

**DÉVELOPPEMENT DE MÉTHODES DE SPECTROSCOPIE ET DE MICROSCOPIE  
PROCHE INFRAROUGE (PIR) POUR L'ANALYSE DU LAIT**

**Lygens JOSEPH**

**TRAVAIL DE FIN D'ETUDES PRESENTE EN VUE DE L'OBTENTION DU DIPLOME  
DE MASTER COMPLEMENTAIRE EN SCIENCES ET TECHNOLOGIE DES  
ALIMENTS**

**ANNEE ACADEMIQUE 2015-2016**

**CO-PROMOTEURS: Vincent BAETEN (CRA-W; UCL) et Ouissam ABBAS (CRA-W)**

## RESUME

Cette étude a été réalisée dans le but de développer deux méthodes d'analyses rapides pour le lait, l'une en microscopie proche infrarouge et l'autre en spectroscopie proche infrarouge (PIR).

La première a été établie sur l'extrait sec de 24 échantillons de laits entiers, demi-écrémés et écrémés séchés sur 2 plaques hydrophobes de 96 puits. L'optimisation de la technique d'analyse a été faite en mode transmission sur un lait demi-écrémé. Elle consistait à déposer un volume idéal d'une goutte sur un puits, de choisir une durée optimale de séchage (23 min) ainsi qu'un nombre suffisant de gouttes cumulées (5 gouttes). Les études de répétabilité et de reproductibilité ont été conduites et évaluées par des tests ANOVA. Les différences entre les puits et entre les deux plaques ne sont pas significatives contrairement aux résultats obtenus lors de la comparaison des spectres collectés en fonction du temps à savoir, une différence significative indiquant une évolution du lait au cours de la conservation. Une ACP effectuée sur l'ensemble des spectres montre une nette discrimination des 3 catégories de lait et confirme que la principale variable de différenciation entre elles est bel et bien la matière grasse.

Dans la seconde méthode, il s'agissait de calibrer un spectromètre portatif dit NIR4FARM pour déterminer la composition des principaux constituants du lait. Les performances des équations de calibration pour l'analyse du lait obtenues avec le NIR4FARM sont en-deçà de celles obtenues avec le Bruker MPA qui est un spectromètre de laboratoire bien adapté à l'analyse du lait, vu sa configuration au CRA-W. Dans un premier temps, la différence de gamme spectrale entre les 2 instruments (NIR4FARM : 920-1658 nm, MPA : 800-2500 nm) nous laissait supposer une trop grande perte d'information spectrale pour le NIR4FARM. Cependant, en n'utilisant que la gamme du NIR4FARM avec la base de données du MPA, les performances des équations sont toujours excellentes. Ceci confirme que la gamme spectrale du NIR4FARM est bien riche en information et que la cause du manque de précisions des équations de calibrations développées avec le NIR4FARM n'est pas liée à la gamme spectrale mais certainement aux conditions d'analyse et de présentation de l'échantillon. En effet, pour le MPA, les mesures sont réalisées en transmission, à la température de 40°C, et l'échantillon est homogénéisé avant l'analyse, tandis que pour le NIR4FARM, les mesures sont réalisées en transréflexion, dans une cellule en or qui n'est à priori pas dédiée à l'analyse du lait, à température ambiante et sans homogénéisation.

**MOTS-CLES:** Spectroscopie PIR, microscopie, lait, extrait sec, ACP, calibration, répétabilité.

## **ABSTRACT**

This study was conducted to develop two methods of rapid tests for milk, one in microscopy (MicroNIR) and the other in near infrared spectroscopy (NIR).

The first was realized on the dry extract of 24 samples of whole milk, semi-skimmed and skimmed on 2 hydrophobic 96 well plates manufactured on silica. The optimization of the analysis technique was made in transmission mode on semi-skimmed milk. It was to prepare the plate with a convenient volume of a drop deposited on a well (5  $\mu$ l), optimal drying time (23 min) and a sufficient number of accumulated drops (5 drops). Repeatability and reproducibility studies have been undertaken and evaluated by ANOVA tests. Differences between spectra collected in various wells and those of the two plates are not significant inversely to the finding obtained with spectra collected during storage time showing significant difference and thus an evolution of the sample during storage. PCA applied on all spectra shows a clear discrimination of the 3 categories of milk and confirms that the main variable of differentiation between them is indeed fat.

This device was originally designed for the analysis of fresh fodder directly on the farm. The performance of calibration equations for analysis of milk obtained with the NIR4FARM are well below those obtained with the Bruker MPA. MPA is a laboratory spectrometer particularly suitable according to his configuration at CRA-W to milk analysis. The difference in spectral range between 2 instruments (NIR4FARM: 920-1658 nm, MPA: 800-2500 nm) let us suppose at first to a very great loss of spectral information for NIR4FARM. However, using only the range of NIR4FARM (920-1658 nm) with the MPA database, the performance of calibration equations are always excellent. This confirms that the spectral range of NIR4FARM is rich in spectral information and the lack of precision of calibration equations developed with NIR4FARM is not related to the spectral range but certainly the conditions of analysis and presentation of the sample. Indeed, for the MPA, the measurements are made in transmission, at the temperature of 40 °C, and the sample is homogenized prior to analysis. While for NIR4FARM, the measurements are performed in transreflexion, in a gold cell that is not dedicated to milk analyze at room temperature and without homogenization.

**KEYWORDS:** NIR spectroscopy, microscopy, milk, dry extract, ACP, calibration, repeatability.

## **REMERCIEMENTS**

Je tiens à remercier mon professeur Vincent BAETEN et Ouissam ABBAS, co-promoteurs de ce travail pour leur leadership dans l'élaboration de ce document.

Mes sincères remerciements s'adressent à Pierre DARDENNE, chef de département et plus particulièrement à Olivier MINET et Bernard LECLER, mes maîtres de stage, pour l'assistance, la disponibilité et la promptitude sans borne dont ils ont toujours fait montre.

Je remercie vivement les techniciens Quentin ARNOULD, Stéphane BRICHARD et Zio DRISSA pour leur irréprochable présence. Enfin, je suis reconnaissant à toute l'équipe du CRA-W et spécialement des laboratoires de spectroscopie, de microscopie et de microscopie proche infrarouge pour leur contribution d'une façon ou d'une autre à la réalisation de ce laborieux travail de fin d'étude.

J'exprime une reconnaissance particulière

- aux responsables de Master les professeurs François BERA et Yvan LARONDELLE
- à mes camarades de stage au CRA-W: Philogène VILSAINT, Aurélien TAIRA, Marie-Babeth EBEN, Hélène MWENGE BINTI YUMA et Soundous EL-HAJJAJI
- et à ma famille en Haïti.

## TABLE DES MATIERES

RESUME .....	i
ABSTRACT.....	ii
REMERCIEMENTS.....	iii
LISTE DES FIGURES .....	vii
LISTE DES TABLEAUX.....	viii
LISTE DES ANNEXES .....	ix
LISTE DES ABREVIATIONS.....	x
INTRODUCTION .....	1
MATERIELS ET METHODES .....	2
PARTIE I .....	2
I. Développement de la méthode d'analyse du lait en microscopie PIR .....	2
1.1. Objectif.....	2
1.2. Description des échantillons de lait.....	2
1.3. Matériel utilisé.....	2
1.4. Procédure de développement de la méthode de microscopie PIR .....	3
1.4.1. Optimisation des paramètres d'analyse.....	4
1.4.1.1. Collecte de spectres PIR de l'extrait sec du lait .....	4
1.4.1.2. Choix du volume de remplissage d'un puits d'une plaque .....	4
1.4.1.3. Définition du temps de séchage d'une goutte de lait.....	5
1.4.1.4. Choix du nombre de goutte au sein d'un puits .....	5
1.4.1.5. Répétabilité au sein d'un puits de la plaque .....	5
1.4.1.6. Répétabilité au sein de plusieurs puits de la plaque.....	5
1.4.1.7. Reproductibilité lors d'un processus de conservation .....	5
1.4.1.8. Reproductibilité entre deux plaques.....	6
1.4.2. Analyse des échantillons .....	6
1.4.3. Prétraitement des données .....	6
1.4.3.1. Correction de la ligne de base .....	6
1.4.3.2. Déviation normale standardisée (SNV) .....	7
1.4.3.3. Dérivation .....	7

1.4.4. Traitement des données .....	7
1.4.4.1. Etude de répétabilité et de reproductibilité par RMS et ANOVA .....	7
1.4.4.2. Discrimination des échantillons de lait par ACP .....	8
1.4.4.2.1. Discrimination des échantillons de lait selon les différentes catégories .....	8
1.4.4.2.2. Discrimination des échantillons de lait en fonction du temps de conservation .....	8
1.5. Etablissement d'un modèle de calibration pour la caractérisation du lait .....	8
PARTIE II .....	9
II. Prédiction des principaux constituants du lait en spectroscopie PIR .....	9
2.1. Objectif .....	9
2.2. Description des échantillons de lait .....	9
2.3. Matériels utilisés et mode opératoire .....	9
2.3.1. Le NIR4FARM .....	9
2.3.2. Le spectromètre Bruker MPA .....	11
2.3.2.1. Transfert de la Base de données MPA vers NIR4FARM .....	11
2.3.2.2. Prédiction de valeurs de référence pour la calibration du NIR4FARM .....	12
2.4. Méthodes de référence utilisées .....	12
2.5. Prétraitement des données .....	12
2.6. Traitement des données .....	12
2.6.1. L'erreur standard de calibration (SEC) et de cross-validation (SECV) .....	13
2.6.2. L'erreur standard de prédiction (SEP) et de cross validation (SECV) .....	13
2.6.3. Le coefficient de détermination ( $R^2$ ) .....	13
2.6.4. Le RPD ou Ratio of Prediction Deviation .....	13
RESULTATS ET DISCUSSION .....	14
PARTIE I .....	14
III. Développement de la méthode d'analyse du lait en microscopie PIR .....	14
3.1. Optimisation des paramètres d'analyse des échantillons .....	15
3.1.1. Collecte de spectres de l'extrait sec du lait .....	15
3.1.2. Choix du volume de remplissage d'un puits de la plaque .....	16
3.1.3. Définition du temps de séchage d'une goutte .....	16
3.1.4. Choix du nombre de gouttes au sein d'un puits .....	17
3.1.5. Répétabilité au sein d'un puits de la plaque .....	18

3.1.6. Répétabilité au sein de plusieurs puits de la plaque .....	19
3.1.7. Reproductibilité lors d'un processus de conservation.....	19
3.1.8. Reproductibilité entre deux plaques .....	19
3.2. Discrimination des échantillons de lait .....	20
3.2.1. Discrimination des échantillons de lait suivant les différentes catégories .....	20
3.2.2. Discrimination des échantillons de lait en fonction du temps de conservation.....	21
3.3. Etablissement d'un modèle calibration pour la caractérisation du lait .....	23
PARTIE II.....	25
IV. Prédiction de valeurs sur le lait en spectroscopie PIR.....	25
4.1. Transfert de la base de données MPA vers le NIR4FARM .....	25
4.2. Calibration avec les spectres du NIR4FARM .....	26
4.2.1. Equations obtenues en régression MLR et PLS pour le NIR4FARM.....	26
4.2.2. Comparaison des équations du NIR4FARM avec les équations du Bruker MPA .....	27
4.2.3. Comparaison de la performance des équations du MPA obtenues avec la gamme complète du MPA et avec la gamme du NIR4FARM.....	29
CONCLUSION ET PERSPECTIVES.....	30
BIBLIOGRAPHIE.....	A
ANNEXE .....	B

## LISTE DES FIGURES

Figure 1: Microscope PIR (Hyperion 3000) .....	3
Figure 2: Procédure d'optimisation de l'analyse des échantillons de lait.....	3
Figure 3: Plaque hydrophobe en silicium à 96 puits contenant des échantillons de laits séchés....	4
Figure 4: Spectromètre NIR4FARM .....	10
Figure 5: Spectromètre PIR (Bruker MPA) équipé d'un homogénéisateur.....	11
Figure 6: Spectres PIR des laits entiers, demi-écrémés et écrémés .....	14
Figure 7: Spectres du lait demi-écrémé: Lait liquide ( <i>Normalisation - rouge</i> ) Vs Extrait sec du lait ( <i>Baseline + Normalisation-bleu</i> ).....	16
Figure 8: Variation de l'information spectrale sous l'effet de la durée de séchage .....	17
Figure 9: Spectres selon le nombre de gouttes séchées cumulées ( <i>correction de la ligne de base</i> ) .....	17
Figure 10: Rapport de surface de la bande spectrale allant de 2234 à 2385 nm selon le nombre de gouttes séchées cumulées.....	18
Figure 11: Résultats de l'ACP (PC-1 (59%) vs PC-2(30%)) appliquée sur les spectres NIR des 3 catégories de laits (EN, DE, EC), <i>traités par SNV, Baseline, 1ère dérivée et Moyenne</i> .....	20
Figure 12: Résultats de l'ACP appliquée sur les spectres NIR de chaque catégorie de lait (EN, DE, EC) <i>collectés en 3 analyses (1<sup>er</sup>, 3<sup>e</sup> et 5<sup>e</sup> jour) et traités par SNV, baseline et 1<sup>ère</sup> dérivée</i> ..	22
Figure 13: Corrélation entre les valeurs prédites et les analyses de référence pour la matière grasse (A), les protéines (B) et la matière sèche (C).....	24
Figure 14: Spectres d'une même série d'échantillons de lait ( <i>1 entier, 1 demi- écrémé, 1 écrémé et 1 cru</i> ) collectés sur le MPA ( <i>bleu</i> ) et sur le NIR4FARM ( <i>rouge</i> ) .....	25
Figure 15: Comparaison des Erreurs Standard de Prédiction (SEP) obtenues pour le NIR4FARM <i>en MLR</i> et le BRUKER MPA <i>en PLS</i> .....	27
Figure 16: Spectre du lait sur le MPA: Indication du range utilisé par le NIR4FARM ( <i>rectangle rouge</i> ) et zones utilisées pour le développement des équations pour les variables matière grasse ( <i>Fat</i> ) <i>en vert</i> , protéines <i>en mauve</i> et matière sèche ( <i>Dry matter</i> ) <i>en marron</i> .....	28
Figure 17: Comparaison des équations de prédiction du MPA dans la gamme allant de 800 à 2500 nm (gamme MPA) avec celle allant de 920 à 1658 nm (gamme NIR4FARM) .....	29



## **LISTE DES TABLEAUX**

Tableau 1: Les différents paramètres optimisés.....	15
Tableau 2: Quelques paramètres de l'équation obtenue pour la méthode .....	23
Tableau 3: Différence entre le MPA et le NIR4FARM.....	25
Tableau 4: Equations en régressions MLR vs PLS pour le NIR4FARM .....	26

## LISTE DES ANNEXES

Annexe 1: Quantités et types de laits de commerce utilisés .....	B
Annexe 2: Photos des laits de commerce utilisés .....	B
Annexe 3: Fonctionnement schématisé du spectromètre à transformée de Fourier (FTIR).....	C
Annexe 4: Les fonctions statistiques.....	D
Annexe 5: Les caractéristiques affichées sur l'emballage des échantillons de lait de commerce utilisés dans l'étude .....	F
Annexe 6: Résultats de Répétabilité au sein d'un puits d'une plaque .....	G
Annexe 7: Résultats de Répétabilité au sein de plusieurs d'une plaque.....	G
Annexe 8: Résultats de reproductibilité au cours de la conservation .....	H
Annexe 9: Résultats du "test post-hoc" sur la reproductibilité des spectres collectés entre le 1er et le 3e jour de conservation .....	H
Annexe 10: Résultats du "test post-hoc" sur la reproductibilité des spectres collectés entre le 1er et le 3e jour de conservation .....	I
Annexe 11: Résultats de reproductibilité entre deux plaques .....	I
Annexe 12: Valeurs de prédiction utilisées comme valeurs de référence en Microscopie pour la prédiction PLS .....	J
Annexe 13: Attribution des bandes spectrales en spectrométrie PIR pour les principaux composants du lait et des produits laitiers .....	K
Annexe 14: Loading de la 1ère composante de l'ACP appliquée sur les spectres NIR de la catégorie des lait écrémés (spectres traités par SNV, baseline et 1ère dérivée) .....	L
Annexe 15: Valeurs prédites corrigées du biais utilisées comme valeurs de référence en Spectroscopie pour la calibration du NIR4FARM.....	M
Annexe 16: Paramètres exhaustifs des équations obtenues en régression MLR pour le NIR4FARM .....	M
Annexe 17: Paramètres exhaustifs des équations obtenues en régression PLS pour le NIR4FARM .....	N
Annexe 18: Paramètres des équations obtenues en régression PLS pour le MPA .....	N

## LISTE DES ABREVIATIONS

ACP	: Analyse en Composante Principale
ANOVA	: Analyse of Variance
CBL	: Confédération Belge de l'Industrie Laitière
CRA-W	: Centre Wallon de Recherches Agronomiques
DE	: Demi-écrémé
DM	: Dry Matter
EC	: Ecrémé
EN	: Entier
FTIR	: Spectroscopie Infrarouge à Transformé de Fourier
G	: Goutte
GLU	: Glucose
min	: Minute
MG	: Matière grasse
MLR	: Multi Linear Regression
MS	: Matière sèche
mm	: Millimètre
mm <sup>2</sup>	: Millimètre carré
NIR	: Near Infrared
nm	: Nanomètre
PIR	: Proche Infrarouge
PLS	: Partial Least Square
PRO	: Protéines
R <sup>2</sup>	: Coefficient de détermination
RMS	: Root Mean Square

RPD	: Ratio of prediction Deviate
SD	: Standard Deviation
SEC	: Erreur Standard de Calibration
SECV	: Erreur Standard de Cross Validation
SEP	: Erreur Standard de Prédiction
SEPC	: Erreur Standard de Prédiction corrigée du biais
PC	: Principale Composante
SNV	: Standard Normal Variate
UHT	: Ultra Haute Température
VS	: Versus
μl	: Microlitre
°C	: Degré Celsius
%	: Pourcentage

## INTRODUCTION

Le lait a toujours été un élément vital dans l'alimentation de l'homme depuis 12000 ans, époque à laquelle celui-ci a commencé à domestiquer les animaux (Courtet-Leymaryos, 2010).

A l'ère actuelle, une consommation massive de cette denrée est constatée. La Confédération Belge de l'industrie Laitière (CBL) (2015) affirme qu'en 2014, près de 805 milliards de litres de lait dont près de 90 % issus de vaches laitières sont produits à travers le monde. Selon la même source, la production des pays de l'Union Européenne s'élevait à 167 milliards de litres dont 2,1 % pour la Belgique.

En Europe, l'industrie laitière rétribue les fermiers selon la qualité microbiologique et la richesse nutritionnelle du lait (CBL, 2015). Cette dernière induit à l'utilisation de méthodes d'analyses rapides comme la spectroscopie infrarouge, technique existant depuis environ 40 ans (Bertrand *et al.*, 2000). Celle-ci garantit non seulement d'importantes rentrées économiques mais se veut un outil rapide, fiable et efficace pour le contrôle qualité et la détection de fraude. En témoigne la crise de la mélamine en 2008, où des producteurs chinois ont ajouté au lait une substance révélée nocive (mélamine) dans le but de le faire apparaître plus riche en protéines.

Cependant, pour plusieurs raisons, l'échantillonnage et l'analyse peuvent s'avérer un véritable challenge pour les techniciens. Dans certains cas, les quantités de lait dont les techniciens disposent sont insuffisantes et l'analyse ne peut être réalisée en double car les spectromètres de routine requièrent généralement un grand volume d'aliquote. Dans d'autres cas, certains laits de mammifères d'intérêt scientifique comme celui du porc sont peu abondants, difficilement accessibles ou exigent d'être échantillonnés en petites quantités sur plusieurs jours. De là découle l'intérêt de développer une méthode d'analyse rapide de microquantités de lait, soit l'un des objectifs de cette étude.

De même, le secteur de la spectrométrie proche-infrarouge fait d'énormes avancées, notamment dans la miniaturisation et la portabilité des instruments. La société anglaise AB Vista a développé sur base d'un spectromètre miniaturisé VIAVI, un instrument nommé NIR4FARM, conçu pour l'analyse des fourrages en frais directement sur le terrain. Dans le souci d'apporter un outil d'analyse rapide à l'agriculteur, nous essayons par cette étude, d'explorer les possibilités d'utilisation de ce matériel portatif au sein de la ferme pour l'analyse du lait.

## **MATERIELS ET METHODES**

Ce travail concerne l'analyse du lait par les méthodes analytiques rapides. Deux méthodes ont été développées; la première concerne l'analyse de microquantités de lait par microscopie proche infrarouge alors que la seconde concerne l'analyse du lait par un mini-spectromètre proche infrarouge dans un modèle miniaturisé (de type VIAVI). Ce dernier est utilisé dans le but de déterminer les principaux paramètres du lait, à savoir: protéines, matière grasse et matière sèche. Dans le souci d'être le plus clair possible, nous avons traité ces deux méthodes séparément.

### **PARTIE I**

#### **I. Développement de la méthode d'analyse du lait en microscopie PIR**

##### **1.1. Objectif**

Cette partie du travail a pour objectif le développement d'une méthode de microscopie proche infrarouge pour l'analyse de microquantités de lait.

##### **1.2. Description des échantillons de lait**

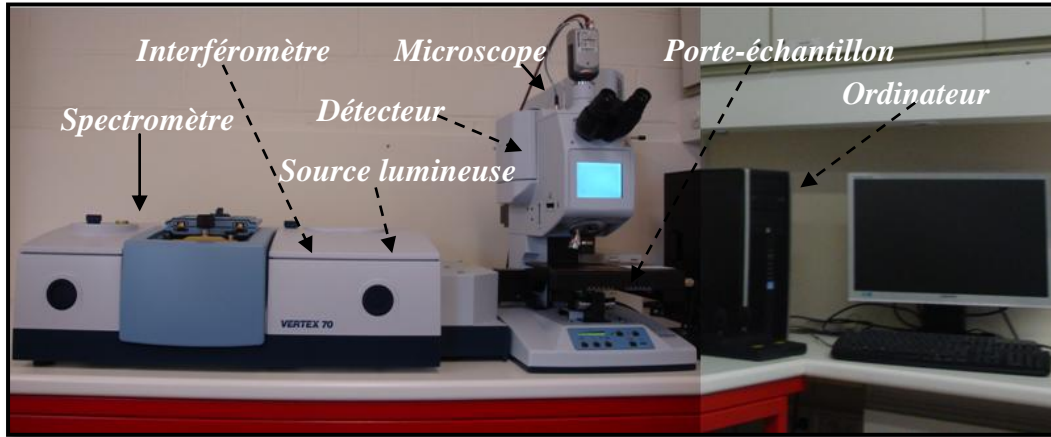
Nous avons utilisé 24 échantillons de lait du commerce. Ils ont été initialement homogénéisés, stérilisés en UHT et conditionnés en boîte ou en bouteille de 1 litre (Annexe 2). L'achat a été fait à Gembloux à partir des rayons de centres commerciaux: Colruyt, Match et Red Market. La méthode d'échantillonnage aléatoire stratifiée a été utilisée. Un total de 11 marques différentes, toutes européennes dont 8 d'origine belge, regroupées en 3 catégories théoriquement différentes (9 entiers, 11 demi-écrémés et 4 écrémés) ont constitué nos 24 échantillons (Annexe 1).

##### **1.3. Matériel utilisé**

Un système (Figure 1) formé d'un spectromètre FTIR (Vertex 70), d'un microscope (Hyperion 3000) et d'un ordinateur a été utilisé pour l'analyse des échantillons de lait. Nos analyses ont été faites en mode de transmission dans la partie du spectre PIR. Quelques composants de l'instrument combiné sont indiqués dans la figure 1. On peut en citer:

1. une source lumineuse PIR dont la lampe est en "Tungsten Allogen"
2. un interféromètre qui est un dispositif de mesure simultanée des fréquences par production d'interférences (Annexe 3)
3. un porte-échantillon qui supporte la plaque contenant l'échantillon à analyser,

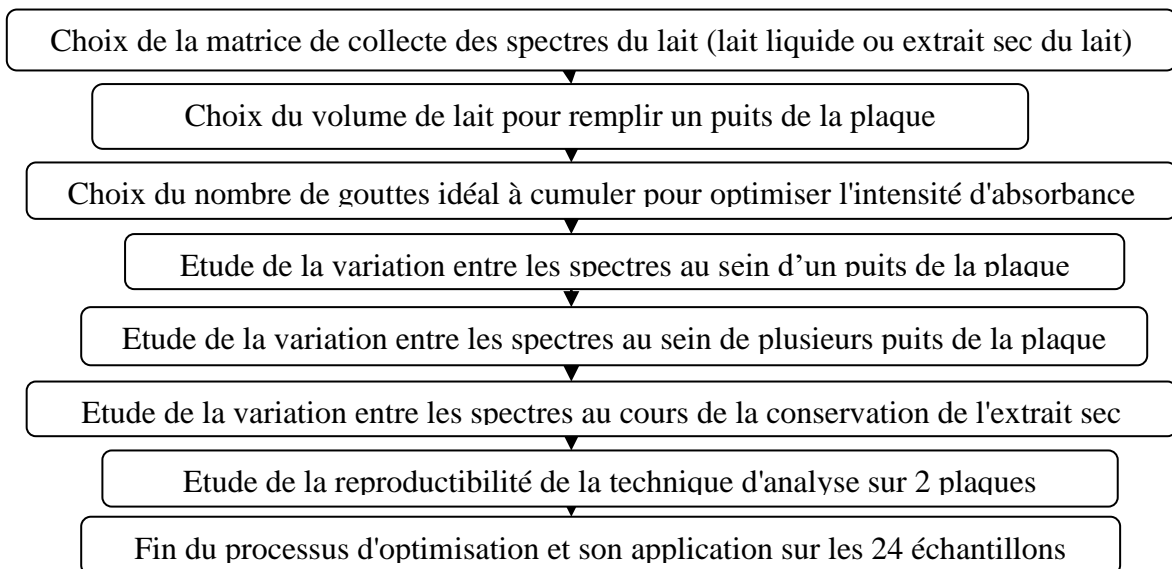
4. un détecteur (InGaAs-Diode) qui travaille dans la gamme 780-2500 nm
5. un ordinateur qui est le système d'acquisition des données par conversion de signal analogique en signal numérique.



**Figure 1: Microscope PIR (Hyperion 3000)**

#### 1.4. Procédure de développement de la méthode de microscopie PIR

Avant d'analyser l'ensemble des échantillons de lait, nous avons d'abord procédé à l'optimisation de la méthode. Etant donné les multiples facteurs à prendre en compte, la procédure d'établissement de la méthode est relativement complexe (Figure 2). Avant de la décrire, voici un schéma qui la résume:



**Figure 2: Procédure d'optimisation de l'analyse des échantillons de lait**

### 1.4.1. Optimisation des paramètres d'analyse

Afin de trouver un compromis entre les caractéristiques des 3 catégories (entier, demi-écrémé, écrémé) de lait dont nous disposons, le processus de préparation a été optimisé sur un lait demi-écrémé.

#### 1.4.1.1. Collecte de spectres PIR de l'extrait sec du lait

D'après Bertrand *et al*, (2000), des interférences qui sont dues au spectre de l'eau sont observées sur les spectres du lait. Aussi de notre côté nous avons fait le même constat sur des échantillons de lait analysés dans leur état liquide (plus bas sur la figure 7). Afin d'éviter ou de minimiser cette interférence et d'accentuer l'empreinte spectrale propre au lait, nous avons choisi de faire l'analyse sur l'extrait sec.

#### 1.4.1.2. Choix du volume de remplissage d'un puits d'une plaque

A l'aide d'une micropipette pouvant prélever entre 2 à 20  $\mu$ l, des gouttes de lait ont été mises à sécher sur une plaque hydrophobe en silicium comptant 96 puits de 7 mm de diamètre chacun (Figure 3). Des volumes de 3; 5 et 7  $\mu$ l de lait ont été déposés sur des puits de la plaque. Le volume choisi pour la suite des analyses est celui qui avait garanti un étalement optimal d'une goutte c'est-à-dire celle occupant uniformément la surface d'un puits tout en ne pas débordant ou dépassant son contour.

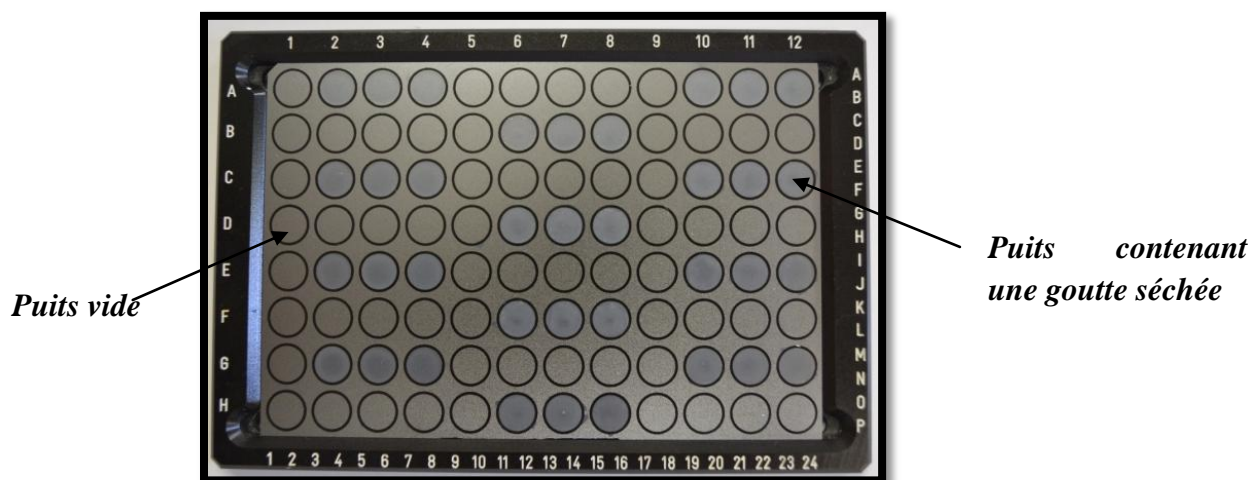


Figure 3: Plaque hydrophobe en silicium à 96 puits contenant des échantillons de laits séchés



#### **1.4.1.3. Définition du temps de séchage d'une goutte de lait**

Pour choisir la durée optimale de séchage d'une goutte sur la plaque à température ambiante (21.5°C en moyenne), les spectres des gouttes enregistrés respectivement à 11; 17; 23; 26; 30; 40; 60; 90 et 120 minutes de séchage ont été comparés entre eux afin de vérifier l'évolution (augmentation et stabilité) de la réponse spectrale. La durée de séchage minimale à partir de laquelle cette information a tendance à se stabiliser est choisie comme durée optimale.

#### **1.4.1.4. Choix du nombre de goutte au sein d'un puits**

Dans le but de réduire les bruits aléatoires susceptibles de perturber l'information spectrale et d'optimiser l'intensité d'absorbance, nous avons cumulé 2, 3, 4, 5, 6 et puis 7 gouttes, l'ajout de la goutte suivante ayant été réalisé sur l'extrait sec des gouttes cumulées précédemment. Le nombre optimal de gouttes cumulées séchées fera l'objet de l'analyse des 24 échantillons. Ce choix sera un compromis entre le temps de séchage des gouttes pour la préparation de la plaque, de la répartition uniforme de l'extrait sec cumulé au sein d'un puits et de la stabilité de l'information spectrale entre différentes quantités de gouttes cumulées.

#### **1.4.1.5. Répétabilité au sein d'un puits de la plaque**

Le microscope PIR (Hyperion 3000) a été programmé pour prendre une mesure dans 5 points différents de chaque puits, chaque mesure étant en elle-même une moyenne de 8 scans. Au total, 10 spectres ont été collectés en 5 différents points du puits à raison de 2 répétitions par point. Le but est de juger de la variabilité des spectres au sein d'un puits.

#### **1.4.1.6. Répétabilité au sein de plusieurs puits de la plaque**

L'échantillon-témoin a été analysé dans trois puits. La répétabilité de la technique au sein de ces trois puits est valable pour toute la plaque. On disposait d'un total de 15 spectres à raison de 5 spectres par puits.

#### **1.4.1.7. Reproductibilité lors d'un processus de conservation**

Afin d'étudier la stabilité de l'extrait sec du lait dans le temps après le premier jour d'analyse, la plaque a été conservée à l'étuve à 22°C. L'échantillon a été analysé le 3e jour puis stocké à nouveau pour être soumis à une dernière collecte de spectres au 5e jour de conservation. Un total de 45 spectres à raison de 15 spectres par jour d'analyse a été enregistré.

#### **1.4.1.8. Reproductibilité entre deux plaques**

Dans le souci de gagner du temps, nous avons été tentés d'utiliser une seconde plaque. L'utilisation simultanée de deux plaques pourrait nous permettre d'économiser un mois d'analyse. Cette deuxième plaque a exactement les mêmes caractéristiques que la première. A ce niveau, il était indispensable de vérifier la variation de l'information spectrale de l'échantillon-témoin au sein de ces deux plaques jumelles, d'où juger de la reproductibilité de la méthode entre elles. Au total, 30 spectres ont permis d'établir ce paramètre à raison de 15 spectres collectés dans 3 puits de chacune des plaques.

#### **1.4.2. Analyse des échantillons**

La technique d'analyse optimisée sur le témoin a été appliquée pour analyser les 24 échantillons. Nous avons optimisé l'utilisation de la plaque à 96 puits en y étalant pas plus que 4 échantillons à raison de 3 puits par échantillon. Une distance d'environ 4 puits a été laissée entre les échantillons pour éviter toute possible interaction ou contamination croisée entre les échantillons. Les 3 catégories (entier, demi-écrémé et écrémé,) ont été représentées à chaque préparation de la plaque en vue de randomiser les analyses, soit de diminuer l'effet jour qui aurait pu influencer l'analyse spectrale de l'un ou l'autre des échantillons.

#### **1.4.3. Prétraitement des données**

Une fois les analyses achevées sur les 24 échantillons et les données spectrales disponibles via le logiciel OPUS/LAB 7.5 (Bruker, Allemagne), celles-ci ont été prétraitées. Cette étape qui précède le traitement permet de mettre en évidence l'information spectrale, laquelle pourrait avoir été perturbée par les variations dues par l'environnement d'analyse ou par les fluctuations du spectromètre. Pour ce faire, le logiciel "The Unscrambler X 10.3 (Camo, Norvège) " a été utilisé.

##### **1.4.3.1. Correction de la ligne de base**

Nous avons corrigé la ligne de base (baseline) des spectres collectés pour éviter les valeurs négatives observées dans les absorbances des données spectrales (dues au très faible volume d'échantillon analysé). En l'appliquant sur la matrice, les valeurs d'absorbances deviennent positives et les spectres qui en résultent sont remontés à partir de la ligne de base. C'est le seul prétraitement qu'ont subi les spectres lors de l'optimisation des paramètres de la technique.

#### **1.4.3.2. Déviation normale standardisée (SNV)**

Le SNV (Standard Normal Variate) a été appliqué pour corriger les effets des interférences spéculaires multiples et celles de la taille des particules sur les spectres des 24 échantillons d'analyse. Cette fonction standardise chaque spectre séparément et ne tient donc pas compte de la moyenne des spectres de sorte qu'un problème survenu dans un spectre ne peut être réparti sur l'ensemble des spectres.

#### **1.4.3.3. Dérivation**

L'application d'une dérivée sur un spectre ou une matrice a pour rôle de séparer les bandes qui se chevauchent. Une dérivée première a été appliquée sur nos spectres. Cette opération permet de mesurer la pente d'un spectre en chacun de ses points<sup>1</sup>.

#### **1.4.4. Traitement des données**

Les fonctions mentionnées ci-dessous ont été appliquées sur les données en guise de traitement. Le logiciel Excel 2010 a été utilisé pour établir les paramètres de répétabilité et de reproductibilité tandis que le logiciel Unscrambler X 10.3 a été utilisé pour établir l'ACP (Analyse en Composantes Principales).

##### **1.4.4.1. Etude de répétabilité et de reproductibilité par RMS et ANOVA**

La moyenne quadratique (RMS: Root Mean Square) et l'analyse de variance (ANOVA) sont les tests qui nous ont permis de mettre en évidence la répétabilité et la reproductibilité de la méthode. Le RMS compare les absorbances entre les spectres d'un même échantillon. Pour une série de 5 spectres, 10 RMS ont été obtenus en combinant les spectres deux à deux. L'ANOVA sur ces 10 répétitions nous a permis de voir si les spectres sont significativement différents entre eux en comparant la valeur de Fisher F à la valeur critique pour F et en considérant la probabilité qui est comparée au seuil de signification alpha (0,05), seuil en dessous duquel la répétabilité ou la reproductibilité n'est plus assurée (Urbano-Cuadrado *et al.*, 2004).

---

<sup>1</sup> Baeten V. (2016). Note de cours. Contrôle technologique de qualité: Chap.4. Université Catholique de Louvain. Chap 6.

#### **1.4.4.2. Discrimination des échantillons de lait par ACP**

L'Analyse en Composante Principale (ACP) montre la variation existant dans un jeu de données multivariées. Elle recombine les variables originales en un jeu de nouvelles variables appelées composantes principales. Celles-ci contiennent un maximum de l'information spectrale contenue dans les données originales<sup>2</sup>.

##### **1.4.4.2.1. Discrimination des échantillons de lait selon les différentes catégories**

Une ACP nous a permis de discriminer les différentes catégories (entier, demi-écrémé, écrémé) de lait. Pour chaque échantillon, nous avons moyenné les 5 spectres de chaque puits, obtenant au final 3 spectres par échantillon (sachant qu'on a 3 puits pour chacun d'eux).

##### **1.4.4.2.2. Discrimination des échantillons de lait en fonction du temps de conservation**

L'extrait sec a été mis à l'étuve (22°C) et analysé deux fois au cours de 5 jours. Une ACP a été appliquée sur la matrice prétraitée de chacune des 3 catégories de lait, les spectres n'ont pas été moyennés. Cette opération a été réalisée afin de vérifier toute variation dans l'échantillon au cours de la conservation et pouvoir décider si l'analyse du lait effectuée sur l'extrait sec peut s'étaler dans le temps en fonction d'une éventuelle variation des constituants.

#### **1.5. Etablissement d'un modèle de calibration pour la caractérisation du lait**

La régression PLS (Partial Least Squares regression) dite "méthode de régression basée sur les moindres carrés partiels" est ici utilisée en vue d'établir une équation de calibration pour la méthode développée. Cette méthode cherche à prédire une variable dépendante Y en utilisant des variables latentes ( $t_1, t_2, \dots$ ) qui sont des combinaisons linéaires des variables prédictives ( $x_1, x_2, \dots, x_p$ ) (Bertrand *et al.*, 2000).

Ne disposant pas de valeurs de références pour nos échantillons, nous les avons analysés sur un instrument (Bruker MPA) (sous leur forme liquide) puis, les avons prédits (Annexe 12) par les équations (plus bas sur la figure 17) extrêmement performantes de ce même instrument utilisé et décrit dans la seconde partie du travail. Ces valeurs prédites ont été utilisées comme valeurs de référence dans l'établissement du modèle de calibration.

---

<sup>2</sup> Baeten V. (2016). Note de cours. Contrôle technologique de qualité: Chap.4. Université Catholique de Louvain. Chap 6.

## **PARTIE II**

### **II. Prédiction des principaux constituants du lait en spectroscopie PIR**

#### **2.1. Objectif**

L'objectif de cette partie est de prédire les principaux constituants du lait (matière grasse, protéines et matière sèche) à l'aide d'un micro spectromètre PIR dit NIR4FARM, de type VIAVI, commercialisé par la société AB Vista pour l'analyse des fourrages en frais sur le terrain.

#### **2.2. Description des échantillons de lait**

Un total de 38 échantillons de lait a été analysé au cours de cette étude. Certains (Annexe 1) provenaient des supermarchés (Colruyt, Match et Red Market) et d'autres étaient issus de la ferme expérimentale du CRA-W: Département production et filière. Ces échantillons étaient répartis comme suit:

- 34 échantillons regroupés en 3 set ont été utilisés pour la calibration. 19 d'entre eux (9 entiers, 8 demi-écrémés et 2 écrémés) (1er set) sont des laits de commerce (Annexe 1) et 15 sont des laits de ferme frais. De ces derniers, 10 (2e set) étaient issus d'une gamme de calibration reconstituée (rétentat: concentré de protéines) et 5 (3e set) étaient enrichis en matière grasse (perméat). Les 34 échantillons de calibration ont été scannés chacun 4 fois avec le spectromètre du NIR4FARM couplé à la tablette.
- 4 échantillons ont servi à la validation. 3 sont des laits de commerce (1 entier, 1 écrémé et 1 demi-écrémé) et l'autre est du lait frais (cru) issu de la ferme. Ils ont été analysés plus de 55 fois chacun avec le spectromètre du NIR4FARM couplé à un ordinateur (sans la tablette).

#### **2.3. Matériels utilisés et mode opératoire**

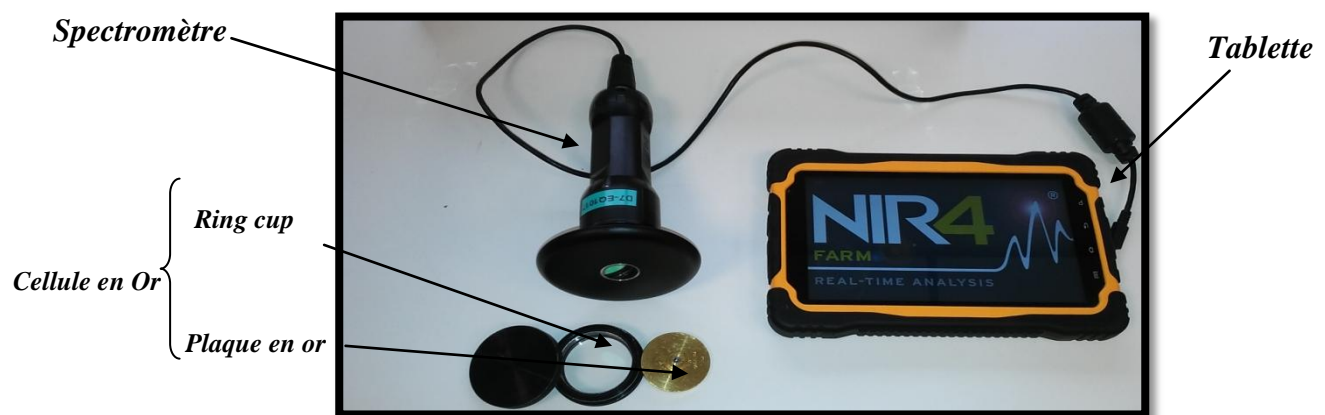
Deux spectromètres PIR ont été utilisés dans l'étude: le NIR4FARM qui utilise une partie de la gamme du PIR soit de 920 à 1658 nm et le Bruker MPA qui utilise toute la gamme du PIR allant de 850 à 2500 nm.

##### **2.3.1. Le NIR4FARM**

Le NIR4FARM (Figure 4) est le spectromètre que nous devons calibrer pour l'analyse du lait. Cet instrument est utilisé pour l'analyse des fourrages en frais sur le terrain. Ici, nous essayons de

voir s'il a du potentiel pour servir d'alternative moins dispendieuse pour l'agriculteur par rapport au prix qu'il consent régulièrement pour l'analyse du lait dans les laboratoires accrédités. Ainsi l'agriculteur disposerait de toute sa batterie d'analyse en un seul instrument. Le NIR4FARM pourrait être éventuellement utilisé pour analyser le lait online soit directement dans la salle de traite, ce qui permettrait d'établir un suivi en temps réel de la production.

Le NIR4FARM est un matériel portatif constitué d'un spectromètre connecté à une tablette numérique. Anciennement appelé JDSU, il est conçu par le groupe américain VIAVI et piloté par la société anglaise AB-Vista.



**Figure 4: Spectromètre NIR4FARM**

Un ensemble dit "cellule en or" est formé d'un ring cup, qui est une cellule de mesure des spectromètres de la marque Foss, pourvu d'une vitre en quartz et d'une plaque en or agissant comme réflecteur, entre lesquelles une goutte de lait d'un volume de 0,5 ml est déposée. Le NIR4FARM travaille ici en transréflexion.

La récupération des spectres n'est ni directe ni immédiate. En effet, après chaque série d'analyse, une demande est lancée via un e-mail à la société anglaise qui se charge d'envoyer les spectres dans les 2 ou 3 jours. Cependant, les données reçues n'ont pas été toujours complètes et certains spectres ont même parfois été aberrants (notamment le 6<sup>e</sup> spectre d'un même échantillon à chaque fois). Ce qui nous a contraints de coupler dans la suite de ce travail le spectromètre à un ordinateur du laboratoire CRA-W qui utilise un logiciel d'acquisition immédiate de spectres (JDSU MicroNIR\_1\_5\_7) et qui est connexe au logiciel utilisé pour le NIR4FARM.

### 2.3.2. Le spectromètre Bruker MPA

Le MPA est un spectromètre FT-NIR dont la zone de mesure est située entre 800 à 2500 nm.

Une pompe homogénéisatrice est reliée au spectromètre par un réseau de tubes transparents, de 0,8 mm<sup>2</sup> de section. Ces tubes assurent le remplissage et la vidange du support de l'aliquote qui est aspirée dans une cellule vitrée de 1 mm d'épaisseur et qui se trouve dans le module de mesure du spectromètre. La mesure est faite en transmission et en double. A chaque fois, un volume de lait est aspiré depuis le contenant de l'échantillon puis homogénéisé et chauffé à 40°C à même l'appareil (Figure 5) dont le nettoyage est automatique avec des solutions chimiques entre les échantillons.



Figure 5: Spectromètre PIR (Bruker MPA) équipé d'un homogénéisateur

#### 2.3.2.1. Transfert de la Base de données MPA vers NIR4FARM

La première idée était de transférer la base de données lait du MPA (279 échantillons) vers le NIR4FARM. Le transfert ou standardisation consiste à transposer la base de données du MPA prise dans la gamme entre 800 et 2500 nm vers le NIR4FARM utilisant la gamme de 920 à 1658 nm, ou en d'autres mots, à la simuler mesurée sur le NIR4FARM. Cette base de données du MPA (279 échantillons dont on dispose des valeurs de référence) servirait pour la calibration du NIR4FARM.

### **2.3.2.2. Prédiction de valeurs de référence pour la calibration du NIR4FARM**

Après l'échec du transfert dû à une trop grande différence de signal entre les 2 instruments, seuls de nouveaux échantillons ont été considérés pour la calibration du NIR4FARM. A ce niveau, le MPA a été utilisé pour prédire des valeurs pour les 19 échantillons (1er set de calibration) pour lesquels nous n'avions pas de valeurs labo. Ces valeurs de prédictions ont été utilisées comme valeurs de référence dans la calibration du NIR4FARM (Annexe 15). Les données du MPA ont été obtenues dans les conditions décrites ci-dessus. Enfin, les équations obtenues sur le NIR4FARM ont été comparées à celles du MPA que l'on sait déjà très performantes.

### **2.4. Méthodes de référence utilisées**

Les méthodes de référence utilisées au CRA-W pour nos constituants d'intérêt sont: la méthode de Rose Gottlieb pour la teneur en matière grasse totale, la méthode Kjeldahl pour les protéines et la conservation en étuve à 102°C pour la teneur en matière sèche totale.

### **2.5. Prétraitement des données**

Après la collecte de spectres sur les logiciels *OPUS* (pour le MPA) et *MicroNir.vi* (pour le VIAVI) le prétraitement des données a été finalisé sur WINISI et Unscrambler.

Afin de voir s'il y avait suffisamment d'information dans la gamme spectrale du NIR4FARM (920-1658nm), la base de données MPA a été trimée puis traitée par *SNV* (Standard Normal Variate) et dérivée première, opérations déjà explicitées dans la première partie du travail.

### **2.6. Traitement des données**

Des équations de calibration et de validation ont été obtenues pour la prédiction de valeurs sur la matière grasse totale (MG), les protéines (PRO) et la matière sèche totale (MS).

Ces équations ont été établies en MLR (multi Linear Regression) (Annexe 16) et en PLS (Partial least Square) (Annexe 17) pour la calibration de l'instrument (NIR4FARM) avec les 38 échantillons. La régression MLR utilise quelques longueurs d'onde tandis que la régression PLS utilise toutes les longueurs d'onde. Leur performance a été évaluée sur base de différents paramètres dont les équations sont figurées en annexe 4. Le modèle avec les meilleurs paramètres sera retenu pour la calibration.



### **2.6.1. L'erreur standard de calibration (SEC) et de cross-validation (SECV)**

La SEC (Standard Error of Calibration) (Annexe 4; Equation 3) exprime la somme des erreurs associées au laboratoire (instrument, environnement, etc.) et au proche infrarouge (PIR). Il doit être idéalement le plus petit possible et tendre vers zéro (Dardenne, 2000). La technique de la validation croisée est utilisée afin de déterminer le nombre optimal de facteurs PLS.

### **2.6.2. L'erreur standard de prédiction (SEP) et de cross validation (SECV)**

L'erreur standard prédiction (SEP) (Annexe 4; Equation 4.1) s'exprime par la racine carrée de la moyenne quadratique des écarts entre la valeur de référence et les valeurs prédites. Idéalement, elles doivent être le plus petit possible et du même ordre de grandeur. Le SECV est idéalement inférieur au SEP.

Le SEPC est l'erreur standard de prédiction corrigée du biais. Quand le biais est significatif, le SEPC est largement supérieur au SEP. On corrige du biais si le SEP est significativement différent du SEPC (Dardenne, 2000).

### **2.6.3. Le coefficient de détermination ( $R^2$ )**

Le coefficient de détermination ( $R^2$ ) (Annexe 4; Equation 5) décrit la corrélation entre les valeurs prédites et les valeurs de références. Il doit tendre vers 1 (Dardenne, 2001). La corrélation est parfaite si elle est égale à 1.

### **2.6.4. Le RPD ou Ratio of Prediction Deviation**

Le RPD (Annexe 4; Equation 6) est le rapport entre la SD et le SECV. Il doit être le plus grand que possible et idéalement supérieur à 3. S'il est inférieur à 1, les prédictions qui en résulteront seront inutilisables (Dardenne, 2000).

## RESULTATS ET DISCUSSION

Fidèle à la méthodologie, nous traiterons les résultats issus des deux parties du travail séparément. Nous les interpréterons et les discuterons successivement.

### PARTIE I

Dans cette première partie, nous restituons les résultats conformément à l'objectif fixé à savoir *"le développement d'une méthode de microscopie proche infrarouge pour l'analyse de microquantités de lait"*.

#### III. Développement de la méthode d'analyse du lait en microscopie PIR

Comme attendu, l'échantillon-témoin demi-écrémé qui a servi à l'optimisation des paramètres d'analyse a donné des spectres dont l'intensité d'absorbance et la signature spectrale sont intermédiaires entre celles des deux autres catégories (entier et écrémé). En effet, conformément aux informations de Jeantet *et al.* (2008) il contient théoriquement deux fois moins de matière grasse que le lait entier et plus de matière grasse que le lait écrémé qui peut aussi ne pas en contenir du tout. En travaillant sur la fraction de matière sèche, cette variation impacte sur l'intensité d'absorbance des spectres.

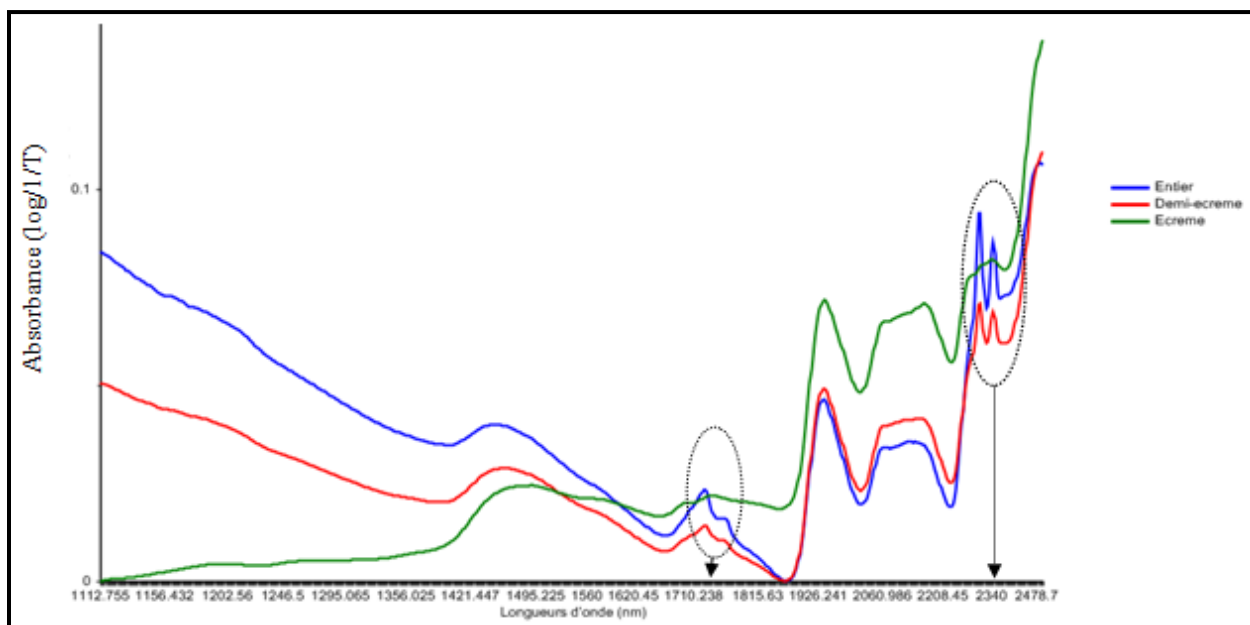


Figure 6: Spectres PIR des laits entiers, demi-écrémés et écrémés

Conformément aux déclarations de Bertrand *et al.*, (2000) (Annexe 13) nous pouvons observer sur la figure 6, dans les zones du PIR des longueurs d'onde de 2310 à 2350 nm et de 1700 à 1800 nm, l'empreinte spectrale caractéristique de la matière grasse. En effet il y a une similitude dans les spectres des laits entiers (bleu) et demi-écrémé (rouge) qu'on ne retrouve pas dans le spectre des laits écrémés (vert). Ce qui confirme l'absence de la matière grasse dans ces derniers.

### 3.1. Optimisation des paramètres d'analyse des échantillons

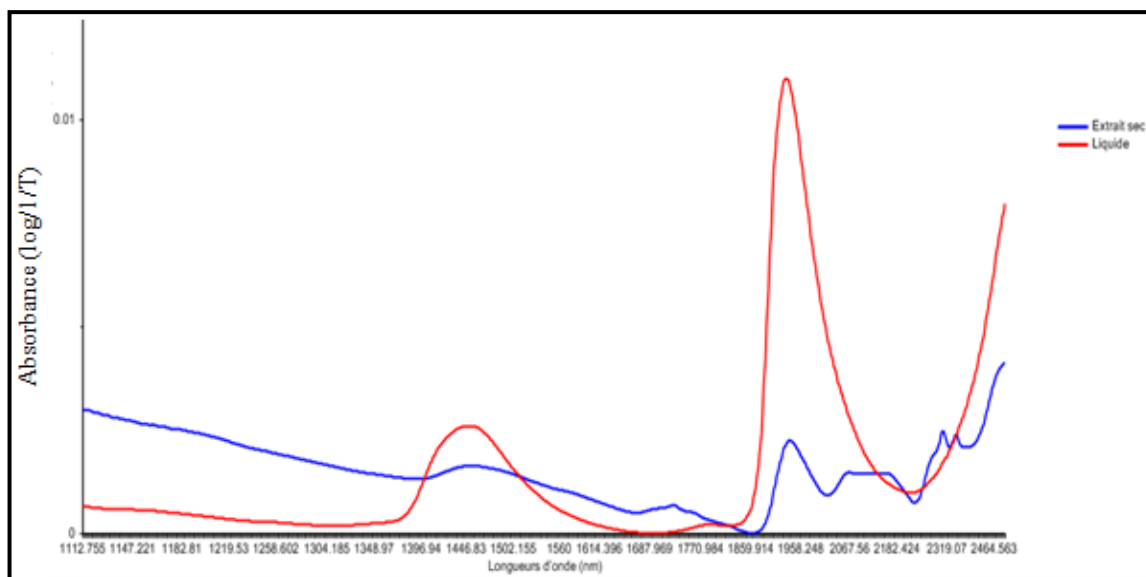
Les différents paramètres testés sont résumés dans le tableau 1. Les résultats de répétabilité et de reproductibilité ont été mis en évidence en comparant la valeur de signification (Probabilité) au seuil de signification (alpha: 0,05) révélés par le test ANOVA sur les RMS des spectres. Alpha étant la limite à ne pas dépasser pour que la répétabilité de la technique soit assurée.

**Tableau 1: Les différents paramètres optimisés**

PARAMETRES D'OPTIMISATION	PARAMETRES TESTES	CHOIX
1. Collecte de spectres	Lait liquide Vs Extrait sec	Extrait sec
2. Volume de remplissage d'un puits	3 vs 5 vs 7 µl	5 µl
3. Temps de séchage d'une goutte	11 vs 15; 17; 23; 26; 30; 40; 60; 90; vs 120 min	23 min
4. Choix du nombre de gouttes au sein d'un puits	1 vs 2; 3; 4; 5; 6 vs 7 Gouttes	5 Gouttes
	<b>Probabilité vs Seuil alpha (0,05)</b>	<b>Différence significative</b>
5. Répétabilité au sein d'un puits d'une plaque	0,98 > 0,05	Non
6. Répétabilité au sein de 3 puits d'une plaque	0,54 > 0,05	Non
7. Variation des spectres dans le temps	0,0008 < 0,05	Oui
8. Reproductibilité entre deux plaques	0,28 > 0,05	Non

#### 3.1.1. Collecte de spectres de l'extrait sec du lait

Lors de la comparaison du spectre du lait liquide et du lait séché, nous avons constaté que l'eau que contient le lait liquide absorbe la majeure partie de l'énergie (Figure 7), ce qui est conforme aux affirmations de Bertrand *et al.*, (2000). Pour minimiser les effets des interférences de l'eau, nous avons choisi de développer la méthode sur l'extrait sec du lait. En effet, comme le montre la figure 7, l'empreinte spectrale pris sur extrait est plus nette que lorsqu'elle est prise sur le lait liquide. Dans le cas de l'extrait sec, il est possible de distinguer les différentes bandes composant le spectre du lait et donc il est désormais possible de les attribuer aux constituants du lait.



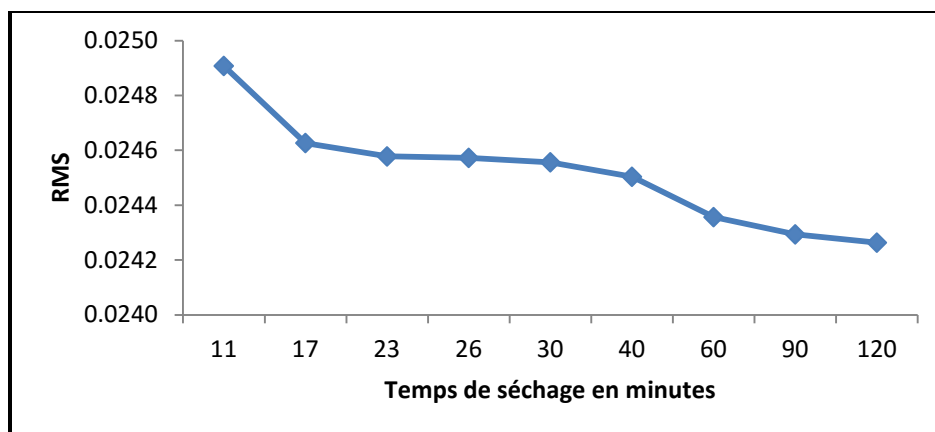
**Figure 7: Spectres du lait demi-écrémé: Lait liquide (*Normalisation - rouge*) Vs Extrait sec du lait (*Baseline + Normalisation-bleu*)**

### 3.1.2. Choix du volume de remplissage d'un puits de la plaque

Parmi les différents volumes (3; 5 et 7  $\mu$ l) de lait déposés au sein d'un puits de la plaque, celui de 5  $\mu$ l a été jugé optimal pour remplir la surface d'un puits car il n'y avait pas de débordement sur le contour ni n'était insuffisant tel était le cas pour respectivement les volumes de 7 et 3  $\mu$ l.

### 3.1.3. Définition du temps de séchage d'une goutte

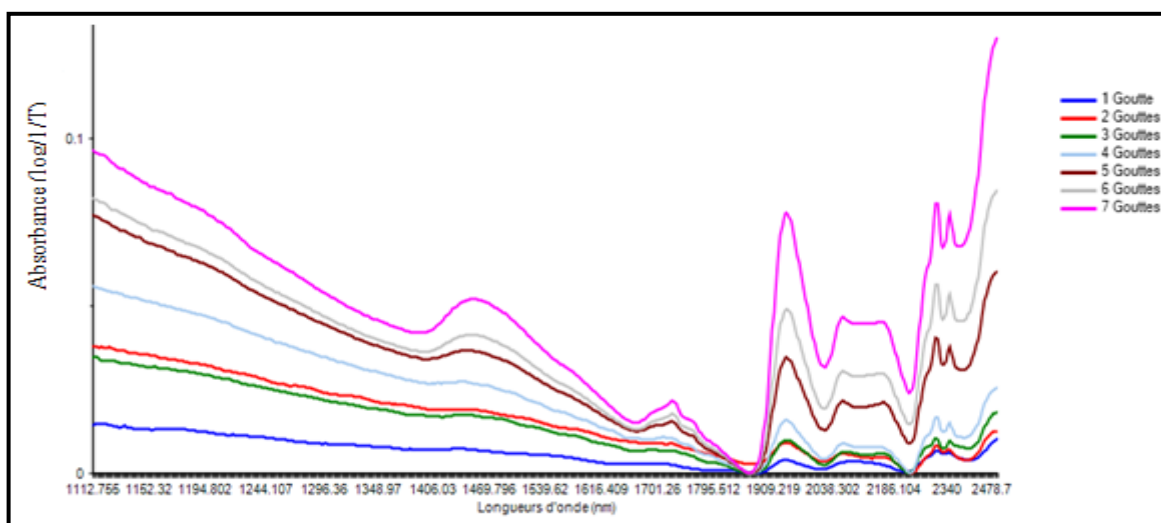
Visuellement une goutte prend en moyenne 23 minutes pour sécher à la température de la pièce ( $21.5 \pm 1^{\circ}\text{C}$ ). Pour valider ce constat, des spectres des gouttes pris respectivement entre 0; 11; 17; 23; 26; 30; 40; 60; 90 et 120 minutes ont été collectés. Un RMS a été calculé pour chaque temps de séchage comme évoqué plus haut mais en soustrayant, dans ce cas, le spectre moyen au spectre de chaque point mesuré. Au regard de leur RMS respectif, une stabilité dans les données spectrales commence à s'établir à partir de 17 minutes pour varier à partir de 40 minutes de séchage (Figure 8). En guise de compromis entre le temps d'analyse et un séchage complet, les 23 minutes ont été retenues comme durée de séchage optimale.



**Figure 8: Variation de l'information spectrale sous l'effet de la durée de séchage**

### 3.1.4. Choix du nombre de gouttes au sein d'un puits

Des analyses réalisées séparément sur un puits contenant entre une à sept gouttes séchées cumulées ont montré que les spectres avec les moins de gouttes séchées sont non seulement plus bruyants mais aussi présentent de plus faibles intensités que ceux avec un nombre de gouttes supérieures. L'addition de 5 gouttes cumulées a été sélectionnée pour la suite des analyses, ce qui ramène à un volume de 25  $\mu$ l par puits. L'intensité d'absorbance était pourtant plus élevée avec 6 ou 7 gouttes (Figure 9).



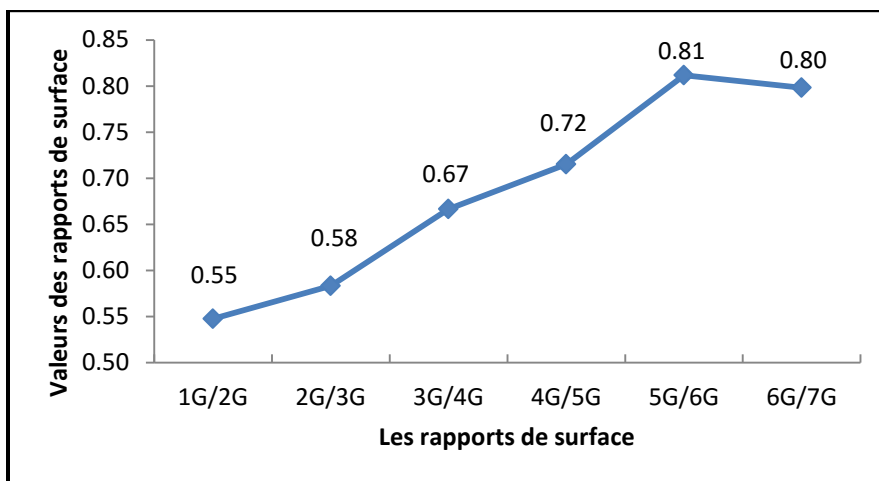
**Figure 9: Spectres selon le nombre de gouttes séchées cumulées (correction de la ligne de base)**

Plusieurs raisons ont motivé le choix de 5 gouttes:

a) A partir de la 6e goutte ajoutée sur les 5 déjà séchées, la surface de l'extrait sec cumulé commence à devenir convexe, ce qui induit une non-uniformité de sa répartition sur la surface du puits. Dans une telle situation, les mesures prises proche de la circonférence du puits risquent de donner des spectres avec des intensités d'absorbance nettement plus faibles que ceux pris au beau milieu du puits.

b) 5 gouttes mettent moins de temps pour sécher (115 minutes) que 6 ou 7 gouttes (138 ou 161 minutes).

c) L'ajout d'une goutte (G) supplémentaire sur l'extrait sec augmente le rapport des surfaces des spectres (surface de la bande entre 2385 et 2234 nm) d'une manière croissante jusqu'à un certain nombre de gouttes. En effet, ce rapport tend à se stabiliser autour des cumuls 5G/6G et 6G/7G tel que l'aire d'un spectre de 5 gouttes vaut 0,81 fois celle de 6 gouttes qui vaut 0,80 fois celle de 7 gouttes (Figure 10).



**Figure 10: Rapport de surface de la bande spectrale allant de 2234 à 2385 nm selon le nombre de gouttes séchées cumulées**

### 3.1.5. Répétabilité au sein d'un puits de la plaque

Le test ANOVA réalisé sur les RMS (Annexe 6) des spectres des 5 points, analysés en double au sein d'un puits affiche une valeur de F de 0,08 inférieure à la valeur du seuil critique 2,3 défini pour F. On peut dire qu'il n'y a pas une différence significative entre les spectres. De plus la

Probabilité de 0,98 qui est supérieur à 0,05 (alpha) confirme cette non-significativité. On en déduit que la technique est répétable au sein d'un puits de la plaque.

### **3.1.6. Répétabilité au sein de plusieurs puits de la plaque**

L'analyse de variance appliquée sur les RMS des spectres collectés au sein des 3 puits accuse une valeur F de 0,63 inférieure au seuil critique de 3,35 définis pour F (Annexe 7). On en déduit qu'il n'y a pas une différence significative entre les données spectrales. De plus la Probabilité de 0,54 qui est supérieure à 0,05 (alpha) vient renforcer cette non-significativité. On en déduit que la méthode est répétable entre les 3 puits et donc au sein de toute la plaque.

### **3.1.7. Reproductibilité lors d'un processus de conservation**

Les 3 analyses effectuées sur l'échantillon-témoin ont eu lieu le jour du dépôt, le 3e et le 5e jour après conservation à l'étuve à 22°C. Le test ANOVA sur les RMS des spectres (Annexe 8) montrent une F de 7,69 supérieure au F critique qui est de 3,10. Il y a donc une différence significative entre les spectres. De plus la Probabilité de 0,0008, inférieure au seuil de signification alpha qui est 0,05 vient confirmer ce constat. On a procédé à un "test post-hoc" afin de savoir laquelle des comparaisons de groupes, pris deux à deux, est significativement différente. On obtient respectivement des Probabilités de 0,0009 et 0,001 pour les comparaisons entre les 1er et 3e (Annexe 9) jours puis entre les 3e et 5e jour (Annexe 10). Comme ces valeurs sont inférieures ou égales à 0,001, on peut dire qu'il existe une différence hautement significative entre les spectres collectés selon les différents jours d'analyse. Cela s'explique par le fait que le produit évolue au cours de la conservation et perd notamment de l'eau.

### **3.1.8. Reproductibilité entre les deux plaques**

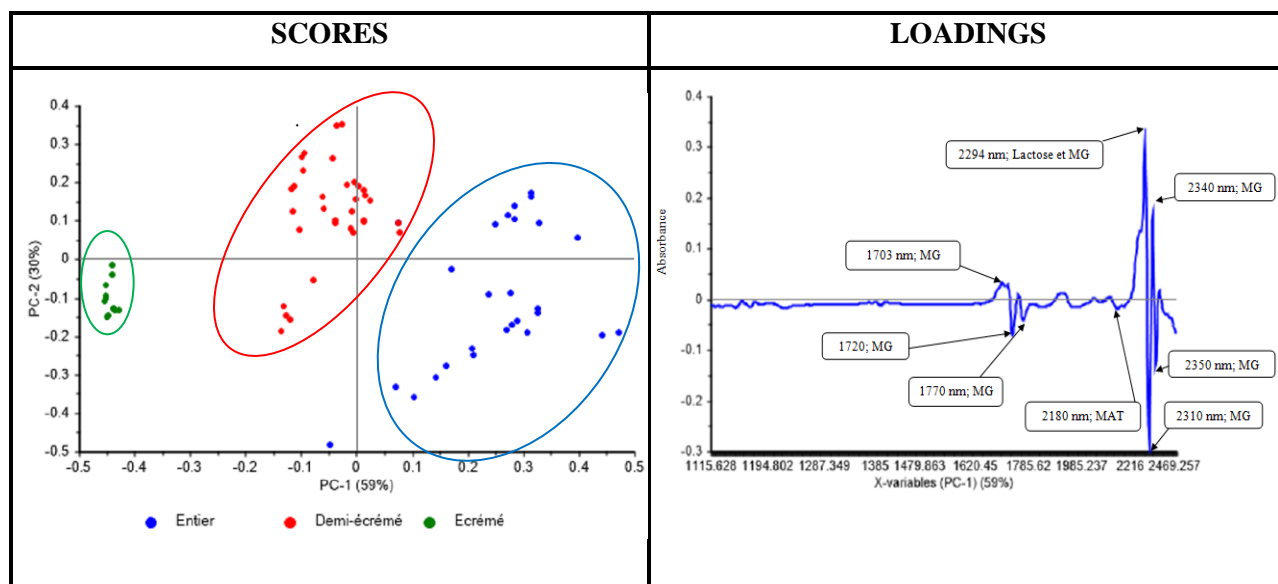
L'ANOVA effectuée sur les RMS des spectres (annexe 11) exprime un F de 1,17 inférieur au seuil critique de 4,01 défini pour F. Il n'y a donc pas de différence entre les spectres collectés sur les deux plaques. De plus, la Probabilité de 0,28 est supérieure au seuil de signification alpha de 0,05. Sur la base de ces résultats, nous pouvons conclure que la méthode est reproductible entre les deux plaques. Leur utilisation simultanée nous a permis d'économiser 1 mois de temps d'analyse.

### 3.2. Discrimination des échantillons de lait

Après avoir optimisé les paramètres sur l'échantillon témoin, la technique a été reproduite pour l'analyse sur les 24 échantillons. Après collecte des spectres et leur prétraitement, une ACP nous a permis de faire ressortir d'une part, la discrimination entre les différentes catégories de lait et d'autre part la variation à l'intérieur de chaque catégorie par rapport à la durée de conservation.

#### 3.2.1. Discrimination des échantillons de lait suivant les différentes catégories

La première et la deuxième composantes de l'ACP entre les laits entiers (EN), demi-écrémés (DE) et écrémés (EC) révèle une nette discrimination entre ces 3 catégories (Figure 11).



**Figure 11: Résultats de l'ACP (PC-1 (59%) vs PC-2(30%)) appliquée sur les spectres NIR des 3 catégories de laits (EN, DE, EC), traités par SNV, Baseline, 1ère dérivée et Moyenne**

Les laits entiers se sont repartis principalement dans la partie droite des scores de la première composante correspondant à la combinaison de ses coordonnées positives. A l'inverse, les laits EC se sont agglomérés exclusivement dans le quadrant inférieur gauche correspondant à la combinaison des coordonnées positives de la première et les coordonnées négatives de la deuxième composante bien qu'une partie s'étale dans la partie positive de cette dernière. Les laits DE sont intermédiaires et occupent essentiellement le quadrant supérieur gauche faisant référence à la partie positive de la deuxième composante avec quelques échantillons qui se positionnent dans la partie négative de cette dernière.

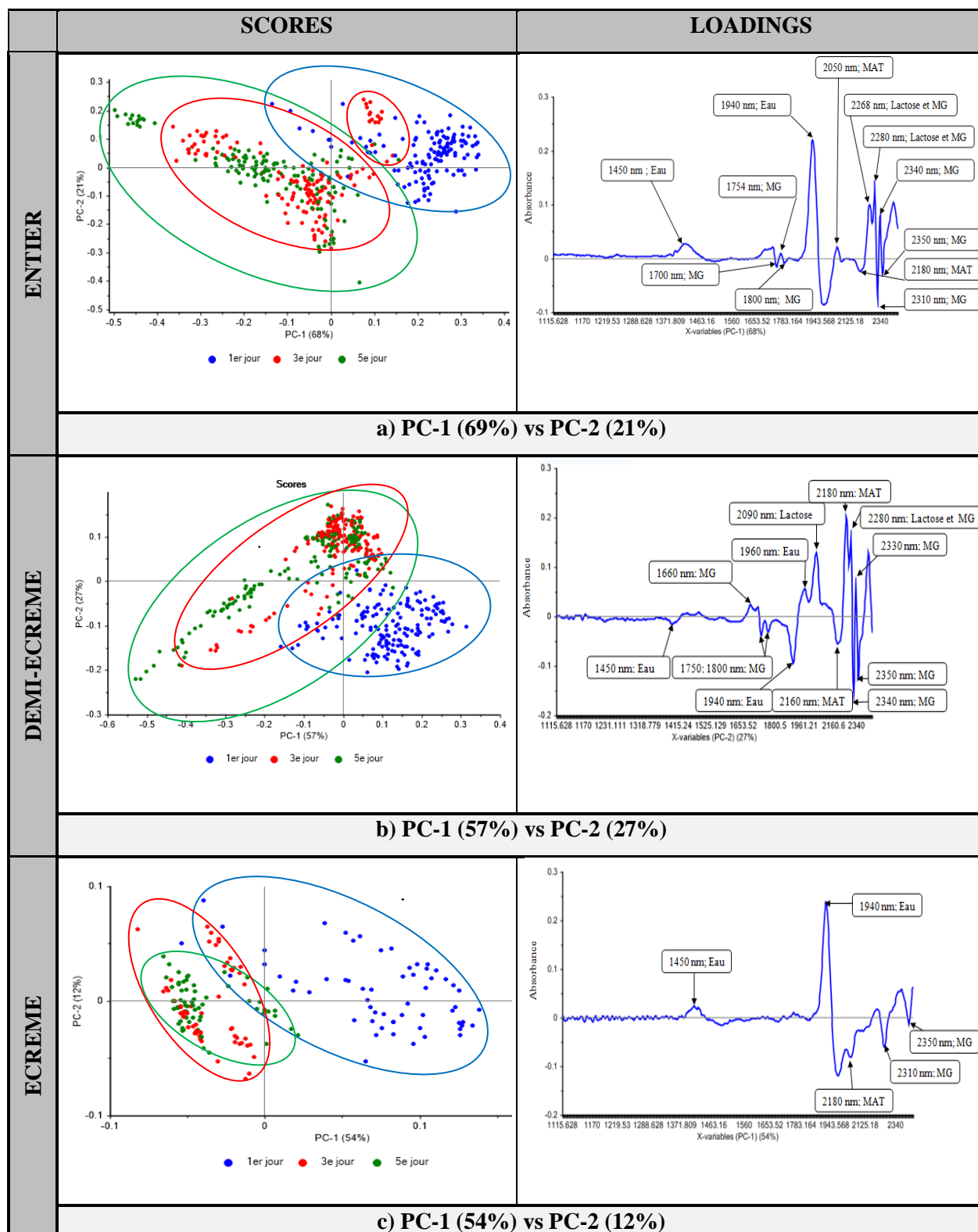


Le tableau (Annexe 13) élaboré en 2000 par Bertrand *et al.*, atteste que les bandes identifiées dans les loadings correspondent à la principale composante discriminative des 3 catégories de lait qui est la matière grasse (MG). En effet, la partie positive de la composante 1 est dominée par environ 3 pics dont les longueurs d'onde (1703; 2294 et 2340 nm) sont identifiées comme étant toutes des zones spectrales de vibration des molécules C-H correspondant à la présence de MG. La partie négative de la composante 1, dominée par les laits EC, comporte essentiellement 4 pics dont les longueurs d'onde (1720; 1770; 2310 et 2350 nm) se réfèrent aux liaisons C-H correspondant à de la MG. Enfin, la position des laits DE (partie positive de la première composante et négative de la seconde) implique que ceux-ci contiennent moins de MG que les laits EN (partie droite) et les laits EC (partie inférieure gauche). Au final, on en déduit par les spectres, les RMS, et la PCA que la principale variable discriminative entre les 3 catégories de lait est bel et bien la MG, ce qui est en bon accord avec la littérature (Bertrand *et al.*, (2000).

### **3.2.2. Discrimination des échantillons de lait en fonction du temps de conservation**

Les ACP réalisées sur les différentes catégories de lait ayant subi la conservation ont démontré que l'eau et la matière grasse constituent les principales variables discriminatives. Pour les catégories EN et EC, les spectres des laits pris au 1<sup>er</sup> jour d'analyse (points bleus) sont essentiellement répartis dans la partie droite des scores. Cette partie correspond à la combinaison des coordonnées positives pour la première composante. Ces points ont tendance à se dissocier du lot tandis que les points des spectres pris au 3<sup>e</sup> (points rouges) et au 5<sup>e</sup> (points verts) jour s'agglomèrent en se situant majoritairement dans la partie gauche; combinaison négative pour la première composante. Cette agglomération montre que l'évolution de la composition des échantillons de lait se serait stabilisée à partir du 3<sup>e</sup> jour de conservation.

La nature des constituants qui évoluent dans l'extrait sec du lait pendant les premiers jours de conservation est révélée par les loadings (Figure 12). Pour la partie positive de la composante 1, les laits EN et EC partagent 2 pics (1450 et 1940 nm) se référant à l'eau. Ce qui implique une évaporation progressive de l'eau dans l'échantillon au jour d'analyse par rapport au 3<sup>e</sup> et 5<sup>e</sup> jour de conservation. D'autres pics se référant simultanément au lactose et à la MG (2268 et 2280 nm) ou à la MG seule (1754; 2340 nm) expriment une différence dans les laits entiers pendant la conservation. Par rapport à la partie négative de la composante 1, les laits EN et EC partagent 3 pics caractérisant la MG (2310; 2350 nm) et les matières azotées totales (MAT- 2180 nm).



**Figure 12: Résultats de l'ACP appliquée sur les spectres NIR de chaque catégorie de lait (EN, DE, EC) collectés en 3 analyses (1<sup>er</sup>, 3<sup>e</sup> et 5<sup>e</sup> jour) et traités par SNV, baseline et 1<sup>ère</sup> dérivée**

Pour le lait DE, on se limite à l'observation de la 2e composante. La disposition des points ne favorisent pas une bonne discrimination. En effet, il est difficile de séparer les spectres suivant l'effet jour car il y a chevauchement entre les spectres du 1er, du 2e et du 3e jour le long de la 1ère composante (loading associé présenté sur l'annexe 14). Par contre, selon la 2e, malgré qu'on arrive à séparer les spectres des 1er jours dans la partie négative et ceux des 3e et 5e jours dans la partie positive, un certain nombre d'échantillons de ce dernier se retrouve aussi dans la partie négative de la PC2 (Figure 12, loadings DE). L'examen du loading correspondant à PC2 ne permet pas d'attribuer certaines des longueurs d'ondes discriminatives à chaque jour d'analyse.

### 3.3. Etablissement d'un modèle calibration pour la caractérisation du lait

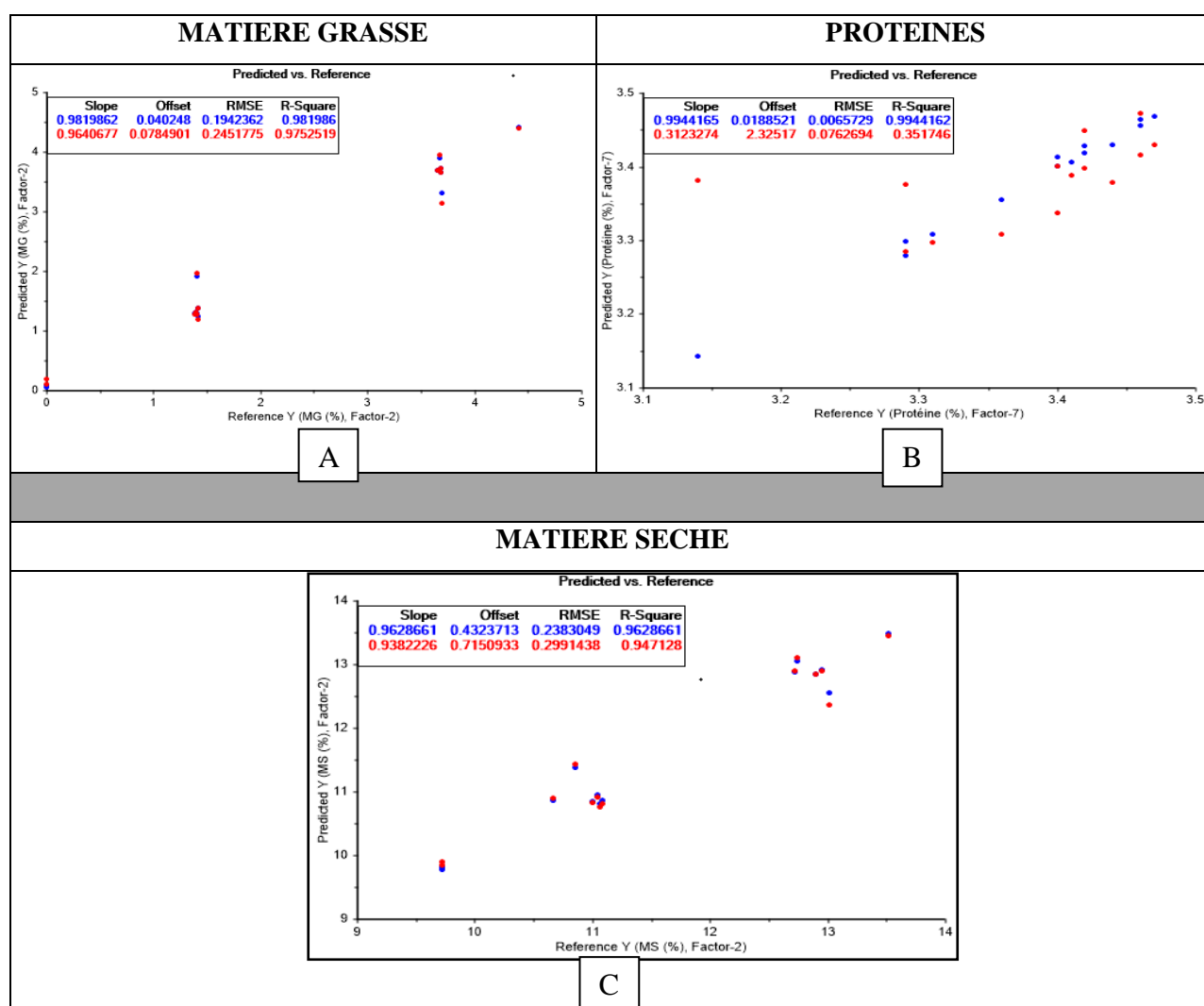
Habituellement les équations de calibration des constituants du lait sont réalisées avec des spectres de laits sous leur forme liquide. Nous nous sommes intéressés à investiguer les performances de ce type d'équation, une fois réalisée, en utilisant des spectres d'extraits secs. Ce qui pourrait être très utile si la quantité de lait dont nous disposons est très faible et que nous soyons obligés de réaliser les analyses sur les extraits secs. Suite donc au développement de la méthode, le modèle mathématique obtenue en régression PLS pour la calibration s'est révélé très performante avec des coefficients de détermination ( $R^2$  ou R-square) respectifs de 0,98; 0,99 et 0,96 pour la matière grasse, les protéines et la matière sèche (Figure 13). Ce paramètre compris entre 0 et 1 doit tendre vers 1. Il existe donc une bonne corrélation entre les données. Au regard des erreurs standards de calibration (SEC) proches des erreurs standards de validation croisée SECV, les équations sont performantes. Par contre, pour les protéines, l'équation accuse un faible coefficient de détermination ( $R^2$ ) de 0.31 (Tableau 2).

**Tableau 2: Quelques paramètres de l'équation obtenue pour la méthode**

Paramètres	Calibration		Validation croisée	
	R2	SEC	R2	SECV
<b>MG</b>	0,98	0,19	0,97	0,24
<b>PRO</b>	0,99	0,01	0,35	0,07
<b>MS</b>	0,96	0,24	0,95	0,29

En considérant ce problème lié exclusivement aux protéines (figure 13, B), nous pouvons affirmer que la non performance de l'équation est liée à un manque de variabilité dans les

données. En effet nous disposons de 3 catégories de lait dont la principale différence est la MG. Ce constituant varie de 0 à 4,41 dans les valeurs de référence et proportionnellement la matière sèche varie avec elle pour des valeurs s'étendant de 9,71 à 13,01. Par contre, les protéines étant les constituants dont les quantités n'ont pas été modifiées dans les échantillons, accusent une faible intervalle de proportion allant de 3,31 à 3,47. La variabilité entre les données de référence est trop faible pour établir une équation performante pour cette variable. Aussi ne peut-on pas déclarer à priori que la méthode ne fonctionnera pas pour cette variable avant de disposer des valeurs de référence contenant une large gamme de protéines. La figure 13 présente les corrélation entre les valeurs prédites et les analyses de référence pour les trois constituants.



**Figure 13: Corrélation entre les valeurs prédites et les analyses de référence pour la matière grasse (A), les protéines (B) et la matière sèche (C)**

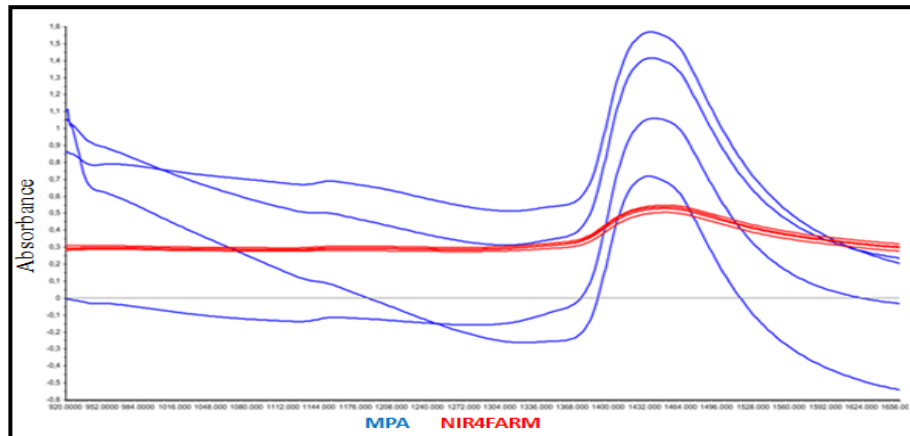
## PARTIE II

### IV. Prédiction de valeurs sur le lait en spectroscopie PIR

Rappelons que l'objectif de cette partie est de *prédire les principaux constituants du lait* à l'aide d'un micro spectromètre NIR de type VIAVI, commercialisé par la société AB Vista sous la dénomination NIR4FARM et conçu pour l'analyse des fourrages en frais sur le terrain.

#### 4.1. Transfert de la base de données MPA vers le NIR4FARM

Le transfert de la base de données du Bruker MPA vers le NIR4FARM n'est pas une réussite. En effet, les spectres entre les deux instruments sont très différents. En les superposant, ceux du NIR4FARM apparaissent très aplatis (en rouge) et ceux du MPA présentent de fortes pentes (en bleu). En témoigne la figure 14 présentant les allures d'un lait entier, 1 demi-écrémé, 1 écrémé et 1 cru sur les deux instruments.



**Figure 14:** Spectres d'une même série d'échantillons de lait (1 entier, 1 demi-écrémé, 1 écrémé et 1 cru) collectés sur le MPA (bleu) et sur le NIR4FARM (rouge)

Quatre principales raisons liées à la différence entre les deux instruments peuvent être à l'origine de l'échec de ce transfert. Elles sont mentionnées dans le tableau 3.

**Tableau 3: Différence entre le MPA et le NIR4FARM**

Paramètres de différences	MPA	NIR4FARM
1. Mode de mesure NIR	Transmission	Transréflexion
2. Température d'analyse du lait	40°C	T° ambiante (21.5°C)
3. Traitement au sein de l'appareil	Homogénéisation	Pas d'homogénéisation

## 4.2. Calibration avec les spectres du NIR4FARM

Après l'échec du transfert, il n'était plus possible d'utiliser les valeurs de référence des échantillons de la base de données du MPA dans la calibration de l'autre instrument. L'une des alternatives était d'établir des calibrations avec uniquement les spectres que l'on possède du NIR4FARM et de les comparer ensuite avec les performances de l'équation Bruker MPA.

### 4.2.1. Equations obtenues en régression MLR et PLS pour le NIR4FARM

Les équations développées pour le NIR4FARM sont plus performantes en régression MLR (Annexe 16) qu'en PLS (Annexe 17). La principale raison est que, pour développer une équation pour une variable donnée (protéines, matière grasse ou matière sèche), le MLR utilise des zones spécifiques du spectre tandis que le PLS utilise toutes les longueurs d'onde. Des deux côtés, les équations (Tableau 4) sont très performantes pour la matière grasse et la matière sèche avec une corrélation quasi parfaite pour les deux variables ( $R^2 = 0,99$  en PLS et en MLR) et un RPD (Ratio of Prediction Deviate) dépassant 3. Par contre, la principale différence réside dans les protéines. En effet, nous enregistrons en PLS des valeurs aberrantes tels un coefficient de détermination nul et un RPD (0,04) largement inférieur à 3. En MLR, les valeurs sont meilleures mais les équations ne sont pas performantes pour cette variable.

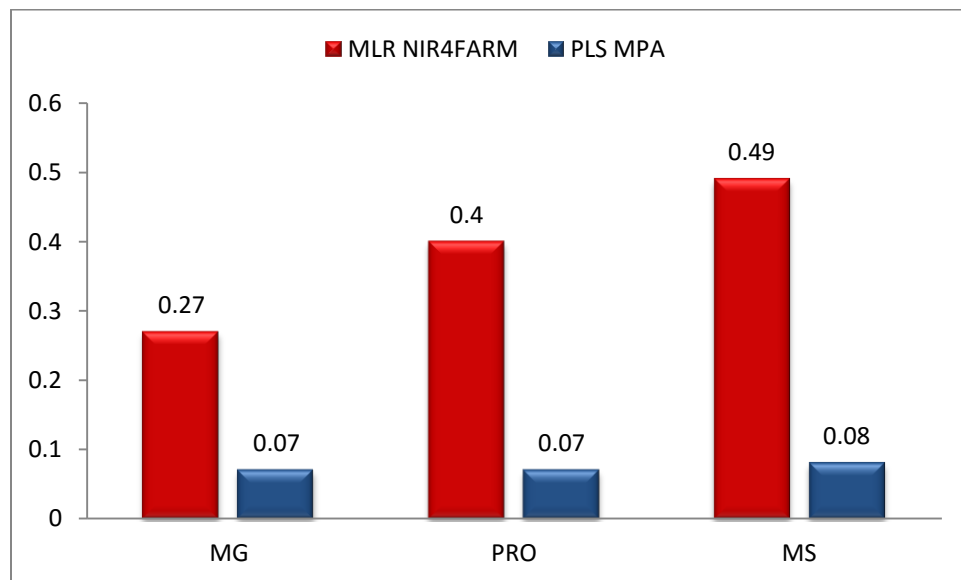
L'une des raisons de la "non performance" des équations obtenues pour les protéines est la même explicitée dans la première partie du travail, à savoir, il n'existe pas assez de variabilité dans les valeurs de référence dont on dispose pour cette variable.

**Tableau 4: Equations en régressions MLR vs PLS pour le NIR4FARM**

Variables	PLS Toutes les longueurs d'onde			MLR Quelques longueurs d'onde		
	SEP doit tendre vers 0	$R^2$ Entre [0,1] et tendre vers 1	RPD doit être > 3	SEP doit tendre vers 0	$R^2$ Entre [0,1] et tendre vers 1	RPD doit être > 3
MG	0,79	0,99	2,55	0,27	0,99	7,33
PRO	5,70	0,00	0,04	0,40	0,10	1,12
MS	6,04	0,98	0,32	0,49	0,99	4,18

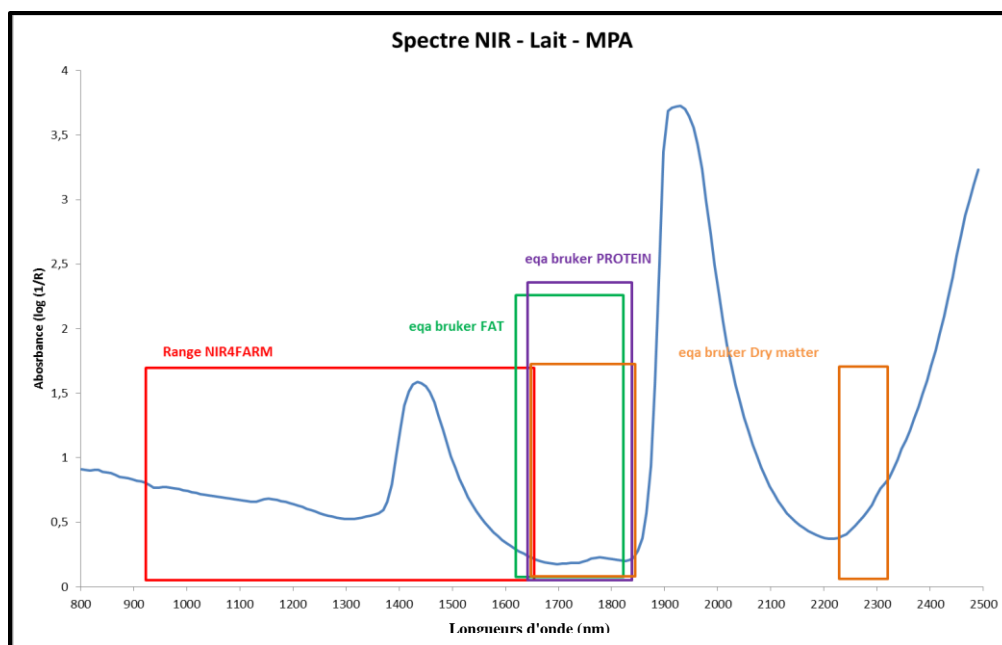
#### 4.2.2. Comparaison des équations du NIR4FARM avec les équations du Bruker MPA

L'erreur standard de prédiction (SEP) est ici utilisée pour comparer l'équation du NIR4FARM à celle du MPA. La figure 15 compare les SEP obtenus pour le NIR4FARM en MLR et ceux obtenus pour le MPA en PLS. Rappelons que le paramètre SEP doit idéalement tendre vers zéro. Au regard de la figure 15, bien que les valeurs de SEP du NIR4FARM sont satisfaisantes, elles sont moins performantes que celles du MPA. Le MPA accuse des SEP extrêmement faibles allant de 0,06; 0,04 et 0,08 respectivement pour la matière grasse, les protéines et la matière sèche, ce qui justifie la performance de ces équations. Pour ces mêmes variables, le NIR4FARM a respectivement 0,27; 0,40 et 0,49. Les autres paramètres sont restitués en annexe 18.



**Figure 15: Comparaison des Erreurs Standard de Prédiction (SEP) obtenues pour le NIR4FARM en MLR et le BRUKER MPA en PLS**

La différence entre les équations obtenues pour le NIR4FARM par rapport et celles obtenues pour le MPA se justifie par la figure 16.



**Figure 16: Spectre du lait sur le MPA: Indication du range utilisé par le NIR4FARM (rectangle rouge) et zones utilisées pour le développement des équations pour les variables matière grasse (Fat) en vert, protéines en mauve et matière sèche (Dry matter) en marron**

Sur la figure 16, nous pouvons observer que le NIR4FARM (rectangle rouge) utilise seulement une partie (920-1658 nm) de la gamme (range) du MPA (800-2500 nm). Nous constatons aussi que les zones les plus spécifiques avec lesquelles les performantes équations du MPA ont été établies ne se retrouvent quasiment pas dans les zones du NIR4FARM. Elles se situent entre 1615 et 1820 nm pour la matière grasse (vert), entre 1640 et 1850 nm pour les protéines (mauve) et enfin entre 2225 et 2325 nm puis entre pour la matière sèche (marron).

On peut faire deux déductions à partir des différences observées dans les équations. La première est que les conditions de mesure au MPA (Tableau 3 plus haut), très différentes de celles du NIR4FARM interfèrent beaucoup dans l'information spectrale. La seconde est qu'il manque d'information spectrale dans la zone utilisée par le NIR4FARM pour une bonne calibration pour le lait.

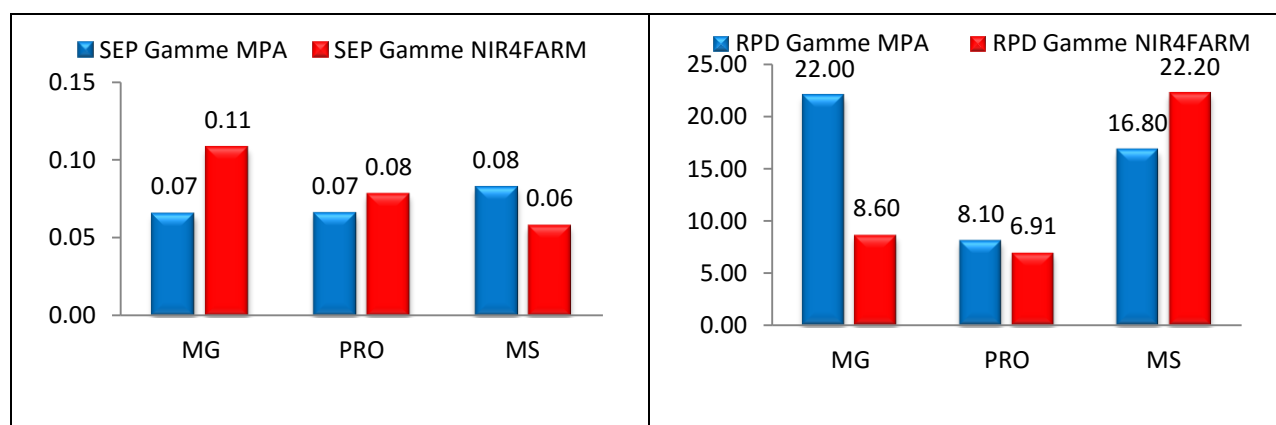
Pour vérifier la première hypothèse, il faudrait que l'on essaie de mettre le NIR4FARM dans les mêmes conditions de mesure que le MPA, à savoir collecter les spectres sur du lait chauffé à 40°C et homogénéisé.



La vérification de la seconde hypothèse se fera en comparant les équations obtenues avec la gamme complète du MPA (920-1658 nm) avec les équations obtenues avec la gamme du NIR4FARM (800-2500 nm). Ainsi, si celles-ci s'avèrent performantes, nous pourrions conclure que la gamme du NIR4FARM contient assez d'informations spectrales pour établir une bonne équation. Dans ce cas, l'hypothèse qu'il manque d'informations ne sera pas soupçonnée d'être à l'origine de la "non performance" des équations obtenues pour le NIR4FARM car la gamme 920-1658 aura suffi à elle seule pour le développement de bonnes équations.

#### 4.2.3. Comparaison de la performance des équations du MPA obtenues avec la gamme complète du MPA et avec la gamme du NIR4FARM

Deux paramètres ont été choisis pour comparer les deux équations: le SEP et le RPD (figure 17). Idéalement les SEP sont correctes en variant de 0,06 à 0,11 pour toutes les variables. De même, les RPD sont largement supérieur à 3 en variant de 6,91 à 22,20.



**Figure 17: Comparaison des équations de prédiction du MPA dans la gamme allant de 800 à 2500 nm (gamme MPA) avec celle allant de 920 à 1658 nm (gamme NIR4FARM)**

Aux regards de la figure 17, il ressort que les différences ne paraissent pas significatives en fonction que l'on optimise en utilisant l'entièreté de la gamme spectrale (800-2500 nm) ou seulement la gamme spectrale correspondant au NIR4FARM (920-1658 nm). Ceci tend à prouver que théoriquement l'information spectrale est bien là pour le NIR4FARM et que le problème viendrait plutôt des conditions de mesure et de la présentation de l'échantillon.

L'analyse du lait sur le NIR4FARM dans les mêmes conditions que chez le MPA reste une piste intéressante à explorer.

## CONCLUSION ET PERSPECTIVES

Dans ce travail, nous avons essayé de développer deux méthodes d'analyse rapide pour le lait: la première basée sur la microscopie PIR et la seconde basée sur la spectroscopie PIR.

La première a été développée sur une gamme de 24 échantillons de laits entiers, demi-écrémés et écrémés. Compte tenu de la rareté des échantillons de laits d'intérêt scientifique de certains mammifères comme le porc, nous avons développé la méthode sur des microquantités (25µl) de lait en utilisant deux plaques jumelles, en silice et comptant 96 puits hydrophobes. Les spectres du lait ont été collectés sur des extraits secs cumulés et ceci dans le but de minimiser non seulement les interférences dues à l'eau mais aussi d'augmenter l'intensité d'absorbance de ces spectres. La technique a été optimisée sur un lait demi-écrémé par l'établissement de différents paramètres d'analyses dont le volume idéal d'une goutte au sein d'un puits d'une plaque (5 µl), le temps de séchage optimal d'une goutte (23 min) et le nombre de gouttes séchées cumulées (5 gouttes). Les paramètres répétabilité au sein d'un puits, la répétabilité au sein de plusieurs puits, la reproductibilité entre les deux plaques et la reproductibilité au cours de la conservation de l'extrait sec (5 jours), ont été mesurés. Les résultats obtenus par la moyenne quadratique (RMS) et l'ANOVA montrent qu'il n'y a pas de différence entre les spectres d'un puits, entre les puits d'une plaque et non plus entre les spectres de deux plaques. En effet, les valeurs de F sont inférieures au seuil critique défini pour F. En plus, les Probabilités obtenues sont toutes supérieures au seuil de signification alpha qui est de 0,05. Elles sont respectivement 0,98; 0,54 et 0,28. Par ailleurs, on constate une forte différence entre les spectres selon la durée de conservation de l'extrait sec du lait à 22 °C à l'étuve vu que la valeur de F est supérieure au seuil critique défini pour F. De plus, la probabilité est de 0,0008, inférieure donc à 0,05 indiquant une différence hautement significative. Cela s'explique par le fait que l'eau contenue dans l'échantillon s'évapore continuellement et cette évaporation impacte les spectres pendant la conservation.

Une ACP (Analyse en Composante Principale) appliquée à l'ensemble des spectres les a discriminés en 3 groupes nets et distincts correspondant aux 3 catégories de lait utilisées (entier, demi-écrémé, écrémé). Les loadings ont confirmé par différents pics (1703; 1720; 2294; 2310; 2340 nm,...) que c'est bel et bien la matière grasse qui est la principale variable de différenciation des 3 types de lait. D'autres ACP réalisées sur chacune des catégories de lait confirment une

variation au niveau de l'échantillon à partir du 3<sup>e</sup> jour de conservation à l'étuve à 22°C. D'après les loadings, la matière grasse et surtout l'eau sont les variables qui ont notamment varié pendant la conservation.

Les équations de prédiction obtenues pour la méthode sont performantes pour la matière grasse et la matière sèche mais pas pour les protéines. Ces dernières accusent un trop faible coefficient de détermination ( $R^2$ ) soit 0,01. Cela s'explique par le fait qu'il y a un manque de variabilité dans les données liées aux protéines. En effet, parmi les 3 catégories de lait analysé, seules les quantités de matière grasse et de matière sèche ont une grande variabilité. La proportion dans la matrice laitière va de 0 à 4,41 pour la MG et de 9,71 à 13,01 pour la MS. Par contre la proportion des protéines n'a pratiquement pas varié dans les laits entiers, demi-écrémés ou écrémés. Elle va de 3,31 à 3,47, ce qui explique le manque de précision de l'équation.

En perspectives, pour valider la méthode, il serait d'un grand intérêt de reprendre l'expérience avec des laits reconstitués en différentes proportions en protéines. De même, la reproductibilité de la méthode développée est à confirmer dans d'autres laboratoires.

La deuxième partie consistait à calibrer pour l'analyse du lait, un spectromètre portatif de type VIAVI dénommé NIR4FARM. Il est utilisé généralement sur la ferme, pour l'analyse des fourrages en frais sur le terrain. Le calibrer pour l'analyse du lait pourrait permettre au producteur laitier de minimiser les dépenses pour les analyses de ses échantillons dans les laboratoires accrédités tout en disposant d'un seul spectromètre pour toutes ses analyses à la ferme. Des équations ont été obtenues d'une part en MLR (Régression linéaire multiple); méthode qui utilise quelques longueurs d'onde cibles du spectre et d'autre part en PLS où on utilise toutes les longueurs d'onde. Les résultats montrent que les équations en MLR sont plus performantes que celles obtenues en PLS. Pour la matière grasse par exemple, on obtient en MLR une valeur de SEP (Erreur Standard de Prédiction) de 0,27 plus faible qu'en PLS où on obtient 0,79. Par contre, pour les protéines, les équations ne sont pas performantes pour les deux méthodes. Avec un coefficient de détermination de 0,10 en MLR, cela prétend qu'il y a très peu de corrélation entre les valeurs labo et les valeurs prédites. Cela peut s'expliquer par les mêmes raisons explicitées précédemment: il existe un manque de variabilité pour les valeurs liées aux protéines parmi les valeurs de références.

Toutefois, en essayant de comparer les équations en MLR obtenues pour le NIR4FARM avec celles obtenues en PLS pour le "MPA", un spectromètre de routine dont on dispose au CRA-W, nous avons constaté une nette différence entre les deux équations. En effet, avec des SEP de 0,07; 0,07 et 0,08 respectivement pour la matière grasse, les protéines et la matière sèche, on peut déduire que le MPA est très performante par rapport au NIR4FARM. En essayant d'identifier le problème, nous avons pensé que la gamme du PIR utilisée par les deux instruments pourraient être à la base des différences observées dans les équations. Autrement dit, il y aurait un manque d'information spectrale dans la gamme du NIR4FARM car celui-ci utilise seulement une partie du PIR (920 à 1658 nm) alors que le MPA utilise toute la gamme du PIR (800 à 2500 nm). Ainsi, nous avons développé des équations pour le MPA avec sa propre base de données mais dans la gamme du NIR4FARM c'est-à-dire entre 920 et 1658 nm. Celles-ci se montrent très performantes avec des SEP qui varient sont de 0,11; 0,08 et 0,06 respectivement pour les trois variables. Ce qui confirme que la gamme 920 à 1658 nm est assez riche en information spectrale et ne peut être mise en cause du manque de précision des équations obtenues pour le NIR4FARM.

En perspective, il y a de forte chance d'obtenir des équations performantes avec le NIR4FARM s'il travaille dans les mêmes conditions que le MPA à savoir, mesure faite en transmission avec un lait chauffé à 40°C et homogénéisé.

Couplé à un meilleur système de mesure, le NIR4FARM a certainement un intérêt dans l'analyse du lait à la ferme.

## **BIBLIOGRAPHIE**

- Baeten V. (2016). Note de cours. Contrôle technologique de qualité: Chap.4. Université Catholique de Louvain. Louvain. Belgique.
- Confédération Belge de l'Industrie laitière (CBL) (2015). Rapport d'activités annuelles de 2014.
- Courtet-Leymaryos F., (2010). Qualité nutritionnelle du lait de vache et de ses acides gras: Voies d'amélioration par l'alimentation. Bibliothèque de la Faculté de Médecine de Créteil. France. Paris. 128 p.
- Bertrand D., Dufour E., (2000). La spectroscopie infrarouge et ses applications analytiques. Collection Sciences et Techniques Agroalimentaires. Paris. 497-540 p.
- El Hajji A., S. Zaydoun. (2000). La spectroscopie Infrarouge. 119 p.
- Jeantet R., Croguennec T., Schuck P. et Brule G., (2008). Les produits laitiers ,2ème édition, Lavoisier. 171-185 p.
- Dardenne P., (2000). Understanding the NIR statistics. Centre Wallon de Recherches Agronomiques. 6 p.
- Urbano-Cuadrado M., Luque de Castro M.D., Perez-Juanb P.M., Garcia-Olmoc J., Gómez-Nieto M.A., (2004). Near infrared reflectance spectroscopy and multivariate analysis in enology. Determination or screening of fifteen parameters in different types of wines. University of Cordoba, Spain. 1-8 p.

## ANNEXE

### Annexe 1: Quantités et types de laits de commerce utilisés

Laits utilisés en Microscopie PIR					Laits utilisés en Spectroscopie PIR			
	Entier	Demi-écrémé	Ecrémé	Quantité	Entier	Demi-écrémé	Ecrémé	Quantité
<b>Boni</b>	1	1	0	2	1	1	0	2
<b>Campina</b>	1	1	0	2	1	1	0	2
<b>Carrefour</b>	1	1	1	3	1	1	1	3
<b>Lola</b>	1	1	0	2	1	1	0	2
<b>Everyday</b>	1	1	1	3	1	1	0	2
<b>Fairebel</b>	1	1	0	2	0	1	0	1
<b>Joyvalle</b>	1	1	0	2	1	1	0	2
<b>La bande des Félaït</b>	0	1	0	1	0	0	0	0
<b>Lactel</b>	0	1	0	1	1	0	0	1
<b>Lait belge</b>	1	1	1	3	0	1	0	1
<b>Lait 365</b>	1	1	1	3	1	1	1	3
<b>Quantité</b>	<b>9</b>	<b>11</b>	<b>4</b>	<b>24</b>	<b>8</b>	<b>9</b>	<b>2</b>	<b>19</b>

### Annexe 2: Photos des laits de commerce utilisés



### Annexe 3: Fonctionnement schématisé du spectromètre à transformée de Fourier (FTIR)

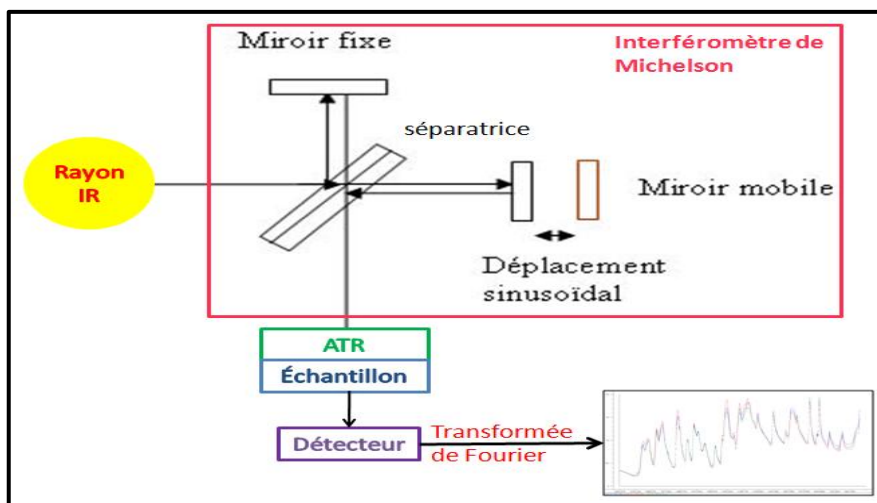


Schéma du spectromètre à transformée de Fourier

L'interféromètre de Michelson fut créé en 1891. Lorsqu'un rayonnement infrarouge arrive sur l'échantillon, il mesure les longueurs d'onde auxquelles l'échantillon absorbe ainsi que les intensités d'absorption. L'interféromètre comporte un miroir semi transparent appelé *séparatrice*, qui divise le faisceau infrarouge en deux. L'un est réfléchi vers un miroir fixe et va parcourir un chemin optique fixe tandis que l'autre est transmis vers un miroir mobile qui avance et recule à une vitesse constante. Son trajet optique, de longueur variable, va moduler chaque longueur d'onde à une fréquence différente. Ces deux miroirs sont placés perpendiculairement l'un à l'autre de sorte que ces deux faisceaux puissent se recombiner sur la séparatrice. Ils vont alors interférer entre eux. Un interférogramme est le résultat de l'ensemble des interférences positives et négatives, de l'addition de toutes les fréquences du faisceau. Il peut être défini comme un signal, c'est-à-dire comme une signature de l'intensité en fonction de la position du miroir. Chaque point qui constitue le signal, est fonction de la position du miroir mobile et possède une information sur chaque fréquence infrarouge provenant de la source. Mesuré très rapidement (de l'ordre de grandeur de la seconde), l'interférogramme contient tous les renseignements requis pour produire un spectre. Le faisceau sortant de l'interféromètre est réfléchi vers l'échantillon, où des absorptions auront lieu pour qu'ensuite, le détecteur les enregistre (El Hajji A., S. Zaydoun, 2000).

## Annexe 4: Les fonctions statistiques

### Equation 1: La Moyenne quadratique (RMS)

$$RMS = \sqrt{\frac{\sum_{i=1}^n (Y_{ij} - \bar{Y}_i)^2}{n}} \quad \text{Equation: 1.1}$$

où **n** est le nombre de longueurs d'onde, **X<sub>ij</sub>** le log (1/T) pour l'échantillon j correspondant à la longueur d'onde i et **ŷ<sub>i</sub>** le log (1/T) pour la moyenne des spectres de l'échantillon à la longueur d'onde i.

### Equation 2: La méthode des moindres carrés partiels (PLS)

L'équation du modèle de la régression PLS est donnée par :

$$Y = b_1 t_1 + e_1 = b_1 t_1 + b_2 t_2 + e_2$$

où, **b<sub>1</sub>** est le coefficient de la régression linéaire (simple) de Y sur t<sub>1</sub>, **b<sub>2</sub>** le coefficient de la régression linéaire (simple) de e<sub>1</sub> sur t<sub>2</sub> et où e<sub>1</sub> et e<sub>2</sub> sont les vecteur des résidus.

### Equation 3: L'erreur standard de calibration (SEC)

$$SEC = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n - p - 1}}$$

où **y<sub>i</sub>** est la valeur de référence; **ŷ<sub>i</sub>** : la valeur prédite; "**n**" le nombre d'échantillon dans l'équation et **p** le nombre de facteurs (PLS) ou de longueurs d'onde dans l'équation.

### Equation 4 : L'erreur standard de prédiction (SEP) et de cross validation (SECV)

$$SEP = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}} \quad \text{Equation: 4.1}$$

où **n** est le nombre d'échantillon prédit dans le groupe

$$SECV = \sqrt{\frac{SEP_1^2 + SEP_2^2 + \dots + SEP_n^2}{n}} \quad \text{Equation: 4.2}$$



où **n** est le nombre de groupe d'échantillon

**Equation 5: Le coefficient de détermination (R<sup>2</sup> ou RSQ)**

$$R^2 = \frac{SD_y^2 - SEC^2}{SD_y^2} \quad \text{Equation: 5.1}$$

avec

$$SD = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}} \quad \text{Equation: 5.2}$$

où **SD<sub>y</sub>** est la déviation standard et, **SEC** l'erreur standard de calibration.

**Equation 6: Le RPD**

$$RPD = \frac{SD}{RMSECV}$$

**Annexe 5: Les caractéristiques affichées sur l'emballage des échantillons de lait de commerce utilisés dans l'étude**

#	Marque	Catégorie	MG	GLU	PRO	EAU	DM
1	Boni	Entier	3.60	4.90	3.30	88.20	11.80
2	Boni	Demi-écrémé	1.50	4.80	3.20	90.50	9.50
3	Campina	Entier	3.60	4.70	3.50	88.20	11.80
4	Campina	Demi-écrémé	1.50	4.80	3.60	90.10	9.90
5	Carrefour	Entier	3.60	4.80	3.20	88.40	11.60
6	Carrefour	Demi-écrémé	1.50	4.80	3.20	90.50	9.50
7	Carrefour	Ecrémé	0.00	4.80	3.20	92.00	8.00
8	Lola	Entier	3.60	4.80	3.20	88.40	11.60
9	Lola	Demi-écrémé	1.00	4.80	3.20	91.00	9.00
10	Everyday	Entier	3.60	4.70	3.40	88.30	11.70
11	Everyday	Demi-écrémé	3.60	4.70	3.40	88.30	11.70
12	Everyday	Ecrémé	0.10	4.70	3.40	91.80	8.20
13	Fairebel	Entier	4.10	4.90	3.40	87.60	12.40
14	Fairebel	Demi-écrémé	1.60	5.00	3.30	90.10	9.90
15	Joyvalle	Entier	3.60	4.80	3.40	88.20	11.80
16	Joyvalle	Demi-écrémé	1.50	4.80	3.50	90.20	9.80
17	Bande des Félaït	Demi-écrémé	1.60	4.80	3.50	90.10	9.90
18	Lactel	Demi-écrémé	1.50	4.80	3.20	90.50	9.50
19	Lait belge	Entier	3.60	4.70	3.40	88.30	11.70
20	Lait belge	Demi-écrémé	1.60	4.80	3.50	90.10	9.90
21	Lait belge	Ecrémé	0.50	4.90	3.50	91.10	8.90
22	Lait 365	Entier	3.60	4.70	3.40	88.30	11.70
23	Lait 365	Demi-écrémé	1.60	4.80	3.50	90.10	9.90
24	Lait 365	Ecrémé	0.10	4.90	3.50	91.50	8.50

## Annexe 6: Résultats de Répétabilité au sein d'un puits d'une plaque

Test d'Analyse de variance (ANOVA) sur les RMS des spectres au sein d'un puits							
Groupes	Nombre d'échantillons	Somme	Moyenne	Variance	F	Probabilité	Valeur critique de F
1er point	2	0,015	0,007	0,00000	0,08	0,98	2,3
2e point	2	0,021	0,011	0,00004			
3e point	2	0,017	0,009	0,00010			
4e point	2	0,018	0,009	0,00004			
5e point	2	0,014	0,007	0,00003			

## Annexe 7: Résultats de Répétabilité au sein de plusieurs d'une plaque

Test d'Analyse de variance (ANOVA) sur les RMS des spectres des 3 puits						
Groupes	Nombre d'échantillons	Somme	Moyenne	Variance		
1er puits	10	29,3	2,93	1,15		
2e puits	10	21,84	2,184	1,18		
3e puits	10	27,01	2,701	4,59		
Source des variations	Somme des carrés	Degré de liberté	Moyenne des carrés	F	Probabilité	Valeur critique pour F
Entre Groupes	2,92	2	1,46	0,63	0,54	3,35
A l'intérieur des groupes	62,22	27	2,30			
Total	65,14	29				

## Annexe 8: Résultats de reproductibilité au cours de la conservation

Test d'Analyse de variance (ANOVA) sur les RMS des spectres pendant les 5 jours de conservation						
Groupes	Nombre d'échantillons	Somme	Moyenne	Variance		
1er jour	30	78,15	2,61	2,25		
3e jour	30	35,47	1,18	2,72		
5e jour	30	79,63	2,65	3,21		
Source des variations	Somme des carrés	Degré de liberté	Moyenne des carrés	F	Probabilité	Valeur critique pour F
Entre Groupes	41,93	2	20,97	7,69	0,0008	3,10
A l'intérieur des groupes	237,18	87	2,73			
Total	279,11	89				

## Annexe 9: Résultats du "test post-hoc" sur la reproductibilité des spectres collectés entre le 1er et le 3e jour de conservation

Test d'Analyse de variance (ANOVA) sur les RMS et test post-hoc entre 1er et 3e jour						
Groupes	Nombre d'échantillons	Somme	Moyenne	Variance		
1er jour	30	78,15	2,61	2,25		
3e jour	30	35,47	1,18	2,72		
Source des variations	Somme des carrés	Degré de liberté	Moyenne des carrés	F	Probabilité	Valeur critique pour F
Entre Groupes	30,36	1	30,36	12,22	0,001	4,01
A l'intérieur des groupes	144,08	58	2,48			
Total	174,44	59				

**Annexe 10: Résultats du "test post-hoc" sur la reproductibilité des spectres collectés entre le 1er et le 3e jour de conservation**

Test d'Analyse de variance (ANOVA) sur les RMS et test post-hoc entre 3e et 5e jour						
Groupes	Nombre d'échantillons	Somme	Moyenne	Variance		
3e jour	30	35,47	1,18	2,72		
5e jour	30	79,63	2,65	3,21		
Source des variations	Somme des carrés	Degré de liberté	Moyenne des carrés	F	Probabilité	Valeur critique pour F
Entre Groupes	32,50	1	32,50	10,96	0,002	4,01
A l'intérieur des groupes	172,04	58	2,97			
Total	204,55	59				

**Annexe 11: Résultats de reproductibilité entre deux plaques**

Test d'Analyse de variance (ANOVA) sur les RMS des spectres collectés sur 2 plaques						
Groupes	Nombre d'échantillons	Somme	Moyenne	Variance		
Plaque 1	30	78,15	2,61	2,25		
Plaque 2	30	91,95	3,07	3,19		
Source des variations	Somme des carrés	Degré de liberté	Moyenne des carrés	F	Probabilité	Valeur critique pour F
Entre Groupes	3,17	1	3,17	1,17	0,28	4,01
A l'intérieur des groupes	157,77	58	2,72			
Total	160,94	59				

**Annexe 12: Valeurs de prédiction utilisées comme valeurs de référence en Microscopie pour la prédiction PLS**

Valeurs de prédictions des laits séchés le 1er jour				Valeurs de prédictions des laits séchés jusqu'au 3e jour			Valeurs de prédictions des laits séchés jusqu'au 5e jour		
Catég,	MG	PRO	MS	MG	PRO	MS	MG	PRO	MS
002 EN	3,70	3,47	13,01	3,39	3,47	12,59	3,88	3,31	13,01
004 DE	1,40	3,32	10,90	1,60	3,20	11,10	1,53	3,23	11,05
006 EN	3,69	3,47	11,02	4,29	3,18	13,35	3,91	3,30	13,04
008 DE	1,42	3,40	12,90	1,87	3,14	11,34	1,97	3,18	11,42
009 EN	3,68	3,40	11,07	2,38	3,12	11,76	4,27	3,20	13,34
012 DE	1,39	3,29	12,74	1,76	3,16	11,23	1,77	3,17	11,25
014 EC	0,00	3,41	11,00	0,19	3,42	9,89	0,16	3,36	9,86
016 EN	3,66	3,46	9,72	4,11	3,25	13,21	4,10	3,28	13,20
017 DE	1,41	3,29	12,72	2,38	3,12	11,76	2,44	3,15	11,81
020 EN	3,61	3,35	9,54	3,93	3,28	13,06	3,76	3,32	12,92
022 DE	1,42	3,31	10,86	1,99	3,22	11,44	1,86	3,39	11,32
023 EC	0,00	3,37	12,78	0,03	3,36	9,76	0,07	3,39	9,79
026 EN	4,41	3,42	11,04	5,07	3,20	14,00	5,24	3,23	14,15
027 DE	1,40	3,46	9,72	2,00	3,14	11,43	1,82	3,31	11,28
029 EN	3,69	3,36	13,52	4,30	3,25	13,37	3,69	3,36	12,84
031 DE	1,41	3,14	10,67	1,83	3,13	11,29	1,91	3,18	11,37
033 DE	1,40	3,42	12,95	1,95	3,20	11,40	1,88	3,37	11,33
035 DE	1,40	3,44	11,09	1,78	3,24	11,24	1,82	3,16	11,29
039 DE	1,42	3,35	10,90	1,86	3,22	11,33	1,70	3,29	11,20
043 EN	3,64	3,44	11,05	4,13	3,23	13,22	3,71	3,33	12,86
045 DE	1,39	3,32	10,87	1,78	3,46	11,25	1,89	3,34	11,35
047 EC	0,00	3,34	12,81	0,20	3,36	9,89	0,24	3,39	9,93

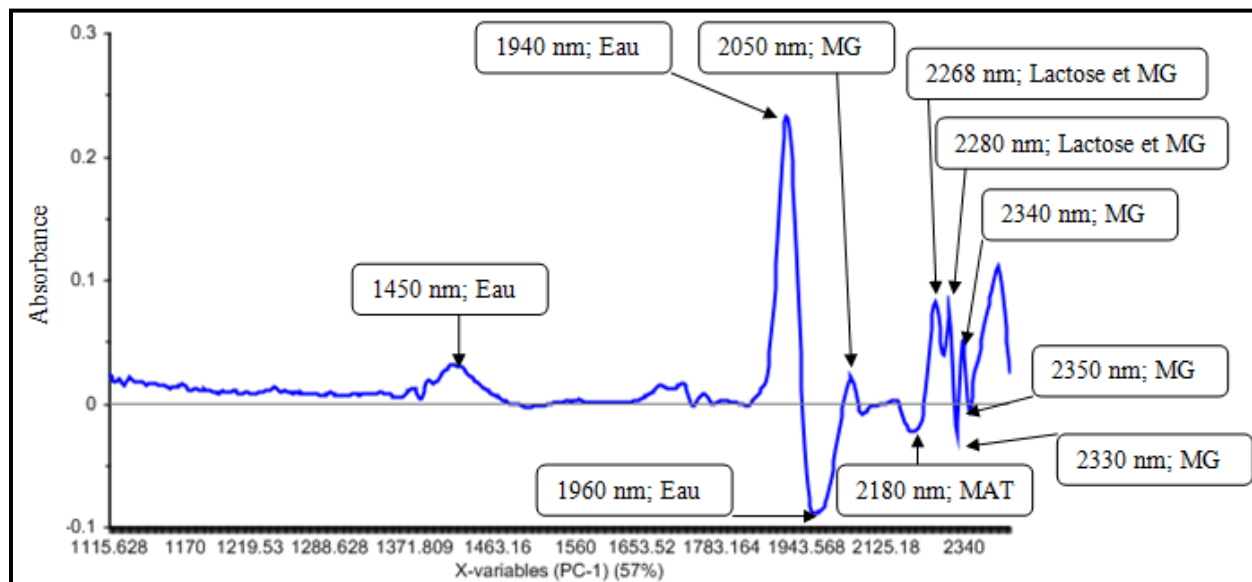
**Annexe 13: Attribution des bandes spectrales en spectrométrie PIR pour les principaux composants du lait et des produits laitiers**

<b>Bandes spectrales en SPIR pour les principaux composants du lait et produits laitiers</b>		
<b>Longueur d'onde (nm)</b>	<b>Attribution</b>	<b>Composants concernés</b>
930	Elongation C-H 3 <sup>e</sup> harmonique	Matière grasse
975	Elongation O-H 2 <sup>e</sup> harmonique	Eau
1025	Elongation N-H 2 <sup>e</sup> harmonique	Matière azotée totale
1150-1250	Elongation C-H 2 <sup>e</sup> harmonique	Matière grasse
1450	Elongation O-H 1 <sup>ère</sup> harmonique	Eau
1700-1800	Elongation C-H 1 <sup>ère</sup> harmonique	Matière grasse
1940	Combinaison élongation + déformation O-H	Eau
2050	Combinaison N-H + amide II	Matière azotée totale
2090	Combinaison élongation O-H + déformation O-H	Lactose
2180	Combinaison 2 * amide I +amide III	Matière azotée totale
2270	Combinaison élongation C-O + élongation O-H Combinaison élongation C-H + déformation C-H	Lactose <i>Matière grasse</i>
2310	Combinaison élongation C-H + déformation C-H	<i>Matière grasse</i>
2350	Déformation C-H 2 <sup>e</sup> harmonique	<i>Matière grasse</i>

Source: Bertrand *et al.*, 2000

En italique, bandes non spécifiques

**Annexe 14: Loading de la 1ère composante de l'ACP appliquée sur les spectres NIR de la catégorie des lait écrémés (spectres traités par SNV, baseline et 1ère dérivée)**





**Annexe 15: Valeurs prédites corrigées du biais utilisées comme valeurs de référence en Spectroscopie pour la calibration du NIR4FARM**

Valeurs Prédites corrigées du biais				
	Codification	MG	PRO	MS
1	DQ160099-001	3,70	3,47	13,01
2	DQ160099-003	1,40	3,47	11,02
3	DQ160099-005	3,69	3,40	12,90
4	DQ160099-007	1,42	3,40	11,07
5	DQ160099-010	3,68	3,29	12,74
6	DQ160099-011	1,39	3,41	11,00
7	DQ160099-013	0,00	3,46	9,72
8	DQ160099-015	3,66	3,29	12,72
9	DQ160099-018	1,41	3,31	10,86
10	DQ160099-019	3,61	3,37	12,78
11	DQ160099-021	1,42	3,42	11,04
12	DQ160099-028	1,40	3,14	10,67
13	DQ160099-030	3,69	3,42	12,95
14	DQ160099-032	1,41	3,44	11,09
15	DQ160099-036	1,40	3,44	11,05
16	DQ160099-040	1,42	3,32	10,87
17	DQ160099-044	3,64	3,34	12,81
18	DQ160099-046	1,39	3,32	10,90
19	DQ160099-048	0,00	3,35	9,54

**Annexe 16: Paramètres exhaustifs des équations obtenues en régression MLR pour le NIR4FARM**

	NIR4FARM Calibration MLR					NIR4FARM Validation MLR <i>sur spectres moyennés</i>							
	N	SEC	SEC V	SD	R <sup>2</sup>	N	SEP	Biais	SEPC	SD	R <sup>2</sup>	RPD (SEP)	RPD (SEPC)
MG	152	0,30	0,31	1,25	0,94	4	0,27	0,23	0,17	1,98	0,99	7,33	11,64
PRO	155	0,46	0,46	0,50	0,18	4	0,40	-0,13	0,43	0,45	0,10	1,12	1,04
MS	157	0,55	0,56	1,35	0,84	4	0,49	-0,37	0,36	2,05	0,99	4,18	5,69

**Annexe 17: Paramètres exhaustifs des équations obtenues en régression PLS pour le NIR4FARM**

	<b>NIR4FARM Calibration PLS</b>					<b>NIR4FARM Validation PLS <i>sur spectres moyennés</i></b>							
	N	SEC	SEC V	SD	R <sup>2</sup>	N	SEP	SEPC	Biais	SD	R <sup>2</sup>	RPD (SEP)	RPD (SEPC)
MG	150	0,34	0,39	1,24	0,93	4	0,79	0,12	-0,79	2,02	0,99	2,55	16,83
PRO	159	0,25	0,37	0,56	0,80	4	5,70	0,32	5,68	0,27	0,00	0,04	0,84
MS	152	0,31	0,45	1,35	0,95	4	6,04	0,33	6,04	1,97	0,98	0,32	5,96

**Annexe 18: Paramètres des équations obtenues en régression PLS pour le MPA**

	<b>BRUKER MPA Equation de prédiction Full range(850-2500 nm)</b>			<b>BRUKER MPA Equation de prédiction Range NIR4FARM (920-1658 nm)</b>		
	SEP	R <sup>2</sup>	RPD	SEP	R <sup>2</sup>	RPD
MG	0,06	0,99	10,67	0,11	0,99	8,60
PRO	0,04	0,99	27,50	0,08	0,98	6,91
MS	0,08	0,99	14,79	0,07	0,99	21,1