



ISARA-Lyon

23, rue Jean Baldassini

69364 LYON CEDEX 07



DUMOULIN SA

Parc Industriel 18

B-5300 SEILLES

L'entreprise *DUMOULIN S.A.* exige une confidentialité des travaux réalisés pendant une
durée de 3 ans.

Analyse en proche infrarouge d'aliments naturels pour animaux (spécifiquement extrudés et floconnés)

Mémoire de Fin d'Etudes

43^{ème} promotion (2014-2015)

Soutenance : 10/07/2015

Enseignant responsable : HALLIER Arnaud

GAVENC Laurent

Elève ingénieur ISARA Lyon

Directeur de mémoire : VANVOLSEM Thibaut

Ce document ayant été réalisé par des Elèves-Ingénieurs de l'ISARA-Lyon dans le cadre d'une convention avec DUMOULIN S.A., toute mention, communication ou diffusion devra faire état de l'origine ISARA-Lyon.

Sommaire

Introduction.....	1
Présentation de l'entreprise.....	3
Partie Bibliographique.....	5
I. Les procédés d'extrusion et de floconnage.....	6
I.1. Les extrudés.....	6
I.2. Les flocons.....	6
II. La spectrométrie proche infrarouge (SPIR).....	7
II.1. Historique de la SPIR.....	7
II.2. Les principes fondamentaux de la SPIR.....	7
II.2.1. Le rayonnement électromagnétique.....	8
II.2.2. L'interaction entre le rayonnement et la matière.....	8
II.2.3. La spectroscopie vibrationnelle.....	9
II.3. L'application de la spectrométrie proche infrarouge pour la prédiction en alimentation animale.....	11
II.4. Les avantages de la SPIR.....	11
II.5. Les contraintes de la SPIR.....	11
III. L'appareillage.....	12
III.1. Généralités.....	12
III.2. Configuration instrumentale.....	13
III.2.1. La source lumineuse.....	13
III.2.2. Le système de sélection des longueurs d'onde.....	13
III.2.3. Le détecteur.....	14
III.2.4. L'enregistreur.....	14
III.3. Les facteurs influant sur les mesures spectrales.....	14
IV. Le traitement des données : la chimiométrie appliquée à la SPIR.....	15
IV.1. Définition de la chimiométrie.....	15
IV.2. Analyse qualitative : l'analyse en composantes principales (ACP).....	16
IV.3. Les prétraitements des spectres.....	16
IV.3.1. La moyenne spectrale.....	16
IV.3.2. La normalisation des spectres : traitement par SNV (Standard Normal Variate).....	16
IV.3.3. Les dérivées spectrales.....	17
IV.3.4. Le lissage spectral.....	18
IV.4. La calibration.....	18
IV.4.1. Les méthodes de régression.....	19
IV.4.2. Les paramètres d'appréciation de la calibration.....	19
IV.5. La validation.....	19

Partie Expérimentale	22
I. Matériel et méthode	23
I.1. Instrumentation	23
I.1.1. Le spectromètre	23
I.1.2. Les logiciels de traitements des données spectrales et de prédiction.....	25
I.2. Echantillonnage et définition des paramètres chimiques.....	25
I.2.1. Les extrudés.....	26
I.2.2. Les flocons	27
I.2.2. Autres aliments et matières premières.....	28
I.3. Méthodologie de la construction des équations de calibration	29
II. Présentation et interprétation des résultats des calibrations	32
II.1. Les aliments extrudés.....	32
II.2. Les correcteurs protéiques.....	38
II.3. Les aliments pour porcs.....	43
II.4. Les matières premières	48
II.5. Les aliments floconnés	56
III. Discussion sur la réglementation européenne relative à l'étiquetage des matières premières et des aliments composés pour animaux.....	57
IV. Discussion sur la méthodologie.....	60
Conclusion	72
Bibliographie	74

Remerciements

Mes remerciements s'adressent tout d'abord à mon directeur de mémoire, Monsieur Thibaut VANVOLSEM, pour sa disponibilité et son suivi tout au long du projet.

Je remercie Monsieur Denis BIEMAR, Responsable qualité, pour sa disponibilité et les nombreux conseils qu'il m'a fournis tout au long de ce mémoire.

Je remercie Monsieur Werner REUTER, directeur Recherche et Développement, pour l'aide qu'il m'a apportée au cours de ce mémoire et ses nombreux conseils qui m'ont permis d'avancer.

Je tiens également à remercier Messieurs Bernard LECLER et Stéphane BRICHARD, travaillant au département NIR de la valorisation des productions au CRA-W pour leur expertise qu'ils ont apportée lors de la réalisation de ce projet.

Je remercie également Messieurs Xavier BAUDOIN et Lionel ROCZEN, avec qui j'ai travaillé sur le projet infrarouge au sein de *Dumoulin*.

Mes remerciements vont également à Monsieur Arnaud HALLIER, qui, en sa qualité d'enseignant suiveur, a su se rendre disponible pour assurer le suivi du projet et me fournir de précieux conseils.

Enfin, je remercie toutes les personnes qui ont contribué de près ou de loin à la réussite du projet : opérateurs et autres collaborateurs de l'entreprise *Dumoulin*.

Table des figures

Figure 1 : Répartition des sites de production de Dumoulin en Belgique

Figure 2 : Présentation des types d'aliments fabriqués par Dumoulin

Figure 3 : Evolution des prix des protéagineux de 2007 à 2013

Figure 4 : Evolution des prix des céréales de 2007 à 2013

Figure 5 : Organigramme de l'usine d'Andenne selon le principe de différenciation

Figure 6 : Vue d'ensemble du spectre électromagnétique et zoom sur le domaine du proche infrarouge

Figure 7 : Représentation schématique de l'interaction de la lumière avec la matière

Figure 8 : Modèle de l'oscillateur

Figure 9 : Exemples de modes de vibrations moléculaires

Figure 10 : Modèle de l'oscillateur anharmonique

Figure 11 : Aperçu des bandes d'absorption dans le proche infrarouge

Figure 12 : Configuration des instruments de spectrométrie

Figure 13 : Classification des spectromètres proche infrarouge

Figure 14 : Principe d'un système à diodes émettrices de lumière

Figure 15 : Principe d'un interféromètre de Michelson (spectromètre à transformée de Fourier)

Figure 16 : Exemple de spectre infrarouge par transformée de Fourier

Figure 17 : Définition de la chimiométrie et ses applications

Figure 18 : Effets d'une normalisation SNV

Figure 19 : Transformation en dérivée première d'un spectre traité par SNV

Figure 20 : Transformation en dérivée seconde d'un spectre traité par SNV

Figure 21 : Aperçu des différents outliers existants

Figure 22 : Tête de lecture on-line Polytec

Figure 23 : Dispositif expérimental du CRA-W simulant l'analyse on-line

Figure 24 : Dispositif d'analyse du Polytec en laboratoire

Figure 25 : Aperçu de la variation expliquée permettant de choisir le nombre de termes du modèle

Figure 26 : Démarche de la mise en place d'une équation de calibration

Figure 27 : Aperçu des pics de vapeur d'eau dans le cas de l'analyse en laboratoire

Figure 28 : Conséquences spectrales de la modification de la distance focale

Figure 29 : Graphique des coefficients appliqués aux absorbances aux longueurs d'onde correspondantes

Liste des tableaux

Tableau 1 : Caractéristiques techniques des appareils Polytec et XDS FOSS

Tableau 2 : Méthodes d'analyse chimique et coefficients de répétabilité

Tableau 3 : Résultats des analyses chimiques, calibrations et cross-validations sur les aliments extrudés

Tableau 4 : Résultats des validations Polytec par comparaison des valeurs chimiques et des prédictions XDS et Polytec pour les aliments extrudés

Tableau 5 : Résultats des analyses chimiques, calibrations et cross-validations sur les correcteurs protéiques

Tableau 6 : Résultats des validations Polytec par comparaison des valeurs chimiques et des prédictions XDS et Polytec pour les correcteurs protéiques

Tableau 7 : Résultats des analyses chimiques, calibrations et cross-validations sur les aliments pour porcs

Tableau 8 : Résultats des validations Polytec par comparaison des valeurs chimiques et des prédictions XDS et Polytec pour les aliments pour porcs

Tableau 9 : Résultats des analyses chimiques, calibrations et cross-validations sur les DDGS

Tableau 10 : Résultats des validations Polytec par comparaison des valeurs chimiques et des prédictions XDS et Polytec pour les DDGS

Tableau 11 : Résultats des analyses chimiques, calibrations et cross-validations sur les tourteaux de colza

Tableau 12 : Résultats des validations Polytec par comparaison des valeurs chimiques et des prédictions XDS et Polytec pour les tourteaux de colza

Tableau 13 : Résultats des analyses chimiques, calibrations et cross-validations sur les tourteaux de lin

Tableau 14 : Résultats des validations Polytec par comparaison des valeurs chimiques et des prédictions XDS et Polytec pour les tourteaux de lin

Table des abréviations et acronymes

% : Pourcent

λ : longueur d'onde

ACP/PCA : Analyse en composantes principales / Principal Component Analysis

ADA : Automated Digital Analyzer

Aw : Activité de l'eau

cm⁻¹ : centimètre⁻¹

CRA-W : Centre Wallon de Recherche Agronomique

DDGS : Dried Distilled Grains with Solubles – Drêches de Distillerie Humide

GH : Distance de Mahalanobis

GMP : Good Management Practice

Hz : Hertz

J : Joule

L : Litre

LED : Light Emitting Diode – Diode émettrice de lumière

m.s⁻¹ : mètre.seconde⁻¹

mol : mole

MLR : Multiple Linear Regression – Régression linéaire multiple

MSC : Multiple Scatter Correction

NIR : Near Infrared Reflectance

nm : nanomètre

PCR : Principal Component Regression – Régression en composantes principales

PDIA : Protéines Digestibles dans l'Intestin grêle d'origine Alimentaire

PLS : Partial Least Square – Régression par les moindres carrés partiels

R² : Coefficient de détermination

R&D : Recherche et Développement

RMS : écart-type au spectre moyen

RPD : Ratio of Prediction (ou Performance) to Deviation ($RPD = SD/SEC\text{V}$ ou SD/SEP)

RPIQ : Ratio de la Performance InterQuartile

s^{-1} : seconde⁻¹

SD : **Standard Deviation** (écart-type pour un paramètre)

SEC/RMSEC : Standard Error of Calibration – Ecart-type standard de calibration

SEC\text{V}/RMSEC\text{V} : Standard Error of Cross-Validation – Ecart-type standard de cross-validation

SEP : Standard Error of Prediction – Ecart-type standard de prédiction

SNV : Standard Normal Variate

SPIR : Spectrométrie Proche Infrarouge

Introduction

L'entreprise *Dumoulin* est spécialisée dans la fabrication d'aliments composés naturels pour animaux (bovins, caprins, ovins, équins, volailles, porcs). Le métier de fabricant d'aliments consiste à développer des rations en fonction du stade de croissance et/ou de production ainsi que de l'environnement de l'animal. Le but est de répondre à ses besoins physiologiques et d'améliorer ses performances zootechniques. L'amélioration continue des performances en formulant de nouvelles rations durables apparaît comme un levier majeur afin de répondre à la demande croissante et nourrir 9 milliards d'êtres humains à l'horizon 2050.

Dumoulin, grâce à son service R&D, s'engage à répondre aux besoins nutritionnels pour réaliser de telles performances, en développant des solutions alimentaires à même de les couvrir. C'est ainsi qu'une diversification forte est apparue avec les aliments floconnés et extrudés. Ce sont des produits à plus haute valeur ajoutée répondant à des demandes nutritionnelles exigeantes en production animale. Exportés à hauteur de 60% en Europe et dans le monde, ils constituent un des piliers majeurs du développement et de la croissance de l'entreprise. L'innovation et la recherche ont depuis toujours été les leitmotivs de *Dumoulin*. Ainsi, de nombreux produits innovants ont vu le jour au cours de cette dernière décennie : gamme d'aliments extrudés ou encore graines de soja thermotraitées par floconnage et toastage.

Le système qualité GMP (Good Management Practice) a été mis en place dès le départ afin de contrôler la réception, le dosage, le mélange, le broyage, la granulation ou encore la manutention. Il permet de garantir le respect des caractéristiques suivantes des produits : valeurs alimentaires (protéine, sucre, amidon, matière grasse, fibres, minéraux), conservation (humidité, A_w), aspect physique (granulométrie, dureté) et aspect sanitaire (contaminants, résidus chimiques). Cependant, en raison de la spécificité des produits extrudés et floconnés, de leur process de fabrication très différent et plus technique que les aliments classiques, le maintien d'une qualité optimale s'avère plus difficile, et le système qualité actuel se révèle insuffisant pour ces produits. Le contrôle qualité par spectroscopie proche infrarouge semblerait ainsi être une technique adaptée à cette problématique, de par sa nature rapide et non destructrice. L'étude de ces aliments est particulièrement intéressante car l'utilisation de la spectroscopie proche infrarouge pour ces aliments n'a à ce jour fait l'objet que peu de recherches, faisant de *Dumoulin* un pionnier en la matière. De plus, l'application de cette technologie est très intéressante en termes de contrôle des critères techniques, sanitaires et d'achat.

Cela amène donc à se questionner sur **les possibilités d'évaluation de la qualité des aliments naturels pour animaux à l'aide de la spectroscopie infrarouge, et plus spécifiquement des aliments extrudés et floconnés.**

L'objectif de l'étude porte sur la création d'équations de calibration pour l'analyse par spectroscopie proche infrarouge de ces aliments atypiques. Cependant, *Dumoulin* ne possède pas les outils et les

connaissances nécessaires pour mener ce projet seul. Une collaboration a donc été mise en place avec le Centre Wallon de Recherche Agronomique (CRA-W) de Gembloux, mondialement réputé pour son expertise dans le domaine de l'analyse infrarouge. Le défi du Groupe *Dumoulin* est donc de construire lui-même l'outil de contrôle qualité adapté à ces produits. L'entreprise ne s'est cependant pas limitée à ces aliments puisque l'analyse infrarouge a concerné par la suite les autres produits finis ainsi que les matières premières. *Dumoulin* s'est également engagé dans une voie unique en Belgique pour un fabricant d'aliments pour animaux : l'analyse de routine on-line où le contrôle s'effectuera après broyage et mélange.

Une description de l'entreprise et du service dans lequel j'ai travaillé sera dans un premier temps réalisée. Dans la partie bibliographique seront tout d'abord présentés les procédés d'extrusion et de floconnage. Par la suite, la spectroscopie proche infrarouge sera présentée, son historique, ses principes, ses applications, ses avantages et ses limites. L'instrumentation employée en proche infrarouge sera ensuite développée dans un troisième chapitre. Enfin, un focus sur le traitement des données chimiques par chimiométrie sera réalisé.

La partie expérimentale exposera dans un premier temps le matériel employé ainsi que la méthode appliquée au cours de ce projet. Les principaux résultats tirés des étapes de calibration et de validation seront ensuite exposés puis une interprétation et une discussion des résultats obtenus et de la méthode mise en place seront réalisées. Enfin seront présentées les conclusions et les perspectives de ce projet pour *Dumoulin*.

Présentation de l'entreprise

Le groupe *Dumoulin* appartient au groupe AVEVE, leader de l'agriculture et de l'horticulture en



Belgique. *Dumoulin* emploie 200 personnes réparties sur 4 sites de productions à travers la Belgique : Hombourg, Moorslede, Kortrijk et Andenne (Seilles) (**Figure 1**). Elle est spécialisée dans la création, la production et la commercialisation d'aliments composés naturels pour animaux.

La production s'élève à 500.000 tonnes/an pour un chiffre d'affaires supérieur à 100 millions/an. Les aliments à destination des bovins lait et viande représentent 55% des tonnages, les porcins 35%, les ovins, caprins et volailles 5% et les équins 5%.

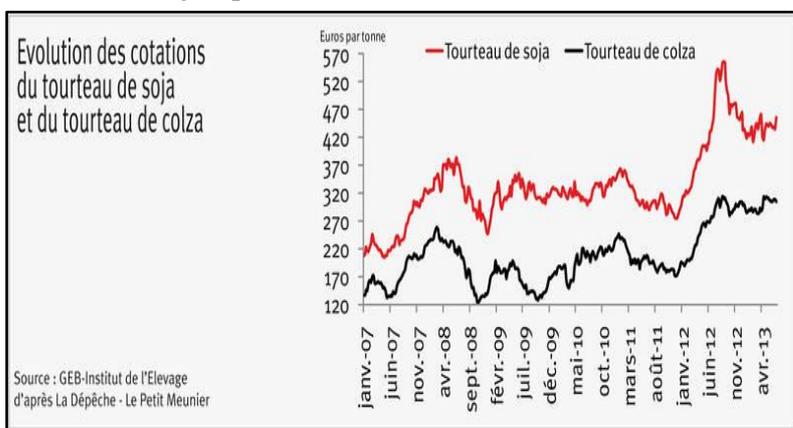
Figure 1 : Répartition des sites de production de Dumoulin en Belgique
Source : Dumoulin, 2013.

Cinq familles de produits composent l'activité de production : farines, granulés, extrudés, flocons et mix (**Figure 2**). Les aliments sont vendus directement aux éleveurs ou à des négociants, soit de manière standard ou selon leurs besoins.



Figure 2: Présentation des types d'aliments fabriqués par Dumoulin
Source : Dumoulin, 2013.

Le groupe doit faire face à de nombreuses contraintes économiques, sociales et environnementales. En



effet, le marché de l'alimentation animale est très concurrentiel, et *Dumoulin* doit également tenir compte de la fragilité économique actuelle des exploitations agricoles et de la fluctuation du prix des matières premières, notamment celui des protéagineux (**Figure 3**) et des céréales (**Figure 4**).

Figure 3 : Evolution des prix des protéagineux de 2007 à 2013

Enfin, *Dumoulin* a des obligations vis-à-vis des consommateurs qui ont des attentes nouvelles au niveau de leur alimentation et du respect de l'environnement.

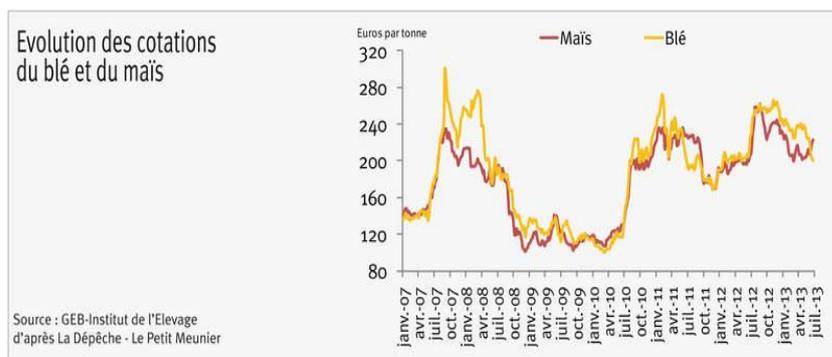


Figure 4 : Evolution des prix des céréales de 2007 à 2013
Source : Dumoulin, 2013.

Le site de production d'Andenne, où j'ai effectué mon mémoire, est situé en bord de Meuse. Cette position lui permet de s'assurer un approvisionnement par bateaux, en plus d'un approvisionnement par camions. La direction du site est composée d'un directeur général et de quatre directeurs délégués : directeur logistique, directeur commercial, directeur recherche et directeur produit et développement. La **Figure 4** présente l'organigramme du site d'Andenne selon le principe de différenciation.

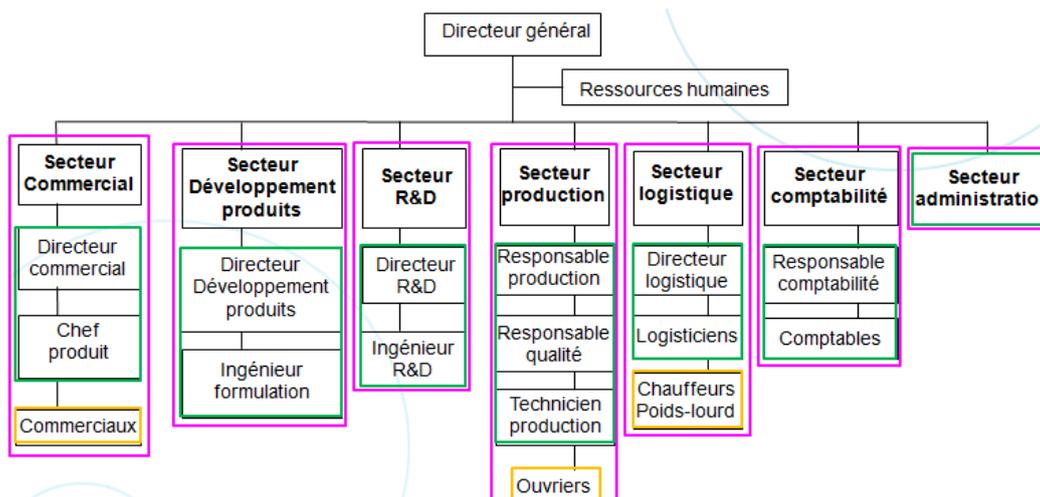


Figure 5 : Organigramme de l'usine d'Andenne selon le principe de différenciation
Source : Dumoulin, 2013.

Le projet infrarouge est ancré dans la gestion de la qualité des aliments et est mené en collaboration avec les Directeurs Développement produits et R&D et le responsable Qualité. Au cours du mémoire, j'ai donc travaillé avec ces personnes afin de mener le projet à son terme, établir les premières courbes afin d'avoir une meilleure connaissance des matières premières, des produits finis et de respecter les engagements faits auprès des clients et apprendre à maîtriser la méthodologie et les logiciels de calibration.

Partie Bibliographique

I. Les procédés d'extrusion et de floconnage

I.1. Les extrudés

Le procédé de cuisson-extrusion consiste à appliquer en continu un traitement thermique couplé à un traitement mécanique (Nutridev, n.d.). Les modifications apportées à l'aliment par l'extrusion lui donnent une forte valeur ajoutée en termes de qualité.

L'extrusion provoque la gélatinisation de l'amidon. Elle se définit comme la « *modification physique de l'amidon en excès d'eau caractérisée par la perte partielle ou totale de la structure granulaire de l'amidon sous l'effet de la température* » (Bouroche et Le Bars, 1996) (**Annexe 1**). Les grains d'amidon gonflent, se dilatent, et perdent leur structure semi-cristalline. A ce stade, le phénomène de gélatinisation est irréversible (Université des Sciences et Technologie de Lille, n.d.). La gélatinisation augmente par ailleurs la digestibilité de l'aliment. L'amidon gélatinisé est dégradé plus rapidement par les enzymes digestives (Roczen, 2014). L'extrusion provoque la destruction des parois cellulaires et l'augmentation de la surface d'échange avec le rumen et l'intestin grêle, ce qui améliore la disponibilité des nutriments (Legoy, n.d.).

Les aliments extrudés possèdent une appétence plus élevée. De plus, le gaspillage est moindre car l'animal n'opère pas de tri parmi les extrudés (Special Diets Service, n.d.). L'extrusion provoque également l'insolubilisation des protéines natives. L'extrusion diminue la quantité d'azote soluble au profit de l'azote non dégradable dans le rumen (PDIA). Ainsi, les protéines échappent à la fermentation dans le rumen et sont digérées dans l'intestin grêle (Demarquilly, n.d.). De plus, elle est à l'origine de l'inactivation des facteurs antinutritionnels et toxiques thermolabiles pouvant jouer sur l'appétit de l'animal (Roczen, 2014). L'extrusion assure la destruction des facteurs antitrypsiques du soja ou la réduction de la teneur en glucosinolates des tourteaux de colza. Elle permet aussi en outre l'inactivation des enzymes sécrétant l'acide cyanhydrique dans le lin (Legoy, n.d.).

Enfin, l'extrusion joue les aspects sanitaires de l'aliment. En raison des hautes températures appliquées, l'extrusion s'apparente à une stérilisation (Special Diets Service, n.d.), assurant la destruction de la majorité des spores de moisissures (Legoy, n.d.). Elle assure d'autre part la diminution de la teneur en eau grâce à l'évaporation en sortie d'extrudeur (5-7%) et au refroidissement/séchage. Cela rend l'aliment moins sensible aux contaminations microbiologiques (Legoy, n.d.).

I.2. Les flocons

Le processus de floconnage est similaire à celui de l'extrusion et présente les mêmes avantages puisque les flocons sont obtenus sous l'effet d'une cuisson sous haute pression. Il assure une meilleure digestibilité de l'animal de part l'amidon gélatinisé et l'augmentation de la surface de contact avec les sucres gastriques, l'inactivation des facteurs antinutritifs, l'amélioration de l'appétence ou encore les aspects sanitaires (Guidolin® Flaked Cereals, n.d.).

II. La spectrométrie proche infrarouge (SPIR)

II.1. Historique de la SPIR

L'étude de la spectrométrie (ou spectroscopie) débute à l'époque d'Isaac NEWTON (1642-1727). Il découvre que la lumière solaire est décomposable grâce à un prisme en un spectre couvrant du rouge au violet : c'est le domaine du visible (Van Belleghem, 2003). En 1800, F.W. HERSCHEL démontre l'existence de l'infrarouge en déplaçant un thermomètre le long du spectre solaire. Il constate que le maximum se trouve au-delà du visible. Ces radiations sont d'abord appelées rayons caloriques, avant que BECQUEREL ne les nomme infrarouges, par opposition aux ultraviolets (Ecole des Mines de Saint Etienne, n.d.).

Avant 1950 et le développement de techniques instrumentales basées sur les propriétés physiques, l'analyse des produits agricoles et agroalimentaires se fait à l'aide de méthodes chimiques. Ces techniques sont coûteuses, et lourdes à mettre en place (Van Belleghem, 2003). La SPIR entre dans une nouvelle dimension grâce à Karl NORRIS, du Département de l'Agriculture du Maryland, qui détermine la teneur en humidité des graines de soja grâce à une série de filtres (CNRIT, n.d.).

L'instrumentation s'améliore d'année en année, et c'est en 1974 qu'est fabriqué le premier appareil de réflexion dans le proche infrarouge, l'Automated Digital Analyzer (ADA). La fin de la décennie voit le développement d'appareils de mesures et de logiciels plus performants, et c'est en 1978 que le premier spectrophotomètre entièrement informatisé est produit (Williams, 2007). C'est enfin en 1980 que les résultats obtenus en proche infrarouge peuvent être analysés avec précision, grâce au couplage de la chimiométrie à la SPIR. Divers outils mathématiques et statistiques peuvent être utilisés pour traiter les données chimiques : application de dérivées, analyse en composantes principales ou encore méthodes de régression. (CNRIT, n.d.)

II.2. Les principes fondamentaux de la SPIR

La SPIR est une « méthode physique basée sur la propriété qu'ont les molécules organiques [...] d'absorber de l'énergie dans le proche infrarouge » (Van Belleghem, 2003). Il s'étend de 800 à 2500 nm. (Bertrand, 2002) (**Figure 6**). L'absorption d'énergie dépend de la composition chimique du produit. Chaque constituant absorbe à des longueurs d'onde spécifiques, correspondant aux fréquences vibratoires des liaisons chimiques. La quantité d'énergie qu'il absorbe est proportionnelle à sa concentration (Bertrand, 2002).

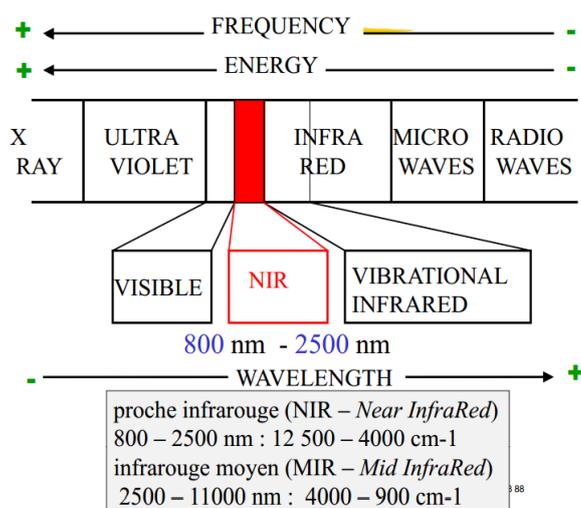


Figure 6 : Vue d'ensemble du spectre électromagnétique et zoom sur le domaine du proche infrarouge.
Source : Dardenne, n.d.

II.2.1. Le rayonnement électromagnétique

Deux théories peuvent expliquer le rayonnement électromagnétique (Lecler, 2014). La radiation électromagnétique a une nature ondulatoire (Roggo, 2003). Elle s'exprime en termes de fréquence de vibration ν (s^{-1}), de longueur d'onde λ (nm) ou de nombre d'onde σ (cm^{-1}). La longueur d'onde est la distance parcourue pendant un cycle complet. Elle est reliée à la fréquence (en Hz) par l'Equation 1 :

Equation 1 $\lambda = c / \nu$ avec c la célérité de la lumière ($3.10^8 m.s^{-1}$)

La radiation électromagnétique a également une nature quantique. Elle est composée de particules élémentaires, les photons. Ces photons contiennent une certaine quantité d'énergie, liée à la fréquence de la radiation et exprimée par la relation (Roggo, 2003) (Equation 2) :

Equation 2 $E = h \cdot \nu$ avec

$E =$ Energie (J)

$h =$ Constante de Planck = $6,626.10^{-34}$ J.s

II.2.2. L'interaction entre le rayonnement et la matière

Lorsqu'une radiation éclaire la matière, elle peut être absorbée, transmise ou réfléchi (Figure 7). La loi de conservation de l'énergie donne la relation suivante (Equation 3) :

Equation 3 $I_0 = I_A + I_T + I_R$ avec

$I_0 =$ Intensité de la lumière incidente

$I_A =$ Intensité de la lumière absorbée

$I_T =$ Intensité de la lumière transmise

$I_R =$ Intensité de la lumière réfléchi

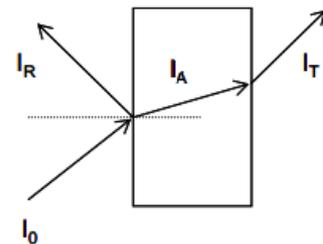


Figure 7 : Représentation schématique de l'interaction de la lumière avec la matière
Source : Roggo, 2003

L'analyse proche infrarouge est basée sur la mesure de l'absorption d'une radiation électromagnétique par la matière dont la longueur d'onde est comprise dans son domaine. L'atténuation de la radiation lorsqu'elle est absorbée est décrite par la loi d'absorption du rayonnement, ou loi de Beer-Lambert (Equation 4) (Fatela, 2012 ; Lecler, 2014 ; Roggo, 2003) :

Equation 4 $A = \log (1/T) = \epsilon.c.L$ avec

$A =$ Absorbance

$T =$ Transmittance = I_T / I_0

$\epsilon =$ coefficient d'extinction molaire ($L.mol^{-1}.cm^{-1}$)

$c =$ concentration de la nature absorbante ($mol.L^{-1}$)

$L =$ épaisseur de l'échantillon (cm)

II.2.3. La spectroscopie vibrationnelle

La région du proche infrarouge correspond à une partie de la spectroscopie vibrationnelle, incluant également le moyen infrarouge (2500-25000 nm). Une molécule peut être modélisée par deux atomes

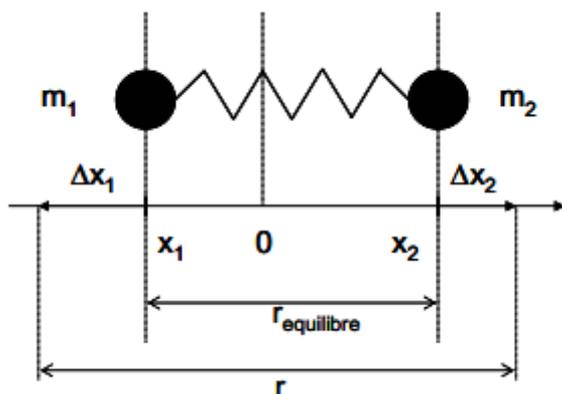
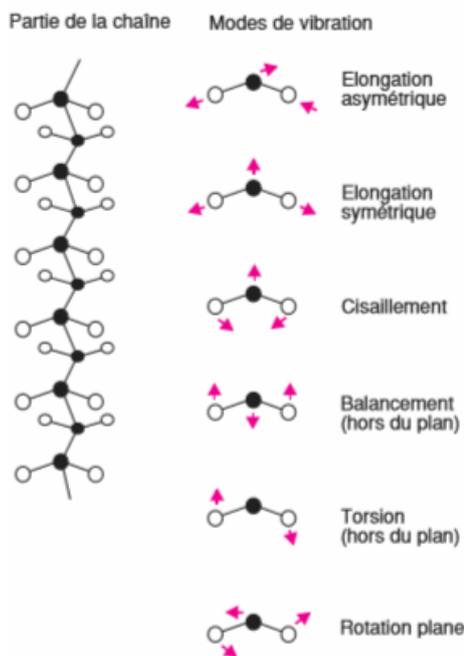


Figure 8 : Modèle de l'oscillateur. Les deux molécules de masse m_1 et m_2 sont reliées par un ressort de constante k .
Source : Roggo, 2003

reliés par un ressort (Figure 8). Ses liaisons sont le résultat de l'équilibre entre les forces de répulsion et les forces d'attraction. Ainsi, les noyaux chargés positivement et les nuages d'électrons chargés négativement se repoussent. Le noyau d'un atome et les électrons d'un autre atome s'attirent, et inversement (Fatela, 2012). La fréquence de vibration de chaque ressort dépend de la nature des groupes chimiques impliqués dans la liaison. Ainsi, l'énergie ne peut être absorbée par une molécule que si les fréquences de la lumière, correspondant à un mode vibrationnel, et des liaisons intermoléculaires sont identiques (Bertrand, 2002).

Il existe deux types de vibrations des liaisons moléculaires, représentés en Figure 9 :

- Les vibrations d'élongation ou élongations de valence (« stretching ») : les atomes s'éloignent et se rapprochent les uns par rapport aux autres le long de leur axe de liaison. Les vibrations d'élongation peuvent être symétriques ou asymétriques
- Les vibrations de déformation (« bending ») : elles entraînent une modification des angles entre les liaisons. Il existe plusieurs modes de déformation : cisaillement, balancement,



hachement ou encore torsion (Fatela, 2012 ; Bertrand, 2002 ; Roggo, 2003).

L'absorption de l'énergie lumineuse par les liaisons entraîne une transition vibrationnelle, l'état d'excitation du système étant trop faible pour engendrer des transitions électroniques. Il est donc possible de passer d'un niveau énergétique à l'autre (Dubuisson, 2009). La variation d'énergie pour passer d'un état à un autre est donnée par la relation de Planck (Equation 5) :

Equation 5 $\Delta E = E_n - E_0 = h \cdot \nu$ avec

ΔE : variation d'énergie (J)

E_n : niveau d'énergie excité (J)

E_0 : niveau d'énergie fondamental

ν : fréquence (s^{-1})

h : Constante de Planck = $6,626 \cdot 10^{-34}$ J.s

Figure 9 : Exemples de modes de vibrations moléculaires
Source : Bertrand, 2002

Les liaisons atomiques ne possèdent pas qu'une seule fréquence de vibration et la distance entre les molécules varie constamment. Cela permet de distinguer le modèle de l'oscillateur harmonique de l'anharmonique qui prend en compte plusieurs fréquences de vibration et non une seule. Ces fréquences de vibration correspondent à différentes bandes d'absorption en infrarouge (Bertrand, 2002). Elles sont au nombre de trois (**Figure 10**) :

- Les bandes fondamentales, correspondant à l'absorption d'énergie nécessaire pour passer de l'état fondamental $v=0$ à l'état excité $v=1$. Elles sont observables dans le moyen infrarouge, c'est-à-dire à partir de 2500 nm.
- Les bandes harmoniques, correspondant à l'absorption d'énergie nécessaire à une transition vibrationnelle d'un état excité à un état encore plus excité. Par exemple, la transition de $v=1$ à $v=2$ est la première harmonique et celle de $v=2$ à $v=3$ la deuxième harmonique.
- Les bandes de combinaisons apparaissent à des fréquences comprises entre celles des harmoniques et des fondamentales, entre 1900 et 2500 nm, lorsque plusieurs liaisons chimiques proches sont excitées simultanément (Baeten et al., 2005 ; Bertrand, 2002 ; Dubuisson, 2009).

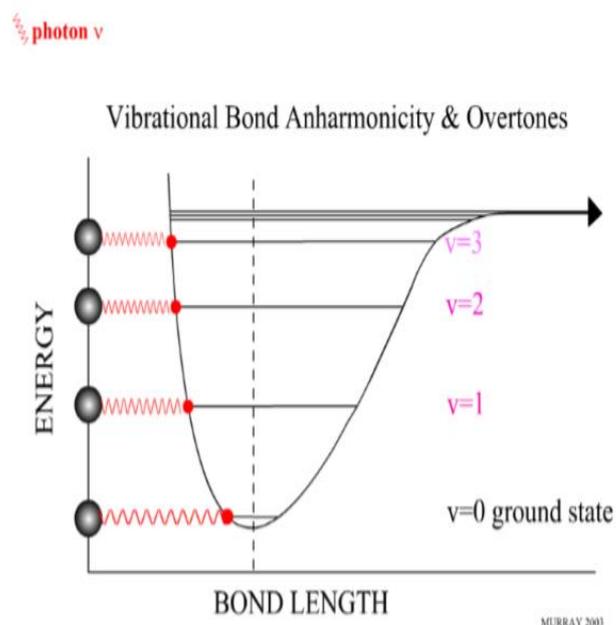


Figure 10 : Modèle de l'oscillateur anharmonique. Source : Lecler, 2014.

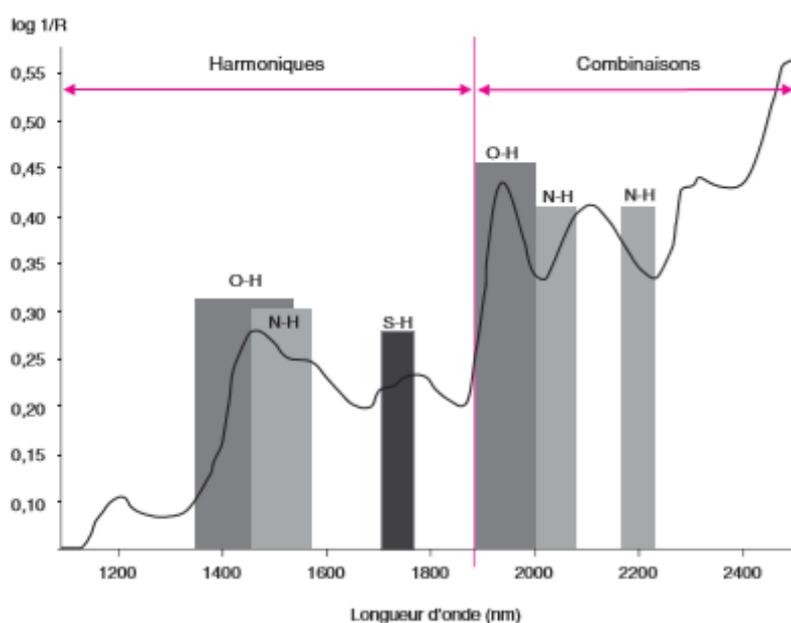


Figure 11 : Aperçu des bandes d'absorption dans le proche infrarouge. Les bandes fondamentales n'apparaissent qu'à partir de 2500 nm. Source : Bertrand, 2002.

Le proche infrarouge est le domaine des bandes harmoniques et des combinaisons d'harmoniques. Les liaisons chimiques dans cette région sont principalement asymétriques, soit de type X-H, où H est un atome d'hydrogène, et X, un atome de carbone C, d'oxygène O, d'azote N ou de soufre S (**Figure 11**) (Bertrand, 2002)

II.3. L'application de la spectrométrie proche infrarouge pour la prédiction en alimentation animale

La SPIR permet de contrôler la qualité des produits et déterminer leur composition chimique (Van Belleghem, 2003). Ainsi, il est courant d'analyser l'humidité afin d'exercer un contrôle sur l'activité en eau. L'Aw est le critère sanitaire principal en alimentation car elle est le témoin du risque de contamination par les moisissures. L'analyse des autres composants, à savoir les protéines, la matière grasse, les hydrates de carbones, permet de contrôler les aspects techniques des produits : apports en énergie et en azote, digestibilité, apports en minéraux (Baeten et al., 2005).

La SPIR apparaît comme une valeur sûre pour la prédiction de la composition et de la valeur alimentaire des matières premières et des fourrages (Bastianelli et Lecomte, 2003), tant en laboratoire que sur le terrain (NIR4FARM, n.d.). *FOSS Analytical* a réalisé des calibrations pour les protéines et l'humidité de tous les types de blé et d'orge, permettant de classer les grains de blé selon leur quantité de protéines et leur assigner un prix. Plus largement, lors de la phase de croissance et après récolte, les céréales sont soumises à divers facteurs : climat, parasites, techniques de récolte, transport, pouvant entraîner une modification des propriétés des grains, et la SPIR apparaît comme un moyen de détection efficace. (Williams, 2007). La SPIR peut également être utilisée pour prédire des « *critères d'appréciations indirects* » (Bastianelli et Lecomte, 2003). Cela peut concerner le rendement d'une prairie via l'analyse des fourrages, la contamination de la matière première par les mycotoxines ou encore l'évaluation de la valeur alimentaire des aliments grâce à l'analyse des fèces en mettant en parallèle leur composition et le régime alimentaire. (Baeten et al., 2005).

II.4. Les avantages de la SPIR

Il s'agit d'une méthode non destructrice, peu coûteuse, rapide (résultats en quelques secondes) et en continu dans le cas d'une analyse en ligne. Elle se révèle également fiable et d'une précision semblable aux méthodes chimiques de référence et ne nécessite que peu voire aucune préparation de l'échantillon. C'est par ailleurs une méthode flexible puisqu'elle permet l'analyse simultanée de plusieurs constituants. Autre avantage, la SPIR ne présente aucun danger pour l'environnement et le personnel car aucun produit chimique n'est utilisé. Enfin, l'appareillage est peu encombrant, durable (plus de 10 ans), et sans danger d'utilisation (Van Belleghem, 2003 ; Williams, 2007).

II.5. Les contraintes de la SPIR

Elle comporte néanmoins quelques désavantages. La SPIR est une méthode indirecte. Un calibrage (ou étalonnage) est ainsi nécessaire pour chaque produit, au contraire des analyses chimiques et elle ne permet pas de détecter directement des composés en faible concentration et les minéraux. Ils se déduisent de l'analyse des groupements CH, OH et NH. Il est de surcroît nécessaire de contrôler régulièrement la précision et la reproductibilité de l'appareillage. Ce dernier est par ailleurs coûteux à l'achat, et un temps d'apprentissage est nécessaire pour comprendre son fonctionnement (Van Belleghem, 2003 ; Williams, 2007).

III. L'appareillage

III.1. Généralités

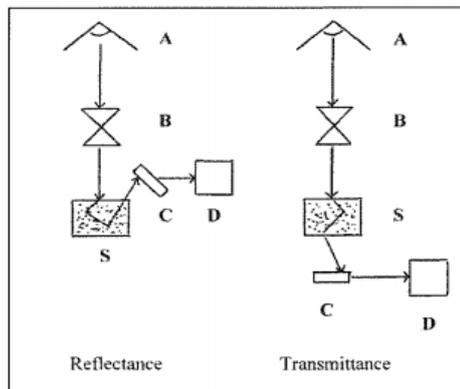


Figure 12 : Configuration des instruments de spectrométrie
A = source lumineuse ; B = système de sélection de longueur d'onde ; C = détecteur ; D = enregistreur ; S = échantillon
 Source : Williams, 2007.

Tous les spectromètres dans le proche infrarouge sont construits sur le même principe. Ils sont composés de quatre éléments : une source lumineuse, un système de sélection des longueurs d'onde, un détecteur et un enregistreur. Le spectromètre peut fonctionner selon deux configurations différentes selon leur mode de détection : en transmittance ou en réflectance (Williams, 2007). En réflectance, la source de lumière et le détecteur sont positionnés du même côté de l'échantillon. En transmittance, la radiation lumineuse est transmise à travers l'échantillon et atteint le détecteur placé de l'autre côté de l'échantillon (Williams, 2007). La **Figure 12** présente les différences observées entre les instruments mesurant en transmittance et en réflectance.

Dans le cadre de ce mémoire, le mode de mesure retenu est la réflectance (NIR ou Near-Infrared Reflectance).

Il est possible de distinguer trois grandes familles de spectromètres analysant le proche infrarouge

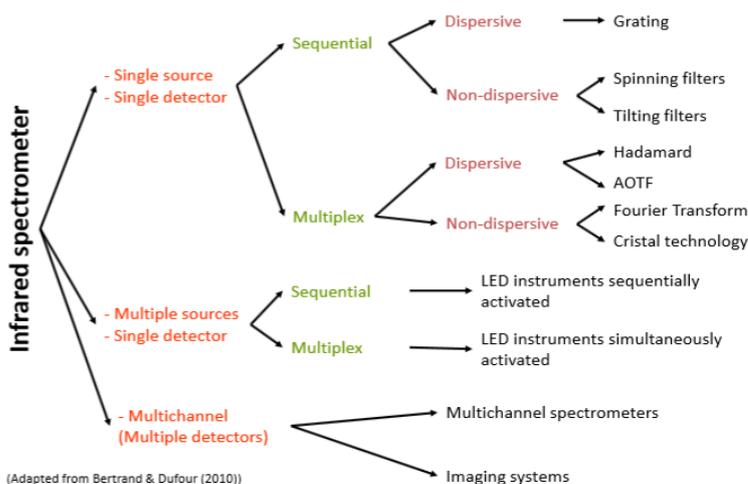


Figure 13 : Classification des spectromètres proche infrarouge
AOTF = acoustico-optico-tunable filter (filtres acoustico-optique modulable) ;
LED = light emitting diodes (diodes émettrices de lumière)
 Source : Lecler, 2014.

(**Figure 13**). Tout d'abord, les appareils séquentiels mesurent une à une les absorptions lumineuses en balayant le proche infrarouge dans un ordre séquentiel. Les spectromètres séquentiels les plus communs sont les spectromètres à filtres et à monochromateur. Il existe également des spectromètres multiplexés tels que les spectromètres à transformée de Fourier, où un capteur unique reçoit les signaux de plusieurs longueurs d'onde. Enfin, les spectromètres multicanaux possèdent plusieurs capteurs pouvant enregistrer l'absorption de plusieurs longueurs d'onde de

manière indépendante. Les spectromètres multicanaux les plus courants sont les systèmes à barrettes de diodes et les caméras (Bertrand, 2002).

III.2. Configuration instrumentale

III.2.1. La source lumineuse

La tête de lecture du spectromètre est composée d'une source lumineuse. En spectrométrie proche infrarouge, la plus utilisée est une lampe halogène à filament de tungstène. Elle émet un rayonnement continu dont les longueurs d'onde s'étendent de l'ultraviolet au moyen infrarouge (Dubuisson, 2010).

III.2.2. Le système de sélection des longueurs d'onde

Deux systèmes de sélection des longueurs d'onde ont été rencontrés au cours du projet.

III.2.2.1. Les systèmes à barrettes de diodes

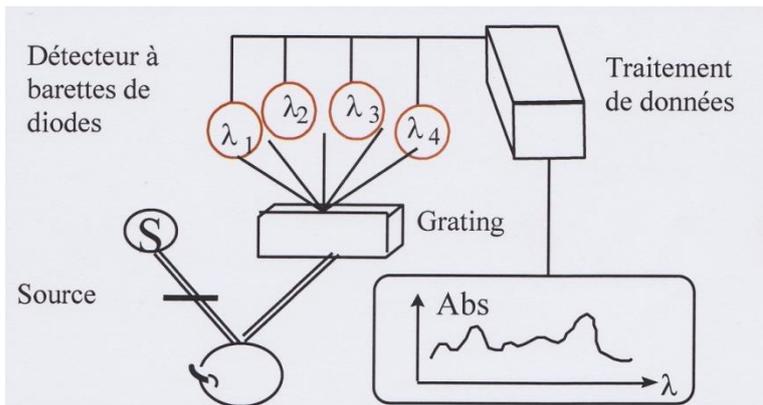


Figure 14 : Principe d'un système à diodes émettrices de lumière
Source : Van Belleghem, 2003.

En effet, certains appareils permettent de réaliser un spectre en quelques millisecondes (Bertrand, 2002 ; Biemar, 2014).

La radiation transmise ou réfléchiée est dispersée à l'aide d'un monochromateur composé d'une série de LED (« light emitting diode » ou diode émettrice de lumière). Couplées à un filtre, elles génèrent différentes intensités de lumière correspondant à des longueurs d'onde spécifiques (Figure 14). Ne possédant aucune pièce mobile, ce type de spectromètre, très robuste, est particulièrement adapté au contrôle de routine en ligne et est de surcroît très rapide.

III.2.2.2. Les interféromètres à transformée de Fourier

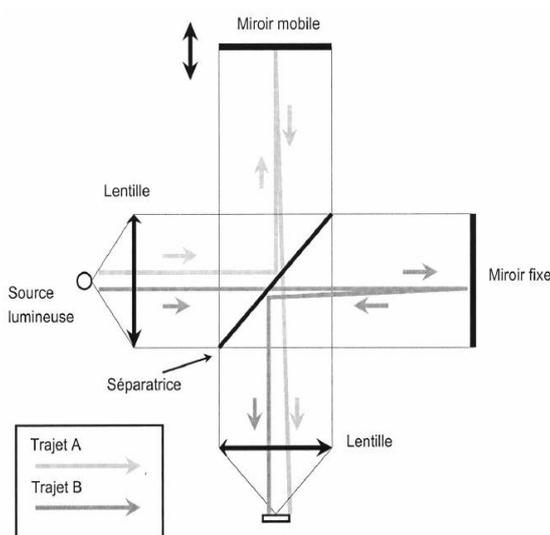


Figure 15 : Principe d'un interféromètre de Michelson (spectromètre à transformée de Fourier)
Source : Bertrand, 2002.

Le dispositif principal de ce type de spectromètre est l'interféromètre de Michelson (Figure 15). Dans cette configuration le rayon lumineux suit deux trajectoires différentes. La lumière atteint une séparatrice. Une moitié est réfléchiée (trajet A) et l'autre moitié est transmise (trajet B). Le rayon correspondant au trajet B atteint un miroir fixe, est renvoyé vers la séparatrice puis atteint à son tour le détecteur. Le rayon correspondant au trajet A atteint un miroir mobile. Il revient ensuite à la séparatrice puis atteint le détecteur.

Le miroir mobile permet de modifier la longueur du trajet optique du rayon A. Lorsque les trajets A et B sont de même longueur, l'interférence est dite constructive. Elle est également considérée comme

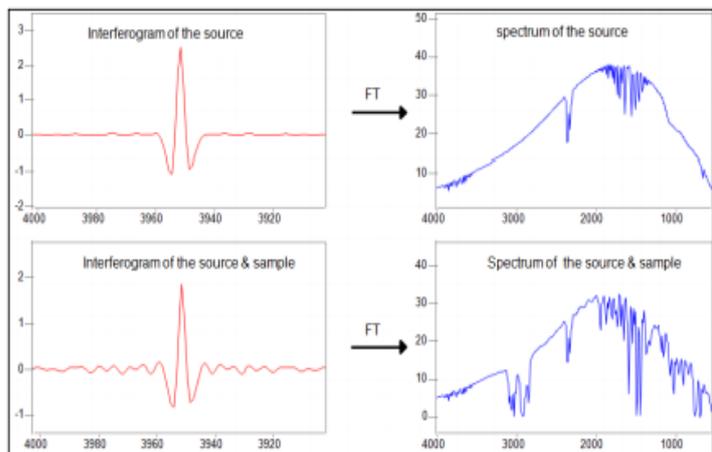


Figure 16 : Exemple de spectres infrarouges par transformée de Fourier
Source : Fatela, 2012.

constructive lorsque pour un rayon de longueur d'onde λ , la différence des trajets δ est un multiple de λ . Au contraire, les rayons interfèrent de façon destructive lorsque $\delta = n \cdot (\lambda/2)$ avec n un entier naturel.

Lorsque qu'un rayon monochromatique atteint le détecteur, un signal électrique sinusoïdal est enregistré. Dans le cas d'une lumière polychromatique, le signal est appelé interférogramme. Si la différence de longueur entre les deux trajets optiques est nulle,

l'interférogramme prend sa valeur maximale. A l'inverse, si δ augmente, alors l'interférogramme tend vers zéro. Ainsi, l'interférogramme correspond à la somme de toutes les fonctions sinusoïdales et sa transformée de Fourier permet d'obtenir des intensités en fonction des fréquences (**Figure 16**) (Biemar, 2014 ; Fatela, 2012).

III.2.3. Le détecteur

Le détecteur mesure la quantité d'énergie lumineuse passant à travers l'échantillon et transforme le signal lumineux en un spectre (Dubuisson, 2009). La qualité d'un détecteur dépend de trois critères (Roggo, 2003) :

- L'efficacité quantique : l'intensité du signal électrique émis doit être proportionnelle à l'intensité de l'énergie reçue.
- La sensibilité : afin d'obtenir des résultats précis, le détecteur doit être le plus sensible possible car les absorbances dans le proche infrarouge sont en général très faibles.
- La constante de temps : en cas de variation de l'intensité de l'énergie reçue, le temps de réponse du détecteur doit être de l'ordre de la milliseconde.

III.2.4. L'enregistreur

L'enregistreur est communément un ordinateur, permettant de convertir le signal analogique du détecteur en un signal binaire, exploitable par les logiciels de traitement des spectres bruts.

III.3. Les facteurs influant sur les mesures spectrales

Lors de l'utilisation d'un spectromètre, plusieurs paramètres physiques et/ou chimiques peuvent modifier un spectre infrarouge (Fatela, 2012 ; Lecler, 2014 ; Williams, 2007) :

- Le spectromètre : passer plusieurs fois le même échantillon permet d'évaluer la répétabilité de l'appareil et son bruit de fond. Le bruit de fond correspond à la différence observée entre deux

mesures consécutives. Plus il est faible, plus le spectromètre est de qualité. Dans le cas d'absorbances faibles, tel que dans le proche infrarouge, le bruit de fond doit être le plus faible possible afin de pouvoir étudier les variabilités inter-échantillons.

- La distance focale : elle correspond à la distance entre l'échantillon et la tête de lecture du spectromètre. La distance optimale est comprise entre 15 et 20 cm, mais dépend de la nature de l'échantillon (épaisseur, forme). Elle ne doit cependant jamais dépasser 50 cm.
- La température : les variations influent sur les liaisons hydrogènes, entraînant une modification de l'agitation moléculaire, et à fortiori de l'allure spectrale. Une dérive au niveau des extrema est alors observée.
- L'environnement : si le spectromètre est placé au dessus de l'échantillon, les particules de l'air peuvent interférer sur la mesure.

Ces facteurs de variations peuvent entraîner des déplacements ou des élargissements des bandes spectrales, du bruit aléatoire ainsi que des modifications au niveau de l'absorbance, conduisant à des erreurs de mesure et de quantification (Fatela, 2012 ; Lecler, 2014).

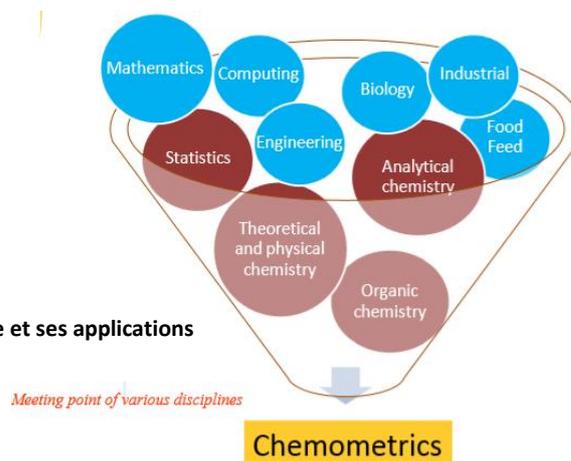
IV. Le traitement des données : la chimiométrie appliquée à la SPIR

IV.1. Définition de la chimiométrie

Le professeur D.L. MASSART a décrit la chimiométrie comme étant *“the chemical discipline that uses mathematics and statistics to design or select optimal experimental procedures, to provide maximum relevant chemical information by analyzing chemical data, and to obtain knowledge about chemical systems”* (Lecler, 2014).

Ainsi, la chimiométrie apparaît comme une science multidisciplinaire applicable dans de nombreux domaines : biologie, agronomie, agroalimentaire (**Figure 17**).

Figure 17 : Définition de la chimiométrie et ses applications
Source : Lecler, 2014.



L'application des traitements chimiométriques suit un ordre logique. Dans un premier temps, il est possible d'avoir recours à une analyse qualitative via une analyse en composantes principales (ACP) permettant de comparer les spectres bruts entre eux et d'observer la variabilité du set. La création des équations de calibration proprement dite débute après cette étape facultative par l'application de prétraitements aux spectres, suivi des phases classiques de calibration et de validation des modèles.

IV.2. Analyse qualitative : l'analyse en composantes principales (ACP)

L'ACP permet de décrire les variations observées dans les spectres à l'aide d'un nombre restreint de variables non corrélées : on parle d'axes principaux ou de composantes principales. Ainsi, on combinant les variables d'origine dans un certain nombre de ces nouvelles variables afin de garder le maximum de la variation spectrale présente, on réduit le nombre de dimensions afin de faciliter la lecture des données et trouver de nouvelles corrélations entre les données (Lecler, 2014).

L'ACP crée de nouvelles matrices afin d'avoir un aperçu sur la dispersion, les tendances, les divers groupes d'échantillons et les outliers. Ainsi, les *scores* et les *loadings* peuvent être vus dans les graphiques linéaires et de dispersion. Les *scores* correspondent à la projection des données sur les composantes principales. Généralement, deux ou trois axes suffisent à décrire la majorité de la variation spectrale. Les *loadings* permettent d'indiquer le poids des principaux axes selon la variation expliquée (Dubuisson, 2009).

IV.3. Les prétraitements des spectres

Suite aux interférences des facteurs externes décrits dans le paragraphe III.3. *Les facteurs influant sur les mesures spectrales*, plusieurs traitements mathématiques, appelés prétraitements (Dubuisson, 2009), peuvent être appliqués aux spectres bruts afin d'augmenter la précision de l'analyse et d'améliorer les données spectrales (Van Belleghem, 2003). Ces prétraitements permettent de supprimer les effets des biais d'ordre non-chimique et de préparer les données spectrales en vue des traitements mathématiques et statistiques ultérieurs. Leur avantage réside dans le fait qu'ils ne changent ni n'affectent le spectre brut d'origine et dépendent uniquement des améliorations que l'on souhaite apporter au signal (Lecler, 2014).

IV.3.1. La moyenne spectrale

La moyenne spectrale est obtenue en scannant plusieurs fois un même échantillon et en faisant la moyenne de ceux-ci afin de recueillir un spectre moyen. Le nombre de répétitions dépend de la nature de l'échantillon et de sa variabilité (Roggo, 2003). Elle permet de réduire le bruit de fond et d'améliorer le ratio signal/bruit (Lecler, 2014).

IV.3.2. La normalisation des spectres : traitement par SNV (Standard Normal Variate)

Le traitement par SNV vise à supprimer les effets des interférences dues à la dispersion et à la taille des particules de l'échantillon. Le SNV standardise individuellement chaque spectre en soustrayant leur moyenne et en les mettant à l'échelle en fonction d'un écart-type de 1,0 (Lecler, 2014) selon l'équation 8. La **Figure 18** présente les effets d'une normalisation SNV en comparant les spectres bruts originaux et les spectres obtenus après SNV. Une partie des différences d'ordre non chimique est corrigée, les spectres sont plus rapprochés et adoptent une allure plus facile à interpréter.

Equation 8
$$X_{ij(SNV)} = \frac{(x_{ij} - \bar{x}_i)}{\sqrt{\frac{\sum_{i=1}^p (x_{ij} - \bar{x}_i)^2}{p-1}}}$$
 avec

x_{ij} : valeur de l'absorbance à longueur d'onde i

p : nombre de longueurs d'onde

\bar{x}_i : valeur moyenne de l'absorbance

$X_{ij(SNV)}$: valeur corrigée à la longueur d'onde i

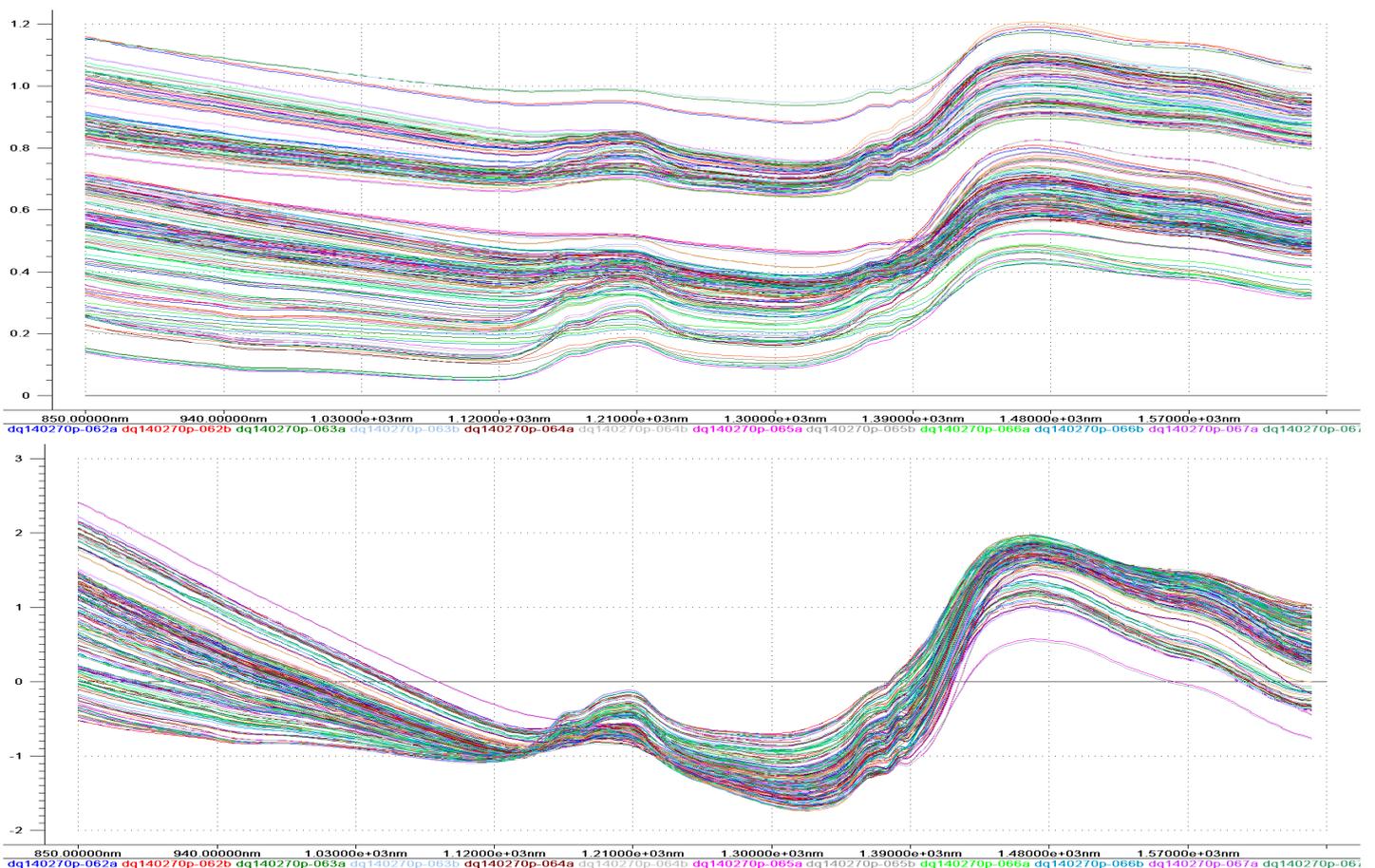


Figure 18 : Effets d'une normalisation SNV. En haut : spectres bruts. En bas : spectres traités par SNV. En abscisse : longueur d'onde (nm) ; en ordonnée : absorbance (log 1/R).

IV.3.3. Les dérivées spectrales

Certaines différences sont trop faibles entre les spectres et demeurent presque invisibles. Le prétraitement par les dérivées spectrales permet de réduire la déviation de la ligne de base et fait apparaître davantage de variations entre les spectres (Lecler, 2014).

La dérivée première (**Figure 19**) consiste à étudier la pente d'une courbe et permet de révéler de nouvelles pentes qui pouvaient être cachées par une pente plus importante due à une absorption plus importante. Elle augmente ainsi les variations dans les spectres et dévoilent davantage de différences spectrales. Chaque maxima du spectre original se retrouve au point 0 sur l'axe des ordonnées, et un pic au niveau de la dérivée correspond à un changement de pente.

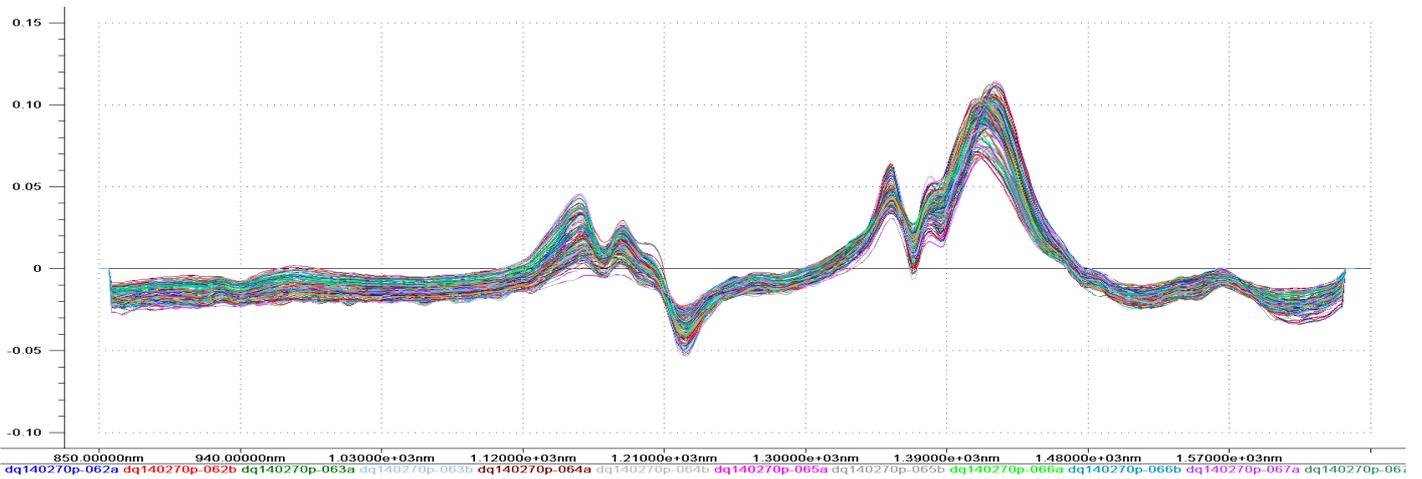


Figure 19 : Transformation en dérivée première d'un spectre traité par SNV. En abscisse : longueur d'onde (nm) ; en ordonnée : absorbance ($\log 1/R$)

La dérivée seconde (Figure 20) consiste à mesurer les variations de la pente de la courbe spectrale. Elle permet de supprimer une ligne de base inclinée à partir d'un spectre. La dérivée seconde est plus facile à interpréter car à chaque pic négatif correspond un pic positif du spectre d'origine.

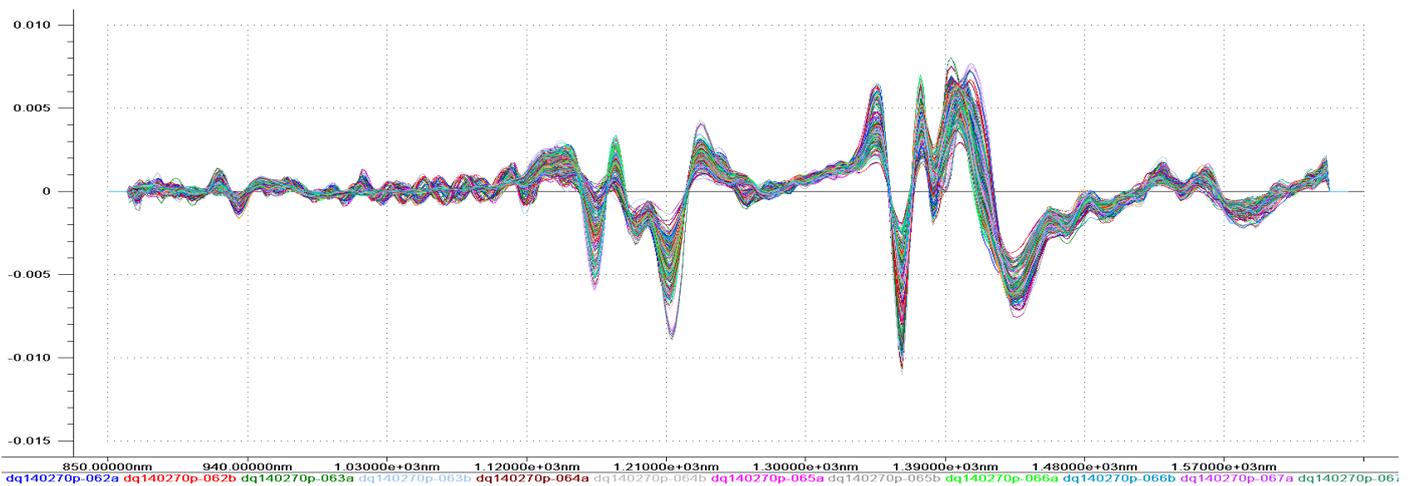


Figure 20 : Transformation en dérivée seconde d'un spectre traité par SNV. En abscisse : longueur d'onde (nm) ; en ordonnée : absorbance ($\log 1/R$)

IV.3.4. Le lissage spectral

Le lissage spectral (*Smoothing*) est généralement appliqué sur les spectres dérivés (Roggo, 2003). Les dérivées spectrales diminuent le rapport du signal par rapport au bruit (Lecler, 2014). Afin de réduire le bruit de fond des spectres, un smoothing est réalisé à l'aide de l'algorithme de Savitsky-Golay qui consiste à remplacer le point central d'un segment de longueur d'onde par la moyenne de tous les points de ce segment. On parlera alors de moyenne mobile.

IV.4. La calibration

La phrase de calibration consiste à établir une relation entre les valeurs chimiques de référence et les valeurs du spectre infrarouge. Pour chaque type d'échantillon possédant le maximum de variabilité, une calibration s'avère nécessaire. Pour développer et valider un modèle, il est nécessaire de se constituer plusieurs sets d'échantillons : un set pour la calibration et un set pour la validation à partir d'un set général. La sélection des échantillons peut s'effectuer de manière aléatoire, à l'aide de la

méthode de Kennard et Stone ou dans l'objectif d'avoir le maximum de variabilité (Lecler, 2014). Une fois le set de calibration retenu, la calibration peut s'effectuer l'aide de méthodes de régression.

IV.4.1. Les méthodes de régression

Plusieurs méthodes de régression existent, de la régression linéaire simple ou multiple à la régression en composantes principales (PCR). Mais la méthode la plus utilisée demeure la régression par les moindres carrés partiels (PLS). Elle combine les caractéristiques de la régression en composantes principales et de la régression linéaire multiple. Elle permet de prédire un ensemble de variables dépendantes (critères d'analyse) d'un grand nombre de variables indépendantes (données spectrales). La méthode PLS fournit un modèle général qui englobe à la fois régression linéaire multiple, analyse multivariée de la variance et analyse en composantes principales. Elle est par ailleurs utilisable avec un set d'échantillons de petite taille (CAMO, n.d.).

IV.4.2. Les paramètres d'appréciation de la calibration

Deux paramètres statistiques permettent d'évaluer la qualité et la robustesse d'une équation de calibration. L'écart-type résiduel de calibration, ou SEC, correspond à la moyenne quadratique des écarts des points par rapport à la droite estimée. Le SEC doit être le plus faible possible et est donné par la formule suivante (Lecler, 2014) (Equation 11) :

$$\text{Equation 11} \quad \text{SEC} = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n - p - 1}} \quad \text{avec}$$

y_i = valeur de référence (chimique)

n = nombre d'échantillons

\hat{y}_i = valeur prédite par le spectromètre

p = nombre de termes de l'équation

On peut également parler de RMSEC si l'on prend en compte le biais de la mesure. Dans certains cas, aucun biais n'existe et le SEC et le RMSEC sont identiques.

Enfin le coefficient de détermination R^2 exprime la part de variation des valeurs de référence expliquée par l'équation de calibration. Il doit être le plus proche de 1. Il se calcule par la formule (Equation 12) :

$$\text{Equation 12} \quad R^2 = \frac{SDy^2 - SEC^2}{SDy^2} \quad \text{avec } SDy^2 \text{ la variance de la variable à expliquer}$$

La complexité de l'équation dépend du nombre de termes (composantes principales). Augmenter le nombre de composantes principales permet d'expliquer une plus grande part de la variation et ainsi augmenter le R^2 et diminuer le SEC. Il faut néanmoins éviter de prendre un trop grand nombre de termes au risque d'avoir un surparamétrage de l'équation et un SEC trop faible.

IV.5. La validation

L'équation de calibration doit ensuite être validée. Deux possibilités existent : soit utiliser un set de validation dont les échantillons n'ont pas contribué à la calibration (set d'échantillons indépendants), soit utiliser la procédure de cross-validation. La cross-validation utilise les échantillons ayant servi à la

calibration mais d'une manière alternative. Cela consiste à supprimer du set de validation un échantillon ou un groupe d'échantillons, puis de construire le modèle avec les échantillons restants. Il ne reste ensuite qu'à prédire l'échantillon ou le groupe d'échantillons et répéter cette opération jusqu'à ce que tous les échantillons ou groupes d'échantillons aient été supprimés et prédits une fois (Lecler, 2014).

La qualité de la validation est alors appréciée à l'aide de l'écart-type standard de cross-validation (SECV) en cas d'application d'une procédure de cross-validation ou de l'écart-type standard de prédiction (SEP) si le set de validation est indépendant de la calibration (Lecler, 2014). Il représente la déviation standard des différences observées entre les valeurs prédites par le spectromètre et les valeurs chimiques de référence (Williams, 2007). Elle se calcule par la formule (Equation 13) :

$$\text{Equation 13} \quad \text{SEP} = \text{SECV} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}}$$

On peut également parler de RMSEP ou RMSECV si l'on prend en compte le biais de la mesure. Dans certains cas, aucun biais n'existe et le SEP ou SECV et le RMSEP ou RMSECV sont identiques. La valeur du SEP ou SECV doit être proche de celle du SEC. Si l'écart est trop important, cela indique une mauvaise calibration. De même que lors de la calibration, il est possible d'améliorer le SEP/SECV et le R^2 en jouant sur le nombre de coefficients de l'équation.

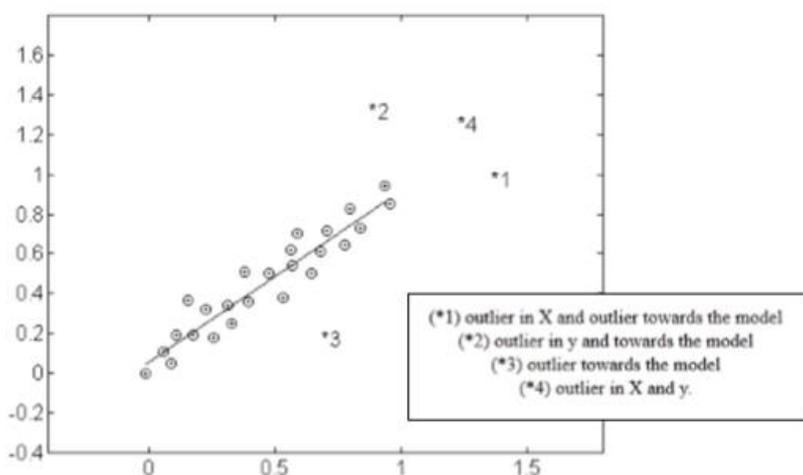


Figure 21 : Aperçu des différents outliers existants.
Source : Lecler, 2014.

C'est également lors de l'étape de validation du modèle que la suppression des outliers est intéressante. Les outliers sont des points qui se révèlent être non pertinents, incorrects ou anormaux, en comparaison de la majorité des données (Williams, 2007). Malgré leur faible nombre, ils ont souvent une forte incidence sur l'établissement du modèle. Ces points peuvent survenir suite à des erreurs de manipulation ou

inhérentes au processus d'analyse. Dans ce cas là, on considère qu'il s'agit de mauvais outliers. Ils doivent donc être éliminés du modèle. Cependant, certaines valeurs aberrantes peuvent n'être que des valeurs extrêmes, et peuvent apporter de la robustesse au modèle, la gamme de valeur étant alors plus large. Ces points sont alors considérés comme de bons outliers (Lecler, 2014). Les outliers en X et Y doivent être détectés avant la modélisation afin que le modèle ne soit pas biaisé par ces valeurs aberrantes. Une fois le modèle créé, les outliers vers le modèle doivent être supprimés afin de supprimer les biais éventuels (Figure 21) (Lecler, 2014).

La détection des outliers s'effectue généralement lors de l'étape de la cross-validation ou la validation du modèle, les prétraitements effectués mettant davantage en évidence les variabilités entre les échantillons. Toutefois, il arrive que des outliers puissent être détectés en amont si certains spectres s'écartent de la moyenne, ou lors de l'ACP. En supprimant les échantillons aberrants, cela suffit parfois à obtenir un SEP/SECV et un R^2 intéressants.

Un dernier paramètre intéressant à prendre en compte est le RPD (Ratio of Prediction (ou Performance) to Deviation). Il se calcule en divisant le SD (Standard Deviation) des valeurs de référence utilisées lors de la validation, par le SEP ou le SECV. Si le SEP est égal ou supérieur au SD, alors le modèle ne prédit pas bien les valeurs. Un RPD faible peut également provenir d'un SD faible en raison d'une variance faible. Dans ce cas, une analyse infrarouge n'est pas nécessaire et un simple contrôle qualité s'avère suffisant. Le RPD doit être au minimum égal à 3, et idéalement égal à 5 ou plus (Williams, 2007).

Partie Expérimentale

I. Matériel et méthode

Cette partie est consacrée à la présentation du matériel utilisé au cours du projet, ainsi qu'à la démarche mise en place afin d'obtenir les spectres infrarouges, traiter les données chimiques recueillies, et installer les équations de calibration.

I.1. Instrumentation

I.1.1. Le spectromètre

Au cours du projet, deux spectromètres ont été utilisés. Chacun possède des caractéristiques différentes et les avantages qu'ils offrent ont été exploités afin que les équations de calibration créés à partir des spectres recueillis soient les meilleures possibles.

Le spectromètre utilisé à Andenne est un spectromètre à barrettes de diodes de la marque Polytec. Deux têtes de lecture ont été installée : une en laboratoire pour les analyses de produits finis, de matières premières et de routine, et une tête de lecture au niveau de la trémie après mélangeuse pour les analyses on-line.

En parallèle, chaque échantillon a été passé sur le spectromètre XDS FOSS du CRA-W. Il s'agit d'un spectromètre utilisant la transformée de Fournier. Le XDS est un appareil très puissant et donnant des prédictions très proches des valeurs chimiques de référence. Cette caractéristique a été démontrée lors de l'analyse d'aliments extrudés. La démarche est présentée en **Annexe 2**. Cela constitue un double avantage en termes de temps et d'argent, la majorité des analyses chimiques pouvant être remplacées par les prédictions XDS. Celles-ci se révèlent toutefois toujours indispensables car une valeur chimique sera toujours plus précise qu'une prédiction. Les spectres obtenus grâce au XDS ont été ajoutés aux bases de données du CRA-W, composées de plusieurs milliers d'échantillons, qui a ensuite utilisé une de ses équations afin de prédire les valeurs des différents critères d'analyse pour chaque échantillon, sur la base des valeurs de référence obtenues via l'analyse chimique. Ces valeurs prédites ont ensuite été ajoutées aux valeurs chimiques de référence et incrémentées aux spectres obtenus sur le Polytec. Les caractéristiques des deux spectromètres sont reprises en **Tableau 1**.

Appareil Caract.	Polytec (Analyse à distance)	Polytec (Analyse de contact)	XDS (FOSS)
Longueur d'onde de travail	850-1650 nm	850-1650 nm	400-2500 nm
Présentation de l'échantillon	Mesure à distance (15-50 cm)	Mesure au contact de l'échantillon	coupelle
Faisceau par le haut/le bas ?	Haut	Bas	Bas
Technologie	instrument à barrettes de diodes	instrument à barrettes de diode	instrument à dispersion (transformée de Fourier)
Vitesse d'acquisition	1500 scan/s	1500 scan/s	64 scan/min
Appareil pour labo ou online	Labo	Online	Labo

Tableau 1 : Caractéristiques techniques des appareils Polytec et XDS FOSS

Les analyses chimiques de référence ont été obtenues auprès du laboratoire *Feed and Food Quality* de Merckem. Sur la base des coefficients de répétabilité des critères d'analyse retenus, le laboratoire donne de bonnes valeurs chimiques (**Tableau 2**) (Biemar, 2014).

Paramètre	Méthode	RSDR (%) (répétabilité)	Moyenne (%)	SDR (écart-type)
Humidité	Méthode 402 : 4h à 103°C	1,4	12	0,168
Protéine Brute	Méthode 423 : Dumas	1,0	35	0,35
Cendre Brute	Méthode 431 : 4h à 550°C	2,3	10	0,23
Cellulose Brute	Méthode 442	6,2	10	0,62
Matière Grasse Brute	Méthode 452 : Soxhlet après extraction à l'éther	1,9	20	0,38
Amidon	Méthode 445 : méthode polarimétrique (Ewers)	0,8	50	0,4

Tableau 2 : Méthodes d'analyse chimique et coefficients de répétabilité



Figure 22 : Tête de lecture on-line Polytec
Source : Biemar, 2014.

Au niveau on-line, la tête de lecture est reliée par fibre optique à l'enregistreur. Le dispositif doit être placé de façon à être en contact direct avec la matière. La mesure s'effectue ainsi en réflectance et la source lumineuse retenue est une lampe au tungstène-halogène (**Figure 22**). Les premières simulations ont été réalisées au CRA-W, qui a mis au point un dispositif expérimental permettant de simuler manuellement le passage de la matière (**Figure 23**).

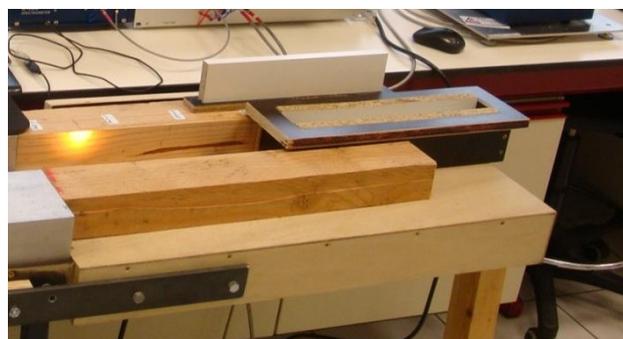
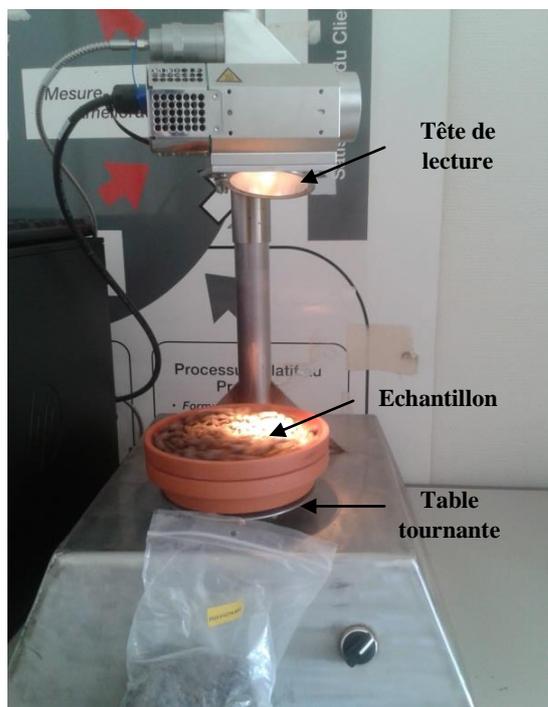


Figure 23 : Dispositif expérimental du CRA-W simulant l'analyse on-line
Source : Biemar, 2014.



Pour les mesures en laboratoire, une tête de lecture différente a été installée. La radiation lumineuse s'effectue par le haut. Le dispositif est composé d'un pied tenant la tête de lecture au-dessus de l'échantillon. L'appareil Polytec installé à l'usine d'Andenne possède une distance focale de 27 cm. L'échantillon est placé sur une table tournante, permettant ainsi d'analyser une plus grande surface de ce dernier (**Figure 24**).

Figure 24 : Dispositif d'analyse du Polytec en laboratoire

1.1.2. Les logiciels de traitements des données spectrales et de prédiction

Le logiciel WinISI (FOSS, Infracore-International) permet de visionner les spectres bruts obtenus après analyse sur le spectromètre. Il assure en outre le management de fichiers, permettant de rentrer les valeurs de référence pour chaque paramètre et de prédire les valeurs théoriques à partir des spectres sur la base des valeurs chimiques de référence.

Le logiciel UNSCRAMBLER version 9.8 de la société CAMO SOFTWARE est ensuite utilisé pour la phase de calibration. Les spectres bruts et les valeurs prédites y sont dans un premier temps transférés avant que les premiers prétraitements (SNV, dérivées) ne soient appliqués. L'étape suivante consiste à construire les courbes à l'aide des méthodes de régression.

Les modèles de chaque paramètre d'analyse ainsi créés peuvent ensuite être transférés sur le logiciel de prédiction OLUP du Polytec de *Dumoulin* afin de prédire les valeurs chimiques d'échantillons inconnus ou de valider les modèles.

1.2. Echantillonnage et définition des paramètres chimiques

Dumoulin fabrique des aliments pour de nombreuses espèces animales (bovins, caprins, ovins, porcins, volailles), et chaque gamme est déclinée en des dizaines de produits. Ainsi, sélectionner l'intégralité des produits était irréalisable pour créer les courbes de calibration. Le choix des produits a donc été réalisé selon deux paramètres : le volume de vente et la variabilité. En effet, un produit dont les ventes sont importantes confère deux avantages : un volume de production conséquent permet de rapidement constituer une base d'échantillons, ainsi qu'une meilleure observation des déviations d'une production à l'autre et une correction rapide des biais. Enfin, le CRA-W a conseillé de sélectionner des échantillons ayant un maximum de variabilité afin d'élargir la plage de valeurs. Même en disposant d'un grand nombre d'échantillons, si les valeurs chimiques sont identiques pour un même critère d'analyse, la calibration ne sera pas robuste. La fiabilité des valeurs prédites par l'équation dépend donc de la variabilité du set d'échantillons.

Au-delà de la variabilité inter-échantillons, il était important de vérifier la répétabilité de la mesure. C'est pourquoi deux mesures ont été réalisées sur chaque échantillon (A et B). Entre deux mesures, l'échantillon a été retiré de la cellule d'analyse, puis replacé, afin de ne pas scanner deux fois la même partie de l'échantillon. Cependant, une moyenne spectrale n'a pas été réalisée afin d'obtenir un spectre moyen (le Polytec prenant 1500 scans/secondes et en faisant la moyenne), et la calibration a été effectuée avec des échantillons en double exemplaires.

Par ailleurs, le CRA-W a également mis à disposition de *Dumoulin* des numéros uniques (DQ), afin d'établir une liste de correspondances avec les échantillons analysés en fonction de leur nature (extrudés, flocons, matières premières, etc.), facilitant ainsi le management de fichiers. La répétition des mesures n'apparaît pas dans ces listes de correspondances.

1.2.1. Les extrudés

Lors de la production d'aliments extrudés, les graines sont tout d'abord broyées pour être réduites à l'état de farine. Elles sont ensuite passées dans un conditionneur où sont ajoutés les liquides à incorporer. Le mélange est ensuite monté en température (100°C) via l'injection de vapeur d'eau. La matière passe ensuite dans l'extrudeur, où elle s'écoule au travers d'un orifice de petit diamètre. Sous l'action de fortes pressions obtenues à l'aide d'une vis d'Archimède à spire variable, elle est comprimée et cisailée à une température de 130°C (Roczen, 2014). La matière est alors soumise à une cuisson, et l'eau contenue à l'intérieur se vaporise instantanément une fois repassée à pression atmosphérique normale. Cette chute brutale de pression provoque l'expansion de la matrice et lui donne une structure plus aérée. Le produit est ensuite transféré dans un refroidisseur/sécheur afin de réduire son taux d'humidité (Nutridev, n.d.). La dernière étape est le conditionnement des aliments extrudés en sacs ou en vrac, ou l'acheminement vers l'usine des mix pour les intégrer aux flocons et aux granulés (Roczen, 2014). (**Annexe 3**).

Les aliments extrudés analysés peuvent être séparés en deux catégories selon leur teneur en matière grasse : les extrudés possédant de 15 à 35% de matières grasses et les extrudés possédant moins de 15% de matières grasses.

Cinq critères d'analyse ont été retenus : humidité (teneur en eau), protéines, matières grasses (FAH avec hydrolyse ou méthode B), cellulose brute et amidon. Il a été décidé que les courbes seraient établies à partir de graines de lin, dont le CRA-W possède des valeurs chimiques pour tous les paramètres excepté l'amidon.

L'humidité est un paramètre important à suivre car c'est un critère technique. Le suivi du taux d'humidité est primordial car l'étape de séchage est coûteuse. Cependant, si l'humidité demeure trop élevée, la conservation des produits s'en retrouve directement impactée. Si la teneur en eau et l'activité de l'eau sont trop élevées, cela permettra le développement de moisissures et impactera la bonne conservation des aliments.

Les teneurs en protéine et en matière grasse constituent des critères d'achat. L'apport en protéine est un besoin essentiel chez l'animal et assure sa croissance et sa productivité. Ainsi, en cas de chute de production de lait chez une vache laitière, la teneur en protéine de l'aliment consommé est directement remise en cause. Enfin, la matière grasse est une des principales sources d'énergie et permettent aux animaux de se constituer des réserves, notamment chez les monogastriques. La matière grasse est également un précurseur d'éléments présents dans les produits finis, tels que les acides gras oméga 3. On distingue deux méthodes d'analyse de la matière grasse en fonction de la concentration de matière grasse liée. Pour les produits contenant plus de 8% de matière grasse, la technique consiste à appliquer le procédé d'extraction B (OIB) par hydrolyse. On parlera alors de FAH. Cette méthode sera utilisée pour l'analyse des aliments extrudés. Pour les produits contenant moins de 8% de matière grasse, le procédé A (OIA) sans hydrolyse sera appliqué. On parlera alors de FAT.

La cellulose brute et l'amidon ont également été analysés même s'ils sont considérés comme secondaires par rapport aux critères précédents. L'amidon constitue la première source d'énergie pour les animaux et le processus d'extrusion entraîne sa gélatinisation. Le contrôle de ce paramètre s'avère donc néanmoins important pour *Dumoulin*.

Ces cinq critères d'analyse seront repris pour la majorité des aliments étudiés. Des analyses sur les cendres et les sucres ont également été réalisées en fonction des catégories de produits, ces paramètres étant plus ou moins importants d'un aliment à l'autre.

Un premier set de 54 échantillons a été avec une distance focale fixée à 45 cm. Une première courbe de calibration a été créée pour les paramètres suivants : humidité, protéines et matières grasses. 93 nouveaux échantillons d'aliments extrudés ont été scannés sur le XDS du CRA-W et sur la nouvelle tête de laboratoire Polytec de *Dumoulin*. Le deuxième set d'échantillons a été ajouté aux courbes préexistantes. Le passage de nouveaux aliments extrudés dans le futur afin de valider la robustesse de l'équation permettra de remplacer progressivement les échantillons du premier set.

Au total, les équations de calibration ont été réalisées pour les aliments extrudés à partir de 262 échantillons pour la protéine et l'humidité, 264 échantillons pour l'amidon et la cellulose et 168 échantillons pour la matière grasse (en comptant les répétitions). En fonction des critères d'analyse, toutes les prédictions n'ont pas nécessairement été retenues, certaines étant trop différentes du reste de la population.

Lors de l'analyse du premier set d'échantillons, il a été démontré que le XDS pouvait servir de référence pour l'humidité et les teneurs en protéines et matières grasses, les valeurs qu'il prédisait étant très proches des valeurs chimiques. C'est pourquoi peu d'analyses chimiques ont été réalisées sur le deuxième set d'échantillons. Au total, 22 échantillons ont été envoyés en analyse chimique pour connaître la teneur en protéine, 26 pour la matière grasse (FAH), 9 pour la cellulose brute et l'amidon et 18 pour la teneur en humidité. Les échantillons envoyés à l'analyse chimique ont été sélectionnés en fonction des teneurs théoriques pour chaque constituant. L'objectif était de disposer de valeurs chimiques d'échantillons extrêmes (valeurs attendues minimales et maximales) et d'échantillons moyens afin d'avoir une variabilité et une représentativité des échantillons maximales.

Concernant les prétraitements, les paramètres choisis afin d'obtenir les meilleurs résultats sont un traitement SNV et une dérivée première avec un gap size = 3 et un segment size = 3. Ces paramètres seront détaillés en *1.3. Méthodologie de la construction des équations de calibration*. Ils ont été appliqués pour l'ensemble des critères d'analyse de tous les aliments étudiés.

1.2.2. Les flocons

Lors de la production d'aliments floconnés, les grains passent dans une mélangeuse où est ajoutée l'eau, afin d'obtenir un produit à 21% d'humidité. Ils sont transférés dans un conditionneur où ils sont cuits sous pression de vapeur d'eau. Les grains transitent ensuite à travers la floconneuse, leur donnant

leur forme spécifique (Cedus, n.d.). Les flocons sont enfin refroidis et séchés, avant d'être conditionnés en sacs ou en vrac, ou envoyés à l'usine des mix (**Annexe 4**).

4 aliments floconnés différents ont été retenus : orge floconnée, maïs floconné, froment floconné et Golden Soy (graines de soja). Le frein majeur de ces aliments est leur faible volume de production. Une première courbe regroupant l'orge, le maïs et le froment doit être établie à partir de céréales classiques. Il a été décidé de créer une seconde courbe spécifique au Golden Soy, les flocons de graines de soja étant trop différents des autres aliments. Cette courbe doit être établie à partir du Full Fat Soya.

Un total de 28 Golden Soy, 60 maïs floconnés, 60 orges floconnées et 8 froments floconnés ont été scannés sur le XDS et Polytec de *Dumoulin*. Les Golden Soy étant envoyés en analyse chimique chaque semaine dans le cadre du plan de contrôle, l'obtention de valeurs chimiques est aisée. La méthode de sélection des échantillons à envoyer en analyse chimique est identique à celle appliquée aux aliments extrudés. Six critères d'analyses ont été retenus : humidité, protéines, matières grasses, cellulose brute, cendres et amidon.

1.2.2. Autres aliments et matières premières

Dumoulin a également choisi d'analyser des aliments « traditionnels », les granulés et les farines représentant des volumes importants de production.

Ainsi, le projet s'est également axé sur l'analyse d'aliments pour porcs. Des équations de calibration ont été réalisées pour les aliments pour porcs à partir de 106 échantillons pour la protéine brute, la matière grasse (méthode A sans hydrolyse), l'humidité, la cellulose brute, l'amidon et les cendres. Selon les critères d'analyse, toutes les prédictions n'ont pas été systématiquement retenues. Parmi ces échantillons, 15 ont été envoyés en analyse chimique pour connaître les teneurs en protéines, matières grasses, humidité, cendres et cellulose brute, et 9 pour connaître la teneur en amidon.

Les aliments pour bovins ont également été analysés. Une courbe regroupe les correcteurs protéiques avec et sans urée (35-44% de protéine), les correcteurs farineux avec urée, les correcteurs à haute teneur en protéines (>50%) et des mélanges contenant entre 44 et 52% de protéines afin de combler l'écart entre les correcteurs avec urée et les correcteurs à plus haute teneur en protéines. L'ajout d'urée permet notamment d'augmenter la teneur initiale en protéines. Des équations de calibration ont été réalisées pour les correcteurs protéiques à partir de 216 échantillons de correcteurs protéiques avec urée, 65 échantillons de correcteurs protéiques sans urée, 88 échantillons de correcteurs protéiques à forte teneur en urée, 14 échantillons de correcteurs protéiques farineux avec urée et 32 échantillons de correcteurs protéiques mélangés (mix de correcteurs protéiques avec urée et de correcteurs protéiques à forte teneur en urée). Au total, 415 échantillons ont été scannés, même si la totalité des prédictions n'ont pas été retenues selon les critères d'analyse, certaines valeurs étant trop différentes de la moyenne de la population. Parmi ces échantillons, 196 ont été envoyés en analyse chimique pour

connaître la teneur en protéines, 15 pour la teneur en matières grasses (FAT ou méthode A sans hydrolyse) et 1 pour la teneur en cellulose.

Dumoulin s'est également intéressé à l'analyse infrarouge de ses matières premières. Trois matières premières ont pour l'instant été analysées.

Des équations de calibration ont été réalisées pour les DDGS de maïs à partir de 77 échantillons pour la protéine brute, la matière grasse (méthode B avec hydrolyse), l'humidité, la cellulose brute, les cendres et l'amidon. Selon les critères d'analyse, toutes les prédictions n'ont pas été retenues, certaines valeurs étant trop différentes du reste de la population. Parmi ces échantillons, 48 ont été envoyés en analyse chimique pour connaître la teneur en protéine, en cellulose brute, en cendres et en humidité, 44 pour la teneur en matière grasse, 36 pour la teneur en amidon et 39 pour la teneur en sucres.

Des équations de calibration ont également été réalisées pour les tourteaux de colza à partir de 98 échantillons pour la protéine brute, la matière grasse (méthode A sans hydrolyse), l'humidité, la cellulose brute, les cendres et l'amidon. Selon les critères d'analyse, toutes les prédictions n'ont pas été retenues. Parmi ces échantillons, 21 ont été envoyés en analyse chimique pour connaître les teneurs en protéines et matières grasses, 19 pour connaître les teneurs en humidité et en cendres, 9 pour connaître la teneur en cellulose brute, 1 pour connaître la teneur en amidon et 21 pour connaître la teneur en matières grasses.

Enfin, 86 échantillons de tourteaux de lin ont été utilisés pour créer des équations de calibration pour la protéine brute, la matière grasse (méthode A sans hydrolyse), l'humidité, la cellulose brute et les cendres. Parmi ces échantillons, 12 ont été envoyés en analyse chimique pour connaître les teneurs en protéines et matières grasses et humidité.

D'autres aliments ont été analysés mais ne seront pas présentés dans ce mémoire, les modèles étant encore en cours de réalisation. Cela concerne les autres aliments bovins contenant moins de 35% de protéine, ce qui permettra d'exercer un contrôle sur l'entièreté de la gamme bovin selon la protéine, les aliments pour volailles ainsi que d'autres matières premières : céréales (séparation du maïs grain et de l'avoine-froment-triticales-orge) et autres tourteaux principalement. Enfin, des échantillons on-line à l'état de farine après mélange de chaque catégorie d'aliment ont été scannés et classés en cinq catégories : bovins classiques, correcteurs protéiques, porcs, volailles et extrudés.

I.3. Méthodologie de la construction des équations de calibration

La méthode appliquée afin de créer les modèles a été identique à tous les aliments afin de ne pas multiplier les protocoles de calibration et d'obtenir des résultats cohérents pour l'ensemble des produits.

En amont de la calibration, des scans de nombreux échantillons ont été effectués sur les appareils Polytec et XDS, ce dernier étant suffisamment précis pour prédire les valeurs des échantillons. En

parallèle, une partie de ces échantillons, en fonction de leur représentativité dans le set (teneurs minimales, maximales et moyennes), a été envoyée en analyse chimique afin d'obtenir des valeurs de référence afin de prédire les valeurs des autres échantillons à l'aide du XDS. La précision de la prédiction par la XDS peut à ce moment être vérifiée en consultant la valeur du GH de chaque échantillon. Le GH, ou distance de Mahalanobis, correspond à l'exactitude de la prédiction. Chaque spectre est converti en un point et est comparé à la moyenne de tous les spectres. La prédiction est considérée comme correcte si le GH est inférieur à 3. Cependant, en cas de forte variabilité, le GH peut être plus élevé et la prédiction correcte. A l'inverse, un manque de variabilité inter-échantillons donnera un GH faible mais une prédiction peu intéressante.

Le choix des méthodes de prétraitements à appliquer aux spectres bruts et de régression a été émis par les professionnels de l'analyse infrarouge et de la chimiométrie du CRA-W, en cohérence avec les informations trouvées dans la bibliographie. Ainsi, deux prétraitements ont été systématiquement appliqués avant de commencer les calibrations : la transformation des spectres bruts par SNV puis par dérivée première à l'aide de la méthode Gap Segment afin d'étudier la pente d'une courbe et de révéler davantage d'information. Cette méthode de dérivation consiste à choisir un ensemble de fenêtres (ou segment) dans lesquelles sont présents un certain nombre de longueur d'onde. Une fenêtre peut être assimilée à une étude de pente. Ces fenêtres sont elles mêmes séparées par plusieurs longueurs d'onde (ou gap) et la dérivée de chaque segment est extrapolée avec les deux gaps situés aux longueurs d'onde suivante et précédente. Les paramétrages retenus sont un gap size = 3 et un segment size = 3.

La calibration peut ensuite être réalisée à l'aide de nombreuses méthodes de régression. La méthode retenue est la régression PLS, qui est la plus adaptée pour la prédiction de plusieurs critères d'analyse pour une large gamme spectrale. Lors de la calibration PLS, une première validation a été effectuée en parallèle par la technique de cross-validation et d'obtenir les premiers indicateurs de la précision de la calibration et de la cross-validation (SEC, R^2 et SECV et RPD) à l'aide de la projection des échantillons sur une droite de régression, les valeurs de référence en abscisse et les valeurs prédites en ordonnée. La précision du modèle a pu ensuite être améliorée en sélectionnant des outliers de calibration, plus intéressante que la sélection des outliers avant les prétraitements ou après ACP, car cette sélection vise uniquement les valeurs éloignées de la moyenne de la calibration et non vis-à-vis

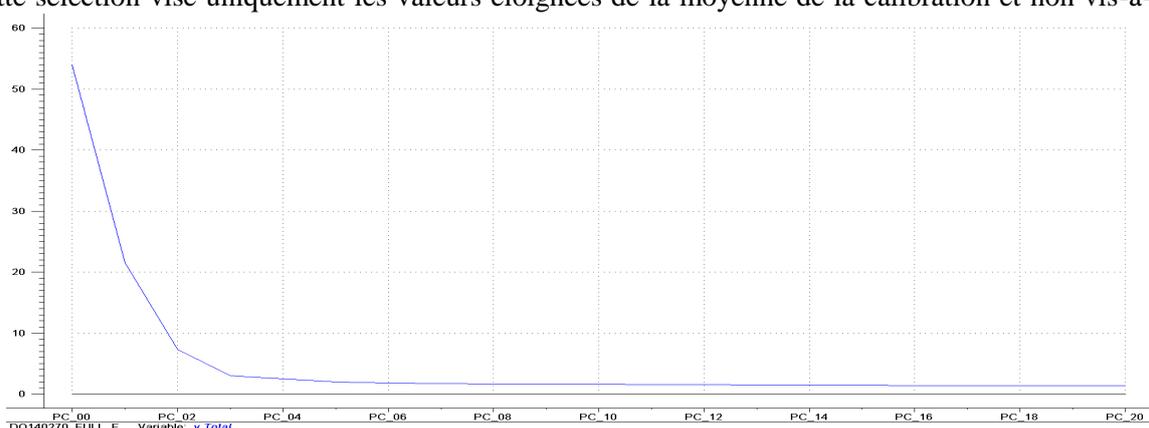


Figure 25 : Aperçu de la variation expliquée permettant de choisir le nombre de termes du modèle (ici 9).

de l'allure spectrale. Cette méthode de sélection, basée en grande partie sur la subjectivité de l'opérateur et de son expérience, ne doit pas pénaliser le modèle. Ainsi, il a été décidé de ne pas supprimer plus de 10 à 15% des échantillons ayant servi à la calibration. Enfin, un dernier moyen d'améliorer le modèle est modifier le nombre de termes (ou composantes principales) de l'équation. On considère généralement que le nombre de termes idéal correspond au moment où la variation expliquée ne varie plus, avant un pic, ou lorsque la variation du SEP d'un terme à l'autre est minimale (**Figure 25**).

Une fois le modèle recalculé sans les outliers de calibration et validé, la robustesse du modèle peut être testée en passant sur le Polytec des échantillons ayant servi à la calibration et des échantillons inconnus. Pour ces derniers, il est possible d'analyser l'exactitude de la prédiction soit en les envoyant en analyse chimique, soit en comparant les prédictions avec les valeurs théoriques de formulation. A nouveau, il est possible d'utiliser le GH pour vérifier l'exactitude de la prédiction. La démarche de la construction d'une équation de calibration par régression PLS est résumée en **Figure 26**.

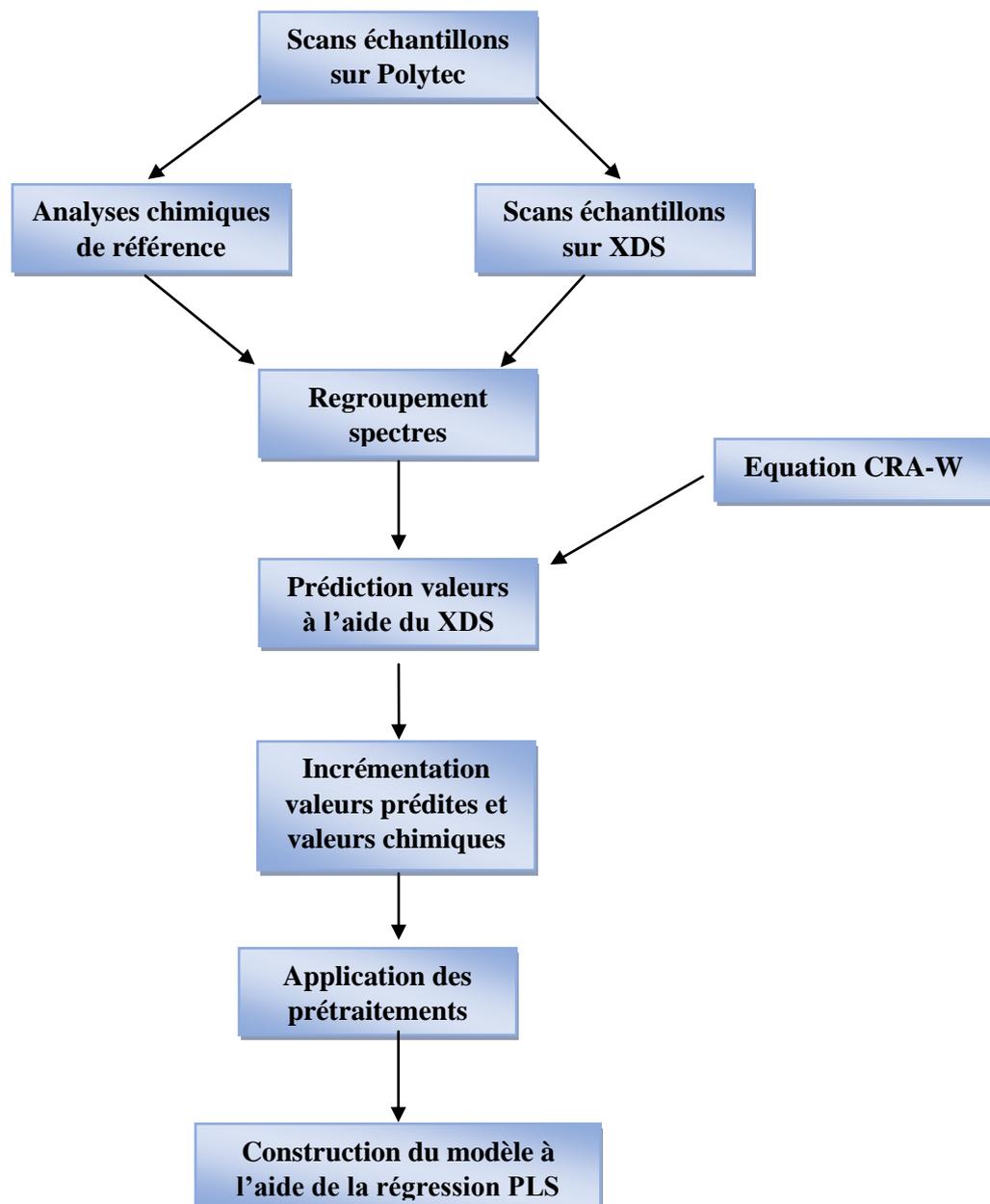


Figure 26 : Démarche de la mise en place d'une équation de calibration

II. Présentation et interprétation des résultats des calibrations

Cette partie est consacrée à la présentation des résultats obtenus lors de la création des équations de calibration pour les produits suivants : aliments extrudés, correcteurs protéiques, aliments pour porcs, matières premières et aliments floconnés. Elle sera suivie de la phase de vérification de la précision des modèles de calibration par le passage sur le Polytec de plusieurs échantillons ayant servi à la calibration pour comparer les valeurs prédites aux valeurs chimiques et prédictions XDS. Afin de faciliter la lecture du document, les résultats obtenus seront directement interprétés et commentés. L'ensemble des valeurs chimiques et des valeurs prédites sont exprimées en %.

II.1. Les aliments extrudés

La liste de correspondance présentant les échantillons sélectionnés pour la calibration ainsi que les valeurs chimiques est présentée en **Annexe 5**. La répétition des mesures n'apparaît pas dans cette liste.

Les résultats des calibrations pour les 5 critères d'analyse retenus pour les aliments extrudés sont présentés en **Tableau 3**. Ils présentent le détail des résultats des valeurs chimiques de références obtenues et des phases de calibration et de cross-validation.

Critères d'analyse	Valeurs chimiques de référence					Résultats des calibrations et cross-validations									
	N	Moy	SD	Min	Max	N	Prétraitements	Outliers	R ² c	SEC	R ² cv	SECV	RPD	Nb termes	
Humidité	18	7,28	1,06	5,60	9,50	262	SNV-1-3-3	4	0,92	0,39	0,90	0,43	3,36	8	
Protéine brute	22	20,50	2,94	11,60	24,90	262	SNV-1-3-3	7	0,92	0,72	0,89	0,85	3,23	11	
Matière grasse (méth B)	25	20,46	6,86	6,80	33,30	168	SNV-1-3-3	10	0,98	0,98	0,97	1,06	6,97	7	
Cellulose brute	9	9,07	2,76	6,10	13,40	264	SNV-1-3-3	6	0,89	0,91	0,88	0,95	3,05	6	
Amidon	9	15,67	4,33	9,50	22,80	264	SNV-1-3-3	17	0,90	1,13	0,89	1,20	3,61	6	

Tableau 3 : Résultats des analyses chimiques, calibrations et cross-validations sur les aliments extrudés.

N : nombre d'échantillons ; Moy : moyenne des valeurs chimiques ; SD : écart-type ; Min : valeur minimum ; Max : valeur maximum
R²c : coefficient de corrélation de calibration ; SEC : erreur standard de calibration ; R²cv : coefficient de corrélation de cross-validation ;
SECV : erreur standard de cross-validation ; RPD : Ratio Performance to Deviation (SD/SECV)

Dans le cas des extrudés, *Dumoulin* a placé sa priorité la matière grasse. Viennent ensuite l'humidité et la protéine, puis la cellulose et les cendres.

Les graphiques présentant les droites de régression pour la calibration et la cross-validation pour chaque paramètre sont disponibles en **Annexe 6**. Ils permettent de comparer les valeurs chimiques de référence/valeurs prédites par le XDS (en abscisse) aux valeurs NIR (en ordonnée).

- Matière grasse

Les résultats de la calibration de la matière grasse montrent une corrélation de la régression excellente (R²c = 0,98 et R²cv = 0,97). Cependant, elle s'accompagne d'un SECV très élevé (SECV = 1,06), ce qui indique une perte au niveau de la précision de la prédiction. Toutefois, le SECV est très proche du

SEC (SEC = 0,98), ce qui montre que la calibration est de qualité. 95% des échantillons sont prédits avec une marge d'erreur égale à 2 fois le SECV, soit 2,12.

Plusieurs facteurs peuvent expliquer les valeurs élevées du SEC et du SECV. Tout d'abord, dans le cas des aliments extrudés, une partie de l'information spectrale n'a pas pu être utilisée pour la calibration de la matière grasse. En effet, celle-ci se trouve dans une région de l'infrarouge qui n'est pas prise en compte par la SPIR. La perte de l'information influe ainsi directement sur la qualité de la prédiction, et ce même si de nombreuses analyses chimiques ont été réalisées. Enfin, les valeurs de ces indicateurs sont élevées en raison de la dispersion des échantillons de la population pour ce paramètre. Les aliments extrudés analysés possèdent une teneur en matière grasse comprise entre 5% et 35%. L'écart-type de la population est par conséquent très élevé et une grande variabilité au niveau des échantillons rend certes le modèle plus robuste, mais augmente dans le même temps le SEC et le SECV. Un SECV élevé ne posera pas de problème pour des teneurs en matière grasse élevée, en revanche, une erreur de prédiction élevée pour de faibles teneurs peut être dérangeante.

C'est pourquoi d'autres calibrations ont été réalisées afin d'améliorer le SEC et le SECV de ce paramètre. Tout d'abord, en regardant le graphique de la régression, on observe un creux entre 12% et 15% de matière grasse. Ensuite, pour les aliments extrudés, la différence de hauteur de tête entre le premier set d'échantillons et le second peut avoir un impact sur la précision de la calibration, deux mêmes aliments possédant deux spectres différents à deux plages d'absorbances différentes. Les effets de la hauteur de tête seront discutés *en IV. Discussion sur la méthodologie – L'acquisition spectrale*.

C'est pourquoi plusieurs calibrations ont été réalisées. En ne prenant que les nouveaux échantillons et en supprimant les échantillons avec une teneur inférieure à 15%, on obtient un SECV de 0,85. En supprimant les échantillons avec une teneur supérieure à 15%, le SECV est de 0,90. En prenant en compte l'intégralité des extrudés et en enlevant les teneurs les plus basses, le SECV est de 1,15, contre 0,91 si l'on enlève les teneurs supérieurs à 15%. Ainsi, il a été retenu de ne conserver que les nouveaux extrudés et l'intégralité de la plage de valeurs. Les nouveaux extrudés ont été prédits en tenant compte des anciens extrudés, des analyses chimiques ayant été réalisées sur ces échantillons. Les anciens extrudés ont ensuite été supprimés de la calibration.

Il est également possible de regarder le RPD. Ce dernier est égal à 6,97, ce qui indique que le modèle est excellent. Cependant, son interprétation est sujette à caution et l'intérêt de cet indicateur sera discutée plus tard (*IV. Discussion sur la méthodologie – Discussion sur les indicateurs de performance des calibrations et des cross-validations*).

- Humidité

La corrélation de la régression pour l'humidité est également excellente puisqu'elle est supérieure ou égale 0,90. Les valeurs du SEC et SECV sont relativement faibles, ce qui peut s'expliquer par la

précision du modèle et la faible variabilité entre les échantillons. L'humidité est un paramètre important car elle constitue un indicateur de la conservation du produit. Cependant, la faible variabilité du set d'échantillons et l'étude du RPD qui montre que le modèle est intéressant ($RPD = 3,36$), ne permettra pas d'améliorer significativement le modèle. 95% des échantillons sont prédits avec une marge d'erreur égale à 2 fois le SECV, soit 0,86, ce qui est acceptable compte tenu de la variabilité et de la précision relative d'une prédiction réalisée sur un tel paramètre

Un des intérêts d'être précis au niveau de l'humidité pour *Dumoulin* est d'être inférieur à 14% afin de respecter le cadre législatif. Ainsi, le premier objectif est surtout de ne pas dépasser ce seuil. Cependant, la mesure de l'humidité sur un aliment est difficile puisqu'elle peut évoluer au cours du temps. Le deuxième intérêt du contrôle de l'humidité est de voir quelle quantité de liquides peut être ajoutée au maximum. Ces points seront davantage détaillés en *III. Discussion sur la réglementation européenne relative à l'étiquetage des matières premières et des aliments composés pour animaux – Le cas particulier de l'humidité.*

- Protéine

Les résultats de la calibration de la protéine montrent une corrélation de la régression excellente ($R^2c = 0,92$ et $R^2cv = 0,89$). Toutefois, tout comme la matière grasse, elle s'accompagne d'un SECV élevé ($SECV = 0,85$), en partie expliqué par la forte variabilité observée entre les échantillons. Cependant, en parallèle, le nombre de termes de l'équation est très élevé (11 termes) ce qui peut engendrer un risque de surparamétrage et avoir une incidence sur la précision du modèle. 95% des échantillons sont prédits avec une marge d'erreur égale à 2 fois le SECV, soit 1,70. Cela est acceptable compte tenu de la plage de valeurs pour la protéine des aliments extrudés.

De plus, l'ensemble des aliments extrudés a été conservé pour la calibration de la protéine, avec les problèmes au niveau de l'absorbance que cela implique. Il est ainsi possible de conclure à ce stade que le modèle pourrait être plus performant, par exemple, en écartant comme dans le cas de matière grasse les anciens extrudés afin de supprimer l'éventuel biais qu'a pu engendrer la modification de la distance focale.

- Cellulose brute

La corrélation de la régression est à nouveau excellente ($R^2c = 0,89$ et $R^2cv = 0,88$), cependant le SECV est très élevé ($SECV = 0,95$) ce qui indique une faible précision du modèle, la variabilité inter-échantillons n'étant pas très marquée dans le cas de la cellulose brute. 95% des échantillons sont prédits avec une marge d'erreur égale à 2 fois le SECV, soit 1,90. La marge d'erreur est très élevée puisqu'il n'existe pas une grande variabilité entre les échantillons et les teneurs en cellulose brute sont relativement faibles. Cependant, la calibration de la cellulose brute dans le cas d'aliments extrudés est très secondaire et le modèle a dû être créé à partir d'une équation différente des autres critères physicochimiques, le CRA-W ne possédant pas de bases de données pour ce paramètre. En l'état,

améliorer le modèle n'est pas primordial dans la mesure où la teneur en cellulose brute ne varie que peu d'un aliment extrudé à l'autre et qu'il ne s'agit pas d'un critère principal pour *Dumoulin*.

- Amidon

A nouveau, la corrélation de la régression est très bonne ($R^2c = 0,90$ et $R^2cv = 0,89$), mais le SEC et le SECV sont très élevés ($SEC = 1,13$ et $SECV = 1,20$). Comme pour tous les paramètres, la calibration est de qualité puisque ces deux indicateurs sont proches l'un de l'autre. En revanche, la précision du modèle est sujette à discussion. 95% des échantillons sont prédits avec une marge d'erreur égale à 2 fois le SECV, soit 2,40. Cela ne pose pas de problème dans le cas d'aliment à haute teneur en amidon, mais le modèle doit cependant être plus précis pour les faibles teneurs. Les valeurs élevées de ces indicateurs peut s'expliquer tout d'abord par la dispersion de la population pour ce paramètre, l'écart-type des valeurs chimiques de référence étant de 4,33, elle peut ensuite s'expliquer par la nature de ces aliments.

Il a été vu dans la bibliographie que le processus d'extrusion provoquait la gélatinisation de l'amidon. L'amidon étant modifié au cours du processus, et le CRA-W ne disposant pas de données pour ce paramètre, prédire avec précision ce paramètre s'avère très difficile, et il semble peu probable que le modèle puisse être amélioré. L'amidon constituant pour *Dumoulin* un paramètre moins important que la matière grasse, la protéine et l'humidité, passer davantage de temps sur ce paramètre n'est pas pertinent.

Plusieurs échantillons dont les valeurs chimiques de référence étaient connues ont ensuite été repassés sur le Polytec une fois l'équation installée afin d'observer la précision du modèle (**Tableau 4**). Cela permet de vérifier que le Polytec sort des prédictions proches des valeurs chimiques de référence et des prédictions XDS. Cela ne remplace pas la prédiction d'échantillons indépendants, mais il s'agit d'une première étape dans la vérification du modèle.

N° échantillon	Nom	Humidité		Protéine			Matière grasse (mét b)		
		Prédiction XDS	Prédiction Polytec	Valeur chimique	Prédiction XDS	Prédiction Polytec	Valeur chimique	Prédiction XDS	Prédiction Polytec
DQ150121-025	Nutex 85	7,14	6,80		19,21	19,40	33,30		33,20
DQ150121-059	Nutex 68	7,74	7,20		18,37	18,50		27,99	28,40
DQ150121-060	Megalín	8,11	7,60		22,56	23,60		21,13	21,50
DQ150121-068	Linafiber Moulu	7,54	6,90		15,12	19,30		16,13	16,10
DQ150121-071	Nutex Elit	8,92	8,50		22,72	23,40		21,74	24,00
DQ150121-076	Nutex Excell	8,10	6,30		17,53	19,50		24,05	25,60
DQ150121-079	Nutex DLR	5,86	6,20		18,47	17,80		16,03	16,90
DQ150121-083	Nutex Sweet	7,18	7,00		19,24	20,20		11,49	11,00
DQ150121-086	Gr lin extr 70%	8,80	7,70		18,89	19,40		26,48	25,80
DQ150121-087	Nutex Excell	8,68	7,44	19,10		19,56	23,40		23,80
DQ150121-088	Nutex Excell	8,18	7,58				23,40		25,03
DQ150121-089	Nutex Excell	8,25	7,64				22,60		22,63
DQ150121-090	Nutex Excell	7,78	7,16				23,60		24,23
DQ150121-091	Nutex Elit	7,40	6,86	22,70		22,80		20,68	21,09
DQ150121-092	Nutex Elit	8,01	7,92	22,50		22,49	20,20		21,05
DQ150121-093	Austrolein								
DQ150121-031	Nutex Sweet								
DQ150121-033	Nutex Elit								
DQ150121-034	Excell								
DQ150121-035	Nutex 68								
DQ150121-004	Gr lin extr 60%								

N° échantillon	Nom	Cellulose			Amidon	
		Valeur chimique	Prédiction XDS	Prédiction Polytec	Prédiction XDS	Prédiction Polytec
DQ150121-025	Nutex 85		5,29	5,70	15,10	14,50
DQ150121-059	Nutex 68		6,21	5,80	15,50	17,00
DQ150121-060	Megalín	13,40		12,60	9,50	9,20
DQ150121-068	Linafiber Moulu	6,70		8,20	18,60	19,80
DQ150121-071	Nutex Elit		11,05	11,20	10,30	10,30
DQ150121-076	Nutex Excell		7,52	8,50	14,70	11,80
DQ150121-079	Nutex DLR		5,76	5,90	20,20	19,30
DQ150121-083	Nutex Sweet		11,24	11,10	14,30	13,40
DQ150121-086	Gr lin extr 70%	6,20		6,50		
DQ150121-087	Nutex Excell					
DQ150121-088	Nutex Excell					
DQ150121-089	Nutex Excell					
DQ150121-090	Nutex Excell					
DQ150121-091	Nutex Elit					
DQ150121-092	Nutex Elit					
DQ150121-093	Austrolein	10,40		6,80		
DQ150121-031	Nutex Sweet	10,00		11,20		
DQ150121-033	Nutex Elit	12,70		12,90		
DQ150121-034	Nutex Excell	9,00		9,30		
DQ150121-035	Nutex 68	7,10		6,20		
DQ150121-004	Gr lin extr 60%	6,10		6,30		

Tableau 4 : Résultats des validations sur le Polytec par comparaison des valeurs chimiques et des prédictions XDS et Polytec pour les aliments extrudés

Le contrôle de l'humidité a un premier but : vérifier que les valeurs restent en-dessous du seuil de 14%. En effet, il est difficile de comparer la précision de la prédiction du Polytec avec celle du XDS et les valeurs chimiques de référence, dans la mesure où l'humidité de l'échantillon a pu évoluer entre les différentes mesures. Ainsi, une différence importante ne sera pas significative d'une mauvaise

prédiction, mais plutôt de l'évolution du produit dans le temps. Cependant, ce seuil de 14% réglementaire n'est pas immuable, et selon les aliments extrudés, la teneur en humidité attendue est comprise entre 9% et 10.5%. Les prédictions XDS et Polytec sous-estiment ainsi largement l'humidité réelle du produit. Ainsi, la difficulté d'obtenir un modèle précis par l'infrarouge va nécessiter de continuer les mesures de l'humidité à l'aide d'un dessiccateur.

La comparaison des valeurs chimiques et des prédictions Polytec pour la protéine permet de constater que la prédiction est bonne, l'écart le plus important relevé étant de 0,46, ce qui est inférieur à la marge d'erreur contenant 95% des échantillons et égale à $2 \cdot \text{SECV}$ ($0,85 \cdot 2 = 1,70$). De même, les écarts observés entre les prédictions Polytec et les prédictions des spectres XDS sont très faibles, à l'exception de l'échantillon de Linafiber Moulu. Cet écart pose alors la question de la présentation de l'échantillon, les aliments extrudés étant à la fois sous forme de farines ou d'agglomérats. Cette différence se retrouve au niveau des spectres, les absorbances des farines étant généralement plus faibles. Elle se retrouve également sur le graphique de distribution des échantillons de la régression PLS, les extrudés sous forme de farines formant un groupe à part.

De même que la protéine, les prédictions des extrudés pour la matière grasse sont très proches des valeurs chimiques de référence, l'écart le plus important étant de 1,63, tout comme des prédictions des spectres XDS, l'écart le plus important étant de 2,26 et constituant une exception. Ainsi, en dépit d'une perte de l'information à l'acquisition spectrale et une forte variabilité entre les échantillons, les prédictions sont très précises et le modèle ne nécessite pas d'être amélioré pour l'instant. Son optimisation s'effectuera d'année en année lors de la vérification de la robustesse du modèle par l'apport de nouveaux échantillons dont les teneurs chimiques seront connues.

Les prédictions Polytec de la cellulose sont en revanche moins bonnes. Les écarts observés entre les valeurs chimiques de référence et les prédictions Polytec sont plus importants pour davantage d'échantillons. Les écarts les plus importants sont de 1,5 pour le Linafiber Moulu et 3,6 pour l'Austrolein. En revanche, les prédictions Polytec sont plus proches de celles du XDS. Ceci s'explique par le fait que des prédictions sur des aliments extrudés n'ont jamais été réalisées, et d'autres équations d'aliments différents tout en étant suffisamment proches des extrudés, ont été utilisées afin d'établir les premières prédictions XDS.

Enfin, pour l'amidon, malgré une marge d'erreur au niveau de la prédiction très importante, les prédictions Polytec sont relativement proches des prédictions XDS, à l'exception d'un Nutex Excell où l'écart entre les prédictions est de 2,90. Toutefois, comme expliqué précédemment, chercher à obtenir une grande précision au niveau de la teneur en amidon n'est pas pertinent puisque l'amidon subit de grandes modifications au cours du processus et il est possible que la prédiction ne prenne pas en compte la totalité de l'amidon présent dans l'aliment. Il s'agira dans ce cas plutôt d'un indicateur afin de vérifier si l'aliment contient une teneur en amidon proche de la formulation.

II.2. Les correcteurs protéiques

La liste de correspondance présentant les échantillons sélectionnés pour la calibration ainsi que les valeurs chimiques est présentée en **Annexe 7**.

Les résultats pour les 5 critères d'analyse retenus pour les correcteurs protéiques sont présentés en **Tableau 5**. Ils présentent le détail des résultats des valeurs chimiques de références obtenues et des phases de calibration et de cross-validation.

Critères d'analyse	Valeurs chimiques de référence					Résultats des calibrations et cross-validations								
	N	Moy	SD	Min	Max	N	Prétraitements	Outliers	R ² c	SEC	R ² cv	SECV	RPD	Nb termes
Humidité	/	/	/	/	/	307	SNV-1-3-3	0	0,91	0,51	0,89	0,57	3,12	9
Protéine brute	196	44,43	5,19	33,80	63,10	407	SNV-1-3-3	0	0,98	0,66	0,97	0,79	6,81	10
Matière grasse (méth A)	14	3,65	1,23	2,70	6,60	251	SNV-1-3-3	0	0,92	0,14	0,88	0,17	2,96	10
Cellulose brute	1	10,60	0,00	10,60	10,60	357	SNV-1-3-3	0	0,87	0,58	0,84	0,66	2,53	10
Cendres	/	/	/	/	/	357	SNV-1-3-3	12	0,76	0,42	0,70	0,48	1,95	10

Tableau 5 : Résultats des analyses chimiques, calibrations et cross-validations sur les correcteurs protéiques

Dans le cas des correcteurs protéiques, la priorité a été placée sur la protéine, car il s'agit de la particularité de ces aliments pour bovins. La haute teneur en protéine de ces aliments constitue un critère d'achat important et une des avantages nutritionnels apportés par les aliments *Dumoulin*. Viennent ensuite les teneurs en humidité et en matière grasse. Enfin, les teneurs en cellulose et en cendres sont des critères très secondaires pour *Dumoulin*.

Les graphiques présentant les droites de régression pour la calibration et la validation sont disponibles en **Annexe 8**. Ils permettent de comparer les valeurs chimiques de référence/valeurs prédites par le XDS(en abscisse) aux valeurs NIR (en ordonnée).

- Protéine

Les résultats de la calibration de la matière grasse montrent une corrélation de la régression excellente ($R^2c = 0,98$ et $R^2cv = 0,97$). Cependant, elle s'accompagne d'un SECV assez élevé ($SECV = 0,79$). 95% des échantillons sont prédits avec une marge d'erreur égale à 2 fois le SECV, soit 1,58, ce qui est tout à fait acceptable. Ceci s'explique par la très forte variabilité entre les échantillons pour la protéine, les valeurs chimiques allant de 33% à plus de 60% et l'écart-type étant de 5,19. Il sera néanmoins difficile d'améliorer le modèle, ce dernier étant par ailleurs particulièrement bon ($RPD = 6,81$).

Cependant, un SECV élevé peut poser problème dans la mesure où il existe plusieurs catégories de correcteurs en fonction de leur teneur en protéine, et une mauvaise prédiction peut faire passer un échantillon d'une catégorie à l'autre. Ainsi, par ordre croissant de teneur en protéine, on retrouve les correcteurs sans urée, les correcteurs avec urée et les correcteurs à haute teneur en protéine ayant la particularité d'être des aliments extrudés.

Au début du projet, il était convenu de créer deux courbes de correcteurs : une courbe regroupant les correcteurs sans urée, et une autre courbe regroupant les correcteurs avec urée et les correcteurs à haute teneur en protéine. Toutefois, un écart existait entre les correcteurs avec urée et les correcteurs extrudés (45%-50%), c'est pourquoi des mélanges entre ces deux catégories ont été effectués pour combler l'écart. En effet, les équations retenues par le CRA-W pour la prédiction de la protéine ne contiennent pas d'échantillons à de telles teneurs en protéine. Comblé l'écart était donc une solution pour disposer d'une plage de valeurs homogène et ainsi obtenir de bonnes prédictions.

Il a été en parallèle décidé de mettre ensemble tous les correcteurs, l'ajout d'urée dans le produit ne provoquant pas de changement au niveau spectral. Ainsi, une grande précision au niveau de la prédiction est indispensable afin qu'il n'y ait pas d'erreur sur la catégorie de correcteurs considérée. Toutefois, on peut supposer que les prédictions sont bonnes puisqu'aucun outlier n'a été détecté lors de la calibration.

En revanche, le nombre de termes de l'équation a été fixé à 10 afin d'améliorer le SEC et le SECV avec le risque d'un surparamétrage du modèle. Des pistes d'amélioration du modèle peuvent ainsi être envisagées sur ce point à l'avenir, lorsque de nouveaux échantillons seront ajoutés au modèle pour vérifier la robustesse de l'équation d'une année à l'autre.

- Humidité

Comme pour l'ensemble des aliments, la précision au niveau de la teneur en humidité n'est pas recherchée tant que cette dernière demeure inférieure à 14%. C'est pourquoi aucune analyse chimique n'a été réalisée et les spectres ont été entièrement prédits grâce aux équations du CRA-W qui dispose de vastes bases de données pour les aliments granulés. Cette stratégie a été payante puisque les résultats de la calibration de l'humidité montrent une corrélation de la régression excellente ($R^2c = 0,91$ et $R^2cv = 0,89$), ainsi que des valeurs de SEC et SECV moyennes et proches l'une de l'autre ($SEC = 0,51$ et $SECV = 0,57$), ce qui prouve que la calibration est correcte.

- Matière grasse

Les résultats de la calibration de la matière grasse montrent une corrélation de la régression excellente ($R^2c = 0,92$ et $R^2cv = 0,88$). Elle s'accompagne également d'un SEC et un SECV très faibles ($SEC = 0,14$ et $SECV = 0,17$). Cependant, l'écart-type des échantillons de référence est de 1,23 pour une plage de valeurs de 2.70 à 6.60. De plus, le RPD est de 2,96, ce qui signifie que le modèle n'est pas très performant et le nombre de termes de l'équation a été fixé à 10 afin de diminuer au maximum le SECV, même si ce modèle permet de prédire 95% des échantillons avec une marge d'erreur de 0,34. La comparaison ci-après des valeurs chimiques de références et des prédictions XDS aux prédictions Polytec permettra de dégager des pistes d'améliorations du modèle.

- Cellulose brute

Les résultats de la calibration de la cellulose montrent une corrélation de la régression très bonne ($R^2c = 0,87$ et $R^2cv = 0,84$). Cependant, le modèle n'est pas très performant car le SECV est très élevé (SECV = 0,60) alors que théoriquement, il n'y a que peu de variabilité entre les échantillons au niveau de la cellulose. De plus, le SECV a été diminué en augmentant significativement le nombre de termes de l'équation. Même si la cellulose est un paramètre secondaire dans le cas des correcteurs protéiques, il serait intéressant d'optimiser le modèle. La comparaison des valeurs chimiques de références et des prédictions XDS aux prédictions Polytec permettra de dégager plusieurs pistes d'améliorations.

- Cendres

Les résultats de la calibration des cendres montrent une corrélation de la régression moyenne ($R^2c=0,76$ et $R^2cv = 0,70$). Le modèle n'est ainsi que peu performant, d'autant plus que le SECV est très élevé (SECV = 0,48) alors qu'il n'existe quasiment aucune variabilité au niveau des cendres pour les correcteurs protéiques. De plus, le nombre de termes de l'équation est très élevé (10 termes), et un risque de surparamétrage a été pris afin de diminuer le SECV, et le modèle a été construit en supprimant un grand nombre d'échantillons (outliers de calibration), ce qui signifie qu'une grande partie des échantillons n'a pu être prédite précisément. Cependant, d'après l'expertise du CRA-W, l'analyse infrarouge des cendres sur des aliments destinés aux animaux est peu pertinente et il est très difficile d'obtenir un modèle de prédiction précis. La comparaison des valeurs chimiques et des prédictions XDS aux prédictions Polytec permettront de conclure sur la précision du modèle pour les cendres, à savoir si le modèle en l'état doit être amélioré ou non.

Plusieurs échantillons dont les valeurs chimiques de référence étaient connues ont ensuite été repassés sur le Polytec une fois l'équation installée afin d'observer la précision du modèle (**Tableau 6**). Cela permet de vérifier que le Polytec sort des prédictions proches des valeurs chimiques de référence et des prédictions XDS. Cela ne remplace pas la prédiction d'échantillons indépendants, mais il s'agit d'une première étape dans la vérification du modèle.

		Humidité		Protéine			Matière grasse (mét a)		
N° échantillon	Nom	Prédiction XDS	Prédiction Polytec	Valeur chimique	Prédiction XDS	Prédiction Polytec	Valeur chimique	Prédiction Polytec	
DQ150033-106	CorrectMaïs 3544	11,85	9,8	43,70		43,80	3,08	3,60	
DQ150033-107	Startomix	10,80	9,3	42,50		42,70	3,59	3,70	
DQ150033-108	Protamix	11,37	10,2	40,10		40,80	3,92	4,00	
DQ150033-109	Start Protéine	11,06	9,1	43,40		43,70	3,45	3,40	
DQ150033-170	Aminostart	10,77	11		39,30	39,00	2,90	3,30	
DQ150033-189	Hipro 58	5,77	5,8	55,30		54,00	6,60	6,30	
DQ150033-191	Regamine Farine		7,63	51,00		50,20		3,83	
DQ150033-192	Regamine Farine		6,82	47,60		47,90		4,47	
DQ150033-194	Hipro 58	7,81	6,04	52,50		52,75	6,19	6,28	
DQ150033-258	Hipro 58		6,9	45,50		48,50	6,70	6,40	
DQ150033-281	Correctmaïs 3544	10,49	9,5		45,36	45,30	2,70	3,30	
DQ150033-282	Startomix	10,81	10		42,92	43,10	2,90	3,50	
DQ150033-283	Protamix	10,15	9,6		40,90	41,20	3,80	4,00	
DQ150033-284	Start Protéine	9,78	8,5		44,24	43,80	3,00	3,40	
DQ150033-287	Start Protéine Farine	10,25	9,4		43,70	43,10	3,10	3,40	
DQ150033-289	Protamix Farine	10,60	10,4		40,70	40,90	3,00	3,30	
		Cellulose			Cendres				
N° échantillon	Nom	Valeur chimique	Prédiction XDS	Prédiction Polytec	Valeur chimique	Prédiction XDS	Prédiction Polytec		
DQ150033-106	CorrectMaïs 3544		15,25	14,90		8,58	8,10		
DQ150033-107	Startomix		14,60	14,40		8,92	7,90		
DQ150033-108	Protamix		12,22	11,80		8,39	8,10		
DQ150033-109	Start Protéine		13,79	13,50		8,76	7,80		
DQ150033-170	Aminostart		12,55	13,00		7,93	8,20		
DQ150033-189	Hipro 58	10,60		12,70	3,90		7,70		
DQ150033-191	Regamine Farine			13,85			7,88		
DQ150033-192	Regamine Farine			14,29			6,93		
DQ150033-194	Hipro 58		15,42	12,48		8,96	7,54		
DQ150033-258	Hipro 58	10,60		12,70	3,80		8,00		
DQ150033-281	Correctmaïs 3544		14,34	14,00		9,41	8,20		
DQ150033-282	Startomix		14,07	13,70		9,45	8,20		
DQ150033-283	Protamix		11,63	10,90		8,26	7,80		
DQ150033-284	Start Protéine		14,41	13,80		8,78	7,90		
DQ150033-287	Start Protéine Farine		12,78	12,10		6,08	7,80		
DQ150033-289	Protamix Farine		12,06	10,80		6,74	7,90		

Tableau 6 : Résultats des validations sur le Polytec par comparaison des valeurs chimiques et des prédictions XDS et Polytec pour les correcteurs protéiques

La comparaison de la teneur en humidité entre les prédictions XDS et Polytec est peu pertinente car elles n'ont pas été réalisées dans le même laps de temps. Ainsi, la teneur en humidité a diminué au cours du temps. C'est le constat qui ressort de cette comparaison. A quelques exceptions près, les prédictions Polytec sont systématiquement inférieures à celles du XDS, tout en étant suffisamment proches, et comprises dans la marge d'erreur propre à la prédiction infrarouge. Le modèle de prédiction de l'humidité est ainsi tout à fait correct et ne nécessite pour l'heure aucune amélioration conséquente.

Le modèle de prédiction de la protéine est très performant puisque les écarts observés entre les valeurs chimiques de référence et les prédictions Polytec sont très faibles. Seul un Hipro 58 possède un écart de 3 points. Cependant, l'analyse chimique sort l'aliment à 45.5 alors qu'il est censé être à 52% si l'on regarde les valeurs de formulation, ce qui peut suggérer qu'une erreur s'est produite lors de l'analyse chimique.

Au niveau de la matière grasse, on constate que les prédictions Polytec sont presque toujours supérieures aux valeurs chimiques, ce qui amène à penser que le Polytec tend à surestimer la valeur de la matière grasse. Le modèle ainsi créé est relativement performant, mais il peut être amélioré. Peu d'aliments à haute teneur en protéine tels que les Hipro 58 ont été utilisés pour la calibration de la matière première alors que ce sont les aliments possédant la plus haute teneur en matière grasse. Ceci peut entraîner un biais au niveau de la prédiction. Ainsi, en prédisant davantage d'Hipro 58 et en les intégrant à la calibration, le modèle pourrait être considérablement optimisé.

Le modèle de prédiction de la cellulose est en revanche mauvais. Sur la base de deux valeurs chimiques, les écarts observés entre ces valeurs et les prédictions sont supérieurs à 2 points. De même, en se référant aux valeurs théoriques de formulation, on s'aperçoit que les prédictions XDS et Polytec surestiment de plusieurs points les teneurs en cellulose des aliments. Le modèle doit ainsi être entièrement repris pour ce paramètre. Pour cela, de nouveaux échantillons doivent être envoyés en analyse chimique puis intégrés à la calibration afin de construire un nouveau modèle de régression.

Le constat est similaire pour les cendres. Le modèle surestime largement la teneur en cendres de l'Hipro 58 : le chimique donne une valeur inférieure à 4%, et la prédiction une valeur proche de 8%. En s'intéressant aux références de formulation, il ne s'agit ici pas d'une erreur au niveau de l'analyse chimique, l'Hipro 58 étant attendu à 3,5% de cendres. De nouveaux échantillons d'Hipro 58 doivent ainsi être envoyés en analyse chimique afin de corriger l'erreur de prédiction du modèle sur ces aliments particuliers.

II.3. Les aliments pour porcs

La liste de correspondance présentant les échantillons sélectionnés pour la calibration ainsi que les valeurs chimiques est présentée en **Annexe 9**.

Les résultats pour les 6 critères d'analyse retenus pour les aliments pour porcs sont présentés en **Tableau 7**. Ils présentent le détail des résultats des valeurs chimiques de référence obtenues et des phases de calibration et de cross-validation.

Critères d'analyse	Valeurs chimiques de référence					Résultats des calibrations et cross-validations									Nb termes
	N	Moy	SD	Min	Max	N	Prétraitements	Outliers	R ² c	SEC	R ² cv	SECV	RPD		
Humidité	15	11.78	1.03	10.10	13.13	106	SNV-1-3-3	0	0.90	0.23	0.87	0.26	2.86	5	
Protéine brute	15	16.22	0.94	13.95	17.59	106	SNV-1-3-3	9	0.94	0.24	0.90	0.32	8.35	11	
Matière grasse (méth A)	15	4.90	1.90	2.46	7.21	106	SNV-1-3-3	0	0.96	0.21	0.95	0.24	4.96	7	
Cellulose brute	15	4.25	1.53	2.50	7.63	106	SNV-1-3-3	0	0.95	0.36	0.94	0.41	4.11	4	
Cendres	15	5.11	0.73	4.24	6.40	106	SNV-1-3-3	2	0.94	0.33	0.91	0.41	3.38	7	
Amidon	9	40.68	2.86	35.71	44	94	SNV-1-3-3	9	0.97	0.47	0.94	0.74	9.96	11	

Tableau 7 : Résultats des analyses chimiques, calibrations et cross-validations sur les aliments pour porcs

Dans le cas des aliments pour porcs, la priorité a été fixée pour la protéine, la matière grasse, l'amidon et la cellulose. Les teneurs en humidité et en cendres sont des paramètres secondaires dans les aliments pour porcs.

Les graphiques présentant les droites de régression pour la calibration et la validation sont disponibles en **Annexe 10**. Ils permettent de comparer les valeurs chimiques de référence/valeurs prédites par le XDS (en abscisse) aux valeurs NIR (en ordonnée).

- Protéine

Les résultats de la calibration de la protéine montrent une corrélation de la régression excellente ($R^2c = 0,94$ et $R^2cv = 0,90$). Les valeurs du SEC et du SECV sont relativement faibles et proches l'une de l'autre ($SEC = 0,24$ et $SECV = 0,32$), ce qui suggère que la prédiction du modèle est bonne. 95% des échantillons sont prédits avec une marge d'erreur égale à 2 fois le SECV, soit 0,64, ce qui est acceptable compte tenu de la variabilité entre les échantillons. Cependant, le fait d'avoir retenu 11 termes tend à relativiser le SECV, puisqu'un risque de surparamétrage de l'équation a été pris afin d'avoir un SECV faible. Enfin, l'étude du RPD est intéressante car elle montre la variabilité qui existe dans les aliments pour porcs au niveau de la protéine sans que cela ne se remarque au niveau du SEC et du SECV ($RPD = 8,35$).

Si l'on se base sur le RPD, le modèle est excellent. Cependant, en analysant plus en profondeur les valeurs chimiques de référence et les prédictions XDS, on remarque que deux échantillons d'un même aliment sont très différents du reste de la population. C'est le cas des aliments pour truies en gestation. La teneur en protéine de cet aliment est fixée à moins de 10%, et le XDS prédit ces aliments 30% de protéine, alors que la teneur moyenne du reste de la population est comprise entre 15 et 19%. Le RPD

prend en compte l'intégralité des valeurs prédites et des valeurs chimiques puisqu'il correspond au ratio de l'écart type des valeurs prédites sur le SECV. Or le SECV ne prend en compte que les échantillons qui ont été inclus dans la calibration, et ces aliments atypiques ont été considérés comme des outliers et retirés de la calibration, ce qui explique un SEC et un SECV si faible car la variabilité entre les échantillons est faible.

En l'état, le modèle est fiable, cependant, il exclut une partie des aliments pour porcs. Il pourrait être optimisé de la même manière que les correcteurs protéiques, à savoir créer de nouvelles formules permettant de combler l'écart observé.

- Amidon

Les résultats de la calibration de l'amidon montrent une corrélation de la régression excellente ($R^2=0,97$ et $R^2_{cv} = 0,94$). En revanche, les valeurs du SEC et du SECV sont élevées et très éloignées (SEC = 0,47 et SECV = 0,74), ce qui peut indiquer une mauvaise calibration alors que la plage des valeurs est relativement restreinte. 95% des échantillons sont prédits avec une marge d'erreur égale à 2 fois le SECV, soit 1,48, ce qui est trop important du fait de la faible variabilité entre les échantillons.

On constate le même phénomène au niveau du RPD que pour la protéine. Il s'agit en effet des mêmes aliments pour truies en gestation possédant une teneur en amidon très différente du reste de la population (9% en moyenne contre 35 à 47% pour le reste de la population), ce qui explique que le RPD soit de 9,96, sans pour autant que le modèle soit excellent. Cette remarque sur le RPD sera expliquée plus longuement en *IV. Discussion sur la méthodologie – Discussion sur les indicateurs de performance de la calibration et de la cross-validation.*

De plus, le nombre de termes retenus pour ce modèle afin de diminuer le SECV est très élevé, et il existe un risque de surparamétrage. En l'état, le modèle mérite d'être amélioré par l'ajout de nouveaux échantillons et de valeurs chimiques.

- Matière grasse

Les résultats de la calibration de la matière grasse montrent une corrélation de la régression excellente ($R^2_c = 0,96$ et $R^2_{cv} = 0,95$). De plus, les valeurs du SEC et du SECV sont très proches l'une de l'autre et très faibles (SEC = 0,21 et SECV = 0,24) ce qui signifie que le modèle est juste même si la faible variabilité entre les échantillons explique que ces valeurs soient faibles. L'étude du RPD montre que le modèle est très bon (RPD =4,96) et permet de prédire les aliments pour porcs avec une marge d'erreur de 0,48.

- Cellulose brute

Les résultats de la calibration de la cellulose montre une corrélation de la régression excellente ($R^2c=0,95$ et $R^2cv = 0,94$). Le SECV est suffisamment faible ($SECV = 0,41$) compte tenu de la faible variabilité des échantillons et du faible nombre de termes qui composent l'équation. Ainsi, 95% des échantillons sont prédits avec une marge d'erreur de 0,82. De plus, l'étude du RPD permet de conclure sur la précision du modèle. Avec une valeur de 4,11, le modèle est très bon.

- Humidité

Les résultats de la calibration de l'humidité montrent une corrélation de la régression excellente ($R^2c=0,90$ et $R^2cv = 0,87$). La plage de teneurs en humidité est très restreinte, ce qui permet d'obtenir un SECV très faible ($SECV = 0,26$). 95% des échantillons sont prédits avec une marge d'erreur de 0,52, ce qui est acceptable compte tenu de la variabilité des échantillons.

- Cendres

Les résultats de la calibration des cendres montrent une corrélation de la régression excellente ($R^2c=0,94$ et $R^2cv = 0,91$). Cependant, le SECV est très élevé ($SECV = 0,41$) alors que 2% à peine séparent les extrema. Même si le RPD témoigne que le modèle est bon ($RPD = 3,38$), les faibles teneurs en cendres et la faible variabilité relativisent ce résultat. De même que l'humidité, la création d'équation de calibration pour ce paramètre et le travail que nécessite l'entretien du modèle n'est peut être pas pertinente.

Plusieurs échantillons dont les valeurs chimiques de référence étaient connues ont ensuite été repassés sur le Polytec une fois l'équation installée afin d'observer la précision du modèle (**Tableau 8**). Cela permet de vérifier que le Polytec sort des prédictions proches des valeurs chimiques de référence et des prédictions XDS. Cela ne remplace pas la prédiction d'échantillons indépendants, mais il s'agit d'une première étape dans la vérification du modèle.

		Humidité		Protéine		Matière grasse (mét A)	
N° échantillon	Nom	Valeur chimique	Prédiction Polytec	Valeur chimique	Prédiction Polytec	Valeur chimique	Prédiction Polytec
DQ150202-007	Porcelet	12,30	12,50	16,40	16,60	4,70	4,60
DQ150202-011	Porc engrais	12,14	12,30	16,18	16,70	2,46	2,60
DQ150202-012	Truies	12,49	13,60	14,58	15,90	3,35	3,20
DQ150202-013	Truies	12,83	12,30	16,01	16,00	2,58	1,70
DQ150202-014	Porc engrais	12,70	12,70	16,52	16,90	4,34	4,40
DQ150202-015	Porcelet 15-35	13,13	13,10	17,59	17,80	2,75	2,40
		Cellulose		Amidon		Cendres	
N° échantillon	Nom	Valeur chimique	Prédiction Polytec	Valeur chimique	Prédiction Polytec	Valeur chimique	Prédiction Polytec
DQ150202-007	Porcelet	4,35	4,30	39,10	41,70	4,51	4,70
DQ150202-011	Porc engrais	4,79	5,30	43,98	43,60	4,33	3,90
DQ150202-012	Truies	7,29	6,50	38,64	39,10	4,33	4,00
DQ150202-013	Truies	7,63	7,10	35,71	37,50	4,74	4,20
DQ150202-014	Porc engrais	4,85	4,10	43,50	43,90	5,17	4,70
DQ150202-015	Porcelet 15-35	4,81	5,40	39,50	41,20	4,80	3,70

Tableau 8 : Résultats des validations sur le Polytec par comparaison des valeurs chimiques et des prédictions XDS et Polytec pour les aliments pour porcs

La comparaison des valeurs chimiques et des prédictions Polytec pour l'humidité permet de témoigner de la robustesse du modèle. En effet, malgré le temps écoulé entre l'analyse chimique et les mesures par le Polytec, les valeurs sont très proches, l'écart le plus important étant de 1,11 et est compris dans la marge d'erreur de la prédiction. Elles sont systématiquement en dessous du seuil de 14% que *Dumoulin* ne doit pas dépasser.

Les prédictions des teneurs en protéine sont très bonnes puisque l'écart maximum entre les valeurs chimiques et les valeurs prédites est de 1,32, ce qui est cependant très supérieur à la marge d'erreur tolérée par le NIR. La valeur théorique de formulation est comprise entre la valeur chimique et la valeur prédite. Il est par conséquent impossible de conclure sur cet échantillon. En revanche, les prédictions pour les autres aliments sont tout à fait correctes. En ajoutant de nouveaux échantillons et de nouvelles valeurs chimiques, le modèle gagnera en robustesse et en précision.

Le constat est identique pour la prédiction de la matière grasse. Les valeurs chimiques sont très proches des valeurs prédites par le XDS, l'écart le plus important étant de 1,88. Cependant, il s'agit d'un des deux aliments truies analysés. L'autre aliment pour truie a une valeur chimique de 3,35 et une prédiction Polytec de 3,20. Or l'autre aliment pour truies sort avec une valeur chimique inférieure de 1 point, ce qui peut laisser penser qu'il y a eu une erreur au niveau de l'analyse chimique ou que cet aliment était très différent, ce qui peut expliquer une prédiction moins précise de la part du Polytec.

La prédiction de la cellulose brute est également très satisfaisante puisque l'écart le plus élevé entre les valeurs chimiques et les valeurs prédites par le Polytec est de 0,79. Cette valeur est comprise dans la

marge d'erreur NIR admissible pour la prédiction de la cellulose avec ce modèle. Le modèle ne nécessite pas d'être amélioré immédiatement. Le passage de nouveaux échantillons pour vérifier la qualité de la prédiction d'année en année pourra servir à l'optimiser.

Dans le cas de l'amidon, le constat est identique à celui de la matière grasse. Les prédictions sont très proches des valeurs chimiques à l'exception d'un aliment pour truies. Il peut à nouveau s'agir d'une erreur au niveau de l'analyse chimique, ou de la composition physicochimique réelle de cet aliment. Il existe un écart de 3 points entre les deux aliments pour truies, et un seul des deux est correctement prédit. Cependant, la prédiction du deuxième aliment pour truies est proche des valeurs chimiques et prédites de l'autre aliment pour truies, ce qui laisse penser qu'il peut s'agir d'une erreur au niveau de la valeur chimique.

Enfin, les prédictions des cendres sont très bonnes. Cependant, la variabilité des échantillons pour ce paramètre est très faible et ne permet pas d'interpréter de manière pertinente les prédictions. Cependant, la création d'un modèle de prédiction pour les cendres est indispensable puisque la teneur en cendres est indiquée sur l'étiquette du produit, il est donc important d'exercer un contrôle sur ce paramètre.

II.4. Les matières premières

3 courbes de matières premières ont pour l'heure été construites : DDGS de maïs, tourteaux de colza et tourteaux de lin.

L'analyse des matières premières permet d'exercer un contrôle en amont de la production et ajuster les recettes en fonction des caractéristiques physicochimiques des matières premières. Dans l'idéal, une même matière première ne devrait pas présenter de variabilité au niveau des teneurs de ces constituants. Cependant, en fonction de son origine ou des conditions de transport, une certaine variabilité peut être observée. Les résultats de chaque matière première seront interprétés succinctement. Un commentaire global sur l'analyse infrarouge des matières premières sera effectué à la fin de cette partie.

- DDGS de maïs

La liste de correspondance présentant les échantillons sélectionnés pour la calibration ainsi que les valeurs chimiques est présentée en **Annexe 11**.

Les résultats pour les 7 critères d'analyse retenus pour les DDGS sont présentés en **Tableau 9**. Ils présentent le détail des résultats des valeurs chimiques de référence obtenues et des phases de calibration et de cross-validation.

Critères d'analyse	Valeurs chimiques de référence					Résultats des calibrations et cross-validations									Nb termes
	N	Moy	SD	Min	Max	N	Prétraitements	Outliers	R ² c	SEC	R ² cv	SECV	RPD		
Humidité	48	10.17	0.50	9.2	11.3	66	SNV-1-3-3	0	0.84	0.19	0.73	0.26	1.94	6	
Protéine brute	48	26.89	0.90	25.6	29.7	68	SNV-1-3-3	3	0.89	0.27	0.81	0.36	2.30	7	
Matière grasse (méth B)	44	12.08	0.91	9.9	14.3	59	SNV-1-3-3	3	0.94	0.17	0.88	0.25	3.36	7	
Cellulose brute	48	6.78	0.59	6	8.5	67	SNV-1-3-3	2	0.72	0.26	0.56	0.34	1.73	5	
Cendres	48	4.17	0.43	2.5	5	68	SNV-1-3-3	0	0.61	0.26	0.50	0.29	1.42	4	
Amidon	36	3.79	0.64	2.6	5	56	SNV-1-3-3	2	0.90	0.20	0.74	0.33	1.98	8	
Sucres	39	2.02	0.52	0.75	3.1	59	SNV-1-3-3	9	0.26	0.32	0.26	0.33	1.53	1	

Tableau 9 : Résultats des analyses chimiques, calibrations et cross-validation sur les DDGS

Les graphiques présentant les droites de régression pour la calibration et la validation sont disponibles en **Annexe 12**. Ils permettent de comparer les valeurs chimiques de référence/valeurs prédites par le XDS (en abscisse) aux valeurs NIR (en ordonnée).

- Humidité

Les résultats de la calibration de l'humidité montrent une bonne corrélation de la régression ($R^2c=0,84$ et $R^2cv = 0,73$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre ($SEC = 0,19$ et $SECV = 0,26$), ce qui indique que la calibration est bonne. En revanche, si l'on s'intéresse au RPD, le modèle est peu intéressant ($RPD = 1,94$).

- Protéine

Les résultats de la calibration de l'humidité montrent une corrélation de la régression excellente ($R^2_c=0,89$ et $R^2_{cv} = 0,81$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre (SEC = 0,27 et SECV = 0,36), ce qui indique que la calibration est bonne. En revanche, si l'on s'intéresse au RPD, le modèle est peu intéressant (RPD = 2,30).

- Matière grasse

Les résultats de la calibration de l'humidité montrent une corrélation de la régression excellente ($R^2_c=0,94$ et $R^2_{cv} = 0,88$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre (SEC = 0,19 et SECV = 0,25), ce qui indique que la calibration est bonne. Si l'on s'intéresse au RPD, le modèle est intéressant (RPD = 3,36).

- Cellulose brute

Les résultats de la calibration de l'humidité montrent une corrélation de la régression moyenne ($R^2_c=0,72$ et $R^2_{cv} = 0,56$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre (SEC = 0,26 et SECV = 0,34), ce qui indique que la calibration est bonne. En revanche, si l'on s'intéresse au RPD, le modèle est peu intéressant (RPD = 1,73).

- Cendres

Les résultats de la calibration de l'humidité montrent une corrélation de la régression moyenne ($R^2_c=0,61$ et $R^2_{cv} = 0,50$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre (SEC = 0,26 et SECV = 0,29), ce qui indique que la calibration est bonne. En revanche, si l'on s'intéresse au RPD, le modèle est peu intéressant (RPD = 1,42).

- Amidon

Les résultats de la calibration de l'humidité montrent une bonne corrélation de la régression ($R^2_c=0,90$ et $R^2_{cv} = 0,74$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre (SEC = 0,20 et SECV = 0,33), ce qui indique que la calibration est bonne. En revanche, si l'on s'intéresse au RPD, le modèle est peu intéressant (RPD = 1,98).

- Sucres

Les résultats de la calibration de l'humidité montrent une corrélation de la régression très mauvaise ($R^2_c=0,26$ et $R^2_{cv} = 0,26$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre (SEC = 0,32 et SECV = 0,33), ce qui indique que la calibration est bonne. En revanche, si l'on s'intéresse au RPD, le modèle est peu intéressant (RPD = 1,53).

Plusieurs échantillons dont les valeurs chimiques de référence étaient connues ont ensuite été repassés sur le Polytec une fois l'équation installée afin d'observer la précision du modèle (**Tableau 10**).

		Humidité		Protéine		Matière grasse			
N° échantillon	Nom	Valeur chimique	Prédiction Polytec	Valeur chimique	Prédiction Polytec	Valeur chimique	Prédiction Polytec		
DQ140983-007	DDGS	10,30	10,30	27,30	26,90	12,30	12,50		
DQ140983-008	DDGS	9,70	8,80	27,60	27,20	12,40	12,20		
DQ140983-010	DDGS	11,30	10,30	28,50	27,90	10,30	10,70		
		Cellulose		Amidon		Cendres		Sucres	
N° échantillon	Nom	Valeur chimique	Prédiction Polytec	Référence Formulation	Prédiction Polytec	Valeur chimique	Prédiction Polytec	Référence Formulation	Prédiction Polytec
DQ140983-007	DDGS	6,70	6,80	4,10	3,30	4,10	4,10	2,50	1,70
DQ140983-008	DDGS	6,70	6,50	4,10	2,70	4,30	4,40	2,50	1,70
DQ140983-010	DDGS	7,30	6,90	4,10	4,50	4,50	4,10	2,50	1,80

Tableau 10 : Résultats des validations Polytec par comparaison des valeurs chimiques et des prédictions XDS et Polytec pour les DDGS

Au niveau de l'humidité, les prédictions Polytec sont assez proches des valeurs chimiques de référence. Toutefois, dans la mesure où ces mesures ont été réalisées à plusieurs mois d'intervalle, on peut penser que le taux d'humidité a diminué au cours du temps, ce qui peut expliquer les écarts observés. Toutefois, les valeurs prédites par le Polytec sont bonnes puisqu'elles demeurent très inférieures à la limite sur la teneur en humidité autorisée (14%).

Un même constat peut être dressé pour les teneurs en protéine, matière grasse, cellulose et cendres. Les prédictions des équations de calibration sont très bonnes et proches des valeurs chimiques de référence.

Dans le cas de l'amidon et des sucres, aucune valeur chimique de référence ou de prédictions XDS n'étaient disponibles pour les échantillons analysés, c'est pourquoi les prédictions Polytec ont été comparées aux valeurs théoriques de formulation. Même si ces dernières donnent des valeurs optimales et sans doute éloignées des valeurs réelles des échantillons, force est de constater que les prédictions Polytec sont relativement bonnes, 1 point séparant en moyenne la valeur de formulation et la prédiction Polytec pour l'amidon et 0.8 point pour les sucres.

- Tourteaux de colza

La liste de correspondance présentant les échantillons sélectionnés pour la calibration ainsi que les valeurs chimiques est présentée en **Annexe 13**.

Les résultats pour les 6 critères d'analyse retenus pour les tourteaux de colza sont présentés en **Tableau 11**. Ils présentent le détail des résultats des valeurs chimiques de référence obtenues et des phases de calibration et de cross-validation.

Critères d'analyse	Valeurs chimiques de référence					Résultats des calibrations et cross-validations									
	N	Moy	SD	Min	Max	N	Prétraitements	Outliers	R ² c	SEC	R ² cv	SECV	RPD	Nb termes	
Humidité	19	11.6	0.54	9.9	12.4	94	SNV-1-3-3	2	0.90	0.13	0.85	0.16	2.55	5	
Protéine brute	21	33.98	0.74	32.7	35.3	96	SNV-1-3-3	0	0.95	0.18	0.88	0.28	2.91	12	
Matière grasse (méth A)	21	3.13	0.81	1.5	5.2	94	SNV-1-3-3	3	0.95	0.11	0.92	0.14	3.60	10	
Cellulose brute	15	13.26	0.89	11.6	14.6	71	SNV-1-3-3	4	0.93	0.16	0.92	0.18	3.83	4	
Cendres	19	6.82	0.37	6.3	7.5	92	SNV-1-3-3	2	0.91	0.08	0.83	0.12	2.53	8	
Amidon	1	6.6	/	6.6	6.6	64	SNV-1-3-3	7	0.95	0.09	0.89	0.15	3.04	8	

Tableau 11 : Résultats des analyses chimiques, calibrations et cross-validations sur les tourteaux de colza

Les graphiques présentant les droites de régression pour la calibration et la validation sont disponibles en **Annexe 14**. Ils permettent de comparer les valeurs chimiques de référence/valeurs prédites par le XDS (en abscisse) aux valeurs NIR (en ordonnée).

- Humidité

Les résultats de la calibration de l'humidité montrent une bonne corrélation de la régression ($R^2c=0,90$ et $R^2cv = 0,84$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre (SEC = 0,13 et SECV = 0,16), ce qui indique que la calibration est bonne. En revanche, si l'on s'intéresse au RPD, le modèle est peu intéressant (RPD = 2,55).

- Protéine

Les résultats de la calibration de l'humidité montrent une corrélation de la régression excellente ($R^2c=0,95$ et $R^2cv = 0,88$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre (SEC = 0,18 et SECV = 0,28), ce qui indique que la calibration est bonne. En revanche, si l'on s'intéresse au RPD, le modèle est peu intéressant (RPD = 2,91).

- Matière grasse

Les résultats de la calibration de l'humidité montrent une corrélation de la régression excellente ($R^2c=0,95$ et $R^2cv = 0,92$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre (SEC = 0,11 et SECV = 0,14), ce qui indique que la calibration est bonne. Si l'on s'intéresse au RPD, le modèle est intéressant (RPD = 3,60).

- Cellulose brute

Les résultats de la calibration de l'humidité montrent une corrélation de la régression excellente ($R^2c=0,93$ et $R^2cv = 0,92$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre (SEC = 0,16 et SECV = 0,18), ce qui indique que la calibration est bonne. Si l'on s'intéresse au RPD, le modèle est intéressant (RPD = 3,83).

- Cendres

Les résultats de la calibration de l'humidité montrent une bonne corrélation de la régression ($R^2c=0,91$ et $R^2cv = 0,83$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre (SEC = 0,08 et SECV = 0,12), ce qui indique que la calibration est bonne. En revanche, si l'on s'intéresse au RPD, le modèle est peu intéressant (RPD = 2,53).

- Amidon

Les résultats de la calibration de l'humidité montrent une corrélation de la régression excellente ($R^2c=0,95$ et $R^2cv = 0,89$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre (SEC = 0,09 et SECV = 0,15), ce qui indique que la calibration est bonne. Si l'on s'intéresse au RPD, le modèle est intéressant (RPD = 3,04).

Plusieurs échantillons dont les valeurs chimiques de référence étaient connues ont ensuite été repassés sur le Polytec une fois l'équation installée afin d'observer la précision du modèle (**Tableau 12**)

		Humidité		Protéine		MG (mét A)	
N° échantillon	Nom	Valeur chimique	Prédiction Polytec	Valeur chimique	Prédiction Polytec	Valeur chimique	Prédiction Polytec
DQ140983-032	Tx Colza	11,60	11,10	34,90	36,50	2,60	2,60
DQ140983-033	Tx Colza		12,30	33,40	34,00		2,80
DQ140983-034	Tx Colza	12,00	11,80	34,50	35,50	2,90	3,10
DQ140983-035	Tx Colza		11,90	33,40	34,00		2,40
DQ140983-036	Tx Colza	11,90	11,90	33,80	34,60	1,90	2,20
DQ140983-037	Tx Colza	12,00	11,90	33,60	34,00	3,30	4,00
DQ140983-038	Tx Colza	11,10	11,30	33,80	34,50	3,00	2,70
		Cellulose		Amidon		Cendres	
N° échantillon	Nom	Valeur chimique	Prédiction Polytec	Référence formulation	Prédiction Polytec	Valeur chimique	Prédiction Polytec
DQ140983-032	Tx Colza	12,80	12,40	5,50	6,70	6,80	6,90
DQ140983-033	Tx Colza	13,50	13,10	5,50	5,90	6,55	6,90
DQ140983-034	Tx Colza	13,20	12,10	5,50	5,50	6,90	7,10
DQ140983-035	Tx Colza	14,00	12,40	5,50	6,70	6,50	7,00
DQ140983-036	Tx Colza	14,60	12,40	5,50	6,30	6,60	6,80
DQ140983-037	Tx Colza	13,20	12,90	5,50	5,70	6,40	6,70
DQ140983-038	Tx Colza	11,60	12,60	5,50	6,80	6,90	7,10

Tableau 12 : Résultats des validations Polytec par comparaison des valeurs chimiques et des prédictions XDS et Polytec pour les tourteaux de colza

Les prédictions Polytec sont dans très proches des valeurs chimiques pour l'humidité, la matière grasse, la cellulose et les cendres. Elles sont en revanche un peu trop éloignées dans le cas de la protéine. Si l'on se base sur la marge d'erreur NIR acceptable ($2 \cdot \text{SECV} = 0,56$), on constate que ce seuil est très souvent dépassé. Le modèle pourrait donc être amélioré pour la protéine. En effet, afin de diminuer le SECV, le nombre de termes de l'équation a été fixé à 12, ce qui est beaucoup, et ce surparamétrage peut être la conséquence d'une prédiction moyenne. En réduisant le nombre de termes, la marge d'erreur de la prédiction n'augmenterait sans doute pas significativement et la prédiction pourrait être plus précise.

De même que pour les DDGS, aucune valeur chimique ou prédiction XDS n'était disponible afin d'interpréter les prédictions de l'amidon. En se basant sur les valeurs théoriques de formulation et en sachant qu'elles ne tiennent pas compte de la variabilité que l'on peut observer d'une matière première à l'autre, on remarque que les prédictions sont tout à fait acceptables, l'écart maximum observé étant de 1,30. Dans le cas d'échantillons situés dans la même gamme de valeurs, cet écart pas rédhibitoire quant à la précision de la prédiction.

- Tourteaux de lin

La liste de correspondance présentant les échantillons sélectionnés pour la calibration ainsi que les valeurs chimiques est présentée en **Annexe 15**.

Les résultats pour les 5 critères d'analyse retenus pour les tourteaux de colza sont présentés en **Tableau 13**. Ils présentent le détail des résultats des valeurs chimiques de référence obtenues et des phases de calibration et de cross-validation.

Critères d'analyse	Valeurs chimiques de référence					Résultats des calibrations et cross-validations									
	N	Moy	SD	Min	Max	N	Prétraitements	Outliers	R ² c	SEC	R ² cv	SECV	RPD	Nb termes	
Humidité	12	8.91	1.02	6.9	10.1	86	SNV-1-3-3	1	0.95	0.24	0.95	0.27	4.47	3	
Protéine brute	12	32.99	2.45	30	37.6	86	SNV-1-3-3	6	0.93	0.34	0.90	0.43	3.29	6	
Matière grasse (méth A)	12	8.53	0.7	7.3	9.4	86	SNV-1-3-3	6	0.81	0.55	0.79	0.60	2.47	4	
Cellulose brute	/	/	/	/	/	86	SNV-1-3-3	8	0.87	0.20	0.79	0.26	2.39	9	
Cendres	/	/	/	/	/	86	SNV-1-3-3	4	0.81	0.21	0.73	0.26	1.86	9	

Tableau 13 : Résultats des analyses chimiques, calibrations et cross-validations sur les tourteaux de lin

Les graphiques présentant les droites de régression pour la calibration et la validation sont disponibles en **Annexe 16**. Ils permettent de comparer les valeurs chimiques de référence/valeurs prédites par le XDS (en abscisse) aux valeurs NIR (en ordonnée).

- Humidité

Les résultats de la calibration de l'humidité montrent une corrélation de la régression excellente ($R^2c=0,95$ et $R^2cv = 0,95$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de

l'autre ($SEC = 0,24$ et $SECV = 0,27$), ce qui indique que la calibration est bonne. Si l'on s'intéresse au RPD, le modèle est intéressant ($RPD = 4,47$).

- Protéine

Les résultats de la calibration de l'humidité montrent une corrélation de la régression excellente ($R^2c=0,93$ et $R^2cv = 0,90$). En revanche, les valeurs du SEC et du SECV sont plus élevées mais proches l'une de l'autre ($SEC = 0,34$ et $SECV = 0,43$). Toutefois, la plage de valeurs de la protéine est plus grande que pour les autres paramètres, ce qui indique que la calibration est néanmoins bonne. Si l'on s'intéresse au RPD, le modèle est peu intéressant ($RPD = 3,29$).

- Matière grasse

Les résultats de la calibration de l'humidité montrent une bonne corrélation de la régression ($R^2c=0,81$ et $R^2cv = 0,79$). Les valeurs du SEC et du SECV sont proches l'une de l'autre mais élevées ($SEC = 0,55$ et $SECV = 0,60$), alors qu'il n'existe que peu de variabilité entre les échantillons. En revanche, si l'on s'intéresse au RPD, le modèle est peu intéressant ($RPD = 2,47$).

- Cellulose brute

Les résultats de la calibration de l'humidité montrent une bonne corrélation de la régression ($R^2c=0,87$ et $R^2cv = 0,79$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre ($SEC = 0,20$ et $SECV = 0,26$), ce qui indique que la calibration est bonne. En revanche, si l'on s'intéresse au RPD, le modèle est peu intéressant ($RPD = 2,39$).

- Cendres

Les résultats de la calibration de l'humidité montrent une corrélation de la régression moyenne ($R^2c=0,81$ et $R^2cv = 0,73$). De même, les valeurs du SEC et du SECV sont faibles et proches l'une de l'autre ($SEC = 0,21$ et $SECV = 0,26$), ce qui indique que la calibration est bonne. En revanche, si l'on s'intéresse au RPD, le modèle est peu intéressant ($RPD = 1,86$).

Plusieurs échantillons dont les valeurs chimiques de référence étaient connues ont ensuite été repassés sur le Polytec une fois l'équation installée afin d'observer la précision du modèle (**Tableau 14**).

		Humidité			Protéine			Matière grasse (mét B)		
N° échantillon	Nom	Valeur Chimique	Prédiction XDS	Prédiction Polytec	Valeur Chimique	Prédiction XDS	Prédiction Polytec	Valeur Chimique	Prédiction XDS	Prédiction Polytec
DQ140983-045	Tx Lin	9,00		9,10	31,00		31,00	8,80		9,40
DQ140983-046	Tx Lin	9,60		9,30	30,70		30,80	9,20		8,90
DQ140983-049	Tx Lin	9,90		9,30	35,00		34,60	8,10		8,90
DQ140983-047	Tx Lin	10,10		7,90	33,60		33,90	9,20		8,60
DQ140983-050	Tx Lin	9,80		8,70	34,90		35,10	8,30		7,50
DQ150012-134	Tx Lin		8,63	8,30		33,62	33,10		7,74	8,30
DQ150012-133	Tx Lin		10,66	10,6		31,3	33		9,25	8,3
DQ150012-124	Tx Lin		8,01	8,3		32,28	30,2		9,39	9,3
DQ150012-118	Tx Lin		8,39	8,4		34,82	35,2		8,14	8
DQ150012-127	Tx Lin		7,47	7,6		33,96	34,1		7,77	8,1

		Cellulose		Cendres	
N° échantillon	Nom	Prédiction XDS	Prédiction Polytec	Prédiction XDS	Prédiction Polytec
DQ140983-045	Tx Lin	9,99	9,00	4,54	4,30
DQ140983-046	Tx Lin	10,34	9,50	4,28	4,00
DQ140983-049	Tx Lin	9,19	8,00	4,64	4,90
DQ140983-047	Tx Lin	9,27	8,40	4,70	4,40
DQ140983-050	Tx Lin	8,74	8,00	4,44	4,60
DQ150012-134	Tx Lin	9,09	9,00	4,75	4,60
DQ150012-133	Tx Lin	9,79	8,9	4,99	4,2
DQ150012-124	Tx Lin	9,24	8,6	5,78	5
DQ150012-118	Tx Lin	9,04	8,6	4,64	4,4
DQ150012-127	Tx Lin	9,42	8,7	4,89	4,8

Tableau 14 : Résultats des validations Polytec par comparaison des valeurs chimiques et des prédictions XDS et Polytec pour les tourteaux de lin

Les prédictions Polytec de l'humidité sont très proches et dans la majorité des cas inférieures aux valeurs chimiques de référence et des prédictions XDS, ce qui témoigne de la précision du modèle et de la perte d'humidité du produit dans le temps.

Pour la protéine et les cendres, les prédictions sont également très proches des valeurs de référence et des prédictions XDS, et sont comprises dans la marge d'erreur propre à l'analyse infrarouge (2*SECV). Les prédictions Polytec de la matière grasse sont plus éloignées des valeurs de référence, mais la marge d'erreur infrarouge est plus importante pour ce paramètre. Cependant, il n'y a que peu de variabilité au niveau de ce paramètre pour les tourteaux de lin, la précision du modèle pourrait ainsi être améliorée.

Enfin, concernant la cellulose, les prédictions XDS sont assez éloignées des prédictions XDS et se trouvent au-delà de la marge d'erreur acceptable. La précision de la prédiction pourrait être améliorée

en ajoutant de nouveaux échantillons au modèle et réduire le nombre de termes de l'équation afin d'augmenter légèrement le SECV.

Les résultats obtenus sur les matières premières ne peuvent pas être interprétés de la même manière que les produits finis. En effet, interpréter les résultats en se basant sur les indicateurs de performance est moins pertinent puisqu'ils dépendent de la nature des échantillons de la population. Le contrôle de la variabilité des matières premières par la SPIR est le principal enjeu pour *Dumoulin*. En contrôlant les approvisionnements, les impacts sur les produits finis seront à fortiori moins nombreux.

Ainsi, dans le cas des produits finis, la variabilité qui peut être rencontrée influence grandement ces indicateurs, ce qui n'est pas le cas des matières premières. La variabilité au sein des matières premières doit être très faible, afin de ne pas procéder à des changements au niveau des formulations de manière récurrente. C'est pourquoi il est difficile de conclure qu'un modèle de prédiction est peu précis comme dans le cas DDGS avec un R^2 , un SECV et un RPD faibles.

La précision du modèle, contrairement aux produits finis, ne peut se démontrer exclusivement qu'avec la comparaison des prédictions Polytec et des valeurs de référence. Après analyse des résultats de ces trois matières premières, on constate que les prédictions sont très bonnes et permettent de remarquer les différences observées d'un approvisionnement à l'autre. En effet, dans le cas des tourteaux de colza, un des échantillons est nettement plus riche en protéine que le reste des autres échantillons et cela apparaît au niveau de la prédiction : 34,90% en analyse chimique, 36,50% en prédiction Polytec alors que le colza contient généralement de 33 à 34% de protéine. ou bien encore dans le cas des tourteaux de lin où la teneur moyenne en protéine est de 31, et l'analyse infrarouge a permis de détecter des tourteaux contenant jusqu'à 35% de protéine.

Cependant, la faible variabilité observée pour tous les paramètres amène à se poser la question de la pertinence de l'analyse de certains critères, entraînant ainsi un travail supplémentaire dans le futur pour l'entretien de ces courbes. C'est notamment le cas du sucre pour les DDGS où la variabilité n'est que de 1 point. Le modèle, composé d'un unique terme, ne semble en effet que peu pertinent.

II.5. Les aliments floconnés

Aucun modèle suffisamment précis n'a pu être créé à ce jour en raison de la haute spécificité des aliments floconnés. Leur processus de fabrication et leur forme très particulière aboutissent à des allures spectrales dont ne dispose pas le CRA-W. Ainsi, dans le cas des Golden Soy, le processus de floconnage a un effet direct sur le spectre, puisqu'en utilisant l'équation initiale du Full Fat Soya, il a été démontré que cette courbe pouvait davantage s'appliquer aux graines de soja qu'aux flocons de soja (même si l'analyse des graines de soja en tant que telle est impossible car la coque ne permet pas une bonne acquisition spectrale). Il reste donc du travail à effectuer sur les flocons, et une nouvelle équation devra être trouvée au sein de la base du CRA-W afin de pouvoir créer les modèles.

III. Discussion sur la réglementation européenne relative à l'étiquetage des matières premières et des aliments composés pour animaux

Le groupe *Dumoulin* s'est engagé dans l'analyse infrarouge afin de contrôler de manière plus efficace la qualité de ces produits et s'assurer d'être dans les limites tolérées pour les constituants chimiques indiqués sur l'étiquette des produits.

Dumoulin certifie sur les étiquettes de ses aliments respecter les teneurs annoncées en protéine brute, matière grasse brute, cellulose brute et cendres brutes.

Les tolérances réglementaires vis-à-vis de ces quatre constituants sont précisées dans le Règlement (UE) 939/2010 qui modifie l'annexe IV du Règlement (CE) 767/2009 sur les tolérances admises pour les indications d'étiquetage relatives à la composition des matières premières pour aliments des animaux ou des aliments composés pour animaux (Journal officiel de l'Union européenne, 2010). L'annexe IV du Règlement (UE) 939/2010 est consultable en **Annexe 17**.

La teneur indiquée sur l'étiquette dispose d'une limite inférieure et d'une limite supérieure à ne pas dépasser. En cas de dépassement, on parlera de non-conformité et *Dumoulin* sera susceptible de poursuites pour fraude. Cependant, les exigences des éleveurs afin de correspondre aux demandes de performances zootechniques sont bien plus étroites que celles de la législation. En effet, si d'un point de vue réglementaire, un aliment à 32% de protéine peut contenir 29% à 35%, un éleveur ne pourra accepter une teneur à 29%. En plus d'affecter les performances zootechniques, les dépassements des teneurs peuvent jouer sur l'intégrité physiologique des animaux ainsi que sur l'environnement. En effet, une teneur en protéine inférieure à celle indiquée sur l'étiquette provoquera une diminution de la production de l'animal et une dégradation de son état de santé. Une teneur en protéine trop élevée entraînera également un rejet plus important de composés azotés dans l'atmosphère. Le contrôle de ce paramètre dans le cas des aliments pour porcs est ainsi particulièrement important puisque c'est dans ce type d'élevage que les surplus en protéine sont les plus importants.

La spectroscopie proche infrarouge, grâce à l'étude du SEP/SECV du modèle, permet de vérifier si les limites fixées sont respectées. Si un certain écart est toléré quant à la valeur inscrite sur l'étiquette, chaque valeur infrarouge possède également une marge d'erreur, 95% des échantillons étant compris dans un intervalle de $\pm 2 \cdot \text{SEP/SECV}$. Lors d'une analyse infrarouge, il est ainsi possible de vérifier si le produit respecte la réglementation en comparant les limites inférieures et supérieures de la valeur inscrite sur l'étiquette en tenant compte des tolérances réglementaires, et la valeur prédite par infrarouge en tenant compte de la marge d'erreur du modèle.

Dumoulin s'intéresse surtout aux limites inférieures décrites dans la réglementation car c'est à ce niveau que sont concernées la majorité des plaintes clients, suite à une diminution de la production par exemple (non-conformité). Cependant, le dépassement des limites supérieures peut également s'avérer problématique pour *Dumoulin*. En effet, si un éleveur ne se plaindra pas d'un surplus en protéine dans

ses aliments, cela représente un manque à gagner très important pour une entreprise de la taille de *Dumoulin* (production de 500.000 tonnes/an), à l'heure où le marché est très concurrentiel et les sources de protéines d'origine européenne devenant plus rares face à l'augmentation de la demande et des prix.

Cela donne un indicateur sur la qualité du produit, mais en aucun cas un résultat définitif. En effet, dans le cas où les limites réglementaires sont dépassées par la marge d'erreur infrarouge, l'échantillon analysé peut tout à fait se trouver dans les limites acceptables. De plus, cela nécessite d'avoir un modèle très robuste, ce qui ne s'obtient qu'après plusieurs années et l'ajout de nouveaux échantillons au modèle. De plus, certains paramètres, en particulier les cendres, s'avèrent difficiles à quantifier sur des aliments composés pour animaux puisque les teneurs sont généralement très faibles, ne dépassant que rarement 8%, et il existe peu de variabilité entre les échantillons. Sur le plan de la SPIR, la création de modèles pour des paramètres possédant peu de variabilité est peu pertinente. Cependant, afin d'avoir un œil sur le respect de la réglementation, il est intéressant pour *Dumoulin* d'avoir ces modèles de prédiction à disposition.

Plusieurs prédictions Polytec des aliments composés et des matières premières ont ainsi été réutilisées afin de vérifier si les valeurs prédites respectaient la législation. Les résultats et conclusions sont présentées en **Annexe 18**. Cela permet de donner une première interprétation sur la qualité des produits analysés et la robustesse du modèle. Par exemple, la majorité des paramètres des échantillons porcs analysés sortent du cadre de la réglementation, ce qui peut signifier que le modèle n'est sans doute pas assez robuste et mériterait d'être amélioré. En revanche, dans le cas des tourteaux de lin, si la majorité des échantillons sortent du cadre légal, cela peut s'expliquer par la variabilité qu'il peut y avoir sur les matières premières, conditionnant ainsi une reformulation des recettes d'un approvisionnement à l'autre.

- Le cas particulier de l'humidité

Enfin, dans le cas de l'humidité, *Dumoulin* n'inscrit pas la teneur de ce paramètre sur ses étiquettes. En alimentation animale, la teneur maximale autorisée est de 14%. Sur l'ensemble des produits finis étudiés, cette limite n'a jamais été dépassée, ce qui prouve que *Dumoulin* maîtrise ce paramètre. L'étude de ce paramètre prend tout son sens pour les matières premières, ces dernières ne devant pas être contaminées par des moisissures et impacter la qualité du produit fini.

Il est cependant possible de faire le lien avec l' A_w des produits et la quantité de liquides ajoutés lors de la fabrication. Pour se développer, les bactéries, champignons et levures nécessitent une A_w minimum en dessous de laquelle aucun développement n'est possible. La stabilité microbienne n'est pas un problème en cas d'une A_w inférieure à 0.60 (Lowe et Kershaw, 1995). Les seuils de développement sont les suivants :

- ✓ $A_w = 0,75 \rightarrow$ Seuil minimal pour la croissance bactérienne

- ✓ $A_w = 0,62$ → Seuil minimal pour la croissance des levures
- ✓ $A_w = 0,65$ → Seuil minimal pour la croissance des champignons xérophiles
- ✓ $A_w = 0,61$ → Seuil minimal la croissance des champignons
- ✓ $A_w = 0,60$ → Aucun risque de prolifération bactérienne et champignons (Timmons, 2007)

Le contrôle de l' A_w est important sur les matières premières, car les moisissures sont adaptées à des A_w faibles et peuvent se développer dans les grains stockés, entraînant mauvaises odeurs, perte de matière sèche et de nutriments, durcissement des grains et apparition de mycotoxines. L' A_w a un rôle important dans la stabilité chimique. Elle peut entraîner un brunissement non enzymatique, une oxydation lipidique, une dégradation des vitamines et des nutriments, des réactions enzymatique, une dénaturation des protéines ainsi qu'une gélatinisation de l'amidon. (Maciorowski et al., 2007)

Ainsi, en utilisant des isothermes de sorption, il est possible de relier l' A_w et la teneur en eau des aliments, afin de voir quelles quantités de liquides peuvent être ajoutées jusqu'à atteindre le seuil de développement des moisissures tout en ayant un œil sur l'humidité du produit, et en particulier la mélasse. Il est en effet plus rentable pour l'entreprise de monter en A_w jusqu'à être limitant en ajoutant des liquides permettant d'augmenter la teneur en protéine par exemple, que d'augmenter la teneur en protéine en ajoutant des matières premières solides, dont les prix d'achats sont bien plus élevés. C'est là tout l'objectif de l'analyse on-line. Les résultats des équations de calibration qui seront obtenues sur l'humidité permettront de déterminer les quantités de mélasse et de vinasse qui peuvent être ajoutées lors de la fabrication.

En parallèle, le contrôle de l' A_w et de l'injection des liquides permettra de vérifier la qualité du séchage ainsi que l'aspect physique des produits via l'analyse de la dureté (par compression) des granulés. Selon l'aliment, la dureté se doit d'être plus ou moins élevée mais ne doit jamais être inférieur à 7 Newton.

IV. Discussion sur la méthodologie

Cette partie sera consacrée à l'étude des limites de la méthodologie mise en place pour construire les différentes équations de calibration. L'objectif est de dégager certaines pistes d'améliorations afin de rendre les modèles créés à ce jour et ceux à venir plus robustes et plus précis. Afin de rendre la discussion cohérente et logique, l'argumentation sera effectuée de façon chronologique, depuis la prise d'échantillons jusqu'à la création des modèles.

- L'échantillonnage

Lors du choix des échantillons qui serviraient à la construction des équations de calibration, trois problèmes ont été rencontrés. Tout d'abord, si la variabilité entre les échantillons est un critère recherché, cela peut constituer un obstacle lorsque plusieurs aliments se révèlent être très différents des autres au niveau de leur composition chimique. Ce fut notamment le cas au niveau des correcteurs protéiques pour la protéine. Si les correcteurs protéiques avec et sans urée se trouvent dans une même gamme de teneur en protéine, ce n'est pas le cas des aliments Hipro 58, qui possèdent une teneur en protéine supérieure due en partie à un apport en urée plus important. Il existait donc un trou entre les correcteurs protéiques avec urée et les Hipro 58 (45-50%). Cet écart a pu être comblé en créant un mix d'aliments. Cette technique a fonctionné car l'écart, même s'il entraînait des perturbations au niveau du modèle, était relativement faible. Ce n'est néanmoins pas le cas des aliments pour porcs, et plus particulièrement les aliments correcteurs à destination des truies en gestation. Ces aliments ayant la particularité d'être riche en protéine (30% contre 16% en moyenne pour le reste de la population) et très pauvre en amidon (9% contre 40% en moyenne pour le reste de la population), il présentait un risque de perturber les modèles pour ces deux paramètres car trop différents de la population moyenne des échantillons. Dans un souci d'être le plus exhaustif possible au niveau des aliments pour porcs, les échantillons de cet aliment particulier ont été considérés comme des outliers de calibration pour ces deux paramètres, mais ont été conservés pour l'analyse des autres critères.

Un autre moyen de combler un écart entre des échantillons atypiques et le reste de la population est de prélever davantage de ces échantillons afin de constituer un groupe suffisamment conséquent pour ne pas perturber le modèle. Cependant, cela requiert de disposer d'une grande disponibilité au niveau des échantillons et cela constitue le deuxième problème rencontré lors de la phase de prélèvement. Pour une même catégorie d'aliments, tous ne sont pas produits en égales quantités, ce qui a constitué un frein à la constitution d'une base d'échantillons conséquente et variée. Face à cette impasse, il a alors été convenu que, dans le cas d'échantillons peu nombreux et proches de la moyenne de la population, ils pourraient toutefois être prédits, les modèles étant composés d'aliments similaires. Pour les aliments trop différents, ils seraient pour l'instant écartés des analyses infrarouges.

La variabilité observée témoigne de la grande spécificité de chaque équation de calibration. Cela amène à penser que de nouvelles calibrations devront être réalisées afin d'inclure tous les aliments et ainsi exercer un contrôle qualité similaire pour tous les produits.

Enfin, le troisième problème lié à l'échantillonnage provient de la composition chimique propre à chaque échantillon d'un même aliment. En effet, si les formulations permettent de définir les teneurs des constituants de chaque aliment, celles-ci varient légèrement d'une production à l'autre, selon les quantités produites, les quantités d'ingrédients utilisées ou les problèmes techniques éventuels rencontrés. Il peut également arriver qu'une production soit mauvaise et que l'aliment ne présente pas les caractéristiques attendues. De même, dans le cas de l'analyse des matières premières, leurs compositions chimiques peuvent être très différentes selon la période de l'année à laquelle elles sont livrées, leur origine et leurs conditions de transport (humidité en particulier). Ces écueils ne peuvent être détectés lors de l'échantillonnage et sont inhérents au milieu industriel. Ainsi, la qualité de la calibration va dépendre à la fois de la variabilité entre les aliments mais aussi de la variabilité entre deux aliments censés être identiques. C'est par exemple le cas de l'Hipro 58 dont la teneur en protéines a été évaluée à 45.5% grâce à l'analyse chimique, alors qu'il est attendu à 52%. Cet écart peut à la fois provenir d'une erreur dans l'analyse chimique, ou du fait que ce produit est non conforme. Ainsi, les conformités pour construire les modèles de calibration ne peuvent être détectées à l'avance et ont une incidence sur la calibration.

- L'acquisition spectrale

Plusieurs facteurs pouvant perturber l'acquisition spectrale par les spectromètres Polytec et XDS FOSS ont été relevés. Tout d'abord, concernant le Polytec, l'acquisition s'effectue par le haut et prend en compte la vapeur d'eau présente dans l'air entre la tête de lecture et l'échantillon. Les pics de vapeur d'eau apparaissent sur les spectres entre 1350 et 1400 nm et peuvent masquer une partie de l'information contenue dans cette région spectrale (**Figure 27**).

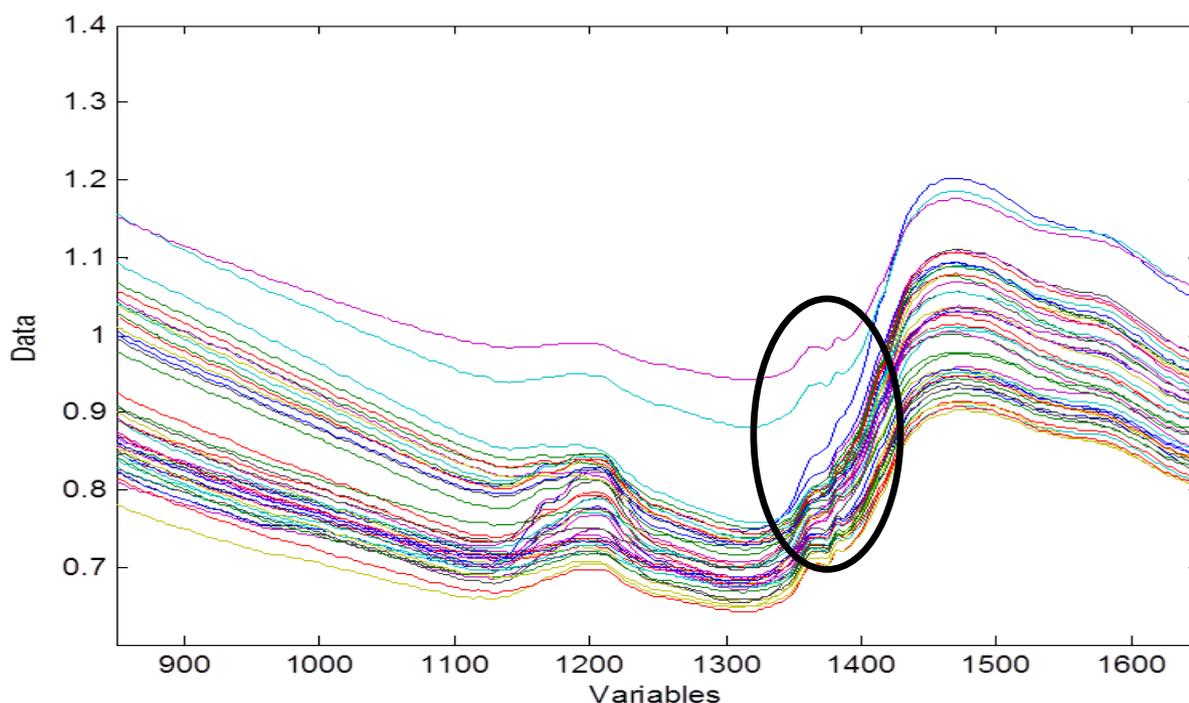


Figure 27 : Aperçu des pics de vapeur d'eau dans le cas de l'analyse en laboratoire

En dépit de ces déviations spectrales, l'appareil Polytec est un des spectromètres répondant au mieux aux attentes de *Dumoulin*. Le temps d'analyse est très court (quelques secondes), ce qui constitue un atout majeur pour les analyses on-line où le produit est continuellement en mouvement. Les spectromètres à barrettes de diodes sont par ailleurs composés d'un faible nombre de pièces mobiles, augmentant la robustesse de l'appareil. Il est donc particulièrement adapté au milieu industriel.

D'autres paramètres ont pu influencer la mesure spectrale du Polytec. L'appareil a été installé dans une pièce soumise à des variations en termes de températures, de luminosité et d'hygrométrie. L'échantillon analysé étant à l'air libre, des déviations supplémentaires ont pu se produire sur les spectres.

Concernant le XDS FOSS, l'acquisition spectrale s'effectue par le bas. Cela présente l'avantage de ne pas analyser les particules de vapeur d'eau, mais une telle analyse prend en compte les particules fines qui tombent par gravité au fond de la cellule d'analyse et peuvent perturber le scan. Cela ne pose pas de problèmes dans le cas des farines pour porcs, ou les granulés qui contiennent peu de particules fines. En revanche, l'acquisition spectrale peut être biaisée pour les aliments extrudés. Par ailleurs, une acquisition par le bas nécessite un nettoyage systématique de la cellule entre le passage de deux échantillons pour éviter les contaminations. Cependant, il est difficile d'éliminer l'intégralité des particules fines et il n'est pas impossible qu'une partie d'un échantillon ait pu être en contact avec un échantillon différent. C'est pourquoi, les deux mesures de chaque échantillon ont été systématiquement comparées. En utilisant l'écart-type au spectre moyen, aussi appelé RMS, cela permet de vérifier si les spectres sont similaires. Plus cette valeur est élevée, moins la répétition sera bonne. La meilleure solution en cas d'un RMS trop grand est généralement de repasser l'échantillon ou de le supprimer s'il n'est pas assez homogène, comme c'est le cas des farines pour volailles.

Avant l'installation du spectromètre à Andenne, un autre appareil Polytec était disponible au CRA-W. La distance focale, c'est-à-dire la distance séparant la tête de lecture et l'échantillon, avait été fixée au CRA-W à 45 cm. Cependant, cette distance a ensuite été réduite à 27 cm une fois l'appareil acquis. Si la bibliographie (Williams, 2077) précise que la distance ne doit pas dépasser 50 cm, une modification de cette distance peut avoir une incidence sur l'absorbance. Cela a été constaté sur les aliments extrudés. L'absorbance pour un même échantillon est différente selon la distance focale (**Figure 28**). Cette différence a obligé à repenser les résultats qui avaient été obtenues sur le premier Polytec. Ainsi, les premiers échantillons d'aliments pour porcs et une partie des correcteurs protéiques ou des matières premières n'ont pas été retenus. De même, pour les extrudés, comme précisé précédemment, les premières mesures spectrales n'ont pas été retenues pour la matière grasse, le biais étant trop important en les conservant.

La variabilité observée lors de l'acquisition spectrale peut ainsi influencer sur la précision de la mesure, être à l'origine des décalages observés dans les spectres et diminuer la précision de la calibration et de ses indicateurs de performance.

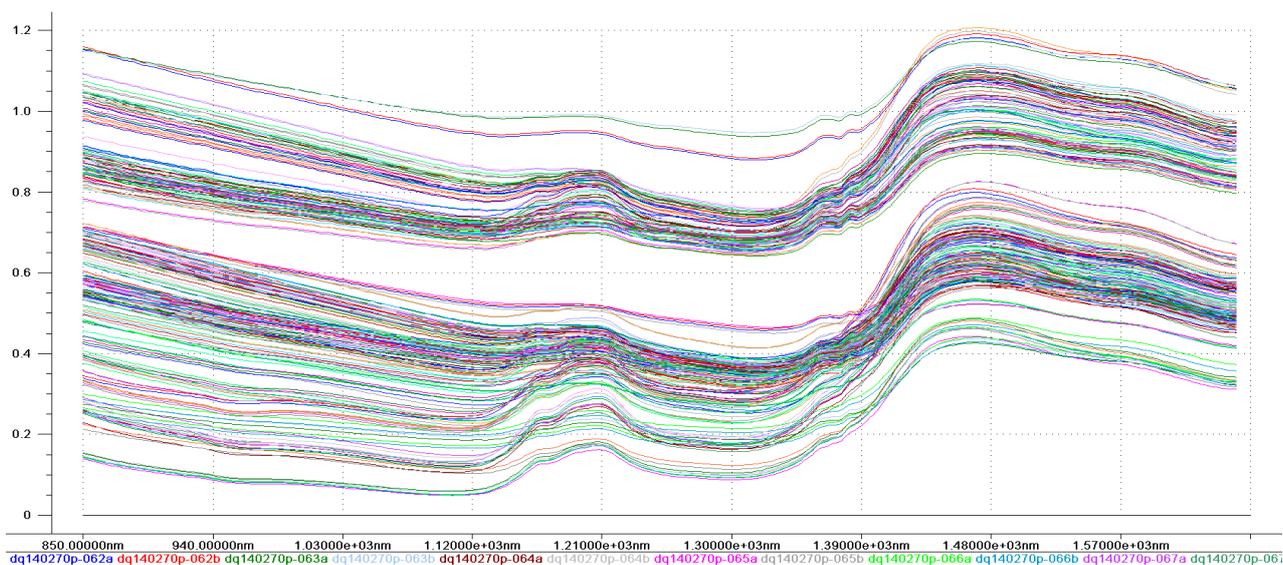


Figure 28 : Conséquences spectrales de la modification de la distance focale. En haut : spectres acquis avec une hauteur de tête fixée à 45 cm. En bas : spectres acquis avec une hauteur de tête à 27 cm.

- Les analyses chimiques de référence

L'envoi d'échantillons en analyse chimique est indispensable pour construire des équations de calibration robustes permettant d'établir des prédictions de grande précision. Pour autant, le nombre d'analyse chimique a néanmoins été limité puisque le XDS FOSS possède l'avantage d'être un outil très puissant et capable de sortir des prédictions très proches des valeurs obtenues par analyse chimique grâce à l'intégration des spectres issus du XDS FOSS aux équations du CRA-W comprenant plusieurs milliers d'échantillons recueillis depuis plus de dix ans.

Dans de rares cas, des erreurs se sont glissées dans les analyses chimiques. Toutefois, rien ne remplace l'exactitude d'une valeur chimique, et pour disposer d'équations plus robustes, il aurait pu être judicieux de faire davantage appel à l'analyse chimique, notamment dans le cas des critères d'analyse secondaires (cendres, cellulose, sucre, amidon).

- Remarques sur les phases de calibration et de cross-validation

Les échantillons choisis, les valeurs chimiques de référence ainsi que les vastes bases de données du CRA-W de plusieurs milliers d'échantillons recueillis depuis dix ans ont permis de construire des modèles de calibration pour cinq à sept paramètres en fonction des aliments. Les modalités des paramétrages appliquées sur les bases de données du CRA-W ont pu être transférées sur les aliments de *Dumoulin* dans un souci d'uniformisation des données.

- Le choix des prétraitements appliqués aux spectres bruts

Comme précisé dans la partie bibliographie au paragraphe « IV.2. Les prétraitements des spectres », divers traitements mathématiques ont été appliqués aux spectres afin d'améliorer la précision de l'acquisition spectrale. Le choix des combinaisons des prétraitements dépend de la subjectivité de l'opérateur et de son expérience. Les mêmes modalités des prétraitements ont été appliquées à l'ensemble des équations de calibration créées. La multiplication des méthodes de prétraitements n'est pas pertinente dans la mesure où les allures spectrales ne diffèrent que peu d'un aliment à l'autre.

Deux prétraitements ont donc été appliqués systématiquement, à savoir un traitement SNV et une dérivée première. La moyenne spectrale n'a pas été appliquée dans la mesure où le Polytec effectue déjà une moyenne de 1500 scans lors de l'acquisition spectrale. De plus, le nombre d'échantillons retenus n'étant pas forcément très important, conserver les répétitions était intéressant afin d'avoir un nombre de spectres suffisants.

La méthode Gap Segment a été utilisée pour la dérivée première, avec un gap size = 3 (taille de saut) et un gap segment = 3 (taille de la fenêtre). Selon l'expertise du CRA-W, essayer d'autres combinaisons n'est pas pertinent. Les valeurs retenues pour ces paramétrages sont des valeurs moyennes que le CRA-W utilise dans la majorité des cas. En effet, la dérivée première correspondant à une étude de pente, en choisissant des valeurs de gap plus élevées, une partie de l'information serait perdue.

L'intérêt de ces prétraitements réside dans le fait qu'ils permettent de repérer plus efficacement des résultats aberrants ou différant de la moyenne de la population, facilitant ainsi le travail lors de la calibration.

- La sélection des outliers de calibration

Les outliers peuvent être mis en évidence de plusieurs façons. Il est tout d'abord possible de les détecter sur les spectres bruts. Généralement, un outlier présentera soit une allure spectrale différente des autres échantillons, soit les valeurs des absorbances aux longueurs d'onde sélectionnées seront plus ou moins importantes que le reste des échantillons destinés à la calibration. La sélection des outliers à ce moment est cependant peu pertinente car à ce stade, aucun prétraitement n'a été effectué sur les spectres bruts, ces derniers pouvant atténuer les écarts observés. De plus, il peut s'agir d'un échantillon pouvant apporter de la robustesse au modèle en élargissant la gamme de valeurs prédites par exemple.

Les outliers peuvent également être observables à l'aide d'une ACP. Ils peuvent être détectés à l'aide de la statistique T^2 de Hotelling, permettant de déterminer l'éloignement des échantillons sur les axes de l'ACP et ainsi montrer leur influence sur le reste la population. Les outliers peuvent être identifiés sur le graphique des *scores* ou sur la projection graphique des valeurs, les outliers étant situés au dessus de la limite statistique fixée (Lecler, 2014). Cependant, aucun outlier n'a été supprimé lors

d'une ACP car leur sélection s'effectue selon la position de l'échantillon vis-à-vis de la moyenne de la population. Cette méthode de sélection ne s'intéresse pas aux allures spectrales et aux absorbances et l'information intéressante pour la calibration d'un paramètre peut se trouver dans la moyenne de la population. Ainsi, un échantillon supprimé lors de l'ACP peut tout à fait se révéler bénéfique à la robustesse du modèle.

C'est pourquoi il est préférable de procéder à la sélection des outliers une fois la régression effectuée en s'aidant des graphiques « Reference vs Predicted » présentant les droites de régression des calibrations et des cross-validations. L'étude des outliers a été menée sur les droites de régression des cross-validations, les résultats attendus témoignant de la précision de l'équation résidant dans l'analyse des indicateurs de performance de la cross-validation (R^2_{cv} , SECV et RPD). Cette opération a été mise en place pour chacun des critères d'analyse de tous les aliments. Dans chaque cas, les échantillons s'éloignant de la droite de régression ont été étudiés sans toutefois être toujours supprimés. En effet, certains échantillons se trouvent sur une même ligne, et en supprimer un n'améliorerait pas le modèle. Pour cela, il faudrait tous les supprimer, ce qui n'est pas forcément réalisable compte tenu du fait que la population n'est composée que de quelques dizaines d'échantillons. C'est à cette étape que le fait d'avoir gardé les répétitions des mesures s'avère utile. En effet, il est possible de comparer les deux mêmes échantillons et de supprimer celui qui s'écarte le plus de l'axe de la droite tout en conservant au moins un exemplaire de l'aliment considéré. De même, les échantillons s'éloignant de l'axe de la droite de régression mais constituant des extrema n'ont jamais été supprimés car ils permettent d'augmenter la gamme de valeurs prédites. Enfin, dans le cas de gammes de valeurs très restreintes, notamment dans le cas des cendres, très peu d'outliers ont été supprimés. En effet, même si certains s'éloignent de l'axe de la droite de régression, l'écart entre les valeurs minimum et maximum est tellement faible que cela n'affectera nullement la précision du modèle.

Cependant, la sélection des outliers demeure une technique relativement confuse. En effet, elle dépend de la subjectivité de l'opérateur et de son expérience en la matière. Enfin, l'existence d'outliers de calibration ne permet pas de préciser la nature de l'erreur constatée. Il est impossible de déterminer s'il s'agit d'une erreur au niveau de la valeur chimique de référence, si l'échantillon concerné est trop différent du reste de la population, ou si le spectre est aberrant en raison d'un problème survenu lors de l'acquisition spectrale.

- Discussion sur les indicateurs de performance des calibrations et des cross-validations

Quatre indicateurs de performances de la calibration ont été utilisés pour interpréter les résultats : le coefficient de corrélation R^2 qui caractérise la part de variation des valeurs de référence expliquée par la calibration et la cross-validation, les SEC et SEP qui caractérisent la précision théorique de la calibration et de la prédiction, le biais n'étant pas pris en compte, et le RPD qui permet d'évaluer la précision de la prédiction en la comparant à la dispersion (écart-type) de la population. Toutefois, les

coefficients de répétabilité des analyses chimiques (erreurs standards de laboratoire) n'ont pas été utilisés dans la mesure où tous les échantillons n'ont pas été envoyés en analyse chimique.

Plusieurs modèles créés possèdent un R^2 , un SEC et un SECV faibles. Cela ne signifie pas pour autant que le modèle doit être abandonné, mais plutôt que la faible plage de valeurs ne permet pas d'établir des prédictions plus précises. Dans le cas des cendres ou du sucre par exemple, un R^2 inférieur à 0.70 et un SECV de 0.20 sont sans doute les meilleurs résultats que l'on puisse obtenir dans la mesure où les valeurs sont comprises dans un intervalle de 2 à 4%. De la même façon, un SEC et un SECV élevé ne signifient pas que le modèle n'est pas précis puisqu'ils dépendent de l'écart-type des échantillons de la population utilisée pour la calibration. Une gamme de valeurs très élevée, comme c'est le cas pour la protéine pour les correcteurs protéiques (35-50%) ou la matière grasse des aliments extrudés (5-35%), augmentera significativement ces indicateurs.

Cependant, l'utilisation du RPD ($RPD = SD/SECV$) en tant qu'indicateur de performance s'est heurtée à plusieurs limites. Tout d'abord, le RPD dépend de l'écart-type des valeurs des échantillons ayant servi à la calibration. Ainsi, si la gamme de valeurs chimiques (ou prédites par le XDS FOSS) est étroite, l'écart-type diminuera, et le RPD également. Pour autant, cela ne signifiera pas que le modèle n'est pas intéressant. De même, une large gamme de valeurs entrainera automatiquement une augmentation du SEC et du SECV, mais également du SD. Mais dans le cas où le SD n'augmente pas aussi vite que le SECV, le RPD sera faible et le modèle sera alors une nouvelle fois considéré comme peu intéressant.

Ces constatations sur le RPD ont d'ailleurs été largement reprises dans les écrits scientifiques. C'est le cas de l'article *Why you don't need to use RPD* (Minasny et McBratney, 2013). Ainsi, cet article remet en cause l'interprétation que Phil Williams a fait du RPD et qui a été reprise lors de l'interprétation des résultats exposés précédemment, à savoir qu'un $RPD > 3$ permet de qualifier le modèle d'intéressant, et qu'un $RPD > 5$ permet d'utiliser ledit modèle pour les contrôles qualité (Williams et Norris, 1987). Afin de montrer que les opinions de la communauté scientifique divergent concernant le RPD, d'autres écrits mentionnent l'existence de catégories définies arbitrairement afin d'interpréter le RPD. Ainsi, un $RPD > 2$ qualifie le modèle d'excellent, un RPD compris entre 1.4 et 2 le qualifie de bon, et un $RPD < 1.4$ le qualifie de mauvais et donc non utilisable.

L'article reprend enfin les remarques de Bellon Maurel et al. (2010), qui précisent que pour interpréter correctement le RPD, les valeurs doivent suivre une loi normale et être centrées autour d'une valeur moyenne. Dans le cas des aliments analysés, cette condition ne peut être remplie dans la mesure où les échantillons sont très différents les uns des autres, tant au niveau de la composition chimique que de leur aspect, soit farineux, soit granulé, soit extrudé. Dans l'optique où les échantillons ne suivent pas une distribution normale, il est recommandé d'utiliser un autre indicateur que le RPD pour mieux représenter la dispersion de la population d'échantillons, à savoir le RPIQ, correspondant au Ratio de la Performance avec la distance InterQuartile ($IQ = Q_{75} - Q_{25}$) (Bellon Maurel et al., 2010).

Ainsi, il est plus pertinent de discuter de la précision des calibrations et des cross-validations en s'intéressant au R^2 , au SEC et au SECV, la largeur des plages de valeurs étant très différentes selon les produits et les paramètres, et de les analyser séparément. Cependant, compte tenu du fait qu'en plus des valeurs chimiques, les autres échantillons ont été prédits grâce aux spectres XDS sur la base d'équations regroupant des milliers d'échantillons, cette particularité de nos calibrations peut apporter des premiers éléments de réponse sur la précision des calibrations.

- Discussion sur la précision des prédictions

Lorsqu'un échantillon est prédit à l'aide d'un modèle de calibration ne contenant pas ce dernier, un autre indicateur peut servir à évaluer la précision de la précision. Il s'agit de la distance de Mahalanobis, ou GH. Chaque spectre correspond à un point dans un domaine à X dimensions. Le GH est la distance qui sépare l'échantillon considéré et le point moyen de la population. Généralement, la valeur du GH ne doit pas dépasser 3 ou l'échantillon est considéré comme trop différent du reste de la population et ne peut être prédit correctement par cette équation. Cependant, cette valeur n'est pas immuable. Dans le cas d'une large gamme de valeurs, le GH peut tout à fait dépasser 3 et la prédiction pourra être considérée comme bonne. A l'inverse, un GH faible ne garantit pas une prédiction précise.

Dans le cas du projet, le GH a été utilisé à deux moments. Tout d'abord de la prédiction des spectres XDS en utilisant les équations du CRA-W. Les spectres ne faisant pas partie de la base, utiliser le GH pour vérifier si les spectres pouvaient être prédits avec cette équation s'est révélé particulièrement pertinent. Enfin, le GH a également été utilisé lors de la prédiction des échantillons ayant servi à la calibration et d'autres échantillons indépendants sur le Polytec. Sur le Polytec, il est possible de fixer une valeur de GH à ne pas dépasser : c'est la *leverage limit*. Si cette limite est dépassée, le Polytec signale que la prédiction n'est pas précise.

- Discussion sur les coefficients appliqués aux variables lors de la régression PLS

La mise en place de modèles de calibration repose sur l'analyse des absorbances spécifiques aux liaisons chimiques de l'échantillon. Cette corrélation est quantifiable par la PLS en s'intéressant à l'analyse des coefficients des variables appliqués à absorbances aux longueurs d'onde associées (**Figure 29**). Lorsque les coefficients sont de 0, cela signifie que l'on pourrait supprimer les longueurs d'onde correspondantes sans modifier la précision du modèle. Des coefficients élevés expliquent la part de l'information expliquée à ces longueurs d'onde pour un paramètre chimique donné. Les coefficients ne doivent cependant pas être trop élevés afin d'avoir un modèle stable sur l'ensemble des longueurs d'onde analysées.

La particularité de l'étude des coefficients réside dans le fait qu'augmenter le nombre de termes de l'équation permet de diminuer le SECV et donc rendre le modèle plus précis, mais augmente dans le même temps les coefficients appliqués aux absorbances aux longueurs d'onde correspondantes. Et augmenter trop fortement les valeurs des coefficients rend à terme le modèle moins précis.

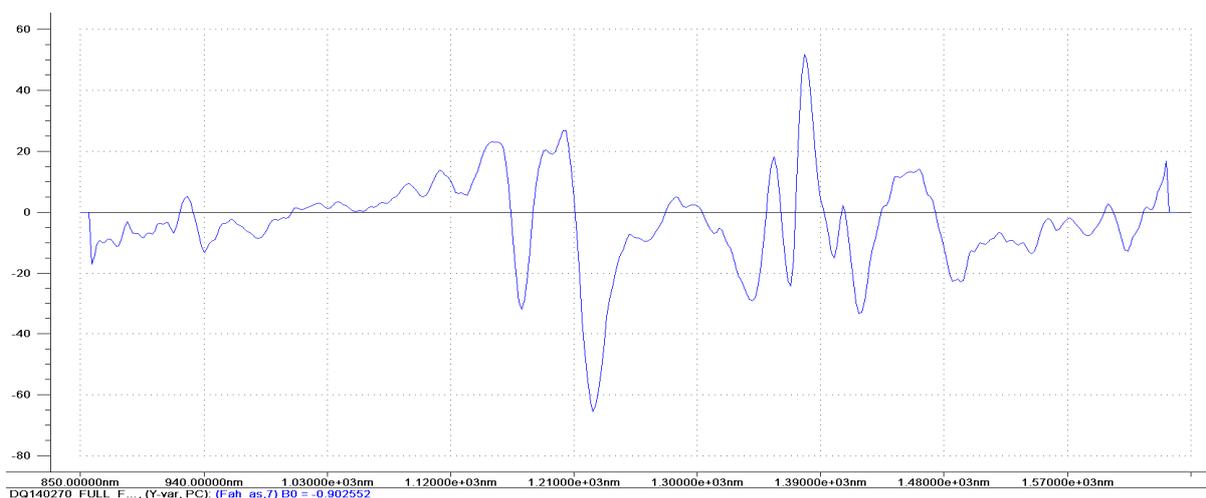


Figure 29 : Graphique des coefficients appliqués aux absorbances aux longueurs d'onde correspondantes

Cette méthode n'a pas été utilisée mais aurait pu être appliquée pour les paramètres chimiques principaux selon les aliments, par exemple la matière grasse pour les aliments extrudés ou la protéine pour les correcteurs.

L'étude des coefficients appliqués à chaque longueur d'onde lors de la régression PLS pourrait servir à expliquer les résultats obtenus, et notamment le cas de la matière grasse pour les aliments extrudés, la majorité de l'information se trouvant au-delà du proche infrarouge, et ainsi proposer des pistes d'amélioration des modèles afin de mettre en évidence l'information utilisable pour la calibration des divers paramètres. De même, il s'avère que les valeurs des coefficients retenus pour la calibration des porcs sont beaucoup trop importante (300 pour l'humidité et 150 pour la protéine) ce qui explique que le modèle ne soit pas très précis. Ainsi, en s'intéressant aux coefficients, il serait possible de trouver un équilibre dans la précision du modèle entre la valeur du SECV et le nombre de termes de l'équation.

- Discussion sur le choix de l'application d'une cross-validation

Le principe de la cross-validation est d'appliquer une première validation en utilisant les mêmes échantillons ayant servi à la calibration. Il ne s'agit généralement que d'une première validation, la véritable validation s'effectuant sur un set d'échantillons indépendants. Pour chaque cross-validation effectuée, un total de cinq sets de validations a été choisi. En fonction du nombre d'échantillons recueillis pour la calibration, le nombre d'échantillons choisis aléatoirement pour chaque set variait d'une quinzaine à près d'une centaine dans le cas des correcteurs. Cependant, une quinzaine d'échantillons apparaît être faible pour établir une cross-validation pertinente. Ainsi, ajouter à l'avenir davantage d'échantillons dans les sets de calibration permettra d'effectuer des cross-validations d'une qualité et d'une précision supérieure.

Peu de validations sur des sets indépendants ont été réalisées pour une raison majeure : la prédiction des échantillons par les équations du CRA-W. Leur robustesse et le nombre d'échantillons les constituant fait foi quant à la qualité de la prédiction qu'elles permettent d'obtenir.

Il aurait cependant été intéressant de tester un set d'échantillons indépendant complet pour au moins un aliment afin de vérifier la précision des prédictions. Cela pourra être réalisé à l'avenir sur les autres produits finis et matières premières qu'il reste à analyser (flocons, aliments pour volailles, etc.).

- Discussion sur les pistes d'amélioration des méthodes de calibration

Les équations de calibration ont été construites à partir de la régression PLS. Cependant, d'autres méthodes de calibration existent et peuvent se révéler plus précises. C'est le cas de la méthode LOCAL, mise en place par les chercheurs de FOSS. Le principe de cette méthode est simple. Un échantillon inconnu est scanné et un spectre est obtenu. Le spectre est projeté dans une vaste banque de données et est comparé aux autres spectres. Des valeurs prédites temporaires lui sont alors appliquées. Une régression classique (PLS, linéaire) est ensuite appliquée afin d'identifier le spectre inconnu. Une calibration LOCAL présente l'avantage d'améliorer la précision de la prédiction, qui tend à diminuer au fil du temps, puisqu'elle s'effectue sur de grandes banques de données couvrant de larges plages de valeurs chimiques. Cette méthode permettrait donc de contourner les problèmes rencontrés au cours du projet concernant les performances des calibrations, cross-validations et prédictions.

Cependant, la méthode LOCAL se heurte à deux problèmes. Elle nécessite tout d'abord la disponibilité d'une grande banque de données (plusieurs dizaines voire centaines de milliers d'échantillons). Le CRA-W dispose de plusieurs banques de cette taille, mais pas pour tous les aliments étudiés. Enfin, cette méthode, créée par FOSS, n'est applicable que sur le logiciel WinISI (conçu par FOSS) et sur les spectromètres de la marque FOSS. Ainsi, elle ne peut être applicable sur l'appareil Polytec, qui utilise le logiciel de prédiction OLUP.

Une autre méthode de calibration peut être utilisée. Il s'agit de la méthode EXPENDABLE qui permet de ne rajouter que des spectres qui n'appartiennent pas à la base générale afin d'éviter des doublons. Une fois encore, cela nécessite de posséder de vastes bases de données, mais cette méthode permet d'éviter les redondances dans les modèles et d'augmenter leur variabilité, et à fortiori leur précision.

- Perspectives et continuité du projet

L'objectif est d'aboutir à la création d'un outil de contrôle de la qualité des produits finis et des matières premières efficace, précis et rapide dans le but d'optimiser le temps de réaction de *Dumoulin*. Le travail va se poursuivre sur les aliments pour volailles. Cependant, les volumes de production étant faibles, la création d'un modèle de calibration prendra du temps. Le problème des aliments floconnés devra également être résolu. Il sera sans doute nécessaire d'envoyer une grande quantité d'échantillons en analyse chimique et tester plusieurs équations du CRA-W afin de trouver un modèle de prédiction

correct. Le reste du travail va également concerner les matières premières. Il est important pour *Dumoulin* d'exercer un contrôle sur elles dès leur arrivée, car les résultats des analyses chimiques arrivent parfois après l'approvisionnement et ne laisse pas la possibilité à l'entreprise d'ajuster sa production.

Plusieurs modèles de matières premières pourront être mis en commun. Cela pourra notamment être le cas des tourteaux. Pour l'instant, chaque type de tourteaux a été prédit séparément, mais à terme, afin de limiter le nombre d'équation, il serait envisageable de les regrouper. Il en sera de même pour les céréales. A ce jour, deux équations sont envisagées : une pour le maïs grain et une regroupant le froment, l'orge, l'avoine et le triticale. En s'appuyant sur les bases du CRA-W, il serait possible de ne créer qu'une seule courbe. Enfin, d'autres matières premières pourraient être prédites grâce à des courbes préexistantes. Par exemple, le germex (tourteaux de germes de maïs) pourrait être prédit avec les DDGS de maïs. Seule la teneur du germex en amidon poserait problème, celle-ci étant bien plus élevée que celle des DDGS (24% au lieu de 4%).

L'ajustement de la production en temps réel sera également un objectif grâce à l'analyse on-line. Cela permettra d'avoir un contrôle au cours de la production et d'ajuster les paramètres de production plus rapidement. De plus, *Dumoulin*, en étudiant l'humidité des farines à ce stade de la production et des produits finis, pourra voir quelle quantité de mélasse pourra être ajoutée avant que le produit n'ait une Aw trop élevée et que sa conservation soit comprise par un risque de développement de bactéries ou de moisissures.

Ensuite, un travail devra être réalisé sur les modèles déjà existants. Il sera nécessaire de les optimiser au fil des années. Le prélèvement de nouveaux échantillons sera indispensable afin d'augmenter la taille des bases de données afin d'améliorer les modèles. Des valeurs chimiques de référence et des prédictions XDS peuvent être ajoutées au modèles afin d'améliorer leur précision. C'est par exemple le cas de la matière grasse pour les Hipro 58 afin de posséder davantage de valeurs au dessus de 6% de matière grasse. Concernant les matières premières, il faudra choisir de les regrouper ou non, ou de prédire d'autres matières premières sur des modèles préexistants. C'est notamment le cas des tourteaux. Ils ont été calibrés à ce jour séparément, mais il pourrait être judicieux de réussir à les regrouper afin de limiter le nombre d'équations.

Les modèles devront par ailleurs être vérifiés chaque année. C'est pourquoi un plan d'échantillonnage devra être mis en place afin d'envoyer des échantillons en analyse chimique et d'autres échantillons en analyse spectrométrique et procéder à de nouvelles calibrations. Les modèles deviendront ainsi de plus en plus performants d'année en année.

De plus, il serait judicieux de s'intéresser aux coefficients appliqués à chaque longueur d'onde afin de mieux caractériser et interpréter les informations contenues dans les spectres et faire un lien entre les différents paramètres physicochimiques des aliments.

Enfin, l'objectif pour *Dumoulin* sera à moyen et long terme de devenir autonome dans la gestion de ses équations de calibration, et s'éloigner peu à peu du CRA-W afin de mener le contrôle qualité par elle-même.

Conclusion

L'objectif du projet est d'utiliser la spectroscopie proche infrarouge et de créer diverses équations de calibration pour l'ensemble des aliments et matières premières de *Dumoulin* pour voir si le système qualité actuel peut être optimisé grâce à cette technique afin de mieux contrôler les différents paramètres physicochimiques des produits : humidité, protéine, matière grasse, cellulose, amidon, cendres et sucres.

L'étude a porté sur l'analyse des aliments caractérisés par une grande variabilité au niveau de leur composition chimique et de leur volume de production. Pour l'ensemble des aliments, plusieurs dizaines d'échantillons ont été prélevés et scannés sur les spectromètres Polytec et XDS FOSS. En se basant sur les vastes banques de données du CRA-W en matière d'aliments pour animaux, les prédictions du XDS FOSS sont très proches des valeurs chimiques, et ces prédictions ont été réutilisées afin de prédire les spectres issus du Polytec, ce qui a permis de limiter le nombre d'analyses chimiques. Les premières calibrations réalisées sur les aliments extrudés, les correcteurs protéiques, les aliments pour porcs et certaines matières premières sont de bonne qualité et valident l'orientation choisie.

Bien que les prélèvements d'échantillons aient été au final peu nombreux pour la majorité des aliments, la faible variabilité parfois observée a néanmoins permis d'obtenir des résultats permettant de caractériser les possibilités d'évaluation de la SPIR pour l'analyse des aliments et des matières premières.

Les premiers résultats obtenus sont très encourageants pour l'ensemble des paramètres physicochimiques étudiés, les prédictions étant très proches des valeurs chimiques de référence. Cela s'aperçoit tout d'abord par l'analyse des indicateurs de performance des calibrations et des cross-validation, dont les valeurs sont globalement faibles, même si le manque de variabilité entre les échantillons permet de relativiser ce constat. De plus, la quasi totalité des équations possèdent une très fortes corrélation entre les valeurs prédites et les valeurs de référence, ce qui permet de disposer de résultats utilisables immédiatement.

Des premiers changements sont ainsi d'ores et déjà apparus au sein du système qualité, le plus marquant étant le contrôle de teneur en protéine des correcteurs protéiques grâce à l'analyse infrarouge et non au chimique. Les équations de calibration des matières premières ont également permis de mettre en évidence la variabilité qu'il peut exister d'un approvisionnement à l'autre, notamment dans le cas des tourteaux. L'analyse infrarouge assure ainsi à *Dumoulin* une meilleure réactivité au niveau des ajustements des recettes en fonction de la composition des matières premières, les résultats des analyses chimiques arrivant souvent après que le stock ait commencé à être utilisé.

Certains modèles méritent néanmoins d'être améliorés. En tenant compte des valeurs des coefficients appliqués aux longueurs d'onde et du nombre de termes affectés à chaque équation, plusieurs

équations pourraient être optimisées au niveau de la qualité de la prédiction. Cela concerne notamment les aliments pour porcs, où les valeurs des coefficients atteignent 150 pour la protéine et 300 pour l'humidité, mais l'ensemble des modèles créés peuvent être concernés, car il est possible de réduire significativement le nombre de termes de chaque équation sans augmenter de manière significative le SECV et ainsi augmenter la marge d'erreur de l'analyse infrarouge.

Les modèles créés sont cependant utilisables, car bien qu'il s'agit d'une première expérimentation, les prédictions des compositions chimiques sont proches des valeurs attendues. Les modèles seront constamment améliorés dans le futur, de nouvelles calibrations étant réalisées chaque année avec de nouveaux sets d'échantillons afin de vérifier leur robustesse. Les modèles seront ainsi composés de davantage d'échantillons année après année. L'analyse infrarouge apparaît ainsi comme une technique de contrôle devenant de plus en plus performante au fil du temps. Cependant, les prédictions provenant de l'exploitation de banques de données de grande quantité et qualité, les modèles ont été construits sur des bases solides et permettent dès à présent à *Dumoulin* de se situer par rapport à la législation relative à l'étiquetage.

D'un point de vue personnel, cette expérience m'a permis de me confronter aux exigences de la qualité en production d'aliments et à élaborer un nouvel outil de contrôle. Cela m'a permis d'améliorer mes connaissances et compétences en analyse infrarouge. Le stage présentait la particularité de partir de zéro pour chaque aliment afin d'arriver à construire un modèle permettant de prédire la composition physicochimique d'un vaste panel d'aliment. Le projet a apporté une dimension méthodologique très importante ainsi qu'une approche concrète grâce à l'apport en termes de contrôle de la qualité des produits qu'ont pu apporter les résultats obtenus. Depuis la sélection des populations d'échantillons de référence à envoyer en analyse chimique et à intégrer aux modèles de calibration jusqu'au traitement des données spectrales par la chimiométrie, j'ai pu avoir un regard complet sur le projet. Ce dernier peut être divisé en trois parties. La première partie consistait à établir les plans d'échantillonnage grâce à l'étude des particularités de chaque aliment afin d'être le plus exhaustif possible au niveau de la variabilité. La seconde partie consistait à mettre en application les connaissances en chimiométrie que j'ai pu acquérir grâce à l'expertise du CRA-W afin d'étudier le logiciel de calibration et de prédiction UNSCRAMBLER® pour construire les modèles. Enfin, la dernière partie du travail était de déterminer quelles applications concrètes *Dumoulin* pouvait mettre en place grâce à ces modèles et ce que cela pouvait apporter au système qualité. Cependant, le temps a manqué pour appliquer le contrôle par infrarouge à davantage de produits, et le projet s'est également heurté à la spécificité des aliments floconnés puisqu'aucune équation du CRA-W utilisée n'a permis de les prédire avec précision.

En conclusion, cette démarche appliquée au cours du projet basée sur la construction d'une méthodologie et de l'utilisation d'outils spécifiques pour répondre aux besoins d'une entreprise a démontré le rôle qui échoie à un ingénieur aujourd'hui et le rend si particulier.

Bibliographie

- BAETEN V., FERNANDEZ PIERNA J.A., MICHOTTE RENIER A., DARDENNE P. Imagerie proche infrarouge : analyse de l'alimentation animale. Techniques de l'ingénieur [en ligne], 10 mars 2005, pp. 34-1 – 34-8. Disponible sur : <https://books.google.be/books?id=YFyH0tcnJgC&pg=PA2&lpg=PA2&dq=SPIR+limites+analyse+aliments&source=bl&ots=FL5uHUSK BX&sig=Rh0qTxwQO9L-zsSCRyoBXjFRUts&hl=fr&sa=X&ei=cLXUVMjhOIOU7AaquIDoDQ&ved=0CDMQ6AEwBDgK#v=onepage&q=SPIR%20limites%20analyse%20aliments&f=false> [Consulté le 06 février 2015].
- BASTIANELLI Denis, LECOMTE Philippe. Applications de la SPIR au Cirad EMVT - Calibration de grandeurs mesurées indirectement. L'exemple des analyses fécales [en ligne]. 18 septembre 2003, 1p. Disponible sur : http://www.heliospir.net/medias/upload/resume_alimentationSPIR_PLDB_CIRAD.pdf [Consulté le 06 février 2015].
- BELLON-MAUREL Véronique, FERNANDEZ-AHUMADA Elvira, PALAGOS Bernard, ROGER Jean-Michel, MCBRATNEY Alex. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. TrAC Trends in Analytical Chemistry, [en ligne]. 2010, Vol. 29, n° 9, pp. 1073-1081. Disponible sur : http://www.researchgate.net/profile/Elvira_Fernandez-Ahumada/publication/244361057_Critical_review_of_chemometric_indicators_commonly_used_for_assessing_the_quality_of_the_prediction_of_soil_attributes_by_NIR_spectroscopy/links/02e7e53590bf3b3ea4000000.pdf [Consulté le 12 juin 2015].
- BERTRAND D. La spectroscopie proche infrarouge et ses applications dans les industries de l'alimentation animale. INRA Productions Animales [en ligne], 2002, vol. 15, n°3, pp. 209-219 ; Disponible sur : <https://www6.inra.fr/productions-animales/2002-Volume-15/Numero-3-2002/La-spectroscopie-proche-infrarouge-et-ses-applications-dans-les-industries-de-l-alimentation-animale> [Consulté le 27 janvier 2015].
- BIEMAR Denis. Rapport d'activité semestriel pour le projet FIRST Entreprise : Etude des possibilités d'évaluation de la qualité des aliments floconnés et extrudés à l'aide de l'analyse par spectroscopie infrarouge. Semestre 1. Rapport d'activité, DUMOULIN S.A.-CRAW, 2014, 18p.
- BIEMAR Denis. Rapport d'activité semestriel pour le projet FIRST Entreprise : Etude des possibilités d'évaluation de la qualité des aliments floconnés et extrudés à l'aide de l'analyse

par spectroscopie infrarouge. Semestre 2. Rapport d'activité, DUMOULIN S.A.-CRAW, 2014, 27p.

- BOUROCHE Annick, LE BARS Michèle. La cuisson-extrusion : vocabulaire français-anglais-allemand. Editions Quae. Paris : INRA, 1996, 96p. (Dictionnaires).s
- CAMO. PLS Regression [en ligne]. Disponible sur : <http://www.camo.com/resources/pls-regression.html> [Consulté le 5 juin 2015].⁷
- CEDUS. Les Céréales. [en ligne]. Rapport de recherche. 11 p. Disponible sur : www.brassageamateur.com/forum/download/file.php?id=1875 [Consulté le 19 janvier 2015].
- CNRIT. Implementing NIRS Fecal Sampling Technology. [En ligne]. Disponible sur : http://cnrit.tamu.edu/elearning/1_history.html [Consulté le 26 janvier 2015].
- DARDENNE P. La spectrométrie infrarouge au service de l'élevage [en ligne]. Disponible sur : <http://www.gembloux.ulg.ac.be/zt/Manifestation/PDF/Carrefour%202010/Dardenne.pdf> [Consulté le 20 janvier 2015].
- DEMARQUILLY C. Eléments nouveaux concernant l'utilisation de l'énergie et de l'azote par les ruminants. [en ligne]. Disponible sur : <http://www.afpf-asso.org/download.php?type=1&id=561&statut=0> [Consulté le 26 janvier 2015].
- DR MAURICE LEGOY AND EXTRUSION TECHNOLOGY. Les transformations provoquées par l'extrusion. [en ligne]. Disponible sur : <http://maurice.legoy.free.fr/extrusion/transformations/index.htm> [Consulté le 19 janvier 2015].
- DUBUISSON Mathieu. Distinction spectrophotométrique des matières grasses. MFE. Université Catholique de Louvain, 2010, 98p.
- DUMOULIN. Que produisons-nous ? [en ligne]. Disponible sur : <http://www.dumoulin.eu/fr/faisons-connaissance/que-produisons-nous.php> [Consulté le 19 janvier 2015].
- DUMOULIN. Dumoulin, une entreprise tournée vers l'innovation pour répondre aux besoins des éleveurs et des consommateurs. Présentation. Novembre 2013.

- ECOLE DES MINES DE SAINT-ETIENNE. Méthodes spectrométriques d'analyse et de caractérisation [en ligne]. Disponible sur : <http://spin.mines-stetienne.fr/sites/default/files/spectroir.pdf> [Consulté le 26 janvier 2015].

- FATELA Laura. Utilisation de la spectroscopie proche infrarouge comme outil PAT pour la quantification de substances en solution [en ligne]. Thèse de doctorat en pharmacie. Grenoble : Université Joseph Fourier, 2012, 95p. Disponible sur : <http://dumas.ccsd.cnrs.fr/dumas-00764264/document> [Consulté le 05 février 2015].

- GUIDOLIN® FLAKED CEREALS. Les avantages du floconnage. [en ligne]. Disponible sur : <http://www.guidolingianni.com/floconnage> [Consulté le 20 janvier 2015].

- LECLER Bernard. 2. Vibrational spectroscopy NIR Présentation. Asmara, Erythrée, 01-05 Décembre 2014.

- LECLER Bernard. Chemometrics Challenges. Présentation. Asmara, Erythrée, 01-05 Décembre 2014.

- LECLER Bernard. Chemometrics Introduction. Présentation. Asmara, Erythrée, 01-05 Décembre 2014.

- LECLER Bernard. Chemometrics Multivariate Analysis. Présentation. Asmara, Erythrée, 01-05 Décembre 2014.

- LECLER Bernard. Chemometrics Multivariate Regression. Présentation. Asmara, Erythrée, 01-05 Décembre 2014.

- LOWE J.A., KERSHAW S.J.. Water activity-moisture content relationship as a predictive indicator for control of spoilage in commercial pet diet components. *Animal Feed Science and Technology*, vol. 56, décembre 1995, pp. 187-194. Disponible sur : [http://www.animalfeedscience.com/article/0377-8401\(95\)00833-0/abstract](http://www.animalfeedscience.com/article/0377-8401(95)00833-0/abstract) [Consulté le 17 juin 2015].

- MACIOROWSKI K.G., HERRERA P., JONES F.T., PILLAI S.D., RICKE S.C. Effects on poultry and livestock of feed contamination with bacteria and fungi. *Animal Feed Science and Technology*, vol. 133, février 2007, pp. 109-136. Disponible sur : <http://www.sciencedirect.com/science/article/pii/S0377840106003087> [Consulté le 17 juin 2015].

2015].

- MINASNY Budiman, MCBRATNEY Alex. Why you don't need to use RPD. Pedometron [en ligne], n°33, août 2013, pp. 14-15. Disponible sur : http://www.academia.edu/4303409/Why_you_dont_need_to_use_RPD [Consulté le 12 juin 2015].
- NIR4FARM. About NIR4FARM [en ligne]. Disponible sur : <http://www.nir4farm.com/about-nir> [Consulté le 25 mars 2015].
- NUTRIDEV. La cuisson-extrusion à très faible coût : un enjeu technologique et marketing. [en ligne]. Disponible sur : http://nutridev.org/IMG/pdf/La_cuisson_extrusion_a_tres_faible_cout_un_enjeu_technologique_et_marketing.pdf [Consulté le 19 janvier 2015].
- ROCZEN Lionel. Etude des possibilités d'évaluation de la qualité des aliments extrudés à l'aide de l'analyse par spectroscopie proche infrarouge. MFE. ISIA HUY, 2014, 75p.
- ROGGO Yves. Détermination de la qualité de la betterave sucrière par spectroscopie proche infrarouge et chimiométrie [en ligne]. Thèse de doctorat en instrumentation et analyses avancées. Lille : Université des Sciences et Technologies de Lille, 2003, 189p. Disponible sur : <https://tel.archives-ouvertes.fr/tel-00006113/document> [Consulté le 21 janvier 2015].
- SPECIAL DIETS SERVICE. Types d'aliments : Pellets & Extrudés. [en ligne]. Disponible sur : http://www.sdsdiets.fr/diet_types/pellets_extrudes/ [Consulté le 19 janvier 2015].
- Spectroscopie infrarouge [en ligne]. Disponible sur : <https://cours.espci.fr/site.php?id=26&fileid=689> [Consulté le 26 janvier 2015].
- TIMMONS R.A. Water activity as a tool for predicting and controlling the stability of pet foods. Animal Feed, 2007. Disponible sur : <http://en.engormix.com/MA-feed-machinery/manufacturing/articles/water-activity-tool-predicting-t616/p0.htm> [Consulté le 17 juin 2015].
- UNIVERSITE DES SCIENCES ET TECHNOLOGIE DE LILLE. Polysaccharides alimentaires : Comportement du grain d'amidon en milieu aqueux [en ligne]. Disponible sur : http://biochim-agro.univ-lille1.fr/polysaccharides/co/Contenu_1_c.html [Consulté le 05

février 2015].

- VAN BELLEGHEM Régine. Application de la spectrométrie dans le proche infrarouge pour le tri de textile. MFE. Institut Supérieur Industriel HUY-GEMBLOUX, 2003, 94p.
- WILLIAMS Phil. Near-infrared Technology – Getting the Best Out of Light : A Short Course in the Practical Implementation of Near-infrared Spectroscopy for the User. Edition 5.0, 2007, 150p.
- WILLIAMS Phil et NORRIS Karl. Near-infrared technology in the agricultural and food industries. 1987, 330 pages.

Auteur : GAVENC Laurent	Année : 2014-2015	Confidentiel : Oui Durée : 3 ans
THEME : Qualité / Sécurité / Environnement		
<p>Titre : Analyse en proche infrarouge d'aliments naturels pour animaux (spécifiquement extrudés et floconnés)</p> <p>Title : Near infrared analysis of natural feed (specifically extruded and flaked)</p> <p>Mots clés : Aliments, Matières Premières, SPIR, Calibration, Prédiction, Qualité</p> <p>Key words : Feed, Raw Materials, NIRS, Calibration, Prediction, Quality</p> <p>Résumé : La production d'aliments composés pour animaux, et plus particulièrement les extrudés et les flocons, nécessite un système d'analyse qualité efficace pour ces produits atypiques. Cet outil peut être la spectroscopie proche infrarouge. L'étude a ensuite été déclinée sur les farines et les granulés, en raison de leurs volumes, et sur les matières premières afin de contrôler les approvisionnements, être plus réactif en cas de déviation et ajuster rapidement les formules.</p> <p>Les analyses ont été réalisées sur deux spectromètres, dont un disponible au Centre Wallon de Recherche Agronomique (CRA-W) de Gembloux (Belgique), avec qui une collaboration a été mise en place pour construire l'outil.</p> <p>Dans une première étape, l'objectif a consisté à réunir des échantillons ayant un maximum de variabilité et les scanner sur les spectromètres. Des analyses chimiques de référence ont du être obligatoirement réalisées afin de construire des modèles les plus proches possibles de la réalité. Cependant, le spectromètre du CRA-W, de par sa puissance et la qualité de ses prédictions, a permis de réduire ce nombre. Ensuite, des équations de calibration ont été réalisées sur les paramètres physicochimiques les plus importants. Les résultats obtenus montrent que l'outil permet de se substituer progressivement aux analyses chimiques classiques qui étaient effectuées jusqu'à présent.</p> <p>Summary: The production of compound feed, especially extruded and flaked, requires an effective quality analysis system for these atypical products. This tool can be Near Infrared Spectroscopy. Then, the study was declined on flours and pellets, due to their volumes, and on raw materials to control supply, be more responsive in the event of deviation and quickly adjust the formulas.</p> <p>The analyses were performed on two spectrometers, one available at the Centre Wallon de Recherche Agronomique (CRA-W) located in Gembloux (Belgium), a collaboration has been established to build the tool with.</p> <p>Firstly, the objective was to gather samples having high variability and scan then on the spectrometers. Reference chemical analysis was also required in order to build the closest possible models to reality. However, the spectrometer available at the CRA-W, by its power and quality of its predictions, has reduced that number. Then, calibration equations were realized on the most important physicochemical parameters. The results show that the tool can gradually replace conventional chemical analysis currently performed.</p>		
Nombre total de volumes : 1		
Nombre total de pages du document principal : 88 pages		
Demandeur : DUMOULIN S.A.		

