and conclusion

Use of outputs in other researches

The current standardization method has been used in several projects to ensure the possibility of merging datasets and transferring developed models. Current or finished projects have defined standardization as a basis for the work with spectral data. Thus, studies focusing on ketosis (Gele et al., 2015; Smith et al., 2016), cellular immune traits (Denholm et al., 2016), pregnancy (Lainé et al., 2017), detection of lameness (Mineur et al., 2017), metabolic status of cows (Grelet et al., 2018; de Koster et al., 2018), nitrogen efficiency (Grelet et al., 2018) and technological properties of milk (Sanchez et al., 2018) have been conducted using standardized spectral data. Consequently, models derived from those works could potentially be applicable to the entire standardized network if the spectral variability is included in the calibration dataset. This method has also been used to implement new data to pre-existing datasets in order to increase the robustness and quality of models. In this manner, additional data have been added to the models predicting fatty acids (Soyeurt et al., 2006; 2011), minerals (Soyeurt et al., 2009), Lactoferrin (Soyeurt et al., 2007), and enteric methane emitted by dairy cows (Dehareng et al., 2012; Vanlierde et al., 2016). Finally, several recent projects integrate the standardization as a basis prior to the model development step, such as Smartcow (H2020; www.smartcow.eu), D4Dairy (COMET-Project; <https://d4dairy.com>), HappyMoo (INTERREG NWE; <http://www.nweurope.eu/projects/project-search/happymoo/>) or Indigess (MOERMAN).

Prospects regarding Master and network stability

Between Chapters 2 and 3, a fundamental change occurred in the standardization procedure. The Master evolved from being a single instrument to a combination of several instruments. This change took place following the observations of instability over time of each individual instrument. The consequence of an unstable reference in time would be the drift of the entire network spectral response. This would consequently induce a bias in the final predictions because of incompatibility between spectra used to develop pre-existing models and new spectra standardized on a slightly different basis. As explained in Chapter 3, a partial solution has been applied by considering the master not as single instrument but as combination of several instruments, to be less influenced by individual variations and more stable in time. However, there is a need for additional research to develop a procedure that ensures a perfect stability of the reference. The first potential solution would be to increase the number of instruments in the combination constituting the master in order to increase the smoothing of individual variation. Nonetheless, this is only conceivable as a short or mid-term solution as each instrument has a limited service life, involving the future replacement of all the instruments currently contributing to the reference. For example, between January and December 2018, 7 Foss FT6000 have been replaced by new Foss FT7. Additionally, it would be necessary to conduct a more profound study on the impact of individual variation on reference variation and on final predictions. Consequently, the ideal solution would be to free the reference from

existing physical instruments and consider a reference which would be perfectly stable in time. Direction for future research could focus on samples with the potential of being preserved or re-created equally over time.

Perspectives regarding timing of sampling

The outputs of this work and associated research are currently being used in the framework of classical milk recording. Individual cows are sampled once a month or every 6 weeks and milk samples analysed on regularly standardized FT-MIR instruments. However, following Berckmans (2006), an essential element for a PLF technology is the possibility of providing continuous (daily) monitoring. In the context of individual management, it is indeed necessary to quickly identify troubles and take the necessary steps to prevent the situation from becoming more critical. In contrast to other technologies, FT-MIR spectrometry is a very fast analysis, capable of scanning hundreds of samples per hour. Consequently, considering the high requirement of information for management and the high potential of this technology to analyse samples, the current use could be considered as suboptimal due to the limited and fixed sampling frequency associated with the classical milk recording.

In order to better reach management requirements, a potential perspective could be to adapt the sampling in order to focus deeper on some critical periods. For example, the risk of diseases and metabolic troubles are not constant over the lactation. Around 75% of diseases occur in the first month after calving in dairy herds (Suthar et al., 2013). Regarding these issues, it seems logical to focus on the early lactation stage. To this end, an adaptation of the milk recording procedure directed at placing more emphasis on the first months, with higher sampling frequency during this period, would be better adapted to early detection of diseases. Additionally, the risk level does not only depend on the stage of lactation but it is also related to the individual status, history and potential of each animal. Further research highlighting risk periods for each individual animal following those criteria would be beneficial in order to define a finest sampling and allow an optimal detection of troubles.

A second strategy to ensure the early detection of problems, is to increase the sampling frequency, ideally to reach a daily frequency. A daily estimation of parameters of interest would provide a very high throughput information, highly valuable as soon as proper analysis methods are used. Longitudinal analysis of data has shown to be an effective way for the detection of perturbations and analysis of resilience of animals (Poppe et al., 2018). However, daily sampling and analysis on classical FT-MIR instruments do not appear feasible for economic and technical reasons. Other technologies are better adapted to daily on-farm measurement, but first commercially available solutions currently provide information limited to one or a few traits, or supply predictions with low accuracy. The automated milking systems only allow conductivity measurements to be used as a mastitis indicator. Another system (Afilab, Afimilk®) consists of analysing milk on each individual milking unit with NIR spectrometers, which are more robust and more economically accessible than classical MIR instruments. It allows analysing for each animal fat, protein and lactose concentrations twice a day, but the results reported show low prediction accuracy as compared to classical MIR instruments (Kaniyamattam and De Vries., 2014). A research project has also focused on the online NIR analysis of milk in the

milking room and reported good accuracy but only for a limited number of fatty acids groups (Nguyen et al., 2011). Currently, manufacturers are also showing interest on the development of low-cost MIR spectrometers, which given their lower cost could be installed in dairy farms. However, these technologies are not yet commercially available. Based on these examples, the development of such tools is still ongoing, and it is highly probable that a higher frequency of sampling will be associated with a lower accuracy of the predictions. Additionally, such instruments are also susceptible to have individual specific instrumental response and to suffer from drifts over time, and standardization procedures should also be considered. Nevertheless, technologies allowing high frequency analysis provide a complementary aspect to the precise but sporadic use of current system and is undeniably opening prospects in the near future.

“Valorization” of raw tools

From the step of development of models predicting new biomarkers of interest, usually done by research institutions, the direct outputs are usually raw predictions of molecules in milk or the status of cows. Final users do not necessarily know how to interpret these values, so these raw results are not directly usable by farmers. Consequently, a crucial step to meet the final expectation, which is to develop tools to help farmers in their daily management work, is the valorization of raw phenotypes into usable indicators to be easily adopted by farmers, with user-friendly presentation. A classical way is to introduce the tools is the presentation of results by displaying traffic lights, e.g. a green light when cows are healthy and a red light when the status of cows show an imbalance. For example, Figure 6-1 shows how the Walloon Breeding Association (AWE, <http://www.awenet.be/>) refines raw predictions of acetone and BHB in milk to provide a final classification of each individual cows following its ketosis status. This tool combines the predictions of BHB and acetone in milk with the ratio fat/proteins and the predictions of the oleic acid (C18:1cis9) to provide a classification of the cows into 4 classes regarding the risk of ketosis: low risk, risk of negative energy balance, moderate risk and high risk.

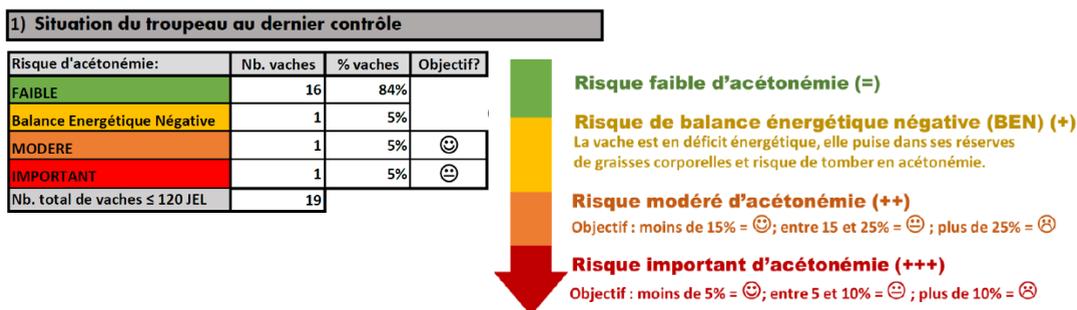


Figure 6-1. Example of valorization of raw predictions (milk BHB, acetone, oleic acid and fat/protein ratio) into usable information by the « Cétolait » tool developed by the Walloon Breeding Association (AWE).

The refining of raw tools is an essential step, although often underestimated, for efficient field implementation. It is also important to provide users with concrete actions to be taken in cases of detection of sub-optimal status. The assessment of the economic impact due to the use of such tools is also a strong argument for the adoption of indicators. This step of tools presentation is not only a marketing matter as the intrinsic precision of the raw tools also need to be considered. The final presentation should be technically adapted to the accuracy of the prediction to avoid overestimating the technical potential of the tool. Some models are not precise enough to provide direct quantitative predictions, for which the probability of being distant from the real value is fairly high, as for BHB and acetone. In such cases, providing a final tool on a quantitative and precise scale would be counterproductive and could mislead the farmer. Combination with additional predictions, using thresholds or relative scales or even averaging at the level of the herd, to diminish randomly distributed errors, are potential solutions to overcome this low accuracy.

Widening the use of IR predictions

In this work, emphasis has been placed on predictions of new phenotypes in order to help dairy farmers in the management of their cows. However, due to the high potential of this technology, it is possible to imagine applications for other uses. Numerous models of interest have been developed linked to applications other than cow management and could be spread through the standardization method.

Indeed, purchasing good quality milk products is clearly a concern for consumers. The potential of FT-MIR milk spectra to generate new predictions is logically of great interest for the dairy industry. The first models from milk spectra have been derived in order to predict fine milk composition. Such models can provide measurement of molecules linked to human health, such as poly-unsaturated fatty acids or calcium content in milk, in order to claim good quality or improve it. Other preoccupations of the dairy industry could potentially be approached through the FT-MIR technology such as authentication of milk regarding geographic origin or species, detection of milk adulteration, or ability of milk to be processed into different products.

The potential of MIR analysis of milk in order to generate new phenotypes, enhanced by the possibility of using models on different instruments through the standardization procedure, is also of great interest for genetic research. Indeed, nowadays the possibility of accessing more and more genetic information for thousands of cows does exist, and the bottleneck in terms of genetic research is the large scale access to phenotypes of interest regarding these dairy cows. Knowing that most of the milk recording and payment laboratories are equipped with FT-MIR spectrometers, there is a possibility to apply robust models on large scale, as soon as this is possible through the standardization of spectral responses, and to generate the needed phenotypes of interest to associate with genetic information.

Conclusion

The objective of this research was to optimize the use of FT-MIR analysis of milk in order to provide new management tools for dairy farmers. Indeed, this technology

has been used for decades to predict fat, protein and lactose, and more recently in the areas of fine milk composition, milk processing qualities and status of cows, but its concrete use by field organizations is still suboptimal due to the impossibility of sharing data and models among spectrometers. Consequently, the sub-objectives of this work were 1) to test a standardization method, well known from the NIR sector, in the framework of MIR spectra of milk, 2) to ensure practically the possibility of using models of interest such as low robustness equations predicting status of cows and 3) to develop models predicting NEB and ketosis indicators, based on standardized spectra from different regions, to be used concretely as management tools.

Chapter 2 has shown the possibility of standardizing a network of FT-MIR instruments dedicated to milk analysis by using the well-known PDS method based on common raw milk samples constituted from the IDF norm (ISO 9622:2013 | IDF 141:2013). Concretely, this was illustrated by the transfer of a robust fat model from a master instrument into 21 secondary instruments. Results show that the use of a unique model was possible on instruments of different brands and models. This work constituted a first step in the optimization of FT-MIR analysis by highlighting the possibility of sharing data and models among trans-national organizations.

Chapter 3 constituted a practical validation of the work done in Chapter 2 and provided assurance that models of interest with low robustness could be transferred from instrument to instrument. This was illustrated by the transfer on 66 instruments of models with limited robustness such as poly-unsaturated fatty acids in milk, methane emitted by dairy cows and fresh cheese yield of milk. This work also highlighted the fact that spectral standardization improves spectral and prediction reproducibility within the network. Concretely, it showed that models of interest for the dairy sector can be used on different instruments, independently of their brands and models. Thus, the method is the basis for data exchange, creation and use of robust models at an international level to generate new phenotypes to help farm management.

Chapter 4 finalized this work by bringing up a practical case of model development predicting new phenotypes arising from the FT-MIR spectra of milk. This work illustrated and used the gains from previous chapters as it was based on standardized spectra coming from 4 European regions and acquired on 10 spectrometers. This last study focused on the development of calibrations aiming to provide indicators of negative energy balance and ketosis, which are important issues in early lactation. Results confirmed the potential of this technology to predict phenotypes of interest such as acetone, BHB, and citrate of milk, which are biomarkers for these metabolic troubles.

This research yielded new knowledge about the possibility of using and sharing spectral data and models. But most importantly, it contributed concrete outputs to provide practically more information to farmers. A standardization procedure has been developed and is being utilized routinely by numerous European milk recording and research organizations. This standardization network has been constituted and represents an exchange platform to share resources and data, and work jointly to create common models and use them within the working group. Based on BHB, acetone and citrate predictions, tools helping farmers to detect negative energy balance or ketosis

have been deployed by milk recording organizations to provide feedback on NEB and ketosis. Finally, this work is not an end in itself but a basis for collaboration, research and concrete development of numerous tools for dairy farms and for the dairy sector in general.

References

Berckmans D., 2006. Automatic on-line monitoring of animals by precision livestock farming. *Livestock production and society*, 287.

Dehareng F., Delfosse C., Froidmont E., Soyeurt H., Martin C., Gengler N., Vanlierde A., Dardenne P., 2012. Potential use of milk mid-infrared spectra to predict individual methane emission of dairy cows. *Animal*. 6:1694–1701.

De Koster J., Salavati M., Grelet C., Crowe M., Opsomer G., Foldager L., the GplusE consortium, Hostens M., Metabolic clustering of cows, indicators of metabolic imbalance and association with production parameters. Accepted in *J. Dairy Sci.*

Denholm S.J., Smith S.L., McNeilly T.N., Hicks V., Wall E., 2016. Investigating the use of Mid Infrared spectral data to predict dairy cow cellular immune traits. Page 432 in *Book of Abstracts of the 67th Annual Meeting of the European Federation of Animal Science*. Wageningen Academic Publishers, Wageningen, the Netherlands. 2016.

Gelé M., Ferrand-Calmels M., Brun-Lafleur L., Werner A. Gollé-Leidreiter F., 2015, Predicting the risk of ketosis using mid infrared spectrometry, *ICAR Technical Series no. 19*, 19.

Grelet, C., Vanlierde, A., Hostens, M., Foldager, L., Salavati, M., Ingvarsten, K. L., Crowe M., Sorensen M.T., Froidmont E., Ferris C.P., Marchitelli C., Becker F., Larsen T., Carter F., GplusE Consortium, Dehareng F., 2019. Potential of milk mid-IR spectra to predict metabolic status of cows through blood components and an innovative clustering approach. *Animal*, 13:3, 649-658.

Grelet C., Froidmont E., Hostens M., Vanlierde A., Foldager L., Salavati M., Ingvarsten K.L., Sorensen M.T., Crowe M., Ferris C.P., Marchitelli C., Becker F., GplusE consortium, Dehareng F., 2018, A first approach to predict nitrogen efficiency of dairy cows through milk FT-MIR spectra, *Book of Abstracts of the 69th Annual Meeting of the European Federation of Animal Science*, Dubrovnik, p 477.

Kaniyamattam K., De Vries A., 2014. Agreement between milk fat, protein, and lactose observations collected from the Dairy Herd Improvement Association (DHIA) and a real-time milk analyzer. *J. Dairy Sci.* 97:5,2896-2908.

Lainé A., Bastin C., Grelet C., Hammami H., Colinet F. G., Dale L. M., Gillon A., Vandenplas J., Dehareng F., Gengler N., 2017. Assessing the effect of pregnancy stage on milk composition of dairy cows using mid-infrared spectra. *J. Dairy Sci.* 100:2863-2876.

Mineur A., Köck A., Grelet C., Gengler N., Egger-Danner C., Sölkner J., 2017. First results in the use of milk mid-infrared spectra in the detection of lameness in Austrian dairy cows. *Agriculturae Conspectus Scientificus*, 82 :2,163-166.

Nguyen H.N., Dehareng F., Hammida M., Baeten V., Froidmont E., Soyeurt H., Niemöller A., Dardenne P., 2011. Potential of near infrared spectroscopy for on-line analysis at the milking parlour using a fibre-optic probe presentation. *NIR news*, 22:7, 11-13.

Poppe H.W.M, Mulder H.A., Veerkamp R.F., 2018. Development of resilience indicators using deviation in milk yield from the lactation curve. Page 268 in *Book of Abstracts of the 69th Annual Meeting of the European Federation of Animal Science*. Wageningen Academic Publishers, Wageningen, the Netherlands.

Sanchez M.P., El Jabri M., Minéry S., Wolf V., Beuquier E., Laithier C., Delacroix-Buchet A., Brochard M., Boichard D., 2018. Genetic parameters for cheese-making properties and milk composition predicted from mid-infrared spectra in a large data set of Montbéliarde cows. *J. Dairy Sci.* 101:1-14.

Smith S., Hicks V., Cooper R., Forrest J., McRae A., Wall E., 2016. Developing a bovine ketosis risk indicator using spectral analysis and animal phenotype data. Page 301 in *Book of Abstracts of the 67th Annual Meeting of the European Federation of Animal Science*. Wageningen Academic Publishers, Wageningen, the Netherlands.

Soyeurt H., Dardenne P., Dehareng F., Lognay G., Veselko D., Marlier M., Gengler N., 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 89:3690– 3695.

Soyeurt H., Colinet F., Arnould V., Dardenne P., Bertozzi C., Renaville R., Portetelle D., Gengler N., 2007. Genetic variability of lactoferrin content estimated by mid-infrared spectrometry in bovine milk. *J. Dairy Sci.* 90:4443–4450.

Soyeurt H., Bruwier D., Romnee J.-M., Gengler N., Bertozzi C., Veselko D., Dardenne P., 2009. Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 92:2444–2454.

Soyeurt H., Dehareng F., Gengler N., McParland S., Wall E., Berry D. P., Coffey M., Dardenne P., 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *J. Dairy Sci.* 94:1657–1667.

Suthar V.S., Canelas-Raposo J., Deniz A., Heuwieser W., 2013. Prevalence of subclinical ketosis and relationships with postpartum diseases in European dairy cows. *J. Dairy Sci.* 96, 2925-2938.

Vanlierde A., Vanrobays M.L., Gengler N., Dardenne P., Froidmont E., Soyeurt H., McParland S., Lewis E., Deighton M. H., Mathot M., Dehareng F., 2016. Milk mid-infrared spectra enable prediction of lactation-stage dependent methane emissions of dairy cattle within routine population-scale milk recording schemes. *Anim. Prod. Sci.* 56:258–264.

Annex

Additional figures

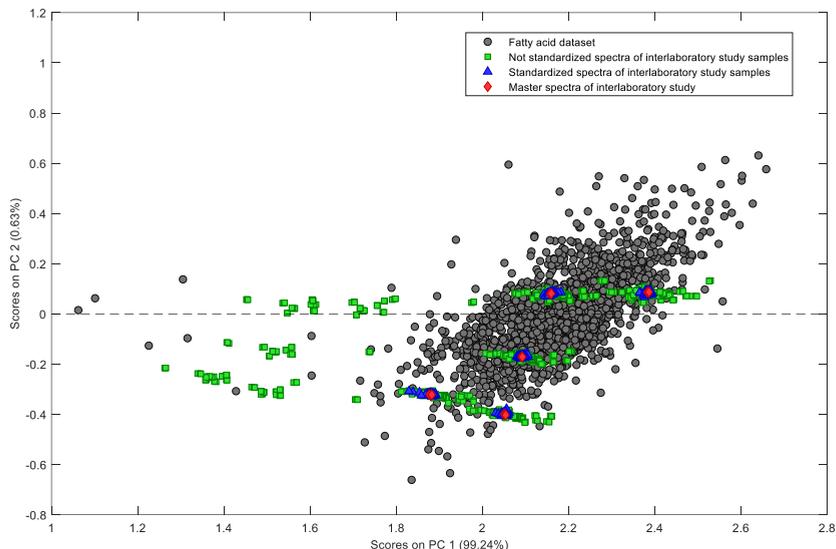


Figure A-1. Principal component analysis of the spectra of the 5 interlaboratory study samples analyzed on 66 instruments, before and after standardization, and projected on the fatty acids dataset containing spectra from 1,822 individual milk samples. All the spectral data were plotted after selection of 212 informative wave numbers. Principal components (PC) 1 and 2 reported on the picture.

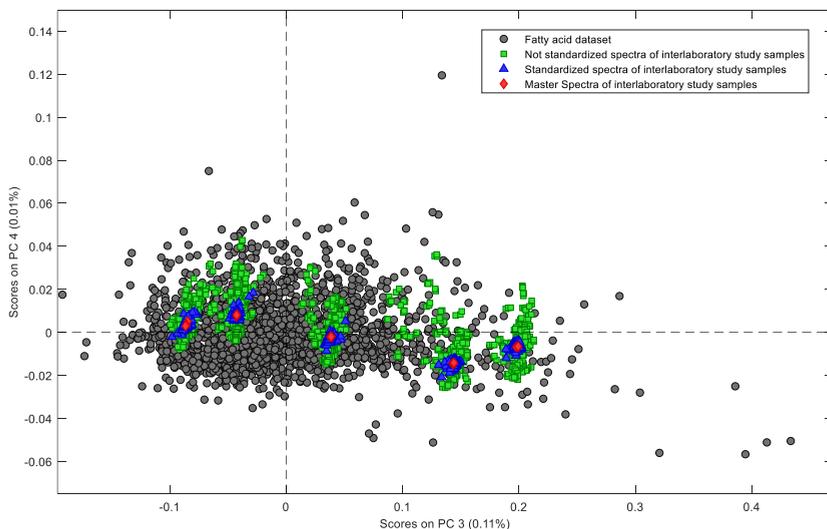


Figure A-2. Principal component analysis of the spectra of the 5 interlaboratory study samples analyzed on 66 instruments, before and after standardization, and projected on the fatty acids dataset containing spectra from 1,822 individual milk samples. All the spectral data were plotted after selection of 212 informative wave numbers. Principal components (PC) 3 and 4 reported on the picture.

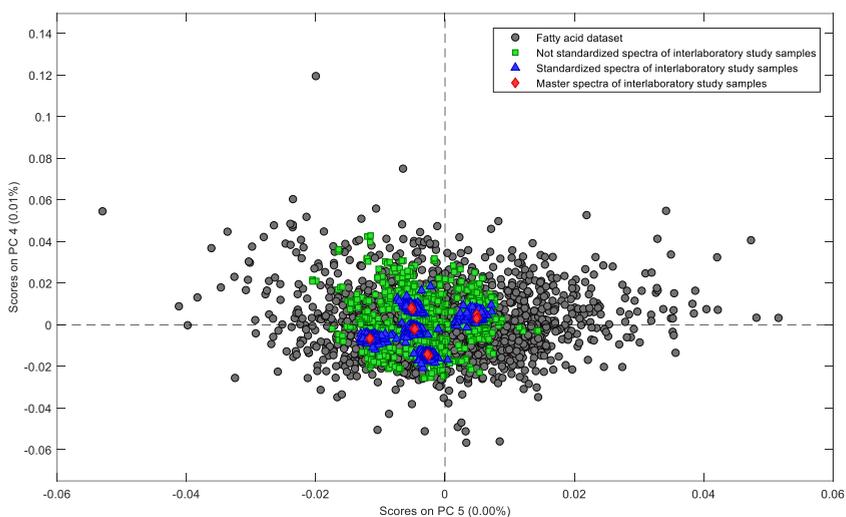


Figure A-3. Principal component analysis of the spectra of the 5 interlaboratory study samples analyzed on 66 instruments, before and after standardization, and projected on the fatty acids dataset containing spectra from 1,822 individual milk samples. All the spectral data were plotted after selection of 212 informative wave numbers. Principal components (PC) 4 and 5 reported on the picture.

Mathematical formula

Determination coefficient

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Where

y_i : value of the reference measure for i

\hat{y}_i : predicted value for i

\bar{y} : average of reference measures

n : number of observations

Root mean square error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

Where

y_i : value of the reference measure for i

\hat{y}_i : predicted value for i

n : number of observations

Relative root mean square error (relative RMSE)

$$Relative\ RMSE = \frac{RMSE}{\bar{y}}$$

Where

\bar{y} : average of reference measures (of the calibration dataset)

Overview of scientific publications

Peer-reviewed scientific publications

Mäntysaari, P., Mäntysaari, E., Kokkonen, T., Mehtiö, T., Kajava, S., Grelet, C., Lidauer, P., Lidauer, M., 2019, Body and milk traits as indicators of dairy cow energy status in early lactation. *J. Dairy Sci.*, accepted

De Koster, J., Salavati, M., Grelet, C., Crowe, M. A., Matthews, E., O'Flaherty, R., Opsomer, G., Foldager, L., GplusE, Hostens, M. 2019. Prediction of metabolic clusters in early-lactation dairy cows using models based on milk biomarkers. *J. Dairy Sci.*, 102:3, 2631-2644.

Grelet, C., Vanlierde, A., Hostens, M., Foldager, L., Salavati, M., Ingvarsten, K., Crowe, M., Sorensen, M., Froidmont, E., Ferris, C., Marchitelli, C., Becker, F., Larsen, T., Carter, F., GplusE Consortium, Dehareng, F. 2019. Potential of milk FT-MIR spectra to predict metabolic status of cows through blood components and an innovative clustering approach. *Animal*, 13:3, 649-658.

Little, M. W., Wylie, A. R. G., O'Connell, N. E., Welsh, M. D., Grelet, C., Bell, M. J., Gordon, A., Ferris, C. P. 2018. Immunological effects of altering the concentrate inclusion level in a grass silage-based diet for early lactation Holstein Friesian cows. *Animal*, 13:4, 799-809.

Leblois, J., Massart, S., Soyeurt, H., Grelet, C., Dehareng, F., Schroyen, M., Li, B., Wavreille, J., Bindelle, J., Everaert, N. 2018. Feeding sows resistant starch during gestation and lactation impacts their faecal microbiota and milk composition but shows limited effects on their progeny. *PLoS ONE*, 13:7,p.e0199568.

Grelet, C., Pierna, J. A. F., Dardenne, P., Soyeurt, H., Vanlierde, A., Colinet, F., Bastin, C., Gengler, N., Baeten, V., Dehareng, F. 2017. Standardization of milk mid-infrared spectrometers for the transfer and use of multiple models. *J. Dairy Sci.*, 100:10, 7910-7921.

Laine, A., Bastin, C., Grelet, C., Hammami, H., Colinet, F., Dale, L. M., Gillon, A., Vandenplas, J., Dehareng, F., Gengler, N. 2017. Assessing the effect of pregnancy stage on milk composition of dairy cows using mid-infrared spectra. *J. Dairy Sci.*, 100:4, 2863-2876.

Mineur, A., Köck, A., Grelet, C., Gengler, N., Egger-Danner, C., Sölkner, J. 2017. First results in the use of milk mid-infrared spectra in the detection of lameness in Austrian dairy cows. *Agriculturae Conspectus Scientificus*, 82:2, 163-166.

Grelet, C., Bastin, C., Gele, M., Davière, J.-B., Johan, M., Werner, A., Reding, R., Colinet, F., Dardenne, P., Gengler, N., Soyeurt, H., Dehareng, F. 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, β -hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. *J. Dairy Sci.*, 99:6, 4816-4825.

Gengler, N., Soyeurt, H., Dehareng, F., Bastin, C., Colinet, F., Hammami, H., Vanrobays, M.-L., Laine, A., Vanderick, S., Grelet, C., Vanlierde, A., Froidmont, E., Dardenne, P. 2016.

Capitalizing on fine milk composition for breeding and management of dairy cows. *J. Dairy Sci.*, 99:5, 4071-4079.

Fernandez Pierna, J. A., Vincke, D., Baeten, V., Grelet, C., Dehareng, F., Dardenne, P. 2016. Use of a multivariate moving window PCA for the untargeted detection of contaminants in agro-food products, as exemplified by the detection of melamine levels in milk using vibrational spectroscopy. *Chemometr. Intell. Lab.*, 152: 157-162.

Grelet, C., Fernandez Pierna, J. A., Dardenne, P., Baeten, V., Dehareng, F. 2015. Standardisation of milk mid-infrared spectra from a European dairy network. *J. Dairy Sci.*, 98, 2150-2160.

Scientific publications without peer-review

Dehareng, F., Grelet, C., Holroyd, S., Van den Bijgaart, H., Warnecke, S., Fernandez Pierna, J. A., Broutin, P., Dardenne, P. 2017. Standardization of FT-MIR instruments for milk analysis. *Bulletin: International Dairy Federation*, IDF 490-2017.

Contributions to international conferences

Grelet, C., Froimont, E., Hostens, M., Vanlierde, A., Foldager, L., Salavati, M., Ingvarsten, K., Sorensen, M., Crowe, M., Ferris, C., Marchitelli, C., Becker, F., GplusE consortium, Dehareng, F. 2018. A first approach to predict nitrogen efficiency of dairy cows through milk FT-MIR spectra. *Book of Abstracts of the 69th Annual Meeting of the European Federation of Animal Science*. Dubrovnik. (oral)

Mehtiö, T., Mäntysaari, P., Kokkonen, T., Kajava, S., Latomäki, T., Nyholm, L., Grelet, C., Pitkänen, T., Mäntysaari, E. A., Lidauer, M. 2018. Developing an indicator for body fat mobilisation using mid-infrared spectrometry of milk samples in dairy cows. *Proceedings of the World Congress on Genetics Applied to Livestock Production*. Auckland. (poster)

Mineur, A., Köck, A., Grelet, C., Egger-Danner, C., Sölkner, J., Gengler, N. 2018. Mid-infrared based biomarkers in milk as a non-invasive tool for early prediction of lameness caused by metabolic disorders. *The fifth and final dairycare conference*, Thessaloniki. (oral)

Foldager, L., Hostens, M., Salavati, M., Grelet, C., Sorensen, M. T., Crowe, M., Ingvarsten, K. L., GplusE Consortium. 2018. Random forests prediction of blood metabolic clusters of dairy cows: comparing three types of milk biomarkers. *Abstract for the Fifth DairyCare Conference*, Thessaloniki. (oral)

Marchitelli, C., Signorelli, F., Napolitano, F., Grelet, C., Gengler, N., Dehareng, F., Soyeurt, H., Ingvarsten, K. L., Sørensen, M. T., Crowe, M., GplusE Consortium. 2018. Milk biomarkers to evaluate health status of mammary gland in high producing dairy cattle. *Proceedings in 5th DairyCare conference*, Thessaloniki. (oral)

Grelet, C., Baeten, V., Dardenne, P., Fernandez Pierna, J. A., Abbas, O., Dehareng, F. 2018. Presentation of a new system to monitor and stabilize mid infrared spectral data. ICAR annual Conference, Auckland. (oral)

Grelet, C., Vanlierde, A., Dehareng, F., Froidmont, E., GplusE Consortium. 2017. Prediction of energy status of dairy cows using MIR milk spectra. Book of Abstracts of the 68th Annual Meeting of the European Federation of Animal Science. Tallin. (oral)

Hammami, H., Colinet, F., Bastin, C., Grelet, C., Vanlierde, A., Dehareng, F., Gengler, N., GplusE Consortium. 2017. Using milk based biomarkers to monitor the physiological state of dairy cows in large populations. Book of Abstracts of the 68th Annual Meeting of the European Federation of Animal Science. Tallin. (oral)

Hammami, H., Colinet, F., Bastin, C., Grelet, C., Vanlierde, A., Dehareng, F., Gengler, N. 2017. Genetic analysis of milk MIR predicted blood and milk biomarkers linked to the physiological status. Book of Abstracts of the 68th Annual Meeting of the European Federation of Animal Science. pp. 403. Tallin. (oral)

Mineur, A., Köck, A., Grelet, C., Gengler, N., Egger-Danner, C., Sölkner, J. 2017. First results in the use of milk MIR spectra in the detection of lameness in Austrian dairy cows. Projektgruppe Genetisch-Statistische Methoden“ der DGfZ, Cuxhaven, Germany. (oral)

Mäntysaari, P., Kokkonen, T., Grelet, C., Mäntysaari, E., Lidauer, M. 2017. Body and milk traits as cow's energy status indicator. Book of Abstracts of the 68th Annual Meeting of the European Federation of Animal Science, Tallin. (poster)

Marchitelli, C., Signorelli, F., Napolitano, F., Buttazzoni, L., Grelet, C., Vanlierde, A., Dehareng, F., Soyeurt, H., Crowe, M., GplusE Consortium. 2017. Expression profiles of immune genes in milk somatic cells and MIR predicted mineral contents in milk as indicators of mastitis. 36th International Society for Animal Genetics Conference, Dublin. (poster)

Vanlierde, A., Grelet, C., Gengler, N., Ferris, C., Sorensen, M. T., Höglund, J., Carter, F., Santoro, A., Hermans, K., Hostens, M., Dardenne, P., Dehareng, F. 2016. Potential of milk MIR spectra to develop new health phenotypes for dairy cows in the GplusE project. Book of Abstracts of the 67th Annual Meeting of the European Federation of Animal Science. pp. 299. Belfast. (oral)

Grandl, F., Vanlierde, A., Vanrobays, M.-L., Grelet, C., Dehareng, F., Gengler, N., Soyeurt, H., Kreuzer, M., Schwarm, A., Münger, A., Dohme-Meier, F., Gredler, B. 2016. Application of a milk MIR methane prediction equation to Swiss dairy cattle population data. Book of Abstracts of the 67th Annual Meeting of the European Federation of Animal Science. pp. 630. Belfast. (oral)

Grelet, C., Gengler, N., Bastin, C., Vanlierde, A., Smith, S., Gelé, M., Soyeurt, H., Massart, X., Dardenne, P., Dehareng, F. 2016. OptiMIR: Use of MIR spectra to predict multiple cow status as advisory tools for dairy farms. Book of Abstracts of the 67th Annual Meeting of the European Federation of Animal Science. pp. 300. Belfast. (oral)

Soyeurt, H., Vanlierde, A., Grelet, C., Vanrobays, M.-L., Colinet, F., Gengler, N., Bastin, C., Dehareng, F. 2016. Improvement of robustness of forward selection of variables using split datasets: Mid-infrared methane equation. Proceedings in: Chimiométrie 2016. Namur. (oral)

Grelet, C., Fernandez Pierna, J. A., Dardenne, P., Baeten, V., Massart, X., Dehareng, F. 2016. Application of a standardisation procedure on an international network of MIR instruments for milk analysis. Proceedings in: Chimiométrie 2016. Namur. (poster)

Fernandez Pierna, J., Vincke, D., Baeten, V., Grelet, C., Dehareng, F., Dardenne, P. 2016. Chemometrics and vibrational spectroscopy for the detection of melamine levels in milk. Proceedings in: Chimiométrie 2016. Namur. (poster)

Grelet, C., Fernandez Pierna, J. A., Dardenne, P., Soyeurt, H., Colinet, F., Vanlierde, A., Gengler, N., Dehareng, F. 2016. Novel innovative possibilities for the dairy industry opened by common format of FT-MIR instruments. IDF world dairy summit 2016, Rotterdam. (poster)

Vanlierde, A., Dehareng, F., Froidmont, E., Gengler, N., Soyeurt, H., Grelet, C., McParland, S., Dardenne, P. 2016. A simple method to predict methane emissions based on milk mid infrared spectra. IDF World Dairy Summit 2016, Rotterdam. (poster)

Grelet, C., Bastin, C., Gelé, M., Davière, J.-B., Reding, R., Werner, A., Darimont, C., Dehareng, F., Gengler, N., Dardenne, P. 2015. Milk biomarkers to detect ketosis and negative energy balance using MIR spectrometry. Book of Abstracts of the 66th Annual Meeting of the European Federation of Animal Science. pp 354. Warsaw. (oral)

Bastin, C., Colinet, F., Dehareng, F., Grelet, C., Hammami, H., Soyeurt, H., Vanlierde, A., Vanrobays, M.-L., Gengler, N. 2015. On the use of novel milk phenotypes as predictors of difficult-to-record traits in breeding programs. Book of Abstracts of the 66th Annual Meeting of the European Federation of Animal Science. pp 550. Warsaw. (oral)

Grelet, C., Fernandez Pierna, J. A., Dardenne, P., Vanlierde, A., Baeten, V., Soyeurt, H., Darimont, C., Dehareng, F. 2015. Standardisation of milk MIR spectra, Development of common MIR equations. Final optimir scientific and expert meeting: from milk analysis to advisory tools. Namur. (oral)

Bastin, C., Grelet, C., Gelé, M., Davière, J.-B., Reding, R., Darimont, C., Dehareng, F., Gengler, N., Dardenne, P. 2015. Mid-infrared prediction of beta-hydroxybutyrate, acetone, and citrate contents in milk. Final optimir scientific and expert meeting: from milk analysis to advisory tools. Namur. (poster)

Grelet, C., Fernandez Pierna, J., Dardenne, P., Baeten, V., Vanlierde, A., Soyeurt, H., Darimont, C., Dehareng, F. 2015. Quality Assurance for new analytical parameters, Optimir standardisation of MIR instruments. Reference Laboratory Network meeting, ICAR technical workshop. Krakow. (oral)

Laine, A., Bastin, C., Grelet, C., Dehareng, F., Massart, X., Gengler, N. 2015. Developing Innovative Health and Welfare Management Tools for Dairy Cows from Optimized Use of Milk Mid-Infrared Spectra (OptiMIR). Second DairyCare Conference. Cordoba. (poster)

Grelet, C., Fernandez Pierna, J., Soyeurt, H., Dehareng, F., Gengler, N., Dardenne, P. 2014. Creation of universal MIR calibration by standardization of milk spectra: example of fatty acids. Book of abstracts of the 65th annual meeting of the European Federation of Animal Science. pp108. Copenhagen. (poster)

Grelet, C., Fernandez Pierna, J. A., Dardenne, P., Dehareng, F. 2014. Standardisation de spectres MIR de lait. Journées 3R 2014. Paris. (poster)

Ninane, V., Grelet, C., Dehareng, F., Romnée, J.-M., Boxus, K. 2014. Effet de la microfiltration d'épuration sur les composants influençant la qualité fromagère du lait. Journées 3R 2014. Paris. (poster)

Grelet, C., Fernandez Pierna, J. A., Dehareng, F., Dardenne, P. 2013. Standardisation of milk MIR spectra: a first step to create new tools of dairy farm management. Book of Abstracts of the 64th Annual Meeting of the European Federation of Animal Science. Nantes. (oral)

Grelet, C., Dehareng, F., Vanlierde, A., Nguyen, H. N., Colinet, F., Dardenne, P. 2013. New tools for the dairy sector based on MIR and NIR spectroscopy. 6th International Symposium on Recent Advances in Food Analysis, Prague. (oral)

Grelet, C., Froidmont, E., Rondia, P., Sinnaeve, G., Dehareng, F. 2013. Influence de l'alimentation des vaches laitières sur la fromageabilité du lait. Journées 3R 2013. Paris. (poster)

Fernandez Pierna, J., Grelet, C., Dehareng, F., Dardenne, P., Baeten, V. 2012. Merging of spectral datasets from different MIR instruments used in the routine analysis of milk. Proceedings of ICAR 2012. Cork. (oral)

Grelet, C., Fernandez Pierna, J., Dardenne, P., Massart, X., Brunscwhig, P., Dehareng, F. 2012. OptiMIR : Standardisation de spectres moyen infrarouge de lait de vache en vue de la création d'une base de données transnationale. Journées 3R 2012. Paris. (poster)

Contributions to national conferences

Mineur, A., Köck, A., Grelet, C., Gengler, N., Egger-Danner, C., Sölkner, J. 2017. First results in the use of milk MIR spectra in the detection of lameness in Austrian dairy cows. 25th International Symposium Animal Science Days, Brandlücken, Austria. (oral)

Grelet, C., Fernandez Pierna, J. A., Dehareng, F., Massart, X., Gengler, N., Dardenne, P. 2013. OptiMIR : de nouveaux outils de management des exploitations laitières grâce au spectre moyen infra-rouge du lait. Carrefour des productions Animales 2013. Gembloux. (oral)