





COMMUNAUTE FRANCAISE DE BELGIQUE  
UNIVERSITE DE LIEGE – GEMBLoux AGRO-BIO TECH

**Spectrométrie moyen infrarouge dans le lait :  
contribution à l'étude des facteurs principaux visant à  
accroître la qualité des modèles de prédiction et à  
assurer leur contrôle**

Frédéric DEHARENG

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

Promoteurs : Professeurs Hélène Soyeurt et Nicolas Gengler

Année civile : 2023

Copyright

© DEHARENG Frédéric, Mars 2023

# Résumé

*Frédéric Dehareng, 2023. Spectrométrie moyen infrarouge dans le lait : contribution à l'étude des facteurs principaux visant à accroître la qualité des modèles de prédiction et à assurer leur contrôle. Thèse de doctorat. Gembloux, Belgique, Gembloux Agro-Bio Tech, Université de Liège, 177 p.*

L'utilisation de la spectrométrie moyen infrarouge à transformée de Fourier (FT-MIR) dans le cadre des analyses du lait a permis récemment de développer des modèles de prédiction pour une large gamme de nouveaux phénotypes. Pour garantir leur robustesse, il est important de couvrir au maximum la variabilité phénotypique. De plus, la qualité des prédictions doit être assurée via la maîtrise de l'ensemble des facteurs pouvant l'influencer. L'objectif de cette thèse est de faire la synthèse de recommandations et pratiques permettant d'assurer la robustesse tout en maintenant la qualité des prédictions FT-MIR des échantillons de lait analysés en routine. Premièrement, la sélection des échantillons est importante. Deux stratégies d'échantillonnage ont été investiguées afin d'accroître la variabilité du set d'échantillons de calibrage. La première reposait sur une méthode originale de tri des échantillons de référence à analyser sur base de leur variabilité spectrale (X). Une deuxième visait à créer de la variabilité dans le set de calibration directement à partir du phénotype (y). Ensuite, le choix de l'analyse de référence est un facteur important qui ne doit pas être exclusivement basé sur la précision. En effet, d'autres critères devront aiguiller ce choix afin de composer avec les contraintes techniques, pratiques, financières et éthiques. L'utilisation d'un prétraitement ou le choix de l'algorithme utilisé pour construire un modèle influencera également ses performances. D'autre part, la plupart des nouveaux modèles utilisent une plus grande quantité de nombres d'ondes, voire l'utilisation de corrélations indirectes. Le contrôle des interférences et des déviations spectrales des appareils sera donc un point important à maîtriser, car toute déviation pourra entraîner des conséquences fâcheuses. Ainsi, l'utilisation de méthodes de standardisation spectrale adéquates semble une solution prometteuse. Les procédures actuellement utilisées en routine pour valider les méthodes, pour assurer leur adaptabilité à une zone géographique donnée, ou encore pour contrôler la qualité au quotidien des appareils de routine, devront subir des adaptations ou de nouvelles procédures devront être créées. En conclusion, ce travail de thèse, reflétant 20 ans d'expérience, a pu mettre en avant de nombreuses solutions et développements futurs pour permettre d'améliorer la robustesse des équations et maintenir la qualité des prédictions en routine. Les résultats obtenus soulignent le besoin de créer des collaborations internationales tant pour le développement des modèles prédictifs que pour éditer des recommandations afin d'utiliser pleinement les potentialités de la spectrométrie FT-MIR pour l'analyse du lait.

# Abstract

*Frédéric Dehareng, 2023. Milk mid-infrared spectrometry: contribution study of the main factors improving the quality of prediction models and to ensure their control. PhD Dissertation. Gembloux, Belgium, Gembloux Agro-Bio Tech, University of Liège, 177 p.*

The use of Fourier Transform Medium Infrared Spectrometry (FT-MIR) in milk has recently allowed the development of prediction models for a wide range of new phenotypes. To ensure their robustness, it is important to cover as much phenotypic variability as possible. Moreover, the quality of predictions must be ensured by controlling all the factors that can influence it. The objective of this thesis is to synthesize recommendations and practices to ensure the robustness while maintaining the quality of FT-MIR predictions of milk samples analyzed in routine. First, sample selection is important. Two sampling strategies were investigated to increase the variability of the calibration sample set. The first was based on an original method of sorting the reference samples to be analyzed on the basis of their spectral variability (X). A second one aimed at creating variability in the calibration set directly from the phenotype (y). Then, the choice of the reference analysis is an important factor that should not be exclusively based on precision. Indeed, other criteria will have to guide this choice in order to deal with technical, practical, financial and ethical constraints. The use of pre-processing or the choice of algorithm used to build a model will also influence its performance. On the other hand, most of the new models use a larger number of wave numbers, or even the use of indirect correlations. The control of interferences and spectral deviations of the devices will be an important point to master, because any deviation could lead to unfortunate consequences. Thus, the use of adequate spectral standardization methods seems a promising solution. The procedures currently used in routine to validate methods, to ensure their adaptability to a given geographical area, or to control the quality of routine equipment, will have to be adapted or new procedures will have to be created. In conclusion, this thesis work, reflecting 20 years of experience, was able to put forward many solutions and future developments to improve the robustness of the equations and maintain the quality of predictions in routine. The results obtained underline the need to create international collaborations both for the development of predictive models and for issuing recommendations in order to fully use the potentialities of FT-MIR spectrometry for milk analysis.

# Remerciements

Si ce travail peut enfin toucher à sa fin, c'est grâce à l'aide et au support de nombreuses personnes. Sans elles, je n'y serais probablement jamais arrivé et je tiens à les en remercier.

Tout d'abord, à mes promoteurs de thèse, Hélène Soyeurt et Nicolas Gengler. Depuis près de vingt ans, j'ai eu la chance de collaborer et travailler avec eux. Merci à Hélène de m'avoir encouragé, soutenu et motivé pour la réalisation de ce travail. Son efficacité, sa sagacité, son pragmatisme et sa rapidité m'impressionneront toujours. Son aide constructive et ses encouragements dans le sprint final, m'ont été très utiles. Merci à Nicolas pour sa confiance répétée dans nos nombreuses collaborations et l'immense savoir qu'il partage sans compter. Chaque conversation avec lui, durant nos si nombreuses missions, est toujours riche d'informations et d'enseignements. Qu'ils reçoivent tous les deux toute mon amitié et gratitude.

Aux membres de ce comité de thèse pour leur soutien et conseils avisés. A Pierre Dardenne, notre mentor dans le domaine de la spectrométrie, qui m'a toujours supporté, guidé et défendu dans mes projets. Ses conseils judicieux, ses doutes qui nous poussent à toujours nous remettre en question et à approfondir notre réflexion, son pragmatisme, le tout enrobé d'une bonne dose d'humanisme me manquent depuis son départ à la retraite. A Georges Sinnaeve, qui, malgré ses responsabilités de Directeur du Centre wallon de Recherches agronomiques, a bien voulu me consacrer une partie de son temps précieux pour me faire part de ses conseils avisés. Sa confiance m'a toujours permis de travailler en autonomie. A Vincent Baeten, pour ses conseils clairvoyants dans le domaine de la spectroscopie et ses idées judicieuses et innovantes. Vincent est un meneur hors pair qui nous pousse à aller plus loin et plus haut, un collègue et un ami précieux. A Marianne Sindic d'avoir accepté d'être membre de ce jury et de partager son expérience. A Georges Lognay pour son altruisme, sa relecture scrupuleuse et son goût de la précision. A Yves Beckers pour avoir accepté de présider ce jury et avoir accompli sa tâche avec bienveillance.

A Gilbert Berben, pour m'avoir permis de faire ce doctorat et son aide dans le suivi de mon stage.

Un merci particulier à Clément Grelet et Amélie Vanlierde. Qu'il est bon et agréable d'être aussi bien accompagné dans ses recherches et son travail au quotidien, et durant nos nombreuses missions. Mille mercis pour les échanges, travaux, conseils, relectures et votre gentillesse. Puissent vos routes être encore longues et remplies de richesses et de découvertes.

A l'ensemble de mon équipe sans qui rien ne serait possible. A Olivier, Petimat, Mohammed, Matthieu, Lora, Octave, Huzaiifa et toutes celles et ceux qui sont passés par le labo lait. Vous effectuez un travail merveilleux au quotidien et vos idées et retours d'expérience sont toujours riches d'enseignements. Merci pour votre patience

à l'égard de ma distraction légendaire. Que la bonne ambiance du labo, la bienveillance et l'entraide continuent à animer notre belle équipe.

A Litsa, Fabienne, pour leur support administratif au quotidien et à Blandine pour sa patience et son merveilleux travail au niveau de la qualité.

A l'ensemble de mes collègues du Département Connaissance et Valorisation des Produits et aux nombreux copains du CRA-W. Tant de chouettes collègues qui font de chaque jour un plaisir à venir y travailler ! A tous les niveaux, vous réalisez un travail formidable. Un merci spécial à mon ami Frédéric Debode d'avoir eu la bonne idée d'attirer le poisson dans le filet du CRA-W.

Merci à mes amis, Carlo Bertozzi, Julie Leblois et toute leur équipe de l'association wallonne de l'élevage (élevéo). Qu'il est agréable et riche de collaborer avec eux. Puisse cette parfaite collaboration perdurer dans le temps et dans de nombreux projets.

Merci à l'ensemble des collaborateurs du European Milk Recording (EMR), et aux très nombreux partenaires des multiples projets européens. Que les nombreuses amitiés nées durant ces projets puissent perdurer, c'est l'essence même de l'Europe.

Aux nombreux collègues du réseau de la Fédération International de Laiterie (FIL), à Philippe, Pierre et Anne et tous les autres, de l'International Committee for Animal Recording (ICAR). On en vient à s'appeler cousin/cousine avec les collègues d'outre atlantique, tant il est vrai que le monde laitier est une grande famille.

Au comité du lait et à la confiance de son directeur Emile Piraux. Sans l'aide et la curiosité de Didier Veselko, beaucoup de nos recherches auraient été bien plus difficiles.

Aux collègues flamands de l'ILVO (Geert, Hadewig, Lieve et Nico) avec qui travailler est toujours un plaisir.

A l'ensemble des éleveurs wallons (et d'ailleurs) pour leur travail quotidien nécessaire à la production de cet aliment fantastique qu'est le lait. En ces temps où l'élevage est si injustement critiqué, j'espère que mes recherches contribueront à faire comprendre et progresser le secteur. Merci pour vos échanges et expériences précieuses qui nous permettent de rester les deux pieds sur terre. Merci en particulier à Eddy Pussemier et Luc Straet qui sont les premiers, il y a bien des années, à m'avoir ouvert les portes de leurs fermes et à m'avoir si gentiment fait part de leurs connaissances et transmis une part de leur passion.

Au regretté professeur Deswysen, homme complexe et attachant, qui m'a donné le goût de la recherche et enseigné la rigueur. Il m'a également transmis sans compter son savoir de la production laitière et de la physiologie des ruminants. Bon nombre de mes recherches sont toujours en lien avec ces acquis. A mes ex-collègues de l'UCL avec qui j'ai tant appris, Christine, Dominique, Daniel, Steve, André et au trop vite disparu El-Hocine.



A mes frères et sœurs adorés, à mes filleuls, neveux et nièces, et mes nombreux amis (trop nombreux pour être tous cités), rencontrés sur les bancs de l'école primaire, au collège, à l'université, au Rhétokot, à l'AGRO, à la Namuroise, amis de colloc, de l'équipe saucisson, des vacances et du Théâtre, des parents des amis des enfants devenus des amis tout court, des lapons, du jogging du dimanche... Merci à tous pour vos encouragements durant ces nombreuses années et pour votre fidélité.

A mes parents trop vite disparus, vous avez cru en moi et vous m'avez fait confiance et dieu sait que cela ne devait pas être facile au départ.

Y, por último, a mis queridos hijos, Inés, Nicolás y Isabel, y a mi mujer Ana, amor de mi vida, que gestionó la casa durante esta tesis y me ha apoyado y soportado durante tantos años. Sin vosotros cariños, no soy nada y nada sería posible.

# Table des matières

<b>Résumé</b> .....	i
<b>Abstract</b> .....	ii
<b>Remerciements</b> .....	iii
<b>Table des matières</b> .....	vi
<b>Liste des figures</b> .....	viii
<b>Liste des tableaux</b> .....	xi
<b>Liste des abréviations</b> .....	xiii
<b>Chapitre 1 : Introduction et objectifs</b> .....	1
1.1. Introduction .....	3
1.2. Objectifs de la thèse.....	22
1.3. Références .....	25
<b>Chapitre 2 : Mid-Infrared Prediction of Bovine Milk Fatty Acids across Multiple Breeds, Production Systems and Countries</b> .....	37
2.1. Points d'attention.....	39
2.2. Abstract.....	43
2.3. Introduction .....	44
2.4. Materials and methods.....	45
2.5. Results and discussion .....	51
2.6. Conclusions .....	57
2.7. Acknowledgments .....	59
2.8. References .....	60
<b>Chapitre 3: Potential use of milk mid-infrared spectra to predict individual methane emission of dairy cows</b> .....	63
3.1. Points d'attention.....	65
3.2. Abstract.....	68
3.3. Introduction .....	68
3.4. Material and methods .....	70
3.5. Results and discussion .....	75

3.6. Conclusion.....	82
3.7. Acknowledgments .....	82
3.8. References .....	83
<b>Chapitre 4: Large-scale phenotyping in dairy sector using milk MIR spectra: Key factors affecting the quality of predictions .....</b>	<b>87</b>
4.1. Points d'attention.....	89
4.2. Abstract .....	90
4.3. Context and framework .....	91
4.4. Material and methods .....	92
4.5. Experiments and results.....	95
4.6. Discussion .....	114
4.7. Conclusions .....	116
4.8. Acknowledgment.....	116
4.9. Funding and disclaimer .....	117
4.10. References .....	118
<b>Chapitre 5: Discussion générale.....</b>	<b>129</b>
5.1. Introduction .....	131
5.2. Echantillonnage .....	132
5.3. Analyses de référence.....	140
5.4. Modèles de prédiction .....	143
5.5. Normalisation .....	145
5.6. Standardisation .....	147
5.7. Contrôle qualité .....	149
5.8. Références .....	155
<b>Chapitre 6 : Conclusion et perspectives .....</b>	<b>167</b>
Références .....	175

# Liste des figures

**Figure 1.1.** Exemple de système de paiement rationnel du lait proposé par la société coopérative de Fléron à ses producteurs sociétaires d'après Collard Bovy A. L'unité de graisse étant payé à 3 centimes de franc belge et le litre de lait à 6 centimes de franc belge. Source : d'après le recueil du Congrès National de Laiterie. 16-17 octobre 1897, Bruxelles.

**Page 4**

**Figure 1.2.** Evolution des quantités livrées de crème (convertie en lait) et de lait aux laiteries, et de lait directement utilisé dans les fermes belges entre 1950 et 1976. D'après Salardi, 1978.

**Page 5**

**Figure 1.3.** Evolution des livraisons de lait aux laiteries de 1984 à 2021 en Belgique en milliers de litres. Source CBL (2022).

**Page 7**

**Figure 1.4.** Nombre d'analyses infrarouge réalisés annuellement par le comité du lait en lien avec le paiement du lait ou les contrôles individuels du contrôle laitier. Source : rapports des activités annuels du Comité du lait de Battice de 2003 à 2022.

**Page 8**

**Figure 1.5.** Productions annuels de lait (en milliers de litres de lait) et du nombre de producteurs laitiers en Wallonie de 2002 à 2021. Source : rapports des activités annuels du Comité du lait de Battice de 2003 à 2022.

**Page 8**

**Figure 1.6.** Principe d'un spectromètre à transformée de Fourier (d'après Bertrand et Dufour, 2000).

**Page 14**

**Figure 1.7.** Spectre moyen infrarouge en transmission d'un échantillon de lait homogénéisé par rapport à l'eau. Les traits pleins de couleurs indiquent les nombres d'ondes des filtres de mesure de la MG (Fat-B en orange et Fat-A en rouge), des protéines (en bleu) et du lactose (vert). Les traits en pointillés indiquent les nombres d'ondes des filtres de référence pour la MG (orange et rouge pour Fat-B et A), pour la protéine (bleu) et pour le lactose (vert) (d'après Bertrand et Dufour, 2000).

**Page 16**

**Figure 1.8.** Représentation schématique des différents niveaux de concentration de MG, MAT et Lactose obtenus par la méthode de séparation/recombinaison de lait permettant la création d'échantillons recombinés (méthode communément appelée FIL 141) pour la calibration et le contrôle qualité des analyses par spectrométrie moyen infrarouge (d'après Leray, 2017).

**Page 20**

**Figure 1.9.** Synthèse des différentes solutions et améliorations techniques visant à améliorer les prédictions MIR des composés majeurs du lait.

**Page 23**

**Figure 2.1.** Effect of first derivative on 2 mid-infrared milk spectra. In order to make an easier reading, the scales of the first derivatives were changed.

**Page 55**

**Figure 3.1.** Possibility to average the spectra milk in function of methane day collect. Vertical lines corresponds to milking time; surface shaded to milk production.

**Page 74**

**Figure 3.2.** Plot of the principal component analysis of milk spectra for individual cows fed three differing diets: corn silage (•), fresh pasture (o) and grass silage (+). PCA = principal component analysis.

**Page 75**

**Figure 3.3.** Infrared CH<sub>4</sub> prediction on basis of milk spectra of the day1.5 for the different diets: corn silage (●), fresh pasture (○) and grass silage (+)

**Page 79**

**Figure 4.2.** Comparison of two models, with different reference values range, when predicting external extreme high values.

**Page 97**

**Figure 4.2.** Comparison of the distribution of the 2 calibration datasets with the distribution of the global population. Second and third principal components of the calibration datasets using random selection (red points, top, n = 200) and selection based on spectral variability (red points, bottom, n = 200), plotted on the global dataset (grey points, n = 3506).

**Page 97**

**Figure 4.3.** R<sup>2</sup> at each wavenumber between absorbance values of a reference instrument and 82 secondary instruments.

**Page 101**

**Figure 4.4.** Results in cross-validation and external-validation for 3 models developed with different spectral ranges.

**Page 102**

**Figure 4.5.** Calibration and validation statistical performances of three models predicting the body weight of cows using different numbers of PLS latent variables, selected after the calibration (left), cross-validation (central) and in validation (right) steps.

**Page 103**

**Figure 4.6.** *Boxplots representation of PUFA (Poly-Unsaturated Fatty Acids) predictions, for a common sample of the inter-laboratories study analyzed on 66 instruments, for master spectra, non-standardized spectra, and spectra after standardization, after removing of aberrant values (Grelet et al., 2017).*

**Page 114**

**Figure 4.7.** *Plot of fat predictions over time originating from a fat model applied on spectra from an instrument analyzing a common batch of UHT milk during a period of 3 months.*

**Page 116**

**Figure 5.1.** *Points d'attention visant à assurer la qualité des prédictions FT-MIR du lait et d'accroître leur robustesse.*

**Page 131**

**Figure 5.2.** *Stratégies permettant d'accroître la variabilité des modèles : grâce à la variabilité spectrale X, grâce à la variabilité des valeurs de référence y et aux échanges de données via des collaborations internationales (d'après Vanlierde et al., 2021).*

**Page 134**

**Figure 5.3.** *Moyenne  $\pm$  ET de 14 troupeaux traits 2 $\times$  pour la production laitière et les pourcentages de matières grasses et de protéines pour chaque période de traite sur une période de 5 jours. D'après Quiest et al. (2008).*

**Page 137**

**Figure 5.4.** *Concentrations (moyennes des moindres carrés) des constituants du lait pendant la traite (% de la traite) en fonction de l'intervalle de traite (6 ou 12 h) et de l'état de santé du quartier échantillonné (sain ou malsain). Les moyennes simples des premiers jets de lait sont indiquées à -10% de la traite. Sain6 (■); Malsain6 (▲); Sain12 (×); Malsain12 (○). D'après Nielsen et al. (2005).*

**Page 138**

**Figure 5.5.** *Nouveaux paramètres prédits par les nouvelles calibrations FT-MIR. Source : Pécou et al., 2020).*

**Page 150**

**Figure 5.6.** *Programme de contrôle qualité appliqué pour les nouveaux paramètres prédits par FT-MIR. Source : Pécou et al., 2020).*

**Page 151**

# Liste des tableaux

**Tableau 1.1.** *Présentation des différentes techniques spectroscopiques selon la gamme de longueurs d'ondes utilisée.*

**Page 9**

**Table 2.1.** *Descriptive statistics of gas chromatographic data for the calibration and validation sample sets.*

**Page 52**

**Table 2.2.** *Comparison of the 6 mathematical approaches (methods 1 to 61) used to develop the calibration equations for fatty acids (FA) in milk (g/dL of milk; n = 267 samples).*

**Page 54**

**Table 2.3.** *Coefficient of determination between the reference GC data and the prediction (Pred.) of fatty acid (FA) contents in fat (g/100 g of fat) obtained from the predictions of fatty acid in milk (g/dL of milk).*

**Page 56**

**Table 2.4.** *Accuracy of mid-infrared predictions for fatty acid (FA) in milk (g/dL of milk) using the entire data set<sup>1</sup>*

**Page 58**

**Table 3.1.** *Estimated statistical parameters for each calibration equation that estimated the concentrations of fatty acid in milk (g / dl milk) used in this study.*

**Page 71**

**Table 3.2.** *Means and range of CH<sub>4</sub> emissions, production and composition of the milk, SAT composition of the milk fat for the two experiments (day 0).*

**Page 76**

**Table 3.3** *Comparison between components prediction for AMS and spectra of average milk in a ratio 50-50 for 37 milk sample during the second experiment.*

**Page 77**

**Table 3.4.** *Evolution of accuracy of the resulting calibration with the different kind of AMS with a ratio 50-50 and a weighted average for the two experiments. The number of available data depends on the chosen day.*

**Page 80**

**Table 3.5.** *Individual correlations between different milk parameters and the CH<sub>4</sub>/kg of milk produced for the day1.5.*

**Page 81**

**Table 3.6.** *Multiple correlations from the parameters appearing in Table 3.5 to predict the CH<sub>4</sub>/kg of milk produced for the day1.5.*

**Page 81**

**Table 4.1.** *Datasets used in the study.*

**Page 93**

**Table 4.2.** *Comparison of the models based on the two different selection methods to constitute the calibration dataset.*

**Page 99**

**Table 4.3.** *Spectral regions used in the development of the 3 models.*

**Page 101**

**Table 4.4.** *Characteristics of the 7 K-mean clusters resulting from the classification of 57 milk MIR models following their mean-centered cross-validation RPD, relative RMSE and R<sup>2</sup>.*

**Page 106**

**Table 4.5.** *Evaluation of 4 models through RMSE and with different validation strategies.*

**Page 109**

**Tableau 6.1.** *Tableau synthétique des facteurs impactant la qualité/robustesse des modèles, les problèmes, difficultés et points d'attention rencontrés et les solutions/recommandations apportées sur la première composante, l'Erreur total de l'échantillonnage (TSE), de l'Erreur Globale d'estimation (GEE).*

**Page 170**

**Tableau 6.2.** *Tableau synthétique des facteurs impactant la qualité/robustesse des modèles, les problèmes, difficultés et points d'attention rencontrés et les solutions/recommandations apportées sur la deuxième composante, l'Erreur analytique globale (TAE), de l'Erreur Globale d'estimation (GEE).*

**Page 171**



# Liste des abréviations

AG: Acides gras

AGV: Acides gras volatils

AMS: Avrage milk spectra

AOAC: Association of Official Agricultural Chemists

BHB:  $\beta$ -hydroxybutyrate

CH<sub>4</sub>: Méthane / Methane

CNIEL: Confédération national de l'interprofession laitière française

CRA-W-DVP : Centre Wallon de Recherches Agronomiques - Département Valorisation des Productions

CV : Cross validation ou validation croisée

DM: Dry matter

DMI: Dry matter intake

EMR: European milk recording

FA: Acides gras

FAO: Food and Agriculture Organization

FCY: Milk fresh cheese yield

FT-MIR: Spectrométrie moyen infrarouge à transformée de Fourier

FIL: Fédération Internationale de Laiterie/ International Dairy Federation IDF

FT: Transformée de Fourier

GC : chromatographie en phase gazeuse / Gas chromatography

GEE: Erreur Globale d'estimation

GH: Standardized Mahalanobis distance or Global-H distance

ICAR: Comité international pour le contrôle des performances en élevage/  
International Committee for Animal Recording

IGF-I: Insulin growth factor

ILVO-T&V: Instituut voor Landbouw en Visserij Onderzoek - eenheid Technologie en Voeding

IR: Infrared

IRMA: Infrared Milk Analyzer  
ISO: Organisation Internationale de normalisation/ International Organization for Standardization  
k20: Time when the curd is firm enough for cutting  
KNN: K-means near neighbor  
LCFA: Acides gras à longues chaînes  
LOO: Leave-one-out cross-validation  
LV: Latent variables  
MADP: Median absolute deviation of predictions  
MAE: Mean absolute error  
MAT: Matière azotée totale  
MCFA: Acides gras à moyennes chaînes  
MG: Matière grasse totale  
MIR: Moyen infrarouge / Mid-infrared  
MLR: Régression multilinéaire  
MONO: Acides gras mono-insaturés  
MSE: Mean square error  
NAGase: N-actetyl- $\beta$ -D-Glucosamidase  
NEFA: Non-esterified fatty acids  
NIR: Proche infrarouge / Near-infrared  
NH: Neighboring H distance  
NUE: Nitrogen efficiency  
OI: Organismes Interprofessionels  
ONL: Office national du lait  
PC: Principal component  
PLS: Régression des moindres carrés partiels / Partial least squares regression  
POLY : Acides gras polu-insaturés  
PUFA: Poly-unsaturated fatty acids  
QA: Assurance qualité/Quality Assurance  
R<sup>2</sup>: Coefficient of determination

$R^2_c$ : Coefficient of determination de la calibration  
 $R^2_{cv}$ : Coefficient of determination de la validation croisée  
 $R^2_v$ : Coefficient of determination de la validation  
RER: Ratio of range to RMSE  
RMSE: Root mean square error  
RMSE<sub>cv</sub>: Root mean square error of cross-validation  
RMSEP: Root mean square error of validation  
RPD: Ratio of standard deviation to RMSE  
SAT: Acides gras saturés  
SCFA: Acides gras à courtes chaînes  
SD: Ecart-type  
SD<sub>y</sub>: Standard deviation of the reference values  
SEC: Erreur type d'étalonnage  
SECV: Standard error of cross-validation  
SEL: Standard error of laboratory  
SEP: Erreur type de prédiction  
SF<sub>6</sub>: Hexafluorure de soufre  
TAE: Erreur analytique globale  
TMR : Total mixed rations  
TSE: Erreur total de l'échantillonnage  
ULiège : Université de Liège  
UNSAT : Acides gras insaturés  
VEM : VoederEenheid Melk  
VFA : Volatile fatty acids  
X: Spectral data  
y: Reference values



# Chapitre 1 :

---

**Introduction et objectifs**



## 1.1. Introduction

### *1.1.1. Nécessité de nouveaux moyens de mesures du lait liés aux demandes analytiques croissantes*

Au cours du dernier siècle, le secteur laitier a été confronté à la nécessité de créer de nouveaux moyens de mesure de la composition et de la qualité du lait. Ce phénomène est notamment consécutif à un accroissement des niveaux de production, le développement de nouvelles techniques laitières, des changements radicaux dans l'organisation économique et sociétale de l'agriculture, au développement de nouvelles coopératives de ramassage et à la création de nouveaux échanges entre producteurs et transformateurs.

Dès le début, le secteur laitier a dû résoudre les problèmes liés à la détection des fraudes (mouillage, ajout de source de graisses non butyriques) et s'outiller pour permettre un paiement équitable et juste du lait. Parmi les conclusions et recommandations du premier congrès national belge de laiterie qui s'est tenu à Bruxelles en 1897 (De Bruyn and Peers, 1898), il est notamment précisé qu'il y a lieu d'effectuer des recherches analytiques dans le but de déterminer la valeur argent ou monétaire du lait et que la législation belge devra être complétée afin de stipuler un minimum de richesse en éléments utiles liés à la composition du lait. Lors du congrès, il est souligné qu'une des difficultés rencontrées et qui contribue à retarder la formation des sociétés coopératives pour la vente du lait, c'est la difficulté de trouver une forme de paiement rationnelle pour le lait livré par les fournisseurs associés. Ceux qui livrent un lait riche, à juste raison, veulent être mieux payés que ceux qui fournissent un lait pauvre ou de qualité médiocre. Déjà à l'époque, on souligne que deux facteurs doivent intervenir dans la fixation du prix du lait : la quantité et la qualité. Les sociétés transformatrices ont avantage à utiliser du lait riche afin de garantir une meilleure qualité du produit fini (essentiellement sous forme de beurre à l'époque). Elles doivent chercher à rémunérer les producteurs proportionnellement à la qualité du lait produit pour les inciter à livrer le meilleur lait possible. Une première idée de système rationnel de paiement du lait lié au degré de graisse du lait fut présentée par Collard-Bovy. Elle repose sur un nouveau système de rémunération appliqué par la Société coopérative laitière de Fléron (Figure 1.1). Au cours de ce même congrès, le secteur laitier avait également marqué le besoin de créer d'une part, une Société Nationale de laiterie de Belgique pour aider à réglementer les échanges et contrôles du lait et produits laitiers, et d'autre part, le besoin de créer une station laitière de l'état afin d'exécuter des expérimentations et vérifier les nouvelles méthodes et nouveaux appareils analytiques. Sous la Direction du Dr. Maurice Henseval, la Station laitière de Gembloux est créée en 1900. Dès ses premières publications, il va s'intéresser à l'étude du dosage de la matière grasse du lait (Henseval, 1904).

Dans son discours d'introduction au premier Congrès International de Laiteries de 1903 (Van der Bruggen and Peers, 1903), Peers de Nieuwburg, président du congrès, indiquait que depuis 1901, lors du Congrès National d'Agriculture de Namur, dans une section spéciale consacrée à l'examen et l'étude des questions laitières, l'idée de la création d'une commission internationale fut retenue. Il était rapidement apparu lors des discussions, que pour certaines questions laitières, notamment relatives aux questions analytiques, les réponses risquaient de rester stériles si l'on ne parvenait pas à obtenir une entente avec les confrères étrangers. Portée par cette volonté de chercher des solutions via une collaboration internationale étendue, la Belgique décida d'organiser ce premier congrès international spécifiquement dédié aux questions laitières. C'est à l'issue de ce premier congrès international qu'il fut décidé de créer une nouvelle Fédération Internationale de Laiterie (FIL ou International Dairy Federation), dont le siège actuel est toujours situé à Bruxelles. Cette fédération a notamment pour but de travailler à la standardisation et à la recherche de nouvelles méthodes analytiques, simples, fiables et reproductibles.

Un lait marquant	2°5	de graisse sera payé	6 + (2.5 x 3) = 13 c.5
"	"	3°	" " 6 + (3 x 3) = 15 c.
"	"	3°5	" " 6 + (3.5 x 3) = 15 c.5

**Figure 1.1.** Exemple de système de paiement rationnel du lait proposé par la société coopérative de Fléron à ses producteurs sociétaires d'après Collard Bovy A. L'unité de graisse étant payé à 3 centimes de franc belge et le litre de lait à 6 centimes de franc belge. Source : d'après le recueil du Congrès National de Laiterie. 16-17 octobre 1897, Bruxelles.

Rapidement, des efforts visant à trouver de nouvelles méthodes indirectes pour doser les composants du lait ont été effectués. A titre d'exemple, une première méthode d'analyse reposant notamment sur l'emploi de la lumière de Wood (rayon ultra-violet filtré) a été testée sur le lait par plusieurs auteurs dès la fin des années vingt (Volmar Y., 1927 ; Schutzler K., 1927). Comme indiqué par Seyewetz, les avantages principaux de cette méthode reposaient sur sa rapidité, sa sensibilité et la possibilité de faire l'analyse sans porter atteinte à l'échantillon (Seyewetz J., 1934).

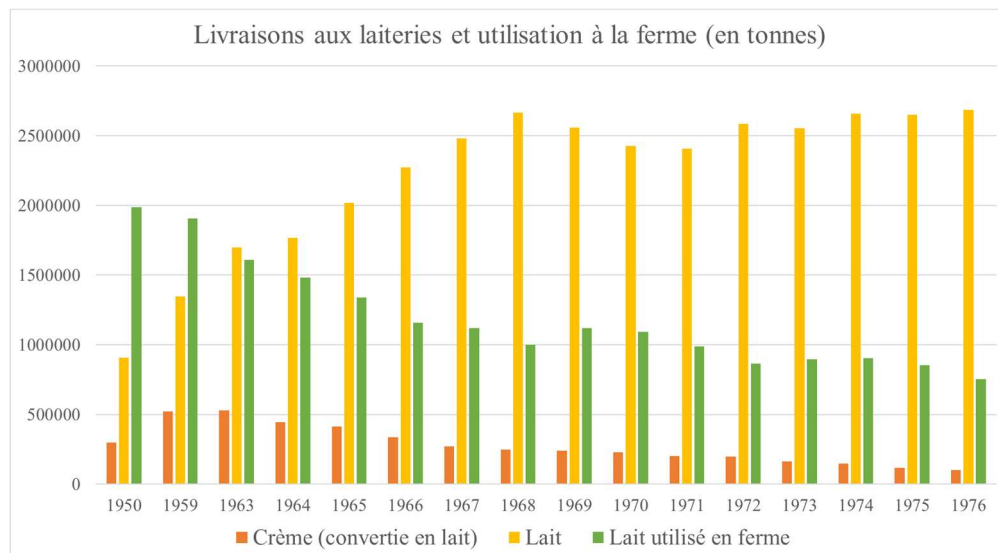
Ces premiers résultats encourageants sont à l'aube de recherches visant à trouver des techniques analytiques permettant de répondre aux demandes croissantes du secteur laitier. Elles vont permettre d'obtenir des solutions de mesures précises, rapides et non destructrices, nécessaires à la réalisation croissante des analyses de routine liées à la détermination de la composition et la qualité du lait. Au cours du vingtième siècle, le secteur laitier et les besoins en données analytiques vont rapidement croître suite notamment au changement structurel du secteur laitier, à l'évolution dans l'utilisation des productions laitières, à la modification des exigences



sanitaires et qualitatives, à l'évolution des méthodes de paiement, et à l'apparition du contrôle laitier.

Les premières laiteries privées ou coopératives sont apparues, en Belgique, aux alentours des années 1883-1887, principalement sous la forme d'unités d'écémage à bras ou à vapeur et de fabrication de beurre. En 1910, on comptait 227 laiteries à vapeur et 301 laiteries à bras. Par la suite, les exigences de qualité des consommateurs issus des villes et la demande croissante en lait de consommation vont jouer en faveur du développement de ces nouvelles laiteries. Lors du recensement de 1937, on comptait 513 laiteries à vapeur et 128 laiteries à bras (De Baere, 1978). Ces dernières auront pratiquement disparu au début de la Seconde guerre mondiale (Salardi, 1978). Peu après la Deuxième Guerre, s'amorce les mouvements de concentration et d'intégration.

En 1965, en Wallonie, le lait de consommation était encore presque exclusivement vendu à la ferme (70 % ou 80 %, en moyenne et jusqu'à 90 % dans le Hainaut) ; en revanche, dans la partie flamande du pays, cette consommation de lait de ferme ne représentait déjà plus que 30 à 40 % du volume des ventes et dans la province de Brabant (Bruxelles compris), seulement environ 10 % (CEE, 1965). En Flandre, 96,4 % des livraisons s'effectuaient sous forme de lait et 3,6 % sous forme de crème fermière. Tandis qu'en Wallonie, 35,3 % des livraisons s'effectuaient sous forme de lait et 64,7 % sous forme de crème fermière nécessaire à la fabrication du beurre (Salardi, 1978).



**Figure 1.2.** Evolution des quantités livrées de crème (convertie en lait) et de lait aux laiteries, et de lait directement utilisé dans les fermes belges entre 1950 et 1976. D'après Salardi, 1978.

Après 1965, on a observé des évolutions dans le domaine des livraisons aux laiteries (Figure 1.2) : les livraisons de lait ont pratiquement doublé et celles de crème

diminuèrent progressivement (elles ne représentent en Wallonie plus que 13 % du total des livraisons en 1975). D'autre part, les volumes de lait utilisés à la ferme diminueront rapidement. Ces évolutions s'expliquent par de multiples facteurs, dont notamment : l'intérêt croissant des consommateurs pour le lait stérilisé en bouteille; l'accroissement du cheptel laitier des exploitations agricoles ; le surplus de travail exigé par l'étape d'écémage, d'autant plus astreignant à mesure de l'augmentation des cheptels; par la diminution des ventes directes à la ferme; par le soutien du prix du lait à partir de la mise en œuvre de la phase transitoire du Marché Commun, qui ne concernait par contre pas la rémunération de la crème; par l'adaptation de nouvelles techniques telles que le recours à la traite mécanique et à l'utilisation des tanks réfrigérés de lait (De Herde, 2020).

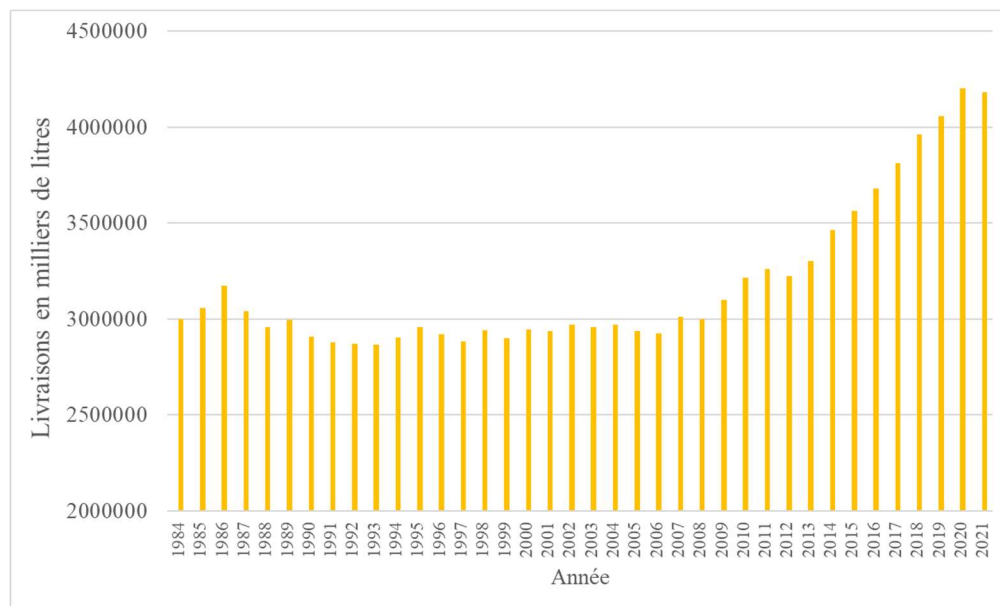
A cela, ajoutons que dès 1953, des efforts avaient été entrepris par l'état belge afin d'accroître la qualité du lait de consommation par l'état belge. Ainsi, les marques de contrôle pour le lait de haute qualité furent instituées par l'Arrêté Royal (A.R.) du 28 janvier 1953, dans lequel deux marques ont été prévues : la marque A et la marque AA. Ensuite, l'A.R. du 31 mars 1954 relatif au commerce du lait de vache réglementa plusieurs critères : la dénomination des laits ; la vente et la qualité du lait de ferme et de laiterie vendu aux consommateurs ; la qualité du lait réceptionné par la laiterie et du lait écrémé et battu ristourné à la ferme ; l'utilisation des récipients ; les mentions et inscriptions sur les récipients ; le transport du lait et l'emploi des ustensiles et appareils. D'autres contrôles officiels seront mis en œuvre à la même époque en Belgique et concerneront la fabrication du beurre, du fromage à pâte dure et des poudres de lait (Salardi, 1978). La transformation du lait en beurre de ferme destiné à la vente a notamment été soumise à l'octroi d'une licence par l'Office National du Lait (ONL). Pour obtenir cette licence, le producteur devait remplir des conditions très sévères de production, d'équipement et d'hygiène. On retrouvait les mêmes tendances pour la transformation en d'autres produits laitiers.

Dès 1964, en Belgique, le lait livré par les producteurs à la laiterie est soumis à plusieurs analyses afin de déterminer sa composition et sa qualité. Ce règlement a été établi par le ministère de l'Agriculture qui a notamment chargé les laboratoires interprofessionnels provinciaux nouvellement créés de Battice, Soignies et Ciney, d'effectuer l'ensemble des analyses de routine liées au paiement du lait pour l'ensemble de la région wallonne. A partir de 1998, la détermination de la qualité du lait cru était régie par les arrêtés royaux et ministériels du 17/03/1994 (M.B. du 06/05/94) et modifiés par la suite (Laloux, 2000). A l'heure actuelle, il n'y a plus que deux organismes interprofessionnels en Belgique qui effectuent les analyses sur le lait cru : le Melkcontrolecentrum-Vlaanderen de Lier pour la Flandre et le Comité du Lait de Battice pour la Wallonie. Ces deux organismes sont officiellement reconnus par les pouvoirs publics. Les résultats de leurs analyses constituent la base du calcul du prix du lait nécessaire à des échanges équitables dans le secteur laitier. Dans le cadre de cette mission, une participation à la guidance scientifique des Organismes Interprofessionnels (OI) est prévue par les instances publiques. Cette guidance est assurée en partenariat entre le CRA-W-DVP (Centre Wallon de Recherches

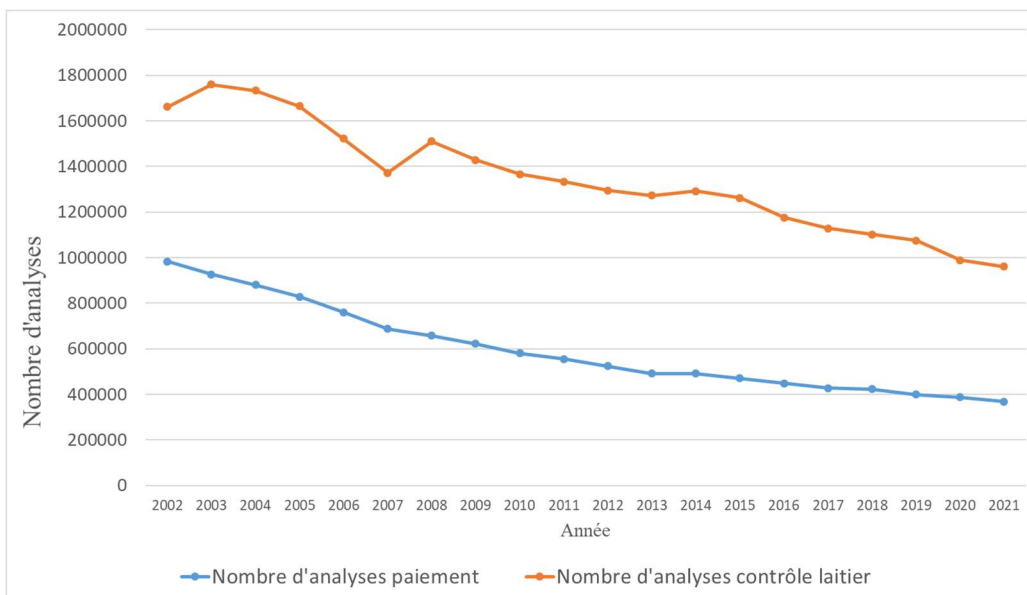
Agronomiques - Département Valorisation des Productions) de Gembloux et l'ILVO-T&V (Instituut voor Landbouw en Visserij Onderzoek - eenheid Technologie en Voeding) de Melle. Elle doit vérifier que les OI travaillent avec justesse et uniformité.

Enfin, signalons que dès 1957, le contrôle des performances des vaches laitières, encore appelé le contrôle laitier, est généralisé sur le territoire belge. Cela entraînera la création une année plus tard du premier centre de calcul national destiné à traiter les données collectées. Les premières évaluations génétiques datent de la fin des années 60. L'accroissement des mesures additionnelles de la composition de laits individuels liées au contrôle laitier ne va être rendu possible que par l'apparition des nouvelles techniques rapides de la spectrométrie moyen infrarouge.

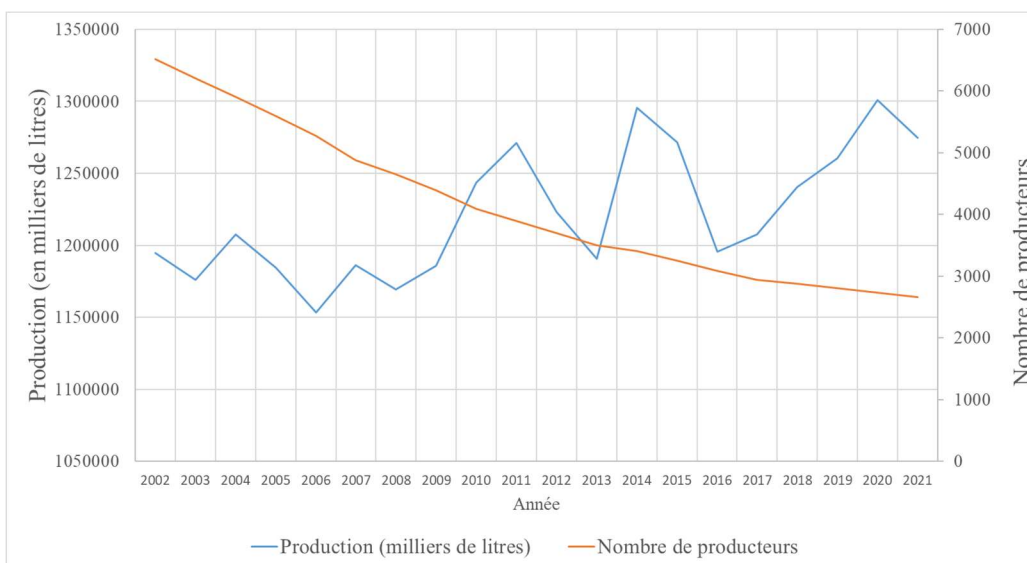
De 1984 à 2021, les livraisons de lait en Belgique ont augmenté de plus d'un million de litres de lait. Cet accroissement a été accentué par la disparition des quotas laitiers en avril 2015 (Figure 1.3). Durant les 20 dernières années, le nombre d'analyses dans l'infrarouge utilisées pour doser notamment les teneurs en matières grasses et protéines s'érode chaque année en Wallonie, tant pour les analyses liées à la composition du lait que celles liées au contrôle laitier (Figure 1.4) et ce malgré une production totale de lait croissante (Figure 1.5). Ceci s'explique d'une part par un nombre décroissant de producteurs laitiers, passant de 6.519 producteurs en Wallonie en 2002 à 2.656 producteurs en 2021, alors que dans le même temps, la production laitière totale croissait d'un peu moins de 7% de 2002 à 2021, soit une augmentation de la production moyenne par producteur passant de 183.263 à 479.866 litres de lait. De plus, le cheptel laitier wallon a été diminué de 2002 à 2021, de 257.559 (CBL, 2004) à environ 193.129 vaches laitières (CBL, 2022).



**Figure 1.3.** Evolution des livraisons de lait aux laiteries de 1984 à 2021 en Belgique en milliers de litres. Source CBL (2022).



**Figure 1.4.** Nombre d'analyses infrarouge réalisées annuellement par le comité du lait en lien avec le paiement du lait ou le contrôle laitier. Source : rapports des activités annuels du Comité du lait de Battice de 2003 à 2022.



**Figure 1.5.** Production annuelle de lait (en milliers de litres de lait) et nombre de producteurs laitiers en Wallonie de 2002 à 2021. Source : rapports des activités annuels du Comité du lait de Battice de 2003 à 2022.

## 1.1.2. Spectrométrie moyen infrarouge à transformée de Fourier

### 1.1.2.1. Une technique spectroscopique

La spectrométrie moyen infrarouge (ou encore MIR, de l'anglais Mid InfraRed spectroscopy) fait partie de la famille des techniques spectroscopiques qui utilisent la capacité d'interaction de rayonnements électromagnétiques avec la matière pour permettre de mesurer rapidement la composition d'un échantillon.

**Tableau 1.1.** Présentation des différentes techniques spectroscopiques selon la gamme de longueurs d'ondes utilisée.

Techniques spectroscopiques	Gamme de longueurs d'ondes
Rayons X	0,5 à 10 nm
Ultraviolets	10 à 350 nm
Visible	350 à 780 nm
Proche infrarouge (NIR)	780 à 2500 nm
Moyen infrarouge (MIR)	2500 à 25000 nm
Micro-ondes	100 $\mu\text{m}$ à 1 cm

Les différentes techniques spectroscopiques sont différenciées grâce à la gamme des longueurs d'ondes qui les constituent (Tableau 1.1). La spectrométrie MIR est appelée spectrométrie moyen infrarouge à transformée de Fourier<sup>1</sup> (ou encore FT-MIR, de l'anglais Fourier Transform Mid InfraRed spectroscopy) lorsque le spectromètre permet de collecter simultanément les données sur un spectre large. Par opposition, les spectromètres à dispersion MIR, communément appelés appareils à filtres, ne peuvent mesurer l'intensité que dans des gammes réduites de longueurs d'ondes et de façon séquentielle. Les filtres utilisés déterminent la ou les longueurs d'ondes mesurées.

Dans un spectromètre FT-MIR, les échantillons sont traversés par des radiations issues d'une gamme spectrale se situant dans la région des longueurs d'ondes du MIR. Les gammes des longueurs d'ondes proposées peuvent être légèrement différentes selon le type et la marque des appareils utilisés, mais sont spécifiquement situées dans la région située entre 2.500 et 25.000 nm ou encore de 4.000 à 400  $\text{cm}^{-1}$  lorsque l'on exprime en nombre d'ondes. En effet, en MIR, on

<sup>1</sup> L'expression « à transformée de Fourier » fait référence au fait qu'une transformation de Fourier est requise pour transformer la donnée brute en un spectre réel

informe couramment de la portée spectrale utilisée en l'exprimant en fréquence spatiale par cm, ce qui équivaut à donner le nombre d'ondes (en anglais wave number) par cm. Le nombre d'ondes ( $\sigma$ ) équivaut à la fonction inverse de la longueur d'onde ( $1/\tilde{\nu}$ ) (Grelet, 2019). Lorsque  $\tilde{\nu}$  est exprimé en nm, on multipliera  $1/\tilde{\nu}$  par  $10^7$  pour exprimer  $\sigma$  en  $\text{cm}^{-1}$  (Williams, 2007).

La spectrométrie MIR suit le principe général de la spectroscopie infrarouge d'absorption qui repose sur les interactions entre la matière et plus spécifiquement sur la mesure de la quantité de lumière absorbée et dispersée par un échantillon, en fonction du nombre d'ondes. La spectrométrie MIR est une technique qualitative et quantitative (Bertrand and Dufour, 2000).

### *1.1.2.2. Une technique qualitative*

Dans la spectrométrie MIR, ce sont les vibrations fondamentales et la structure vibrationnelle des liaisons qui sont concernées. En effet, selon la nature des atomes constituant une liaison, celle-ci peut vibrer à une ou plusieurs fréquences bien précises et peut interagir avec les rayons infrarouges qui ont précisément ces mêmes fréquences (nombre d'ondes) et va alors absorber l'énergie de ce rayonnement (Bertrand and Dufour, 2000). Retenons que les liaisons covalentes ont la capacité d'absorber la lumière à différentes fréquences, car différents types de vibrations sont possibles (élongations et déformations) selon les molécules (Rouessac et al., 2016). La spectrométrie MIR est donc une technique qualitative, qui permet, selon l'étude des bandes d'absorption du spectre en fonction des nombres d'ondes, de déterminer la composition de l'échantillon mesuré (Bertrand and Dufour, 2000).

### *1.1.2.3. Une technique quantitative*

En outre, par l'étude de l'intensité des bandes d'absorption du spectre, on va pouvoir quantifier le ou les composés présents, étant donné que l'absorbance est directement liée à la concentration (Bertrand and Dufour, 2000). Cette relation est régie par la loi de Beer-Lambert-Bouguer (couramment appelée loi de Beer-Lambert) qui est la loi fondamentale de la spectroscopie quantitative (Griffiths and de Haseth, 2007). La Transmittance ( $T$ ) de tout échantillon à la longueur d'onde ( $\tilde{\nu}$ ) est donnée par le rapport entre l'intensité énergétique du rayonnement réfléchi ou transmis ( $I(\tilde{\nu})$ ) et l'intensité énergétique du rayonnement incident ( $I_0(\tilde{\nu})$ ). La Transmittance à travers un échantillon pur dont l'épaisseur est  $b$  (cm) à une longueur d'onde ( $\tilde{\nu}$ ) est donnée par la loi de Beer-Lambert :

$$T(\tilde{\nu}) = \frac{I(\tilde{\nu})}{I_0(\tilde{\nu})} = \exp[-\alpha(\tilde{\nu})b]$$

*D'après Griffiths and de Haseth, 2007*

où  $\alpha(\tilde{\nu})$  est le coefficient d'absorption linéaire (en  $\text{cm}^{-1}$ ) à la longueur d'onde  $\tilde{\nu}$ . L'Absorbance de l'échantillon à la longueur d'onde  $\tilde{\nu}$ ,  $A(\tilde{\nu})$ , est donné par le logarithme en base 10 de  $1/T(\tilde{\nu})$  :

$$A(\tilde{\nu}) = \log_{10} \frac{1}{T(\tilde{\nu})} = \frac{1}{\ln 10} \alpha(\tilde{\nu})b$$

*D'après Griffiths and de Haseth, 2007*

$(1/\ln 10)\alpha(\tilde{\nu})$  est l'absorptivité à  $\tilde{\nu}$ ,  $a(\tilde{\nu})$ .

Si l'échantillon mesuré est un mélange, l'absorbance de chaque composant,  $i$ , à la concentration  $c_i$  est donné par la loi Beer-Lambert :

$$\log_{10} \frac{1}{T(\tilde{\nu})} = A_i(\tilde{\nu}) = a_i(\tilde{\nu})bc_i$$

*D'après Griffiths and de Haseth, 2007*

où  $a_i(\tilde{\nu})$  est en unité de (concentration . longueur de trajet)<sup>-1</sup>.

Pour  $N$ -composés mélangés où plus d'un composé absorbe à  $\tilde{\nu}$ , l'absorbance totale est donnée par :

$$A(\tilde{\nu}) = \sum_{i=1}^N [a_i(\tilde{\nu})bc_i]$$

*D'après Griffiths and de Haseth, 2007*

#### **1.1.2.4. Analyses d'échantillons de lait liquide : MIR ou NIR ?**

Depuis les premiers travaux de Goulden en 1964, la spectrométrie MIR s'est rapidement imposée comme méthode spectrométrique pour l'analyse des laits liquides notamment pour sa capacité à déterminer la composition en éléments majeurs (matières grasses (MG), matières azotées totales (MAT) et lactose) de façon très précise avec une préparation d'échantillon limitée (Bertrand et Dufour, 2000). Comme déjà dit au point 1.1.2.2., la spectrométrie MIR se base sur les vibrations fondamentales et la structure vibrationnelle des liaisons. Les rapports signal/bruit sont intenses même dans l'eau (le lait est composé à environ 88% d'eau) et les informations chimiques spécifiques sont plus claires et plus intenses comparativement au NIR. Dès le début de l'utilisation du MIR, ceci a facilité l'interprétation des résultats et a procuré un avantage pour une utilisation sur la matrice du lait par rapport au NIR. En revanche,

vu que l'énergie de l'onde dépend de sa fréquence (voir la formule ci-dessous (Pu et al., 2020)), l'énergie du rayonnement MIR est relativement basse. Cela oblige d'utiliser en transmission des cellules très fines (quelques microns) et un homogénéisateur accroissant le coût des appareillages. Cependant, l'apparition d'appareils automatiques précis et à haute cadence d'analyse (jusqu'à 600 échantillons par heure) va contribuer à accroître leur utilisation à large échelle dans les laboratoires de routine pour les analyses liées au paiement et au contrôle laitier. La spectrométrie MIR va donc s'imposer comme la méthode couramment utilisée et reconnue pour analyser le lait (Biggs, 1972).

$$E = h \cdot f = \frac{h \cdot c}{\lambda}$$

$E$  est l'énergie (J) ;  $h$  est la constante de Planck (J.s) ;  $f$  est la fréquence de l'onde (Hz ou  $s^{-1}$ ) ;  $c$  est la vitesse de la lumière (m s<sup>-1</sup>) et  $\lambda$  est la longueur d'onde de l'onde

*D'après Pu et al., 2020*

Dans une moindre mesure, d'autres matrices alimentaires liquides, comme les jus de fruits (Sinnaeve et al., 1997), le vin (Patz et al., 2004) ou encore la bière (Cernuda et al., 2017), peuvent également utiliser la spectrométrie MIR comme méthode de mesure.

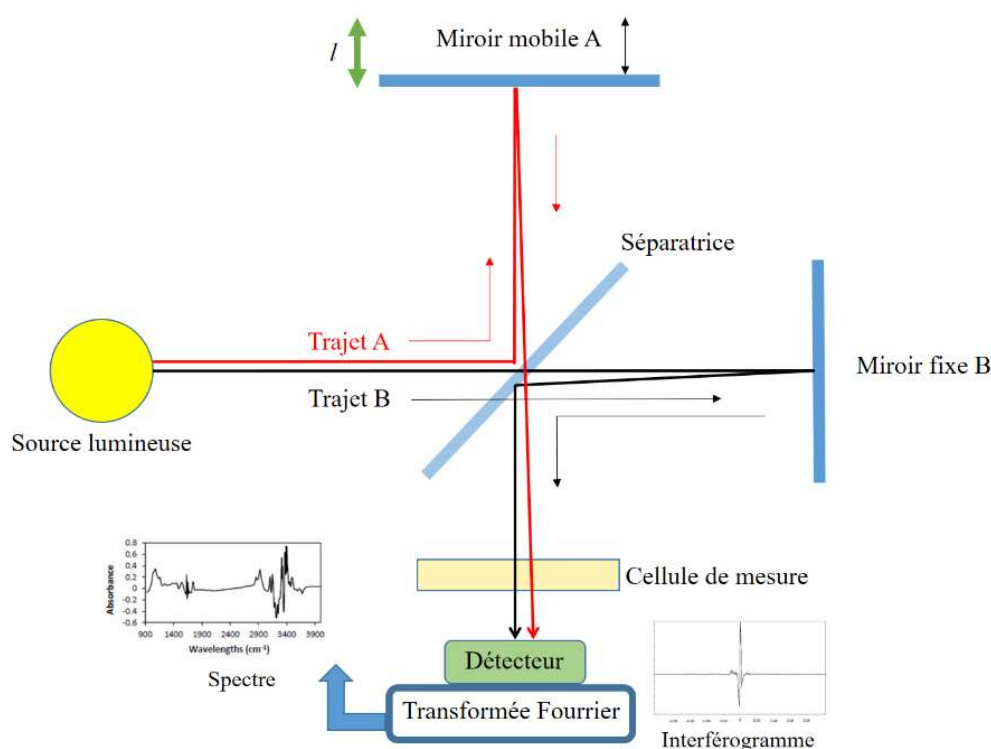
Afin de diminuer le coût de l'appareillage, on pourrait être tenté d'utiliser la spectrométrie NIR. Contrairement au MIR, les spectres NIR sont constitués des absorbances liées aux harmoniques et aux bandes de combinaison des liaisons chimiques. Par rapport aux spectres MIR, les bandes spectrales de la région NIR sont plus larges et se chevauchent, ce qui rend leur interprétation spectrale plus difficile. Selon Norris (2009), la NIR ne peut pas mesurer les constituants inférieurs à 0,1 %. Par contre, vu que la région NIR a une énergie plus élevée que la MIR, les ondes NIR pénètrent plus profondément l'échantillon et permet l'utilisation de cellules de taille plus larges, ce qui permet de mesurer un volume d'échantillon plus important et facilite l'analyse d'échantillons moins homogènes. Les ondes NIR peuvent passer à travers des matériaux tel que le verre ou le plastique en perdant beaucoup moins d'énergie que celles issues du MIR. Cela confère à la spectrométrie NIR un avantage majeur par rapport à la MIR pour des utilisations en ligne ou sur le terrain (Nguyen et al, 2017). Par contre, l'utilisation de la spectrométrie NIR dans le lait est limitée par l'interaction de l'eau avec le signal NIR. On retiendra que son utilisation est particulièrement bien adaptée pour analyser les échantillons laitiers plus sec (fromage, poudres, ...).

#### **1.1.2.5. Spécificités de la FT-MIR**

Contrairement aux appareils monochromatiques dits à filtres qui ne peuvent mesurer qu'un nombre limité d'absorbances à des nombres d'ondes fixes par chacun des filtres utilisés, les spectromètres FT-MIR permettent de mesurer presque simultanément l'ensemble des absorbances de la gamme spectrale du moyen infrarouge. Pour cela, les spectromètres FT-MIR utilisent un élément essentiel appelé



interféromètre de Michelson. Ce dispositif se compose de deux plans perpendiculaires à l'axe optique (Figure 1.6). Comme expliqué par Bertrand et Dufour (2000), ce dispositif est composé d'un miroir fixe et d'un miroir mobile qui peut se déplacer le long de l'axe optique. Une source (généralement de type laser He-Ne) émet un faisceau de lumière MIR qui entre dans l'interféromètre puis est divisée en deux parties grâce à une lame semi-transparente, appelé *séparatrice* (ou *beamsplitter* en anglais). Sous l'action de la séparatrice, la moitié de l'intensité lumineuse va être réfléchi et l'autre moitié sera transmise. La partie réfléchi va atteindre le miroir mobile qui peut se déplacer d'une longueur  $l$  comme indiqué par la flèche verte sur la Figure 1.6. La lumière sera cette fois réfléchi par le miroir mobile pour atteindre à nouveau la séparatrice où elle sera transmise directement vers l'échantillon et le détecteur (trajet A en rouge). L'autre moitié du rayon lumineux initial suit un trajet différent (trajet B, en bleu) : le rayon est cette fois, transmis (et non réfléchi) par la séparatrice puis réfléchi par le miroir fixe, pour être enfin réfléchi lors de son retour par la séparatrice en direction de l'échantillon et du détecteur. Vu la présence d'un miroir mobile se déplaçant d'une longueur  $l$  pour le trajet A, les trajets optiques A et B vont différer d'une longueur  $x$  égale à  $2l$ . Cette longueur  $x$  est appelée retard ou déphasage. Le principe général est que ces deux faisceaux se recombinent, mais vu les différences de chemins optiques parcourus, il peut se créer des interférences constructives et destructives formant l'interférogramme. En effet, on observera l'apparition d'interférences constructives, si  $x$  est un multiple entier des nombres d'ondes ( $\sigma$ ), soit,  $x = \sigma, 2\sigma, 3\sigma, \dots, n\sigma$  ou des interférences destructives si  $x = (n + 1/2)\sigma$ . Ensuite, le faisceau passe à travers l'échantillon à mesurer et est recueilli par un détecteur qui mesure la variation de l'énergie en fonction du temps pour tous les nombres d'ondes simultanément. L'interférogramme, qui représente l'intensité en fonction du temps (ou distance), est ensuite converti en un spectre MIR, représenté par l'intensité en fonction du nombre d'onde, par une transformation mathématique appelée transformée de Fourier, d'où la dénomination « Infrarouge à transformée de Fourier ». Ce procédé technologique nécessite l'utilisation d'un ordinateur puissant nécessaire pour effectuer les calculs de transformation de Fourier de façon instantanée. Comme nous le verrons dans une section ultérieure, c'est d'ailleurs à la fin des années 1980, que les premiers spectromètres FT-MIR ont réellement été proposés à la commercialisation dans les laboratoires de routine, peu après l'apparition des premiers ordinateurs personnels.



**Figure 1.6.** Principe d'un spectromètre à transformée de Fourier (d'après Bertrand et Dufour, 2000).

### 1.1.3. Utilisation de la FT-MIR dans le secteur laitier

#### 1.1.3.1 Contrôle laitier, paiement du lait, et utilisation en laiterie

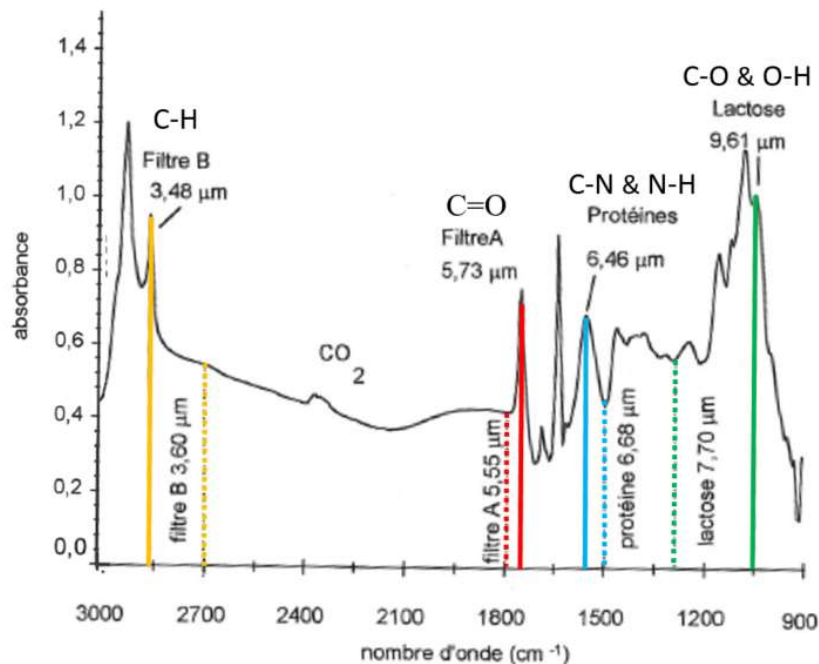
Les techniques rapides reposant sur la spectrométrie MIR se sont rapidement répandues dans l'ensemble des laboratoires effectuant les méthodes d'analyse pour le secteur laitier afin de quantifier les taux en composants majeurs du lait nécessaires à la fixation du prix du lait, au contrôle qualité, aux calculs des bilans des usines, et au contrôle des performances des vaches laitières lié au contrôle laitier. Alors que les méthodes analytiques de référence sont généralement longues et complexes à réaliser, coûteuses en produits et en main d'œuvre, nécessitant souvent l'emploi de réactifs polluants et ne pouvant être réalisées que par un personnel qualifié, les méthodes MIR ont l'avantage non négligeable, d'être rapides et non destructives, évitant ainsi l'utilisation de plusieurs échantillons pour réaliser les analyses de différents paramètres.

Le premier appareil MIR commercialisé pour les analyses laitières fut l'appareil InfraRed Milk Analyzer (IRMA), développé par Grubb Parson, Newcastle upon Tyne, au Royaume-Uni, d'après les travaux initiés par Goulden en 1958 (Adda et al., 1968).

En moins d'une minute, cet appareil à transmission pouvait prédire quantitativement la MG, les protéines ou MAT et le lactose, en utilisant respectivement les mesures des intensités des pics d'absorption de trois nombres d'onde dits « primaires » : 1.745, 1.548, 1.042  $\text{cm}^{-1}$  (ou 5,73, 6,46 et 9,60  $\mu\text{m}$ ). Ajoutons que pour ces appareils conventionnels à filtres, les pics d'absorption fournis pour les 3 nombres d'onde considérés, sont obtenus en calculant pour chacun des 3 nombres d'onde le logarithme du rapport entre les valeurs enregistrées pour le lait et celles enregistrées pour l'eau. En effet, l'absorption d'eau est très forte dans la plage de nombres d'ondes de 1.000 à 5.000  $\text{cm}^{-1}$  et les bandes d'absorption des autres composants vont se superposer à celle de l'eau, ce qui rend les mesures compliquées. Car, l'eau étant le constituant majeur du lait, elle contribue largement à l'absorption infrarouge. Pour pallier ce problème, deux cellules étaient utilisées par le spectromètre à double faisceau de type IRMA, l'une contenant de l'eau (référence) et l'autre l'échantillon de lait, permettant alors de soustraire automatiquement l'absorption liée à l'eau (McGann, 1978). Le MilkoScan 300 (Foss Electric, Hillerød, Danemark) commercialisé dans les années 70, va solutionner le problème différemment (Grappin and Jeunet, 1976). Pour résoudre ce problème lié à l'eau, le MilkoScan va utiliser une seule cellule (Biggs, 1978). Deux mesures d'absorbance vont être réalisées à deux nombres d'ondes. La première correspond au maximum d'absorption de la liaison et la seconde, qui reste proche de la première, est communément appelée référence, et correspond à un minimum d'absorption de la liaison (voir Figure 1.7). Le MilkoScan utilisait donc en réalité deux filtres différents pour chaque composant à mesurer, avec un nombre d'onde spécifique à une liaison moléculaire du composant (E) et un nombre d'onde de référence (R). Par la suite, la majorité des instruments à base de filtres vont utiliser cette technique de cellule unique et utiliser 3 filtres de référence distincts avec des nombres d'onde aux alentours de 1.802, 1.497, 1.299  $\text{cm}^{-1}$  (ou 5,55 ; 6,68 et 7,70  $\mu\text{m}$ ), correspondant respectivement aux filtres de référence MG, MAT et lactose (Lynch et al., 2006). L'avantage de cette méthode réside dans le fait qu'en utilisant une seule cellule, on élimine les sources de variations liées à l'utilisation de deux cellules (McGann, 1978). La détermination de l'absorbance d'un composant à un nombre d'onde donné résulte dans un premier temps du calcul du rapport R/E à la fois pour le lait ( $R_L/E_L=L$ ) et pour l'eau ( $R_O/E_O=Z$ ) pris comme zéro pour l'analyse, puis à la transformation logarithmique du rapport L/Z (Bertrand et Dufour, 2000).

Ces deux instruments étaient déjà équipés d'un homogénéisateur permettant de diminuer et uniformiser la taille des globules gras. En effet, les tailles des globules gras natifs dans le lait cru peuvent varier entre 1 et 20  $\mu\text{m}$  (Pritchard and Kailasapathy, 2011) selon des facteurs biologiques liés par exemple à la race, l'âge, ou la santé de l'animal (Postelmans et al, 2020). Un homogénéisateur est essentiel pour limiter l'impact de deux effets néfastes majeurs qui peuvent se produire lors de l'analyse de lait contenant de gros globules gras. Premièrement, la présence de gros globules gras entraîne une plus grande diffusion de la lumière, qui est interprétée à tort par le détecteur comme une absorption de lumière. En effet, la quantité de diffusion de la lumière dépend de la longueur d'onde et est influencée par la taille des particules

(Postelmans et al, 2020). Or, la diffusion de la lumière devient plus prononcée lorsque le diamètre des globules gras s'approche de la taille de la longueur d'onde mesurée (Guillou et al., 1986).



**Figure 1.7.** Spectre moyen infrarouge en transmission d'un échantillon de lait homogénéisé par rapport à l'eau. Les traits pleins de couleurs indiquent les nombres d'onde des filtres de mesure de la matière grasse (Fat- B en orange et Fat-A en rouge), des protéines (en bleu) et du lactose (vert). Les traits en pointillés indiquent les nombres d'onde des filtres de référence pour la matière grasse (orange et rouge pour Fat-B et A), pour la protéine (bleu) et pour le lactose (vert) (d'après Bertrand et Dufour, 2000).

Un homogénéisateur est essentiel pour réduire le diamètre des globules gras à une taille inférieure à 2  $\mu\text{m}$  pour éviter ce problème (Smith et al., 1993). Notons que la diffusion due aux micelles de caséines, vu leur taille (moins de 0,1  $\mu\text{m}$ ), est par contre négligeable aux longueurs d'onde utilisées dans le MIR (Goulden, 1964). L'effet Christiansen est le deuxième effet néfaste lié à la taille importante des globules gras sur la mesure de la teneur en MG du lait par les appareils infrarouges. C'est Goulden (1964) qui a pu mettre en évidence l'existence de cet effet dans le lait par l'observation d'une diminution de la longueur d'onde du pic d'absorption après l'utilisation d'une homogénéisation efficace (Smith et al., 1993). Normalement, pour la plupart des longueurs d'onde, l'indice de réfraction de la MG du lait est relativement constant à une température donnée. Cependant, aux longueurs d'onde proches de l'absorption maximale qui coïncident avec les groupes carbonyle (C=O) à 5,73  $\mu\text{m}$  et carbone-

hydrogène (C-H) à 3,48  $\mu\text{m}$  de la MG du lait, l'indice de réfraction de la MG du lait peut changer brusquement, alors que l'indice de réfraction du plasma de lait reste quant à lui constant. En raison de ce changement d'indice de réfraction au niveau du pic d'absorption, la longueur d'onde apparente d'absorption maximale de la lumière se déplace vers une longueur d'onde plus élevée (Smith et al., 1993). Ce décalage des longueurs d'onde peut avoir un effet négatif sur la précision de la détermination de la concentration de la MG du lait (Di Marzo and Barbano, 2016). En effet, le résultat net de l'effet Christiansen est une interprétation du signal, et donc une estimation de la teneur totale en MG, qui peut varier en fonction de la taille des globules gras, alors que la teneur totale en MG reste la même. L'effet Christiansen est minimisé en réduisant le diamètre des globules gras à moins d'un tiers de la longueur d'onde à laquelle une mesure est effectuée (Smith et al., 1993).

Outre, ces aspects liés à l'homogénéisation, Biggs (1967) soulignait que d'autres facteurs pouvaient affecter la précision de mesure de l'appareil IRMA : le système de purge de la cellule de mesure (effet de mélange), la température de l'appareil et de l'échantillon, la précision des nombres d'onde et les interférences liées aux variations de la composition du lait dues à la saison, la région, le troupeau ou encore l'individu. En particulier, le manque de précision de l'IRMA pour la quantification de la MG était expliqué par Grappin et Jeunet (1972) comme étant lié à la variation des proportions des longueurs de chaîne des acides gras. Si un lait présente une proportion plus importante en acides gras à longues chaînes, il y a, de surcroît, une diminution des acides gras à courtes et moyennes chaînes et par conséquent une diminution du nombre de liaisons ester par unité de poids. Cette modification se répercutait par une sous-estimation de la MG par l'IRMA par rapport à la méthode de référence pour une MG riche en acides gras longs et de facto par une sur-estimation de la MG pour des laits pauvres en acides gras longs. L'augmentation du nombre d'étalonnages lors des changements de saison pouvait alors aider à partiellement résoudre ce problème. Par contre, les facteurs liés à la région, au troupeau et à l'individu restaient plus compliqués à gérer, surtout dans les laboratoires des contrôles laitiers traitant des échantillons provenant parfois de plusieurs régions éloignées. Malgré cela, grâce à sa rapidité d'analyse, sa rentabilité dix fois supérieure (Biggs, 1967) et sa capacité à fournir des mesures utiles au secteur laitier (MG, MAT et lactose), la spectrométrie MIR est devenue une méthode reconnue par l'Association of Official Agricultural Chemists (AOAC) pour les mesures quantitatives du lait (Biggs, 1972) et a permis à cette technique de se propager rapidement dans les laboratoires laitiers. Elle a fini par remplacer les anciennes techniques coûteuses, non polyvalentes, et plus lentes. Par la suite, et afin de calibrer ces premiers spectromètres MIR, l'AOAC a publié dès 1975 les premières recommandations relatives à la maintenance et aux actions requises pour le maintien des calibrations. La procédure de calibration proposait d'utiliser 8 échantillons de lait provenant de tank à lait de ferme, présentant une large variabilité des taux en MG (au min 2 %) et en MAT (min. de 1 %). La pente et biais étaient alors corrigés par comparaison avec les mesures obtenues via les méthodes chimiques de référence (Randall, 1978)

Vers 1995, de nouveaux appareils automatisés à haute fréquence d'analyses ont été proposés (type modèle MilkoScan 4000 de chez FOSS Electric, Hillerød, Danemark). Leur conception repose sur l'utilisation d'un système infrarouge mono faisceau comprenant une roue à filtres, composée d'une dizaine de filtres, ce qui permet de mesurer plusieurs composés de façon quasi simultanée. Pour la MG, MAT et lactose, les modèles mathématiques de prédiction utilisaient des régressions linéaires multiples (en anglais Multiple Linear Regression : MLR) apparus quelques années plus tôt (Biggs et McKenna, 1989). Ainsi, généralement quatre filtres ( $2.874$ ,  $1.745$ ,  $1.548$ , et  $1.042$   $\text{cm}^{-1}$ ), appariés à leurs quatre filtres de référence ( $2.778$ ,  $1.802$ ,  $1.497$  et  $1.299$   $\text{cm}^{-1}$ ), sont utilisés (Lynch et al., 2006) pour prédire les taux de MG, MAT et lactose par le Milkoscan 4000 utilisant des modèles MLR (Figure 1.7). On notera l'apparition d'un filtre supplémentaire à  $2.874$   $\text{cm}^{-1}$ . En effet, afin de résoudre le problème lié à la variation des proportions des longueurs de chaînes des acides gras lors de la détermination de la MG, Biggs et McKenna (1989) proposent d'utiliser des modèles de prédiction pour la MG ne prenant plus seulement en considération l'absorbance au nombre d'onde A de  $1.745$   $\text{cm}^{-1}$  ( $5,73$   $\mu\text{m}$ ) qui mesure plus spécifiquement les étirements des liaisons C=O, mais également le nouveau nombre d'onde B à  $2.874$   $\text{cm}^{-1}$  ( $3,48$   $\mu\text{m}$ ). Un brevet publié en 1981 par Nexo and Frandsen et possédé par FOSS décrit l'utilisation de ce nombre d'onde spécifique additionnel, qui correspond aux étirements des liaisons C-H de la MG. Comme on a vu plus haut, un modèle de prédiction de la MG à partir d'un modèle utilisant la longueur d'onde A (communément appelé modèle « Fat-A ») ne peut pas différencier et prendre en compte les variations de composition dues à la longueur de la chaîne des acides gras de la MG des laits. En utilisant le modèle Fat-A, on en vient donc à supposer que les MG de tous les laits ont la même longueur de chaîne (Biggs et McKenna, 1989). L'utilisation d'un modèle utilisant la longueur d'onde B (communément appelé modèle « Fat-B ») permet de compenser les différences de longueur de chaîne des acides gras. Par contre, cela revient dans ce cas, à supposer que les MG de tous les laits ont la même quantité de liaisons insaturées. Or, ceci n'est pas vrai en réalité. En effet, de même que pour la longueur des chaînes des acides gras, les laits présentent également des variations au niveau du nombre de liaisons insaturées (Kaylegian et al., 2009a). Cela entraînera des erreurs dans la prédiction de la MG B du lait. Des modèles appelés « approche à filtres fixes » (en anglais « fixed-filter approach ») ont été proposés pour utiliser un mélange des deux MG prédites (Fat-A et Fat-B) afin d'améliorer la précision des modèles. Cependant, le rapport à utiliser entre Fat-A et Fat-B n'a pas rencontré de consensus (Biggs et McKenna, 1989, Kaylegian et al., 2009a).

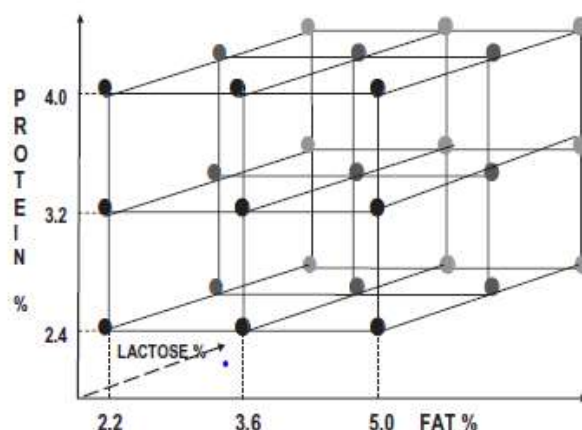
Comme souligné par Barbano et Clark (1989), le passage des techniques traditionnelles reposant sur la chimie vers la spectrométrie MIR va nécessiter l'adoption de nouvelles instructions. Il est nécessaire de vérifier et régler tous les facteurs instrumentaux pouvant modifier la réponse de l'appareil : zéro optique, linéarité, efficacité de l'homogénéisation, purge de la cellule (carry-over), stabilité et répétabilité, avant de procéder au calibrage de l'appareil. Ces éléments sont repris

pour la première fois dans un guide publié par la FIL sur les opérations liées à l'utilisation des appareils MIR (FIL/IDF, 1988). L'étalonnage des MIR est une opération cruciale car elle va déterminer la qualité des résultats analytiques fournis par les laboratoires laitiers. Cependant, la complexité et la variabilité de la composition du lait rend cette opération compliquée (Leray, 2020). Pour obtenir des résultats qui soient corrigés pour les interférences issues de différents composants (MG, MAT et lactose) et de l'influence moyenne de la matrice, il faut établir des équations MLR à partir des signaux primaires, obtenus après les analyses d'échantillons de lait de référence dont les compositions sont connues. Afin d'établir des corrections des interférences (ou inter-corrections) précises (ISO 8196-2:2009), il est nécessaire de disposer à la fois d'échantillons de lait qui couvrent l'étendue normale des concentrations mais également pour lesquels les corrélations entre les teneurs des éléments à mesurer sont faibles ou nulles (échantillons décorrélés en MG, MAT et lactose). Une méthode consistant à utiliser des échantillons de lait recombinaés à partir de la crème, du perméat et du rétentat d'un lait d'origine et d'ajouts de lactose pure cristallisé ou d'eau déminéralisée (voir Figure 1.8) fut proposée suite aux travaux de Leray (2020). Elle permet de vérifier la linéarité de chacun des modèles des composants majeurs du lait et d'étalonner de façon standardisée les appareils MIR utilisés pour les contrôles laitiers et/ou par les laboratoires effectuant les analyses pour le paiement du lait. Cette méthode a été proposée par le sous-comité analytique de l'International Committee for Animal Recording (ICAR), soit, en français, le Comité international pour le contrôle des performances en élevage (Leray, 1990) et par la FIL (Leray and Grappin, 1990). Le détail complet de la méthode fut décrit pour la première fois en annexe de la norme internationale publiée par l'Organisation internationale de normalisation ou International Organization for Standardization (ISO, ISO 9622 : 1999). Elle est mieux connue sous le vocable FIL 141 par les utilisateurs.

Le Galaxy 600, développé par la société française AEGYS (ANADIS) au cours des années 1989-1991, est le premier analyseur automatique destiné à déterminer en routine la composition chimique du lait, par le biais d'une technologie nouvelle : la spectrométrie infrarouge à transformée de Fourier dans le moyen infrarouge (FT-MIR). Plusieurs auteurs (Van de Voort et al., 1992 ; Lefier et al., 1996) ont rapidement souligné l'intérêt à utiliser ce type de spectrométrie. En effet, étant donné que l'instrument FT-MIR fournit plus d'informations spectrales liées à la composition du lait que l'instrument à filtre, les analyses FT-MIR à étalonnage unique d'échantillons de lait collectés à différentes saisons sont plus précises que les méthodes à appareils à filtres (Lefier et al., 1996).

De quelques mesures de points spectraux, on est alors passé à quelques milliers avec certains points spectraux qui présentent de forts effets de colinéarité avec notamment les points voisins. Par conséquent, la méthode MLR n'était plus appropriée pour utiliser l'entièreté de cette information disponible. Ainsi, l'algorithme de régression le plus utilisé pour construire les équations de calibration des appareils FT-MIR à partir du spectre MIR, est devenu la régression des moindres carrés partiels, encore appelée PLS (de l'anglais « Partial Least Square »). L'utilisation de la PLS,

s'inspirant directement de la spectrométrie proche infrarouge (NIR), a été proposée dès la fin des années 80 (Beccard, 1987 ; Fuller et al., 1988). Les premiers modèles PLS ont été proposés dès 1992 pour prédire les composés majeurs du lait (Luinge et al., 1991). Pourtant, certains pays ont fait le choix de continuer à n'utiliser que les méthodes MLR pour prédire les composés majeurs du lait en utilisant que les longueurs d'onde traditionnellement utilisées par les appareils à filtres. C'est le cas notamment de la France, qui en l'absence de différences significatives entre les performances obtenues par les méthodes MLR et PLS, a continué à préférer l'utilisation de ses propres modèles MLR afin de garder son indépendance par rapport aux modèles PLS proposés par les fournisseurs et fabricants de machine. Cela implique de devoir continuer à organiser des étalonnages et réajustements périodiques plus fréquents et contrôlés parfaitement et à continuer de calculer les facteurs d'inter-corréction (les modèles PLS ne le nécessitant plus). Malgré ces rares exceptions, l'utilisation de la PLS s'est généralisée à partir des années 2000 pour la majorité des composés prédits y compris les composés majeurs, suite notamment à son acceptation comme méthode officielle par la FIL et ICAR (Horwitz, 2000 ; FIL 2000). En Belgique, les modèles PLS sont utilisés tant pour les analyses du paiement du lait que pour le contrôle laitier depuis avril 1998.



**Figure 1.8.** Représentation schématique des différents niveaux de concentration de matières grasses (FAT), matières azotées (PROTEIN) et lactose obtenus par la méthode de séparation/recombinaison de lait permettant la création d'échantillons recombinaison pour la calibration et le contrôle qualité des analyses par spectrométrie moyen infrarouge (d'après Leray, 2017).

### 1.1.3.2. Un nouveau potentiel d'utilisation de la FT-MIR

Rapidement, plusieurs études vont mettre en évidence l'intérêt de l'utilisation de l'information spectrale et le potentiel important de l'utilisation de la FT-MIR pour



prédire des paramètres autres que les traditionnels composés majeurs du lait (de Marchi et al., 2014). On sort donc de la dynamique d'amélioration des seules équations de prédiction des composés majeurs qui a perduré tout au long du vingtième siècle (Figure 1.9). Ce changement de paradigme va alimenter un nombre important de nouvelles recherches dans domaine du MIR à partir du spectre de lait, dès la fin du deuxième millénaire. En effet, actuellement, on ne se focalise plus uniquement sur les seuls composés majeurs qui fixent la valeur économique du lait, soit la « valeur argent », pour répondre aux objectifs anciens définis par le premier congrès de laiterie belge de 1897 (De Bruyn and Peers, 1898). Sur base du FT-MIR, on cherche d'une part, à fournir de nouveaux phénotypes utiles pour aider à accroître la rentabilité et la résilience de l'ensemble de la filière laitière. D'autre part, même si les vitesses de séquençages du génome ont été considérablement accélérées ces dernières années, les généticiens restent limités dans leur compréhension de l'interaction génotype-phénotype par l'accès à des données phénotypiques fiables et abondantes. Car, les coûts associés aux mesures de différents paramètres restent encore trop élevés, ce qui tend à freiner les vitesses et volumes d'acquisition de ces données phénotypiques (Juárez et al., 2018). Ce manque d'informations phénotypiques est devenu un grand défi pour la recherche génomique. Cela a conduit Coffey à dire : « à l'ère de la génomique, le phénotype est roi ! » (Coffey, 2011). La FT-MIR est un outil puissant pour répondre à ce besoin de la génomique.

A titre d'exemples, on a pu démontrer qu'il était possible de prédire à partir du spectre MIR du lait : l'urée (Hansen, 1998), les acides gras individuels ainsi que les groupes d'acides gras (Soyeurt et al., 2006 et 2011, Rutten et al., 2009), la longueur et le degré d'insaturation des chaînes d'acides gras (Wojciechowski and Barbano, 2016), la lactoferrine comme indicateur potentiel des mammites (Soyeurt et al., 2012), les minéraux majeurs (Soyeurt et al., 2009 ; Christophe et al., 2021), la composition de la protéine et ses variants (Bonfatti et al., 2011), des biomarqueurs d'intérêt pour évaluer la santé des animaux, tels les taux de beta-hydroxybutyrate (BHB), de citrate ou d'acétone du lait (van Kneegsel et al., 2010 ; Grelet et al., 2016), de composés sanguins liés à l'état métabolique de la vache (Grelet et al., 2019), les propriétés technologiques du lait (de Marchi et al., 2009), l'aptitude à la coagulation (Dal Zotto et al., 2008 ; de Marchi et al., 2013), l'aptitude à la transformation en beurre (Lefébure et al. 2021), la taille de globules gras dans du lait homogénéisé (Di Marzo et al., 2016), la couleur du lait (Scarso et al., 2017), l'état de gestation (Lainé et al., 2014), le statut énergétique corporel (McParland et al., 2011), l'apport et l'efficacité énergétique (McParland et al., 2014), l'efficacité d'utilisation de l'azote (Grelet et al., 2020), l'ingestion de matière sèche (Tedde et al., 2021a), le poids corporel des vaches (Tedde et al., 2021b), le stress chez les animaux à partir des modifications de la composition du lait (Bahadi et al., 2021), la détection de contaminants dans le lait (Fernández Pierna et al., 2016), la présence d'adultérations ou de laits anormaux (Spieß et al., 2021), la prédiction d'indicateurs de la composition de la rations ingérées par les vaches (par ex. Coppa et al., 2021 ; Soyeurt et al., 2022) ou encore la présence de maladies métaboliques (Franceschini et al., 2022).

### ***1.1.3.3. De nouveaux défis analytiques***

En règle générale, il est important que le spectre d'un échantillon à prédire soit inclus dans la variabilité spectrale du modèle utilisé, car, les extrapolations sont dangereuses en spectrométrie infrarouge (Dardenne, 2010). Traditionnellement, pour les modèles qui prédisent les composants majeurs du lait, les absorbances mesurées sont linéaires et directement en lien avec leurs concentrations, comme le stipule la loi de Beer-Lambert. Ceci présente un avantage au niveau de la robustesse par rapport aux nouvelles équations cherchant à prédire des éléments mineurs ou d'autres équations indirectes pour lesquelles les absorbances utilisées ne sont pas en lien direct avec le paramètre prédit mais qui utilisent les intercorrélations existantes. Si l'on veut utiliser ce type de modèles PLS sur de larges bases de données, on doit d'avantage veiller à s'assurer que la variabilité spectrale du modèle couvre un maximum de la variabilité de la population. En effet, le modèle ne peut prédire en dehors de son espace et doit inclure un maximum de la variabilité pour assurer sa robustesse. On dira d'une équation qu'elle est robuste, si elle est capable de générer des prédictions précises dans diverses conditions, en mode « tout-terrain » (Grelet et al., 2021) ou encore, que la robustesse d'un modèle d'étalonnage multivarié est sa capacité à prédire de façon stable (statistiquement), face à des perturbations définies en conditions standard (Zeaiter et al, 2004).

## **1.2. Objectifs de la thèse**

Comme nous venons de le voir, l'analyse du lait par la spectrométrie FT-MIR offre l'opportunité de prédire une large gamme de nouveaux phénotypes, de permettre le développement d'outils visant à améliorer de multiples façons la rentabilité et la pérennité de la filière lait (Gengler et al., 2015 ; Franceschini et al., 2022) et d'augmenter les possibilités au niveau des évaluations génétiques et génomiques (de Marchi et al., 2014 ; Gengler et al., 2016).

Pour une grande majorité de ces nouveaux phénotypes, les solutions apportées pour assurer la robustesse et le maintien de la qualité des modèles de prédiction dans les laboratoires de routine ne sont pas valables ou difficilement applicables. En effet, les normes ISO 9622:1999 et ISO 8196-2:2009, développées à l'origine pour les modèles des composés majeurs du lait, que sont la MG, la MAT et le lactose, ne peuvent continuer à fonctionner que pour des modèles de paramètres directement mesurables dans le lait et en quantité suffisante (c'est-à-dire supérieure au seuil de détection). Ainsi, la création de modèles indirectes, utilisant les intercorrélations existantes entre la composition fine du lait et le phénotype à prédire (p.ex. le méthane éructé, l'efficacité azotée ou encore le statut de bien-être des animaux), nous ont amené à nous poser des questions sur les éléments clés permettant d'assurer leur qualité lors de leur création ou de leur mise en routine.

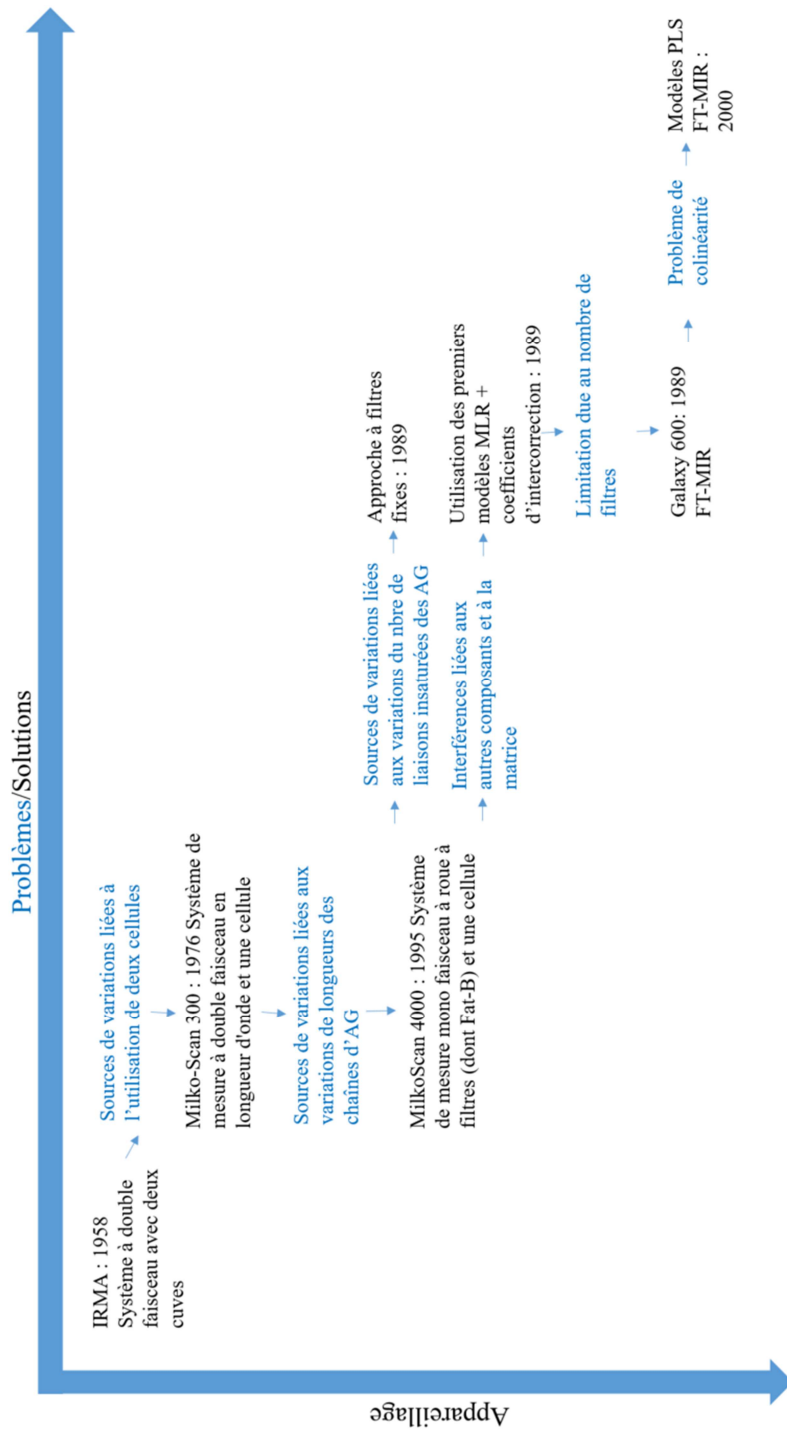


Figure 1.9. Synthèse des différentes solutions et améliorations techniques visant à améliorer les prédictions MIR des composés majeurs du lait.

Pour apprécier la qualité d'un modèle, Esbensen and Paasch-Mortensen (2010) ont défini l'erreur globale d'estimation (GEE) comme étant la somme de deux composantes : l'erreur total de l'échantillonnage (TSE) et l'erreur analytique globale (TAE). Ces deux composantes sont essentielles pour bien comprendre l'ensemble des facteurs qui vont influencer la qualité des modèles. Par conséquent, les premières questions qui seront solutionner dans le présent document de thèse sont en lien avec la TSE qui va directement influencer la robustesse des modèles. Plus spécifiquement, nous répondrons aux questions suivantes :

- Comment peut-on maximiser la variabilité spectrale des nouveaux modèles ?
- A quel moment doit-on réaliser les échantillonnages pour mesurer les spectres et les mesures phénotypiques ?
- Comment doit-on réaliser les mesures spectrales et phénotypiques ?

Une deuxième partie des questions qui seront soulevées dans la présente thèse sont en lien avec la composante TAE à savoir :

- Quelle analyse de référence choisir ?
- Quelle méthode dois-je employer pour créer un nouveau modèle prédictif ?
- Est-il nécessaire de normaliser les données ?
- Quel est l'intérêt à travailler dans des réseaux internationaux ?
- Comment assurer le contrôle qualité du nouveau modèle ou de ses prédictions en routine ?

Même si la réponse à certaines de ces questions peut paraître triviale, l'expérience acquise pendant ces deux dernières décennies au sein du laboratoire laitier du CRA-W a montré que ces réponses variaient considérablement en fonction du phénotype étudié et plus particulièrement si ce dernier est issu directement du lait ou pas. La présente thèse a donc pour objectif principal de faire la synthèse des recommandations et des pratiques qui permettent de maximiser la robustesse et le maintien de la qualité des modèles FT-MIR du lait et de leurs prédictions en routine.

Ainsi, cette thèse montrera deux stratégies différentes d'échantillonnage visant à augmenter la variabilité spectrale et phénotypique selon que le caractère considéré soit issu directement du lait comme pour la prédiction en acides gras (Chapitre 2) ou présente un lien indirect avec la composition du lait comme la prédiction du méthane (Chapitre 3). Ensuite, la thèse insistera sur l'importance de la phase de développement des équations pour assurer la qualité de la prédiction en mettant en évidence les points clés à considérer (Chapitre 4). Finalement, la discussion générale (Chapitre 5) permettra d'élargir le concept à la mise en place au sein d'un laboratoire laitier d'une politique d'assurance qualité pour les prédictions FT-MIR qu'elles soient directement liées à la composition du lait ou pas et les défis que cela représente. Ce document de thèse se terminera par la conclusion de ce travail et les perspectives pour les années futures.

### 1.3 Références

- Adda J., Blanc-Patin E., Jeunet R., Grappin R., Mocquot G., Pousardieu B. and Ricordeau G., 1968. Essais d'utilisation de l'infrarouge Milk Analyser. *Le Lait*, INRA Editions, 1968, 48 (473\_474), pp.145-154.
- Bahadi M., Ismail A.A., and Vasseur E., 2021. Fourier Transform Infrared Spectroscopy as a Tool to Study Milk Composition Changes in Dairy Cows Attributed to Housing Modifications to Improve Animal Welfare. *Foods* 10 (2):450.
- Barbano D.M. and Clark J.L., 1989. Infrared milk analysis: Challenges for the future. *J. Dairy Sci.* 72: 1627-1636.
- Beccard B., 1987. Analyse quantitative de données en spectrophotométrie infrarouge par la méthode des moindres carrés. *Analisis*, 15, pp. 25-30.
- Bertrand D. and Dufour E., 2000. La spectroscopie infrarouge et ses applications analytiques. Lavoisier, 566 p., Sciences et Techniques Agroalimentaires, 2-7430-0375-8.
- Biggs D.A., 1967. Milk composition: milk analysis with the infrared milk analyzer. *J. Dairy Sci.* v.50, p. 799-803.
- Biggs D.A., 1972. Precision and accuracy of infrared milk analysis. *J. AOAC Int.* v.55 p.488-497.
- Biggs D.A., 1978. Instrumental infrared estimation of fat, protein, and lactose in milk: collaborative study. *J Assoc Off Anal Chem.*, 61(5):1015-34.
- Biggs D. A. and McKenna D., 1989. Alternative methods for infrared analysis of fat in milk: Interlaboratory study. *J. AOAC*, 72:724–734.
- 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., 2017. Standardization of milk infrared spectra for the retroactive application of calibration models. *J. Dairy Sci.*, 100:2032–2041.
- CBL, 2004. Rapport annuel 2004. Confédération belge de l'industrie laitière. pp 75.

- CBL, 2022. Rapport annuel 2022. Confédération belge de l'industrie laitière. pp 60.
- CEE, Communauté Economique Européenne, 1965. L'organisation du marché du lait de consommation dans les Etats membres de la CEE. Collections études. Série Agriculture n°20, Bruxelles. 8147\*/1/VIII/1966/5. pp 55.
- Cernuda C., Lughofer E., Klein H. Forster C., Pawliczek M. and Brandstetter M., 2017. Improved quantification of important beer quality parameters based on nonlinear calibration methods applied to FT-MIR spectra. *Anal Bioanal Chem* 409, 841–857
- Christophe O.S., Grelet C., Bertozzi C., Veselko D., Lecomte C., Höeckels P., Werner A., Auer F.-J., Gengler, N., Dehareng F. and Soyeurt H., 2021. Multiple Breeds and Countries' Predictions of Mineral Contents in Milk from Milk Mid-Infrared Spectrometry. *Foods*, 10, 2235.
- Coates J., 2000. Interpretation of infrared spectre, a practical approach R.A. Meyers (Ed.), *Encyclopedia of Analytical Chemistry*, John Wiley & Sons, New York, NY, pp. 10815-10837.
- Coffey M., 2011. Phenotyping that maximises the value of genotyping. In 'General assembly and annual workshop of ICAR 2011. New technologies and new challenges for breeding and herd management'. Bourg-en-Bresse, France, 22–24 June 2011.
- Coppa M., Martin B., Hulin S., Guillemin J., Gauzentes J.V., Pecou A. and Andueza D., 2021. Prediction of indicators of cow diet composition and authentication of feeding specifications of Protected Designation of Origin cheese using mid-infrared spectroscopy on milk. *J. Dairy Sci.*, 104:112–125.
- Dal Zotto R., de Marchi M., Cecchinato A., Penasa M., Cassandro M., Carnier P., Gallo L., 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.
- Dardenne P., 2010. Some considerations about NIR spectroscopy: Closing speech at NIR-2009. *NIR news.*, 21(1):8-14.
- De Baere J., 1978. Un Siècle d'activité Laitière En Belgique. Société nationale de laiterie. *Courrier hebdomadaire du CRISP*, n°796-797, pages 1 à 45.
- De Bruyne L. and Peers L., 1898. Actes constitutifs, liste de membres, rapports et discussions. Congrès National de Laiterie. 16-17 octobre 1897, Bruxelles, Belgium, pp. 318.

- De Herde V., 2020. Narratif historique : évolution des laiteries coopératives wallonnes de 1948 à nos jours, travail d’investigation réalisé dans le cadre du doctorat en sciences agronomiques, avec le soutien du FRIA-FNRS, pp. 140.
- 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(10), 1694-1701.
- de Marchi M., Fagan C.C., Odonnell C.P., Cecchinato A., Dal Zotto R., Cassandro M., Penasa M. and Bittante G., 2009. Prediction of coagulation properties, titratable acidity, and pH of bovine milk using mid-infrared spectroscopy. *J. Dairy Sci.*, 92, 423–432.
- de Marchi M., Toffanin V., Cassandro M. and Penasa M., 2013. Prediction of coagulating and noncoagulating milk samples using mid-infrared spectroscopy. *J. Dairy Sci.*, 96, 4707–4715.
- 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, pp. 1171-1186.
- Di Marzo L. and Barbano D.M., 2016. *J. Dairy Sci.* Volume 99, Issue 12, December 2016, pp. 9471-9482.
- Di Marzo L., Cree P. and Barbano D.M., 2016. Prediction of fat globule particle size in homogenized milk using Fourier transform mid-infrared spectra. *J. Dairy Sci.*, 99, 8549–8560.
- Esbensen K.H. and Paasch-Mortensen P., 2010. Process sampling: Theory of sampling – The missing link in process analytical technologies (PAT). In: *Process Analytical Technology*: 37-80.
- Fernández Pierna J.A., Vincke D., Baeten V., Grelet C., Dehareng F. and 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. *Chemometrics and Intelligent Laboratory Systems*, 152, pp. 157-162.
- FIL/IDF International Dairy Federation. 1988. Whole milk: determination of milkfat, protein, and lactose content – guide for the operation of mid-infrared instruments. IDF Standard 141. Int. Dairy Fed., Brussels, Belgium.

- Franceschini S., Grelet C., Leblois J., Gengler N. and Soyeurt H., 2022. Can unsupervised learning methods applied to milk recording big data provide new insights into dairy cow health? *J. Dairy Sci.*, 105, Issue 8, pp. 6760-6772.
- Fuller M. P., Ritter G. L., and Draper C. S., 1988. Partial Least-Squares Quantitative Analysis of Infrared Spectroscopic Data. Part I: Algorithm Implementation, *Appl. Spectrosc.*, 42, 217-227.
- 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. Proc. EAAP 66th Annual Meeting, Warsaw, Wageningen Academic Publishers, Wageningen, the Netherlands, p. 353.
- 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, pp. 1-9.
- Gengler N., 2022. Collaboration is imperial or how to address this critical issue when developing novel traits. Proc. EAAP 77th Annual Meeting, Porto, Wageningen Academic Publishers, Wageningen, the Netherlands, p. 329.
- Goulden J.D.S., 1961. Quantitative analysis of milk and other emulsions by infra-red absorption. *Nature*, 191, pp. 905-906.
- Goulden J.D.S., 1964. Analysis of milk by infra-red absorption. *Journal of Dairy Research* , Volume 31 , Issue 3, pp. 273 – 284.
- Grappin R., Jeunet R. and Rigogne R., 1972, Facteurs biologiques responsables des variations observées dans la précision des méthodes de routine utilisées pour le dosage de la matière grasse du lait (MilkoTester, Infra-Red-Milk Analyser et Gerber). *Le Lait*, INRA Editions, 52 (515-516), pp.324-346.
- 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. *Le Lait*, 56 (558), pp.498-520.



- Grelet C., Bastin C., Gelé M., Davière J.B., Johan M., Werner A., Reding R., Fernandez Pierna J.A., Colinet F.G., Dardenne P., Gengler N., Soyeurt H. and Dehareng F., 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, b-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. *J. Dairy Sci.*, 99, 4816–4825.
- Grelet C., 2019. Promoting international prediction models through standardization of milk mid-infrared spectra. PhD thesis, University of Liège, Gembloux Agro-Bio Tech Gembloux, Belgium. pp.110.
- 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 and 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., Foldager L., Salavati M., Hostens M., Ferris C.P., Ingvarsten K. L., Crowe M. A., Sorensen M. T., Fernandez Pierna J.A., Vanlierde A., Gengler N., and Dehareng F., 2020. Potential of milk mid-infrared spectra to predict nitrogen use efficiency of individual dairy cows in early lactation. *J. Dairy Sci.*, 103:4435–4445.
- Grelet, C., Dardenne, P., Soyeurt, H., Fernandez Pierna, J. A., Vanlierde, A., Stevens, F., Gengler, N., and Dehareng, F., 2021. Large-scale phenotyping in dairy sector using milk MIR spectra: Key factors affecting the quality of predictions. *Methods*, 186, 97-111.
- Griffiths P. R. and de Haseth J.A., 2007. *Fourier Transform Infrared Spectrometry*, 2e éd., p. 100, Wiley-Interscience, pp. 535.
- Guillou H., Pelissier J.P., and Grappin R.. 1986. Méthodes de dosage des protéines du lait de vache. *Le Lait*, 66 (2), pp.143-175.
- Hansen P.W., 1998. Urea determination in milk using Fourier transform infrared spectroscopy and multivariate calibration. *Milchwissenschaft-milk Science International* 53:251-255.
- Henseval M., 1904. Contribution à l'étude du dosage de la matière grasse du lait. *Revue générale du lait*, III, n°23.
- Horwitz W., 2000. *Official Methods of Analysis of AOAC International*; AOAC International: Gaithersburg, MD, USA.

- ISO 8196-2:2009 | IDF 128-2:2009. Milk – Definition and evaluation of the overall accuracy of alternative methods of milk analysis – Part 2: Calibration and quality control in the dairy laboratory. International Organization for Standardization, Geneva and International Dairy Federation, Brussels.
- ISO 9622:1999 | IDF 141:2000, Whole milk – Determination of milk fat, protein and lactose content – Guidance on the operation of mid-infrared instruments. International Organization for Standardization, Geneva and International Dairy Federation, Brussels.
- ISO 9622:2013 | IDF 141:2013. Milk and liquid milk products, Guidelines for the application of mid-infrared spectrometry. International Organization for Standardization, Geneva and International Dairy Federation, Brussels.
- Juárez M., Roberts J. C., López-Campos O., Prieto N. and Aalhuset J. L., 2018 Focusing on phenomics to unlock the potential of livestock genomics, CABI Reviews. CABI International.
- Kaylegian K. E., Dwyer D. A., Lynch J. M., Bauman D. E., Fleming J. R., and Barbano D. M., 2009a. Impact of fatty acid composition on the accuracy of mid-infrared fat analysis of farm milks. *J. Dairy Sci.*, 92:2502–2513.
- Kaylegian KE, Lynch JM, Fleming JR, Barbano DM, 2009b. Influence of fatty acid chain length and unsaturation on mid-infrared milk analysis. *J Dairy Sci.*, 92 (6):2485-501.
- 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, pp. 33-38.
- Laloux J., 2000. Cours laitier Législation, Confédération belge de l'industrie laitière.
- Lefébure E., Troch T., Noutfia Y., Colinet F., Gérard A., Dehareng F., Baeten V., Gengler N. and Sindic M., 2021. Is it possible to predict milk processing into butter using infrared spectroscopy? *Biotechnol. Agron. Soc. Environ.*, 25 (1), 21-31
- Lefebvre W. and Segers Y., 2003. “Industrialsiering Op Het Platteland. De Coöperatieve Zuivelnijverheid in de Brusselse Rand En Haar Periferie Tijdens Het Interbellum.” *Belgisch Tijdschrift Voor Nieuwste Geschiedenis*, XXXIII edition, sec. 3-4, pp. 485-533.

- Leray O., 1989. Ajustement/calcul des inter-corrrections des spectromètres utilisés pour les dosages TB-TP-TL du lait en moyen infra-rouge. Note Technique / Technical Note ITEB-INRA Poligny n°2, France.
- Leray O., 2017. Recombined milk samples for calibration and quality control in mid-infrared milk analysis. In: Quality assurance tools for Mid Infrared Spectrometry in dairy laboratories – Part 1. Bulletin of the International Dairy Federation, 490/2017.
- Luinge H. J. Hop E., Lutz E. T.G., van der Maas J. H., Holstra A., van Hemert H. A., Koops J., Wooldrik H., de Jong E. A.M. and Ellen G., 1992. Quantitative determination of the fat, protein, lactose, and water content in cow milk using Fourier transform infrared spectrometry, Proc. SPIE 1575, 8th Intl Conf on Fourier Transform Spectroscopy, (1 March 1992).
- Lynch JM, Barbano DM, Schweisthal M and Fleming JR., 2006. Precalibration evaluation procedures for mid-infrared milk analyzers. *J Dairy Sci.* Jul;89(7):2761-74. doi: 10.3168/jds.S0022-0302(06)72353-0. Erratum in: *J Dairy Sci.*, 2006 Aug;89(8):3274.
- McGann T.C.A., 1978. Automated Physico-Chemical Methods for the Analysis of Milk: A Review of Major Advances (1960-1978). *Irish Journal of Food Science and Technology*, Vol. 2, No. 2, pp. 141-155.
- 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, pp. 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, pp. 5863-5871.
- Nexo S.A. and Frandsen A.S., 1980. Method for quantitatively determining fat in a fat-containing sample. US Patent 4236075.
- Nguyen H.N., Dehareng F., Hammida M., Baeten V., Froidmont E., Soyeurt H., Niemöeller A., and 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.
- Norris, K., 2009. Hazards with near infrared spectrometry in detecting contamination. *Journal of Near Infrared Spectrometry*, 17, 165–166.

- Patz C.-D., Blicke A., Ristow R. and Dietrich H., 2004. Application of FT-MIR spectrometry in wine analysis, *Analytica Chimica Acta*, Volume 513, Issue 1, 81-89.
- Pritchard, S. R. and K. Kailasapathy. 2011. Chemical, Physical, and Functional Characteristics of Dairy Ingredients. Pages 35-57 in *Dairy Ingredients for Food Processing*. Wiley-Blackwell.
- Postelmans A., Aernouts B., Jordens J., Van Gerven T. and Saeys W., 2020. Milk homogenization monitoring: Fat globule size estimation from scattering spectra of milk, *Innovative Food Science & Emerging Technologies*, Volume 60, 102311.
- Pu Y.Y., O'Donnell C., Tobin J. and O'Shea N., 2020. Review of near-infrared spectroscopy as a process analytical technology for real-time product monitoring in dairy processing, *International Dairy Journal*, Volume 103, 104623.
- Randall S.Y., 1978. Calibration and Standardization of the Infrared Milk Analyzer. The California Experience. *J Dairy Sci.*, 61:1279-1283.
- Rouessac F., Rouessac A., Cruché D., Duverger-Arfuso C. and Martel A., 2016. *Analyse chimique 8ème édition : Méthodes et techniques instrumentales modernes*. Dunod. Paris, ISBN 987-2-10-074688-0.
- 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, pp. 6202-6209.
- Saldari, C., 1978. Le secteur laitier en Belgique. *Courrier hebdomadaire du CRISP*, 796-797, pages 1-45.
- Scarso S., McParland S., Visentin G., Berry D.P., McDermott A. and De Marchi M., 2017. Genetic and nongenetic factors associated with milk color in dairy cows, *J. Dairy Sci.*, Volume 100, Issue 9, pp. 7345-7361.
- Schutzler K., 1927. Analyse aux rayons ultra-violetes dans l'industrie laitière. *Milchwirtsch. Forsch. t.*, IV, p. 352.
- Seyewetz J., 1934. L'emploi de la lumière de Wood. In : *Les nouvelles méthodes d'analyses, dans les sciences, l'industrie, l'alimentation, la recherche des fraudes*. pp.251.

- Sinnaeve G., Dardenne P., Agneessens R., Lateur M. and Hallet A., 1997. Quantitative Analysis of Raw Apple Juices Using near Infrared, Fourier-Transform near Infrared and Fourier-Transform Infrared Instruments: A Comparison of Their Analytical Performances. *Journal of Near Infrared Spectroscopy*, 5(1):1-17.
- Smith E. B., Barbano D.M. and Lunch J.M., 1993. Performance of Homogenizers in Infrared Milk Analyzers: A Survey, *Journal of AOAC International*, Volume 76, Issue 5, Pages 1033–1041.
- Soyeurt H., Dardenne P., Dehareng F., Lognay G., Veselko D., Marlier M., Bertozzi C., Mayeres P. and Gengler N. 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. *J. Dairy Sci.*, 89, 3690–3695.
- 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.
- Soyeurt H., Bastin C., Colinet F.G., Arnould V.M., Berry D.P., Wall E., Dehareng F., Nguyen H.N., Dardenne P., Scheifers 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. *Anim. Int. J. Anim. Biosci.*, 6, 1830–1838.
- Soyeurt H., Grelet C., McParland S., Calmels M., Coffey M., Tedde A., Delhez P., Dehareng F. and Gengler N., 2020. A comparison of 4 different machine learning algorithms to predict lactoferrin content in bovine milk from mid-infrared spectra. *J. Dairy Sci.*, 103, Issue 12, 11585-11596.
- Soyeurt H, Gerards C, Nickmilder C, Bindelle J, Franceschini S, Dehareng F, Veselko D, Bertozzi C, Gengler N, Marvuglia A, Bayram A, Tedde A., 2022. Prediction of Indirect Indicators of a Grass-Based Diet by Milk Fourier Transform Mid-Infrared Spectroscopy to Assess the Feeding Typologies of Dairy Farms. *Animals*, 12(19):2663.
- Spieß L., de Peinder P., and van den Bijgaart H., 2021. Advances in Atypical FT-IR Milk Screening: Combining Untargeted Spectra Screening and Cluster Algorithms. *Foods*, 10, no. 5: 1111.

- Tedde A., Grelet C., Ho P.N., Pryce J.E., Hailemariam D., Wang Z., Plastow G., Gengler N., Froidmont E., Dehareng F., Bertozzi C., Crowe M.A., and Soyeurt H., on behalf of the GplusE Consortium. 2021a. Multiple Country Approach to Improve the Test-Day Prediction of Dairy Cows' Dry Matter Intake. *Animals*, 11, 1316.
- Tedde A., Grelet C., Ho P.N., Pryce J.E., Hailemariam D., Wang Z., Plastow G., Gengler N., Brostaux Y., Froidmont E., Dehareng F., Bertozzi C., Crowe M.A., Dufrasne I., GplusE Consortium Group, and Soyeurt H., 2021b. Validation of Dairy Cow Bodyweight Prediction Using Traits Easily Recorded by Dairy Herd Improvement Organizations and Its Potential Improvement Using Feature Selection Algorithms. *Animals*, 11(5):1288.
- Tiplady K. M., Sherlock R. G., Littlejohn M. D., Pryce J. E., Davis S. R., Garrick D. J., Spelman R. J., and Harris B. L.. 2019. Strategies for noise reduction and standardization of milk mid-infrared spectra from dairy cattle. *J. Dairy Sci.*, 102:6357–6372.
- Van der Bruggen M. and Peers L., 1903. Rapport du Congrès International de Laiterie, 8-11 septembre 1903, Bruxelles, Belgium, pp. 632.
- van Kneysel 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, pp. 3065-3069.
- Vanrobays M.-L., Vandenplas J., Bastin C., Hammami H., Soyeurt H., Vanlierde A., Dehareng F., Froidmont E. and Gengler N., 2015. Genetic correlations between methane production and milk fatty acid contents of Walloon Holstein cattle throughout the lactation. In Final OptiMIR Scientific and Expert Meeting: From milk analysis to advisory tools (Palais des Congrès, Namur, Belgium, April 16–17, 2015. *Biotechnol. Agron. Soc. Environ.*, 19, pp. 97-124.
- Volmar Y., 1927. Utilisation des phénomènes de fluorescence dans l'analyse des substances alimentaires. *Journ. Pharm. Chim.*, n°5, p.435.
- Williams, P.C. 2007. Near-infrared technology – Getting the best out of light. A short course in the practical implementation of Near-Infrared Spectroscopy for the user. PDK projects, Inc., Nanaimo, Canada.

Wojciechowski K.L. and Barbano D.M. 2016. Prediction of fatty acid chain length and unsaturation of milk fat by mid-infrared milk analysis. *J. Dairy Sci.*, 99, 8561–8570.

Zeaiter M., Roger J.M., Bellon Maurel V., and Rutledge D.N., 2004. Robustness of models developed by multivariate calibration. Part I: The assessment of robustness. *Trends in Analytical Chemistry*, 23 (2), pp.157-170.





# Chapitre 2 :

---

## **Mid-Infrared Prediction of Bovine Milk Fatty Acids across Multiple Breeds, Production Systems and Countries**

**Adapté de:** *Soyeurt H.\**, *Dehareng F.\**, *Gengler N.*, *McParland S.*, *Wall E.*, *Berry D.*, *Coffey M.*, and *Dardenne P.*, 2011. *Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. Journal of Dairy Science*, 94(4), 1657-1667.

*\* co-premiers auteurs*



## 2.1. Points d'attention

Ce second chapitre cherche à illustrer une solution originale permettant d'améliorer la qualité des modèles prédisant des composés mineurs du lait en visant à augmenter la variabilité spectrale du jeu de calibrage. Cette publication vise notamment à confirmer la capacité du FT-MIR à quantifier les teneurs individuelles en acides gras (AG) dans le lait en utilisant une procédure innovante d'échantillonnage sur plusieurs races, plusieurs pays et plusieurs systèmes de production.

La méthode analytique de référence choisie (chromatographie en phase gazeuse ou GC) est une méthode particulièrement chronophage et coûteuse. Le nombre important d'acides gras présents dans le lait (plus de 70 acides gras différents mesurés dans notre cas), volatils pour certains (acides gras à courte chaîne), nous a amené à recourir à des temps de séparation longs, à utiliser une colonne capillaire très longue (100 mètres), à procéder à une extraction à froid, à utiliser de l'hydrogène comme gaz vecteur et un injecteur on-column, ce qui a compliqué fortement le travail. Dès lors, vu ces contraintes et l'impossibilité de procéder à l'analyse d'un grand nombre d'échantillons pour construire le set de données de départ, nous avons dû recourir à une solution originale pour faire un premier tri des échantillons à analyser. Cette solution, fait suite aux travaux que nous avons entrepris quelques années plus tôt (Soyeurt et al., 2006). Afin de construire un set de données de calibrage présentant des variations de concentration en AG les plus larges possibles, on a fait le postulat simple qu'une variabilité spectrale plus grande induisait une variabilité plus grande des composés recherchés. Cette approche simple en soi, via les spectres, n'avait encore, à notre connaissance jamais été appliquée dans le domaine de la spectrométrie MIR. Pourtant, elle présente un double avantage. Premièrement, elle permet d'éviter de recourir à des plans d'échantillonnage coûteux. Deuxièmement, elle permet d'obtenir rapidement des équations robustes applicables à des bases de données importantes. Cette stratégie a été appliquée par la suite pour développer de nouveaux modèles pour lesquels les analyses de référence peuvent se faire à posteriori et dont les frais d'analyse sont particulièrement élevés (p. ex. Soyeurt et al., 2009 ; Christophe et al., 2021 ; Ferrand-Calmels et al., 2014).

Cet article soulève également une importante question relative à l'hétérogénéité de la réponse spectrale d'un instrument par rapport à un autre, même si la marque et le modèle sont identiques. Cette hétérogénéité pourrait entraîner des biais sur les spectres ou lors du transfert des modèles. Or, dans le cas présent, des données spectrales ont été échangées entre deux appareils. La solution proposée ici pour améliorer la reproductibilité entre les instruments, a été de générer un fichier appelé « fichier de répétabilité » ou en anglais « repeatability file » en enregistrant les spectre MIR de plusieurs échantillons de lait communs sur les différents spectromètres utilisés. En revanche, cette question devra être approfondie par la suite, car qu'en sera-t-il de l'utilisation de ces modèles lorsque l'on voudra les utiliser sur d'autres

spectromètres ou les appliquer sur des bases de données regroupant de multiples appareils ? Ce problème est moins important pour les prédictions des composés majeurs (MG, MAT et lactose), car dans ce cas on peut facilement ajuster les prédictions à l'aide de corrections de type biais et pente sur base de standards aux valeurs de référence connues comme expliqué dans le premier chapitre de cette thèse. Mais pour les nouveaux modèles PLS, ces variations risquent d'être amplifiées et la réalisation d'analyses de référence fréquentes sur des échantillons standards est souvent difficile, trop coûteuse, voire impossible à réaliser pour certains paramètres (p.ex. méthane). Nos recherches par la suite nous ont amené à développer la méthode de standardisation spectrale appelée « standardisation directe fragmentaire » ou « Piece-Wise Direct Standardization » en anglais. Cette méthode découle directement des travaux initiés par Shenk et al. (1985) dans le domaine de la spectrométrie NIR. Elle nous a permis de résoudre en partie le problème (Grelet et al., 2015 ; Tiplady et al., 2019), d'échanger des spectres d'autres régions ou pays, d'envisager le transfert de modèles sur d'autres appareils et de créer des bases de données spectrales à partir d'échantillons analysés sur différents appareils et à des périodes différentes.

Par la suite, ces équations AG vont pouvoir être utilisées de façon suffisamment robuste sur de larges bases de données spectrales, pour permettre des premières études sur l'activité de la  $\Delta 9$  désaturase (Soyeurt et al., 2008a), les paramètres génétiques de la teneurs en AG saturés ou insaturés (Soyeurt et al., 2008b), le stress thermique (Hammami et al., 2015), la balance énergétique (Ho et al., 2019), les variants génétiques (p.ex. : Knutsen et al., 2022) ou encore des indicateurs de la composition de la rations ingérées par les vaches (par ex. Coppa et al., 2021 ; Soyeurt et al., 2022).

Ma contribution dans cet article en tant que co-premier auteur a visé à mettre en place le protocole de transfert et de conservation des échantillons ainsi que le protocole d'analyse de ces derniers pour déterminer leurs teneurs de référence en AG du lait. J'ai également réalisé les analyses de référence et contribué à la rédaction de cette publication.

### ***2.1.1 Références***

- Christophe O.S., Grelet C., Bertozzi C., Veselko D., Lecomte C., Höeckels P., Werner A., Auer F.-J., Gengler N., Dehareng F. and Soyeurt H., 2021. Multiple Breeds and Countries' Predictions of Mineral Contents in Milk from Milk Mid-Infrared Spectrometry. *Foods*, 10, 2235.
- Ferrand-Calmels M., Palhière I., Brochard M., Leray O., Astruc J.M., Aurel M.R., Barbey S., Bouvier F., Brunschwig P., Caillat H., Douguet M., Faucon-Lahalle F., Gele M., Thomas G., Trommenschlager J.M. and Larroque H. 2014. Prediction of fatty acid profiles in cow, ewe, and goat milk by mid-infrared spectrometry. *J. Dairy Sci.*, 97, pp. 17-35.
- Grelet C., Fernandez 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(4), 2150-2160.
- Hammami H, Vandenplas J, Vanrobays ML, Rekik B, Bastin C and Gengler N., 2015. Genetic analysis of heat stress effects on yield traits, udder health, and fatty acids of Walloon Holstein cows. *J Dairy Sci.*, 98(7):4956-68.
- Ho P. N., Marett L. C., Wales W. J., Axford M., Oakes E. M. and Pryce J. E., 2019. Predicting milk fatty acids and energy balance of dairy cows in Australia using milk mid-infrared spectroscopy. *Animal Production Science*, 60, 164-168.
- Knutsen T.M., Olsen H.G., Ketto I.A., Sundaasen K.K., Kohler A., Tafintseva V., Svendsen M., Kent M.P. and Lien S, 2022. Genetic variants associated with two major bovine milk fatty acids offer opportunities to breed for altered milk fat composition. *Genet. Sel. Evol.*, 54, 35.
- Shenk J.S., Westerhaus M.O. and Templeton W.C., 1985. Calibration Transfer Between near Infrared Reflectance Spectrophotometers 1. *Crop Sci.*, 25(1), 159-161.
- Soyeurt H., Dehareng F., Mayeres P., Bertozzi C., Gengler N., 2008a. Variation of Delta 9-desaturase activity in dairy cattle. *J Dairy Sci.*, 91(8):3211-24.
- Soyeurt H., Dardenne P., Dehareng F., Bastin C., Gengler N., 2008b. Genetic parameters of saturated and monounsaturated fatty acid content and the ratio of saturated to unsaturated fatty acids in bovine milk. *J Dairy Sci.*, 91(9):3611-3626.

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

Tiplady K.M., Sherlock R.G., Littlejohn M.D., Pryce J.E., Davis S.R., Garrick D.J., Spelman R.J. and Harris B.L., 2019. Strategies for noise reduction and standardization of milk mid-infrared spectra from dairy cattle. *J. Dairy Sci.*, 102(7), 6357-6372.

## 2.2. Abstract

There is an increasing consumer concern over the relationship between food and human health. Because of its known effects on human health, the development of a quick, inexpensive and accurate method to quantify the fatty acid (FA) composition directly in milk would be valuable for milk processors to develop a payment system for milk pertinent to their customer requirements and for farmers to adapt their feeding systems and breeding strategies accordingly. The aim of this study was: 1) to confirm the ability of mid-infrared spectrometry (MIR) to quantify individual FA content in milk by using an innovative procedure of sampling combining multiple breeds, multiple countries and multiple production systems, 2) to compare 6 different mathematical methods to develop robust calibration equations for predicting the contents of individual FA in milk and 3) to show the interest to use the FA equations developed in milk basis to predict FA content in fat without corrections of slope and bias based on reference milk samples with known contents of FA. A total of 517 samples selected based on their spectral variability in 3 countries (Belgium, Ireland, and United Kingdom), from various breeds, cows, and production systems were analyzed by gas chromatography (GC). The samples presenting the largest spectral variability were used to calibrate the prediction of FA by MIR. The remaining samples were used to validate externally the 28 FA equations developed. The six methods were compared based on the cross-validation coefficient of determination ( $R^2_{cv}$ ), the ratio of standard deviation of reference GC values to the standard error of cross-validation (RPD), and the validation coefficient of determination ( $R^2_v$ ). The six methods were: 1) partial least squares regression (PLS), 2) PLS + repeatability file, 3) first derivative of spectral data + PLS, 4) first derivative + repeatability file + PLS, 5) second derivative of spectral data + PLS, and 6) second derivative + repeatability file + PLS. The third and fourth methods resulted, on average, in the highest values of  $R^2_{cv}$ , RPD, and  $R^2_v$ . Finally, the equations were built using all available GC data and the best accuracy was observed for the infrared predictions of C4:0, C6:0, C8:0, C10:0, C12:0, C14:0, C16:0, C18:0, C18:1 *trans*, C18:1 *cis*-9, C18:1 *cis* and for some groups of FA studied in milk (saturated FA, monounsaturated FA, unsaturated FA, short chain FA, medium chain FA, and long chain FA). All of these equations showed  $R^2_{cv}$  greater than 0.95. With  $R^2_{cv}$  equal to 0.85, the MIR prediction of polyunsaturated FA could be used to make a screening of the cow population. As already published, the infrared predictions of FA in fat are less accurate and no better results were obtained by using develop FA contents if no corrections for bias and slope based on reference milk samples with known contents of FA are used. The large number of samples, specifically chosen to cover the natural variation in milk FA present, coupled with the high accuracy of selection for some FA indicate the

usefulness of equations with  $R^2_{cv}$  greater than 95% in milk payment systems and the usefulness of equations with  $R^2_{cv}$  greater than 75% for animal breeding purposes.

## 2.3. Introduction

Traditionally milk quality in dairy cattle refers to fat and protein content or the number of somatic cells or total bacteria in the milk. However, there is an increasing consumer concern over the relationship between food and human diseases/health such as cardiovascular disease and cancer. Therefore, the subject of milk quality should be expanded to include characteristics with proven associations with human health. The nutritional quality of bovine milk is largely related to the milk fat content and composition. Bovine milk fat is composed of, on average, 96% triacylglycerols made up of a molecule of glycerol and 3 esterified fatty acids (FA) ([Jensen](#), 1995). There are 3 classes of FA, based on the number of double bonds present in their chemical structure: saturated fatty acids (SAT, no double bond), monounsaturated fatty acids (MONO, one double bond), and polyunsaturated fatty acids (POLY, more than one double bond). The latter two classes (MONO and POLY) make up the unsaturated group of FA (UNSAT). Many studies have documented the potential effects of dietary FA on human health (e.g., [Williams](#), 2000). Therefore, quick, inexpensive and accurate methods to quantify the FA composition of milk would be valuable for milk processors to develop a payment system for milk pertinent to their customer requirements and for farmers to adapt their feeding systems accordingly. Such tools would also be useful to animal breeders to take the long term strategy of breeding for milk with a FA composition sought by consumers.

Until recently, the only method used to quantify FA content in milk was a method based on gas chromatography (GC) after a step of extraction, followed by a saponification and transmethylation. Although this method is accurate and offers detailed information about the FA of milk sample ([Collomb and Bühler](#), 2000), it is expensive and time consuming and therefore not feasible for routine analysis of the large number of samples necessary for management or breeding purposes. Several studies ([Soyeurt et al.](#), 2008a; 2008b; [Rutten et al.](#) 2009) have documented the potential to use mid-infrared (MIR) spectrometry to quantify the major FA in cow milk. The approach proposed by [Soyeurt et al.](#) (2008a; 2008b) used milk samples from multiple breeds based on the assumption that changes in the milk MIR regions related to FA were proportional to the contents of FA in milk and these relationships were similar across breeds. [Rutten et al.](#) (2009) hypothesized that the changes in the spectral data related to the FA profile (i.e. the relationship among FA) could be different during winter and summer and developed calibration equations within season in order to



improve the accuracy of prediction. However, they concluded that these equations did not improve the predictive ability of milk FA from MIR spectrometry.

The method of choice currently used for relating MIR spectra to milk fatty acid is partial least squares (PLS) without any pre-treatment of the spectral data (Soyeurt et al. 2006, 2008a, 2008b; Rutten et al. 2009). Equations with improved predictive ability for milk FA composition developed using different statistical approaches such as pre-treatment of spectral data is lacking.

The objective of this study was: 1) to confirm the ability of mid-infrared spectrometry (MIR) to quantify individual FA content in milk by using an innovative procedure of sampling combining multiple breeds, multiple countries and multiple production systems, 2) to compare 6 different mathematical methods to improve the robustness of calibration equations for predicting the contents of individual FA in milk and 3) to show the interest to use the FA equations developed in milk basis to predict FA content in fat without corrections of slope and bias based on reference milk samples with known contents of FA.

## 2.4. Materials and methods

### 2.4.1. Collected Samples

Samples available for this study contained 40ml of milk with bronopol preservative and were collected between March 2005 and December 2009 in Belgium, between April 2009 and August 2009 in Ireland, and in August 2009 in the UK. All milk samples were analyzed using MilkoScan FT6000 spectrometers (Foss, Hillerod, Denmark) which outputs one spectra for each milk sample analyzed containing 1,060 data points in the infrared range from 900  $\text{cm}^{-1}$  to 5,000  $\text{cm}^{-1}$ . Belgian milk samples were analyzed in the milk laboratory 'Comité du lait' (Battice, Belgium) and other samples were analyzed at Moorepark (Fermoy, Cork, Ireland). All spectra were generated directly without using the calibration mode of the spectrometer because a direct access to the spectral data was available in the 2 milk labs. The collection of samples was separated through the time.

A total of 1,609 milk samples were collected from 8 herds in the Walloon Region of Belgium between March 2005 and May 2006 during the routine milk recording and frozen at  $-26^{\circ}\text{C}$ . These milk samples containing 50% of morning milk and 50% of evening milk were from 475 cows representing 6 dairy breeds (Dual Purpose Belgian Blue, Holstein, Jersey, Normande, Montbeliarde, and Red and White). Herds were chosen based on their variability in milk fat content, their geographical location, and the number of breeds in each herd. Because it was not

feasible to undertake GC on all samples, a selection of 78 samples were chosen based on maximizing the spectral variability of the samples, determined using a principal component approach (PCA). This approach identifies the spectra that are most useful in development of prediction equations.

To further increase the variability of milk samples with different MIR spectra, additional samples were collected from the Walloon Region of Belgium between June 2006 and December 2009. Using PCA, the MIR spectra of bulk milk samples and individual cow samples (a mix of 50% of morning milk and 50% of evening milk) routinely collected and analyzed in the Belgian milk lab (Comité du Lait, Battice, Belgium) were compared to the spectra of the 78 initially selected samples. Milk samples whose MIR spectra differed, based on the PCA, from the previously collected samples were retained. In total, an additional 210 samples were selected using this method and frozen pending the GC analysis.

The MIR spectra of weekly milk samples from Teagasc research herds in Ireland, and monthly milk samples from the Loughill lines of dairy cattle kept at Crichton research herd, Scotland were also compared using PCA to the samples already stored in Belgium. The number of milk samples chosen to further increase the variability within the dataset of MIR spectra was 144 in Ireland and 106 in Scotland. All of these samples (morning or evening milks) were frozen at -20°C pending the GC analysis. The cows sampled in the Irish research herds were fed predominantly grazed grass and consisted of animals of different breeds ([Prendiville et al., 2010](#)) and different strains of Holstein-Friesian ([Coleman et al., 2009](#)). Cows from Scotland were from 2 genetically divergent lines (divergent for milk solids) and were fed on 2 different diets ([Coffey et al., 2004](#)). Approximately half were fed on a predominantly forage only diet and the other half a diet consisting approximately 60% silage and 40% concentrates.

The multiple breed, multiple countries, and multiple production approach proposed in this study to collect the samples used to develop the calibration equations allows us to cover the natural variability of fatty acids in bovine milk.

#### ***2.4.2. Measurement of Fatty Acids***

In total, 538 frozen milk samples from Belgium, Ireland, and Scotland with MIR spectral data were analyzed by GC. The contents of FA in each milk sample were quantified by GC with a methodology similar to that described by [Soyeurt et al. \(2008a\)](#). GC analyses of all 538 samples were undertaken by the Agricultural Walloon Research Centre (Gembloux, Belgium). The milk fat was extracted according to ISO Standard 14156:2001 ([ISO, 2001](#)). Preparation of FA methyl esters was made

following the ISO 15884:2002 ([ISO](#), 2002). These milk fat samples were analyzed using GC from the method developed by [Collomb and Bühler](#) (2000). The gas chromatograph (model 6890N, Agilent Technologies Inc., Palo Alto, CA) was equipped with a CPSil-88 capillary column (Varian Inc., Palo Alto, CA) with a length of 100 m, an internal diameter of 0.25 mm, and a film thickness of 0.20  $\mu\text{m}$ . The conditions for the chromatographic analyses were as follows: carrier gas, helium; average velocity, 19 cm/s; cold on-column injector; flame ionization detector at 255°C; and a temperature program from 60°C (5 min) to 165°C (at 14°C/min) during 1 min, then 165 to 225°C (at 2°C/min) during 17 min. The volume injected was 0.5  $\mu\text{l}$ . An anhydrous milk fat with a certified FA composition (reference material BCR-164, obtained from the Commission of the European Communities, Brussels, Belgium) was used to determine the FA methyl ester response factors, the repeatability, and the accuracy of this method. For all studied FA, the coefficient of variation ( $(\text{SD}/\text{mean}) \times 100$ ) was less than 3.5% suggesting a good repeatability of GC data.

The FA contents were expressed in g/dl of milk using the content of fat predicted by the MilkoScan FT6000 (Foss, Hillerod, Denmark). This prediction was considered in this publication as the reference fat content for two reasons. The ability of MIR to quantify the content of fat is high and the accuracy of the fat prediction is regularly checked by the milk lab using reference samples with known fat content. The error of prediction of fat content on the Belgian MilkoScan FT6000 spectrometers used from the data collected in 2009 from 179 spectra was equal to 0.2 g/dl of milk. The group of short chain FA (SCFA) included FA with between 4 and 10 carbons (C) in their structure. Medium chain FA (MCFA) included FA whose structures contained between 12 and 16 C, while long chain FA (LCFA) included FA with structures containing between 17 and 22 C.

A PCA was undertaken across all FA and the Mahalanobis distance was calculated. As described by [Williams](#) (2007), an outlier sample or spectrum is a sample or spectrum, which differs from the mean of the population by 3 or more Mahalanobis distance. Using this threshold, 21 samples were deemed to be outliers and were discarded. The final data set contained 517 samples.

### ***2.4.3. Calibration and Validation Set***

An initial calibration dataset of 220 samples was generated using only the samples from the Walloon region of Belgium. To augment the variability of the MIR spectra in this dataset, 47 samples from Belgium (n=31), Ireland (n=12) and Scotland (n=4) were identified using PCA and added to the calibration dataset. The samples from each country were not balanced in the calibration set because the majority of

milk samples collected in Ireland and Scotland showed a spectral variability existing already in the Walloon data set. Only informative samples were added. Therefore, the calibration data set used in this study contained 267 samples and the remaining 250 samples were included in the external validation dataset (16 samples from Belgium, 132 from Ireland and 102 samples from Scotland).

#### ***2.4.4. Development of Prediction Equations***

In this study, six different approaches were used to develop the most accurate prediction equations:

- method 1: PLS and no pre-treatment on the spectral data;
- method 2: PLS + the use of a repeatability file;
- method 3: PLS + the use of a first derivative pre-treatment on the spectral data;
- method 4: PLS + the use of a first derivative pre-treatment + repeatability file;
- method 5: PLS + the use of a second derivative pre-treatment;
- method 6: PLS + the use of a second derivative pre-treatment + repeatability file.

In order to make the spectral data linear, the initial spectral data expressed in transmittance were converted in absorbance by taking the  $\log_{10}$  of the reciprocal of transmittance.

An equation developed for one spectrometer could provide results slightly biased on another instrument. To improve the repeatability across instruments, a file referred to throughout this paper as the ‘repeatability file’ was generated by recording the MIR spectrum of several milk samples provided by different spectrometers. In the present study, the MIR spectra of 108 samples were determined at least on 2 different spectrometers. A total of 5 different MilkoScan FT6000 (Foss, Hillerod, Denmark) were used: 3 located in a Belgian milk lab (Comité du Lait, Battice, Belgium), 1 located in Luxembourg (CONVIS herdbuch, Ettelbruck, Luxembourg) and 1 located in Ireland (Teagasc Moorepark, Co. Cork, Ireland). The approach taken to use the repeatability file in the development of equations was similar to the one described by [Westerhaus](#) (1990). As explained by [Westerhaus](#) (1990), each spectrum was centered by subtracting the average of all spectra for the samples included in the repeatability file (i.e., 108 selected samples). The reference GC values for these centered samples (i.e., composition of FA) were equal to 0. These samples in the repeatability file were then used to extend the initial calibration set. This method creates a decrease in the repeatability error. The repeatability of spectral data for the infrared regions considered was on average lower than 2%.

The derivative of the spectral data was obtained to “sharpen” the absorption bands. First and second derivatives were obtained using the following formula:

$$dx_k = x_{k-\frac{g-1}{2}} - x_{k+\frac{g-1}{2}}$$

Where  $dx_k$  is the value of the derivative for the spectral data point  $k$ ;  $g$  is an odd integer strictly positive entitled ‘gap’ (i.e., 5 consecutive spectral data points in this study; [Hruschka, 1987](#)). The second derivative used the absorbance obtained for the first derivative instead of the absorbance of spectra.

A cross-validation approach using 20 groups of samples from the calibration set was used to determine the most appropriate number of factors for the prediction equation of each FA and to assess the robustness of calibration. Cross-validation leaves out one group at a time and performs a calibration with the remaining 19 groups. Then, the calibration was applied to the one group left out and then repeats that 20 times with a different group left out each time. Therefore, cross-validation is an internal validation because it uses the same samples for validation and calibration. The maximal number of factors per equation was 16.

The calibration coefficient of determination ( $R^2c$ ) was calculated as the square of the correlation coefficient between the reference GC data and their corresponding predicted values calculated from the final model based on all samples included in the calibration set. The standard error of calibration (SEC) was calculated as:

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

where  $\hat{y}_i$  was the predicted values obtained for the sample  $i$  included in the calibration set;  $y_i$  was the reference GC value of the sample  $i$ ,  $N$  was the number of samples in the calibration set; and  $p$  the number of factor used to build each calibration equation.

The standard error of cross-validation (SECV) was calculated as:

$$SECV = \sqrt{\frac{\sum_{j=1}^Z \sum_{i=1}^N (\hat{y}_{i,j} - y_{i,j})^2}{N}}$$

where  $\hat{y}_{ij}$  was the predicted values obtained the for the sample  $i$  in the group of cross-validation  $j$ ;  $y_{ij}$  was the reference GC value the for the sample  $i$  in the group of cross-validation  $j$ ,  $N$  was the number of samples in each group of cross-validation; and  $Z$  was the number of cross-validation groups.

The cross-validation coefficient of determination ( $R^2_{cv}$ ) was calculated as the square of the correlation coefficient between the reference GC data and their corresponding predicted values calculated from the results obtained by the 20 models developed during the cross-validation.

The mean and SD for each FA measured using GC were calculated for the validation and calibration sets. The best calibration equation should have the larger SD for GC data (to cover the natural range of FA variability) and the smallest SECV. So, the ratio of SD from GC data to SECV was calculated (RPD, [Williams and Sobering](#), 1993). If this ratio is greater than 3, the prediction given by the developed calibration equation can be considered good.

By observing the spectral variation ([Soyeurt et al.](#), 2010), it was decided to develop the calibration from 3 specific regions located between 926  $\text{cm}^{-1}$  and 1,600  $\text{cm}^{-1}$ , 1,712  $\text{cm}^{-1}$  and 1,809  $\text{cm}^{-1}$ , and 2,561  $\text{cm}^{-1}$  and 2,989  $\text{cm}^{-1}$ . The selection of MIR regions for calibration is less important with the use of a repeatability file because this one will decrease the PLS coefficient for wavelength with poor repeatability ([Hruschka](#), 1987).

#### ***2.4.5. External Validation***

The equations developed by the six proposed methods were also externally validated using the 250 independent samples. The validation coefficient of determination ( $R^2_v$ ) was calculated using the same formula shown to calculate  $R^2_c$  except that the values were calculated from the validation set. The six studied methods were compared based on  $R^2_{cv}$ , RPD, and  $R^2_v$ .

#### ***2.4.6. Accuracy of Calibration***

To assess the accuracy of mid-infrared prediction of FA, all samples with FA content data available were combined together in a final data set ( $N=517$  samples) and the equations were re-built using the best method derived from the results obtained with the calibration and validation data. To improve the robustness of equations, the critical T outlier value (Winisi software, Foss, Hillerod, Denmark) was set to 2.5 (i.e. recommended default value). If the difference between the predicted and reference FA values was greater than 2.5 times the SEC, the sample was considered to be an outlier and deleted. This methodology was iterated twice. In order to ensure a good

comparison of methods, this methodology of outlier detection was not used before because the samples considered as outliers were not the same for the 6 studied methods. It is logical because this method is based on differences between predicted and reference values. So, this method is influenced by the accuracy of developed equations. As shown later, the 6 methods gave different error values.

## 2.5. Results and discussion

### 2.5.1. Descriptive statistics

A total of 47 samples were selected from the Irish and Scottish samples and the 49 remaining Belgian samples. This low number of additional selected samples demonstrated the high natural spectral variability already covered by the initial Belgian calibration set which comprised 221 milk samples. This was to be expected as the initial data set included milk samples representative of varying seasons, years, and geographical location from cows of different parity, stage of lactation, and breed.

Mean, SD, minimum and maximum values, and the coefficient of variation (CV) of the GC data used in the calibration and validation sample sets are presented in [Table 2.1](#). The variability of FA profile in the calibration set used in this study can be considered as high based on the values of CV calculated for each studied FA. CV ranged from 8.97% to 60.05%. CV, minimal and maximal values shown in [Table 2.1](#) for the external validation set demonstrate that the variability of GC data in this set was lower than the one shown in the calibration set except for the group of C18:2 and branched FA, C18:2 *cis-9,cis-12*, and C18:2 *cis-9,trans-11*. These figures confirm the high range of natural variability of FA present in the sample set used to build the calibration equations. This variation, which is larger than previously published estimates ([Rutten et al., 2009](#)), resulted from the careful selection of samples for entry to the calibration data set based on spectral variability.

### 2.5.2. Comparison of Methods to Develop Calibration Equations for Fatty Acids in Milk

[Table 2.1](#) summarizes the  $R^2_{cv}$ , RPD, and  $R^2_v$  calculated for each calibration equation built from all studied methods and FA contents in milk (g/dl of milk). The RPD ratio is useful to compare the methods studied since, by definition, RPD considers SECV and the variability of the calibration set simultaneously. So, RPD is more sensitive than  $R^2_{cv}$ . For instance, the  $R^2_{cv}$  for fat and SAT calibration equations were equal but RPD were different.



**Table 2.1.** Descriptive statistics of gas chromatographic data for the calibration and validation sample sets

Constituent <sup>1</sup> (g/100 g of fat)	Calibration set (n = 267)					Validation set (n = 250)						
	Mean	SD	Minimum	Maximum	CV	Range	Mean	SD	Minimum	Maximum	CV	Range
C4:0	2.67	0.51	1.26	4.18	19.24	2.92	2.64	0.49	1.26	4.28	18.41	3.02
C6:0	1.87	0.34	0.90	2.97	17.97	2.07	1.79	0.30	0.86	2.47	16.58	1.62
C8:0	1.18	0.27	0.47	2.10	22.45	1.63	1.13	0.20	0.53	1.70	17.60	1.17
C10:0	2.70	0.73	0.79	5.43	26.93	4.64	2.66	0.53	1.44	4.57	20.02	3.14
C12:0	3.30	0.95	0.90	7.05	28.79	6.16	3.23	0.63	1.78	5.79	19.51	4.01
C14:0	10.94	2.08	4.69	16.59	18.97	11.90	11.00	1.42	6.97	14.53	12.88	7.56
C14:1	1.01	0.39	0.30	2.58	39.26	2.29	1.07	0.39	0.34	2.53	36.34	2.19
C16:0	29.13	4.95	20.00	46.03	17.00	26.03	28.03	4.70	18.99	40.20	16.77	21.21
C16:1 <i>cis</i>	1.79	0.61	0.24	5.25	34.35	5.01	1.96	0.48	1.09	3.51	24.27	2.42
C17:0	0.71	0.18	0.00	1.06	24.90	1.05	0.73	0.07	0.47	0.93	10.00	0.46
C18:0	10.50	2.77	1.93	18.67	26.35	16.74	9.89	2.48	4.00	17.39	25.03	13.39
C18:1 <i>trans</i>	3.28	1.47	1.11	8.95	44.72	7.83	4.58	1.70	1.07	9.41	37.22	8.34
C18:1 <i>cis</i> -9	20.10	4.94	8.80	36.36	24.57	27.56	20.39	2.94	12.62	28.28	14.45	15.66
Total C18:1 <i>cis</i>	21.76	5.06	9.84	38.41	23.24	28.57	21.70	3.11	13.44	30.10	14.35	16.66
Total C18:2	2.28	0.58	1.11	6.44	25.30	5.33	2.45	0.63	1.07	4.18	25.52	3.11
C18:2 <i>cis</i> -9, <i>cis</i> -12	1.53	0.49	0.63	3.93	31.93	3.31	1.47	0.64	0.47	2.92	43.33	2.45
C18:3 <i>cis</i> -9, <i>cis</i> -12, <i>cis</i> -15	0.58	0.22	0.18	1.32	38.31	1.14	0.57	0.16	0.02	1.09	28.56	1.07
C18:2 <i>cis</i> -9, <i>trans</i> -11	0.88	0.53	0.21	2.90	60.65	2.68	1.29	0.79	0.21	4.02	61.05	3.81
Saturated FA	67.01	6.01	51.08	79.31	8.97	28.22	65.20	4.35	53.02	77.97	6.67	24.95
Monounsaturated FA	28.57	5.32	17.80	43.94	18.62	26.14	29.92	3.77	19.03	39.98	12.59	20.95
Polyunsaturated FA	4.31	1.05	2.51	8.75	24.47	6.24	4.83	0.98	2.21	7.77	20.29	5.56
Unsaturated FA	32.88	5.89	20.69	48.89	17.91	28.20	34.75	4.47	22.30	46.98	12.87	24.68
Short-chain FA	8.75	1.58	4.73	13.82	18.11	9.09	8.51	1.29	4.33	12.25	15.16	7.92
Medium-chain FA	49.81	7.33	32.97	72.33	14.72	39.36	49.10	6.11	35.36	63.63	12.44	28.27
Long-chain FA	41.33	7.96	18.83	60.90	19.25	42.07	42.34	6.47	26.89	58.34	15.28	31.45
Branched-chain FA	2.56	0.50	1.15	4.51	19.58	3.36	2.59	0.70	1.23	5.47	26.88	4.25
n-3 FA	0.80	0.31	0.24	2.41	38.59	2.17	0.77	0.18	0.24	1.28	24.01	1.05
n-6 FA	2.63	0.64	1.34	7.38	24.39	6.04	2.77	0.67	1.32	4.69	24.07	3.37

<sup>1</sup>Short-chain fatty acids (FA) = C4 to C10; medium-chain FA = C12 to C16; long-chain FA = C17 to C22.



The best calibration equation has to have the highest  $R^2_{cv}$ , RPD and  $R^2_v$ . Method 4 (PLS + first derivative + repeatability file) gave the best prediction results for the majority of studied FA (see results in bold face in [Table 2.1](#)) because this method has simultaneously more often the highest values for  $R^2_{cv}$ , RPD, and  $R^2_v$ . This conclusion was expected. Inclusion of spectra provided by separate spectrometers from milk samples of individual cows of different breeds and countries kept on varying feeding systems resulted in increased noise (e.g., baseline drift) and explains the need for pre-treatment of the data before the development of the calibration equations. As explained previously, the repeatability file used in method 4 allows accounting for the repeatability of each spectral data point.

Even if differences between the use of a repeatability file and no repeatability file (methods 3 and 4, [Table 2.2](#)) were not very large (however the P values between methods used a repeatability file and the others were less than 5%), the use of the first derivatives improved the accuracy of MIR prediction for FA (comparison between methods 2 and 4, [Table 2.2](#)). The first derivatives corrected for the baseline drift evident in the spectra file ([Figure 2.1](#)). The use of first derivative was also suggested by Dal Zotto et al. (2008). These authors mentioned that the use of the first derivative on spectral data improved the accuracy of the equations predicting the milk rennet coagulation time and the curd firmness. The second derivative, which removes the linear trend from the spectral data, did not have a statistically significant effect in this study.

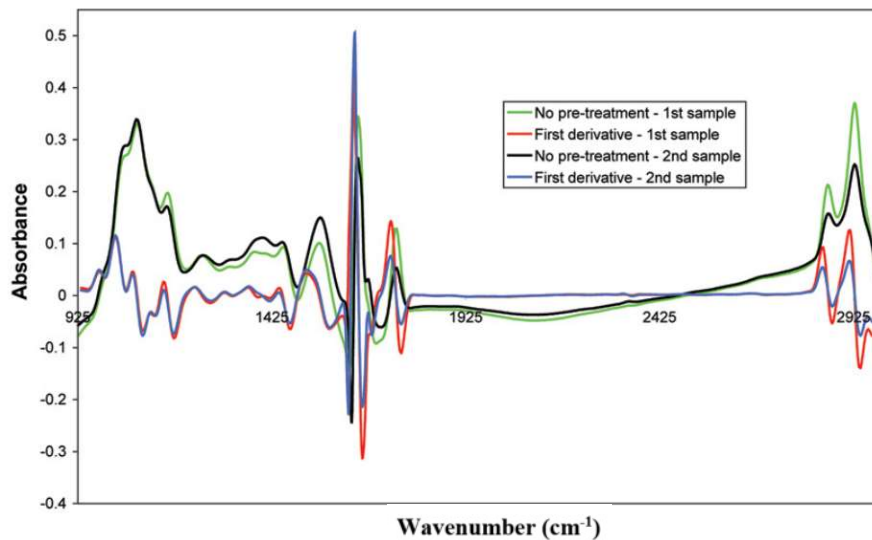
Generally, the highest  $R^2_{cv}$  and RPD were observed for the studied groups of FA ([Table 2.1](#)). Therefore, as shown by [Soyeurt et al. \(2006\)](#) and [Rutten et al. \(2009\)](#), the efficiency of MIR prediction of FA depends on the concentrations of FA in milk. The correlation between the average concentration of FA in the calibration set and the corresponding  $R^2_{cv}$  was positive and equal to 0.59. Moreover, the calibration set for each studied trait should also be variable. In a regression including the mean and SD of GC data expressed in milk, SD explained more than 36% of the variability of  $R^2_{cv}$  obtained. Consequently, the infrared prediction should be improved if the FA yield in milk is high and the content of the considered FA is variable within the calibration set. A similar correlation of 0.55 was obtained using the validation set.

**Table 2.2.** Comparison of the 6 mathematical approaches (methods 1 to 6) used to develop the calibration equations for fatty acids (FA) in milk (g/dL of milk; n = 267 samples).

Constituent (g/dL of milk)	R <sup>2</sup> <sub>cv</sub>						RPD						R <sup>2</sup> <sub>v</sub>					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Fat	1.00	1.00	1.00	1.00	1.00	1.00	21.8	19.4	24.0	22.2	22.4	21.1	0.99	1.00	1.00	1.00	1.00	1.00
C4:0	0.87	0.86	0.88	<b>0.88</b>	0.87	0.87	2.8	2.7	2.9	<b>2.9</b>	2.8	2.8	0.81	0.81	0.82	<b>0.83</b>	0.83	0.81
C6:0	0.93	0.94	<b>0.95</b>	0.94	0.94	0.94	3.9	4.0	<b>4.3</b>	4.2	4.2	4.2	0.87	<b>0.88</b>	0.88	0.88	0.87	0.88
C8:0	0.91	0.91	0.92	0.92	0.91	0.92	3.3	3.4	3.5	3.5	3.4	3.5	0.88	0.89	0.89	0.89	0.90	0.89
C10:0	0.90	0.90	<b>0.91</b>	<b>0.91</b>	<b>0.90</b>	<b>0.91</b>	3.1	3.1	3.2	<b>3.3</b>	3.2	3.3	0.87	<b>0.88</b>	0.88	<b>0.89</b>	0.90	0.89
C12:0	0.89	0.90	<b>0.90</b>	<b>0.90</b>	<b>0.90</b>	<b>0.90</b>	3.0	3.2	<b>3.2</b>	<b>3.2</b>	<b>3.2</b>	3.2	0.87	0.89	<b>0.90</b>	<b>0.90</b>	<b>0.90</b>	0.89
C14:0	0.93	0.94	0.93	<b>0.94</b>	0.93	0.94	3.9	4.0	3.8	<b>4.0</b>	3.9	3.9	0.89	0.90	0.91	<b>0.91</b>	0.90	0.91
C14:1	0.54	0.55	0.55	<b>0.58</b>	0.57	0.57	1.5	1.5	1.5	<b>1.5</b>	1.5	1.5	0.50	0.50	0.49	<b>0.50</b>	0.46	0.47
C16:0	0.91	0.91	0.91	<b>0.92</b>	0.91	0.91	3.4	3.4	3.4	<b>3.4</b>	3.3	3.3	0.85	0.86	0.84	<b>0.86</b>	0.84	0.84
C16:1 cis	0.56	0.56	0.59	0.61	0.60	0.60	1.5	1.5	1.6	1.6	1.6	1.6	0.50	0.52	0.51	0.52	0.54	0.53
C17:0	<b>0.47</b>	<b>0.47</b>	0.46	0.46	0.47	0.47	1.4	1.4	1.4	1.4	1.4	1.4	0.47	0.47	0.47	0.47	0.47	0.47
C18:0	0.86	0.86	<b>0.87</b>	0.87	0.86	0.87	2.7	2.7	2.7	<b>2.8</b>	2.7	2.7	0.72	0.72	0.73	0.72	0.74	0.70
C18:1 trans	0.81	0.81	<b>0.82</b>	0.83	0.81	0.83	2.3	2.3	2.4	<b>2.4</b>	2.3	2.4	0.81	0.80	0.84	<b>0.84</b>	0.74	0.80
C18:1 cis-9	0.90	0.91	<b>0.95</b>	0.95	0.94	0.94	3.2	3.4	4.3	<b>4.4</b>	4.1	4.2	0.89	0.90	0.89	<b>0.90</b>	0.89	0.90
Total C18:1 cis	0.90	0.92	0.95	0.95	0.94	0.95	3.2	3.4	4.3	4.3	4.2	4.4	0.90	0.92	0.90	0.91	0.90	0.90
Total C18:2	0.66	0.66	0.67	0.68	0.68	<b>0.68</b>	1.7	1.7	1.7	1.7	1.8	1.8	0.45	0.46	0.49	0.50	0.53	<b>0.54</b>
C18:2 cis-9, cis-12	0.66	0.66	0.66	0.66	<b>0.67</b>	0.66	1.7	1.7	1.7	1.7	1.7	1.7	0.34	0.34	0.34	0.34	<b>0.39</b>	0.38
C18:3 cis-9, cis-12, cis-15	0.49	0.48	0.51	0.53	0.53	0.54	1.4	1.4	1.4	1.4	1.5	1.5	0.63	0.59	0.60	0.59	0.57	0.60
C18:2 cis-9, trans-11	0.56	0.56	0.52	0.52	0.51	0.54	1.5	1.5	1.4	1.5	1.4	1.5	0.62	0.64	0.64	0.67	0.59	0.64
Saturated FA	0.99	0.99	<b>0.99</b>	0.99	0.99	0.99	8.2	8.4	<b>10.0</b>	9.7	9.6	9.4	0.98	0.98	0.98	<b>0.98</b>	0.98	0.98
Monounsaturated FA	0.94	0.94	<b>0.97</b>	0.97	0.96	0.96	4.1	4.2	<b>5.4</b>	5.4	5.2	5.3	0.95	0.96	0.96	<b>0.96</b>	0.96	0.95
Polysaturated FA	0.77	0.76	<b>0.77</b>	0.77	0.76	0.77	2.1	2.0	<b>2.1</b>	2.1	2.0	2.1	0.81	0.81	0.83	0.82	0.82	0.82
Unsaturated FA	0.95	0.95	<b>0.97</b>	0.97	0.97	0.97	4.3	4.6	<b>5.7</b>	5.6	5.5	5.5	0.96	0.96	0.96	<b>0.96</b>	0.96	0.95
Short-chain FA	0.94	0.94	<b>0.95</b>	0.94	0.94	0.94	4.1	4.2	<b>4.3</b>	4.2	4.2	4.2	0.90	0.91	0.91	<b>0.91</b>	0.91	0.91
Medium-chain FA	0.95	0.95	0.95	<b>0.95</b>	0.95	0.95	4.4	4.5	4.5	<b>4.6</b>	4.4	4.6	0.91	0.91	0.91	<b>0.92</b>	0.91	0.91
Long-chain FA	0.94	0.94	0.95	<b>0.95</b>	0.95	0.95	4.1	4.2	4.6	<b>4.6</b>	4.4	4.4	0.92	0.91	0.92	<b>0.93</b>	0.92	0.92
Iso-ante + iso	0.67	0.68	0.69	<b>0.70</b>	0.65	0.66	1.7	1.8	1.8	<b>1.8</b>	1.7	1.7	0.65	0.61	0.63	<b>0.65</b>	0.60	0.65
n-3 FA	0.46	<b>0.46</b>	0.46	0.46	0.46	0.46	1.3	1.4	1.4	1.4	1.3	1.3	0.69	<b>0.69</b>	0.68	0.67	0.65	0.65
n-6 FA	0.67	0.68	0.68	0.68	0.69	<b>0.69</b>	1.7	1.8	1.8	1.8	1.8	1.8	0.50	0.46	0.50	0.52	0.54	<b>0.55</b>

<sup>1</sup>Method 1 = partial least squares (PLS) and no pretreatment on the spectral data; method 2 = PLS + the use of a repeatability file; method 3 = PLS + the use of a first-derivative pretreatment on the spectral data; method 4 = PLS + the use of a first-derivative pretreatment + repeatability file; method 5 = PLS + the use of a second-derivative pretreatment; method 6 = PLS + the use of a second-derivative pretreatment + repeatability file. Bold face indicates the most interesting calibration equations.

<sup>2</sup>R<sup>2</sup><sub>cv</sub> = cross-validation coefficient of determination; RPD is the ratio of SD of GC data from the calibration set to the standard error of cross-validation; R<sup>2</sup><sub>v</sub> = validation coefficient of determination.



**Figure 2.1.** Effect of first derivative on 2 mid-infrared milk spectra. In order to make an easier reading, the scales of the first derivatives were changed.

The best method differed depending on the studied FA traits even if the use of first derivative seemed to be the best approach with or without the repeatability file (see results in bold face in Table 2). Fat content, C6:0, C18:0, C18:1 trans, C18:1 cis-9, SAT, MONO, POLY, UNSAT, and SCFA seemed to be slightly better predicted by method 3. In contrast, C17:0 and the group of Omega-3 were better predicted using method 2. Moreover, the predictions obtained for the total content of C18:2 cis-9, cis-12 seemed to be more accurate from method 5 and 6, respectively. However, these differences were slight and thus suggested that the use of method 4 could be the best for all studied FA traits.

### 2.5.3. Accuracy of Calibration

In order to give an idea of the accuracy of each infrared prediction, [Table 2.4](#) presents the best approach for each studied FA using the entire data set (N=517) together with the corresponding descriptive statistics, SECV,  $R^2_{cv}$ , and RPD. The RPD calculated using the entire data set was generally greater than the RPD shown in [Table 2.2](#) (calculated using only the samples in the calibration data set). In contrast to [Table 2.3](#), there is no difference observed between methods 3 and 4.

**Table 2.3. Coefficient of determination between the reference GC data and the prediction (Pred.) of fatty acid (FA) contents in fat (g/100 g of fat) obtained from the predictions.**

Constituent (g/100 g of fat) <sup>1</sup>	Calibration set (n = 267)						Validation set (n = 250)							
	Reference fat <sup>2</sup>		Pred. fat with 4 <sup>3</sup>		Pred. fat with 3 <sup>4</sup>		Pred. fat with 4 <sup>5</sup>		Reference fat		Pred. fat with 3		Pred. fat with 4	
	4 <sup>6</sup>	Best <sup>7</sup>	4	Best	4	Best	4	Best	4	Best	4	Best	4	Best
C4:0	0.72	NA <sup>8</sup>	0.73	NA	0.73	NA	0.73	NA	0.58	NA	0.57	NA	0.57	NA
C6:0	0.82	0.85	0.83	0.85	0.84	0.86	0.80	0.80	0.59	0.58	0.61	0.59	0.61	0.55
C8:0	0.82	NA	0.82	NA	0.83	NA	0.82	NA	0.68	NA	0.70	NA	0.69	NA
C10:0	0.80	NA	0.80	NA	0.82	NA	0.83	NA	0.73	NA	0.74	NA	0.75	NA
C12:0	0.85	NA	0.85	NA	0.85	NA	0.81	0.81	0.77	NA	0.78	NA	0.78	NA
C14:0	0.85	NA	0.85	NA	0.86	NA	0.86	NA	0.65	NA	0.68	NA	0.68	NA
C14:1	0.41	NA	0.42	NA	0.43	NA	0.59	0.28	0.28	NA	0.28	NA	0.29	NA
C16:0	0.64	NA	0.64	NA	0.63	NA	0.72	0.48	0.48	NA	0.50	NA	0.52	NA
C16:1 cis	0.36	0.37	0.35	0.37	0.35	0.37	0.45	0.23	0.23	0.18	0.23	0.18	0.24	0.19
C18:0	0.66	NA	0.66	NA	0.66	NA	0.70	0.38	0.38	NA	0.39	NA	0.37	NA
C18:1 trans	0.67	NA	0.68	NA	0.69	NA	0.77	0.70	0.70	NA	0.69	NA	0.69	NA
C18:1 cis-9	0.79	NA	0.80	NA	0.82	NA	0.87	0.64	0.64	NA	0.66	NA	0.65	NA
Total C18:1 cis	0.80	NA	0.82	NA	0.83	NA	0.86	0.64	0.64	NA	0.67	NA	0.67	NA
Total C18:2	0.31	0.26	0.31	0.28	0.31	0.30	0.37	0.07	0.24	0.24	0.08	0.24	0.09	0.26
C18:2 cis-9, cis-12	0.43	0.45	0.45	0.32	0.45	0.45	0.42	0.25	0.31	0.26	0.22	0.28	0.34	0.37
C18:3 cis-9, cis-12, cis-15	0.21	0.18	0.22	0.20	0.24	0.22	0.44	0.24	0.36	0.24	0.24	0.35	0.23	0.30
C18:2 cis-9, trans-11	0.57	NA	0.57	NA	0.57	NA	0.56	0.54	0.54	NA	0.53	NA	0.53	NA
Saturated FA	0.84	0.90	0.85	0.92	0.82	0.90	0.87	0.74	0.72	0.79	0.77	0.77	0.78	0.74
Monounsaturated FA	0.88	NA	0.88	NA	0.89	NA	0.88	0.88	0.75	NA	0.77	NA	0.77	NA
Polysaturated FA	0.40	0.35	0.41	0.36	0.42	0.38	0.59	0.56	0.61	0.61	0.55	0.59	0.56	0.61
Unsaturated FA	0.87	0.86	0.89	0.88	0.90	0.90	0.89	0.78	0.79	0.80	0.81	0.81	0.81	0.78
Short-chain FA	0.85	0.83	0.85	0.85	0.85	0.85	0.81	0.66	0.64	0.69	0.66	0.66	0.68	0.61
Medium-chain FA	0.80	NA	0.81	NA	0.80	NA	0.82	0.62	0.62	NA	0.64	NA	0.65	NA
Long-chain FA	0.83	NA	0.84	NA	0.85	NA	0.85	0.73	0.73	NA	0.74	NA	0.73	NA
Iso-arite + iso	0.42	NA	0.43	NA	0.43	NA	0.42	0.18	0.18	NA	0.17	NA	0.16	NA
n-3 FA	0.09	0.08	0.10	0.08	0.11	0.10	0.27	0.23	0.23	0.40	0.23	0.38	0.22	0.38
n-6 FA	0.26	0.26	0.28	0.28	0.30	0.30	0.37	0.12	0.12	0.19	0.13	0.20	0.15	0.22

<sup>1</sup>The best methods chosen for C16:1 cis, C18:1 cis, C18:3 cis-9, cis-12, cis-15, and C18:2 cis-9, trans-11 were methods 5, 4, 6, and 4, respectively.

<sup>2</sup>Reference fat = fat content provided by the mid-infrared spectrometer and corrected if needed by the bias and the slope.

<sup>3</sup>Fat content predicted using method 4.

<sup>4</sup>Fat content predicted using method 3, the best method for fat prediction.

<sup>5</sup>Prediction = content of FA in fat predicted using the equation built from the method 4, the best for the majority of FA.

<sup>6</sup>FA content in milk predicted using method 4.

<sup>7</sup>FA content in milk using the best method for each FA based on the result of Table 3.

<sup>8</sup>NA = FA for which method 4 is the best.

Method 3 was the best predictor for the majority of FA (data not shown). It can be easily explained by the introduction of new samples in the calibration set. As explained previously, the samples included in the validation (and thus used in the calibration set) came mainly from Ireland and Scotland. The introduction of more spectra coming from other regions in the calibration set decreases the interest to use the repeatability file. The larger data set resulted in an improved ability of MIR spectrometry to quantify the studied polyunsaturated FA.

The second part of [Table 2.4](#) shows the results for the new FA calibration equations built using outlier detection and method 3. The number of factors per equation ranged from 7 to 15. Following the removal of outliers, the RPD calculated for all/most FA was higher, indicating that the removal of outliers allows improvement of the accuracy. The same equations than the one isolated previously from [Table 2.2](#) for potential quantitative and qualitative applications can be considered from [Table 2.4](#) and showed a greater RPD threshold (4 instead of 3). Based on a RPD threshold equal to 3, three additional equations can be used: C4:0, C18:0 and C18:1 equations. With  $R^2_{cv}$  greater than 0.75, the MIR predictions of POLY, Omega-3, and the group of branched FA could be used to make a screening of the cow population. So for all of these equations, the confidence interval for 95% of the population will be 1.96 times SECV.

## 2.6. Conclusions

This study compared 6 different methods to develop calibration equations for predicting the contents of FA in milk by using an innovative approach to collect milk samples combining multiple breeds, multiple countries, and multiple production systems. Globally, the methods using a first derivative pre-treatment of the spectral data gave the best predictions. The best accuracy was observed for the infrared predictions of C4:0, C6:0, C8:0, C10:0, C12:0, C14:0, C16:0, C18:0, C18:1 *trans*, C18:1 *cis-9*, C18:1 *cis* and for the majority of studied groups of FA studied in milk. With  $R^2_{cv}$  greater than 0.75, the MIR predictions of POLY, Omega-3, and the group of branched FA could be used to make a screening of the cow population. As already published, the infrared predictions of FA in fat are less accurate and no better results were obtained by using the developed equations for fat and for FA in milk without corrections of slope and bias based on milk samples with known contents of FA. The large number of samples, specifically chosen to cover the natural variation in milk FA present, coupled with the high accuracy of selection for some FA indicate the usefulness of these equations in milk payment system (if  $R^2_v$  is greater than 95%) and animal breeding (if  $R^2_{cv}$  is greater than 75%).



Table 2.4. Accuracy of mid-infrared predictions for fatty acid (FA) in milk (g/dL of milk) using the entire data set<sup>1</sup>

Constituent (g/dL of milk)	With supposed outliers						Without supposed outliers (method 3)							
	Method	n	Mean	SD	SECV	R <sup>2</sup> <sub>cv</sub>	RPD	n	Factor	Mean	SD	SECV	R <sup>2</sup> <sub>cv</sub>	RPD
Fat	3	517	3.65	1.15	0.04	1.00	25.6	498	15	3.59	1.10	0.04	1.00	30.4
C4:0	4	517	0.10	0.03	0.01	0.89	3.0	490	13	0.10	0.03	0.01	0.94	4.1
C6:0	3	517	0.067	0.024	0.01	0.95	4.3	492	11	0.067	0.024	0.00	0.97	5.7
C8:0	4	517	0.042	0.016	0.00	0.93	3.8	490	14	0.042	0.016	0.00	0.97	6.1
C10:0	4	517	0.10	0.04	0.01	0.92	3.6	495	15	0.10	0.04	0.01	0.96	5.1
C12:0	4	517	0.12	0.05	0.01	0.92	3.6	495	15	0.12	0.05	0.01	0.96	5.2
C14:0	4	517	0.40	0.13	0.03	0.95	4.3	494	14	0.39	0.13	0.02	0.97	5.4
C14:1	4	517	0.037	0.016	0.01	0.61	1.6	493	11	0.036	0.014	0.01	0.68	1.8
C16:0	4	517	1.05	0.39	0.10	0.93	3.8	494	13	1.02	0.37	0.08	0.95	4.6
C16:1 <i>cis</i>	5	517	0.068	0.029	0.02	0.64	1.7	493	9	0.065	0.024	0.01	0.71	1.9
C17:0	2	517	0.026	0.010	0.01	0.61	1.6	484	11	0.026	0.009	0.00	0.89	3.1
C18:0	4	517	0.38	0.18	0.06	0.88	2.8	492	14	0.37	0.17	0.05	0.90	3.2
C18:1 <i>trans</i>	4	517	0.14	0.07	0.03	0.84	2.5	502	15	0.14	0.07	0.02	0.88	2.9
C18:1 <i>cis-9</i>	4	517	0.74	0.29	0.06	0.95	4.5	494	15	0.73	0.28	0.05	0.97	5.9
Total C18:1 <i>cis</i>	4	517	0.80	0.31	0.07	0.95	4.7	495	15	0.79	0.30	0.05	0.97	6.0
Total C18:2	6	517	0.085	0.029	0.02	0.70	1.8	503	7	0.084	0.028	0.01	0.73	1.9
C18:2 <i>cis-9, cis-12</i>	5	517	0.054	0.025	0.01	0.71	1.9	502	10	0.053	0.024	0.01	0.74	2.0
C18:3 <i>cis-9, cis-12, cis-15</i>	6	517	0.021	0.011	0.01	0.60	1.6	489	8	0.021	0.009	0.01	0.71	1.8
C18:2 <i>cis-9, trans-11</i>	4	517	0.037	0.023	0.01	0.63	1.6	488	12	0.036	0.021	0.01	0.74	2.0
Saturated FA	3	517	2.42	0.82	0.08	0.99	10.8	496	14	2.40	0.80	0.05	1.00	15.7
Monounsaturated FA	4	517	1.06	0.37	0.06	0.97	5.8	491	15	1.06	0.37	0.04	0.99	8.9
Polyunsaturated FA	3	517	0.16	0.05	0.02	0.81	2.3	499	11	0.16	0.05	0.02	0.85	2.6
Unsaturated FA	3	517	1.23	0.41	0.07	0.97	6.2	492	15	1.22	0.41	0.04	0.99	9.6
Short-chain FA	3	517	0.32	0.11	0.02	0.95	4.5	486	12	0.31	0.11	0.02	0.98	6.7
Medium-chain FA	4	517	1.80	0.62	0.12	0.96	5.1	496	14	1.78	0.60	0.09	0.98	6.5
Long-chain FA	4	517	1.53	0.57	0.12	0.96	4.8	495	14	1.52	0.57	0.09	0.98	6.5
Branched-chain FA	4	517	0.09	0.03	0.02	0.73	1.9	492	12	0.09	0.03	0.01	0.83	2.4
n-3 FA	2	517	0.029	0.015	0.01	0.55	1.5	485	8	0.028	0.012	0.01	0.75	2.0
n-6 FA	6	517	0.10	0.03	0.02	0.72	1.9	504	7	0.10	0.03	0.02	0.74	2.0

<sup>1</sup>SECV = standard error of cross-validation; R<sup>2</sup><sub>cv</sub> = cross-validation coefficient of determination; RPD = ratio of SD of GC data from the calibration set to the SECV.

## 2.7. Acknowledgments

This research receives a financial support from the European Commission, Directorate-General for Agriculture and Rural Development, under Grant Agreement 211708 and from the Commission of the European Communities, FP7, KBBE-2007-1. This paper does not necessarily reflect the view of these institutions and in no way anticipates the Commission's future policy in this area. H el ene Soyeurt, a Post-doctoral Researcher, and Nicolas Gengler, a Research Associate, acknowledge also the support of the National Fund for Scientific Research (Brussels, Belgium) for these positions and for the additional grants 2.4507.02F (2) and F.4552.05 (2.4.623.08.F). Ministry of Agriculture of the Walloon Region of Belgium [Service public de Wallonie, Direction g en erale de l'Agriculture, des Ressources naturelles et de l'Environnement, Direction de la Recherche] is acknowledged for their financial support through the project NovaUdderHealth (D31-1207). The authors acknowledge the technical support provided by the Walloon Breeding Association (Ciney, Belgium), Convis Herdbuch (Ettelbruck, Luxembourg) and the milk lab (Comit e du Lait, Battice, Belgium).

## 2.8. References

- Arnould V.M.-R., and Soyeurt H.. 2009. Genetic variability of milk fatty acids. *J. Appl. Genet.* 50:29-39.
- Coffey M.P., Simm G., Oldham J.D., Hill W.G., and Brotherstone S., 2004. Genotype and diet effects on energy balance in the first three lactations of dairy cows. *J. Dairy Sci.* 87:4318-4326.
- Coleman J., Pierce K.M., Berry D.P., Brennan A., and Horan B., 2009. The influence of genetic selection and feed system on the reproductive performance of spring-calving dairy cows within future pasture-based production systems. *J. Dairy Sci.* 92:5258-5269.
- Collomb M. and Bühler T., 2000. Analyse de la composition en acides gras de la graisse de lait. *Mitt. Lebensm. Hyg.* 91:306-332.
- Dal Zotto R., De Marchi M., Cecchinato A., Penasa M., Cdoro M., Carnier P., Gallo L., and Bittante G.. 2008. Reproducibility and repeatability of measures of milk coagulation properties and predictive ability of mid-infrared reflectance spectroscopy. *J. Dairy.* 91:4103-4112.
- Hruschka W.R., 1987. Data analysis : wavelength selection methods. In: Williams, P., Norris, K. (Eds), *Near-Infrared Technology in the Agricultural and Food Industries*. American Association of Cereal Chemists, St Paul, MN, pp. 35-55.
- ISO, 2001. Milk and milk products. Extraction methods for lipids and liposoluble compounds. ISO 14156 :2001. IDF 172:2001.
- ISO, 2002. Milk fat. Preparation of fatty acid methyl esters. ISO 15884:2002. IDF 182:2002.
- Jensen R.G., 1995. *Handbook of milk composition*. Academic Press.
- Prendiville R., Lewis E., Pierce K.M., and Buckley F., 2010. Comparative grazing behavior of lactating Holstein-Friesian, Jersey, and Jersey × Holstein-Friesian dairy cows and its association with intake capacity and production efficiency. *J. Dairy Sci.* 93:764-774.



- Rutten M.J.M., Bovenhuis H., Hettinga K.A., Van Vanlenberg H.J.F., and Van Arendonck 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., Dardenne P., Dehareng F., Lognay G., Veselko D., Marlier M., Bertozzi C., Mayeres P., and Gengler N., 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 89: 3690-3695.
- Soyeurt H., Dehareng F., Mayeres P., Bertozzi C., and Gengler N., 2008a. Variation of delta9-desaturase activity in dairy cattle. *J. Dairy Sci.* 91:3211-3224.
- Soyeurt H., Dardenne P., Dehareng F., Bastin C., and Gengler N., 2008b. Genetic parameters of saturated and monounsaturated fatty acid content and the ratio of saturated to unsaturated fatty acids in bovine milk. *J. Dairy Sci.* 91:3611-3626.
- Soyeurt H., Misztal I., and Gengler N., 2010. Genetic variability of milk components based on mid-infrared spectral data. *J. Dairy Sci.* 93:1722-1728.
- Westerhaus M.O., 1990. Improving repeatability of NIR calibrations across instruments. Pages 671-674 in Proc. 3rd Int. Conf. Near Infrared Spectroscopy, Brussels, Belgium. Agricultural Research Centre, Gembloux, Belgium.
- Williams C.M., 2000. Dietary fatty acids and human health. *Ann. Zootech.* 49:165-180.
- 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.
- Williams P.C., 2007. Near-infrared technology – Getting the best out of light. A short course in the practical implementation of Near-Infrared Spectroscopy for the user. PDK projects, Inc., Nanaimo, Canada.



# Chapitre 3:

---

## **Potential use of milk mid-infrared spectra to predict individual methane emission of dairy cows**

*Adapté de: 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(10), 1694-1701.*

*\* co-premiers auteurs*



### 3.1. Points d'attention

Ce troisième chapitre est une bonne illustration des difficultés rencontrées lorsque l'on cherche à prédire un paramètre/phénotype indirect. En effet, dans le cas du méthane (CH<sub>4</sub>), il n'y a nulle présence de CH<sub>4</sub> dans le lait. Dès lors, seules les intercorrélations existantes entre la composition fine du lait et le CH<sub>4</sub> sont recherchées. Cette première étude exploratoire sur la faisabilité de prédire les émissions des vaches laitières à partir du FT-MIR a permis d'adapter le protocole d'échantillonnage aux contraintes de la méthode de référence et d'étudier le biais temporel existant entre la synthèse des acides gras dans la glande mammaire et la production d'acides gras volatiles (AGV) dans le rumen.

En effet, contrairement au deuxième chapitre, un échantillonnage basé sur la variabilité spectrale n'est pas réalisable ici. En effet, la méthode de référence utilisée pour mesurer le CH<sub>4</sub> éructé limite grandement les moyens d'action par son coût et son côté invasif pour l'animal, vu qu'un bolus doit être ingéré par l'animal afin de pouvoir mesurer le gaz traceur lors de chaque éructation. Il y a donc des contraintes techniques, financières et potentiellement éthiques qui limitent le nombre d'individus sur lesquels on peut effectuer la quantification du CH<sub>4</sub>. Pour limiter le nombre d'individus et les périodes de mesure, nous avons cherché à obtenir un set de données de référence présentant la variation la plus large possible de CH<sub>4</sub> en jouant essentiellement sur l'utilisation de régimes alimentaires extrêmes et en proposant un protocole de prise de mesure du CH<sub>4</sub> journalier maximisant la certitude de la mesure.

Concernant le biais temporel, celui-ci s'explique par notre hypothèse de départ qui stipule que des liens existent entre la composition en AG du lait et le CH<sub>4</sub> éructé par l'animal. Or, il existe un biais temporel entre les AGV et la composition en AG du lait. En effet, on va devoir attendre que les AGV produits dans le rumen se déplacent au sein de la glande mammaire pour synthétiser les AG de novo (AG avec une chaîne composée entre 4 et 14 carbones). Cette désynchronisation temporelle a été étudiée dans ce chapitre afin d'en évaluer l'impact sur la qualité de la prédiction finale.

Malgré des performances de prédiction encourageantes, on a constaté par la suite, lorsqu'on appliquait cette équation sur une base de données spectrales importante comme celle relative au contrôle laitier wallon, que les courbes d'émissions de CH<sub>4</sub> au cours de la lactation ne correspondaient pas à celles rapportées par la littérature (Garnsworthy et al., 2012). Ce manque de robustesse s'est expliqué par un manque de considération de la variation intra-lactationnelle du CH<sub>4</sub>. Ainsi, cette découverte a permis de souligner que pour certains modèles, les relations entre le caractère à prédire et les variables qui sont utilisées peuvent être influencées par d'autres facteurs

amenant à avoir des coefficients de régression non constants en fonction des sources de variation. Pour solutionner ce problème, lors des recherches postérieures, on a d'abord augmenté la taille du jeu de données en maximisant la variabilité du CH<sub>4</sub> liée à la race ou encore au régime alimentaire, mais on a également accentué les prises de mesure chez des vaches en début de lactation. Cette démarche a été couplée à l'inclusion d'une modification fondamentale au sein de la structure du modèle prédictif. On a ainsi proposé d'ajouter à la régression des polynômes liés au jour de lactation pour permettre aux coefficients de régression de varier en fonction du stade de lactation (Vanlierde et al., 2015). Ceci a permis de prendre notamment en compte la variation du processus de la lipomobilisation chez la vache au cours d'une même lactation, impliquant des relations changeantes entre le CH<sub>4</sub> et les AG du lait.

En outre, vu la complexité à acquérir les données de référence requises pour développer un modèle robuste, des collaborations internationales ont été mises en place. Celles-ci ont été facilitées grâce à la publication des travaux de ce chapitre. Cet échange de données issues de spectromètres FT-MIR différents a mis en évidence l'importance d'utiliser un processus de standardisation des données spectrales (Grelet et al., 2015) comme expliqué dans le point 2.1. En plus de l'accroissement du nombre de données phénotypiques et de l'intégration du jour de lactation, nous avons également par la suite testé d'autres modèles intégrant d'autres variables explicatives facilement disponibles (quantités de lait produit, parité, race), celles-ci ont encore amélioré la qualité des modèles (Vanlierde et al., 2021).

Enfin, l'application de l'équation CH<sub>4</sub> sur de larges bases de données a permis d'étudier pour la première fois à large échelle les sources de variation du CH<sub>4</sub> liées à la parité de la vache et aux variations annuelles, saisonnières et régionales (Vanlierde et al., 2016). Ces prédictions ont d'ailleurs le potentiel d'être utilisées pour effectuer un dépistage des bovins laitiers faibles et hauts producteurs de CH<sub>4</sub> et pour estimer les premières héritabilités et corrélations phénotypiques et génétiques (Vanrobays et al., 2016 ; Kandel et al., 2017).

Ma contribution dans cet article en tant que co-premier auteur a visé à formuler l'hypothèse de base sur base de la recherche bibliographique et à écrire puis gérer le projet de recherche qui a mené notamment à cet essai. J'ai également contribué à la mise en place du plan expérimental, à la mise au point de la méthode de référence et au protocole d'analyse du CH<sub>4</sub>. J'ai enfin contribué à la réalisation de l'analyse des données, à la rédaction, et à répondre aux questions des examinateurs de cette publication.

### ***3.1.1 Références***

- Garnsworthy P.C., Craigon J., Hernandez-Medrano J.H. and Saunders N., 2012. Variation among individual dairy cows in methane measurements made on farm during milking. *J. Dairy Sci.*, 95 (2012), pp. 3181-3189.
- Grelet C., Fernandez 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(4), 2150-2160.
- Kandel PB, Vanrobays ML, Vanlierde A, Dehareng F, Froidmont E, Gengler N, and Soyeurt H., 2017. Genetic parameters of mid-infrared methane predictions and their relationships with milk production traits in Holstein cattle. *J Dairy Sci.*, 100(7):5578-5591.
- 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, pp. 5740-5747
- 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*, 56, 258-264.
- Vanlierde, A., Dehareng, F., Gengler, N., Froidmont, E., McParland, S., Kreuzer, M., Bell, M., Lund, P., Martin, C., Kuhla, B. and Soyeurt, H., 2021. Improving robustness and accuracy of predicted daily methane emissions of dairy cows using milk mid-infrared spectra. *J Sci Food Agric*, 101: 3394-3403.
- Vanrobays ML, Bastin C, Vandenplas J, Hammami H, Soyeurt H, Vanlierde A, Dehareng F, Froidmont E, Gengler N., 2016. Changes throughout lactation in phenotypic and genetic correlations between methane emissions and milk fatty acid contents predicted from milk mid-infrared spectra. *J Dairy Sci.*, 99(9):7247-7260.

## 3.2. Abstract

This study investigates the feasibility to predict individual methane (CH<sub>4</sub>) emissions from dairy cows using milk mid-infrared (MIR) spectra. To have a large variability of milk composition, two experiments were conducted on 11 lactating Holstein cows (two primiparous and nine multiparous). The first experiment aimed to induce a large variation in CH<sub>4</sub> emission by feeding two different diets: the first one was mainly composed of fresh grass and sugar beet pulp and the second one of maize silage and hay. The second experiment consisted of grass and corn silage with cracked corn, soybean meal and dried pulp. For each milking period, the milk yields were recorded twice daily and a milk sample of 50 ml was collected from each cow and analyzed by MIR spectrometry. Individual CH<sub>4</sub> emissions were measured daily using the sulfur hexafluoride method during a 7-day period. CH<sub>4</sub> daily emissions ranged from 10.2 to 47.1 g CH<sub>4</sub>/kg of milk. The spectral data were transformed to represent an average daily milk spectrum (AMS), which was related to the recorded daily CH<sub>4</sub> data. By assuming a delay before the production of fermentation products in the rumen and their use to produce milk components, five different calculations were used: AMS at days 0, 0.5, 1, 1.5 and 2 compared with the CH<sub>4</sub> measurement. The equations were built using Partial Least Squares regression. From the calculated R<sup>2</sup><sub>cv</sub>, it appears that the accuracy of CH<sub>4</sub> prediction by MIR changed in function of the milking days. In our experimental conditions, the AMS at day 1.5 compared with the measure of CH<sub>4</sub> emissions gave the best results. The R<sup>2</sup> and s.e. of the cross-validation were equal to 0.79 and 5.14 g of CH<sub>4</sub>/kg of milk. The multiple correlation analysis performed in this study showed the existence of a close relationship between milk fatty acid (FA) profile and CH<sub>4</sub> emission at day 1.5. The lower R<sup>2</sup> (R<sup>2</sup> = 0.76) obtained between FA profile and CH<sub>4</sub> emission compared with the one corresponding to the obtained calibration (R<sup>2</sup><sub>c</sub> = 0.87) shows the interest to apply directly the developed CH<sub>4</sub> equation instead of the use of correlations between FA and CH<sub>4</sub>. In conclusion, our preliminary results suggest the feasibility of direct CH<sub>4</sub> prediction from milk MIR spectra. Additional research has the potential to improve the calibrations even further. This alternative method could be useful to predict the individual CH<sub>4</sub> emissions at farm level or at the regional scale and it also could be used to identify low-CH<sub>4</sub>-emitting cows.

## 3.3. Introduction

Livestock farming is right at the forefront of climate change issues. Agriculture is regarded as the biggest producer of anthropogenic methane (CH<sub>4</sub>), mainly through ruminant gas emissions (Food and Agriculture Organization (FAO), 2010). Enteric CH<sub>4</sub> is produced by ruminants during the microbial digestion of feed in the rumen. CH<sub>4</sub> contributes widely to global warming, accounting for 52% of the greenhouse gases emissions in both developing and developed countries (FAO, 2010). CH<sub>4</sub> absorbs 25 times as much infrared radiation than carbon dioxide (CO<sub>2</sub>), enlarging global warming problems (Fuglestad, 2009). In addition to those environmental



concerns, the eructed CH<sub>4</sub> induces a significant loss of gross energy intake for the animal. This gas loss corresponds between 2% and 12% of gross energy intake (Johnson and Johnson, 1995; Jouany, 2008; Beauchemin et al., 2009). These losses are mainly explained by the animal (age and species), the diet (intake level, composition) and the level of milk production (Vermorel, 1995). However, individual variations between animals of the same breed were reported in the range from 30% to 60% in production of CH<sub>4</sub>/unit of feed ingested when these animals received the same diet at equivalent ingestion and production levels (Lassey et al., 1997). They are related to different rumen microbial populations, coming from different ruminal kinetics, such as salivation, rumination (Demeyer and Fievez, 2000) and different feeding behaviors, indicating interaction between animals and their microbiota. Recent results suggest, therefore, a genetic component in variation of individual CH<sub>4</sub> production (Martin et al., 2010). However, most studies showed that ruminants that are genetically different for feed efficiency (e.g. Hegarty et al., 2007) and productivity showed reduced daily CH<sub>4</sub> production (e.g. Bell et al., 2010). Given the complexity of this variation, large-scale studies should be conducted, but no in vivo method to measure CH<sub>4</sub> emissions in ruminants is sufficiently accurate to be applied on a large number of animals (Clark, 2010). Indeed, the available methods, such as the use of calorimeter or CH<sub>4</sub> chambers, the tunnel system and the sulfur hexafluoride (SF<sub>6</sub>) tracer technique, are not easy to use on farm and are laborious to estimate individual CH<sub>4</sub> production of many cows in a short period. Therefore, a simpler and appropriate method is required to predict individual CH<sub>4</sub> emissions. For this, the method investigated here was based on the fact that CH<sub>4</sub> synthesis depends on the kind of rumen fermentation, which also influences many other parameters such as milk composition (Ørskov et al., 1969).

Carbohydrates are the most important source of energy and the primary precursors of fat and lactose in cow's milk. The end products of carbohydrate degradation are the volatile fatty acids (VFA), CO<sub>2</sub>, CH<sub>4</sub> and H<sub>2</sub>. Demeyer and Fievez (2000) have developed an equation that shows clearly the relationship between VFA and CH<sub>4</sub>. During the fermentation process, acetate and butyrate promote CH<sub>4</sub> production, whereas propionate maintains a competitive role for the use of hydrogen. According to Miettinen and Huhtanen (1996), an increased butyrate/propionate ratio decreases the lactose content and increases the fat content in milk and the rumen CH<sub>4</sub> synthesis. Thus, a relationship among CH<sub>4</sub>, lactose and fat contents could be assumed as suggested by Vlaeminck and Fievez (2005). In addition, the rumen VFA composition also influences milk fat composition. In ruminants, the milk fatty acids (FA) come from two sources, uptake from circulation and de novo synthesis within the mammary gland. Approximately one half of the milk FA (molar percent) is derived from de novo synthesis, based on acetate and butyrate (Bauman and Grinari, 2003). Short and medium-chain FA (4 to 14 carbons) arise totally from de novo synthesis. Long-chain FA (>16 carbons) are collected in the circulating lipids, and FA of 16 carbons depend on these two sources (Bauman and Grinari, 2003). Thus, the proportions of different FA reflect the ruminal fermentation via VFA and then CH<sub>4</sub>

production (Weill et al., 2008 and 2009). Moreover, several authors have already reported, for different diets, relationships between milk FA (measured by gas chromatography (GC)) and CH<sub>4</sub> emission of dairy cows (Chilliard et al., 2009; Delfosse et al., 2010; Dijkstra et al., 2010). These results suggest that, in general, CH<sub>4</sub> emissions are linked to milk composition. As the mid-infrared (MIR) spectrum reflects the milk composition (Soyeurt et al., 2006 and 2011), it is logical to postulate that the MIR spectrum could predict directly the individual emission of CH<sub>4</sub>. Therefore, the aim of this study was to have a first evaluation of the potential use of milk spectra obtained by Fourier Transform Infrared analysis (FTIR) apparatus directly to predict the quantities of CH<sub>4</sub> eructated by individual dairy cows. The development of a MIR CH<sub>4</sub> equation will permit to predict and therefore to study at large scale the CH<sub>4</sub> emission of dairy cows, because the MIR technology is already implemented in milk laboratories to quantify the major milk components used for the milk payment and the routine milk recording.

### 3.4. Material and methods

#### 3.4.1. *Animals and diets*

Two experiments were carried out with different experimental conditions in order to maximize the variability of individual CH<sub>4</sub> emission needed to establish a robust calibration model.

In the first experiment, two primiparous and six multiparous (three of second lactation and three of third lactation) lactating Holstein cows were divided into two groups of four cows each. At the beginning of the trial, the groups were similar in terms of lactation stage (180 days in milk) and of average milk production ( $17.4 \pm 3.9$  kg/day). Two isoenergetic experimental diets (17 kVEM) were offered according to a 2x2 crossover design. Experimental diets were fed twice daily at 0900 and 1600 h as total mixed rations (TMR). The diets were fed at an intake level of 18.5 kg/day and aimed at producing 20 kg/day of milk. Diet 1 consisted of: freshly cut pasture grass (third cutting, *Holcus lanatus*), 550 g/kg dry matter (DM); dried beet pulp, 200 g/kg DM; soybean meal, 150 g/kg DM; and soybean hulls, 100 g/kg DM. Diet 2 consisted of: corn silage, 400 g/kg DM; meadow hay, 200 g/kg DM; cracked corn, 130 g/kg DM; rapeseed meal, 150 g/kg DM; palm meal, 55 g/kg DM; soybean meal, 55 g/kg DM; and a 50 : 50 mix of coconut and flaxseed oil, 10 g/kg DM. Both diets contained a mixture of vitamins and minerals.

In the second experiment, three multiparous (two of second lactation and one of fourth lactation) lactating Holstein cows with a similar milk production ( $26.2 \pm 1.9$  kg/day) were fed a same basal diet, with an intake level of 20.3 kg/day. This diet (TMR) consisted of grass silage, 520 g/kg DM; corn silage, 130 g/kg DM; cracked corn, 130 g/kg DM; soybean meal, 110 g/kg DM; and dried beet pulp, 110 g/kg DM.

For both the experiments, the adaptation period was 21 days and milk and CH<sub>4</sub> samples were then collected from days 22 to 28. Fresh water was available at all times.

**Table 3.1.** Estimated statistical parameters for each calibration equation that estimated the concentrations of fatty acid in milk (g / dl milk) used in this study.

Fatty acids (g/dL of milk)	n	Mean	SD	SEC	R <sup>2</sup> <sub>C</sub>	SECV	R <sup>2</sup> <sub>CV</sub>	RPD
C4:0	137	0.11	0.03	0.01	0.82	0.01	0.73	1.93
C6:0	138	0.08	0.02	0.01	0.86	0.01	0.82	2.32
C8:0	131	0.05	0.01	0.00	0.86	0.01	0.81	2.33
C10:0	137	0.11	0.04	0.01	0.90	0.02	0.84	2.52
C12:0	138	0.13	0.05	0.02	0.88	0.02	0.84	2.48
C14:0	136	0.45	0.12	0.03	0.93	0.04	0.88	2.90
C14:1	138	0.04	0.02	0.01	0.66	0.01	0.58	1.52
C16:0	139	1.19	0.35	0.09	0.93	0.11	0.89	3.06
C16:1	143	0.08	0.04	0.02	0.82	0.02	0.77	2.06
C17:0	138	0.03	0.01	0.00	0.84	0.01	0.72	1.89
C18:0	138	0.39	0.15	0.07	0.79	0.07	0.76	2.05
C18:1 <i>trans</i>	138	0.10	0.04	0.02	0.67	0.03	0.55	1.49
C18:1 <i>cis-9</i>	138	0.80	0.42	0.08	0.97	0.09	0.95	4.50
Total C18:1 <i>cis</i>	138	0.87	0.43	0.07	0.98	0.08	0.97	5.38
Total C18:2	143	0.09	0.03	0.01	0.70	0.02	0.6	1.58
C18:2 <i>cis-9,cis-12</i>	142	0.06	0.02	0.01	0.77	0.01	0.7	1.83
C18:3 <i>cis-9,cis-12,cis-15</i>	138	0.02	0.01	0.01	0.52	0.01	0.4	1.29
C18:2 <i>cis-9,trans-11</i>	139	0.03	0.02	0.01	0.5	0.01	0.36	1.24
Saturated FA	132	2.72	0.72	0.06	0.99	0.07	0.99	10.57
Monounsaturated FA	135	1.12	0.46	0.06	0.98	0.07	0.98	6.45

SD : Standard deviation; SEC: Standard error of calibration; R<sup>2</sup><sub>C</sub> : Calibration coefficient of determination; SECV: Standard error of cross-validation; R<sup>2</sup><sub>CV</sub> : Cross-validation coefficient of determination; RPD : Ratio of standard error of cross validation to standard deviation

All the procedures used involving the animals were approved by the regional ethics committee in animal experimentation (protocol CRAW09/01).

### 3.4.2. Sampling and analyses

Milk yield was recorded daily at each sampling period at 0730 and 1630 h. A 50ml aliquot of milk, containing sodium azide (0.32 g/l), was stored at 4°C until infrared analyses. The sample stored at 4°C was analyzed by a FTIR Lactoscope spectrometer (Delta Instruments, Drachten, The Netherlands). The Lactoscope works within 925 to 3000/cm and uses an interferometer. This instrument gave the spectral data and the direct measurement of milk components such as lactose, protein, fat, and non-protein nitrogen (considered to be essentially urea). Similar FA composition predictions were also obtained based on equations specific to this machine. Table 3.1 presents the statistical parameters of some of these equations. For each FA equation, the standard error of cross-validation (SECV) given by the FTIR Lactoscope equations are included between those given by Soyeurt et al. (2006) and those of Soyeurt et al. (2011). A total of 154 samples were collected and analyzed (spectra were in triplicate).

### 3.4.3. Quantification of CH<sub>4</sub>

During the 7-day period, the SF<sub>6</sub> gas tracer technique was used according to the method described by Martin et al. (2008) in order to measure the individual production of enteric CH<sub>4</sub> by the studied cows. In brief, a calibrated permeable tube containing ultra pure SF<sub>6</sub> was placed in the rumen of each cow before the experimental period. The average release rate of SF<sub>6</sub> from the tubes was  $1189.5 \pm 168.6$  ng/min.

A representative of breath gas sample, containing respired and eructated gas, was collected through a capillary tube situated between the nostril and the mouth of each animal, thanks to an halter. This gas sample was stored in a canister placed near the animal (more or less one meter) in a way that the animal could not touch it. The canister was changed daily after morning feeding and the CH<sub>4</sub> and SF<sub>6</sub> concentrations were analyzed using GC (Martin et al., 2008).

A GC (Varian-Chrompack, CP-9003, Les Ulis, France) fitted with an electron capture detector (Perkin Elmer instruments; Autosystem XL, Courtaboeuf, France) and with a flame ionization detector was used to determine the concentrations of SF<sub>6</sub> and CH<sub>4</sub>, respectively. The samples were run on GCs equipped either with a Molecular Sieve 0.5 nm column (3m x 3.2mm i.d.) maintained at 50°C for the SF<sub>6</sub> or with a Porapak N 80-100 mesh column (3m x 3.2mm i.d.) maintained at 40°C for the CH<sub>4</sub>. The flow rate of the carrier gas was 30 ml/min of N<sub>2</sub> for SF<sub>6</sub> and 40 ml/min of He for CH<sub>4</sub>. Chromatographic analyses were performed after calibration with standard gases (Air Liquide, Mitry-Mory, France) for SF<sub>6</sub> (55 and 195 ppt) and CH<sub>4</sub> (100 ppm). To determine the released amount of CH<sub>4</sub> ( $Q_{CH_4}$ ), the concentration of SF<sub>6</sub> and CH<sub>4</sub> ( $C_{CH_4}$  and  $C_{SF_6}$ ) in the canister and the predetermined released rate of SF<sub>6</sub> ( $Q_{SF_6}$ ) were used. In parallel to each daily measurement, the concentration of the atmospheric air ( $C_{CH_4}^b$  and  $C_{SF_6}^b$ ) determined with another canister was subtracted (Johnson et al., 1994):

$$Q_{CH} = \frac{C_{CH_4} - C_{CH_4}^b}{C_{SF_6} - C_{SF_6}^b} Q_{SF_6} \frac{MW_{CH_4}}{MW_{SF_6}}$$

where  $MW_{CH_4}$  and  $MW_{SF_6}$  are the molecular weights of CH<sub>4</sub> and SF<sub>6</sub>, respectively.

The CH<sub>4</sub> emission was therefore expressed in g of CH<sub>4</sub>/day and re-expressed later in g of CH<sub>4</sub>/kg of milk, thanks to the recorded milk yield because the global daily CH<sub>4</sub> production is linked to productivity.

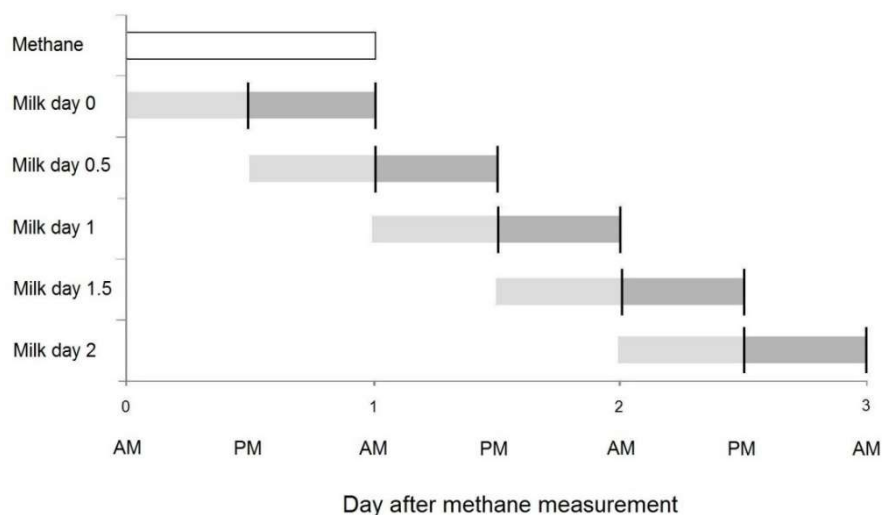
#### ***3.4.4. Spectral data treatment***

Each cow was milked twice daily, providing two individual samples, which were analyzed by MIR, thus two spectra, whereas individual CH<sub>4</sub> emission was measured once daily. Therefore, to build the calibration model, the recorded spectral data were transformed to represent one daily spectrum related to one daily CH<sub>4</sub> record. The methodology used to create the average milk spectra (AMS) was tested by comparing MIR predictions obtained from AMS and the real spectral data obtained from the infrared analysis of a representative daily milk sample (i.e. 50% of morning and 50% of evening milk samples). This ratio was chosen because it corresponds to the averaging used during the majority of milk-recording process. Alternatively, the weighted average was also tested. It corresponded to the average of the two milk spectra of the day in proportion to the amount of milk produced by the cow in each respective milking (morning and evening). The reason for this was that a weighted average represented better the biological background of the process.

A dynamic relationship between milk composition (and therefore spectral data) and CH<sub>4</sub> was considered by assuming that there is a delay between the production of fermentation products and their use to produce milk components. Consequently, different ways exist to define the AMS. Indeed, the spectral data used to create the AMS can be generated by the milk analysis of samples collected at different times compared with the time of CH<sub>4</sub> measurement. The AMS were calculated from five methods averaging always two samples, which are illustrated in Figure 3.1:

- the same day of CH<sub>4</sub> measurement (day 0),
- evening of the same day and morning next day (day 0.5),
- next day (day 1),
- evening next day and morning 2 days later (day 1.5), and
- 2 days later (day 2).

Always two samples were averaged to reflect a whole test-day in the milk as whole day was also represented in the CH<sub>4</sub> measurements.



**Figure 3.1.** Possibility to average the spectra milk in function of methane day collect. Vertical lines corresponds to milking time; surface shaded to milk production.

Given the stability of diets and the high degree of autocorrelation that can be therefore assumed across values for the five methods, the objective of this was not to infer a definitive response what the delay between the production of fermentation products and their use to produce milk components, but to find retrospectively which definition of the AMS gives the best predictions under the condition of this study.

### 3.4.5. Calibration model

As mentioned previously, a calibration set with a large spectral variability is required to develop robust calibration models. Therefore, a principal component analysis was carried out on available spectra at day 0 in order to illustrate the spectral variation observed from the collected milk samples.

The calibration models were developed by Foss WINISI 4 software from the recorded spectral and  $\text{CH}_4$  data using Partial Least Squares regressions. The spectral regions used for that were: 972 to 1589/cm, 1720 to 1782/cm and 2746 to 2970/cm. No additional pretreatment was applied on spectral data because with a derivative, results were not significantly better probably because of the fact that only one spectrometer was used in this experiment. The number of factors used was determined by a full cross-validation (with  $n$  observations, create  $n$  models by removing  $n$  time one sample that is predicted by the others  $n-1$ ), which was also used to estimate its robustness. The accuracy of the resulting calibration models was evaluated by calculating the calibration coefficient of determination ( $R^2_c$ ), cross-validation coefficient of determination ( $R^2_{cv}$ ), standard error of calibration (SEC) and the standard error of cross-validation (SECV). The predictability of the calibration was evaluated through the ratio of performance to deviation (RPD;  $\text{SD}/\text{SECV}$  where SD was the standard deviation of the  $\text{CH}_4$  measures). The RPD should be as high as

possible; values between 5 and 10 are adequate for quality control, and values > 2.5 are satisfactory for screening breeding programs (Williams and Sobering, 1993).

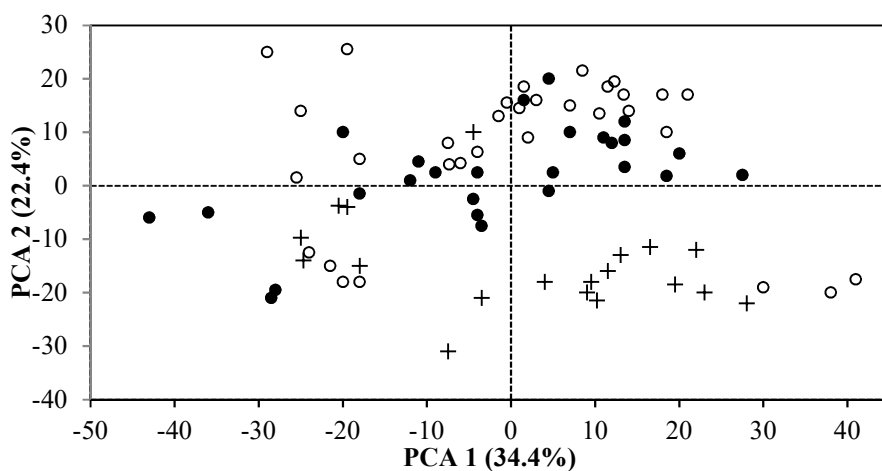
### 3.4.6. Correlations with milk components

To show the interest of using a new direct MIR equation to quantify the CH<sub>4</sub> emissions, the best correlation between milk components already predicted by MIR and CH<sub>4</sub> was calculated. In this way, it was possible to verify whether the predicted CH<sub>4</sub> production data obtained by the developed calibration equation were because of new recombination of global spectral information or just because of specific milk components already predicted by MIR and whose equations are available. If the cross-validation correlation is higher than the correlation between CH<sub>4</sub> and specific MIR milk component, it can be assumed that the developed equation provides additional information than the simple correlation between CH<sub>4</sub> and the already-predicted milk components.

## 3.5. Results and discussion

### 3.5.1. Spectral variation

The first two principal components described 56.8% of the total variation observed in the studied data set (n = 77). These principal components (Figure 3.2) were used to illustrate the spectral differences of studied milk samples. Spectra from cows fed different diets (i.e. grass silage, corn silage and fresh pasture) are separated (Figure 3.2).



**Figure 3.2.** Plot of the principal component analysis of milk spectra for individual cows fed three differing diets: corn silage (•), fresh pasture (o) and grass silage (+). PCA = principal component analysis.

On the basis of the literature (Collomb et al., 2002), this suggests that the use of different diets involved a change in milk composition, and therefore in the milk spectral data. Moreover, Table 3.2, with descriptive statistics, also shows high variations of milk composition primarily fat, protein, lactose and saturated fatty acids. This variability of spectral data is needed to develop a robust calibration equation. Table 3.2 also shows the variability of CH<sub>4</sub> expressed in g/day and g/kg of milk. The coefficients of variation were 29.8% and 33.3%, respectively. This large variation could be explained partly by the use of three different diets. A slightly larger variation was observed for the CH<sub>4</sub> emission expressed in g of CH<sub>4</sub>/kg of milk, because the definition of this trait took into account simultaneously the variability of CH<sub>4</sub> and milk yield.

**Table 3.2.** Means and range of CH<sub>4</sub> emissions, production and composition of the milk, SAT composition of the milk fat for the two experiments (day 0)

Components	n	Mean	Minimum	Maximum	SD
CH <sub>4</sub> (g/d)	77	429	218	653	128
(g/kg <sub>milk</sub> )	77	21.9	10.2	47.1	7.3
Milk, kg/d	77	19.7	12.5	30	5.1
Milk Composition (g/100g <sub>milk</sub> )					
Fat	77	4.11	2.86	5.63	0.61
Protein	77	3.40	2.39	4.82	0.42
Lactose	77	4.67	3.69	5.21	0.27
Milk Fat composition (g/100g <sub>fat</sub> )					
SAT	77	66.1	55.4	76.9	5.2

CH<sub>4</sub> = methane ; SAT = saturated fatty acids ; SD = standard deviation

### 3.5.2. Spectral data treatment

As mentioned previously, for each cow, two spectra measurements and one CH<sub>4</sub> measurement were available. To have one spectrum for one measure of CH<sub>4</sub>, the collected spectral data from the morning and evening milk samples were averaged. To validate the methodology, the contents of major milk components (i.e. fat, protein, lactose, true protein and NPN) predicted from AMS derived from 50% morning and 50% evening spectral data and from the spectral data obtained from the infrared analysis of a representative 24 h milk sample were compared. Table 3.3 shows no significant difference ( $P > 0.05$ ) in milk components predicted by AMS and by the 24 h spectral data. Therefore, AMS was considered similar to the spectrum of the representative 24 h milk sample showing the pertinence of the method used to create the AMS.



### 3.5.3. Calibration model

Different equations were built to predict the quantity of CH<sub>4</sub> produced per day in function of the five different possibilities used to create AMS as shown in Figure 3.1. Table 3.4 presents the statistical parameters obtained for the 20 developed calibration equations. The prediction of CH<sub>4</sub> by MIR, based on R<sup>2</sup><sub>cv</sub>, was better by considering g of CH<sub>4</sub>/kg of milk instead of g of CH<sub>4</sub>/day (R<sup>2</sup><sub>cv</sub> = 0.79 v. R<sup>2</sup><sub>cv</sub> = 0.73). This could be explained by the fact that the expression unit g of CH<sub>4</sub>/kg of milk compared with g of CH<sub>4</sub>/day takes into account the milk production, which is directly related to the release of CH<sub>4</sub> by cows (Vermorel, 1995) and, therefore, reflected in the spectral data, thanks to the dilution effect known for the majority of milk components. For instance, the literature shows that the contents of major milk components such as fat, protein and FAs (Soyeurt et al., 2009) are negatively correlated with the milk yield. Table 3.4 also shows that the use of a ratio taking into account the milk produced during the evening and morning milkings for the calculation of AMS compared with the ratio 50% of morning and 50% of evening spectral data gave better results. This could be explained by the relationship between the milk production and the CH<sub>4</sub> emission by dairy cows (Vermorel, 1995). Therefore, the equations using the CH<sub>4</sub> content expressed in g/kg of milk and the morning and evening milk AMS gave the best results.

**Table 3.3** Comparison between components prediction for AMS and spectra of average milk in a ratio 50-50 for 37 milk sample during the second experiment.

Components	n	AMS	Average milk	P<
Fat milk <sub>predicted</sub> (g/100g milk)	37	4.3	4.4	ns
Protein <sub>predicted</sub> (g/100g milk)	37	3.3	3.3	ns
Lactose <sub>predicted</sub> (g/100g milk)	37	3.7	3.4	ns
True protein <sub>predicted</sub> (g/100g milk)	37	3.3	3.3	ns
NPN <sub>predicted</sub> (mg/100g milk)	37	21.7	20.7	ns

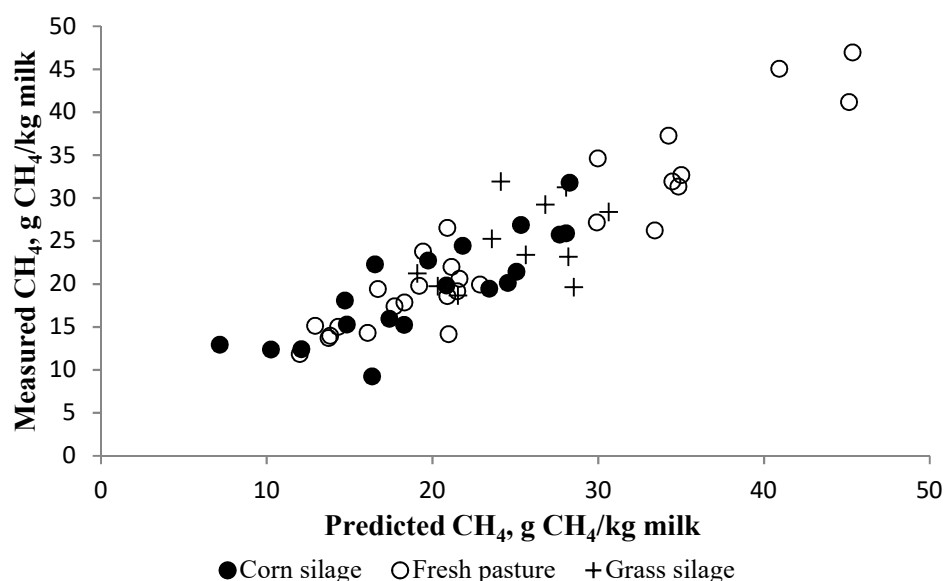
AMS = average milk spectra ; NPN = non-protein nitrogen.

ns: P>0.05

The interval between the measurement of CH<sub>4</sub> by the SF<sub>6</sub> tracer method and the spectral data used (Figure 3.1) gave different results (Table 3.4). By observing the last five lines of Table 2.4, day 1.5 gave the best result for the internal validation (i.e. R<sup>2</sup><sub>cv</sub> = 0.79). The best equation seems to be the one built from the AMS created from

the weighted average, taking into account the milk produced, the CH<sub>4</sub> content expressed in g/kg of milk and an interval between the measurement of CH<sub>4</sub> and the spectral data equal to 1.5 days. This equation also seemed to be the most robust because the difference between R<sup>2</sup>c and R<sup>2</sup>cv was the lowest (0.08). These results are only first indications of the delayed response of milk composition to rumen fermentation. More research is needed in other experimental conditions, for instance, by inducing the inhibition of the rumen methanogenesis and observing the milk spectra modifications in the next few days.

Even if this study must be validated on a larger data set, the obtained results suggest a clear indirect link between the CH<sub>4</sub> emission and the milk composition through the direct use of the MIR spectral data. Chilliard et al. (2009) also showed this relation through the link between the contents of milk FA analyzed by GC and the production of CH<sub>4</sub> by dairy cows fed with diets supplemented or not with linseeds. The present study also shows the feasibility of the MIR prediction of CH<sub>4</sub> produced by dairy cows and a relative robustness of the developed equation. Indeed, despite of the large variation observed for g of CH<sub>4</sub>/kg of milk produced per cow and per diet over the study period and the observed variability of milk composition (Table 3.2), the MIR CH<sub>4</sub> predictions were good and R<sup>2</sup>c and R<sup>2</sup>cv were never below 0.77 and 0.68, respectively. Moreover, CH<sub>4</sub> prediction did not seem to be affected (Figure 3.3) by the different kind of diets used during the experiments (one consisting of corn silage, another of pasture grass and other of grass silage). However, it seems that the pasture grass diet contributes mainly to the large variation in CH<sub>4</sub> emissions. This variation between animals for this diet could be because of the season, and therefore a higher lignification of grass at the end of the experimentation. In addition, individual, for example genetic, differences among the animals could be the reason for these results. To be able to make a generalization of the prediction, it is important to continue to collect data on different diets. Finally, the RPD, which relates the standard error of prediction to the standard deviation of the original reference data, was equal to 2.19, suggesting a good robustness of the prediction. Indeed, Williams and Sobering (1993) showed that a value of 2.5 and above are satisfactory for screening. Therefore, using this equation, it is currently feasible to classify the dairy cows in two groups: low and high CH<sub>4</sub> producers. More variability in the available data can improve the ability of the equation to predict CH<sub>4</sub> emission. Potential use of data obtained from genetically diverse animals, fed on different diets and producing in diverse management, should allow improving the equations.



**Figure 3.3.** Infrared CH<sub>4</sub> prediction on basis of milk spectra of the day1.5 for the different diets: corn silage (●), fresh pasture (○) and grass silage (+).

### 3.5.4. Correlations with milk components

Because of the strong indirect link existing between CH<sub>4</sub> emissions and the overall milk composition, it was expected that CH<sub>4</sub> emissions would correlate strongly with milk production and milk components already predicted by MIR and whose equations are available (Table 3.5). R<sup>2</sup><sub>cv</sub> for the best equation is always superior to the correlation obtained from milk production and MIR milk components, suggesting that the developed equation provides additional information.

A multiple correlation (Table 3.6) was built starting from the parameters appearing in Table 3.5 to predict CH<sub>4</sub> production/kg of milk for the day 1.5. The R<sup>2</sup> (R<sup>2</sup> = 0.76) was lower than the R<sup>2</sup><sub>c</sub> (R<sup>2</sup><sub>c</sub> = 0.87) obtained by the equation for this parameter in Table 3.4. In addition, the results presented in Table 3.6 suggest the existence of a close relationship between FA profiles and CH<sub>4</sub> emission, as previously reported by Chilliard et al. (2009). However, our first results showed that the use of the full spectral data gave more information. In the future, the use of more precise equations for the prediction of the FA could help us better understand the role of each FA and confirm or reverse these observations. In fact, the precision (SECV) of our FA equations could be upgraded in order to meet at least the precision obtained by Soyeurt et al. (2011). This underlines the practical interest to use MIR spectra to predict individual CH<sub>4</sub> emission of dairy cows.

**Table 3.4.** Evolution of accuracy of the resulting calibration with the different kind of AMS with a ratio 50-50 and a weighted average for the two experiments. The number of available data depends on the chosen day.

Ratio	Day	n	R <sup>2</sup> c	R <sup>2</sup> cv	SEC	SECV	RPD	
	0	77	0.84	0.71	73.0	98.3	1.84	
	0.5	71	0.81	0.68	79.4	105.2	1.77	
	g of CH <sub>4</sub>	1	65	0.84	0.72	71.7	98.3	1.88
		1.5	60	0.77	0.70	87.8	100.6	1.83
		2	59	0.77	0.72	86.2	95.2	1.89
Ratio 50-50	0	77	0.80	0.72	4.8	5.7	1.91	
	0.5	71	0.83	0.77	4.4	5.7	2.06	
	g of CH <sub>4</sub> /kg of milk	1	65	0.86	0.69	4.2	6.1	1.81
		1.5	60	0.85	0.75	4.2	5.6	2.01
	2	59	0.93	0.72	3.1	6.0	1.89	
Ratio defined from morning and evening milk yields produced	0	77	0.85	0.72	69.0	95.9	1.90	
	0.5	71	0.83	0.73	74.8	95.4	1.92	
	g of CH <sub>4</sub>	1	65	0.85	0.69	71.5	102.2	1.80
		1.5	60	0.78	0.71	86.9	99.2	1.85
		2	59	0.77	0.72	85.8	83.8	1.89
	g of CH <sub>4</sub> /kg of milk	0	77	0.89	0.75	3.5	5.4	2.00
		0.5	71	0.86	0.78	3.7	5.0	2.12
		1	65	0.84	0.69	4.3	6.2	1.78
		1.5	60	0.87	0.79	4.1	5.1	2.19
	2	59	0.90	0.73	3.6	5.9	1.93	

AMS = average milk spectra; R<sup>2</sup>c = calibration coefficient of determination; R<sup>2</sup>cv = cross-validation coefficient of determination; SEC: Standard error of calibration; SECV: Standard error of cross-validation; RPD: ratio of performance to deviation; CH<sub>4</sub> = methane

**Table 3.5.** Individual correlations between different milk parameters and the CH<sub>4</sub>/kg of milk produced for the day1.5

Milk production and milk components <sup>1</sup>	Correlation CH <sub>4</sub> /Milk
Milk production	-0.45
Protein	0.40
Lactose	0.19
Fat	0.49
C4:0	0.26
C6:0	0.35
C8:0	0.40
C10:0	0.31
C12:0	0.34
C14:0	0.48
C14:1cis-9	0.62
C16:0	0.32
C16:1cis-9	0.48
C17:0	0.53
C18:0	0.25
Tot C18:1 trans	0.11
C18:1 cis-9	0.44
C18:2cis-9,cis-12	0.23
C18:3cis-9;cis-12,cis-15	0.60
c18:2cis-9,trans-11	0.32
R <sup>2</sup> cv	0.79
Rcv	0.89

CH<sub>4</sub> = methane; FTIR = Fourier Transform InfraRed analysis; R<sup>2</sup>c = cross-validation coefficient of determination. Rcv = square root of the R<sup>2</sup>cv value

<sup>1</sup> Milk components were predicted by FTIR spectrometer.

**Table 3.6.** Multiple correlations from the parameters appearing in Table 3.5 to predict the CH<sub>4</sub>/kg of milk produced for the day1.5

Correlation CH <sub>4</sub> /Milk	Name	F test
R <sup>2</sup> : 0.76	1. C18:3 n <sup>-3</sup>	121.83
R : 0.87	2. C4:0	61.87
	3. C12:0	49.48
	4. C6:0	30.57
	5. C8:0	12.42

CH<sub>4</sub> = methane; R = square root of the R<sup>2</sup> value

### **3.6. Conclusion**

The results of our study tended to show that the prediction of CH<sub>4</sub> emissions from milk spectra is feasible. This is partially explained by the fact that the milk spectrum reflects the milk composition, which is considered linked to CH<sub>4</sub>, because of the relationship of both phenomena to ruminal fermentation. Results from this study showed that CH<sub>4</sub> is better predicted, based on R<sup>2</sup><sub>cv</sub>, with the weighted AMS built from spectral data collected at day 1.5 compared with the moment of CH<sub>4</sub> measurement. Additional studies providing data from genetically diverse animals and different breeds, fed different diets and producing in diverse managements, are needed to improve the obtained RPD, and therefore clarify and potentially confirm the initial relationship, which is reported in this study.

The calibration results indicated that the best equation could be used for screening purposes, differentiating high and low CH<sub>4</sub> producers, even if addition research is still required to make it more reliable and potentially implementable on other spectrometers. Once this is achieved, as the existing methods for CH<sub>4</sub> measurement are difficult to apply on a large scale, this alternative method has the potential to be very useful to predict CH<sub>4</sub> emission for dairy herds at individual cow level from milk MIR spectra. These required data are already taken and used for the milk recording and the milk payment. Through the large-scale generation of predicted CH<sub>4</sub> emission data, the method could help improve the knowledge about the sources of CH<sub>4</sub> variation (genetic or not) and about its link to other traits of interest.

### **3.7. Acknowledgments**

The authors gratefully acknowledge the financial support of the Ministry of Agriculture of the Walloon Region of Belgium [Service Public Wallon, Direction générale opérationnelle ‘Agriculture, Ressources naturelles et Environnement’ (DGO3), Direction de la Recherche; Namur, Belgium] and Delta Instruments who provided gracefully a FTIR Spectrometer. The authors thank M. Didelez and the bovine group for their technical assistance, Y. Rochette for the laboratory gas analysis and M. Doreau for his help in the elaboration of the first experiment protocol. Hélène Soyeyrt and Nicolas Gengler, who are Post-doctoral Researcher and Senior Research Associate of the National Fund for Scientific Research (Brussels, Belgium), acknowledge their support.

### 3.8. References

- Beauchemin K.A., McAllister T.A. and McGinn S.M., 2009. Dietary mitigation of enteric methane from cattle. CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources 4, pp. 1–18.
- Bauman D.E. and Griinari J.M., 2003. Nutritional regulation of milk fat synthesis. Annual Review of Nutrition 23, pp. 203–227.
- Bell M.J., Wall E., Russell G., Morgan C. and Simm G., 2010. Effect of breeding for milk yield, diet and management on enteric methane emissions from dairy cows. Animal Production Science 50, pp. 817–826.
- Chilliard Y., Martin C., Rouel J. and Doreau M., 2009. Milk fatty acids in dairy cows fed whole crude linseed, extruded linseed, or linseed oil, and their relationship with CH<sub>4</sub> output. Journal of Dairy Science 92, pp. 5199–5211.
- Clark H., 2010. Animal vs. measurement technique variability in enteric methane production – is the measurement resolution sufficient? In Proceedings of the 4th International Conference on Greenhouse Gases and Animal Agriculture Banff, CA (ed. EJ Mc Geough and SM McGinn), 48pp.
- Collomb M., Bütikofer U., Sieber R., Jeangros B. and Bosset J.-O., 2002. Correlation between fatty acids in cows' milk fat produced in the Lowlands, Mountains and Highlands of Switzerland and botanical composition of the fodder. International Dairy Journal 12, pp. 661–666.
- Delfosse C., Froidmont E., Fernandez Pierna J.A., Martin C. and Dehareng F., 2010. Estimation of methane emission by dairy cows on the basis of milk composition. In Proceedings of the 4th International Conference on Greenhouse Gases and Animal Agriculture Banff, CA (ed. EJ Mc Geough and SM McGinn), 168pp.
- Demeyer D. and Fievez V., 2000. Ruminants et environnement: la méthanogénèse. Annales de Zootechnie 49, pp. 95–112.
- Dijkstra J., Apajalahti J.A., Bannink A., Gerrits W.J.J., Newbold J.R., Perdok H.B., van Zijderveld S.M. and Berends H., 2010. Relationship of milk fatty acid profile with methane production in dairy cattle. In Proceedings of the 4th International Conference on Greenhouse Gases and Animal Agriculture Banff, CA (ed. EJ Mc Geough and SM McGinn), 169pp.

- Food and Agriculture Organization (FAO), 2010. New FAO report assesses dairy greenhouse gas emissions. Retrieved November, 2011, from <http://www.fao.org/news/story/en/item/41348/icode/>
- Fuglestad J., 2009. Impacts of metric choice on analyzing the climate effects of emissions. IPCC Expert Meeting on the Science of Alternative Metrics, Oslo, Norway, pp. 40–47.
- Hegarty R.S., Goopy J.P., Herd R.M. and McCorkell B., 2007. Cattle selected for lower residual feed intake have reduced daily methane production. *Journal of Animal Science* 85, pp. 1479–1486.
- Johnson K.A. and Johnson D.E., 1995. Methane emissions from cattle. *Journal of Animal Science* 73, pp. 2483–2492.
- Johnson K.A., Huyler M., Westberg H., Lamb B. and Zimmerman P., 1994. Measurement of CH<sub>4</sub> emissions from ruminant livestock using a sulphur hexafluoride tracer technique. *Environmental Science & Technology* 28, pp. 359–362.
- Jouany J.P., 2008. Enteric CH<sub>4</sub> production by ruminants and its control. In *Gut efficiency: the key ingredient in ruminant production. Elevating animal performance and health* (ed. S Andrieu and D Wilde), pp. 35–59. Wageningen Academic Publishers, Wageningen, the Netherlands.
- Lassey K.R., Uylatt M.J., Martin R.J., Walker C.F. and Shelton I.D., 1997. Methane emissions measured directly from grazing livestock in New Zealand. *Atmospheric Environment* 31, pp. 2905–2914.
- Martin C., Doreau M. and Morgavi D.P., 2010. Methane mitigation in ruminants: from microbe to the farm scale. *Animal* 4, pp. 351–365.
- Martin C., Rouel J., Jouany J.P., Doreau M. and Chilliard Y., 2008. Methane output and diet digestibility in response to feeding dairy cows crude linseed, extruded linseed, or linseed oil. *Journal of Animal Science* 86, 2642–2650.
- Ørskov E.R., Fraser C. and Kay R.N.B., 1969. Dietary factors influencing the digestion of starch in the rumen and small and large intestine of early weaned lambs. *British Journal of Nutrition* 23, pp. 217–226.
- Soyeurt H., Dardenne P., Dehareng F., Bastin C. and Gengler N., 2009. Genetic parameters of saturated and monounsaturated fatty acid content and the ratio of saturated to unsaturated fatty acids in bovine milk. *Journal of Dairy Science* 91, pp. 3611–3626.



- 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. *Journal of Dairy Science* 94, pp. 1657–1667.
- Soyeurt H., Dardenne P., Dehareng F., Lognay G., Veselko D., Marlier M., Bertozzi C., Mayeres P. and Gengler N., 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. *Journal of Dairy Science* 89, pp. 3690–3695.
- Vermorel M., 1995. Emissions annuelles de méthane d'origine digestive par les bovins en France: variation selon le type d'animal et le niveau de production. *INRA Productions Animales* 8, pp. 265–272.
- Vlaeminck B. and Fievez V., 2005. Milk odd and branched-chain fatty acids to predict ruminal methanogenesis in dairy cows. *Communications in Agricultural and Applied Biological Sciences* 70, pp. 43–47.
- Weill W., Chesneau G., Chilliard Y., Doreau M. and Martin C., 2009. June 24, Method for evaluating the amount of CH<sub>4</sub> produced by a dairy ruminant and method for decreasing and controlling this amount. *World Patent WO/2009/156453*.
- Weill P., Kerhoas N., Chesneau G., Schmitt B., Legrand P. and Renaud J.P., 2008. Existe-t-il un lien entre production de méthane par les vaches laitières et profil en acides gras des laits? *Nutrition clinique et métabolisme* 22, pp. 71–72.
- Williams P.C. and Sobering D.C., 1993. Comparison of commercial near infrared transmittance and reflectance instruments for the analysis of whole grains and seeds. *Journal of Near Infrared Spectroscopy* 1, pp. 25–33.



# Chapitre 4:

---

## **Large-scale phenotyping in dairy sector using milk MIR spectra: Key factors affecting the quality of predictions**

**Adapté de:** Grelet C., Dardenne P., Soyeurt H., Fernandez Pierna J. A., Vanlierde A., Stevens F., Gengler N., and Dehareng, F. (2021). Large-scale phenotyping in dairy sector using milk MIR spectra: Key factors affecting the quality of predictions. *Methods*, 186, 97-111



## 4.1. Points d'attention

Ce quatrième chapitre est le fruit de retours d'expérience accumulés par l'équipe de recherche travaillant sur la spectrométrie FT-MIR du lait durant les 17 dernières années et depuis que l'accessibilité aux spectres bruts des appareils de routine a été rendue possible via l'intervention de Didier Veselko (Comité du lait de Battice). Néanmoins, bon nombre de recommandations, d'observations, de solutions, proviennent directement de l'expérience acquise par le passé en NIR (p.ex. : Dardenne et al., 2010 ; Shenk et al., 1985) et en MIR durant les travaux de mise au point de la guidance scientifique belge au CRA-W (Dardenne and Van Reusel, 2003).

On doit malheureusement constater que dans la plupart des cas, les modèles développés en recherche visent essentiellement à faire la démonstration que la spectrométrie FT-MIR peut être utilisée pour prédire de nouveaux paramètres/phénotypes. Ils sont développés à partir de jeux de données collectés dans des conditions expérimentales spécifiques, et ils cherchent à maximiser les performances du modèle sur ces jeux de données réduits, plutôt que de viser à accroître leur robustesse. Habituellement, et on ne peut les blâmer, aucune tentative n'est faite par les chercheurs pour déterminer si leurs modèles fonctionnent correctement en conditions réelles, car leur objectif principal vise à démontrer la faisabilité du modèle et à le publier. Dès lors, trop souvent, ces modèles ne sont pas suffisamment robustes pour être utilisés en routine ou être appliqués directement sur de larges bases de données. Or, depuis quelques années, la création de base de données spectrales et leurs potentiels à être utilisées pour, par exemple fournir de nouveaux phénotypes nécessaires à la génomique, ou pour aider à améliorer la rentabilité ou la résilience de la filière lait, nous oblige à faire en sorte que ces modèles soient les plus robustes possibles.

Ce chapitre tente de plaider en faveur d'un développement de recommandations au niveau international et à la création de collaborations internationales pour le développement de modèles prédictifs en spectrométrie FT-MIR dans le lait. Nous développerons plus en détail ce point important dans la discussion générale réalisée dans le chapitre suivant.

Cette publication est un travail collaboratif mené par les équipes de recherche actives dans le domaine de la spectrométrie infrarouge du CRA-W et de Gembloux Agro-Bio Tech/ULiège, chacun apportant son expertise respective. Ma contribution dans cet article en tant que dernier auteur a été notamment d'aider à la coordination générale et faire le lien entre les différents projets de recherche sur la problématique de la spectrométrie FT-MIR au CRA-W. J'ai également participé à la genèse et la conception de ce travail et aidé à la rédaction de cette publication.

### ***4.1.1. Références***

- Dardenne P. and Van Reusel A., 2003. Le lait, qualité et guidance: collaborations sans frontières. Proceedings in: Symposium organisé à l'occasion du départ à la retraite de André Van Reusel, chef de section, Gembloux - Belgique, 26/11/2003
- Dardenne P., 2010. Some considerations about NIR spectroscopy: Closing speech at NIR-2009. NIR news. 21(1):8-14.
- Shenk J.S., Westerhaus M.O. and Templeton W.C., 1985. Calibration Transfer Between near Infrared Reflectance Spectrophotometers 1. Crop Sci., 25(1), 159-161.
- Williams P., Dardenne P. and Flinn P., 2017. Tutorial: Items to be included in a report on a near infrared spectroscopy project. J. Near Infrared Spec., 25(2), 85-90.

## **4.2. Abstract**

Methods and technologies enabling the estimation at large scale of important traits for the dairy sector are of great interest. Those phenotypes are necessary to improve herd management, animal genetic evaluation, and milk quality control. In the recent years, the research was very active to predict new phenotypes from the mid-infrared (MIR) analysis of milk. Models were developed to predict phenotypes such as fine milk composition, milk technological properties or traits related to cow health, fertility and environmental impact. Most of models were developed within research contexts and often not designed for routine use. The implementation of models at a large scale to predict new traits of interest brings new challenges as the factors influencing the robustness of models are poorly documented. The first objective of this work is to highlight the impact on prediction accuracy of factors such as the variability of the spectral and reference data, the spectral regions used and the complexity of models. The second objective is to emphasize methods and indicators to evaluate the quality of models and the quality of predictions generated under routine conditions. The last objective is to outline the issues and the solutions linked with the use and transfer of models on large number of instruments. Based on partial least square regression and 10 datasets including milk MIR spectra and reference quantitative values for 57 traits of interest, the impact of the different factors is illustrated by evaluating the influence on the validation root mean square error of prediction (RMSEP). In the displayed examples, all factors, when well set up, increase the quality of predictions, with an improvement of the RMSEP ranging from 12% to 43%. This work also aims to underline the need for and the complementarity between different validation procedures, statistical parameters and quality assurance methods.

Finally, when using and transferring models, the impact of the spectral standardization on the prediction reproducibility is highlighted with an improvement up to 86% with the tested models, and the monitoring of individual spectrometer stability over time appears essential. This list inspired from our experience is of course not exhaustive. The displayed results are only examples and not general rules and other aspects play a role in the quality of final predictions. However, this work highlights good practices, methods and indicators to increase and evaluate quality of phenotypes predicted at a large scale. The results obtained argue for the development of guidelines at international levels, as well as international collaborations in order to constitute large and robust datasets and enable the use of models in routine conditions.

### 4.3. Context and framework

Methods and technologies enabling the measurement or estimation of traits of interest regarding dairy production at large scale are of great interest. Those phenotypes are necessary to improve herd management, animal genetic evaluation, and milk quality control (e.g. nutritional quality, authentication of milk regarding geographic origin or species, detection of milk adulteration, ability of milk to be processed into different products).

Among the potential technologies, analysis of milk by Mid-Infrared (MIR) spectrometry allows a fast and cost-effective quantification of milk chemical composition. It is a method of choice to generate phenotypes at large scale as it is already used in main dairy areas to quantify fat, protein, and lactose contents in the frame of milk payment and milk recording. Moreover, MIR spectrum of milk is considered as a biomarker containing information on fine milk composition, milk technological properties, status of cows and is actively used by research teams to predict additional phenotypes of interest for the dairy sector (De Marchi et al. [1], Gengler et al. [2]). Within only the past three years, numerous researches were carried out to enlarge the number of predicted phenotypes related to many aspects of the dairy production such as energy status (Smith et al. [3]), metabolic status (Belay et al. [4], Pralle et al. [5], Grelet et al. [6], Bonfatti et al. [7], Benedet et al. [8], Luke et al. [9], De Koster et al. [10]), body weight (Soyeurt et al. [11]), feed intake (Shetty et al. [12], Wallen et al. [13], Grelet et al. [14]), feed composition (Klaffenböck et al. [15]), methane emission (Shetty et al. [16], Vanlierde et al. [17,18], Van Gastelen et al. [19], Denninger et al. [20]), nitrogen efficiency (Grelet et al. [14]), fertility status (Lainé et al. [21], Toledo-Alvarado et al. [22], Ho et al. [23], Delhez et al. [24]), mastitis (Rienesl et al. [25]), lameness (Mineur et al. [26,27], Bonfatti et al. [28]), lactoferrin content (Nayeri et al. [29]), fatty acids (Fleming et al. [30]), fat globule size (Fleming et al. [31]), cheese making properties (Manuelian et al. [32], Sanchez et al. [33], El Jabri et al. [34]). In these studies, most of the models predicting the traits of interest

were developed using datasets collected under specific experimental conditions, and by maximizing the model performances on these reduced datasets. Therefore, the validity of these models in other conditions cannot be guaranteed a priori. However, in order to be useful for herd management, genetic studies or milk quality control, MIR models need to be used at a large scale, i.e. with data coming from diverse origins and subject to various external factors of variability. For example, approximately 1,700,000 milk MIR spectra are generated each year in Wallonia (South Belgium) in the framework of milk payment or milk recording and those field data show large variability due to different breeds, genetic backgrounds, feed diets, herd managements, parities, stages of lactation or climates.

Therefore, in order to insure predicted phenotypes of good quality, the factors influencing the ability of models to cope with these different conditions should be carefully considered when developing models. Additionally, methods and indicators are necessary to evaluate and validate the quality of models and of final predictions before their uses for industrial or management purposes. Regarding ability of models to be applied in real field conditions, robustness is an essential parameter as it reflects the capacity of models to be all-terrain, i.e. to generate accurate predictions in various conditions. The impact of the modelling algorithm and pretreatment on robustness and quality of predictions has been discussed in the literature, especially for NIR spectroscopy (Goedhart [35], Zeaiter et al. [36, 37], Balabin et al. [38]). However, the impact of factors such as the variability of the spectral and reference data, the spectral regions used and the complexity of models are less documented, especially for MIR dedicated to milk analysis. Consequently, the first objective of this work is to highlight those important complementary factors influencing the model robustness and therefore the quality of predictions. The second objective is to emphasize methods and indicators to evaluate the performances of models and the reliability of predictions generated under routine conditions. The last objective is to outline the issues and the solutions related with the different spectral responses between instruments and instability of individual instruments spectral response over time. Both objectives aims to highlight good practices, methods and indicators to increase and evaluate quality of phenotypes predicted at a large scale.

## **4.4. Material and methods**

### ***4.4.1. Models, samples and instruments***

To highlight the impact of the different factors, this work aims to provide concrete examples, and not only theoretical views, with real datasets related to phenotypes of interest for the dairy sector. As the impact of the modeling algorithm is not studied in this work, only Partial Least Square (**PLS**) regressions models were used, as this technique is currently the most used with milk MIR spectra.



Ten different datasets including milk MIR spectra and reference quantitative values for 57 traits of interest in total were used in this article, some of which already referenced in published works and others unpublished (Table 4.1). The datasets were selected to cover different data size, model performances, direct and indirect predictions and various types of traits (i.e., fine milk composition, blood composition, cow status, technological properties of milk). Each dataset contains data from winter and grazing seasons and was constituted by individual cow samples, except the casein dataset, which also contains bulk tank samples. When needed, models were cross-validated using cycling 10-fold cross-validation (CV) procedure. During PLS model development, number of latent variables (LV) were selected visually through the point of break of root mean square error of cross-validation (RMSE<sub>cv</sub>) slope, where adding a new LV does not reduce substantially the RMSE<sub>cv</sub>, with a maximum of 16 LV to avoid overfitting.

**Table 4.1.** Datasets used in the study.

Traits	N records	N cows	N countries	Sampling years	References
Milk Fatty acids, 30 models	1822	1822	7	2005-2015	Soyeurt et al. [39,40]
Milk Minerals, 5 models	1340	1340	4	2005-2015	Soyeurt et al. [41]
Milk Lactoferrin	3906	3906	3	2005-2009	Soyeurt et al. [42]
Methane emitted (CH <sub>4</sub> )	1089	299	7	2010-2019	Dehareng et al. [43] and Vanlierde et al. [18]
Milk Fresh Cheese Yield (FCY), Coagulation time $\tau$ , Time when the curd is firm enough for cutting (k20)	283	283	1	2011-2014	Colinet et al. [44]
Milk Casein	996	*	1	2011-2014	Not published
Milk Acetone, $\beta$ -hydroxybutyrate (BHB) and Citrate	566	346	3	2013-2014	Grelet et al. [45]
Blood BHB, Non-Esterified Fatty Acids (NEFA), Insulin Growth Factor I (IGF-I), Glucose	387	241	6	2014-2016	Grelet et al. [6]
Nitrogen efficiency (NUE), Nitrogen losses, Dry matter intake (DMI), Body weight	1034	129	3	2014-2015	Grelet et al. [14]
Milk Glucose free, Glucose-6-phosphate, Uric acid, Iso-citrate, Progesterone	2175	241	6	2014-2016	Not published

\*The casein model was constituted by 790 samples from individual cows and 206 samples from bulk tank.

All milk samples were collected following the guidelines edited by the International Committee for Animal Recording [46] and with the use of preservative to ensure stability of milk physico-chemical properties before analysis. To be representative of the real routine methods, the samples coming from the different countries were preserved using the molecules and concentration classically used within each local milk recording. In majority of cases, bronopol was used with a concentration of 0.03%, with in some places addition of Kathon CG at 0.003%. Spectral data were obtained from Fourier transform MIR spectrometers included in the European Milk Recording network (EMR) or from instruments used punctually for research purposes. The different spectrometers used were FT120, FT2, FT6000, FT+ and FT7 (Foss, Hillerød, Denmark), FTS (Bentley, Chaska, United States) and

Standard Lactoscope FT-MIR automatic (PerkinElmer, Waltham, United States). The wavenumbers ranges of the different brands were from 925.66 to 5,010.15  $\text{cm}^{-1}$  for Foss instruments, from 649.03 to 3,998.59  $\text{cm}^{-1}$  for Bentley instruments and from 397.31 to 4,000  $\text{cm}^{-1}$  for PerkinElmer instruments. The spectral resolution used is 8  $\text{cm}^{-1}$  for PerkinElmer and Bentley Instruments and unknown for Foss instruments. Spectrometers were located in Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Luxembourg, Switzerland and United Kingdom. In the frame of standardization procedure, common sets of 5 milk samples were analyzed monthly as described in Grelet et al. [47] on a total of 111 instruments.

#### 4.4.2. Calculations

The standardized Mahalanobis distance (GH, Global H) and the neighboring H distance (NH) were calculated as described in Shenk and Westerhaus [48, 49]. The GH is the standardized Mahalanobis distance between a sample and the centroid of the dataset. High GH are associated with different or extreme spectra with respect to the dataset. It is calculated based on the principal components of (PC) the pretreated spectra (i.e. after selection of informative wavenumbers and first derivative) as following:

Equation 1. 
$$GH = \sqrt{(x - \mu)^T S^{-1}(x - \mu)} / nPC$$

where  $x$  is the PC scores of a specific spectrum,  $\mu$  is the mean of PC scores estimated from the calibration set (i.e., representing the centroid),  $S$  is the (co-)variance matrix between PC scores estimated from the calibration set and  $nPC$  represents the number of PC considered for the calculation. As the distance is divided by the number of principal components, the average GH for any data set is always 1. To calculate GH for a new sample, the scores are obtained by using the loadings of the calibration dataset. NH is the standardized Mahalanobis distance between a sample and its nearest neighbor, with high NH associated with samples which are not redundant with others and which are filling a gap in the multivariate space of the dataset. NH is calculated similarly as GH but looking to the distance between 2 samples instead of distance of a sample to the centroid.

Model robustness was assessed by the validation root mean square error of prediction (**RMSEP**). The impact on the quality of prediction of the spectral standardization, aiming to harmonize the predictions among spectrometers, was evaluated through the reproducibility, i.e. the standard deviation of predictions between instruments. Computations were carried out with the PLS toolbox v. 4.11 (Eigenvector Research, Inc., Wenatchee, WA, USA), Winisi software (version 4.6; Foss, Hillerød, Denmark) and scripts developed in Matlab v7.5.0. (The Mathworks,

Inc., Natick, MA, USA), and in R studio 1.2.5033 (RStudio, Inc.) using R statistical software 3.6.2 (R development Core Team 2019).

## 4.5. Experiments and results

This article aims to highlight key issues that may influence the quality of predictions, from the development of models to the evaluation of predicted phenotypes. This section is divided into 5 parts: (1) inclusion of variability in the calibration dataset, (2) spectral regions to use, (3) model complexity, (4) evaluation of models and predictions and (5) use and transfer of models. This list inspired from our experience is of course not exhaustive.

### *4.5.1. Inclusion of variability in the calibration dataset*

The inclusion of variability related both to the trait to predict and to the spectral data is a crucial issue. In general, predicting new samples out of the calibration domain (i.e. extrapolation because samples are substantially different to those used to develop the calibration) should be done cautiously as it can lead to high inaccuracies and misleading conclusions. This is particularly true with PLS models in the field of IR spectroscopy, where it has been noted that the systematic control of whether a new sample is within the calibration domain is needed (Dardenne [50]). However, when using a model at a large scale, it faces a large variability in the new samples, regarding both the phenotype of interest, the milk composition and the MIR spectrum. In order to avoid extrapolation it is necessary to maximize the variability in the calibration dataset, so that new samples to predict fall within the range of calibration data (for spectral features and reference values). Additionally, inside the variability of the calibration dataset, both regarding the distribution of the trait to predict and the spectral space, parts of the domain could suffer from low sample density due to non-uniform repartition of samples, and prediction accuracy within those areas would probably be affected. A calibration dataset maximizing the variability of the trait to predict ( $y$ ) and the spectral data ( $X$ ) is therefore of interest. Additionally, collection of new samples associated with reference values is usually time-consuming and expensive. Hence, there is a need to efficiently select new samples in order to maximize the variability of the calibration dataset.

#### *4.5.1.1. Variability of reference value ( $y$ )*

The first important aspect regarding the variability of the calibration dataset is the range of the reference ( $y$ ) data. Models predicting the main milk components (i.e. fat, protein and lactose) are robust to extrapolation as the absorbance response is linear to the concentration of the molecule, as stated by the Beer-Lambert law. However, this may not be the case in the frame of models predicting minor

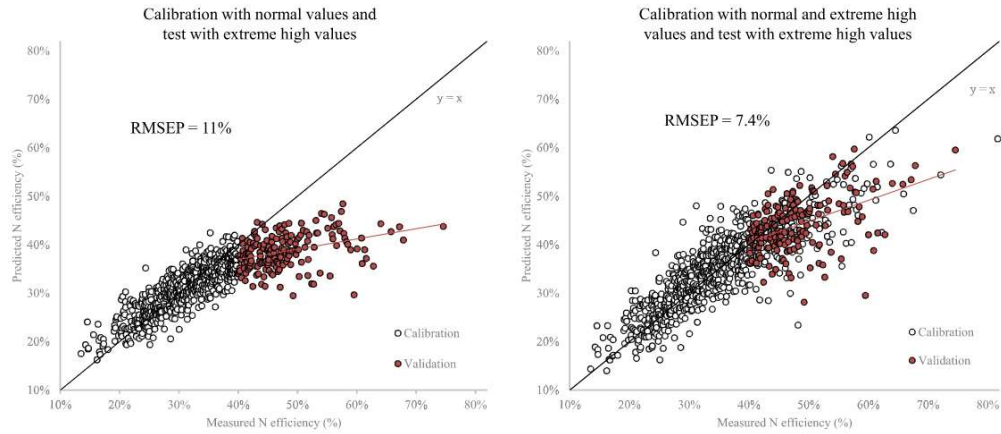
components or animal related traits. Such models may be partly or totally indirect because predicting molecules in low concentration, even being under the reference threshold of detection of 100 ppm (Dardenne et al. [51]), having low, non-usable or no interaction between the chemical element and MIR rays, or because models predict blood composition or the status of the cow.

To illustrate this, the dataset related to the individual nitrogen efficiency is used (Table 1), which is a trait indirectly predicted from milk MIR spectra and calculated as the ratio of nitrogen excreted in milk to nitrogen in the diet of the cow. Two models were developed using different data subsets with different distributions. The first model was calibrated using records included in the expected range of N efficiency, i.e. with 660 samples between 8% to 40% (Castillo et al. [52]). The second model was calibrated using records over the full range of the dataset, with values of N efficiency from 8% to 81.7%. In addition to the 660 measurements of the first model, the second model included 187 extreme high measurements from 40% to 81.7%. For each datasets, a model was developed from milk MIR spectra using PLS regression with 14 latent variables. Then both models were tested with a common external independent validation set containing 186 measurements with high values only, from 40% to 75.6% to highlight the impact of a prediction out of the y calibration range. Although the first model containing no extreme values resulted in a better cross-validation RMSE (3.5% vs. 5.2% when considering the full range of models, and 3.5% vs. 4.3% when considering the normal range only), its RMSEP was higher than for the second model developed from normal and high values (11% and 7.4%, respectively) (Figure 4.1.). Inclusion of extreme high values in the calibration dataset resulted thus in an improvement of the RMSEP by 33%. Those results highlight the poor capacity of the model developed with values included in the expected range to extrapolate and provide good predictions when facing records outside of the expected range. On the opposite, the model containing high values seemed less accurate in the cross-validation step, but was finally more robust and provided better predictions when facing extreme but existing values. Hence, from this small test it can be concluded (1) that extrapolation is dangerous, which argues for the use of large and variable dataset, and (2) that models are more precise when developed on specific ranges. In order to develop robust and precise models, this is arguing for the use of local methods such as shown by Allegrini et al. [53] because it brings the possibility to combine large calibration datasets with important y variability, and the precision of locally developed specific models.

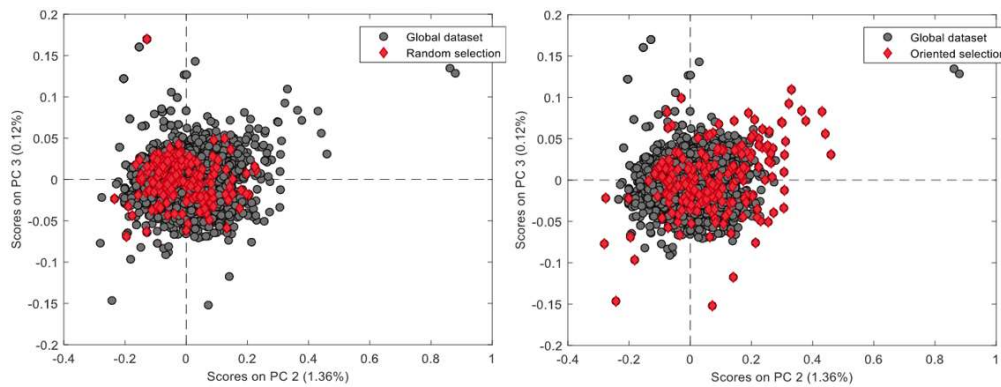
#### ***4.5.1.2. Variability of spectral data (X)***

If the importance of maximizing variability in reference values is easily understandable, this is only the tip of the iceberg. To ensure the robustness of the

prediction model before considering large-scale application, the second aspect, which is often underestimated, is the variability in spectral data (X). This is a more complex issue as many factors have an influence on the IR spectrum measured such the milk composition, the milk production specificities or the IR measure. As for the reference values, there is a need to cover as much as possible the potential spectral variability in the calibration dataset by using an appropriate selection of samples.



**Fig 4.1.** Comparison of two models, with different reference values range, when predicting external extreme high values.



**Figure 4.2.** Comparison of the distribution of the 2 calibration datasets with the distribution of the global population. Second and third principal components of the calibration datasets using random selection (red points, top,  $n = 200$ ) and selection based on spectral variability (red points, bottom,  $n = 200$ ), plotted on the global dataset (grey points,  $n = 3506$ ).

To illustrate this, the lactoferrin dataset (Table 4.1.,  $n=3906$  samples) was used to compare the impact of two sampling methods on the spectral variability of the calibration dataset, and its impact on robustness: (1) simple random selection versus (2) selecting samples based on spectral variability (i.e. oriented selection for extreme and not redundant samples). Beforehand, 400 samples were randomly excluded to constitute a validation dataset and the remaining samples ( $n=3506$ ) were considered as a population of cows. In order to simulate a concrete context of limited financial resources to analyze reference samples, the size of the calibration dataset was fixed to 200 samples. In a first case, 200 samples were selected using a simple random selection by cycling, preserving the proportion of samples per country but selecting in a blind way regarding lactoferrin content, milk composition and spectral variability. In a second case, the GH and the NH were calculated for the 3506 samples. Based on the GH and the NH, 200 samples were selected, by taking the 130 samples with the highest NH to insure a good repartition of the calibration samples within the dataset, and by taking the 70 samples with the highest GH to cover the extreme parts of the calibration dataset. The distribution of the 2 calibration datasets are compared to the distribution of the global population using a principal component analysis. A larger variability for the oriented selection can be observed on Figure 4.2 with a better coverage of the global dataset. Figure 4.2 reports the second and third principal components as the spectral information is more visible than using the first principal component. For both selected datasets, a model predicting lactoferrin content in milk was developed from milk MIR spectra using PLS regression with 8 latent variables. RMSE<sub>cv</sub> of the model developed from the oriented selection was 176 mg/L, which is 40% higher than the RMSE<sub>cv</sub> obtained from the simple random selection, suggesting a lower capacity of this model to provide accurate predictions. Results are displayed in Table 4.2. The model was then tested using the external validation set (i.e., 400 excluded samples). The capacity of the two models to cover the spectral variability of the external dataset was evaluated by calculating the percentage of samples with GH higher than 3, which are considered as not covered by the spectral calibration datasets. The percentage was 5.6% for the random selection and decreased to 1.6% with the oriented selection. The better proportion of samples covered by the calibration dataset suggests thus a reduced risk of error in the predictions, thanks to the reduction of spectral extrapolation. Finally, the RMSEP was calculated with the external dataset and reached 170 mg/L with the simple random selection and 146 mg/l with the oriented selection, i.e. a reduction of 14%. Therefore, the strategy aiming at limiting spectral extrapolation via the maximization of spectral variability in the calibration dataset increased the model robustness in the current example.

As highlighted,  $y$  and  $X$  variabilities are key elements influencing the quality of predictions and both aspects can be influenced by third factors. When constituting a calibration dataset, and complementary to selection of samples, it is consequently of

great interest to maximize the chances to gather variable records by sampling in different geographical areas, countries, years, climates and to look for different diets, management systems, genetics... Moreover, as highlighted by Eskildsen et al. [54], the covariance structure between the trait to predict and the absorbance values may vary between breeds, for instance the covariance between fatty acids and absorbance values differing between Holstein and Jersey. The covariance structure may potentially be affected by many other variables such as seasons, lactation stages or health status of cows. For future samples to predict, it is important that the specific covariance is covered by the calibration dataset by including the different aspects of those variables. Collaboration between countries to create common calibration sets with different and complementary specificities is an optimal solution to achieve those objectives.

**Table 4.2.** Comparison of the models based on the two different selection methods to constitute the calibration dataset.

	Calibration datasets			
	Simple selection	random	Oriented selection	% improvement
N	200		200	
SDy	152		214	
RMSEcv (mg/L)	126		176	
% of external samples with GH > 3	5.6		1.6	71%
RMSEP (mg/L)	170		146	14%

#### 4.5.2. Which spectral regions to use

Reference works derived from the spectroscopy field mentioned that the strong absorption of water in the regions from about 1,600 to 1,700  $\text{cm}^{-1}$  and from 3,000 to 3,500  $\text{cm}^{-1}$  masks the information in relation with milk composition (Bertrand and Dufour [55], Andersen et al. [56], Li-Chan et al. [57]). Additionally, spectral region from 3,500 to 5,000  $\text{cm}^{-1}$  is considered to contain few valuable information and is often discarded from models (Karoui et al. [58]). However, more recent studies derived from genetics evaluated those regions and concluded in the potential presence of physico-chemical information (Bittante and Cecchinato [59], Wang et al. [60], Toledo-Alvarado et al. [22]). To use models at large scale and ensure model robustness, there is a need to know the reproducibility of the absorbance signal at each individual wavenumbers whatever it contains information or not. Indeed, if a model has to be used for prediction at a large scale, with different instruments, it should only rely on

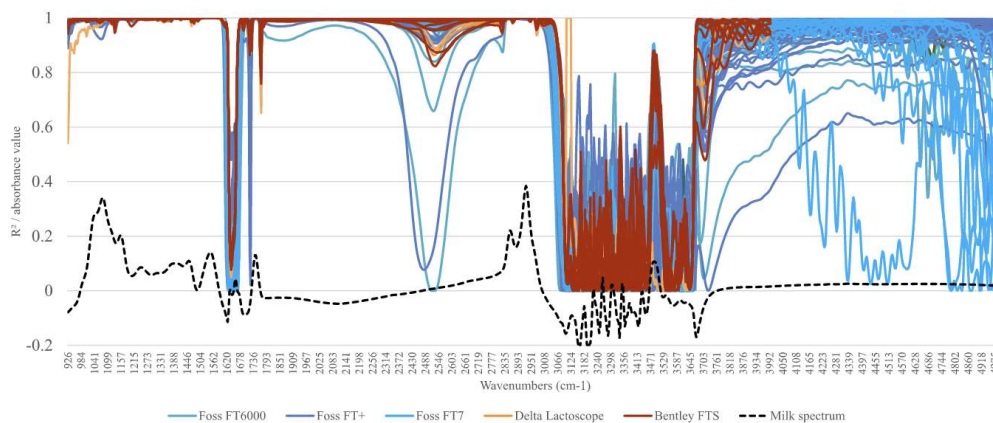
wavenumbers for which the absorbance is highly reproducible between these instruments in order to insure similar and correct predictions among spectrometers.

To illustrate this concept, 5 milk samples with variation in fat and protein were analyzed on 83 MIR instruments dedicated to milk analysis. Instruments used were 19 FT 6000, 41 FT+, and 12 FT7 (Foss, Hillerød, Denmark), 10 FTS (Bentley, Chaska, United States) and 1 Standard Lactoscope FT-MIR automatic (PerkinElmer, Waltham, United States). Those instruments were located in European countries and the proportion of each brand and model may differ from others parts of the world. A fictitious reference instrument was constituted by averaging 18 instruments and the 5 samples were constituted as described in Grelet et al. [47]. A linear interpolation was performed on the Bentley and PerkinElmer spectra to match the Foss range. The five spectra obtained with the reference instrument were compared, wavenumber by wavenumber, to the five spectra obtained with each of the 82 secondary instruments. For each secondary instrument and at each wavenumber, the  $R^2$  was calculated to assess whether information is related between the different spectrometers. The values of  $R^2$  throughout the wavenumbers is shown in Figure 4.3 for each of the secondary instruments. A typical milk MIR spectrum from the reference instrument is also displayed in order to identify the regions of interest. First, Figure 4.3 shows that between  $1,620$  and  $1,670\text{ cm}^{-1}$  and between  $3,080$  and  $3,670\text{ cm}^{-1}$ , i.e., the regions of water absorption, the absorbance of each secondary instruments was poorly related with the absorbance of the reference instrument, with  $R^2$  being erratic and close to zero. Consequently, even if a signal is detectable within those spectral regions with one instrument, it would be counterproductive to include them in a model, as the information is not reproducible between instruments. Besides water absorption regions, studies mentioned earlier focus on spectral regions between  $3,670$  and  $5,000\text{ cm}^{-1}$ . Such spectral areas exclude *de facto* the Bentley and PerkinElmer instruments as their spectral range ends at  $4,000\text{ cm}^{-1}$ . On top of that, the spectral information is not highly reproducible between Foss instruments, making dangerous the use of such wavenumbers in models. Finally, other spectral regions, as between  $925$  and  $960\text{ cm}^{-1}$  and in the flat part between  $1,810$  and  $2,820\text{ cm}^{-1}$ , (especially where  $\text{CO}_2$  affects the spectra, approximately from  $2,450$  and  $2,600\text{ cm}^{-1}$ ), the absorbance values were not highly reproducible between instruments, with  $R^2$  being lower than 0.9. A large number of instruments is necessary to highlight those  $R^2$  variations, which are likely to be underestimated if evaluated with only few spectrometers.

To evaluate the impact of the choice of the spectral regions on model robustness, 3 different models were developed with different spectral regions (M1, M2, and M3) but using the same samples. The oleic acid (c18:1cis9) dataset, being a sub-dataset from the fatty acid dataset (Table 4.1.) and containing 1,832 spectral data was used. To test the inclusion in models of the area between  $4000$  and  $5010\text{ cm}^{-1}$ , only the Foss spectra were retained, resulting in a dataset of 1739 samples. A total of



250 samples were randomly selected and used for calibration and the remaining 1,489 samples were used for validation. The robustness was assessed by the validation RMSEP. M1 was developed using the entire Foss range. M2 was developed after removing noisy water absorption regions and the near infrared spectral region. M3 used only wavenumbers having high reproducibility between the instruments (Grelet et al. [45]), i.e. with  $R^2$  being higher than 0.95. The model specificities are shown in Table 4.3.



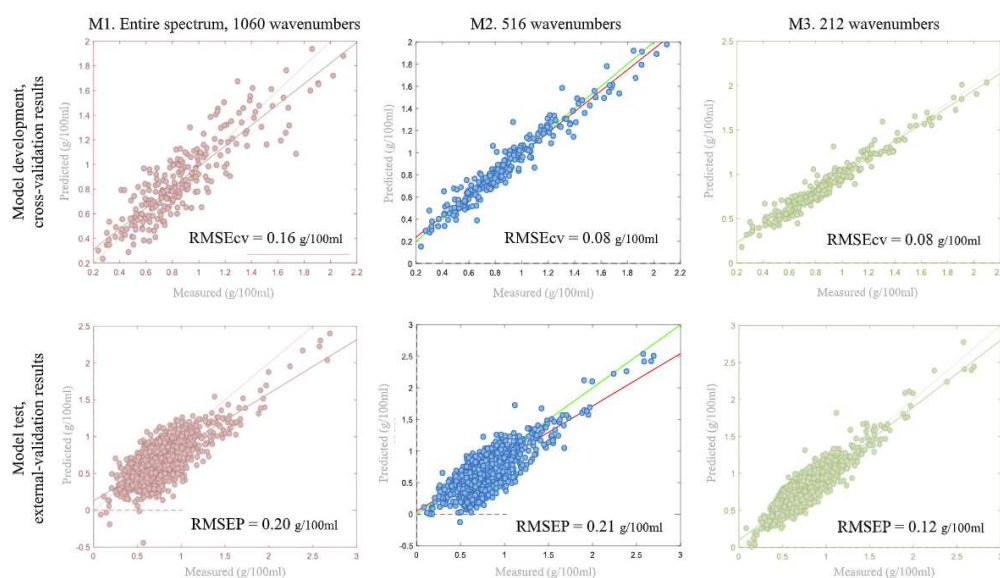
**Figure 4.3.**  $R^2$  at each wavenumber between absorbance values of a reference instrument and 82 secondary instruments.

**Table 4.3.** Spectral regions used in the development of the 3 models.

Model	Spectral regions	Wavenumbers used
M1	925-5,010 $\text{cm}^{-1}$	1,060
M2	925-1,597; 1,693-3,004 $\text{cm}^{-1}$	516
M3	968.1-1,577.5; 1,731.8-1,762.6; 1,781.9-1,808.9; 212 2,831.0-2,966.0 $\text{cm}^{-1}$	

Models were all developed from milk MIR spectra with PLS and 13 latent variables. The cross-validation results were improved by deleting noisy water absorption regions and the near infrared spectral region, and comparable RMSEcv were obtained for M2 and M3 (Figure 4.4). However, the use of high reproducible wavenumbers increased the validation prediction performance. RMSEP decreased to 0.12 g/100ml whereas models developed from 1,060 and 516 wavenumbers provided RMSEP of 0.20 and 0.21 g/100ml of milk, respectively (Figure 4.4). Using a stricter spectral range based on reliable spectroscopic information allowed improving the

RMSEP by 40 and 43% respectively. Theoretically, PLS algorithm should down weight the contribution of wavenumbers with no information. Unfortunately, the size of the calibration set is usually not large enough to fully discard those regions. Only Foss spectra were used for the need of the test, but more and more datasets combine data from spectrometers of different brands. Development of models without taking into account the reproducibility of the information between the instruments and the brands will inevitably decrease model robustness and prediction quality when applied to a large number of diverse spectrometers. Complementary to the use of reliable spectral areas, standardization procedures were developed to harmonize the spectral format among spectrometers. However, those methodologies rely on the fact that the information is correlated at a same wavenumber between the instruments. Standardization methods depend on this reproducibility and cannot improve it, and therefore do not dispense from using reproducible spectral areas in modelling step. Consequently, even if information could be highlighted in the water and NIR regions during genetic studies carried out in specific conditions (e.g. often in only one milk laboratory), the development of models for use at a large scale should rely on spectroscopic basis, i.e. reproducibility of information at specific wavenumbers.

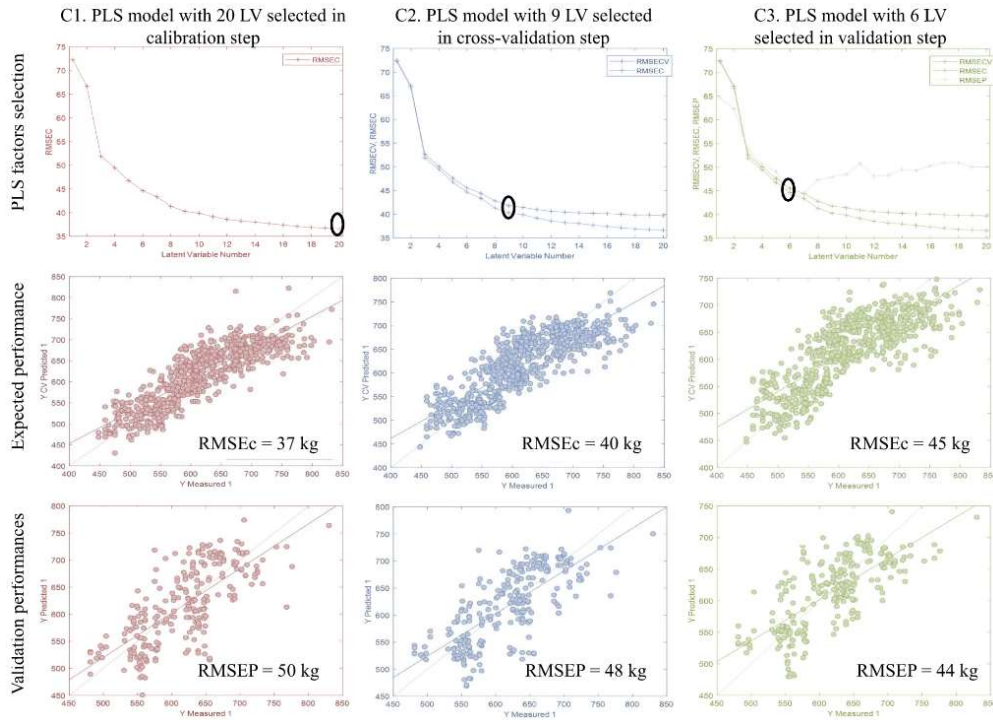


**Figure 4.4.** Results in cross-validation and external-validation for 3 models developed with different spectral ranges.

### 4.5.3. Model complexity

As mentioned in 4.5.1 and 4.5.2, the key element for a model designed to be used at a large scale is its robustness, i.e. its capacity to provide good predictions under

various conditions. However, most of models developed by research institutions use specific datasets with limited size and not representative for real field conditions. Besides, studies generally aim to estimate the best potential of MIR analysis to predict a trait of interest, by giving a strong focus on the maximization of prediction performance (typically the  $R^2$ ) on these datasets. Maximizing model performances on a reduced dataset is usually achieved through an increase of model complexity, which will most likely lead to overfitting, resulting in models that will not generalize well in other conditions.



**Figure 4.5.** Calibration and validation statistical performances of three models predicting the body weight of cows using different numbers of PLS latent variables, selected after the calibration (left), cross-validation (central) and in validation (right) steps.

The effect of model complexity on robustness is illustrated by modifying a very common parameter, the number of PLS factors or latent variables. The dairy cow body weight dataset, containing 1033 records collected from 241 cows was used (see table 4.1). A fraction of 25% of cows (252 records) was randomly excluded to constitute a validation dataset. The remaining 75% of cows (781 records) were used as a calibration dataset. The maximum number of latent variables was fixed to 20. Three models, C1, C2 and C3, were developed using different methods to select the LV number. C1 was developed by fixing the number of LV based on the lowest RMSEc in the calibration step, which resulted in 20 LV. C2 was developed by fixing

the number of LV based on the point of breaking of the RMSEcv slope during the cross-validation step, which resulted in 9 LV. C3 was developed by fixing the number of LV based on the point of breaking of RMSEP slope observed during the validation step, which was also the first local minimum, and resulted in 6 LV. The best apparent calibration was performed with C1 with a RMSEc of 37 kg, compared to 40 kg and 45 kg for C2 and C3 respectively (Figure 4.5). However, when testing the model in a validation step with different cows, C3 obtained the best performance with a RMSEP of 44 kg, compared to 50 kg and 48 kg for C1 and C2. The use of less complex model resulted in an improvement of the RMSEP by 12% and 8% respectively.

This illustrates the fact that model complexity, even if it can improve the prediction on a specific subset, will generally decrease robustness even when the model is applied on a similar subset, i.e. randomly selected cows from the same dataset. The issue of model complexity and overfitting is of course larger than the number of PLS latent variables, and the complexity as well as the inherent structure of the different modeling algorithms strongly affect the robustness of final predictions. Inclusion of additional complexity, steps, layers in models results potentially in additional sources of errors, and for models to be used at a large scale, simplicity can be considered as a safeguard in order to obtain robust predictions.

#### ***4.5.4. Evaluation of models/predictions***

##### ***4.5.4.1. Which indicators to evaluate a model?***

There are plenty of indicators to report in order to evaluate a quantitative model. It is important to describe the characteristics of samples, the reference method used to measure the trait of interest along with an estimate of its standard error of laboratory (SEL), the distribution of the dataset, the data acquisition and the calibration and validation methodologies (Williams et al. [61]). Among them, some are of particular interest. A first important criterion when evaluating a model is its accuracy, which is defined as the closeness of agreement between a test result and the reference value (ISO [62]). Therefore, accuracy combines the random error (precision) and the common systematic error or bias (trueness). Its calculation is based on the difference between reference and predicted values. Those differences can be synthesized using different indicators like the mean absolute error (MAE), the median absolute deviation of predictions (MADP; more robust to extreme values), the mean square error (MSE) or the root mean square error (RMSE). The RMSE is the most frequently used in the literature and allows consequently to compare studies predicting the same trait. The RMSE also allows the users to compare with expected precision limits or thresholds. Fearn [63] stated that the RMSEP is the best parameter to evaluate a model and that it should be sufficient to judge whether the method fit for the intended purpose or not. Nonetheless, in order to compare models with different units and order

of magnitude, he also stated that a dimension-less criteria has to be considered, such as a ratio. In this way, the  $R^2$ , RER, RPD or relative RMSE can be considered.

The  $R^2$  is a well-known dimension-less parameter frequently reported and easy to evaluate. Several studies mentioned accuracy through the  $R^2$  calculation. However, this is not consistent with the ISO definition of accuracy as mentioned earlier because the  $R^2$  implies the distribution of the reference data in the denominator. Hence, the  $R^2$  is a different indicator telling how the predicted values fit with the distribution of the reference data, and is *de facto* highly sensitive to this distribution. To illustrate this, Davies and Fearn [64] simulated a very basic modelling application and showed that, when the range was artificially reduced from 10-20 to 12-18 and then to 13-17, the corresponding  $R^2$  were 0.91, 0.79 and 0.63 whereas the RMSEP stayed at the same level of 0.62. Those results clearly show that evaluation of a model only through the  $R^2$  does not reflect the accuracy and could be misleading. The dependence on the range is particularly detrimental when the distribution of the reference phenotype is narrow, or when the sampling methodology does not allow to gather sufficient variability (e.g. unique farm, similar lactation stage...). Complementary to the accuracy, the evaluation of how the model fit with the reference data is of interest as well, and in this purpose, it seems more relevant to use the  $R^2$  instead of the correlation coefficient  $r$  in order to allow comparison with other studies and avoid confusion for non-statistician readers.

In order to get rid of the dimension of the recorded trait, the RMSE can be combined into a ratio. The frequently used RPD is obtained by dividing the standard deviation of the reference values in the calibration dataset by the RMSE. A high RPD means a precise model associated with large reference variability. However, as mentioned by Esbensen et al. [65], using only the RPD is critical because it is very sensitive to the distribution and could be inflated by few samples, as the  $R^2$ . Alternatively, the RMSE can also be normalized by the range to derive the ratio error range (RER). Unfortunately, it is even more susceptible to be inflated by one extreme sample (Fearn [63]).

Otherwise, the relative RMSE, which divides the RMSE by the mean or the median, is not sensitive to the range, contrary to previous indicators. However, it is strongly impacted by the concentration of the trait, as two models with similar RMSE,  $R^2$  and ranges but at a different level (e.g. a trait ranging from 10 to 20 and a trait ranging from 80 to 90) would have a different relative RMSE despite of similar predictive capacities.

Consequently, in order to compare and classify models, both  $R^2$ , RPD and relative RMSE are imperfect indicators that cannot be used alone but that are highly complementary. Additionally, RPD has an exponential relationship with both RMSE and  $R^2$  (as  $RPD = 1/\sqrt{1 - R^2}$ ), and is therefore more effective to evaluate model with

high statistical performances but less relevant than those parameters to discriminate models with low statistical performances. In order to evaluate and compare models, a well-known classification of model through RPD has been developed for grain analysis by Williams [66] in which models with RPD higher than 3 are considered as fair models. However, Williams also mentioned that such RPD values could be difficult to obtain with respect to the sample to analyze, the precision of the reference method and the spectroscopic methods. Therefore, the classification of models through RPD has been adapted in Williams [67] to specific context of NIR analysis of forage and feed to predict complex phenotypes, and the thresholds were adapted and lowered. However, the global RPD classification by Williams [66] is not fully adapted for the dairy sector as well. For some traits, even when RPD of models are low, it has been mentioned that MIR models could be of great interest because the reference value is very difficult or expensive to measure, e.g. methane emitted by dairy cows or blood composition (Gengler et al. [2]). To overcome their low accuracy, some models can also be useful through the use of thresholds or relative values. It is consequently of interest to derive a classification of models predicting dairy phenotypes regarding the proper characteristics and needs of the dairy sector. Additionally, due to the complementarity of indicators, it seems relevant to attempt the classification of models by combining the RPD, relative RMSE and  $R^2$ . Finally, the classification is done using a non-supervised method on a large number of existing model in order to compare with results deriving from Williams's knowledge and experience.

**Table 4.4.** Characteristics of the 7 K-mean clusters resulting from the classification of 57 milk MIR models following their mean-centered cross-validation RPD, relative RMSE and  $R^2$ .

Cluster	RPD <sub>cv</sub>	Relative RMSE <sub>cv</sub>	$R^2$ <sub>cv</sub>	Interpretation for application
1	> 6	<5%	> 0.97	Any application
2	4.2 - 6	<10%	0.94 - 0.97	Quality control
3	3 - 4.2	<10%	0.89 - 0.94	Quantitative screening
4	2 - 3	<25%	0.74 - 0.89	Rough screening
5	1.5 - 2	<25%	0.55 - 0.74	Allows to compare groups, discriminate high or low values
6	1.5 - 2	>25%	0.55 - 0.74	Highly imprecise, can be used to detect extreme values
7	< 1.5	-	< 0.55	Not recommended

A total of 57 quantitative models predicting fine milk components, blood components, status of dairy cows and technological properties of milk were used in order to perform a non-supervised K-means Near Neighbor (KNN) clustering of models, with seven clusters, following 3 parameters: their mean-centered cross-

validation RPD, relative RMSE and  $R^2$ . KNN classification gather data following Euclidian distances between samples and center of clusters. The seven groups of models are detailed in Table 4.4 with the range of their performance indicators and their interpretation for potential applications. Clustering of all models is shown in appendix 4.1.

Using different models from another context, the derived classification appears very similar to Williams [67] classification. Globally, the classification is mainly guided by RPD and  $R^2$ , except for models with RPD below 2, where relative RMSE discriminate models with high imprecision due to non-normal distributions. The obtained classification is also finer in the low range of RPD. In details, the best models are included in the first 3 clusters and are those allowing the prediction of major fatty acid groups, casein or citrate in milk. Those models are characterized by  $RPD > 3$  and  $relative\ RMSE < 10$ , and quantitative predicted values can be used as such. The models included in cluster 4 allowed the prediction of some minor fatty acids and calcium, with  $RPD > 2$  and  $RMSE < 25\%$ . Such models allow rough screening of samples to approximate the quantity of predicted phenotype. Models in the cluster 5 predicted for instance complex phenotypes such as dry matter intake, N efficiency of cows or methane emissions, with RPD between 1.5 and 2, and  $relative\ RMSE < 25\%$ . Such models are of great interest for the dairy sector, but the predicted values cannot be used as such because of the relative imprecision. Nevertheless, they can be used to compare groups of records or groups of cows and discriminate high or low values. Models in the cluster 6 have the same RPD and  $R^2$  ranges than cluster 5, but have  $relative\ RMSE > 25\%$ , and therefore are highly imprecise. Those models are predicting indirect components in blood or milk which have highly skewed distributions affecting the models. They can only distinguish extreme high values from low values. Finally, the last cluster is grouping models with low RPD and  $R^2$  and high relative RMSE, and the use of such model is uncertain.

This classification is specific to the set of models used and has not the purpose to be used as a reference, but rather to highlight the need of a critical view on the indicators and highlight the interest of combining several indicators. Assessment of accuracy through the RMSE is essential and should stay the gold standard. In order to compare models with different units and magnitudes, combination of several indicators such as RPD and  $R^2$  with relative RMSE seems relevant as well. Besides this, a graphical representation of the measured vs. predicted records is highly valuable as it allows to catch, in one graph, the distribution of reference and predicted data, the quality of the fit and the accuracy of model. These considerations only apply to the quantitative models, whereas more and more studies attempt to classify phenotypes through discriminant methods implying other relevant indicators. Finally, models needs to be evaluated by the final users in regard of their objective, as the

performance needed is totally different whether the aim is to perform quality control, discriminate high and low values or achieve a genetic study on a very large population.

#### ***4.5.4.2. Evaluation of model, cross-validation and validation strategies***

Models with poor performance on calibration dataset will generally produce poor prediction on new data. Nevertheless, good performances on calibration dataset does not automatically means good prediction on new data. Calibration results show how efficiently the model can find an interpretation for the variation present in the calibration dataset. This does not mean, however, that this interpretation is correct and will work on new data. Therefore, relying only on calibration results to assess model quality would provide in many cases overoptimistic results. To evaluate both the accuracy and the robustness of the model, rigorous assessment has to be performed with samples that were not used at any step of model building.

Strictly speaking, in the framework of supervised methods, three steps should be distinguished: calibration, optimization and validation (or testing). During calibration, the modelling algorithm searches for the best values of the parameters. For example, in the case of a PLS regression, the parameters are the combination of the scores and the loadings. Besides this, different model hyperparameters are often compared for the learning algorithm. Hyperparameters are external configurations that are usually set by the user. In the case of PLS regression, the typical setting is the number of latent variables. The step consisting in making the best choice for the value of hyperparameter(s) is called optimization. In practice, the limited size of datasets frequently penalizes the use of truly separate datasets for calibration and optimization, and an intermediate step of cross-validation is used to test the model with internal samples and tune the hyperparameters. Cross-validation includes among others N-fold cross-validation, splitting randomly or by cycling a calibration dataset into n folds to test them iteratively after removing of the calibration dataset, and leave-one-out cross-validation (LOOCV), removing records one by one in order to test them individually. Finally, the validation (or test) step consists in applying the model on samples that were not used in the two previous steps and deriving performance indicators. In datasets collected from dairy cows, several samples are often dependent as coming from the same cow at different lactation stage, or sampled in the same herd with all cows having a common management and feed diet. Testing the model using internal randomly selected samples, with no fundamental difference between calibration and test dataset, is considered as cross-validation. Testing the model by samples being different on at least one criterion is considered as a validation, as testing a model with other cows, herds, countries, seasons or years. The use of those different strategies results in different challenging levels for the model. Therefore, conclusions should be different following the objective of the study, assessing the potential of the model in a precise context or testing its robustness to be used in real field conditions.



To illustrate the importance of the validation strategy, four models were cross-validated and validated using different approaches. These models predicted respectively the quantity of methane emitted by dairy cows, the dry matter intake, the fresh cheese yield and the content of milk  $\beta$ -hydroxybutyrate. Tested approaches were 10-fold cross-validation, LOOCV, cow-validation by removing and validating each cow one by one, and herd-validation, by removing and validating each herd one by one. In the last two approaches, the validation sets are therefore cow and herd independent, respectively. For all models except the fresh cheese yield one, the different herds came from different countries, and herd validation was actually a country validation as well.

**Table 4.5.** Evaluation of 4 models through RMSE and with different validation strategies.

	Methane (g/d/cow)	emitted	Dry Matter (kg/d/cow)	intake	Fresh Cheese Yield (g curd/100g milk)	Milk ( $\mu$ mol/L)	BHB
	12 LV		10 LV		10 LV	13 LV	
	# groups	RMSE	GH # mean groups	RMSE	GH # mean groups	RMSE	GH # mean groups
Calibration	1	64.8	1.00 1	3.6	1.00 1	4.8	1.00 1
Cross-validation LOO	1089	67.4	1.01 1033	3.9	1.01 373	5.3	1.09 2168
Cross-validation 10 folds	10	67.1	1.01 10	3.9	1.02 10	5.3	1.10 10
Validation per cow	299	70.8	1.04 129	3.7	1.04 204	5.4	1.09 225
Validation per herd	7	98.7	1.37 3	5.2	1.44 8	6.0	1.27 6

<sup>1</sup>Number of group of records iteratively excluded

<sup>2</sup>GH mean: mean of standardized Mahalanobis distances between excluded records and calibration samples

The number of latent variables was fixed by selecting visually the point of breaking of the RMSE<sub>cv</sub> slope during the 10 folds cross-validation, within a limit of 20. RMSE were calculated on the entire set of residuals between predicted and measured values instead of averaging individual cow or herd RMSE, which would artificially reduce the final RMSE. Results of calibration, of cross-validation and validation strategies are shown in Table 4.5.

For each model, calibration provided overoptimistic results when compared to cross-validation. LOO and 10-fold cross-validation provided RMSE in the same order of magnitude. The herd validation resulted in the highest RMSE, with values increased from 44%, 37%, 21% and 35% when compared to calibration results, for the 4 datasets respectively. As expected, the cow-validations and the herd-validations were the more challenging methodologies. At first glance, conclusions from such herd validation procedure could be that a model is not working to predict the trait of interest.

For example, Wang and Bovenhuis [68] studied the impact of cross-validation strategies on a methane emission model. According to their conclusions, random cross-validation was indicating overoptimistic results, whereas the lower performances obtained in herd validation showed that the MIR model was not appropriate to predict methane emission. However, a risk in the interpretation of results is to confound the potential of a method to predict a trait in a specific context and its robustness to be used with real field data. Indeed, as explained in 3.1, IR models are sensitive to extrapolation and a model tested on samples being very different from calibration dataset is likely to provide high RMSEP. Dardenne [50] listed the statistical parameters to report in validation step when communicating on an IR model, and concluded that the most important was to report the GH or the NH of the validation samples as they report if the model covering the samples to predict. During the validation procedures, the mean of GH distances between calibration and validation sets were computed. They are reported in Table 4.5. In the present example, a strong correlation was observed between the RMSE of each validation dataset and its GH means, highlighting the link between validation results and capacity of the calibration dataset to cover variability of validation dataset. Poor results in herd validation are resulting of herd specificities (e.g. diets, genetic, management...) that are different between calibration and validation herds due to the lack of representativeness of the global dataset. This confirms that extrapolation regarding zootechnical specificities or spectral data is likely leading to larger predictions errors. Consequently, RMSEP reflects both accuracy and the robustness of the model, contrary to RMSE<sub>cv</sub>, which evaluates a model with similar calibration and test conditions and only reflects accuracy. High RMSEP in validation step does not necessarily means that a method is not suitable, but that the model may not be robust enough to be used under such conditions. Therefore, in order to assess the robustness of a model, it is useful to validate it with independent data (other herd, country, year) and to analyze RMSEP values in the light of GH values. Considering this, in order to use models in routine, model robustness should be updated by adding new samples until covering an acceptable proportion of samples when tested under real field conditions.

#### ***4.5.4.3. Quality assurance methods for phenotyping at a large scale***

As mentioned in 4.5.1.1 and 4.5.1.2, extrapolation is dangerous regarding both reference and spectral data, and is likely to generate predictions of poor quality. Currently, there is no defined quality assurance methodology or indicators to secure the use of models and predictions. However, even if models are robust, when applied to large and variable populations, there is always a proportion of samples not covered by the variability used to build the models and for which the predicted phenotype should not be considered as a tool for management of cows or industrial process due to the risk of erroneous predictions.

An easy solution to prevent extrapolation regarding the reference data range is to use with caution the predicted values outside of the range of the calibration dataset. Obviously, the width of the calibration reference data is directly affecting the proportion of samples to discard, and sampling methodologies to increase this range are of great interest.

Extrapolation regarding the spectral data is a risk that need to be addressed as well. As mentioned by Dardenne [50] and highlighted in section 3.4.2, the quality of prediction is depending of the ability of the calibration dataset to cover the spectral variability of a new sample to predict. Consequently, when applying a model on a new sample to predict a phenotype, there is a need to assess whether the spectra of this sample is included within the spectral variability of the calibration dataset. This can be done through the GH, calculated as defined in section 2. Zhang et al. [69] evaluated the use of a GH threshold as a system of data quality check when using a model into a Chinese population of cows. He concluded that excluding samples with high GH was an appropriate methodology to avoid extrapolation and poor quality of spectral data. When studying the GH threshold level, between 1 and 10, a GH of 5 was considered as a good balance between improving the quality of predictions without losing too many samples. The proportion of excluded samples is directly depending of the spectral variability of the calibration dataset, and sampling procedure to increase it are of great interest. This is also depending of the quality and stability of spectral data and highly complementary with spectral standardization procedures.

Another data cleaning procedure mentioned by Zhang et al. [69] is the comparison between the fat prediction by the manufacturer model and a fat prediction derived from a model developed internally. The fat prediction from the manufacturer model is usually corrected with slope and bias correction at a high frequency and predictions are expected to be close to reference values. On the opposite, a model developed internally and applied on raw spectra data is not corrected and the absolute difference between both predictions is susceptible to reflect the potential spectral outliers. In this study, different thresholds were tested to discard samples, with fat differences from 0.08 to 0.40 g/dL of milk, and it was observed that all threshold improved the predictions, and that lower thresholds induced higher accuracy and higher loss of data. The absolute difference of 0.3 g/dL was considered a good balance between loss of data and gain in accuracy. In another study, the same method was used in order to clean a calibration dataset, to prevent errors arising from analytical issues and incorrect associations between spectra and milk composition data, with the threshold of 0.2 g/dL (Grelet et al. [14]).

### ***4.5.5. Use and transfer of models***

#### ***4.5.5.1. Differences in spectral responses between instruments and spectral standardization***

As stated in introduction, an exponential number of MIR models were developed within the last years, mainly by research organizations. However, the real use of such models by milk recording organizations or industries is not frequent due to the lack of transfer between research and these field organizations. Among others, sharing of data and models is difficult due to differences of spectral responses among spectrometers. The difficulties to share models between all instruments complicates the large-scale phenotyping through MIR analysis. This issue is usually overcome for prediction of main components (i.e. fat, protein, lactose...) by adjusting frequently the predictions using slope and bias correction based on standard samples with known reference values. For phenotypes such as indirect prediction of milk components, phenotype related to cow status or blood composition, it is expensive or impossible to create standard samples with known reference values, and the predictions cannot be adjusted in this way. The use of MIR models predicting such traits would inevitably lead to erroneous predictions when applied on different instruments (Rodriguez et al. [70]). To face this issue, Shenk et al. [71] initiated standardization methodologies from 1985 in order to transfer models on different NIR instruments. More recently, different methods were tested in order to standardize the format of MIR spectra of milk. A first method aimed to create standardization coefficients based on common milk samples analyzed on all instruments using a regression methodology, e.g. Piece-Wise Direct Standardization (Grelet et al. [72], Tiplady et al. [73]). Other methods consisted to harmonize the instruments through the use of historical database in association with regression methods (Bonfatti et al. [74], Tiplady et al. [73]) or by mean-centering the absorbance values (Zaalberg et al. [75]). Those different methods are complementary, but nevertheless, all authors concluded that instruments should be standardized using the methodology associating common samples and regression methods to guarantee correct application of models to generate predictions in routine. To illustrate the impact of spectral standardization, results obtained in a previous work are displayed (Grelet et al. [47]). In this study, three MIR models predicting daily methane emitted by dairy cows, poly-unsaturated fatty acids (PUFA) and fresh individual laboratory cheese yield of milk (FCY) were applied on spectra from common samples analyzed on different spectrometers. The standardization of 66 instruments into a reference instrument resulted in an improvement of 80%, 80% and 86% of the prediction reproducibility between the spectrometers for the 3 models, respectively. Standard deviations between the predictions of the 66 instruments were reduced from 126 to 25 g/day for CH<sub>4</sub>, from 0.0346 to 0.0069 g/100ml for PUFA and from 4.89 to 0.69 g curd/100g milk for FCY. The harmonization of predictions is illustrated in Figure 6 by the plot of PUFA predictions by the reference instrument, the non-standardized

instruments and the standardized instruments. By reducing the bias between spectrometers predictions, the method enables the transfer and the use of multiples models within a network of standardized instruments in order to generate phenotypes at a large scale.

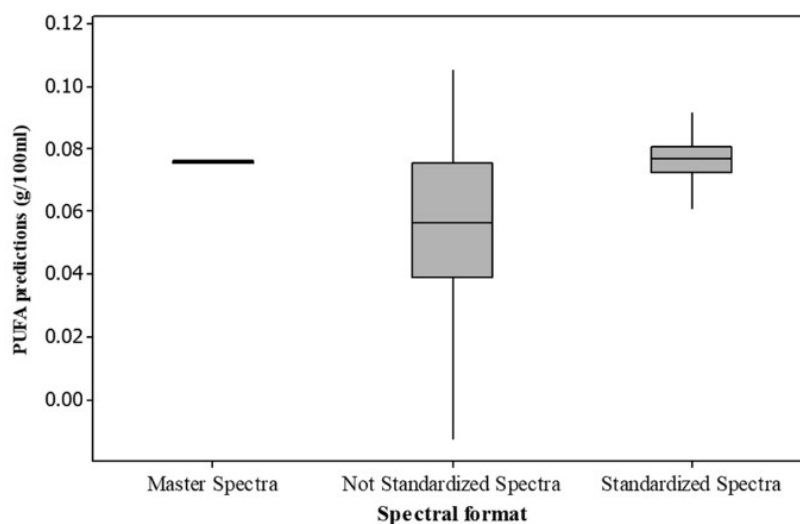
In the objective of undertaking a routine standardization based on the analysis of common standardization samples, the OptiMir network has been constituted in 2011, including 26 spectrometers from 3 countries. The standardization network includes currently 111 MIR spectrometers being standardized monthly, located in Austria, Belgium, France, Germany, Ireland, Luxembourg, Netherlands, Spain, Switzerland and United Kingdom. For research purposes and exchange of data, punctual standardizations of instruments were achieved in Australia, Brazil, Canada, China, Denmark, and Finland.

#### *4.5.5.2. Instability of individual instruments over time*

Deterioration of component pieces, environmental factors, maintenance operations, piece replacements, detector instability and electronic drifts lead to the instability of each individual spectrometer response over time (Bonfatti et al. [74]). Such spectral drifts and perturbations are susceptible to affect all the predictions based on spectra and may lead to providing wrong predicted values to farmers and industries. It is therefore of great interest to monitor the stability of each individual spectrometer over time, whatever the milk analyzed. A simple methodology to observe this instability over time is to analyze daily on an instrument, during a long period, to the extent considered as identical as possible, UHT milk samples coming from a common batch. The application of a model on the spectra, and the plot in time of the predictions reveal drifts and perturbations occurring on the spectrometer as shown in Figure 4.7, where a strong perturbation occurred after a maintenance operation and is detected using a fat model. Even if a model is transferred to an instrument after a spectral standardization, the predictions will inevitably be erroneous after drifts or perturbations occurring from multiples sources mentioned earlier. There is consequently a need to detect and take into account those perturbations in order to generate correct predictions over time. In Figure 4.7, a fat model is used as it is a well-known variable, easy to evaluate by users. However, the spectral drifts are susceptible to affect differently the various spectral areas. Hence it is relevant to detect perturbations by combining several complementary models as each model rely on specific spectral areas through the weight of B coefficients. Any model could be used, as the predicted value is not used in itself and only the stability of the predictions is monitored, through for example visual observation and standard deviation. Nonetheless, inaccuracy of low precision models may mask instability of spectrometers and good models are preferred. The use of models and predictions is only a way to synthetize the stability into familiar variables, and drifts and

perturbations can also be detected directly through changes in spectra, by looking to latent variables or spectral distances. When this methodology is based on UHT milk samples, it is important to note that homogenizer issues and deterioration cannot be detected as UHT milk is already homogenized.

A frequent execution (e.g. monthly) of standardization procedure harmonizing the spectra into a stable reference is an easy way to insure stability of instruments over the long term. Complementary, the combination of a rapid instability detection system with the possibility to run a new analysis of standardization samples, in order to generate new coefficients correcting for this drift, is considered as a solution to overcome the instability in-between two standardization procedures.



**Figure 4.6.** Boxplots representation of PUFA (Poly-Unsaturated Fatty Acids) predictions, for a common sample of the inter-laboratories study analyzed on 66 instruments, for master spectra, non-standardized spectra, and spectra after standardization, after removing of aberrant values (Grelet et al., 2017).

## 4.6. Discussion

The objective of this work was to highlight some elements to take into account when working with MIR data in order to predict dairy phenotypes at a large scale. The different aspects combine to obtain reliable predictions, with improvement of the RMSEP from 12% to 43% in the displayed examples. It also highlights the impact of the spectral standardization on the prediction reproducibility, with an improvement up to 86% with the tested models. Finally, it aims to enhance the complementarity between the different validation procedures, statistical parameters and quality assurance methods. Those results do not aim to quantify the impact of different

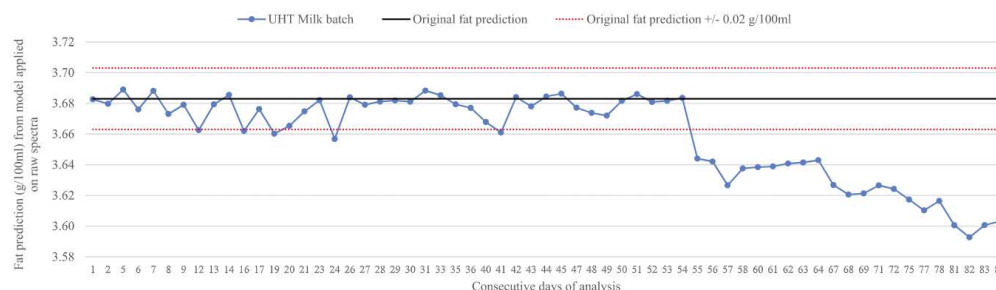
factors, as there is only one example, and further demonstrations would be needed in order to fully demonstrate those different concepts.

This list inspired from our experience is of course not exhaustive. Many other aspects play a role in the quality of final predictions. For example, spectral pre-treatment and the modelling algorithm were not discussed in this document, but for each dataset, it is easily possible to compare methods by testing and assessing their performances using an external validation. Then, this work only consider the data aspect, whereas the practical aspects preceding the data treatment are at least as important because they condition the data quality. Among others, the representativeness of the milk sampling, the storage condition, the sample heating and homogenization, the temperature of analysis, the right association between the sample, the sample ID and the spectra are all important sources of errors that should not underestimated. Those multiples sources of errors and their high frequency in MIR datasets argue for necessity of cleaning procedures in the calibration datasets as well, in order to identify and remove those samples affecting the models performances. In regard with the data acquisition, the quality of the reference measure is a major factor as well. Indeed, the SEL, which can be calculated by analyzing several times the same sample, directly define the maximum  $R^2$  to obtain in the calibration step from the following formula (Dardenne [50]), with  $SD_y$  being the standard deviation of the reference value and SEL being the standard error of laboratory.

$$\text{Equation 2.} \quad R_{max}^2 = (SD_y^2 - SEL^2)/SD_y^2$$

Consequently, in the case of complex phenotypes to records, such as DMI or methane emissions, the high imprecision associated with the reference measure inevitably prevent to obtain highly precise models and it is important to consider all parameters influencing the quality of reference data. More and more researches focus on the prediction of molecules with low concentration in milk or on the status of the cow (health, fertility, efficiency...). As shown earlier, the performance of those models are likely to be limited when compared to models predicting major milk molecules. Despite of their relative imprecision, they can be of great interest when they predict phenotypes not accessible at large scale through gold standard measure. In such cases, it is even more important to communicate on the objective performances of models, using comparable methods and indicators, to avoid misleading uses. Indeed, the final user should be able to easily identify the potential of the model in order to adapt the prediction use. It will bring added values to those models when compared with predictions generated by the instrument manufacturers, for which the statistical performances of equations are generally not communicated, and values such as milk urea are usually considered as such by the final users despite of a high imprecision. Knowing the accuracy associated with predictions, it is possible for the milk recording

organizations, the industry or the geneticists to use directly the value generated, or use thresholds and relative values in order to bypass the inaccuracy. Finally, based on objective evaluation of performances, it is also to the final user to decide whether the models are precise enough for their needs.



**Figure 4.7.** Plot of fat predictions over time originating from a fat model applied on spectra from an instrument analyzing a common batch of UHT milk during a period of 3 months.

## 4.7. Conclusions

The simple examples shown in this work aim to highlight the positive impact on prediction quality of (1) integrating relevant spectral and reference variability in the calibration datasets, (2) selecting the spectral regions based on spectroscopic information, (3) limiting the complexity of models. The second objective was to bring the attention on the relevance of evaluation methods and indicators, for both models and predictions. Finally, when using and transferring models on large number of instruments, the importance of spectral standardization and monitoring of individual instrument stability was highlighted. It encourages for development of guidelines and good practices, to develop and evaluate models with comparable methods, as well as for international collaborations in order to constitute large and robust datasets and enable the use of models in routine conditions.

## 4.8. Acknowledgment

The authors are especially indebted to Octave Christophe, Claire Darimont, Matthieu Dubuisson, Olivier Genard, Petimat Kitaeva, and Mohamed El Morabit for their daily precious work. The authors acknowledge the Walloon partners in Belgium, Elevéo by AWE group and the Comité du lait de Battice. The authors are grateful to European Milk Recording and particularly to Julie Leblois and Xavier Massart for the management of the standardization network as well as to the laboratories constituting it: Gebietslabor der landwirtschaftskammer Vorarlberg, Kärntnermilch, Milchprüfing



Oö, Prüfstele labor Rotholz, Qlm-qualitätslabor St-Michael, Qualitätslabor Niederösterreich (Austria), Comité du lait de Battice (Belgium), Valio oy, regional laboratory of Seinäjoki (Finland), Adecl62, Analis 53, Bentley instruments, Cialso, Galilait, Labilait, Lacolait, Lda39, Lial-fc, Lial-mc, Lidal, Lilano, Lilco, Mael, Mylab 29, Mylab 35, Uriane (France), Mlv-mv, Lkv-nrw, Lkv Sachen, Lkv Schleswig-Holstein, Lkv Rlp-Saar, Lkv Weser-Ems, Mpr-Kirchheim (Germany), Independant milk laboratories, Irish dairy board (Ireland), Convis (Luxembourg), Qlip (Netherlands), Licoval, Universitat politècnica de València (Spain), Suissselab (Switzerland), Nmr Wolverhampton, Nmr Glasgow (United kingdom). The authors thanks the partners that contributed to constitute the datasets: Aarhus University (Denmark), Natural Resources Institute (Finland), IDELE, INRA (France), Leibniz Institute for Farm Animal Biology (Germany), University College Dublin, Irish Cattle Breeding Federation, Teagasc (Ireland), Research Center for Animal Production and Aquaculture (Italy), Agri-Food and Biosciences Institute (Northern Ireland), ETH Zürich, Qualitas (Switzerland), Royal Veterinary College, Scottish Rural University College, (UK).

## **4.9. Funding and disclaimer**

The authors would like to thank the Walloon Region and European Union's INTERREG NWE program for their financial support of the OptiMIR and the HappyMoo project, and the European Union's Seventh Framework Program for funding the GplusE project under grant agreement n° 613689. The views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission.

## 4.10. References

- [1] M. De Marchi, V. Toffanin, M. Cassandro, M. Penasa, 2014. Invited review: Mid-infrared spectroscopy as phenotyping tool for milk traits. *J. Dairy Sci.* 97(3), 1171-1186. <https://doi.org/10.3168/jds.2013-6799>
- [2] N. Gengler, H. Soyeurt, F. Dehareng, C. Bastin, F. Colinet, H. Hammami, M.L. Vanrobays, A. Lainé, S. Vanderick, C. Grelet, A. Vanlierde, E. Froidmont, P. Dardenne, 2016. Capitalizing on fine milk composition for breeding and management of dairy cows. *J. Dairy Sci.* 99(5), 4071-4079. <https://doi.org/10.3168/jds.2015-10140>
- [3] S.L. Smith, S. J. Denholm, M.P. Coffey, E. Wall, 2019. Energy profiling of dairy cows from routine milk mid-infrared analysis. *J. Dairy Sci.* 102(12), 11169-11179. <https://doi.org/10.3168/jds.2018-16112>
- [4] T.K. Belay, M. Svendsen, Z.M. Kowalski, T. Ådnøy, 2017. Genetic parameters of blood  $\beta$ -hydroxybutyrate predicted from milk infrared spectra and clinical ketosis, and their associations with milk production traits in Norwegian Red cows. *J. Dairy Sci.* 100(8), 6298-6311. <https://doi.org/10.3168/jds.2016-12458>
- [5] R.S. Pralle, K.W. Weigel, H.M. White, 2018. Predicting blood  $\beta$ -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. <https://doi.org/10.3168/jds.2017-14076>
- [6] C. Grelet, A. Vanlierde, M. Hostens, L. Foldager, M. Salavati, K.L. Ingvarsen, M. Crowe, M.T. Sorensen, E. Froidmont, C.P. Ferris, C. Marchitelli, F. Becker, T. Larsen, F. Carter, GplusE consortium, F. Dehareng, 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. <https://doi.org/10.1017/S1751731118001751>
- [7] V. Bonfatti, S.A. Turner, B. Kuhn-Sherlock, T.D.W. Luke, P.N. Ho, C.V.C. Phyn, J.E. Pryce, 2019. Prediction of blood  $\beta$ -hydroxybutyrate content and occurrence of hyperketonemia in early-lactation, pasture-grazed dairy cows using milk infrared spectra. *J. Dairy Sci.* 102(7), 6466-6476. <https://doi.org/10.3168/jds.2018-15988>
- [8] A. Benedet, M. Franzoi, M. Penasa, E. Pellattiero, M. De Marchi, 2019. Prediction of blood metabolites from milk mid-infrared spectra in early-lactation cows. *J. Dairy Sci.* 102(12), 11298-11307. <https://doi.org/10.3168/jds.2019-16937>

- [9] T. D. W. Luke, S. Rochfort, W.J. Wales, V. Bonfatti, L. Maret, J.E. Pryce, 2019. Metabolic profiling of early-lactation dairy cows using milk mid-infrared spectra. *J. Dairy Sci.* 102(2), 1747-1760. <https://doi.org/10.3168/jds.2018-15103>
- [10] J. De Koster, M. Salavati, C. Grelet, M.A. Crowe, E. Matthews, R. O'Flaherty, R., G. Opsomer, L. Foldager, GplusE, M. Hostens, 2019. Prediction of metabolic clusters in early-lactation dairy cows using models based on milk biomarkers. *J. Dairy Sci.* 102(3), 2631-2644. <https://doi.org/10.3168/jds.2018-15533>
- [11] H. Soyeurt, E. Froidmont, I. Dufrasne, D. Hailemariam, Z. Wang, C. Bertozzi, F.G. Colinet, F. Dehareng, N. Gengler, 2019. Contribution of milk mid-infrared spectrum to improve the accuracy of test-day body weight predicted from stage, lactation number, month of test and milk yield. *Livest. Sci.* 227, 82-89. <https://doi.org/10.1016/j.livsci.2019.07.007>
- [12] N. Shetty, P. Løvendahl, M.S. Lund, A.J. Buitenhuis, 2017. Prediction and validation of residual feed intake and dry matter intake in Danish lactating dairy cows using mid-infrared spectroscopy of milk. *J. Dairy Sci.* 100(1), 253-264. <https://doi.org/10.3168/jds.2016-11609>
- [13] S.E. Wallén, E. Prestlökken, T.H.E. Meuwissen, S. McParland, D.P. Berry, 2018. Milk mid-infrared spectral data as a tool to predict feed intake in lactating Norwegian Red dairy cows. *J. Dairy Sci.* 101(7), 6232-6243. <https://doi.org/10.3168/jds.2017-13874>
- [14] C. Grelet, E. Froidmont, L. Foldager, M. Salavati, M. Hostens, C.P. Ferris, K.L. Ingvarsen, M.A. Crowe, M.T. Sorensen, J.A. Fernandez Pierna, A. Vanlierde, N. Gengler, GplusE Consortium and F. Dehareng, 2020. Potential of milk mid-infrared spectra to predict nitrogen use efficiency of individual dairy cows in early lactation, *J. Dairy Sci.* In press, <https://doi.org/10.3168/jds.2019-17910>
- [15] M. Klaffenböck, A. Steinwider, C. Fasching, G. Terler, L. Gruber, G. Mészáros, J. Sölkner, 2017. The use of mid-infrared spectrometry to estimate the ration composition of lactating dairy cows. *J. Dairy Sci.* 100(7), 5411-5421. <https://doi.org/10.3168/jds.2016-12189>
- [16] N. Shetty, G. Difford, J. Lassen, P. Løvendahl, A.J. Buitenhuis, 2017. Predicting methane emissions of lactating Danish Holstein cows using Fourier transform mid-infrared spectroscopy of milk. *J. Dairy Sci.* 100(11), 9052-9060. <https://doi.org/10.3168/jds.2017-13014>

- [17] A. Vanlierde, H. Soyeurt, N. Gengler, F.G. Colinet, E. Froidmont, M. Kreuzer, F. Grandl, M. Bell, P. Lund, D.W. Olijhoek, M. Eugène, C. Martin, B. Kuhla, F. Dehareng, 2018. Development of an equation for estimating methane emissions of dairy cows from milk Fourier transform mid-infrared spectra by using reference data obtained exclusively from respiration chambers. *J. Dairy Sci.* 101(8), 7618-7624. <https://doi.org/10.3168/jds.2018-14472>
- [18] A. Vanlierde, 2019. Development of equations to predict methane eructed by lactating cows from milk mid-infrared spectra. Doctoral dissertation, Université de Liège, Liège, Belgique. <http://hdl.handle.net/2268/238609>
- [19] S. van Gastelen, H. Mollenhorst, E.C. Antunes-Fernandes, K.A. Hettinga, G.G. van Burgsteden, J. Dijkstra, J.L.W. Rademaker, 2018. Predicting enteric methane emission of dairy cows with milk Fourier-transform infrared spectra and gas chromatography-based milk fatty acid profiles. *J. Dairy Sci.* 101(6), 5582-5598. <https://doi.org/10.3168/jds.2017-13052Get>
- [20] T.M. Denninger, A. Schwarm, F. Dohme-Meier, A. Münger, B. Bapst, S. Wegmann, F. Grandl, A. Vanlierde, D. Sorg, S. Ortmann, M. Clauss, M. Kreuzer, 2019. Accuracy of methane emissions predicted from milk mid-infrared spectra and measured by laser methane detectors in Brown Swiss dairy cows. *J. Dairy Sci.* 103(2), 2024-2039. <https://doi.org/10.3168/jds.2019-17101>
- [21] A. Lainé, C. Bastin, C. Grelet, H. Hammami, F.G. Colinet, L.M. Dale, A. Gillon, J. Vandenplas, F. Dehareng, N. Gengler, 2017. Assessing the effect of pregnancy stage on milk composition of dairy cows using mid-infrared spectra. *J. Dairy Sci.* 100(4), 2863-2876. <https://doi.org/10.3168/jds.2016-11736>
- [22] H. Toledo-Alvarado, A.I. Vazquez, G. de los Campos, R.J. Tempelman, G. Bittante, A. Cecchinato, 2018. Diagnosing pregnancy status using infrared spectra and milk composition in dairy cows. *J. Dairy Sci.* 101(3), 2496-2505. <https://doi.org/10.3168/jds.2017-13647>
- [23] P.N. Ho, V. Bonfatti, T.D.W. Luke, J.E. Pryce, 2019. Classifying the fertility of dairy cows using milk mid-infrared spectroscopy. *J. Dairy Sci.* 102(11), 10460-10470. <https://doi.org/10.3168/jds.2019-16412>
- [24] P. Delhez, P.N. Ho, N. Gengler, H. Soyeurt, J.E. Pryce, 2020. Diagnosing the pregnancy status of dairy cows: How useful is milk mid-infrared spectroscopy? *J. Dairy Sci.* 103(4), 3264-3274. <https://doi.org/10.3168/jds.2019-17473>

- [25] L. Rienesl, N. Khayatzadeh, A. Köck, L. Dale, A. Werner, C. Grelet, N. Gengler, F.J. Auer, C. Egger-Danner, X. Massart, J. Sölkner, 2019. Mastitis Detection from Milk Mid-Infrared (MIR) Spectroscopy in Dairy Cows. *Acta Univ. Agric. Silvic. Mendelianae Brun.*, 67(5), 1221-1226. <https://doi.org/10.11118/actaun201967051221>
- [26] A. Mineur, A. Köck, C. Grelet, N. Gengler, C. Egger-Danner, J. Sölkner, 2017. First results in the use of milk mid-infrared spectra in the detection of lameness in Austrian dairy cows. *ACS Agric. Conspec. Sci.* 82(2), 163-166. <https://hrcak.srce.hr/191814>
- [27] A. Mineur, H. Hammami, C. Grelet, C. Egger-Danner, J. Sölkner, N. Gengler, 2020. Investigation of the temporal relationships between milk mid-infrared predicted biomarkers and lameness events in later lactation, *J. Dairy Sci.* in press. <https://doi.org/10.3168/jds.2019-16826>
- [28] V. Bonfatti, P.N. Ho, J.E. Pryce, 2020. Usefulness of milk mid-infrared spectroscopy for predicting lameness score in dairy cows. *J. Dairy Sci.* 103(3), 2534-2544. <https://doi.org/10.3168/jds.2019-17551>
- [29] S. Nayeri, F.S. Schenkel, P. Martin, A. Fleming, J. Jamrozik, F. Malchiodi, L.F. Brito, C.F. Baes, M. Sargolzaei, F. Miglior, 2020. Estimation of genetic parameters for mid-infrared-predicted lactoferrin and milk fat globule size in Holstein cattle. *J. Dairy Sci.* 103(3), 2487-2497. <https://doi.org/10.3168/jds.2019-16850>
- [30] A. Fleming, F.S. Schenkel, J. Chen, F. Malchiodi, V. Bonfatti, R.A. Ali, B. Mallard, M. Corredig, F. Miglior, 2017. Prediction of milk fatty acid content with mid-infrared spectroscopy in Canadian dairy cattle using differently distributed model development sets. *J. Dairy Sci.* 100(6), 5073-5081. <https://doi.org/10.3168/jds.2016-12102>
- [31] A. Fleming, F.S. Schenkel, J. Chen, F. Malchiodi, R.A. Ali, B. Mallard, M. Sargolzaei, M. Corredig, F. Miglior, 2017. Variation in fat globule size in bovine milk and its prediction using mid-infrared spectroscopy. *J. Dairy Sci.* 100(3), 1640-1649. <https://doi.org/10.3168/jds.2016-11427>
- [32] C.L. Manuelian, G. Visentin, C. Boselli, G. Giangolini, M. Cassandro, M. De Marchi, 2017. Prediction of milk coagulation and acidity traits in Mediterranean buffalo milk using Fourier-transform mid-infrared spectroscopy. *J. Dairy Sci.* 100(9), 7083-7087. <https://doi.org/10.3168/jds.2017-12707>

- [33] M.P. Sanchez, M. El Jabri, S. Minéry, V. Wolf, E. Beuvier, C. Laithier, A. Delacroix-Buchet, M. Brochard, D. Boichard, 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(11), 10048-10061. <https://doi.org/10.3168/jds.2018-14878>
- [34] M. El Jabri, M.P. Sanchez, P. Trossat, C. Laithier, V. Wolf, P. Groperrin, E. Beuvier, O. Rolet-Répécaud, S. Gavoye, Y. Gaüzère, O. Belysheva, E. Notz, D. Boichard, A. Delacroix-Buchet, 2019. Comparison of Bayesian and partial least squares regression methods for mid-infrared prediction of cheese-making properties in Montbéliarde cows. *J. Dairy Sci.* 102(8), 6943-6958. <https://doi.org/10.3168/jds.2019-16320>
- [35] P.W. Goedhart, 1990. Comparison of multivariate calibration methods for prediction of feeding value by near infrared reflectance spectroscopy. *NJAS-Wagen. J. Life Sc.* 38(3B), 449-460.
- [36] M. Zeaiter, J.M. Roger, V. Bellon-Maurel, D.N. Rutledge, 2004. Robustness of models developed by multivariate calibration. Part I: The assessment of robustness. *TrAC-Trend Anal. Chem.*, 23(2), 157-170. [https://doi.org/10.1016/S0165-9936\(04\)00307-3](https://doi.org/10.1016/S0165-9936(04)00307-3)
- [37] M. Zeaiter, J.M. Roger, V. Bellon-Maurel, 2005. Robustness of models developed by multivariate calibration. Part II: The influence of pre-processing methods. *TrAC-Trend Anal. Chem.*, 24(5), 437-445. <https://doi.org/10.1016/j.trac.2004.11.023>
- [38] R.M. Balabin, R.Z. Safieva, E.I. Lomakina, 2007. Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction. *Chemometr. Intell. Lab.* 88(2), 183-188. <https://doi.org/10.1016/j.chemolab.2007.04.006>
- [39] H. Soyeurt, P. Dardenne, F. Dehareng, G. Lognay, D. Veselko, M. Marlier, C. Bertozzi, P. Mayeres, N. Gengler, 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 89(9), 3690-3695. [https://doi.org/10.3168/jds.S0022-0302\(06\)72409-2](https://doi.org/10.3168/jds.S0022-0302(06)72409-2)
- [40] H. Soyeurt, F. Dehareng, N. Gengler, S. McParland, E. Wall, D.P. Berry, M. Coffey, P. Dardenne, 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *J. Dairy Sci.* 94(4), 1657-1667. <https://doi.org/10.3168/jds.2010-3408>

- [41] H. Soyeurt, D. Bruwier, J.M. Romnee, N. Gengler, C. Bertozzi, D. Veselko, P. Dardenne, 2009. Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 92(6), 2444-2454. <https://doi.org/10.3168/jds.2008-1734>
- [42] H. Soyeurt, C. Bastin, F. G. Colinet, V. M.R. Arnould, D. P. Berry, E. Wall, F. Dehareng, H. N. Nguyen, P. Dardenne, J. Schefers, J. Vandenplas, K. Weigel, M. Coffey, L. Théron, J. Detilleux, E. Reding, N. Gengler, S. McParland, 2012. Mid-infrared prediction of lactoferrin content in bovine milk: potential indicator of mastitis. *Animal*, 6(11), 1830-1838. <https://doi.org/10.1017/S1751731112000791>
- [43] F. Dehareng, C. Delfosse, E. Froidmont, H. Soyeurt, C. Martin, N. Gengler, A. Vanlierde, P. Dardenne, 2012. Potential use of milk mid-infrared spectra to predict individual methane emission of dairy cows. *Animal*, 6(10), 1694-1701. <https://doi.org/10.1017/S1751731112000456>
- [44] F. Colinet, T. Troch, V. Baeten, F. Dehareng, P. Dardenne, M. Sindic, N. Gengler, 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. <https://doi.org/10.3920/978-90-8686-816-2>
- [45] C. Grelet, C. Bastin, M. Gelé, J.B. Davière, M. Johan, A. Werner, R. Reding, J.A. Fernandez Pierna, F.G. Colinet, P. Dardenne, N. Gengler, H. Soyeurt, F. Dehareng, 2016. Development of Fourier transform mid-infrared calibrations to predict acetone,  $\beta$ -hydroxybutyrate and citrate contents in bovine milk through a European dairy network. *J. Dairy Sci.* 99(6), 4816-4825. <https://doi.org/10.3168/jds.2015-10477>
- [46] ICAR Dairy Cattle Milk Recording Working Group, 2017, Section 2 -Guidelines for Dairy Cattle Milk Recording, <https://www.icar.org/Guidelines/02-Overview-Cattle-Milk-Recording.pdf>, accessed in march 2020.
- [47] C. Grelet, J.A. Fernández Pierna, P. Dardenne, H. Soyeurt, A. Vanlierde, F. Colinet, N. Gengler, V. Baeten, F. Dehareng, 2017. Standardization of milk mid-infrared spectrometers for the transfer and use of multiple models. *J. Dairy Sci.* 100(10), 7910-7921. <https://doi.org/10.3168/jds.2017-12720>
- [48] J.S. Shenk, M.O. Westerhaus, 1991. Population definition, sample selection, and calibration procedures for near infrared reflectance spectroscopy. *Crop Sci.*, 31(2), 469-474. doi:10.2135/cropsci1991.0011183X003100020049x

- [49] J.S. Shenk, M.O. Westerhaus, The application of near infrared reflectance spectroscopy (NIRS) to forage analysis, in: G.C. Fahey (Eds); Forage quality, evaluation, and utilization, Madison, Wisconsin, USA, 1994, pp. 406-449, doi:10.2134/1994.foragequality.c10
- [50] P. Dardenne, 2010. Some considerations about NIR spectroscopy: Closing speech at NIR-2009. NIR news. 21(1):8-14. <https://doi.org/10.1255/nirn.1165>
- [51] P. Dardenne, D. Vincke, V. Baeten, J.A. Fernandez-Pierna, 2015. Untargeted multivariate characterization of contaminants in powdered milk. Final OptiMIR Scientific and Expert Meeting: From milk analysis to advisory tools. Biotechnol. Agron. Soc. Environ. 2015, 19(2), 97-124
- [52] A.R. Castillo, E. Kebreab, D.E. Beever, J.H. Barbi, J.D. Sutton, H.C. Kirby, J. France, 2001. The effect of protein supplementation on nitrogen utilization in lactating dairy cows fed grass silage diets. J. Anim. Sci. 79(1):247-253. <https://doi.org/10.2527/2001.791247x>
- [53] F. Allegrini, J.A. Fernandez Pierna, W.D. Fragoso, A.C. Olivieri, V. Baeten, P. Dardenne, 2016. Regression models based on new local strategies for near infrared spectroscopic data. Anal. Chim. Acta, 933, 50-58. <https://doi.org/10.1016/j.aca.2016.07.006>
- [54] C.E. Eskildsen, M.A. Rasmussen, S.B. Engelsen, L.B. Larsen, N.A. Poulsen, T. Skov, 2014. Quantification of individual fatty acids in bovine milk by infrared spectroscopy and chemometrics: Understanding predictions of highly collinear reference variables. J. Dairy Sci. 97(12), 7940-7951. <https://doi.org/10.3168/jds.2014-8337>
- [55] D. Bertrand, E. Dufour, La spectroscopie infrarouge et ses applications analytiques, Collection Sciences et Techniques Agroalimentaires. Second ed., Editions Tec et Doc, London, Paris, New York, 2000.
- [56] S.K. Andersen, P.W. Hansen, H.V. Andersen, Vibrational spectroscopy in the analysis of dairy products and wine. In: J.M. Chalmers, P.R. Griffiths (Eds), Handbook of vibrational spectroscopy, John Wiley & Sons, Chichester, United Kingdom, 2002. DOI: 10.1002/0470027320
- [57] Li-Chan, E., Chalmers, J. M., and Griffiths, P. R. (Eds.). Applications of vibrational spectroscopy in food science. John Wiley & Sons. Chichester, United Kingdom, 2010.



- [58] R. Karoui, G. Downey, C. Blecker, 2010. Mid-infrared spectroscopy coupled with chemometrics: a tool for the analysis of intact food systems and the exploration of their molecular Structure– Quality relationships– a review. *Chem. Rev.*, 110(10), 6144-6168. <https://doi.org/10.1021/cr100090k>
- [59] G. Bittante, A. Cecchinato, 2013. Genetic analysis of the Fourier-transform infrared spectra of bovine milk with emphasis on individual wavelengths related to specific chemical bonds. *J. Dairy Sci.* 96(9), 5991-6006. <https://doi.org/10.3168/jds.2013-6583>
- [60] Q. Wang, A. Hulzebosch, H. Bovenhuis, 2016. Genetic and environmental variation in bovine milk infrared spectra. *J. Dairy Sci.* 99(8), 6793-6803. <https://doi.org/10.3168/jds.2015-10488>
- [61] P. Williams, P. Dardenne, P. Flinn, 2017. Tutorial: Items to be included in a report on a near infrared spectroscopy project. *J. Near Infrared Spec.*, 25(2), 85-90. <https://doi.org/10.1177/0967033517702395>
- [62] ISO 5725-1 - Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions, International Organization for Standardization, Geneva. 1994.
- [63] T. Fearn, 2002. Assessing calibrations: SEP, RPD, RER and R<sup>2</sup>. *NIR news*, 13(6), 12-13. <https://doi.org/10.1255/nirn.689>
- [64] A. Davies, T. Fearn, 2006. Back to basics: calibration statistics. *Spectroscopy Europe*, 18(2)31-32. [https://www.spectroscopyeurope.com/system/files/pdf/TD\\_18\\_2.pdf](https://www.spectroscopyeurope.com/system/files/pdf/TD_18_2.pdf)
- [65] K.H Esbensen, P. Geladi, A. Larsen, 2014. The RPD myth... *NIR news*, 25(5), 24-28. <https://doi.org/10.1255/nirn.1462>
- [66] P.C. Williams, D.C. Sobering, 1993, Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *J. Near Infrared Spec.* 1, 25-32. <https://www.osapublishing.org/jnirs/abstract.cfm?URI=jnirs-1-1-25>
- [67] P. Williams, 2014, Tutorial: The RPD statistic: a tutorial note, *NIR news*, 25, 22–26, <https://doi.org/10.1255/nirn.1419>
- [68] Q. Wang, H. Bovenhuis, 2019. Validation strategy can result in an overoptimistic view of the ability of milk infrared spectra to predict methane emission of dairy cattle. *J. Dairy Sci.* 102(7), 6288-6295. <https://doi.org/10.3168/jds.2018-15684>

- [69] Zhang, L., Dehareng, F., Grelet, C., Colinet, F., Gengler, N., Brostaux, Y., Soyeurt, H., (unpublished results, submitted in 2020). Ensuring high quality milk mid-infrared predictions through appropriate data quality checks.
- [70] J.D. Rodriguez, B.J. Westenberger, L.F. Buhse, J.F. Kauffman, 2011. Standardization of Raman spectra for transfer of spectral libraries across different instruments. *Analyst* 136:4232–4240. <https://doi.org/10.1039/C1AN15636E>
- [71] J.S. Shenk, M.O. Westerhaus, W.C. Templeton, 1985. Calibration Transfer Between near Infrared Reflectance Spectrophotometers 1. *Crop Sci.*, 25(1), 159-161. doi:10.2135/cropsci1985.0011183X002500010038x
- [72] C. Grelet, J.A. Fernandez Pierna, P. Dardenne, V. Baeten, F. Dehareng, 2015. Standardization of milk mid-infrared spectra from a European dairy network. *J. Dairy Sci.* 98(4), 2150-2160. <https://doi.org/10.3168/jds.2014-8764>
- [73] K.M. Tiplady, R.G. Sherlock, M.D. Littlejohn, J.E.Pryce, S.R. Davis, D.J. Garrick, R.J. Spelman, B.L. Harris, 2019. Strategies for noise reduction and standardization of milk mid-infrared spectra from dairy cattle. *J. Dairy Sci.* 102(7), 6357-6372. <https://doi.org/10.3168/jds.2018-16144>
- [74] V. Bonfatti, A. Fleming, A. Koeck, F. Miglior, 2017. Standardization of milk infrared spectra for the retroactive application of calibration models. *J. Dairy Sci.* 100:2032–2041. <https://doi.org/10.3168/jds.2016-11837>
- [75] R.M. Zaalberg, A.J. Buitenhuis, U.K. Sundekilde, N.A. Poulsen, H. Bovenhuis, 2020. Genetic analysis of orotic acid predicted with Fourier transform infrared milk spectra. *J. Dairy Sci.* 103(4), 3334-3348. <https://doi.org/10.3168/jds.2018-16057>

**Appendix 4.1.** Details of the 7 K-mean clusters resulting from the classification of 57 milk MIR models following their normalized cross-validation RPD, relative RMSE and R<sup>2</sup>.

Phenotype	Min	Max	Mean	SD	R <sup>2</sup> cv	Relative RMSEcv	RPDcv	Cluster
Milk SAT FA(g/dL)	0.31	6.97	2.70	0.75	0.99	3%	10.22	1
Milk C18_1cis9 (g/dL)	0.08	2.69	0.76	0.29	0.95	8%	4.35	2
Milk Casein (g/100g)	1.61	4.05	2.66	0.34	0.95	3%	4.46	
Milk LCFA (g/dL)	0.19	4.79	1.59	0.52	0.95	7%	4.52	
Milk MCFA (g/dL)	0.22	5.48	2.00	0.60	0.97	5%	5.53	
Milk MONO FA (g/dL)	0.12	3.42	1.08	0.35	0.97	5%	5.83	
Milk Tot18_1cis (g/dL)	0.09	2.77	0.82	0.31	0.95	8%	4.58	
Milk Total_C18_1 (g/dL)	0.10	2.98	0.94	0.33	0.96	7%	5.18	
Milk UNSAT (g/dL)	0.14	3.86	1.25	0.39	0.97	5%	5.75	
Milk C10 (g/dL)	0.02	0.32	0.11	0.04	0.91	9%	3.37	3
Milk C12 (g/dL)	0.02	0.41	0.13	0.04	0.92	9%	3.62	
Milk C14 (g/dL)	0.05	1.20	0.45	0.13	0.93	7%	3.88	
Milk C16 (g/dL)	0.12	3.32	1.20	0.40	0.94	8%	4.18	
Milk C4 (g/dL)	0.01	0.23	0.10	0.03	0.93	8%	3.67	
Milk C6 (g/dL)	0.01	0.16	0.07	0.02	0.91	9%	3.32	
Milk C8 (g/dL)	0.01	0.11	0.05	0.01	0.91	9%	3.29	
Milk Citrates (mmol/L)	3.88	16.12	9.04	2.21	0.89	8%	3.04	4
Milk SCFA (g/dL)	0.05	0.80	0.35	0.10	0.93	7%	3.88	
Milk C17 (g/dL)	0.00	0.09	0.03	0.01	0.80	13%	2.24	
Milk C18 (g/dL)	0.05	1.32	0.40	0.15	0.84	14%	2.51	
Milk Calcium (mg/kg)	593	1743	1149	135	0.82	5%	2.34	
Milk Odd fatty acids (g/dL)	0.03	0.50	0.16	0.04	0.83	10%	2.41	
Milk PUFA (g/dL)	0.02	0.53	0.16	0.05	0.77	13%	2.10	
Milk Total_Trans (g/dL)	0.02	0.75	0.16	0.08	0.80	19%	2.26	
Tot18_1trans (g/dL)	0.01	0.57	0.13	0.06	0.79	21%	2.17	
Cheese process r (s)	319	1653	906	231	0.58	16%	1.54	5
Dry matter intake (kg/d)	8.8	36.2	19.9	4.5	0.71	12%	1.83	
Fresh cheese yield (g curd/100g milk)	7.40	47.93	26.76	6.45	0.73	12%	1.91	
Methane emitted (g/d)	180	786	413	102	0.68	14%	1.79	
Milk C14_1 (g/dL)	0.00	0.15	0.04	0.02	0.68	21%	1.78	
Milk C16_1c (g/dL)	0.01	0.24	0.07	0.03	0.73	20%	1.91	
Milk C18_2c9c12 (g/dL)	0.00	0.17	0.06	0.02	0.72	19%	1.91	
Milk C18_3c9c12c15 (g/dL)	0.00	0.09	0.02	0.01	0.68	22%	1.77	
Milk isoanteiso FA (g/dL)	0.02	0.28	0.09	0.03	0.75	14%	2.00	
Milk Magnesium (mg/kg)	61	157	100	13	0.72	7%	1.88	
Milk omega3 (g/dL)	0.00	0.11	0.03	0.01	0.66	22%	1.73	
Milk omega6 (g/dL)	0.01	0.33	0.10	0.03	0.72	14%	1.89	
Milk Phosphorus (mg/kg)	509	1447	999	124	0.75	6%	1.99	
Milk Potassium (mg/kg)	819	1985	1524	147	0.55	6%	1.48	
Milk Tot18_2 (g/dL)	0.01	0.32	0.10	0.03	0.69	15%	1.79	
N efficiency (%)	9.8	81.7	36.9	10.3	0.71	15%	1.87	
N losses (kg/d)	0.04	0.81	0.31	0.11	0.65	20%	1.69	
Weight of cows(kg)	448	832	617	73	0.70	6%	1.83	
Blood BHB (mmol/L)	0.19	3.46	0.77	0.48	0.70	35%	1.81	6
Blood IGF-I (mg/L)	13	436	107	71	0.61	42%	1.59	
Lactoferrin (mg/L)	7	1248	299	222	0.66	44%	1.71	
Milk BHB (mmol/L)	0.05	1.60	0.22	0.17	0.75	46%	1.97	
Milk C18_2c9t11 (g/dL)	0.00	0.14	0.03	0.02	0.74	37%	1.95	
Blood Glucose (mmol/L)	1.93	4.51	3.47	0.47	0.44	10%	1.33	7
Blood NEFA (µekv/L)	26	1956	672	440	0.39	51%	1.28	
Cheese process k20 (s)	160	386	225	39	0.34	13%	1.24	
Milk Glucose Free (mmol/L)	0.00	0.69	0.24	0.11	0.50	32%	1.41	
Milk Glucose6Phosphate (mmol/L)	0.00	0.93	0.16	0.08	0.49	36%	1.40	
Milk IsoCitrate (mmol/L)	0.02	2.90	0.17	0.10	0.11	55%	1.06	
Milk Natrium (mg/kg)	234	1273	356	91	0.44	15%	1.34	
Milk Progesterone (ng/ml)	0.50	22.44	5.22	2.74	0.08	50%	1.05	
Milk Uric Acid (µmol/L)	2.4	348.5	158.8	54.6	0.32	28%	1.22	



# Chapitre 5:

---

**Discussion générale**



## 5.1. Introduction

Comme on a pu le voir au Chapitre 4, on dit d'une équation qu'elle est robuste quand elle est capable de générer des prédictions précises dans diverses conditions, en mode « tout-terrain » ou encore, que la robustesse d'un modèle d'étalonnage multivarié est sa capacité à prédire de façon stable (statistiquement), face à des perturbations ciblées sous des conditions standards (Zeaiter et al, 2004). Les facteurs qui influencent la qualité et la robustesse de la prédiction sont nombreux comme nous l'avons déjà mentionné dans les chapitres précédents. Cependant, tout n'a pas encore été abordé et l'expérience acquise durant les deux dernières décennies peut être utile pour compléter l'information. C'est pourquoi, cette discussion visera à expliciter les points d'attention à apporter lors de la construction d'une équation mais aussi lors de l'utilisation de cette équation par les laboratoires laitiers ou les organismes d'élevage. Ainsi, cette discussion sera structurée en six grandes thématiques relatives : à l'échantillonnage, aux analyses de référence, aux modèles de prédiction, à la normalisation, à la standardisation et aux méthodes de contrôle qualité pouvant être appliquées. Par souci de clarté, chacun de ces points sera discuté séparément (Figure 5.1.).



**Figure 5.1.** Points d'attention visant à assurer la qualité des prédictions FT-MIR du lait et d'accroître leur robustesse.

Cependant, les points d'attention peuvent varier en fonction du type de modèles visés. Il est donc important avant de commencer cette discussion de les définir. On pourrait les résumer en quatre grandes familles : les modèles directs, les modèles pseudo-directs, les modèles indirects et les modèles particuliers.

Les modèles directs sont des modèles prédictifs relatifs à des caractères dont la concentration est supérieure à 100 ppm dans le lait et dont les analyses de référence permettant de les quantifier peuvent être directement appliquées aux échantillons de lait ayant servis aux prises de spectres FT-MIR. Cette limite de 100 ppm est une estimation de la limite de détection de la spectrométrie FT-MIR selon Dardenne et al. (2015). Cette famille de modèles furent les premiers développés en spectrométrie FT-MIR dans le lait et regroupe l'ensemble des prédictions pour les composés majeurs (MG, MAT et lactose), mais également d'autres éléments mineurs, dont par exemple : l'urée (Godden et al., 2000 ; Hansen, 1998), les citrates (Grelet et al., 2016), ou les principaux AG (Soyeurt et al., 2006 et 2011, Rutten et al., 2009).

Les modèles pseudo-directs ont certaines similarités avec les modèles directs mais présentent la particularité de concerner des caractères ayant une concentration dans le lait inférieure ou égale à 100 ppm. Vu ces concentrations inférieures au seuil limite de détection, on ne peut expliquer le potentiel à prédire de ces modèles que par des interactions possibles avec d'autres composés présents plus massivement dans le lait. A titre d'exemples, ces modèles concernent la prédiction de l'acétone (van Knegsel et al., 2010; Grelet et al., 2016), du BHB du lait (Grelet et al., 2016) ou encore des minéraux majeurs (Soyeurt et al., 2009 ; Christophe et al, 2021).

Les modèles indirects reprennent l'ensemble des modèles prédictifs relatifs à des phénotypes dont les analyses de référence pour les quantifier ne peuvent pas être directement appliquées sur les échantillons de lait ayant servis aux prises de spectres FT-MIR. Ces phénotypes sont multiples. A titre d'exemples, ces modèles concernent la prédiction du stress chez les animaux (Bahadi et al., 2021), du poids corporelle (Soyeurt et al, 2019), de la quantité de CH<sub>4</sub> érucée (Dehareng et al., 2012 ; Vanlierde et al., 2015 et 2021), ou encore de l'efficacité azotée (Grelet et al., 2020).

Certains modèles ne pouvant être classés dans aucune des 3 familles précédentes, une dernière famille reste à définir. Celle-ci est relative aux modèles particuliers dont aucune quantification de référence pour le phénotype visé ne peut être obtenue. C'est, par exemple, le cas pour les modèles visant à détecter les échantillons anormaux ou encore la présence d'une possible adultération (Cassoli et al., 2011 ; Hansen and Holroyd, 2019 ; Spieß et al., 2021).

## 5.2. Echantillonnage

Si on aborde le point relatif à l'échantillonnage en premier, c'est parce qu'il est encore trop souvent à l'origine de nombreux problèmes mal maîtrisés et qu'il

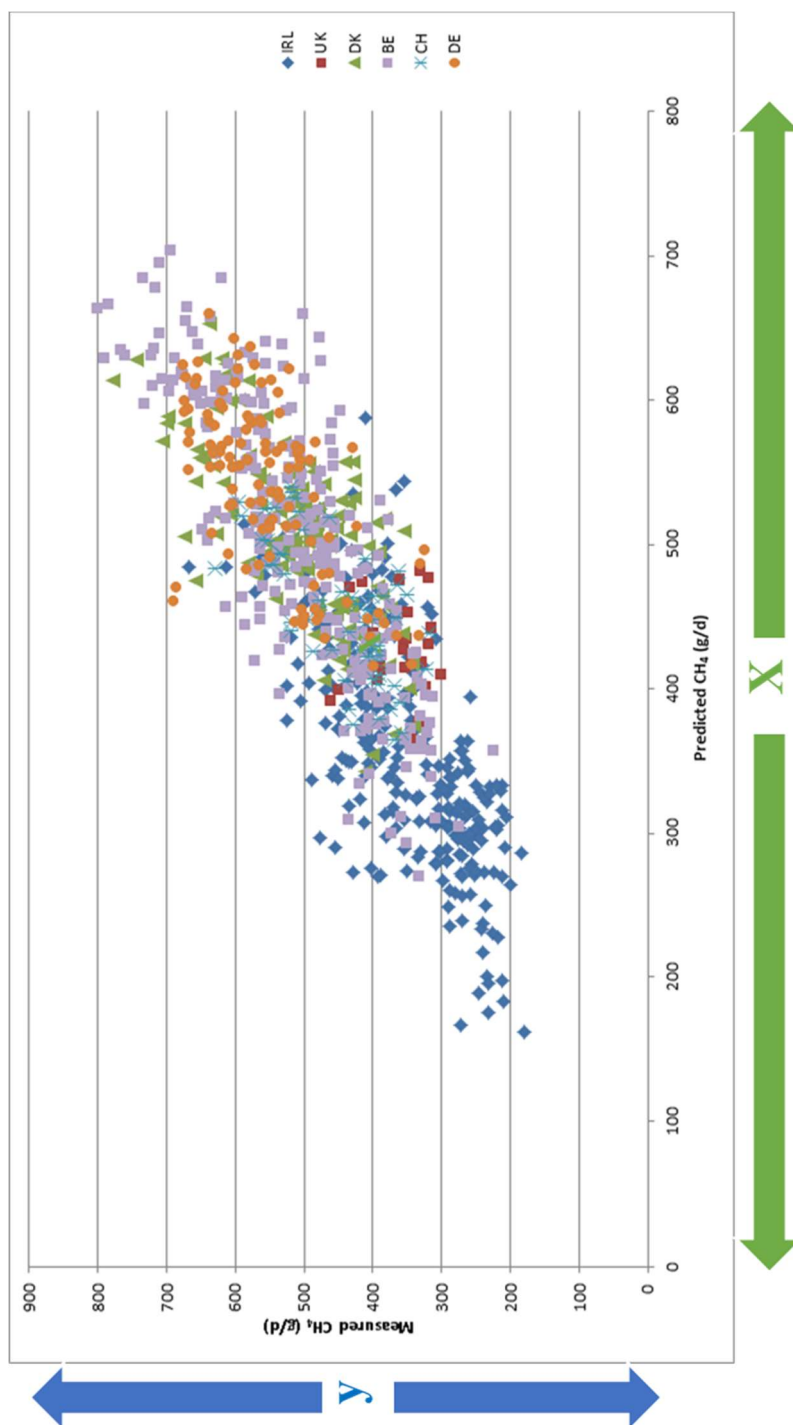


contribue de manière importante à l'incertitude du modèle. La qualité de l'échantillonnage est en quelque sorte la colonne vertébrale de toute tentative de construction d'un nouveau modèle de prédiction, et se doit donc d'être particulièrement soignée. Un plan d'échantillonnage mal construit donnera systématiquement des modèles de faible qualité. Ce plan d'échantillonnage devra être représentatif de la population étudiée mais aussi tenir compte des contraintes techniques, éthiques ou financières relatives à l'acquisition des données elles-mêmes. Essayons ici de lister les points les plus importants.

### ***5.2.1. Couvrir un maximum de variabilité***

Comme nous l'avons vu au Chapitre 4, afin d'augmenter la robustesse du modèle, le plan d'échantillonnage devra être élaboré en couvrant le maximum de la variabilité du phénotype visé par la modélisation au sein de la population de référence. En d'autres termes, cela implique de collecter les données en quantité et variabilité suffisantes pour couvrir l'ensemble des gammes attendues en routine pour le phénotype étudié. Pour atteindre cet objectif, deux stratégies ont été envisagées dans le présent travail de thèse et ce sont révélées être très efficaces (Figure 5.2).

La première stratégie consiste en un échantillonnage cherchant à maximiser la variabilité spectrale ( $X$ ), comme présenté et discuté aux Chapitres 2 et 4, en calculant les valeurs de la distance standardisée de Mahalanobis (distance GH ou en anglais Global-H distance) et la distance H entre voisins (distance NH ou en anglais neighboring H distance) (Shenk and Westerhaus, 1991). Concrètement, comme mentionné dans le Chapitre 4, on sélectionnera de préférence une partie des échantillons qui présentent les valeurs de NH les plus élevées pour assurer une bonne répartition des échantillons de calibration dans l'ensemble des données, et en prenant des échantillons avec des valeurs de GH élevées pour essayer de bien couvrir les parties extrêmes de l'ensemble des données de calibration. Cette méthode est très intéressante surtout si on satisfait aux 2 contraintes suivantes : 1) disposer d'une base de données spectrales massive relative par exemple aux analyses effectuées dans le cadre du contrôle laitier ou liée aux paiements du lait permettant ainsi de passer au crible la variabilité spectrale de toute la population et 2) avoir un phénotype ayant une empreinte dans le signal spectral FT-MIR ou du moins des caractères qui lui sont liés. Cette stratégie fut particulièrement intéressante pour améliorer la robustesse des modèles prédictifs des teneurs en AG du lait (Chapitre 2).



**Figure 5.2.** Stratégies permettant d'accroître la variabilité des modèles : grâce à la variabilité spectrale X, grâce à la variabilité des valeurs de référence y et aux échanges de données via des collaborations internationales (d'après Vanliende et al., 2021)

La seconde stratégie repose sur un échantillonnage cherchant à maximiser la variabilité du phénotype ( $y$ ) en utilisant ses sources de variation. Ainsi, cette stratégie repose d'abord sur une recherche bibliographique afin de collecter toute information utile qui pourrait expliquer les facteurs de variation du paramètre recherché. Cette stratégie est intéressante lorsque le phénotype à modéliser n'a pas directement d'empreinte dans le lait. Celle-ci a, par exemple, été utilisée pour développer le modèle de prédiction de la quantité de  $\text{CH}_4$  éructé (Chapitre 3). On a ainsi particulièrement veillé à couvrir au maximum l'ensemble de la période de lactation, les parités, les races, les régimes alimentaires, les systèmes de gestion des troupeaux, différentes régions, etc. Cette deuxième stratégie est universelle et peut répondre à n'importe quel cas de figure. Cependant, des contraintes éthiques, pratiques ou financières peuvent nous empêcher d'acquérir suffisamment de données sur la population visée (par ex. la population des vaches laitières wallonnes). Citons, par exemple, le développement d'un modèle permettant de détecter de façon précoce les mammites en utilisant le biomarqueur N-acetyl- $\beta$ -D-glucosaminidase ou NAGase. Dans ce cas, il n'existe pas de base de données spectrales utilisable car les enregistrements présents dans les bases de données conventionnelles ne sont pas relatifs à l'analyse spectrale de laits mammitiques. De plus, la NAGase n'est pas mesurée dans le lait mais dans le sérum sanguin et sa quantification doit être immédiate pour éviter toute détérioration. Cette urgence ne permet donc pas de stocker l'échantillon pour une analyse ultérieure en fonction de la réponse spectrale obtenue. D'autre part, chercher à accroître la variabilité dans ce cas de figure, nous obligerait à provoquer des mammites, ce qui éthiquement parlant est complexe et difficile à réaliser. La seule solution trouvée pour pallier ce problème a été d'accroître le set d'échantillons via de nombreuses collaborations internationales (Grelet et al., 2022). En sortant des contraintes de la population de référence (par ex. les vaches laitières wallonnes), on a pu augmenter la variabilité du jeu de données. Créer des collaborations internationales permet de composer avec les contraintes financières, pratiques et/ou éthiques. Cela s'est aussi vérifié dans le développement des modèles prédictifs liés aux AG (Chapitre 2), et durant les phases d'amélioration des modèles  $\text{CH}_4$  (Vanlierde et al., 2021). En combinant une approche multi-pays et multi-races, nous avons pu augmenter la variabilité de la seule population de référence. Cela a permis de renforcer la robustesse des modèles développés par l'augmentation du jeu de données et d'étendre leur utilité au-delà des frontières wallonnes. En revanche, l'accumulation de données provenant de nombreuses sources internationales nécessite au préalable de trouver des stratégies innovantes en termes de construction de consortiums prêts à accepter d'échanger des données (Gengler et al., 2022).

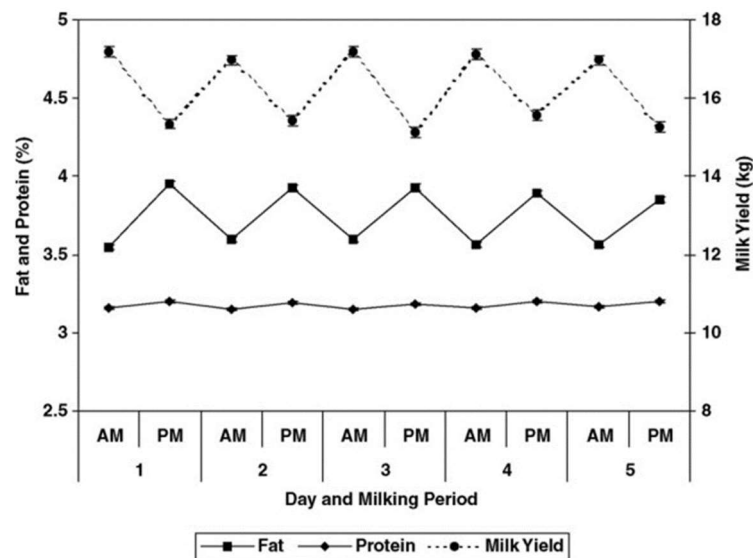
Pour se faciliter la tâche liée à l'échantillonnage, l'envie peut être grande de créer des sets d'échantillons dopés par l'élément à prédire en laboratoire. La création de modèle à partir d'un set d'échantillons dopés devra être réalisé sur base d'un

protocole particulièrement strict, afin d'éviter toute erreur ou création d'artefact qui pourrait mener à des interprétations erronées. C'est d'autant plus vrai pour des modèles de type PLS qui utilisent les relations existantes entre les points spectraux. En effet, en créant ces échantillons dopés, on pourrait déstructurer les relations qui existent entre les points spectraux et rendre in fine le modèle PLS inutilisable sur de vrais échantillons de lait. Citons, par exemple, les études liées au développement de modèles de détection de la tétracycline dans le lait, avec des concentrations bien inférieures à la limite de détection de 100 ppm, utilisant des gammes variant de 4 à 2000 ppb (Sivakesava and Irudayaraj, 2002) ou de 10 à 400 ppb (Casarrubias-Torres et al., 2018). Bien que ces auteurs aient obtenu des résultats encourageants ( $R^2 = 0.98$  et  $0.99$  ; SEP = 387 et 1.21 ppb), nous n'avons jamais pu reproduire en interne ces résultats. De plus, aucune publication n'a validé ces résultats à partir d'échantillons de lait issus de vaches traitées avec de la tétracycline et pour lesquels on mesure la concentration en tétracycline par analyse de référence. Les bons résultats obtenus de ces auteurs pourraient s'expliquer par les protocoles qu'ils auraient utilisés pour créer leurs échantillons dopés. En effet, on peut soupçonner que ce serait finalement les ajouts d'eau, dont la réponse spectrale est particulièrement forte, qui seraient détectés et non la tétracycline. Cet exemple montre l'importance de limiter l'utilisation d'échantillons dopés même si cela facilite la création de gamme de variation pour un phénotype laitier spécifique.

### ***5.2.2. Choix du moment d'échantillonnage***

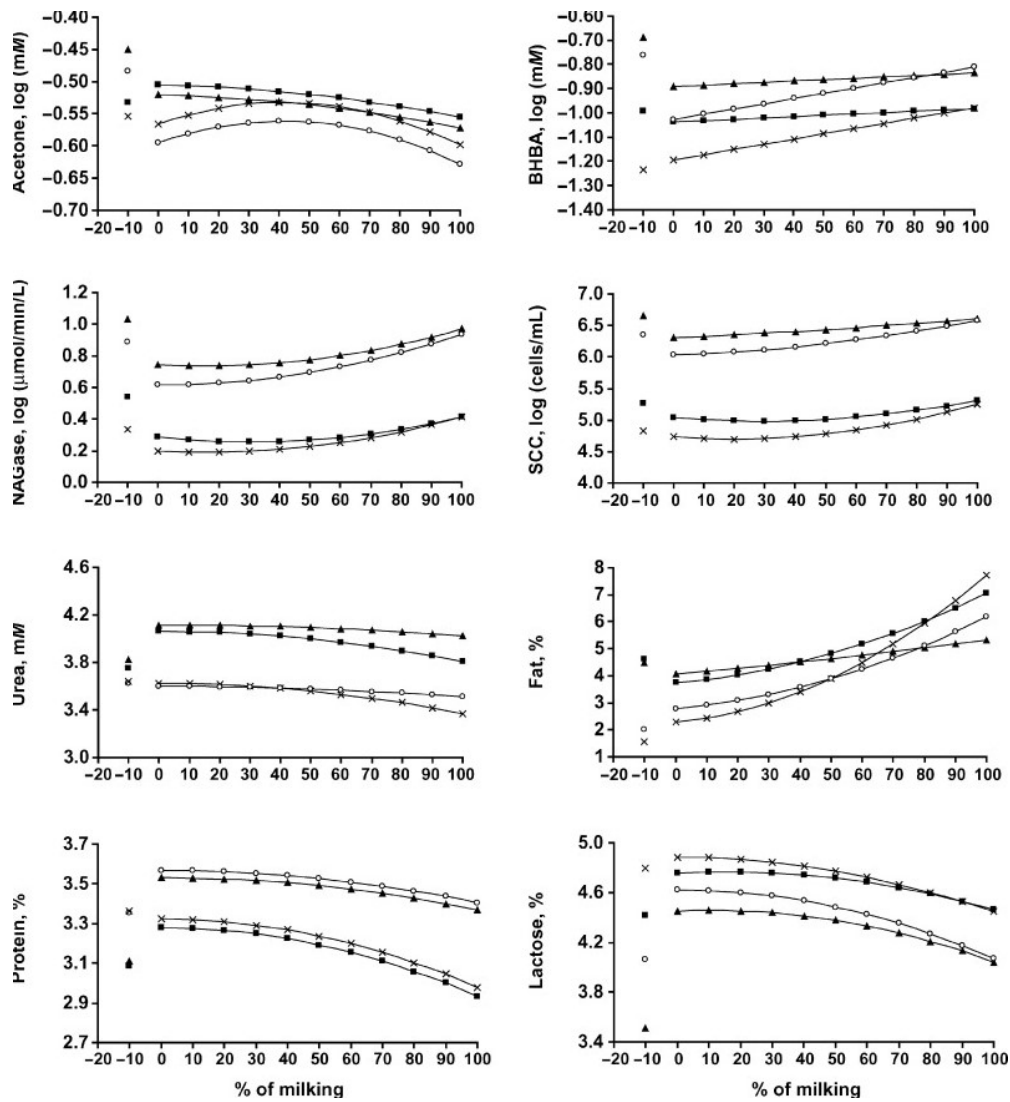
Le choix du moment d'échantillonnage est souvent passé sous silence. Pourtant, il peut se révéler particulièrement important tant pour la donnée spectrale que la valeur de référence.

Il est bien connu que la composition du lait varie entre traites (voir Figure 5.3) et qu'un lait issu de la traite du matin est plus riche en MG et en MAT que celui issu de la traite du soir (p.ex. Lee and Wardrop, 1984 ; Quiest et al., 2008). Ces variations de composition concernent également d'autres composés (Arnould et al., 2015 ; Nielsen et al., 2005). Cela peut s'expliquer par de nombreux facteurs liés notamment à la taille de l'intervalle entre les traites (Aladi et al., 2004), à la production de lait et l'utilisation de l'énergie (Sutton et al., 1985), aux effets du régime alimentaire et à la fréquence et aux moments des repas (Klusmeyer et al., 1990), ou encore le caractère nyctéméral de l'ingestion et de la rumination chez les bovins (Deswysen et al., 1989 et 1993). Comme le montre la Figure 5.4, des variations de la composition du lait sont également présentes au cours d'une même traite et dépendent de l'état de santé du quartier de la mamelle échantillonné (Nielsen et al., 2005 ; Forsbäck et al., 2010). L'ensemble de ces variations doit être pris en compte dans le schéma d'échantillonnage.



**Figure 5.3.** Moyenne  $\pm$  ET de 14 troupeaux traits  $2 \times$  pour la production laitière et les pourcentages de matières grasses et de protéines pour chaque période de traite sur une période de 5 jours. D'après Quiest et al. (2008).

Il faudra donc s'assurer de l'adéquation entre les moments de collecte des différents échantillons composant le set de calibrage du modèle prédictif. Cette adéquation doit aussi être présente entre le spectre (X) et le phénotype de référence (y). Pour les modèles indirects ou pseudo-directs, comme les données de référence sont mesurées directement sur les échantillons ayant servis à l'analyse FT-MIR, l'adéquation est parfaite, comme par exemple, pour les AG (Chapitre 2). Cependant, dans d'autres cas, la synchronisation entre les données de référence et les spectres peut parfois être décalée dans le temps. C'est particulièrement le cas pour les modèles indirects dont fait partie le modèle prédictif du  $\text{CH}_4$  (Chapitre 3). Il faut alors effectuer des études permettant de quantifier le biais dû à la désynchronisation et évaluer son importance au regard de l'incertitude de prédiction globale. C'est ce qui a été fait dans le cas du  $\text{CH}_4$ . Pour rappel, cette désynchronisation s'explique par le fait qu'il existe un biais temporel entre les acides gras volatils (AGV) et la composition en AG du lait. Lors de nos études, nous avons réalisé que ce biais était minime au regard de l'erreur de prédiction globale. Vu la difficulté logistique additionnel qu'entraînait cette désynchronisation entre la mesure de  $\text{CH}_4$  et la prise d'échantillon de lait et le peu de biais que cela entraînait nous avons par la suite utilisé les données de référence en  $\text{CH}_4$  mesurées le même jour que la prise d'échantillon de lait. Cependant, au-delà de la mesure du  $\text{CH}_4$ , nous avons également fait attention à obtenir un spectre représentatif des 24 heures de mesure du  $\text{CH}_4$ .



**Figure 5.4.** Concentrations (moyennes des moindres carrés) des constituants du lait pendant la traite (% de la traite) en fonction de l'intervalle de traite (6 ou 12 h) et de l'état de santé du quartier échantillonné (sain ou malsain). Les moyennes simples des premiers jets de lait sont indiquées à -10% de la traite. Sain6 (■); Malsain6 (▲); Sain12 (×); Malsain12 (○). D'après Nielsen et al. (2005).

D'autres auteurs ont également relevé l'importance du moment d'échantillonnage pour la confection du set de calibrage. Ce fut le cas, notamment, pour le BHB (Duplessis et al., 2019) ou pour les données de CH<sub>4</sub> collectées par des GreenFeed<sup>®2</sup> pendant de longues périodes de prises de mesure (Coppa et al., 2022). D'autre part, ajoutons que le moment d'échantillonnage pour les modèles indirects devrait idéalement couvrir l'ensemble de la variabilité intra et inter-lactationnelle. En effet, pour rappel, nous avons pu observer à la suite de nos premiers modèles méthane (Chapitre 3), que les relations entre le caractère à prédire et les variables pouvaient avoir des coefficients de régression non constants en fonction du temps et que nous devons en tenir compte (Vanlierde et al., 2015). Afin de permettre de détecter et corriger ce type de modèles variables dans le temps, on doit veiller à établir un plan d'échantillonnage qui couvre l'ensemble de la lactation en veillant particulièrement à couvrir les périodes de début, pic et fin de lactation. En conclusion, les modèles indirects sont parfois dépendants des rythmes biologiques et idéalement ce facteur doit être maîtrisé.

### ***5.2.3. Procédure d'échantillonnage***

Au-delà des lieux et moments à couvrir pour maximiser la variabilité et assurer la synchronisation entre le spectre et les données de référence, il est important également de collecter les échantillons de manière rigoureuse et en respectant des règles strictes. Ainsi, des procédures ont été publiées par ICAR ou ISO (ICAR, 2020 ; ISO 707:2008) pour collecter les échantillons de lait représentatif de l'ensemble de la traite (matériel certifié, étalonné et parfaitement propre, vitesse contrôlée évitant les interférences avec les échantillonnages suivants, évitant les contaminations, respectant les températures afin de permettre une parfaite homogénéisation et conservation des échantillons, utilisant des conservateurs approuvés, échantillons parfaitement identifiés, etc...). Ceci est important pour assurer un spectre de qualité et, dans le cas des modèles directs ou pseudo-directs, une quantification de qualité du phénotype visé. Cependant, si on doit être très strict lors de la prise d'échantillons en routine pour fournir par exemple une donnée prédictive à l'éleveur, cela pourrait être moins problématique dans le cadre du développement d'un modèle. En effet, citons, par

---

<sup>2</sup> Le système GreenFeed<sup>®</sup> est un système mobile permettant de mesurer les gaz respiratoires émis par les ruminants. Il est peu invasif pour l'animal et ne présente aucune gêne pour celui-ci. Il se compose d'un distributeur automatique d'aliments avec identification électronique de l'animal, dans lequel l'air (air respiratoire et air ambiant) est aspiré au niveau des naseaux et de la bouche de l'animal pendant que celui-ci consomme l'aliment distribué. Le volume d'air aspiré est mesuré et les teneurs en CH<sub>4</sub> (et en dioxyde de carbone) sont déterminées en continu. Cette méthode permet de relever la quantité de CH<sub>4</sub> et de dioxyde de carbone émises pendant les visites au distributeur.

exemple, le cas où on effectue un échantillonnage à partir d'une traite incomplète pour créer un modèle de prédiction de la MG. Cet échantillon ne sera pas représentatif de la traite mais le contenu en MG de cet échantillon sera correct par rapport à l'échantillon collecté. Il sera donc en phase avec la donnée spectrale générée. Par conséquent, ce biais de prise d'échantillon n'impactera pas la qualité du modèle. Cependant, pour d'autres modèles, et en particulier les modèles indirects, si cette collecte partielle de l'échantillon de lait lors de la traite a engendré une composition du lait non représentative d'un échantillon normal, cela pourrait impacter la qualité du modèle final. Par exemple, si l'on cherche à faire le lien entre la quantification du CH<sub>4</sub> émis sur 24h avec un spectre réellement représentatif de 24h, et qu'un mauvais échantillonnage est réalisé, on risque de détruire partiellement ou complètement les liens entre le lait et le phénotype, et introduire des imprécisions dans le set de calibrage. En effet, on risque de briser certaines relations entre les pics du spectre et le phénotype, et donc de biaiser la construction du modèle final. En conclusion, comme il est difficile de connaître à l'avance l'ensemble des relations, il sera toujours préférable de bien respecter les procédures d'échantillonnage des échantillons de lait et ce sera d'autant plus primordial pour les modèles indirects dont on doit assurer la synchronicité et la représentativité avec le phénotype collecté.

Dans cette section, nous nous sommes concentrés essentiellement sur l'échantillonnage du lait mais n'oublions pas les phénotypes qui y sont associés et qui ne sont pas directement mesurés sur ce même échantillon. Dans ce cas-là aussi, il faudra également respecter un protocole strict d'échantillonnage. Malheureusement dans de nombreux cas, ces protocoles ne sont pas publiés. Il faudra donc s'assurer notamment dans le cadre de collaborations internationales que ce protocole soit standardisé et suffisamment détaillé pour que l'ensemble des prises de mesure soient effectuées de façon parfaitement uniforme. Par exemple, récemment, dans le cadre d'une campagne de collecte d'échantillons visant à la création de modèles liés au stress chronique, la collecte d'échantillons de sérum sanguin et de poils du toupillon de la vache (Grelet et al., 2022a) à travers le réseau européen de partenaires du projet HappyMoo (<https://www.nweurope.eu/projects/project-search/happymoo/>) a été réalisé selon un protocole strict, parfaitement documenté, auquel on a adjoint des vidéos afin d'éviter toute mécompréhension ou interprétation qui entraînerait des différences dans la prise d'échantillon.

### 5.3. Analyses de référence

Le choix de l'analyse de référence est un autre point critique lorsque l'on cherche à développer un modèle prédictif. De la précision de la méthode de référence utilisée va dépendre la performance du modèle, suivant la formule suivante:  $R_{max}^2 = (SD_y^2 - SEL^2)/SD_y^2$ , où  $R_{max}^2$  est le coefficient de détermination maximum pouvant



être obtenu,  $SD_y$  est l'écart-type des valeurs de référence, et SEL est l'erreur de la méthode de référence (Dardenne, 2010). On cherchera donc à minimiser au maximum les erreurs liées aux analyses de référence ou à l'échantillonnage. A titre d'exemple, pour les analyses de  $CH_4$  par la technique du gaz traceur avec de l'hexafluorure de soufre ( $SF_6$ ), des phénomènes non contrôlés peuvent subvenir lors de la collecte des échantillons de gaz : bris ou écrasement du tube de collecte par l'animal, colmatage du filtre, mauvais positionnement du capteur proche de l'orifice bucco-nasal. Ceci peut affecter de façon aléatoire et non contrôlée la qualité de l'échantillonnage et augmenter les erreurs liées aux analyses de référence. Le système de collecte a donc été adapté afin d'être dédoublé (deux systèmes d'échantillonnage indépendants reliés chacun à une bouteille de stockage des gaz), ce qui nous a permis de répéter les mesures chaque jour. Sur base des résultats obtenus, si le coefficient de variation du  $CH_4$  quantifié entre les deux répétitions était inférieur à 15%, la moyenne des deux valeurs était calculée et considérée comme la valeur de référence du  $CH_4$ . Dans le cas contraire, les valeurs étaient considérées comme valeur aberrantes (Vanlierde, 2019). Enfin, on cherchera toujours à utiliser des mesures répétées qui permettent d'estimer la répétabilité de la mesure de référence.

Cependant, dans le cas où plusieurs méthodes de référence sont disponibles, le choix ne se basera pas exclusivement sur la précision de la méthode. En effet, d'autres critères devront aiguiller ce choix afin de composer avec les contraintes techniques, pratiques, financières et éthiques tout en assurant une large variabilité.

Le coût de l'analyse de référence est souvent le premier critère auquel on pense. Cela n'est pas anodin car de celui-ci va dépendre le nombre d'analyses potentiellement réalisables et donc les possibilités de couvrir un maximum de variabilité. Il est parfois préférable de choisir une méthode de référence moins coûteuse et moins précise (quitte à répéter les mesures en double ou triple) afin de permettre d'augmenter le nombre d'analyses de référence et mieux couvrir la variabilité de la population.

Pour augmenter le nombre de données de référence, la facilité de mesure du phénotype à large échelle sera un élément crucial. Sa facilité à pouvoir être mesuré dans des fermes commerciales devra être prise en compte car ces dernières représentent mieux la variabilité existante dans la population dans laquelle sera utilisée le modèle. A titre d'exemple, la mesure des quantités ingérées d'aliments par les vaches laitières n'est possible chez nous que dans des fermes expérimentales. Cela rend difficile l'accès à une large gamme de mesures. Cependant, pour ce caractère, on peut espérer à l'avenir de nouveaux développements pour la mesure de l'ingestion individuelle (Borchersen et al., 2020). Il en va de même pour les mesures de  $CH_4$  en chambres respiratoires ou via la méthode  $SF_6$  qui sont des méthodes complexes, qui requièrent une expertise et un suivi important. L'utilisation de Greenfeed® permet d'envisager une utilisation d'un nombre plus élevé d'animaux sur des périodes de mesure plus longues (Coppa et al., 2022) pour mesurer le  $CH_4$ . Mais il faut garder à l'esprit que cette technique va obliger à devoir répéter les mesures pendant plusieurs

jours pour obtenir une précision suffisante (Arbre et al., 2016 ; Rischewski et al., 2019), ce qui va rendre plus complexe l'acquisition des données.

Le côté invasif de la méthode de référence sera également un élément à prendre en considération notamment pour des raisons de bien-être animal. Par exemple, il est complexe de mesurer correctement les paramètres relatifs à la digestion ruminale sans avoir recours à l'utilisation de fistules ruminales, ce qui exclut leur utilisation dans des fermes commerciales et oblige à faire une demande d'autorisation auprès d'un comité d'éthique (Castillo and Hernández, 2021). L'utilisation de la méthode SF<sub>6</sub> pour quantifier le CH<sub>4</sub> nécessite également l'ingestion d'un bolus qui même s'il est inoffensif pour l'animal peut rebuter l'éleveur et donc contraindre les prises de mesure. D'autre part, même si cela peut paraître anecdotique, il faut veiller à toute interférence de la méthode de mesure avec le comportement de la vache. Par exemple, les mesures en chambre respiratoire pour mesurer le CH<sub>4</sub>, pourrait entraîner des changements de comportement, notamment liés au stress, chez les vaches laitières (Tremblay and Massé, 2016 ; Garnsworthy et al., 2019). Cela peut entraîner un changement de la composition du lait qui ne sera pas représentative d'une vache en pleine santé. Cela pourrait donc aussi biaiser le set de calibrage et donc in fine le développement du modèle.

Un point essentiel sera de toujours préférer la méthode standardisée ou normalisée, plutôt qu'une méthode interne, aussi précise soit elle. Idéalement, elle pourrait même être réalisée dans un laboratoire accrédité (p.ex. ISO 17025). En effet, comme nous l'avons déjà vu plus haut, la possibilité d'échanger des données de référence est un point particulièrement important. D'autre part, lors d'échanges de mesure, on doit s'assurer, par le biais d'essais inter laboratoires, que les conditions utilisées par tous les laboratoires de mesures soient strictement identiques. Ceci n'est pas toujours le cas et peut causer des problèmes de reproductibilité ou d'interprétation des résultats. A titre d'exemple, lors des mesures du profil en AG, les conditions chromatographiques peuvent changer ainsi que le matériel ou les colonnes utilisés, ce qui impactent les résultats. De mes expériences passées, ce point de reproductibilité entre méthodes de référence est un problème à résoudre. En effet, il y a peu de temps, des efforts conjoints ont été réalisés au sein de l'ISO, FIL et AOAC, à partir des travaux de Golay et al. (2009) et Destailats et al. (2007) afin de permettre de trouver des accords sur une méthode commune pour quantifier les AG. Ces travaux ont abouti à la publication de la norme ISO 16958:2015. Néanmoins, trop souvent, nous constatons toujours des divergences entre laboratoires quant au matériel utilisé, aux modes de calcul pour grouper les familles d'AG par exemple, ou encore à l'utilisation d'autres normes ou méthodes. Ce constat est malheureux surtout si on veut échanger des données. Cela peut s'expliquer notamment par le fait que des équations pour quantifier les AG du lait sont déjà vendues par des constructeurs ou utilisées par des organisations d'élevage. Ainsi, les constructeurs de ces modèles souhaiteront que les méthodes de référence qu'ils ont utilisées pour faire leurs modèles soient reconnues comme le standard. Ceci leur permettrait de surcroit de ne pas devoir réinvestir dans de nouvelles analyses de référence qui sont coûteuses. Cela demanderait aussi à

certaines laboratoires de se mettre aux normes ce qui représenterait également un coût. Il y a donc un frein économique/commercial à standardiser ces méthodes. Pour faire avancer les choses, il faudra donc dans ce cas, plus qu'une simple publication de norme.

D'autre part, certaines méthodes commerciales peuvent fonctionner en mode *boîte noire*, en utilisant des prétraitements, des étalonnages ou des corrections non ou mal maîtrisées. Cela peut diminuer la fiabilité ou la précision des mesures (Schori F. and Münger A., 2022). De plus, le manque d'accessibilité aux données brutes des méthodes commerciales peut également être un facteur limitant à l'échange de données.

Un dernier point relatif aux méthodes de référence concerne également la communication des résultats. En effet, les unités des méthodes de référence utilisées sont importantes. Ainsi, pour certains paramètres, elles expriment des choses bien différentes qui exerceront une influence sur la capacité des modèles à prédire un paramètre. Par exemple, si l'on exprime les teneurs en AG en g/dl de lait ou g/100g de MG du lait, les performances seront différentes (Soyeurt et al., 2006). On observe la même chose dans le cas du CH<sub>4</sub> où l'utilisation de différentes unités donneront des modèles différents si l'on parle en termes de production journalière (en litres ou g de CH<sub>4</sub>/jour), d'intensité (en g de CH<sub>4</sub>/Kg de lait), de rendement (en litres ou g de CH<sub>4</sub>/Kg de matières sèches ingérées) ou encore de concentration de CH<sub>4</sub> dans l'air (ppm) (Garnsworthy et al., 2019 ; Negussie et al., 2017). De par notre expérience, il est intéressant de communiquer la donnée de référence dans son unité de départ : g/100g MG dans le cas des acides gras ou encore en g/jour dans le cadre du CH<sub>4</sub> afin de laisser le modélisateur choisir l'unité qui sera la plus utile pour lui. Par exemple, dans le cas du CH<sub>4</sub>, les nutritionnistes préféreront exprimer le CH<sub>4</sub> en g/kg de matière sèche ingérée alors que les généticiens seraient plus intéressés par l'unité exprimée en g/jour car ce n'est pas le fruit d'un quelconque ratio.

## 5.4. Modèles de prédiction

Sans entrer dans le détail des algorithmes de modélisation, je m'attarderai juste dans cette section à présenter quelques points importants à retenir qui concernent la robustesse des modèles.

Pour les calibrations multivariées, un prétraitement peut permettre d'accroître la robustesse d'un modèle de prédiction (Zeaiter et al., 2005), car il élimine les effets indésirés comme les artefacts instrumentaux et permet de se concentrer davantage sur les variations d'intérêt (Engel et al., 2013). Les techniques de prétraitement les plus courantes sont celles qui ne nécessitent pas de valeurs de référence pour les opérations de prétraitement, comme les méthodes de dérivation (p.ex. : dérivées premières ou secondes, via l'algorithme de Savitzky-Golay) et les méthodes de correction de la dispersion (p.ex. correction du signal multiplicatif). Nous avons utilisé une dérivée première dans le Chapitre 2, en revanche, aucun prétraitement n'a été appliquée dans

le Chapitre 3. Cela s'explique par le fait qu'à l'époque nous n'avions, pour ce dernier modèle, utilisé qu'un seul spectromètre. Lors des développements ultérieurs des équations CH<sub>4</sub>, utilisant des spectres provenant de multiples appareils, nous avons appliqué un prétraitement en utilisant une dérivée première (Vanlierde et al., 2016). Actuellement et pour toutes nos équations, nous utilisons systématiquement une dérivée première qui va corriger la ligne de base de l'information spectrale. Les dérivées secondes n'ont jamais montré d'intérêt dans les modèles développés.

Concernant la modélisation, les algorithmes utilisés pour construire les modèles sont liés à des apprentissages supervisés impliquant que la cible soit connue (ici, la donnée de référence). Comme mentionné au Chapitre 1, les premières modélisations utilisaient la MLR. Ensuite, avec l'essor des nouveaux spectromètres, la modélisation PLS a vu le jour afin de permettre de considérer l'ensemble du signal spectral au sein du modèle. Ces modélisations concernaient toutes des modèles directs. C'est relativement récent que d'autres familles de modèles ont vu le jour. Par contre, comme montré dans ce travail de doctorat, la PLS reste intéressante et ceci d'autant plus pour les modèles indirects ou pseudo-directs dont l'information extraite par le modèle concerne des caractères liés au phénotype visé. Cependant, dans le domaine de l'apprentissage machine (ou en anglais Machine Learning), il n'y a pas de recette fixée. Le mieux est de tester les approches. Ainsi, Luinge et al. (1993) avaient montré qu'il n'y avait pas de différences significatives entre les méthodes MLR et PLS pour les prédictions des composés majeurs du lait. Mais, leurs essais étaient uniquement réalisés à partir d'échantillons de lait de tank collectés pendant le même hiver et dans la même région. Trop peu de variabilité était donc présente que pour pouvoir comparer correctement les deux méthodes. C'est d'ailleurs en élargissant le set de données à une variabilité bien plus large (couvrant plusieurs pays et une période de 10 mois), que Lefier et al. (1996) ont pu démontrer que les modèles PLS étaient plus robustes que les modèles MLR et ce d'autant plus lorsque les laits analysés présentaient des compositions non courantes. Ceci doit expliquer pourquoi depuis presque 20 ans maintenant, les algorithmes que nous utilisons pour tous les modèles développés sont de type PLS. Cependant, récemment, Soyeurt et al. (2020) a montré qu'un réseau neuronal était légèrement plus intéressant que la PLS pour quantifier la lactoferrine dans le lait. Grelet et al. (2020) ont également montré qu'un modèle de type Support Vector Machine était plus intéressant que la PLS pour doser les efficacités individuelles d'utilisation de l'azote. Ces améliorations pourraient vraisemblablement s'expliquer par des jeux de données plus importants qui permettent d'utiliser des modélisations plus complexes permettant notamment de modéliser des relations non-linéaires. Même si les modèles utilisés étaient plutôt de type régression, on a remarqué dernièrement que pour certains paramètres prédits par le FT-MIR dans le lait, notamment liés à certains métabolites du sang (Grelet et al., 2022b), la distribution non normale des données de référence, préférait l'utilisation d'un modèle discriminant (PLS-DA dans ce cas) de type classification. D'autres modèles, dont notamment des modèles bayésiens, ont été testés par d'autres équipes afin d'améliorer la qualité et la

robustesse des modèles (Bonfatti et al., 2017, El Jabri et al., 2019). Meza Ramirez et al. (2020) décrivent ces différents modèles dans une revue bibliographique.

Pourtant même si l'intérêt de la PLS s'est avéré performant comme mentionné dans le Chapitre 1, certains pays comme la France font le choix de continuer à utiliser les modèles MLR en FT-MIR. La raison invoquée pour conserver ces équations MLR (MG, MAT, lactose) est qu'ils les ont développés en interne et que cela leur permet de conserver une plus grande indépendance face aux constructeurs. Ils ont également des craintes quant à la perte du contrôle et de la visibilité face aux modifications des modèles développés à l'étranger. Pourtant, les modèles proposés par l'ensemble des constructeurs pour les éléments majeurs que sont la MG, MAT et lactose sont suffisamment robustes et validés par des organismes internationaux tels que ICAR. En effet, les certifications des nouveaux appareils intègrent tant la machine que les équations fournies par chaque constructeur selon le protocole établi et une validation externe impliquant de nombreux pays est prévue dans celui-ci (ICAR, 2019). D'autre part, même si un constructeur venait à modifier un modèle après la certification, les dérives seraient alors observées lors des validations bisannuelles de ces modèles et des essais inter laboratoires internationaux proposés par ICAR. Si aucune dérive n'est observée alors même que le modèle a subi des modifications, ce n'est pas un problème en soi, selon moi, vu que la prédiction reste fiable et précise. Néanmoins, ce problème d'accessibilité et d'indépendance face aux constructeurs est important. En effet, idéalement, les utilisateurs devraient facilement accéder en totale transparence aux modèles utilisés par les spectromètres, pouvoir tester leurs performances et si besoin, avoir accès aux spectres brutes afin d'améliorer les modèles avec leurs propres données de référence. Le développement de modèles grâce à l'intervention d'équipes universitaires ou de centres de recherches publiques est une garantie de cette indépendance.

Concernant, les critères de performance, le Chapitre 4 est une bonne base de travail pour permettre de proposer un certain nombre de solutions et de critères nécessaires à la comparaison et au jugement de la qualité des modèles. Cependant, vu l'internationalisation dans l'utilisation des modèles et la création de modèles pour des mêmes phénotypes par différentes équipes, il est temps de mettre aussi en place des recommandations afin d'harmoniser la communication des performances d'un modèle et peut-être aussi leur validation. Cette discussion concertée devrait permettre de proposer à l'échelle internationale un guide ayant pour objectif de permettre la comparaison des nouveaux modèles FT-MIR en toute indépendance et transparence.

## 5.5. Normalisation

Les résultats obtenus dans le Chapitre 4 et la discussion déjà menée jusqu'à présent plaident pour le développement de recommandations au niveau international, mais également à la création et à l'intensification des collaborations internationales, afin de permettre la constitution d'ensembles de données les plus larges possibles afin

d'augmenter la robustesse des modèles et permettre leur utilisation dans des conditions de routine.

Actuellement, les organisations internationale (ISO, FIL et/ou ICAR) proposent un ensemble de normes et protocoles d'évaluation destiné aux méthodes alternatives dont la spectrométrie FT-MIR. Ces normes définissent un certain nombre de termes statistiques généraux (ISO 3534-1 :2006), relatifs à l'exactitude (justesse et fidélité) (ISO 5725-1) ou, devant être utilisés lors de l'évaluation de nouveaux appareils infrarouge (ISO 8196-1:2009). Un protocole précis permet la validation des nouveaux appareils FT-MIR (ISO 8196-3:2022), mais malheureusement, ce protocole n'est applicable actuellement qu'aux validations des appareils FT-MIR associés aux seuls modèles de prédiction pour les composés majeurs du lait (MG, MAT, Lactose et urée). Des recommandations devraient être faites pour permettre de correctement évaluer l'ensemble des nouveaux modèles. Cette nouvelle norme ou guide relatif à l'évaluation des nouveaux modèles, devrait idéalement, sur base du Chapitre 4, inclure des informations concernant la région spectrale utilisée, la complexité et la méthode de validation du modèle. De plus, les critères d'évaluation liés au coefficient de corrélation ( $R^2$ ), à l'erreur type d'étalonnage (SEC), à l'erreur type de prédiction (SEP), à l'étendue du ratio de l'erreur (RER), au ratio de l'erreur type de prédiction par rapport à l'écart-type (RPD) et à l'erreur quadratique (RMSE) devraient également être rapportés. Actuellement, une nouvelle initiative conjointe des trois organisations internationales ISO, FIL et ICAR, appelée ExtraMIR (<https://fil-idf.org/our-work/methods-of-analysis-and-sampling/extramir/>), va tenter de résoudre une partie du problème et s'intéresser spécifiquement aux AG. Le choix des AG est judicieux, car ce sont certainement les nouvelles équations issues de la FT-MIR qui ont le plus haut potentiel d'utilisation et qui suscitent le plus d'intérêt. ExtraMIR travaillera sur 4 grands axes. Le premier axe se focalisera sur les analyses de référence, afin de trouver des solutions permettant l'échange de données de référence AG et de s'assurer que les résultats entre laboratoires sont reproductibles. Ce sera donc une suite des travaux conduits précédemment et mentionné au point 5.3. Le deuxième axe s'intéressera aux modèles de prédiction. Un inventaire des modèles existants pour prédire les AG sera réalisé avec l'idée de faciliter l'échange des données de référence et de spectres entre équipes et créer in fine des modèles communs. Ces données seront également utiles pour valider les performances de prédiction des différents modèles existants sur le marché. Cela permettra d'apprécier les biais entre modèles et par rapport à la référence. Le troisième axe visera à trouver une solution pour évaluer la qualité et la robustesse des équations à l'échelle mondiale. Pour y arriver, une base de données spectrales représentative de la variabilité spectrale existante dans le monde entier sera créée. Celle-ci pourra alors être utilisée pour tester la robustesse des modèles existants. Deux sous tâches supplémentaires s'intéresseront à la comparaison des solutions de standardisation spectrale et à développer un outil permettant de détecter les spectres aberrants ou atypiques (outliers) sur base des valeurs des distances GH ou NH. Finalement, le quatrième axe concernera la communication et la rédaction de différents guides à l'attention des développeurs de modèles, des

laboratoires, des producteurs et de l'industrie laitière. Avec les autres personnes impliquées dans ce projet, nous espérons que les réponses apportées par ExtraMIR pourront répondre aux défis actuellement rencontrés. Les solutions apportées dans le cadre des axes 2 et 3 pourraient certainement être une excellente base pour pouvoir permettre, par la suite, de poursuivre le travail vers des modèles prédisant d'autres phénotypes. Car, il faudra poursuivre les efforts pour répondre aux problèmes soulevés par les autres familles de modèles (pseudo-directs, indirects et particuliers). Actuellement, aucun consensus n'existe à l'échelle mondiale pour ces modèles, alors que l'intérêt à leur égard est réel de la part du secteur laitier, et que les échanges de données relatifs à ces nouveaux paramètres sont utiles voire nécessaires à bien des égards.

## 5.6. Standardisation

Afin de bénéficier à toutes les familles de modèles, un des défis les plus importants à relever, selon moi, concerne la standardisation spectrale à l'échelle mondiale afin de s'assurer que la réponse spectrale est parfaitement égale d'un appareil FT-MIR à l'autre à tout moment. L'utilisation de spectres non standardisés, provenant de différentes machines, parfois de marques différentes, et à des périodes différentes, crée indubitablement des interférences et donc des biais tant lors de la confection des modèles que lors de l'obtention des prédictions. D'autre part, même si on arrive à s'accorder sur un modèle commun, par exemple pour prédire un ou une famille d'AG, le transfert de celui-ci sur les machines risque de présenter un cas de figure identique à celui donné au Chapitre 4 (voir Figure 4.6.), qui concerne l'équation relative à la prédiction des PUFA. Au travers de cette figure, on observe que si aucune standardisation n'est réalisée sur les spectres, on accroît fortement la variabilité des prédictions effectuées à partir des spectres non standardisés. Le problème général, lié au transfert des modèles est discuté plus en détail dans Grelet et al., 2017. Or, la nature des modèles indirects, ne fera que compliquer ce problème. Lorsque l'on passera à l'échelle mondiale, les différences entre appareils risquent d'être encore bien plus marquées, générant encore plus de variabilité entre les réponses des appareils, engendrant encore plus de bruit, ce qui impactera la précision finale des prédictions.

Autrefois, ce problème était relativement simplement résolu pour les composés majeurs en utilisant une approche d'étalonnage adaptée permettant des corrections quotidiennes liées au biais et à la pente grâce à un échantillon pilote aux données de référence connues (Lynch et al., 2006). Le fait que les composés majeurs, soient prédits uniquement à partir de 4 ou 5 nombres d'ondes, et que la réalisation de gammes d'étalonnage et de pilote soit relativement aisée, facilite la tâche. Concernant les nouveaux phénotypes à prédire, il n'en est plus de même, les étalonnages et corrections « biais et pente » sont impossibles pour certains. Pour tenter de répondre à ce problème différentes solutions ont été développées.

Tout d'abord, des solutions ont été apportées par les constructeurs, comme par exemple, l'utilisation d'une solution appelée Equalizer chez Foss (Hillerød,

Danemark). Cette solution liquide chimique ajuste le spectre par rapport à une réponse spectrale connue, car la solution est sensée être stable et reproductible au cours du temps (Winning et al., 2014). Bentley Instruments (Chaska, E-U) a également développé une solution de standardisation propre à chaque appareil via une solution mécanique. Ils utilisent un film de polystyrène pour ajuster les décalages de fréquence laser de l'interféromètre dans le temps (Gupta et al., 1994) et les informations liées à la cellule de flux infrarouge pour ajuster les mesures d'absorbance (Parsons and Lyder, 2017). Ces approches sont intéressantes, car elles laissent entrevoir des possibilités d'ajustements automatiques intra-instruments au cours du temps. Par contre, il n'existe pas d'études indépendantes permettant de valider leur efficacité et d'évaluer si elles offrent réellement des possibilités de standardisation entre instruments de modèle identique, de la même marque ou encore de marques différentes, pour permettre l'échange de spectres entre spectromètres FT-MIR afin d'aider à la création de modèles communs et aider à leur transfert.

Une deuxième solution, comme nous l'avons vu au Chapitre 2, est de procéder à un échange de laits communs au moment des prises de mesure servant à la création des modèles afin de créer un fichier de répétabilité (ou en anglais repeatability file). Cette solution n'est cependant que temporaire (lors de la création du modèle) et ne permet pas de corriger les déviations éventuelles au cours du temps, liées par exemple à l'usure, au remplacement de pièces, etc...

Une troisième solution a été proposée par Bonfatti et al. (2017). Elle vise à harmoniser des spectres collectés à partir d'instruments non standardisés dans le cas où l'on cherche à créer des bases de données historiques pour permettre l'utilisation rétroactive de modèles d'étalonnage ou l'utilisation directe des variables spectrales. Comme le souligne l'auteure, cette solution ne doit être considérée que comme complémentaire aux procédures classiques de standardisation qui utilisent des spectres acquis sur des échantillons de référence communs.

La standardisation spectrale développée par Grelet et al. (2015) présentée au Chapitre 4, repose sur l'analyse d'échantillons de référence identiques sur tous les instruments d'un même réseau. Cette solution est déjà appliquée avec succès en Europe dans le cadre de l'European Milk Recording (EMR) (<https://www.milkrecording.eu/>) et permet de standardiser un réseau d'une centaine d'appareils. En plus de cette activité de standardisation des appareils de routine, notre laboratoire offre la possibilité aux équipes de recherche de participer à des échanges des gammes de lait. Ceci permet d'effectuer la standardisation de leurs appareils selon le standard EMR, ce qui permet d'envisager des échanges de spectres standardisés associés à de nouvelles données de référence et/ou d'envisager le transfert de leurs nouveaux modèles sur les appareils du réseau EMR. Néanmoins, le succès de cette méthode, et son étendue à d'autres pays à l'extérieur du continent européen, nécessitent de résoudre plusieurs problèmes clefs, liés à la logistique, la préservation des échantillons, et aux risques liés à la biosécurité (Tiplady et al., 2020).



## 5.7. Contrôle qualité

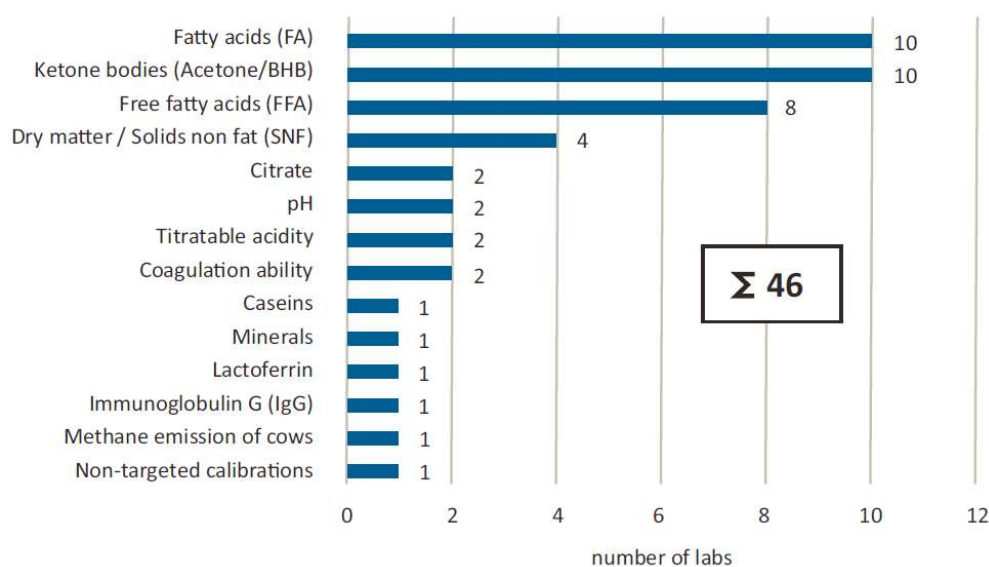
Si l'on se réfère à la norme ISO 8196-2, le contrôle qualité est défini comme étant l'ensemble des vérifications effectuées par le laboratoire laitier afin d'assurer la parfaite qualité des prédictions analytiques de ses appareils en routine. Néanmoins, il est important d'ajouter une dimension et de sortir de la simple logique de laboratoire. En effet, l'étape de prédiction peut se faire directement au sein du laboratoire lors de l'analyse de l'échantillon (prédiction interne), mais également, indépendamment du moment de l'analyse du lait, si par exemple, le spectre est conservé dans une base de données (prédiction externe) et que l'on applique à posteriori un nouveau modèle. Dans ces deux cas, on va devoir s'assurer de la notion de transférabilité du modèle (section 5.7.1) et du maintien de la qualité des prédictions (section 5.7.2).

### 5.7.1. Contrôle qualité du modèle

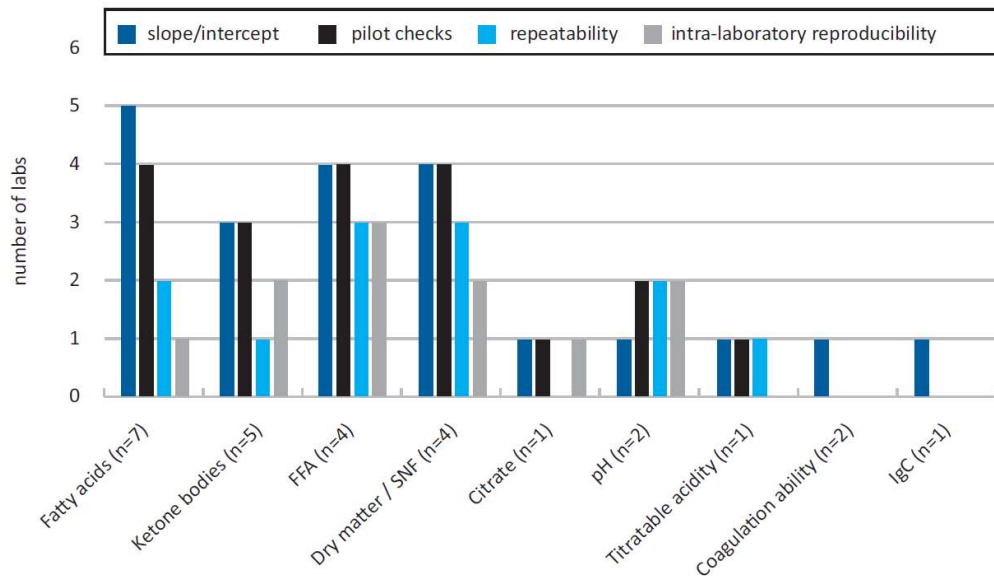
Lors du transfert d'un modèle développé dans une région 1 vers une autre région 2 n'ayant pas participé à la création de ce modèle, on sera particulièrement prudent car cela pourrait donner des prédictions erronées. En effet, si la variabilité des données spectrales de la région 1 utilisée par ce modèle ne couvre pas la variabilité spectrale observée des laits dans cette nouvelle région 2, on risque de faire des extrapolations qui peuvent biaiser la prédiction finale (Dardenne, 2010). Dès lors, un contrôle systématique devra être fait pour vérifier l'absence d'extrapolation. En effet, un modèle n'est pas figé mais dynamique. En intégrant des échantillons qui permettent d'accroître la variabilité spectrale, on augmente la robustesse du modèle. On peut élargir cette variabilité tant que ce n'est pas préjudiciable à ses performances. Pour ce faire, on pourrait proposer une solution basée sur la collecte d'un set d'échantillons représentatifs de cette nouvelle région 2. Ensuite, les distances GH de ces échantillons par rapport à ceux utilisés lors du calibrage du modèle seraient estimées (voir section 4.5.2). Si l'on obtient un taux plus élevé d'échantillons avec une valeur GH supérieure à 3 que celui observé dans la région 1, c'est que le modèle couvre mal la variabilité de cette nouvelle région 2. Pour résoudre le problème, avant toute utilisation en routine, il serait recommandé d'effectuer de nouvelles analyses de référence, jusqu'à l'obtention de proportion équivalente. Idéalement, on effectuerait préférentiellement ces analyses sur des échantillons présentant des valeurs de GH élevées afin de compléter la variabilité manquante du set de calibrage. Ensuite, sur base du nouveau set, le modèle serait affiné afin de permettre une mise à jour de ses coefficients pour intégrer cette nouvelle variabilité. Cela permettrait alors d'accroître la robustesse du modèle. A l'inverse si le taux d'échantillons avec un GH supérieur à 3 est plus faible, le modèle pourrait être utilisé en routine par les laboratoires laitiers et les organismes d'élevage. D'autres méthodes que celle reposant sur l'estimation de la valeur GH pour estimer la distance de Mahalanobis ou mesurer les distances associées pourraient être utilisées pour détecter les extrapolations (Garrido-Varo et al., 2019).

### 5.7.2. Contrôle qualité de la prédiction

Les paramètres prédits par FT-MIR doivent être surveillés conformément à un plan de contrôle qualité. Les éléments traditionnellement utilisés pour garantir la fiabilité des résultats sont les critères liés au contrôle d'étalonnage, à la répétabilité, à la reproductibilité et à la surveillance de la stabilité. Pour respecter ces critères dans le domaine de la spectrométrie FT-MIR pour les composés majeurs, des normes ont été publiées et comprennent notamment les procédures pour l'étalonnage en routine (ISO 9622:2015) et le contrôle de la qualité dans les laboratoires (ISO 8196-2:2009). Dans le cadre de ma participation aux travaux de la FIL, nous avons organisé une enquête permettant d'interroger 106 laboratoires dans le monde. Ainsi, 27 laboratoires (25%), répartis sur 4 continents différents (Europe, Amérique du nord, Asie et Océanie), avaient déclaré utiliser une ou plusieurs nouvelles calibrations FT-MIR (voir Figure 5.5) montrant l'intérêt des nouveaux modèles développés. Lors de cette enquête, les laboratoires ont été interrogés sur le système mis en place pour assurer le contrôle qualité de ces nouveaux phénotypes prédits. Sur 46 équations prédisant un nouveau phénotype, il existait pour 26 de celles-ci une solution déclarée d'étalonnage. D'autre part, pour 31 équations, un programme de contrôle qualité était appliqué régulièrement, sous formes variées, consistant généralement à vérifier la pente/ordonnée à l'origine et à utiliser des échantillons pilotes réguliers. Dans certains laboratoires, la répétabilité et la reproductibilité intra-laboratoire étaient également évaluées (voir Figure 5.6.).



**Figure 5.5.** Nouveaux paramètres prédits par les nouvelles calibrations FT-MIR. Source : Pécou et al., 2020).



**Figure 5.6.** Programme de contrôle qualité appliqué pour les nouveaux paramètres prédits par FT-MIR. Source : Pécou et al., 2020).

L'étape de suivi au quotidien d'un échantillon pilote est simple à mettre en œuvre et devrait être systématiquement proposée afin de construire facilement des cartes de contrôle efficaces. En revanche, l'approche traditionnelle de vérification de l'étalonnage est inadaptée pour l'ensemble des nouveaux modèles. Par exemple, pour le groupe particulier des modèles prédisant les AG, le nombre important d'intercorrélations possibles entre les AG eux-mêmes, rend particulièrement complexe les étalonnages. Ainsi, l'utilisation de la méthode ISO 9622:1999, communément appelée FIL 141, permettant de recombinaison des échantillons sur base de méthodes de séparation/recombinaison (Figure 1.8) afin d'obtenir une gamme d'échantillons dont les paramètres sont totalement décorrélés est dans ce cas techniquement difficile, voire impossible à réaliser. A défaut, les laboratoires ont dû s'adapter afin de proposer des solutions. A titre d'exemple, une approche expérimentale a été évaluée par l'interprofession laitière française (CNIEL) qui visait à utiliser des gammes mensuelles d'étalonnages de leurs appareils (Pécou et al., 2020) à partir d'un set de 5 échantillons variables en concentration, fabriqués à partir d'un seul et unique lait, lui-même issu d'un mélange de laits collectés dans plusieurs régions, ou sous régions, et utilisant des régimes alimentaires différents. Les différentes concentrations étaient obtenues par dilutions simples (variation de la MG totale et par association, variation des différents AG qui la compose) et les analyses de référence étaient réalisées uniquement sur l'échantillon de mélange de départ (les concentrations étant recalculées à partir des taux de MG). Sept modèles différents ont été étalonnés par cette méthode : le total des acides gras saturés, insaturés,

monoinsaturés, polyinsaturés, l'acide palmitique (C16:0), l'acide stéarique (C18:0) et l'ensemble des isomères du C18:1. Les résultats présentés à la FIL, montrent que les corrections appliquées semblent permettre un certain contrôle des déviations au cours du temps (Pécou et al., 2020). Cette solution est simple et économique, mais, le fait de n'utiliser uniquement que de simples dilutions, ne permet que de faire varier la MG totale et il n'y a par conséquent aucune décorrélation entre les différents AG, ce qui n'est pas idéal (Lefier et al., 1996). Dès lors, cette approche n'apporte pas de variations supplémentaires à la simple variation linéaire des valeurs de concentration de la MG totale. Les résultats présentés par Pécou et al. (2020) ne seraient donc que la conséquence des corrélations existantes entre AG et MG totale (Soyeur et al., 2006) et ne permettraient pas de tirer profit de l'information supplémentaire apportée par les équations AG. Pour étayer ces propos, une validation devrait être réalisée sur des sets d'échantillons individuels présentant des taux d'AG variables et si possible décorrélés. Une nouvelle procédure d'étalonnage des AG vient d'être proposée par la société Lactanet au Canada (communication personnelle, J-P. Angers, octobre 2022). Elle repose sur un set de 8 échantillons indépendants les uns des autres, choisis sur base de leur potentiel à présenter de la variabilité au niveau des concentrations en différents groupes d'AG d'intérêt : les AG provenant de la lipogenèse de novo ayant une chaîne carbonée variant de C4 à C14, les AG préformés, provenant de l'alimentation ou des réserves corporelles de la vache et ayant des chaînes carbonées égales ou supérieures à C18 et enfin, les AG dits « mélangés », c'est-à-dire provenant soit de la lipogenèse ou soit préformés, il s'agit essentiellement des AG de type C16 (Castro et al., 2022). Nous n'avons pas encore eu l'occasion de tester cette méthode, néanmoins, la plus grande variabilité spectrale et le fait que les AG ne sont pas corrélés d'un échantillon à l'autre devrait certainement être plus adaptée à l'étalonnage des modèles AG, mais une validation indépendante permettant de comparer ces deux méthodes reste à organiser.

La réalisation de gammes d'étalonnage pour les modèles AG n'est certainement pas le cas de figure le plus compliqué. Comment devrions-nous procéder pour les modèles pseudo-directs, indirects ou particuliers ? Devrions-nous créer des gammes adaptées pour chaque modèle ? Cela représenterait un coût non négligeable et serait parfois impossible comme avec le CH<sub>4</sub>. La solution apportée par la standardisation spectrale EMR (section 5.6.) pourrait permettre de s'affranchir de ce problème. Elle vise à permettre de fournir des réponses spectrales identiques pour un même échantillon mesuré par tous les appareils ayant été standardisés, et ce indépendamment de la date d'analyse. Pour réaliser cette étape, des facteurs de correction sont calculés à partir d'une gamme de standardisation pour chacun des nombres d'ondes d'intérêt (Grelet et al., 2015). Si de nouveaux modèles sont construits à partir de spectres standardisés, la standardisation spectrale mensuelle permet de s'affranchir d'une étape supplémentaire d'étalonnage pour l'ensemble de ces modèles. Grelet et al. (2017) a pu démontrer que ce processus permettait le transfert d'équations difficiles à étalonner

(prédiction des émissions de CH<sub>4</sub>, des concentrations en acides gras polyinsaturés ou du rendement en fromage frais).

Certains défis importants, liés à cette méthode, doivent cependant encore être résolus, on doit par exemple s'assurer que le ou les appareils maîtres, selon qu'on en utilise un ou plusieurs, ne dévient pas dans le temps. Un processus que nous appelons ancrage du ou des appareils maîtres reste encore à améliorer. La difficulté majeure est de trouver une solution permettant de corriger les appareils maîtres de toute déviation dans le temps. Des solutions sont actuellement à l'étude. Elles reposent sur l'utilisation de gammes identiques de lait de standardisation conservé sur de longues périodes, à l'aide de la congélation à très basse température. Ces congélations à très basse température (-196°C) permettent d'éviter une cristallisation trop lente, qui peut détériorer la structure physique du lait, en produisant de trop gros cristaux (Zhiwei et al., 2019) qui peuvent alors influencer négativement la qualité du spectre. La possibilité d'utiliser des solutions chimiques présentant des spectres proches du lait est également une alternative à l'étude. Celle-ci aurait également pour avantage supplémentaire par rapport aux laits congelés à très basse température de nous aider à résoudre les problèmes liés à la biosécurité qui empêchent parfois les échantillons de lait de passer les frontières lors de transferts trans-continentaux.

En complément à la solution de standardisation, on étudie une procédure de suivi des appareils, via l'utilisation d'un échantillon pilote au quotidien ou encore appelé daily check (Grelet et al., 2018), pour nous aider à surveiller et contrôler la stabilité entre les étalonnages. Pour l'instant dans sa version « routine » tel qu'établi au sein de mon laboratoire, le daily check permet de détecter les perturbations spectrales en appliquant une prédiction MG sur les spectres bruts (Figure 4.7). Nos recherches actuelles nous montrent que l'utilisation d'un seul lait n'est pas suffisante. Afin de résoudre ce problème et affiner la méthode, nous essayons actuellement de détecter les éventuelles perturbations en utilisant deux laits ayant des spectres différents et en utilisant non plus un seul paramètre prédictif (MG) mais plusieurs paramètres prédits pour mieux couvrir l'ensemble des plages d'absorbances du spectre (résultats non encore publiés). Une approche purement spectrale est également étudiée comme moyen de détection des perturbations, sur base du calcul des différences d'absorbances enregistrées au cours du temps pour chacun des nombres d'ondes d'intérêt. En outre, si une perturbation est détectée, on teste une correction automatique des perturbations à partir de 3 laits différents. L'ensemble de ces recherches sont en cours au sein de notre laboratoire et restent encore à tester dans des laboratoires de routine. On peut donc espérer à moyen terme des solutions permettant d'assurer la qualité de la prédiction.

Lors des analyses de routine, comme déjà évoqué précédemment, on devra s'assurer pour ces nouveaux modèles, que la variabilité de chaque spectre à prédire est couverte par la variabilité spectrale du modèle en calculant systématiquement leurs valeurs de GH avec les échantillons ayant servis à élaborer le modèle. Idéalement, une méthode de routine devrait systématiquement calculer ces valeurs de GH et les

prédictions avec un GH dépassant un seuil supérieur à 3 (Vanlierde et al., 2021) devraient ne pas être fournies au client et être signalées comme valeurs suspectes ou anormales. Si la proportion d'échantillons dépassant ce seuil GH devenait important, cela serait alors le signe que le modèle ne serait plus adéquat et qu'il devrait alors être nécessaire d'analyser de nouveaux échantillons en référence pour mettre à jour les coefficients du modèle. En effet, ce n'est pas parce qu'on a vérifié qu'un modèle était adéquat à un moment donné, qu'il le sera toujours quelques années plus tard. Des changements dans l'alimentation des vaches laitières, dans leur génétique ou encore dans le système de gestion des troupeaux pourraient engendrer de nouvelles compositions laitières dont la variabilité ne serait pas comprise dans le set de calibrage précédemment créé pour construire le modèle.

Comme on l'a dit précédemment, les composés majeurs du lait sont corrigés par le laboratoire par la correction « biais et pente », cette correction n'a donc aucun impact sur le spectre. Pourtant, ces spectres sont conservés dans des bases de données pour des prédictions ultérieures. Ainsi, il peut être intéressant, comme l'a montré Zhang et al. (2021), de combiner l'estimation de la distance GH avec l'estimation d'une erreur quadratique entre la MG prédite sur le spectre stocké dans la base de données et la MG communiquée par le laboratoire laitier qui a été corrigée. Si cette erreur est grande, cela veut dire que la correction appliquée a été forte et donc que probablement le spectre est de faible qualité. Un spectre de faible qualité engendrera une prédiction de faible qualité. Cette manière de procéder est intéressante pour vérifier l'état du spectre avant l'application de modèles indirects, par exemple.

## 5.8. Références

- Arbre M., Rochette Y., Guyader J., Lascoux C., Gómez L. M., Eugène M., Morgavi D. P., Renand G., Doreau M., and Martin C., 2016. Repeatability of enteric methane determinations from cattle using either the SF6 tracer technique or the GreenFeed system. *Anim. Prod. Sci.* 56:238–243.
- Arnould V.M., Reding R., Bormann J., Gengler N. and Soyeurt H., 2015. Predictions of Daily Milk and Fat Yields, Major Groups of Fatty Acids, and C18:1 cis-9 from Single Milking Data without a Milking Interval. *Animals (Basel)*, 5(3):643-61.
- Ayadi, M., Caja, G., Such, X., Rovai, M., and Albanell, E., 2004. Effect of different milking intervals on the composition of cisternal and alveolar milk in dairy cows. *Journal of Dairy Research*, 71(3), 304-310.
- Bahadi M, Ismail AA, and Vasseur E., 2021. Fourier Transform Infrared Spectroscopy as a Tool to Study Milk Composition Changes in Dairy Cows Attributed to Housing Modifications to Improve Animal Welfare. *Foods* 10 (2):450.
- Bastin, C, Gengler, N, and Soyeurt, H., 2011. Phenotypic and genetic variability of production traits and milk fatty acid contents across days in milk for Walloon Holstein first-parity cows. *J. Dairy Sci.*, 94(8), 4152-63.
- Baumont R., Dewhurst R.J., Mesragan S.D., Hurtaud C., Orchard A., Kuhla B., Martin C., Munksgaard L., Reynolds C., O'Donovan M. and Esmeim B., 2019. SmartCow: a project aimed at improving phenotyping capacity across cattle research infrastructures in Europe. ICAR Conference 2019, NA, France.
- Bonfatti V., Tiezzi F., Miglior F., Carnier P., 2017. Comparison of Bayesian regression models and partial least squares regression for the development of infrared prediction equations. *J Dairy Sci.*, 100(9):7306-7319.
- Borchersen S., Worsoe N. and Borggaard H.C., 2020. System for determining feed consumption of at least one animal. Patent EP2983465B1.
- Brand W., Wells A.T., Smith S.L., Denholm S.J., Wall E. and Coffey M.P., 2021. Predicting pregnancy status from mid-infrared spectroscopy in dairy cow milk using deep learning. *J. Dairy Sci.*, Volume 104, Issue 4, pp. 4980-4990.

- Casarrubias-Torres L.M., Meza-Márquez O.G., Osorio-Revilla G. and Gallardo-Velazquez T., 2018. Mid-infrared spectroscopy and multivariate analysis for determination of tetracycline residues in cow's milk. *Acta Vet Brno.*, 87:181–8.
- Cassoli L.D., Sartori B., Zampar A. and Machado P.F., 2011. An assessment of Fourier Transform Infrared spectroscopy to identify adulterated raw milk in Brazil. *International Journal of Dairy Technology*, 64: 480-485.
- Castillo C. and Hernández J., 2021. Ruminant Fistulation and Cannulation: A Necessary Procedure for the Advancement of Biotechnological Research in Ruminants. *Animals (Basel)*, 11(7):1870.
- Castro M.M.D., Matson R.D., Santschi D.E., Marcondes M.I. and DeVries T.J., 2022. Association of housing and management practices with milk yield, milk composition, and fatty acid profile, predicted using Fourier transform mid-infrared spectroscopy, in farms with automated milking systems. *J Dairy Sci.*, 105 (6) :5097-5108.
- Coppa M., Vanlierde A., Bouchon M., Jurquet J., Musati M., Dehareng F., Martin C., 2022. Methodological guidelines: Cow milk mid-infrared spectra to predict reference enteric methane data collected by an automated head-chamber system. *J Dairy Sci.*, 105(11):9271-9285.
- Dardenne P., 2010. Some considerations about NIR spectroscopy: Closing speech at NIR-2009. *NIR news.*, 21(1):8-14.
- Dardenne P., Vincke D., Baeten V. and Fernandez-Pierna J.A., 2015. Untargeted multivariate characterization of contaminants in powdered milk. Final OptiMIR Scientific and Expert Meeting: From milk analysis to advisory tools. *Biotechnol. Agron. Soc. Environ.*, 19(2), 97-124.
- Destailats F., Golay P.A., Joffre F., de Wispelaere M., Hug B., Giuffrida F., Fauconnot L. and Dionisi F., 2007. Comparison of available analytical methods to measure trans-octadecenoic acid isomeric profile and content by gas-liquid chromatography in milk fat. *Journal of Chromatography A*, 1145 (1-2), pp. 222-228.
- Deswysen A.G., Dutilleul P.A. and Ellis W.C., 1989. Quantitative analysis of nycterohemeral eating and ruminating patterns in heifers with different voluntary intakes and effects of monensin. *J Anim Sci.*, 67(10):2751-61.



- Deswysen A.G., Dutilleul P., Godfrin J.P. and Ellis W.C., 1993. Nycterohemeral eating and ruminating patterns in heifers fed grass or corn silage: analysis by finite Fourier transform. *J Anim Sci.*, 71(10):2739-47.
- Duplessis, M., Santschi, D., Plante, S., Bergeron, C., Lefebvre, D., Durocher, J., and Cue, R., 2019. Milk  $\beta$ -hydroxybutyrate concentration measured by Fourier-transform infrared and flow-injection analyses from samples taken at different times relative to milking. *Journal of Dairy Research*, 86(2), 208-210.
- Egger-Danner C., Suntiger M., Grandl F., Majcen P., Mayerhofer M., Papst F., Saukh O., Fallast M., Turkaspa A., Steiniger F., Linke K., Duda J., Wittek T., Fuerst-Waltl B., Auer F. J., and Kofler J., 2019. Opportunities and Challenges of New Technologies for Performance Recording with Focus on Claw Health and Metabolism. Paper presented at ICAR Conference, Prague, Czech Republic.
- Engel, J., Gerretzen J., Szymanska E., Jansen J.J., Downey G., Blanchet L., and Buydens L.M., 2013. Breaking with trends in pre-processing ? *Trends Anal. Chem.* 50:96–106.
- El Jabri M., Sanchez M-P., Trossat P., Laithier C., Wolf V., Grosperin P., Beuquier E., Rolet-Répécaud O., Gavoye S., Gaüzère Y., Belysheva O., Notz E., Boichard D. and Delacroix-Buchet A., 2019. Comparison of Bayesian and partial least squares regression methods for mid-infrared prediction of cheese-making properties in Montbéliarde cows. *J Dairy Sci.*, 102(8):6943–6958.
- Fernández Pierna J.A., Vincke D., Baeten V., Grelet C., Dehareng F. and 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, *Chemometrics and Intelligent Laboratory Systems*, Volume 152, Pages 157-162.
- Forsbäck L., Lindmark-Månsson H., Andrén A., Åkerstedt M., Andrée L. and Svennersten-Sjaunja K., 2010. Day-to-day variation in milk yield and milk composition at the udder-quarter level, *J Dairy Sci.*, 93 (8), 3569-3577.
- Friedrichs P., Bastin C., Dehareng F., Wickham B. and Massart X., 2015. OptiMIR— A project aiming the development of novel mid-infrared based management tools for dairy herds. In *Final OptiMIR Scientific and Expert Meeting: From milk analysis to advisory tools* (Palais des Congrès, Namur, Belgium, April 16–17, 2015, *Biotechnol. Agron. Soc. Environ.*, 19, pp. 97-124.

- Garrido-Varo A., Garcia-Olmo J. and Fearn T., 2019. A note on Mahalanobis and related distance measures in WinISI and The Unscrambler. *Journal of Near Infrared Spectroscopy*. 27(4):253-258.
- Garnsworthy P.C., Difford G.F., Bell M.J., Bayat A.R., Huhtanen P., Kuhla B., Lassen J., Peiren N., Pszczola M., Sorg D., Visker M.H.P.W. and Yan T., 2019. Comparison of Methods to Measure Methane for Use in Genetic Evaluation of Dairy Cattle. *Animals (Basel)*, 9(10):837.
- Gengler N., Grelet C., Soyeurt H. and Dehareng F., 2019. GplusE: Mid-infrared milk analysis based technologies adding value to gene banks. *Proceedings in: Book of Abstracts of the 70th Annual Meeting of the European Federation of Animal Science, Ghent, August 2019*, 443.
- Gengler N., 2022. Collaboration is imperial or how to address this critical issue when developing novel traits. *Proc. EAAP 77th Annual Meeting, Porto, Wageningen Academic Publishers, Wageningen, the Netherlands*, p. 329.
- Godden SM, Lissemore KD, Kelton DF, Lumsden JH, Leslie KE, and Walton JS., 2000. Analytic validation of an infrared milk urea assay and effects of sample acquisition factors on milk urea results. *J Dairy Sci.*, 83(3):435-42.
- Golay P.A., Giuffrida F., Dionisi F., and Destailats F., 2009. Streamlined Methods for the Resolution and Quantification of Fatty Acids including Trans Fatty Acids Isomers in Food Products by Gas-Chromatography. Special edition of *Journal of AOAC International on "Trans Fats: Update on Health Effects, Methodology and Levels in Processed Foods - Journal of AOAC International"*, 92 (5), pp. 1301-1309.
- Grappin, R., 1987. Definition and evaluation of the overall accuracy of indirect methods of milk analysis – Application to calibration procedure and quality control in dairy laboratory (IDF Provisional Standard 128:1985). *Bulletin of the International Dairy Federation 208/1987*, 3 –12. IDF, Brussels.
- Grelet C., Pierna J.F., Dardenne P., Baeten V. and Dehareng F., 2015. Standardization of milk mid-infrared spectra from a European dairy network. *J. Dairy Sci.*, 98 (2015), pp. 2150-2160
- Grelet C, Bastin C, Gelé M, Davière JB, Johan M, Werner A, Reding R, Fernandez Pierna JA, Colinet FG, Dardenne P, Gengler N, Soyeurt H and Dehareng F., 2016. Development of Fourier transform mid-infrared calibrations to predict acetone,  $\beta$ -hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. *J Dairy Sci.*, 99(6):4816-4825.

- Grelet C., Pierna J.F., Dardenne P., Soyeurt H., Vanlierde A., Colinet F., Bastin C., Gengler N., Baeten V. and Dehareng F., 2017. Standardization of milk mid-infrared spectrometers for the transfer and use of multiple models. *J. Dairy Sci.*, 100pp. 7910-7921.
- Grelet, C., Baeten, V., Dardenne, P., Fernandez Pierna, J. A., Abbas, O. and Dehareng, F., 2018. Presentation of a new system to monitor and stabilize mid infrared spectral data. Oral presentation at ICAR annual Conference, february 2018, Auckland, N-Z.
- Grelet C, Froidmont E, Foldager L, Salavati M, Hostens M, Ferris CP, Ingvarsten KL, Crowe MA, Sorensen MT, Fernandez Pierna JA, Vanlierde A, Gengler N; GplusE Consortium and Dehareng F., 2020. Potential of milk mid-infrared spectra to predict nitrogen use efficiency of individual dairy cows in early lactation. *J Dairy Sci.*,103(5):4435-4445.
- Grelet C., Vanden Dries V., Leblois J., Wavreille J., Mirabito L., Soyeurt H., Franceschini S., Gengler N., Brostaux Y., HappyMoo Consortium and Dehareng F., 2022a. Identification of chronic stress biomarkers in dairy cows. *Animal.*, 16(5):100502.
- Grelet C., Larsen T., Fernandez Pierna J.A., Stevens F., Chen Y., Christophe O., Vanlierde A., Schuler U., Auer F.J., Egger-Danner C., Werner A., Sölkner J., Wathes J. D.C., Ingvarsten K.L., Crowe M.A., Ferris C.P., Marchitelli C., Becker F., Hummel, Calmels M., Reding R., Gelé M., Gengler N., Leblois J., GplusE consortium and Dehareng F., 2022b. Prediction of key milk biomarkers in dairy cows through milk MIR spectra. Proceeding in Book of Abstract, p. 12. First Dair'innov Congress– Innovations to benefit cow welfare and dairy farming sustainability, Namur, Belgium, April 2022.
- Guidance, 2019. Protocole commun aux DVP-CRAW Gembloux et ILVO-T&V Melle de la guidance scientifique belge - Version 24.10.2019.
- Gupta D., Wang L., Hanssen L.M., Hsia J.J. and Datla R.U., 1995. Standard reference materials: Polystyrene films for calibrating the wavelength scale of infraredspectrophotometers - SRM 1921. Boulder (CO): U.S. Department of Commerce. Report No.: NIST spec publ. 260-122.
- Hansen, P. W. 1998. Urea determination in milk using Fourier transform infrared spectroscopy and multivariate calibration. *Milchwissenschaft* 53:251–255.
- Hansen P.W. and Holroyd S., 2019. Development and application of Fourier transform infrared spectroscopy for detection of milk adulteration in practice. *Int. J. Dairy Technol.*, 72, 321–333.

- Ho P. and Pryce J., 2020. Predicting the likelihood of conception to first insemination of dairy cows using milk mid-infrared spectroscopy. *J. Dairy Sci.*, 103:11535–11544.
- ICAR, 2019. Procedure 1 of Section 12 of ICAR. Guidelines-Protocol for the Evaluation of Milk Analysers for ICAR Approval. v19.02.docx. ICAR, Utrecht, The Netherlands.
- ICAR, 2020. Section 11 - Guidelines for Testing, Approval and Checking of Milk Recording Devices. ICAR, Utrecht, The Netherlands.
- ISO 16958 :2015 | IDF 231:2015. Lait, produits laitiers, formules infantiles et produits nutritionnels pour adultes - Détermination de la composition en acides gras - Méthode de chromatographie en phase gazeuse sur colonne capillaire. International Organization for Standardization, Geneva and International Dairy Federation, Brussels.
- ISO 3534-1:2006. Statistique - Vocabulaire et symboles - Partie 1: Termes statistiques généraux et termes utilisés en calcul des probabilités. International Organization for Standardization, Geneva.
- ISO 5725-1 :1994. Exactitude (justesse et fidélité) des résultats et méthodes de mesure - Partie 1 : Principes généraux et définitions. International Organization for Standardization, Geneva.
- ISO 707:2008| IDF 50:2008, Milk and milk products - Guidance on sampling. International Organization for Standardization, Geneva and International Dairy Federation, Brussels.
- ISO 8196-1:2009 | IDF 128-1:2009. Milk – Definition and evaluation of the overall accuracy of alternative methods of milk analysis – Part 1: Analytical attributes of alternative methods. International Organization for Standardization, Geneva and International Dairy Federation, Brussels.
- ISO 8196-2:2009 | IDF 128-2:2009. Milk – Definition and evaluation of the overall accuracy of alternative methods of milk analysis – Part 2: Calibration and quality control in the dairy laboratory. International Organization for Standardization, Geneva and International Dairy Federation, Brussels.
- ISO 8196-3:2022 | IDF 128-3:2022. Milk – Definition and evaluation of the overall accuracy of alternative methods of milk analysis – Part 3: Protocol for the evaluation and validation of alternative quantitative methods of milk analysis. International Organization for Standardization, Geneva and International Dairy Federation, Brussels.

- ISO 9622 :2013 | IDF 141:2013. Milk and liquid milk products – Guidelines for the application of mid-infrared spectrometry. International Organization for Standardization, Geneva and International Dairy Federation, Brussels.
- Kandel P.B., Vanrobays M.-L., Vanlierde A., Dehareng F., Froidmont E., Gengler N., Soyeurt H., 2017. Genetic parameters of mid-infrared methane predictions and their relationships with milk production traits in Holstein cattle. *J. Dairy Sci.* 100, Issue 7, pp. 5578-5591.
- Klusmeyer T.H., Cameron M.R., McCoy G.C. and Clark J.H., 1990. Effects of feed processing and frequency of feeding on ruminal fermentation, milk production, and milk composition *J. Dairy Sci.*, 73, pp. 3538-3543.
- Leblois J., Bertozzi C., Calmels M., Dale L., Grelet C., Lecomte C., Mirabito L., Tran M.-N., Vanden Dries V., Werner A., Wolf V., Dehareng F. and HappyMoo Consortium, 2021. HappyMoo project – using milk mid-infrared spectrum to predict dairy cows' welfare status, Book of Abstracts of the 72th Annual Meeting of the European Federation of Animal Science, Davos
- Lee A.J., Wardrop J., 1984. Predicting Daily Milk Yield, Fat Percent, and Protein Percent from Morning or Afternoon Tests<sup>1</sup>, *J. Dairy Sci.*, 67(2):351-360.
- Lefier D., Grappin R., Pochet S., 1996. Determination of fat, protein, and lactose in raw milk by Fourier transform infrared spectroscopy and by analysis with a conventional filter-based milk analyzer. *J AOAC Int.*, 79(3):711-7.
- Lynch J.M., Barbano D.M., Schweisthal M. and Fleming J.R., 2006. Precalibration evaluation procedures for mid-infrared Milk analyzers. *J Dairy Sci.*, 89(7):2761–2774.
- Meza Ramirez C.A., Greenop M., Ashton L. and Rehman R.I., 2021. Applications of machine learning in spectroscopy, *Applied Spectroscopy Reviews*, 56:8-10, 733-763.
- Moore J. C., Spink J. and Lipp M., 2012. Development and application of a database of food ingredient fraud and economically motivated adulteration from 1980 to 2010. *Journal of Food Science* 77 R108– R116.
- Negussie E., de Haas Y., Dehareng F., Dewhurst R.J., Dijkstra J., Gengler N., Morgavi D.P., Soyeurt H., van Gastelen S., Yan T. and Biscarini F., 2017. Invited review: Large-scale indirect measurements for enteric methane emissions in dairy cattle: A review of proxies and their potential for use in management and breeding decisions. *J. Dairy Sci.*, 100(4):2433-2453.

- Nielsen N.I., Larsen T., Bjerring M. and Ingvarsten K.L., 2005. Quarter Health, Milking Interval, and Sampling Time During Milking Affect the Concentration of Milk Constituents. *J. Dairy Sci.*, 88 (4):3186-3200.
- Parsons C. and Lyder H., 2017. Determining a size of cell of a transmission spectroscopy device. Patent US9829378B1.
- Pécou A., Holroyd S., Müller B., Ransborg Pedresen M., Broutin P., Schwarz D., and Dehareng, F., 2020. New applications of MIR spectrometry: Quality Assurance practices with new parameters in raw milk analysis. *Bulletin: International Dairy Federation*, 504, 50.
- Quist M.A., LeBlanc S.J., Hand K.J., Lazenby D., Miglior F., Kelton D.F., 2008. Milking-to-Milking Variability for Milk Yield, Fat and Protein Percentage, and Somatic Cell Count. *J. Dairy Sci.*, 91 (9):3412-3423.
- Rischewski, J., Bielak A., Nürnberg G., Derno M., and Kuhla B., 2017. Rapid Communication: Ranking dairy cows for methane emissions measured using respiration chamber or GreenFeed techniques during early, peak and late lactation. *J. Anim. Sci.* 95:3154–3159.
- Roukh A., Nolak Fote F., Mahmoudi S.A. and Mahmoudi S, 2020. Big Data Processing Architecture for Smart Farming, *Procedia Computer Science*, Volume 177, pp. 78-85.
- 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, pp. 6202-6209.
- Scarso S., McParland S., Visentin G., Berry D.P., McDermott A. and De Marchi M., 2017. Genetic and nongenetic factors associated with milk color in dairy cows, *J. Dairy Sci.*, Volume 100, Issue 9, pp. 7345-7361.
- Schori F. and Münger A., 2022. Assessment of two wireless reticulo-rumen pH sensors for dairy cows. *Agrarforschung Schweiz*, 13, 11-16.
- Sivakesava S. and Irudayaraj J., 2002. Rapid determination of tetracycline in milk by FT-MIR and FT-NIR spectroscopy. *J Dairy Sci.* 2002 Mar;85(3):487-93.
- Soyeurt H., Dardenne P., Dehareng F., Lognay G., Veselko D., Marlier M., Bertozzi C., Mayeres P. and Gengler N. 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. *J. Dairy Sci.*, 89, 3690–3695.

- 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., Froidmont E., Dufrasne I., Hailemariam D., Wang Z., Bertozzi C., Colinet F.G., Dehareng F. and Gengler N., 2019. Contribution of milk mid-infrared spectrum to improve the accuracy of test-day body weight predicted from stage, lactation number, month of test and milk yield. *Livest. Sci.* 227, 82-89.
- Soyeurt H., Grelet C., McParland S., Calmels M., Coffey M., Tedde A., Delhez P., Dehareng F. and Gengler N., 2020. A comparison of 4 different machine learning algorithms to predict lactoferrin content in bovine milk from mid-infrared spectra. *J. Dairy Sci.*, 103(12):11585-11596,
- Spieß L., de Peinder P. and van den Bijgaart H., 2021. Advances in Atypical FT-IR Milk Screening: Combining Untargeted Spectra Screening and Cluster Algorithms. *Foods*, 10, 1111.
- Sutton J.D., Broster W.H., Napper D.J. and Siviter J.W., 1985. Feeding frequency for lactating cows: Effects on digestion, milk production and energy utilization *Br. J. Nutr.*, 53, pp. 117-130.
- Tedde A., Grelet C., Ho P.N., Pryce J.E., Hailemariam D., Wang Z., Plastow G., Gengler N., Brostaux Y., Froidmont E., Dehareng F., Bertozzi C., Crowe M.A., Dufrasne I., GplusE Consortium Group, and Soyeurt H., 2021. Validation of Dairy Cow Bodyweight Prediction Using Traits Easily Recorded by Dairy Herd Improvement Organizations and Its Potential Improvement Using Feature Selection Algorithms. *Animals*, 11(5):1288.
- Tiplady K.M., Sherlock R.G., Littlejohn M.D., Pryce J.E., Davis S.R., Garrick D.J., Spelman R.J. and Harris B.L., 2019. Strategies for noise reduction and standardization of milk mid-infrared spectra from dairy cattle. *J Dairy Sci.*, 102(7):6357-6372.
- Tiplady K.M., Lopdell T.J., Littlejohn M.D. and Garrick D.J., 2020. The evolving role of Fourier-transform mid-infrared spectroscopy in genetic improvement of dairy cattle. *J Anim Sci Biotechnol.*, 11:39.
- Tremblay F.J.B. and Massé D.I., 2016. Detailed Design of Respiratory Chamber that Provides Highly Accurate Measurements of Enteric Methane Emissions. *Applied Engineering in Agriculture.*, 32(5): 651-658.

- van Kneegsel 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, pp. 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, pp. 5740-5747
- 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*, 56, 258-264.
- Vanlierde A., Dehareng F., Gengler N., Froidmont E., McParland S., Kreuzer M., Bell M., Lund P., Martin C., Kuhla B. and Soyeurt H., 2021. Improving robustness and accuracy of predicted daily methane emissions of dairy cows using milk mid-infrared spectra. *J Sci Food Agric.*,101(8):3394-3403.
- Vanlierde A, Dehareng F, Gengler N, Froidmont E, McParland S, Kreuzer M, Bell M, Lund P, Martin C, Kuhla B and Soyeurt H., 2021. Improving robustness and accuracy of predicted daily methane emissions of dairy cows using milk mid-infrared spectra. *J Sci Food Agric.*, 101(8):3394-3403.
- Winning H., Mulawa K.M. and Selberg T. Standardization of FT-IR instruments, 2014. White Paper from Foss A/S. 1(1):7.
- Wojciechowski, K. L, Melilli, C. and Barbano, D. M., 2016. A proficiency test system to improve performance of milk analysis methods and produce reference values for component calibration samples for infrared milk analysis. *J. Dairy Sci.* 99: 6808–6827.
- Zhang L., Li C., Dehareng F., Grelet C., Colinet F., Gengler N., Brostaux Y., Soyeurt H., 2021. Appropriate Data Quality Checks Improve the Reliability of Values Predicted from Milk Mid-Infrared Spectra. *Animals (Basel)*, 11(2):533.
- Zeaiter M., Roger J.M., Bellon Maurel V., and Rutledge D.N., 2004. Robustness of models developed by multivariate calibration. Part I: The assessment of robustness. *Trends in Analytical Chemistry*, 23 (2), pp.157-170.



- Zeaiter M., Roger J.M. and Bellon Maurel V., 2005. Robustness of models developed by multivariate calibration. Part II: The influence of pre-processing methods. *Trends in Analytical Chemistry*, 24 (5), pp.437-445.
- Zhiwei Z., Qianyun Z. and Da-Wen S., 2019. Measuring and controlling ice crystallization in frozen foods: A review of recent developments, *Trends in Food Science & Technology*, Volume 90, Pages 13-25.



# Chapitre 6 :

---

Conclusion et perspectives



Cette thèse est le fruit de l'expérience acquise ces deux dernières décennies lors du développement de nouveaux modèles FT-MIR et la recherche de solutions permettant leur contrôle qualité. De nombreux facteurs vont impacter la robustesse et la qualité des nouveaux modèles développés récemment dans le domaine de la spectrométrie FT-MIR du lait.

Comme nous avons pu le voir en introduction, le secteur laitier utilise depuis de nombreuses années la spectrométrie MIR pour prédire les composés majeurs du lait (MG, MAT et lactose) nécessaires au paiement du lait, aux contrôles qualité des usines ou des performances des vaches laitières. Ainsi, de nombreuses solutions ont été apportées pour améliorer la robustesse des modèles et leur contrôle qualité. En revanche, l'arrivée de la spectrométrie FT-MIR et le développement de nouveaux modèles permettant d'offrir de nouvelles perspectives d'utilisation, font apparaître de nouveaux défis et questions. Cette thèse a permis de présenter des solutions et a mis l'accent sur les développements à réaliser à l'avenir pour assurer la qualité des prédictions FT-MIR du lait et accroître leur robustesse. Une synthèse de ces solutions ou recommandations est présentée dans le Tableau 6.1 pour ceux impactant l'erreur totale de l'échantillonnage (TSE) et dans le Tableau 6.2 pour ceux influençant l'erreur analytique globale (TAE).

Lors du développement de nouveaux modèles de prédiction pour les AG du lait, nous avons rapidement été confrontés à une augmentation vertigineuse de la variabilité à couvrir par les modèles afin de garantir leur robustesse. Une solution originale a été proposée via une sélection des échantillons de référence à analyser sur base de leur variabilité spectrale (X). Pour la création d'un modèle de prédiction du méthane par contre, une telle sélection des échantillons sur base de leur variabilité spectrale était impossible. Nous avons dû chercher à créer de la variabilité directement à partir du phénotype (y). La seule issue pour poursuivre notre travail d'amélioration de ces modèles, tant pour les acides gras que pour le méthane a été par la suite de passer par la recherche de collaborations internationales nécessaires à l'échange de données. Cette voie est nécessaire pour résoudre le problème lié à une couverture suffisante de la variabilité afin d'améliorer la robustesse. Ceci est d'autant plus vrai que le modèle est indirect.

La même nécessité de collaboration est apparue par la suite à tous les niveaux : tant sur le point relatif à l'échantillonnage, afin qu'il soit réalisé dans de parfaite condition et de façon standardisée ; que sur le besoin d'harmoniser les analyses de référence, tant sur les méthodes de mesures, que sur les unités ; sur les modèles de prédiction ; la normalisation ; la standardisation spectrale et les méthodes de contrôle qualité. Les questions en suspens sont encore nombreuses, la rapidité et efficacité des solutions dépendront bien souvent de la volonté des différents acteurs du secteur à trouver des terrains d'entente et le courage ou l'envie de partager leurs savoirs. Ce travail ne pourra se faire que via une forte implication des institutions internationales telles qu'ICAR, ISO ou encore la FIL.

**Tableau 6.1. Tableau synthétique des facteurs impactant la qualité/robustesse des modèles, les problèmes, difficultés et points d'attention rencontrés et les solutions/recommandations apportées sur la première composante, l'Erreur total de l'échantillonnage (TSE), de l'Erreur Globale d'estimation (GEE).**

Facteurs impactant la qualité/robustesse	Problèmes/difficultés/Points d'attention	Solutions/recommandations
Couvrir un maximum de variabilité	Afin d'augmenter la robustesse du modèle, Celui-ci doit couvrir le maximum de la variabilité de la population à analyser	Si base de données spectrales massive accessible et si le phénotype a une empreinte dans le signal spectral FT-MIR ou du moins des caractères qui lui sont liés, on peut créer/améliorer le modèle à partir d'un échantillonnage effectué sur base de la variabilité spectrale (X) en utilisant par exemple les distances GH et NH  Si pas de base de données spectrale accessible ou si pas d'empreinte dans le signal spectral (p.ex. méthane), on doit chercher à maximiser la variabilité du phénotype (y) en utilisant ses sources probables de variation
Choix du moment d'échantillonnage	Variations possibles de la composition du lait en fonction de multiples facteurs qui peuvent entraîner des désynchronisations entre les mesures du phénotype et l'empreinte spectrale	Prévoir des études permettant d'estimer et corriger les éventuels biais dans le temps qui expliquent les désynchronisations  S'assurer que la mesure phénotypique soit représentative de la période considérée et du moment opportun de l'échantillonnage
Procédure d'échantillonnage	S'assurer de la représentativité de l'échantillonnage du lait  S'assurer de la représentativité de l'échantillonnage pour la mesure du phénotype	S'assurer de bien couvrir la variabilité intra et inter-lactationnelle  Respecter les règles et procédures liées à l'échantillonnage du lait (p.ex. ICAR ou ISO). Ceci est essentiel pour les modèles de prédiction indirects  Utiliser des méthodes d'échantillonnage standardisées et/ou parfaitement documentées

**Tableau 6.2. Tableau synthétique des facteurs impactant la qualité/robustesse des modèles, les problèmes, difficultés et points d'attention rencontrés et les solutions/recommandations apportées sur la deuxième composante, l'Erreur analytique globale (TAE), de l'Erreur Globale d'estimation (GEE).**

Facteurs impactant la qualité/robustesse		Problèmes/difficultés/Points d'attention	Solutions/recommandations
<b>Analyses de référence</b>	Precision et reproductibilité de la méthode	On doit chercher à estimer la répétabilité et reproductibilité de la méthode de mesure	
	Coût de la méthode de mesure	Parfois préférer une méthode de référence moins coûteuse et moins précise qui permette d'augmenter le nombre d'analyses de référence et mieux couvrir la variabilité de la population	
	Facilité de la méthode	La méthode de référence devra idéalement permettre de maximiser le nombre d'analyses à large échelle et en toute occasion	
	Méthode invasive/bien-être animal	Veiller à ce que le choix de méthode phénotypique n'influence pas le comportement de l'animal, ni son métabolisme et préférer une méthode éthiquement acceptable pour faciliter les mesures à large échelle	
	Méthode standardisée ou normalisée	Afin de permettre l'échange de données, préférer les méthodes normalisées/standardisées	
<b>Modèles de prédiction</b>	Utilisation d'une méthode commerciale pour mesurer le phénotype	S'assurer de l'accessibilité aux données brutes et éviter les méthodes en "boîte noire" aux étalonnages et corrections non ou mal maîtrisées	
	Unité de mesure	Veiller à l'unité de mesure, sa signification et son interprétation	
	Utilisation d'un prétraitement	Tester leur utilisation afin d'éliminer des effets indésirés	
<b>Normalisation</b>	Choix du modèle	En fonction du set d'échantillons disponible, de leur variabilité, du phénotype	
	Evaluation des modèles de prédiction pour les composés majeurs du lait	Normes ISO, FIL/ICAR existantes	
<b>Standardisation</b>	Evaluation des nouveaux modèles de prédiction non couverts par normes actuelles	Des solutions et recommandations internationales sont nécessaires. L'évaluation des modèles devrait inclure les informations sur la région spectrale utilisée, la complexité et méthode de validation et intégrer l'estimation des R <sup>2</sup> , SEC, SEP, RER, RPD, RMSE...	
	Comment s'assurer de l'équivalence et stabilité des réponses spectrales des tous les appareils FT-MIR ?	Méthodes apportées par les constructeurs : pas d'études indépendantes et pas de solutions entre constructeurs Utilisation d'un fichier de répétabilité entre deux appareils (set d'échantillons communs) : utilisation temporaire	
<b>Contrôle qualité</b>	Comment assurer le contrôle qualité des nouveaux modèles et éviter les extrapolations ?	Standardisation spectrale (Grelet et al., 2015) : méthode actuelle la plus adaptée, mais doit être améliorée Via l'estimation de la valeur GH des spectres, on pourrait voir si la variabilité du modèle couvre la variabilité des spectres à prédire	
	Comment assurer au quotidien le contrôle qualité des prédictions des nouveaux modèles ?	Via la standardisation spectrale + utilisation d'un daily check, associée au contrôle systématique des GH	

Erreur analytique globale (TAE)

Une des clefs importantes du problème, restera l'utilisation d'un système unique de standardisation spectrale, ce qui est loin d'être évident. En effet, dans l'absolu, lorsque les problèmes liés au système d'ancrage et au daily check auront pu être résolus, que nous pourrions assurer que les déviations et dérives de l'ensemble des appareils sont parfaitement maîtrisées, nous aurons alors un outil puissant, permettant de faire le lien entre les différentes équipes de recherches et laboratoires du monde entier pour créer des modèles communs et maîtriser le contrôle qualité des appareils au quotidien de l'ensemble des modèles en une étape. La tâche des laboratoires sera grandement facilitée et les coûts d'étalonnage fortement diminués. C'est probablement un argument important pour inciter l'ensemble des laboratoires à participer. Par contre, des réticences pourraient naître de la part de certaines parties réalisant la vente de gammes d'étalonnage.

Je m'en voudrais de clôturer ce travail, sans évoquer cette belle réussite dans la création de base de données spectrales, née à l'aube de mon expérience dans le domaine de la FT-MIR et d'une très belle collaboration. Cette approche innovante fut proposée par Soyeurt et al. (2006) et visait à stocker l'ensemble des informations spectrales issues des analyses d'échantillons de lait de vache individuels prélevés dans le cadre du contrôle laitier et de créer une base de données spectrales. Cette idée s'est concrétisée dès 2008 via la création de la première plateforme de collecte de l'information spectrale laitière, nommée Futurospectre. Cette dernière regroupe des acteurs de la filière associée au contrôle laitier et des acteurs de la recherche en Wallonie, tant dans le domaine de l'analytique, du contrôle de qualité, que de l'analyse de données, de la génétique et de la génomique (Dehareng et al., 2013). A l'ensemble des spectres d'absorbance mesurés par les appareils FT-MIR ont été associés toutes les données disponibles permettant l'identification et la caractérisation des animaux et toutes les prédictions infrarouges issues des spectromètres. D'autres valeurs de référence ont pu également être associées au fil des ans. Par la suite, une deuxième base de données a été créée afin de récolter l'ensemble des spectres mesurés sur les laits de tank prélevés lors de chaque ramassage par les camions des laiteries. Ces derniers échantillonnages sont réalisés selon les protocoles établis en Belgique pour le paiement du lait. Fin août 2022, le nombre de spectres wallons étaient d'environ 6.625.000 pour les laits individuels du contrôle laitier et de 5.470.000 pour les tanks à lait. D'autres initiatives de construction de bases spectrales ont par la suite été réalisées dans d'autres pays en prenant exemple sur cette réussite wallonne, par exemple, au Canada (Bonfatti et al., 2017), en Nouvelle-Zélande (Tiplady et al., 2019), au Royaume-Uni (Brand et al., 2021), en Irlande (Scarso et al., 2017) ou encore en Australie (Ho and Pryce, 2020). Actuellement, de nombreuses organisations d'élevage ont mis en place des pipelines pour stocker cette nouvelle source d'information sur la composition du lait. Comme le souligne Tiplady et al. (2019), la création de telles bases de données spectrales requiert l'utilisation d'une approche de standardisation adaptée, qui préférera l'utilisation d'une procédure impliquant des échanges de lait (p.ex. Grelet et al., 2015) à d'autres stratégies utilisées à priori, telle que celle proposée par exemple par Bonfatti et al. (2017).



Les avantages liés à la création de bases de données spectrales sont multiples. L'utilisation de nouveaux modèles sur une large base spectrale permet d'évaluer rapidement si les effets prédictions concordent avec ce qui est attendu et/ou déjà observé dans la littérature. Ceci a permis de notamment déceler le problème de coefficients de prédiction non constants en fonction de l'évolution du stade de lactation du premier modèle de prédiction du CH<sub>4</sub> (Vanlierde et al., 2015). D'autre part, cela aide à la création de nouveaux modèles via une réutilisation de paramètres enregistrés dans le passé dans la base de données. A titre d'exemple, on a pu des années plus tard, réutiliser des spectres provenant d'essais expérimentaux passés afin de chercher à prédire le poids des animaux (Tedde et al., 2021). L'utilisation de nouvelles prédictions sur une base spectrale historique permet également d'étudier les tendances observées sur une large population et ce pendant plusieurs années. Par exemple, en appliquant un modèle de prédiction de CH<sub>4</sub> sur une base de données de plusieurs millions de spectres, Vanlierde a ainsi pu étudier les variations intra et inter lactations, saisonnières, annuelles, et entre exploitations (Vanlierde et al., 2016). La disponibilité à grande échelle, de phénotypes, liés à de nouveaux caractères, directement prédits à partir de ces bases de données, va permettre d'effectuer les évaluations génétiques (Kandel et al., 2017 ; Scarso et al., 2017) et étudier les modifications phénotypiques et génétiques au cours de la lactation, soulignant la relation entre l'état physiologique de la vache et la composition du lait (Bastin et al., 2011). L'utilisation de telles bases, accroît le nombre de données phénotypiques et rend possible l'utilisation de génotypes plus anciens pour des caractères phénotypiques nouveaux, ceci accroît fortement les possibilités d'évaluations génomiques (Talpaly et al., 2020). Enfin, l'utilisation de ces bases de données spectrales vont aider au développement d'outils d'aide à la décision (Brand et al., 2021). L'association des spectres, comme source d'information importante, à d'autres variables et données de référence, outils ou technologies, tels des capteurs en ferme, la télédétection, le cloud computing ou l'internet des objets ou « Internet of Things » (IoT), devrait permettre de contribuer au développement du concept d'agriculture intelligente ou « Smart Farming ». Le projet WALLeSMART en région wallonne, cherche déjà à développer de tels outils (Roukh et al., 2020).

Enfin, l'utilisation de base de données spectrales permet également d'utiliser des modèles particuliers, cités en introduction de la discussion, que sont les modèles qui visent à détecter les échantillons anormaux ou permettant de détecter la présence d'éventuels adultérant. Le risque d'adultération est moins présent en Europe, car le système de contrôle est bien développé, mais, à l'échelle mondiale, on estimait encore entre 1980 et 2010, que le lait était le deuxième type d'aliment le plus souvent frelaté avec 14 % des cas (Moore et al., 2012). L'utilisation de la FT-MIR pour détecter ces échantillons anormaux est donc intéressante. La difficulté est de construire une base de données d'échantillons totalement exempt d'adultérant, y compris présent par accidents. D'autre part, il serait bon de standardiser au préalable l'ensemble des spectres de référence utilisés pour construire la base de données spectrales de référence, ainsi que les spectres devant être comparés à celle-ci. A ma connaissance,

l'ensemble des outils proposés dans la littérature actuelle n'utilise aucune standardisation spectrale (p.ex. Fernández Pierna et al., 2016, Hansen and Holroyd, 2019 ; Spieß et al., 2021), or, cette dernière ne pourrait qu'améliorer la fiabilité de ces outils. Par exemple, malgré l'utilisation de la solution Equalizer de la société Foss (Hillerod, Danemark) pour corriger l'ensemble des spectres de la base de données spectrale utilisée, Spieß et al. (2021) n'ont pas pu éviter des problèmes d'artefacts liés à des problèmes supposés de machine, ce qui aurait pu être évité avec l'utilisation d'une standardisation spectrale (Grelet et al., 2015).

Enfin, une étape importante à la création commune de toute base de données, est la recherche d'une solution d'accord régulant les échanges de données spectrales et éventuellement de données de référence. Cette étape essentielle ne doit pas être sous-estimée, car elle inclut souvent des nombreux acteurs ayant des objectifs parfois bien différents. C'est souvent une des étapes les plus difficiles et complexes à laquelle les chercheurs sont confrontés. Dans ces accords, on devra tout particulièrement veiller à ce que l'accessibilité aux nouveaux modèles découlant de ces bases de données soient disponibles pour le plus grand nombre et du moins pour l'ensemble des fournisseurs de données. Il ne serait pas acceptable que des efforts communs soient accaparés par des sociétés privées ou constructeurs.

C'est donc à nouveau une histoire de compromis, d'accords, de collaborations, et d'échanges coordonnés à créer. Comme Nicolas Gengler l'a très bien suggéré en septembre 2022, en complément de la phrase de Mike Coffey citée dans la partie introductive de cette thèse en référence au manque d'informations phénotypiques : « à l'ère de la génomique, le phénotype est roi ! » (Coffey, 2011), il serait bon d'ajouter « Comme, sont impériales les nouvelles collaborations et coordinations structurées qui cherchent à combler ce déficit » (Gengler, 2022).

## Références

- Bastin, C, Gengler, N, and Soyeurt, H., 2011. Phenotypic and genetic variability of production traits and milk fatty acid contents across days in milk for Walloon Holstein first-parity cows. *J. Dairy Sci.*, 94(8), 4152-63.
- Bonfatti V., Tiezzi F., Miglior F., Carnier P., 2017. Comparison of Bayesian regression models and partial least squares regression for the development of infrared prediction equations. *J Dairy Sci.*, 100(9):7306-7319.
- Brand W., Wells A.T., Smith S.L., Denholm S.J., Wall E. and Coffey M.P., 2021. Predicting pregnancy status from mid-infrared spectroscopy in dairy cow milk using deep learning. *J. Dairy Sci.*, Volume 104, Issue 4, pp. 4980-4990.
- Coffey M., 2011. Phenotyping that maximises the value of genotyping. In 'General assembly and annual workshop of ICAR 2011. New technologies and new challenges for breeding and herd management'. Bourg-en-Bresse, France, 22–24 June 2011.
- Dehareng, F., Soyeurt, H., Gengler, N., Bastin, C., Veselko, D., Bertozzi, C., Baeten, V., Sinnaeve, G., and Dardenne, P., 2013. Qualité du lait: la Wallonie à la pointe grâce à l'utilisation de la spectrométrie infrarouge. Proc. In 18ème Carrefour des productions animales: Nouvelles approches pour une optimisation de nos élevages laitiers. Février 2013, Gembloux, Belgium. pp. 93-97.
- Fernández Pierna J.A., Vincke D., Baeten V., Grelet C., Dehareng F. and 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, *Chemometrics and Intelligent Laboratory Systems*, Volume 152, Pages 157-162.
- Gengler N., 2022. Collaboration is imperial or how to address this critical issue when developing novel traits. Proc. EAAP 77th Annual Meeting, Porto, Wageningen Academic Publishers, Wageningen, the Netherlands, p. 329.
- Grelet C., Pierna J.F., Dardenne P., Baeten V. and Dehareng F., 2015. Standardization of milk mid-infrared spectra from a European dairy network. *J. Dairy Sci.*, 98 (2015), pp. 2150-2160
- Hansen, P. W. 1998. Urea determination in milk using Fourier transform infrared spectroscopy and multivariate calibration. *Milchwissenschaft* 53:251–255.

- Ho P. and Pryce J., 2020. Predicting the likelihood of conception to first insemination of dairy cows using milk mid-infrared spectroscopy. *J. Dairy Sci.*, 103:11535–11544.
- Kandel P.B., Vanrobays M.-L., Vanlierde A., Dehareng F., Froidmont E., Gengler N., Soyeurt H., 2017. Genetic parameters of mid-infrared methane predictions and their relationships with milk production traits in Holstein cattle. *J. Dairy Sci.* 100, Issue 7, pp. 5578-5591.
- Moore J. C., Spink J. and Lipp M., 2012. Development and application of a database of food ingredient fraud and economically motivated adulteration from 1980 to 2010. *Journal of Food Science* 77 R108– R116.
- Roukh A., Nolack Fote F., Mahmoudi S.A. and Mahmoudi S, 2020. Big Data Processing Architecture for Smart Farming, *Procedia Computer Science*, Volume 177, pp. 78-85.
- Scarso S., McParland S., Visentin G., Berry D.P., McDermott A. and De Marchi M., 2017. Genetic and nongenetic factors associated with milk color in dairy cows, *J. Dairy Sci.*, Volume 100, Issue 9, pp. 7345-7361.
- Soyeurt H., Dardenne P., Dehareng F., Lognay G., Veselko D., Marlier M., Bertozzi C., Mayeres P. and Gengler N. 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. *J. Dairy Sci.*, 89, 3690–3695.
- Spieß L., de Peinder P. and van den Bijgaart H., 2021. Advances in Atypical FT-IR Milk Screening: Combining Untargeted Spectra Screening and Cluster Algorithms. *Foods*, 10, 1111.
- Tedde A., Grelet C., Ho P.N., Pryce J.E., Hailemariam D., Wang Z., Plastow G., Gengler N., Brostaux Y., Froidmont E., Dehareng F., Bertozzi C., Crowe M.A., Dufrasne I., GplusE Consortium Group, and Soyeurt H., 2021. Validation of Dairy Cow Bodyweight Prediction Using Traits Easily Recorded by Dairy Herd Improvement Organizations and Its Potential Improvement Using Feature Selection Algorithms. *Animals*, 11(5):1288.
- Tiplady K.M., Sherlock R.G., Littlejohn M.D., Pryce J.E., Davis S.R., Garrick D.J., Spelman R.J. and Harris B.L., 2019. Strategies for noise reduction and standardization of milk mid-infrared spectra from dairy cattle. *J Dairy Sci.*, 102(7):6357-6372.
- Tiplady K.M., Lopdell T.J., Littlejohn M.D. and Garrick D.J., 2020. The evolving role of Fourier-transform mid-infrared spectroscopy in genetic improvement of dairy cattle. *J Anim Sci Biotechnol.*, 11:39.

- 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, pp. 5740-5747
- 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*, 56, 258-264.
- Vanlierde A., Dehareng F., Gengler N., Froidmont E., McParland S., Kreuzer M., Bell M., Lund P., Martin C., Kuhla B. and Soyeurt H., 2021. Improving robustness and accuracy of predicted daily methane emissions of dairy cows using milk mid-infrared spectra. *J Sci Food Agric.*,101(8):3394-3403.