
RESUME en français

La digestion anaérobie de déchets solides est un moyen de traitement de la pollution. Elle a l'avantage de produire de l'énergie sous forme de méthane. Afin que les installations industrielles soient rentables, l'optimisation de la production de méthane doit être réalisée. Différentes voies d'optimisation sont possibles, notamment la sélection en amont de déchets à fort potentiel méthanogène. Actuellement, la mesure du potentiel méthane est réalisée par le test BMP (Biochemical Methane Potential). Ce test repose sur une fermentation d'une petite quantité de déchet et dure plus de 30 jours. Dans un contexte d'optimisation d'une production industrielle, ce délai d'attente est trop important et impose soit, de stocker le déchet, soit de prendre le risque de l'utiliser immédiatement. Dans le cas, notamment, d'un déchet à faible potentiel méthane, la production pourra diminuer, et la biomasse bactérienne pourra décroître faute de nutriments. Dans ce contexte, une méthode rapide de détermination du potentiel méthane réduirait ce risque et permettrait une sélection en amont des déchets à fort potentiel méthanogène en vue de l'optimisation de la production. Cette sélection améliorerait également la co-digestion (mélange de deux déchets à fort et à faible potentiel méthane) afin d'accroître les capacités de production. La spectroscopie proche infrarouge, semble favorable à cette utilisation, grâce à un certain nombre d'avantages : analyse globale (quantitative et qualitative) de la matière organique, mesure rapide, non destructive, peu de préparation d'échantillon et possibilité d'utiliser des fibres optiques pour déporter la mesure. Ce travail est principalement axé sur l'application de la spectroscopie proche infrarouge pour la prédiction du potentiel méthane de déchets solides tout en montrant les différentes potentialités offertes par cette technique.

TITRE en anglais

Study of the use of the near infrared spectroscopy for predicting the methane potential of solid waste.

RESUME en anglais

Anaerobic digestion of solid waste is a solution to treat pollution. It has the advantage to supply energy by methane production. To ensure industrial plants are gainful, optimization of methane production has to be performed. Different ways of optimization are possible, including the selection of wastes with high methane potential. Currently, to extent the methane potential, the BMP (Biochemical Methane Potential) test is conducted. This test is based on a fermentation process of a small amount of waste. It is time consuming, sometimes, lasting over 30 days. This waiting time is too long in order to optimize the production at industrial level. It requires, either to store the waste or to take the risk of using it immediately. In the case where the waste has a low methane potential, it can cause many problems. Production will decrease and bacterial biomass may be loose due to lack of nutrient. In this context, a rapid method for the determination of methane potential would reduce this risk. The optimization of production would also be increase by rapid selection of waste with high methane potential. Knowledge of the potential methane would also improve the co-digestion, *i.e.* the mixing of wastes with high and low methane potential. In this aim, the method studied in this work is the near-infrared (NIR) spectroscopy. His many qualities seem to be favorable to its use: global analysis (quantitative and qualitative) of the organic matter, fast measurement, nondestructive, few or no sample preparation and optical fiber could be used to measure on-line. We will show the real potential of NIR spectroscopy to predict the methane potential of solid wastes.

DISCIPLINE : Génie des Procédés

MOTS-CLES : Spectroscopie proche infrarouge ; digestion anaérobie ; méthane ; méthanisation ;

INTITULE ET ADRESSE DE L'U.F.R. OU DU LABORATOIRE :

Institut National de la Recherche Agronomique
Laboratoire de Biotechnologie de l'Environnement - (INRA – LBE)
Avenue des Etangs – 11100 Narbonne, France

Ecole des Mines d'Alès
Laboratoire Génie de L'Environnement Industriel - (EMA – LGEI)
6 Avenue de Clavières – 30319 Alès Cedex, France

Cemagref,
Informations et Technologies pour les Agro-Procédés – (Cemagref – UMR ITAP)
BP 5095 – 34033 Montpellier Cedex 1, France

Remerciements

Voilà, trois années se terminent durant lesquelles une aventure humaine a permis d'aboutir à ce travail.

Je tiens à remercier tout particulièrement mes directrices de thèse, Véronique Bellon Maurel et Catherine Gonzalez, pour leurs disponibilités, leurs encouragements (surtout vers la fin) et pour le partage de votre savoir. Je remercie chaleureusement mes encadrants, Eric Latrille, Jean-Michel Roger, Jean-Philippe Steyer et Guillaume Junqua, qui m'ont beaucoup apporté pendant ces trois années. Leurs conseils, leurs soutiens, leurs connaissances ont été de précieuses aides pour mener à bien ce travail.

Je remercie l'ensemble du jury d'avoir accepté d'être présent le jour de la soutenance.

Je remercie également Jean Philippe Delgenes et Jean Philippe Steyer, en tant qu'ancien et actuel Directeur du LBE pour m'avoir accueilli dans leur laboratoire pendant ces trois années.

Je tiens également à remercier Gaël Rolland pour son aide et sa présence au début de la thèse et aux comités de thèse. Je remercie également Pierre Buffière pour sa disponibilité aux comités de thèse et pour sa relecture de mon article bibliographique.

Merci à l'Ecole des Mines d'Alès et au Cemagref, de m'avoir soutenu...financièrement.

Merci à Audrey pour m'avoir aidé dans certaines de mes manips et aux élèves de l'EMA pour leur projet d'école.

Enfin, je tiens à remercier la plate-forme régionale GEPETOS Ecotech-LR et la région Languedoc Roussillon pour leur soutien et leur financement.

Je tiens également à remercier tout le personnel administratif, la Dream Team : Annie, Nadine, Sylvie, Véronique, Alexandra pour leur patience. Merci à Michèle Brun de l'EMA. A Michèle Egéa, merci pour ta disponibilité et tes petits mots envers mon fils, Timoté lors de nos échanges de mails administratifs.

Merci à toutes les personnes travaillant au LBE pour cette ambiance agréable au quotidien.

Merci à Emanuelle et à Eddy, miss et mister Informatique...pour avoir récupéré mes données lors de mes trop nombreux crashes d'ordi...et merci aux sauvegardes...

Je voudrais remercier toutes les personnes avec qui j'ai partagé le bureau, où qui ne faisaient que passer. D'abord, Thierry, merci de ta bonne humeur et de ta joie de vivre. Merci à Romain et Jean Batman pour les bons moments passés sur le Carlit. Merci à Jürgen, César, Juan, Ivan et Marianne. Merci également à Bruno pour nos échanges photographiques et un grand merci à Amel, Gwen, pour les bons moments passés à plaisanter sur les ragots. Merci à Monique, Sarah, Florian, Adeline,

Remerciements

Merci Céline, Gaël, Laurie de m'avoir permis d'encadrer les TP de l'IUT et ainsi réaliser ma première expérience d'enseignement, le plus beau métier du monde.

Merci à Romain, Olivier, Stéphanie, Emilie, BrunoM, pour m'avoir permis d'utiliser vos échantillons.

Merci à l'ensemble des doctorants et post-doc de m'avoir fait confiance pour vous représenter au conseil des scientifiques du LBE durant cette dernière année, expérience enrichissante.

Ils n'étaient plus au LBE depuis longtemps, mais merci à Maxime et Co, Maia, Alexis, Magali, Jérémy pour les bons moments passés ensemble, notamment à l'Art'Bar.

Merci à tous mes amis de l'ENITAB, avec qui je décompressais et passais de bon moments : Marie, Eglantine, Véro, Virginie, Fanny, Vince, Alex, Guigui, Mika, Coco, Julien...et les amis d'AgroTuc.

Merci à tous mes amis narbonnais, les Bon'Art & Co...Math (mon jumeau, le seul à comprendre mon humour), Val (merci de me laisser ta place à la batterie sur "A ma place"), Seb (nous oublie pas pour la prochaine soirée jeux), Romy (merci pour tes gâteaux...pense à changer ton four), Yan (arrête de mettre le même T-shirt que moi), Flo(merci de ta patience pour ces répétitions cacophoniques), Marie (merci d'être à l'initiative de ces dîners presque parfaits mémorables).

Merci à mes beaux parents et ma belle sœur de m'avoir accueilli chaleureusement dans leur famille.

Merci à mon père qui m'a donné le goût de la Science par ses notions de chimie et sa mythique aventure à Crozet. Merci à ma mère de m'avoir poussé comme personne d'autre dans les études. Merci à mes sœurs pour m'avoir supporté...Merci au noyau dur de la famille. J'aimerais également remercier tout particulièrement mon oncle Jean Pierre, pour les nombreuses discussions que l'on a ensemble et qui nourrissent l'esprit.

Merci à Lucia, Matias et mon petit Timoté pour leur amour.

Et enfin, un immense merci à ma moitié, Sophie, pour m'avoir aidé, encouragé, soutenu lors des moments difficiles, surtout en fin de thèse. Merci d'avoir été patiente...Merci pour tout ce que tu es et tout ce que tu m'apportes...

Publications dans des journaux à comité de lecture :

Lesteur, M., Bellon-Maurel, V., Gonzalez, C., Latrille, E., Roger, J.M., Junqua, G. and Steyer, J.P. (2010). "Alternative methods for determining anaerobic biodegradability: A review". *Process Biochemistry*. 45(4): 431-440.

Lesteur, M., Latrille, E., Bellon-Maurel, V., Roger, J.M., Gonzalez, C., Junqua, G. and Steyer, J.P. (2011) "First step towards a fast analytical method for the determination of biochemical methane potential of solid wastes by near infrared spectroscopy". *Bioresource Technology*. 102(3): 2280-2288.

Lesteur, M., Bellon-Maurel, V., Roger, J.M., Latrille, E., Gonzalez, C., Junqua, G., Steyer, J.P. "Parallel study of the PLS b coefficients and the near infrared wavelengths from the PLS regression of the methane potential of municipal solid wastes". (en préparation)

Communications orales :

Lesteur, M., Roger, J.M., Latrille, E., Junqua, G., Steyer, J.P., Gonzalez, C., Bellon-Maurel, V. (2009). "Prediction of biochemical methane potential of municipal solid wastes and understanding of their biodegradability by NIRS". The 14th International Conference on Near Infrared Spectroscopy - NIR2009, Bangkok, Thailand, 2009/11/07-2009/11/16, 053-O, p59.

Lesteur, M., Latrille, E., Bellon-Maurel, V., Roger, J.M., Gonzalez, C., Junqua, G., Steyer, J.P. (2010). "Fast prediction of the methane potential of solid waste by Near Infrared Spectroscopy". The 12th World Congress on Anaerobic Digestion – AD12, Guadalajara, Mexico, 2010/10/31 – 2010/11/04.

Posters :

Lesteur, M., Latrille, E., Bellon-Maurel, V., Roger, J.M., Gonzalez, C., Junqua, G., Steyer, J.P. Prédiction rapide du potentiel méthane de déchets solides par spectroscopie proche infrarouge. Journée de l'Ecole Doctorale - Montpellier, Juin 2010 Prix du meilleur poster.

Lesteur, M., Guo, X.M., Latrille, E., Trably, E., Roger, J.M., Bellon-Maurel, V., Gonzalez, C., Junqua, G., Steyer, J.P. (2010). "Fast prediction of the hydrogen potential of solid waste by Near Infrared Spectroscopy". The 12th World Congress on Anaerobic Digestion – AD12, Guadalajara, Mexico, 2010/10/31 – 2010/11/04.

Prix :

Prix des Techniques Innovantes lors du Salon Pollutec (Paris, 1 – 4 December 2009).

Lesteur, M., Latrille, E., Bellon-Maurel, V., Roger, J.M., Gonzalez, C., Junqua, G., Steyer, J.P. "Mesure rapide du potentiel méthane (Biochemical Methane Potential) d'ordures ménagères et interprétation de leur biodégradabilité par spectroscopie proche infrarouge pour l'optimisation des procédés de digestion anaérobie". Salon Pollutec (Paris, 1 – 4 December 2009).

Liste des Abréviations

ACP :	Analyse en Composante Principale
ADF :	Acid Detergent Fiber
ADL :	Acide Detergent Lignin
AG :	Algorithme Génétique
AGLC :	Acide Gras Longue Chaîne
AGV :	Acide Gras Volatil
ATR :	Attenuated Total Reflection
BD :	Biodégradabilité
BMP :	Biochemical Methane Potential
BRS :	Bioréacteur stabilisateur
COT :	Carbone Organique Total
DBO ₅ :	Demande Biologique en Oxygène en 5 jours
DCO :	Demande Chimique en Oxygène
DO ₂₀ :	Demande en Oxygène d'un test aérobie de 20h
DM :	Déchet Ménager
ETr :	Ecart Type de répétabilité
ETR :	Ecart Type de Reproductibilité
FFOM :	Fraction Fermentescible des Ordures Ménagères
GES :	Gaz à Effet de Serre
GS ₂₁ :	Indice de production de biogaz sur 21 jours
GWh :	GigaWatt-heure
Ktep :	Kilo Tonne Equivalent Pétrole
LML :	Litter Mass Losses
MIR :	Moyen Infrarouge
MV :	Matière Volatile
NDF :	Neutral Detergent Fiber
nm :	nanomètre
NTK :	Nitrogen Total Kjeldahl
PLS :	Partial Least Square
PMT :	Potentiel Méthane Théorique
RI ₄ :	Indice de Respiration sur 4 jours
RMSEP :	Root Mean Square Error of Prediction
RPD :	Ratio of Performance to Deviation (Ratio Ecart Type / SEP)
SECV :	Standard Error of Cross-Validation
SEP :	Standard Error of Prediction
SNV :	Standard Normal Variate
SPIR :	Spectrophotométrie Proche Infrarouge
STEP :	Station d'Épuration
STP :	Condition Standard de Température et de Pression
UTN :	Unité de Turbidité Néphélométrique
UV :	Ultra-Violet

SOMMAIRE

Résumé

Remerciement

Liste des Abréviations

Sommaire

I.	Contexte et problématique	1
I.1.	Contexte de la thèse	2
I.1.1.	La valorisation des déchets	2
I.1.2.	La méthanisation	3
I.1.3.	Optimisation de la production de méthane	7
I.2.	Problématique de la thèse	10
I.3.	Plan de la thèse	11
II.	Etude Bibliographique	13
II.1.	Introduction.....	14
II.2.	La biodégradabilité anaérobie	15
II.2.1.	La mesure de la biodégradabilité anaérobie	15
II.2.2.	Le problème de la mesure de la DCO	15
II.2.3.	Quelques notions proches de la biodégradabilité anaérobie.....	17
II.2.4.	Le test BMP : un test respirométrique anaérobie	17
II.2.5.	Sensibilité du test BMP	19
II.2.6.	Conclusion	20
II.3.	Le potentiel méthane corrélé avec les tests respirométriques aérobies rapides	22
II.4.	Prédiction de la valeur BMP basée sur l'étude de la matière organique des déchets	23
II.4.1.	Les tests enzymatiques	24
II.4.2.	Analyse de la composition élémentaire	25
II.4.3.	Analyse de la composition chimique (glucides, protéines, fibres, etc.).....	26
II.4.4.	Conclusion	33
II.5.	Utilisation de méthodes non-destructives : les méthodes spectrophotométriques	33
II.5.1.	La spectrophotométrie UV-Visible	33
II.5.2.	La spectrophotométrie de fluorescence	37
II.5.3.	La spectrophotométrie moyen-infrarouge : MIR.....	37
II.5.4.	La spectrophotométrie proche infrarouge : SPIR.....	39

II.5.5. Conclusion	43
II.6. Conclusion : choix d'une méthode rapide.....	44
III. Faisabilité de la spectroscopie proche infrarouge pour la prédiction du potentiel méthane de déchets solides	47
III.1. Introduction.....	48
III.2. Matériel et méthodes.....	49
III.2.1. Les échantillons utilisés.....	49
III.2.2. Les tests BMP.....	51
III.2.3. La spectroscopie proche infrarouge.....	51
III.2.4. Traitement des données et paramètres statistiques	51
III.3. Résultats et discussion	56
III.3.1. Choix du meilleur prétraitement chimiométrique.....	56
III.3.2. Discussion autour du modèle sélectionné.....	61
III.3.3. Validation avec les échantillons indépendants	65
III.3.4. Comparaison des résultats avec la méthode basée sur la caractérisation chimique	66
III.3.5. Discussion sur la faisabilité de remplacer le test BMP par les mesures PIR	67
III.4. Conclusion	70
IV. Etude conjointe des b coefficients PLS et des longueurs d'onde d'intérêt pour les modèles de prédiction du BMP	73
IV.1. Introduction.....	74
IV.2. Théorie.....	76
IV.2.1. Les b coefficients de la PLS	76
IV.2.2. Les algorithmes génétiques	78
IV.2.3. L'algorithme CovSel	79
IV.3. Matériel et méthodes.....	80
IV.4. Résultats et discussion	81
IV.4.1. Observation des b coefficients PLS.....	81
IV.4.2. Conclusion sur l'observation des b coefficients	92
IV.4.3. Sélection de variables.....	92
IV.5. Conclusion	97
V. Etude de la robustesse du modèle vis à vis de la nature des échantillons	101

V.1. Introduction.....	102
V.2. Matériel et méthodes.....	102
V.2.1. Les échantillons.....	102
V.2.2. Traitement des données et paramètres statistiques	104
V.3. Résultats et discussion	105
V.3.1. Prédiction des nouveaux échantillons à l'aide du modèle précédent.....	105
V.3.2. Analyse multi-variée des échantillons.....	107
V.3.3. Modèle global de régression PLS.....	113
V.4. Conclusion	116
VI. Conclusion et perspectives.....	119
Références Bibliographiques.....	125
Liste des Figures	137
Liste des Tableaux.....	139

I.
Contexte et problématique

I.1. Contexte de la thèse

I.1.1. *La valorisation des déchets*

La directive de la Communauté Economique Européenne 75/442/CEE, du 15 juillet 1975, relative aux déchets, spécifie la nécessité « de promouvoir la valorisation des déchets par recyclage, réemploi, récupération ou toute autre action visant à obtenir des matières premières secondaires, ou, à leur utilisation comme source d'énergie ». La valorisation des déchets, consiste donc à réintroduire les déchets, dans un cycle de production, en les considérant comme des ressources. Différents types de valorisation existent actuellement, comme le recyclage, l'incinération, le compostage ou la digestion anaérobie.

- Le recyclage

Le recyclage est le procédé de traitement des déchets (déchets industriels ou ordures ménagères) qui permet de réintroduire, dans le cycle de production d'un produit, des matériaux qui composaient un produit similaire arrivé en fin de vie. Son utilisation est devenue courante pour divers matériaux : le papier et le carton usagés sont réutilisés depuis le 18^{ème} siècle pour faire du « papier recyclé », le verre est refondu, les métaux, les textiles, le caoutchouc et la majorité des plastiques sont recyclés avec une multitude de réutilisations possibles comme la création de vêtements par exemple. Depuis 2008, les plastiques destinés aux contacts alimentaires peuvent également être réutilisés dans ce même domaine.

Les eaux usées sont également traitées dans les stations d'épuration pour être recyclées directement dans des activités d'irrigation ou ne nécessitant pas d'eau potable. Certaines villes ré-injectent l'eau recyclée directement dans les rivières ou les nappes phréatiques d'où est captée l'eau potable, notamment depuis 1965, en Afrique du Sud à Johannesburg et aux Etats-Unis dans la ville de Los Angeles, ou encore depuis 1968 à Windhoek la capitale namibienne (Anderson, 2003). Mais globalement, un frein psychologique persiste à l'utilisation directe des eaux traitées en eau potable. La tendance actuelle est donc un usage indirect, avec un passage dans le milieu naturel des eaux traitées et une dilution de celle-ci avec les eaux naturelles.

- L'incinération

Certains déchets sont, dans l'état actuel des technologies, non recyclables (huiles, peintures, déchets liquides à composantes minérales), ils sont alors incinérés. L'incinération a l'avantage de réduire, de façon importante, les quantités de déchets et de produire de l'énergie sous forme de chaleur et/ou d'électricité. L'incinération a cependant d'importants inconvénients comme la nocivité des fumées générées (Wey et al., 2001 ; McKay, 2002). Ces fumées issues du procédé apparaissent souvent toxiques (CO, oxydes d'azote (NO, NO₂), gaz acides (HCl, SO₂), dioxines et furanes) et contiennent d'importantes teneurs de dioxyde de carbone, qui viennent augmenter les quantités de gaz à effet de serre. Certains mâchefers (résidu de l'incinération), ne présentant pas de métaux lourds, sont alors utilisés comme matériau de recouvrement des chaussées. Les mâchefers ne répondant pas aux normes en termes de métaux lourds sont alors stockés.

- Le compostage

Le compostage est un procédé biologique de transformation et de valorisation des matières organiques issues des déchets, en un produit stabilisé, riche en composés humiques pouvant servir d'apport pour les cultures. Il se déroule en présence de micro-organismes sous condition aérobie. Les déchets concernés sont ceux possédant une part de matière organique biodégradable : déchets verts, fraction fermentescible des ordures ménagères. L'utilisation des méthodes de compostages sont en plein essor, tant dans le secteur économique que domestique.

Une autre forme de valorisation est la digestion anaérobie, également appelée méthanisation. Etant au cœur de cette étude, elle fait l'objet d'une description plus détaillée

1.1.2. La méthanisation

1.1.2.1. Un peu d'histoire (Moletta, 2008)

La méthanisation a été mise en évidence, pour la première fois, dans les marais du Lac Majeur en Italie en 1776 par Alessandro Volta. Il démontra notamment que le gaz qui s'en échappait était inflammable. En 1787, Antoine Lavoisier nomme le gaz des marais : « gaz hidrogenium carbonatrum ». Mais ce n'est que plus tard, en 1856, que le nom de « méthane » est proposé. Il est enfin adopté définitivement en 1892, lors du congrès de nomenclature chimique. En

1884, un élève de Louis Pasteur, Ulysse Gayon, évoque déjà l'utilisation de ce gaz de fermentation comme source de chauffage et d'éclairage. Au début du XX^{ème} siècle, la première installation produisant du méthane exploitable, voit le jour à Exeter (Grande Bretagne). Le méthane y est réutilisé pour l'éclairage urbain. Le premier essor de la méthanisation débuta après la seconde guerre mondiale. Le second eut lieu après les chocs pétroliers de 1973 et 1979. En prévision de probables pénuries de pétrole, les recherches sur les énergies renouvelables furent entreprises. Cependant, la majorité des projets n'aboutirent pas, en raison de l'essor de l'énergie nucléaire, notamment en France. Aujourd'hui, dans un contexte de développement durable et de réduction des gaz à effet de serre (GES), les énergies renouvelables telles que la méthanisation reprennent une place importante dans les recherches.

I.1.2.2. Quelques chiffres au niveau national et européen

Dans un contexte de développement durable, la méthanisation répond à de nombreuses problématiques :

- dépollution : réduction de la quantité de déchet et des nuisances olfactives,
- production d'énergie : le méthane engendré servant à la production d'électricité et/ou de chaleur (Figure 1),
- valorisation des résidus de digestion : le digestat (sous produit de la digestion) stabilisé est utilisé comme amendement en agriculture.

Le tableau 1 présente les productions primaires de biogaz des pays de l'Union Européenne en 2007 (EurObserv'ER, 2009), suivant les trois modes de production possible :

- le gaz issu de la mise en décharge des déchets,
- la digestion anaérobie des boues issues des stations d'épurations,
- et le gaz produit à partir des déchets d'installations d'agro-industrie et à la ferme.

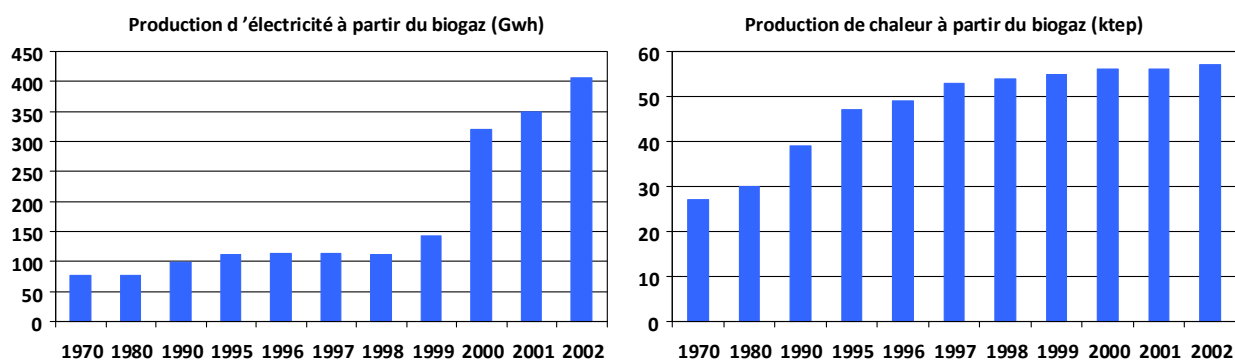


Figure 1: Production d'électricité (GWh) et de chaleur (ktep) à partir du biogaz (France) (Observatoire de l'énergie, 2005).

Les quantités produites sont très inégales suivant les pays européens. L'Allemagne comptabilise 50% de la production totale européenne, avec une origine de production essentiellement basée sur les installations à la ferme (80%). La Grande Bretagne occupe la seconde place avec 22% et une production essentiellement basée sur les décharges (88%). La France arrive en quatrième position avec seulement 421 ktep (6% de la production de l'Union Européenne). Les pays qui ont une part importante de leur production issue de digesteurs à la ferme ou sur site industriel sont l'Allemagne (80%), l'Autriche (95%), le Danemark (70%) et la Hollande (45%).

La production de méthane en digesteur est maintenant amenée à se développer, surtout grâce à une rentabilité due à un prix intéressant lors du rachat de l'électricité produite et - suivant les pays - à la possibilité d'injecter le gaz produit directement dans le réseau de gaz de ville. A noter que l'utilisation de réacteurs dit « familiaux » devient très courante et depuis de nombreuses années, chez le particulier, dans certains pays asiatiques tels que la Chine, l'Inde, et l'Asie du Sud-Est (He, 2010). La production dans des méthaniseurs en co-digestion, avec des mélanges de différents types de déchets, comme par exemple, un déchet d'abattoir et un déchet agricole de type lisier, se développera également dans les années à venir.

Pays	Décharges	STEP(1)	Autres(2)	Total
Allemagne	346	387	2926	3659
Grande Bretagne	1393	191	0	1584
France	340	52	29	421
Italie	315	2	71	388
Autriche	5	6	206	217
Hollande	48	48	80	176
Espagne	116	50	27	193
Pologne	21	43	1	65
Suède	25	53	19	97
Danemark	7	21	66	94
Rep. Tchèque	31	31	14	76
Belgique	49	4	26	79
Finlande	28	12	2	42
Irlande	24	8	2	34
Grèce	30	5	0	35
Portugal	0	0	16	16
Slovénie	8	1	4	13
Hongrie	2	1	3	6
Luxembourg	0	0	9	9
Slovaquie	0	7	1	8
Lettonie	5	2	0	7
Lituanie	0	2	1	3
Estonie	3	1	0	4
Roumanie	0	0	1	1
Total UE	2796	927	3504	7227

Tableau 1 : Production primaire de biogaz dans différents pays de l'Union Européenne en 2007 (en ktep) (EurObserv'ER, 2009).

1 : Issue du traitement des boues de stations d'épuration urbaines et industrielles.

2 : Issue d'unités décentralisées de biogaz agricole, unités de méthanisation des déchets municipaux solides, unités centralisées de co-digestion.

I.1.2.3. La méthanisation : un processus biologique

La digestion anaérobie est donc un processus naturel se déroulant dans les marais, les sédiments, mais également dans le système digestif (rumen) des animaux tels que les mammifères et certains insectes (termites). La méthanisation consiste en la dégradation de la matière organique complexe en méthane et dioxyde de carbone par des micro-organismes en condition d'anaérobie. Les différentes étapes de la digestion anaérobie sont présentées Figure 2.

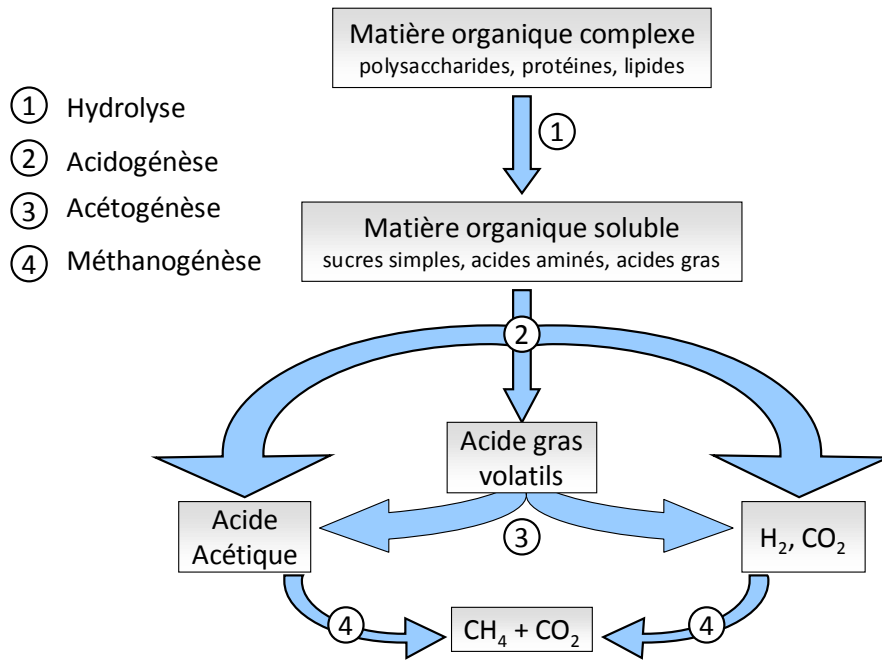


Figure 2 : Les étapes de la digestion anaérobie (Mata-Alvarez, 2003).

Les déchets utilisables en digestion anaérobie sont ceux contenant de la matière organique biodégradable :

- fraction fermentescible d'ordures ménagères (déchets de cuisine, papier, carton),
- déchets verts (tontes des pelouses, déchets issus des collectivités),
- déchets agricoles (lisiers, fumiers, résidus de cultures, cultures énergétiques dédiées...),
- déchets issus de l'industrie agro-alimentaire,
- boues de station d'épuration,
- macro et micro-algues...

1.1.3. Optimisation de la production de méthane

L'optimisation de la production de méthane dans les digesteurs repose sur de nombreux critères : physico-chimiques, avec la régulation du pH et de la température aux valeurs optimales respectivement 6,5-7,2 et 35 ou 55°C, mécaniques avec le broyage qui augmente l'accessibilité de la matière organique aux micro-organismes (Mata-Alvarez et al., 2000 ; Angelidaki et Sanders, 2004 ; Ward et al., 2008).

En outre, la qualité et la quantité de matière organique biodégradable contenue dans le déchet joue un rôle prédominant pour l'optimisation de la production du biogaz. En effet, l'optimisation de la production se réalise aussi en amont du digesteur, par le choix des déchets

à méthaniser. L'utilisation de déchets fortement biodégradables augmente la production de méthane. A l'inverse, la dégradation de déchets faiblement biodégradables engendre obligatoirement une baisse de rendement. Elle peut ainsi amener à une perte de la biomasse microbienne par manque de substrats métabolisables. Ainsi, en digestion anaérobie, la qualité d'un déchet correspond à sa capacité à être biodégradable. Dans le cas de la co-digestion, la connaissance de la qualité des déchets permet d'effectuer des mélanges adéquats, augmentant la production (Alvarez et al., 2010).

Actuellement, la biodégradabilité anaérobie d'un déchet est évaluée par la mesure de son potentiel méthane. Le test le plus couramment effectué pour mesurer le potentiel méthane d'un déchet est le test « Biochemical Methane Potential ». Il est plus communément appelé « test BMP » (Angelidaki et al., 2009 ; Hansen et al., 2004). Ce test de fermentation consiste à placer une quantité connue de déchet en présence d'un inoculum microbien adéquat en condition anaérobie. A intervalles fréquents, la composition et quantité de biogaz produit sont mesurées afin d'obtenir la production de méthane en fonction du temps (Figure 3). La Figure 3 représente le méthane produit (cumulé sur la durée de fermentation) par 4 déchets différents. La hauteur du plateau correspond à la valeur BMP. Plus le déchet est biodégradable, plus son potentiel méthane (la valeur BMP) sera important. Les mêmes conditions opératoires ont été utilisées pour ces 4 déchets différents.

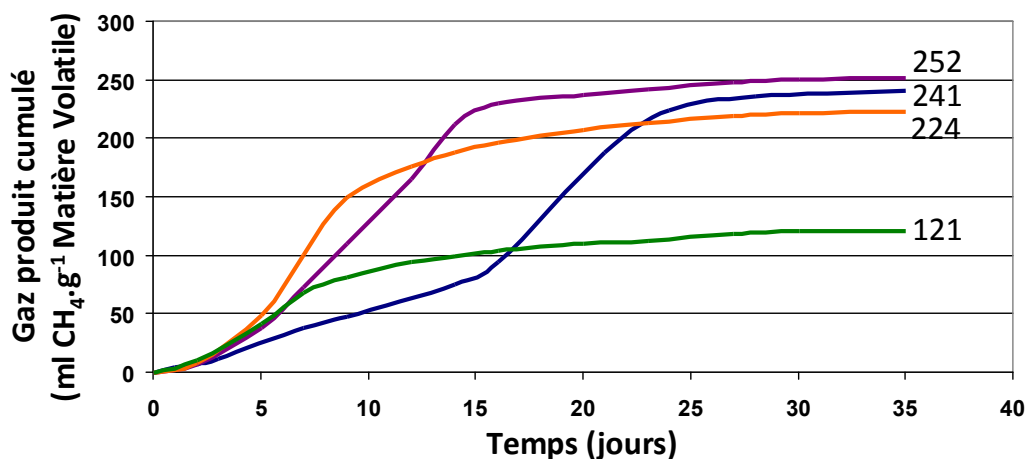


Figure 3 : Représentation de résultats obtenus lors du test BMP. Les chiffres indiquent la valeur BMP.

Deux informations essentielles sont extraites de ce graphique :

- la relation qui lie la valeur BMP à la matière organique d'un déchet. La valeur BMP est uniquement liée à la composition de la matière organique contenue dans le déchet à la fois d'un point qualitatif et quantitatif. La matière organique peut

être composée de glucides, protéines, lipides et fibres en différentes quantités suivant les déchets.

- le temps nécessaire pour obtenir la valeur BMP. Comme le montre la Figure 3, le temps d'analyse nécessaire à l'obtention de la valeur BMP est plus ou moins long, généralement entre 15 et 30 jours, plus suivant le type de déchet.

Le test BMP est complexe dans sa mise en œuvre nécessitant un suivi régulier de la production de biogaz. Ces contraintes ne sont pas compatibles ni avec une optimisation en temps réel du procédé de méthanisation et ni avec la conduite de digesteurs industriels.

En effet, dans le cas d'une exploitation de méthanisation deux cas se présentent lors de l'arrivée d'un nouveau déchet :

- soit l'exploitant utilise ce déchet sans connaître son potentiel méthane. Il prend donc le risque de voir le rendement de l'exploitation diminuer, notamment dans le cas où le déchet est faiblement biodégradable. Par ailleurs, il y a un risque de « perdre » la biomasse déjà présente dans le réacteur et par voie de conséquence de réduire le rendement sur la période nécessaire à la reconstitution de cette biomasse.
- Soit l'exploitant attend de connaître le potentiel méthane avant d'utiliser le déchet. Il devra donc stocker le déchet pendant la durée de l'évaluation du test BMP.

Nous pouvons conclure qu'une méthode rapide de prédiction du potentiel méthane est donc nécessaire.

L'utilisation d'une méthode rapide pour prédire le potentiel méthane aurait donc plusieurs avantages :

- une sélection rapide des meilleurs déchets à méthaniser, permettant une optimisation de la production de méthane,
- une réduction du risque de méthaniser un déchet faiblement biodégradable et de perdre la biomasse bactérienne,
- une production de méthane améliorée en co-digestion par un meilleur choix des mélanges,
- une meilleure gestion des intrants afin d'éviter le risque de déséquilibre au sein du réacteur, voir son blocage.

I.2. Problématique de la thèse

Cette thèse a pour objet d'élaborer une méthode de caractérisation rapide de la matière organique des déchets solides afin de prédire plus rapidement et plus facilement leur potentiel méthane. La méthode a pour finalité d'être utilisée directement sur le terrain, au plus près des digesteurs. Les différents critères du cahier des charges sont :

- **la rapidité de l'analyse et de l'obtention du résultat,**
- la simplicité d'utilisation,
- un entretien simple,
- une préparation d'échantillon minimale.

Les méthodes spectroscopiques pourraient satisfaire ce cahier des charges.

Cet objectif opérationnel génère plusieurs questions de recherche :

- Les méthodes spectroscopiques rapides comme la spectrométrie UV/Visible ou proche- infrarouge, sont-elles appropriées pour la détermination du BMP, grandeur hautement intégratrice ?
- Une prédiction directe du BMP à partir de spectres proche-infrarouge est-elle pertinente et pourquoi ?
- L'écueil de l'étalonnage sur des échantillons, très hétérogènes et couvrant une gamme très large, peut-il être évité ?
- Jusqu'à quel stade, la préparation des échantillons doit être poussée, afin d'obtenir la meilleure prédiction du potentiel méthane ?

La démarche a consisté dans un premier temps, au travers d'une étude bibliographique, à comprendre le lien entre la matière organique et la valeur BMP. L'intérêt potentiel de méthodes de caractérisations rapides de la matière organique sera également discuté quant à leurs aptitudes pour la prédiction du potentiel méthane. On verra que la spectroscopie proche infrarouge apparaît être la méthode la plus appropriée à une prédiction rapide du potentiel méthane de déchets solides. Un premier modèle de prédiction du BMP sera élaboré et testé dans des conditions simples, c'est à dire en restreignant la base de déchets à un seul type, les ordures ménagères. Les coefficients du modèle seront étudiés afin de le valider d'un point de vue chimique, c'est à dire en relation avec la composition chimique. La robustesse de ce

modèle vis à vis de la nature et des origines des échantillons sera testée et étudiée à l'aide d'échantillons hétérogènes. Enfin, un modèle incluant de nombreux échantillons hétérogènes sera construit afin d'apporter un élément de réponse à ce problème de robustesse.

Sur la base de ces objectifs et problématiques scientifiques présentés ci-dessus, le plan de la thèse choisi est présenté ci-dessous.

I.3. Plan de la thèse

A la suite de ce contexte, une étude bibliographique sera réalisée afin d'étudier des méthodes alternatives pour la prédiction du potentiel méthane.

A la suite de ce travail bibliographique, une troisième partie aura comme objectif de montrer l'intérêt et la faisabilité de la spectroscopie proche infrarouge pour la prédiction du potentiel méthane de déchets solides.

Lors du quatrième chapitre, une étude conjointe des b coefficients PLS et des longueurs d'onde d'intérêts sera conduite afin de valider le modèle précédemment établi d'un point de vue chimique. Elle permettra également de vérifier la pertinence du modèle au regard des connaissances actuelles sur la biodégradabilité des déchets ménagers et de leur composition.

Des algorithmes de sélection de variables seront également utilisés pour compléter et valider l'étude.

Un cinquième chapitre sera dédié à l'étude de la robustesse de la méthode vis à vis de la nature des déchets, pour répondre à l'objectif d'un modèle opérationnel en conditions industrielles.

Enfin, les principales conclusions de ce travail seront reprises et des perspectives seront envisagées, quant à son utilisation en conditions industrielles et d'autres méthodes seront proposées

II.
Etude Bibliographique

Ce travail bibliographique est la traduction d'une publication dans *Process Biochemistry*, traitant « des méthodes alternatives de mesure de la biodégradabilité » (Lesteur et al., 2010).

II.1. Introduction

L'optimisation de la production de méthane dans les digesteurs industriels nécessite de caractériser les déchets. Cette caractérisation peut être d'ordre physico-chimique (taux de matière sèche, taux de matière organique, granulométrie, quantité de glucides, lipides, protéines, fibres), mais la méthode employée le plus souvent est la mesure de la biodégradabilité anaérobie du déchet par le test BMP. Cependant, la biodégradabilité anaérobie étant obtenue après un test d'environ 30 jours, des méthodes plus rapides de cette estimation doivent être mise au point. Ceci afin d'avoir un rôle cohérent dans l'optimisation des procédés de méthanisation.

Dans ce contexte, ce travail bibliographique s'attache d'abord au rappel de la définition et du processus de mesure de la biodégradabilité anaérobie avec le test BMP. Certaines méthodes alternatives sont d'ores et déjà proposées :

- des tests rapides de respirométrie en conditions aérobies utilisés afin d'accélérer le processus biologique,
- des modèles de prédiction de la valeur BMP basés sur la caractérisation chimique des déchets.

Ces méthodes sont présentées et évaluées principalement sur leurs capacités à prédire rapidement la valeur BMP.

Enfin des méthodes essentiellement spectrophotométriques, qui analysent la matière organique rapidement, ont été étudiées afin d'estimer leur possibilité de prédiction du potentiel méthane des déchets solides.

II.2. La biodégradabilité anaérobie

II.2.1. La mesure de la biodégradabilité anaérobie

La biodégradabilité anaérobie (BD) est actuellement définie comme le rapport de la quantité de méthane produit sur la quantité maximale théorique de méthane du déchet (Buffière et al., 2006) (1) :

$$BD = \frac{\text{méthane produit}}{350 \times DCO_{\text{déchet}}} \quad (1)$$

Avec *BD*, la biodégradabilité anaérobie, exprimée en pourcentage de la matière organique initiale du déchet, *le méthane produit*, en ml CH_{4,STP}.g⁻¹ Matière Volatile, exprimé dans les conditions standard de température et de pression (STP) respectivement 273,5 K (0°C) et 100 kPa (1 atm) et la *DCO*, la Demande Chimique en Oxygène, qui mesure la quantité d'oxygène nécessaire à l'oxydation totale de la matière organique contenue dans le déchet, exprimée en g O₂.g⁻¹ Matière Volatile. La valeur de 350 ml CH₄.g⁻¹ DCO exprime la quantité théorique de méthane produit par unité de DCO.

II.2.2. Le problème de la mesure de la DCO

Cependant, la mesure de la DCO d'un déchet solide étant contraignante à mettre en œuvre et faisant appel à des produits dangereux (dichromate de potassium et sulfate de mercure), son utilisation sur des déchets solides non solubles est moins courante. Cette méthode a d'ailleurs été interdite en Suède (OCDE, 2008). De plus, une étude récente réalisée par 26 laboratoires sur la mesure de la DCO de déchet solide précise le manque d'homogénéité des résultats entre laboratoires. Elle met en évidence des écarts importants entre la valeur théorique de référence et les résultats de DCO des différents laboratoires (Figure 4) (Raposo et al., 2009).

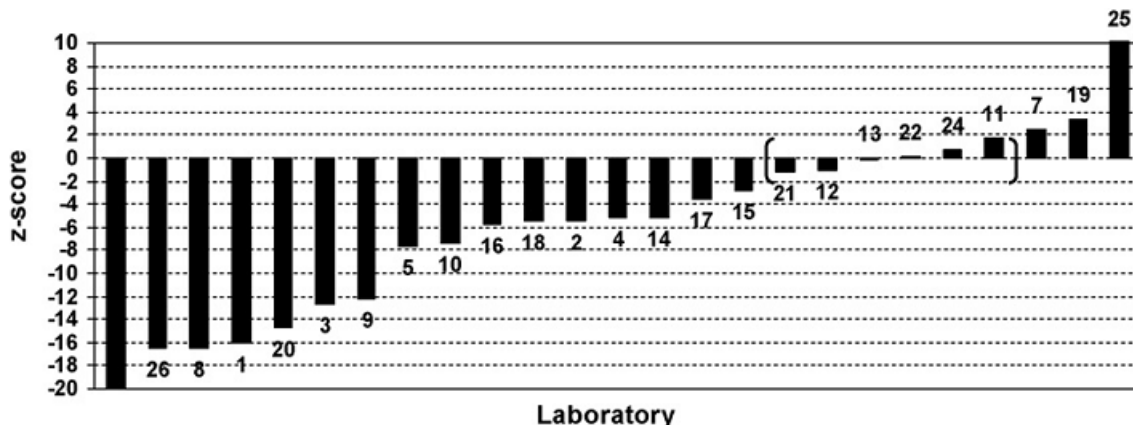


Figure 4 : Exemple de résultats de l’étude réalisée entre 26 laboratoires, afin de comparer des résultats de dosage de la DCO.

z-score est l’écart entre valeur mesurée et valeur vraie de DCO rapporté à l’écart acceptable. Les déchets testés sont des graines de Quinoa (Raposo et al., 2009).

Le paramètre regardé, Z-score, est calculé par la formule suivante (2) :

$$z - score = \frac{(X_{ev} - X_{av})}{ESD} \tag{2}$$

avec X_{ev} , la valeur expérimentale de DCO du laboratoire, X_{av} , la valeur théorique de référence calculée à partir de la composition élémentaire et ESD, l’écart acceptable admis pour cette étude. Comme expliqué par Raposo et al., (2009), dans le cas d’un déchet solide et afin d’éviter un nombre trop important de laboratoires avec un z-score supérieure à 2, l’ESD a été fixé à 2.5%. Un z-score inférieur à 2 (en valeur absolue) est considéré comme satisfaisant. Lorsqu’il est compris entre 2 et 3, il est alors nécessaire de revoir, par exemple, le protocole de mesure et lorsque le z-score est supérieur à 3, les résultats ne sont pas inacceptables. Comme l’indique la figure 4, la plupart des laboratoires génèrent des valeurs supérieures à 3, montrant le manque de justesse des mesures.

La DCO de déchet solide n’est pas une méthode permettant d’estimer avec justesse sa biodégradabilité, il est donc nécessaire d’améliorer la reproductibilité interlaboratoire de cette méthode, voire d’envisager d’autres méthodes.

II.2.3. Quelques notions proches de la biodégradabilité anaérobie.

Certaines notions, utilisées dans le domaine de l'environnement et du traitement des déchets, sont proches de la biodégradabilité anaérobie. La notion de digestibilité, par exemple, est à rapprocher de la biodégradabilité anaérobie. La capacité d'un aliment à être digéré est un terme couramment utilisé en zootechnie pour parler de la digestion des aliments pour bétail. Or, la digestion de fourrage par les ruminants se déroule en condition anaérobie dans le rumen, où du méthane est également produit. Les notions de production de méthane, biodégradabilité anaérobie et de digestibilité sont donc équivalentes : elles traitent toutes de la dégradation de la matière organique (de déchet ou de fourrage) par des micro-organismes en condition anaérobie. Les réacteurs où se déroule la méthanisation sont également appelés digesteurs.

II.2.4. Le test BMP : un test respirométrique anaérobie

Le test Biochemical Methane Potential (BMP) (Angelidaki et al., 2009) est le test actuellement le plus utilisé pour déterminer le potentiel méthane des déchets. La valeur de potentiel méthane, ou valeur BMP, est la valeur maximale de méthane produit en condition anaérobie, exprimée en ml $\text{CH}_4 \cdot \text{g}^{-1}$ Matière Volatile (MV). La cinétique de production est également étudiée afin de déterminer la vitesse de biodégradabilité des déchets. Ces valeurs sont utilisées pour évaluer la biodégradabilité anaérobie des déchets, et permettent de dimensionner les réacteurs industriels (Buffière et al. 2006), mais aussi de contrôler et d'optimiser la production. Ce test a l'avantage de pouvoir déceler les inhibitions du métabolisme microbien, dues à la présence dans le déchet de composés inhibiteurs, tels que des antibiotiques, des métaux lourds, etc. (Hansen et al., 2004 ; Chen et al., 2008). Il met également en évidence une surcharge ou une adaptation à un type de substrat (Hansen et al., 2004).

Le test BMP consiste à placer une certaine quantité de déchets en présence de micro-organismes en condition anaérobie dans un milieu tamponné et thermostaté (pH entre 6,5 et 7,2, température à 35°C ou 55°C) (Figure 5).

La pression dans l'espace de tête est mesurée à l'aide d'un système de fiole renversée où un déplacement de liquide est mesuré. La composition du biogaz (méthane et dioxyde de carbone)

est régulièrement analysée par chromatographie en phase gazeuse. Reposant sur une fermentation et les déchets nécessitant une étape d'hydrolyse, la durée du test est souvent longue, environ 30 jours. La Figure 6 donne quelques résultats de test BMP obtenus pour différents échantillons.

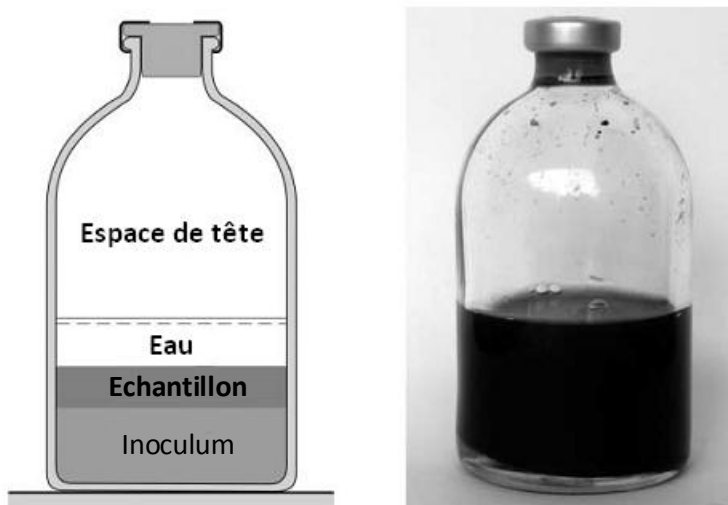


Figure 5 : Mise en œuvre des fioles à plasma couramment utilisées pour le test BMP (Angelidaki et al., 2009).

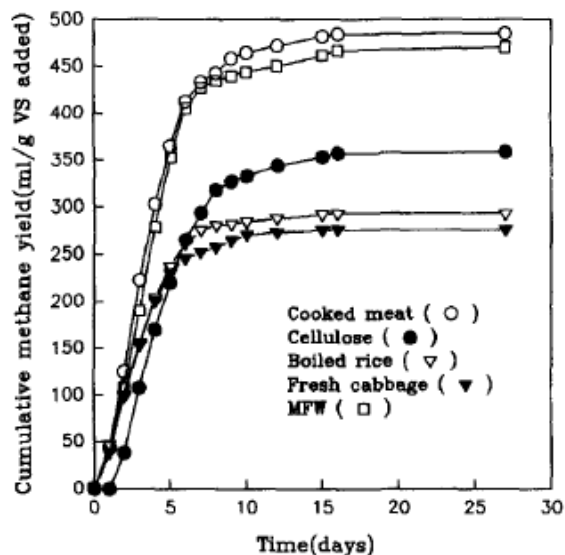


Figure 6 : Production de méthane cumulé de divers échantillons (Cho et al., 1995).
MFW: Mixed Food Waste (mélange de déchet de cuisine).

Du fait de sa durée, ce test ne peut pas être utilisé pour la régulation du procédé de digestion, ni même pour l'optimisation des mélanges des matières premières entrantes.

II.2.5. Sensibilité du test BMP

Plusieurs paramètres sont susceptibles de faire varier la valeur BMP mesurée sur un déchet : les plus importants sont la température, le pH, la taille des particules et le ratio inoculum/substrat.

- **La température** agit directement sur les réactions enzymatiques, en modifiant la conformation des enzymes. La température est généralement réglée à 35°C (condition mésophiles). Cependant, certaines études sont réalisées à 55°C (conditions thermophiles) conditions plus favorables à une vitesse de réaction plus élevée et une meilleure solubilisation des composés (Angelidaki et Sanders, 2004).
- **Le pH** joue également un rôle dans la configuration des enzymes, agissant indirectement sur la vitesse et l'efficacité des réactions. Normalement tamponné en début de réaction, le pH évolue lors de la fermentation sous certaines conditions. En effet, en cas d'accumulation d'acides gras volatils (AGV), le pH diminuera, engendrant une inhibition des bactéries méthanogènes. A contrario, une concentration importante des ions ammonium (NH_4^+), suite à l'hydrolyse de substrat azoté (protéines), augmentera le pH intracellulaire inhibant les voies métaboliques des bactéries méthanogènes (Chen et al., 1996 ; Angelidaki et Sanders, 2004 ; Chen et al., 2008).
- **La taille des particules** semble influencer à la fois la vitesse de réaction, et le rendement de production de méthane. Les fines particules seront plus facilement solubles et rapidement métabolisables par les micro-organismes, car directement accessibles. Les grosses particules devront être hydrolysées, ce qui est l'étape limitante de la digestion anaérobie. Plus la taille des particules est fine, plus l'étape d'hydrolyse sera courte et la vitesse globale de réaction diminuée. La taille des particules du déchet modifie parfois la valeur BMP. Lorsque le déchet est composé de fibres, le broyage augmente la surface accessible par les bactéries et libère également

de la matière organique facilement biodégradable renfermée dans le réseau de fibres. Le broyage améliore ainsi la digestibilité des déchets (Angelidaki et Sanders, 2004, Palmowski et Müller, 2000).

- **Le ratio inoculum/substrat (X/S)** a, en théorie, uniquement un effet sur la cinétique et le temps de latence de la réaction, mais n'en a pas sur la valeur de potentiel méthane (Chen et al., 1996 ; Nallathambi Gunaseelan, 1995 ; Raposo et al., 2006). Un faible ratio est préférable afin d'avoir le minimum de respiration endogène, qui peut biaiser les résultats. Chaque substrat a un ratio optimal prenant en compte sa capacité à produire des AGV et son pouvoir tampon par la production des ions ammonium, lors de l'hydrolyse des matières azotées. Le choix du ratio se présente donc comme un compromis entre le temps de latence, la respiration endogène et le pouvoir tampon du substrat (Angelidaki et Sanders, 2004).

De nombreuses publications traitent de la digestion anaérobie de différents types de déchets. Le Tableau 2 présente une liste - non exhaustive- de déchets et leurs valeurs BMP associées. La valeur BMP peut varier fortement suivant le type de déchet (déchet d'orge : 20 ml CH₄.ml⁻¹ MV et des déchets de cuisine : 472 ml CH₄.ml⁻¹ MV). On remarque également que la fraction organique des ordures ménagères présente une gamme importante allant de 275 à 410 ml CH₄.ml⁻¹ MV dépendant en partie de la zone géographique de production des déchets et de la saison.

II.2.6. Conclusion

Le test BMP est majoritairement utilisé pour évaluer la biodégradabilité anaérobie des déchets par la mesure du potentiel méthane. Ce test est utilisable pour tous les types de déchets, qu'ils soient solides ou liquides. Il a l'avantage de permettre la mise en évidence de composés inhibant la méthanisation. L'inconvénient majeur de ce test est sa durée de mise en œuvre, environ 30 jours. Ceci résulte du fait qu'il repose sur une fermentation et que la majorité des déchets devant être assimilables par les micro-organismes, doivent être hydrolysés. Avec un temps d'expérimentation aussi long, le test BMP n'est pas une méthode adaptée pour décider ou non de méthaniser un déchet, avec les risques qui en découlent : méthaniser un déchet avec peu de matière organique biodégradable, entraînerait la diminution du rendement dans la production de méthane, et réduirait la population microbiennes par manque de matière biodégradable.

Déchet	BMP (écart-type) (ml CH ₄ .g ⁻¹ VS)	Références	Déchet	BMP (écart-type) (ml CH ₄ .g ⁻¹ VS)	Références
Paille	195 (5,9)	6	Pelures de pommes	317	2
Pelures d'asperges	219 ^a	1	Haricots verts	343 ^a	1
Raisins	231 (15)	4	Lisier	353 (28)	6
Betteraves	231 (8)	4	Pomme de pin	356 (9)	4
Marc de café	255	3	Canne à sucre	357 (13)	4
Laminaria sp.	260-280	5	Pelures de carottes	388 (35)	2
Bois	260 (5) ^a	8	Herbes de tontes	388 (35)	2
Fraises	262 ^a	1	Macrocystis	390 – 410	5
Fraction fermentescible d'ordure ménagère	275 – 410	7	Petits pois	390 (13)	4
Lisier	275 (36)	6	Pelures de pomme de terre	390 (25)	2
Purée de pomme	279 ^a	1	Pelures d'oignons	400 (14)	4
Peau de banane	289 (16)	2	Sorgho	404 (36)	4
Laitue	294 (30)	2	Pellures de carottes	418 ^a	1
Tomates	297 (12)	4	Cellulose	419 (19)	4
Pelures d'oranges	297 (26)	2	Pelures de mangues	454 (13)	4
Pommes	306 ^a	1	Déchet de cuisine	472	9
Epinards	314 ^a	1	Citron	473 (11)	4

Tableau 2 : Comparaison des différents déchets avec leur valeur de potentiel méthane associée.

^a Valeurs calculées à partir des données de références des articles.

1 : Knol et al., (1978) ; 2 : Buffière et al., (2006) ; 3 : Neves et al., (2006) ; 4 : Nallathambi (2004) ; 5 : Chynowet et al., (1993) ; 6 : Moller et al., (2004) ; 7 : Davidsson et al., (2007) ; 8 : Turick et al., (1991) ; 9 : Cho et al., (1995).

D'autres méthodes doivent être envisagées pour prédire au mieux et le plus rapidement possible le potentiel méthane des déchets. Une alternative serait de réduire la durée du test respirométrique.

II.3. Le potentiel méthane corrélé avec les tests respirométriques aérobies rapides

Les tests respirométriques aérobies sont basés sur la mesure de l'oxygène consommé ou du dioxyde de carbone dégagé (Bellon-Maurel et al., 2003). Ces tests sont couramment utilisés afin d'estimer la biodégradabilité aérobie de déchet, et leur cinétique de dégradation ou la maturité et la stabilité des composts (Lasaridi et al., 1998 ; Calmon et al., 1999). Ces tests aérobies durent en général 30 jours. Cependant, certains tests aérobies peuvent être réalisés en quelques jours, certains en quelques heures. Deux types de protocoles peuvent être appliqués en conditions aérobies :

- soit sans aération, la fiole est alors fermée et le gaz produit est mesuré en fin d'expérimentation (test dit « closed bottle »),
- soit avec une aération constante, le gaz étant mesuré à intervalle réguliers.

Chaque méthode fournit un indice qui lui est propre, comme par exemple le test DO_{20} (Demande en Oxygène d'un test aérobie de 20h), ou encore le test DBO_5 (Demande Biologique en Oxygène en 5 jours) (Barena et al., 2009 ; Wagland et al., 2009 ; Sanchez, 2009 ; Schievanio et al., 2008)). La biodégradabilité, définie au II.2.1., est également exprimée comme le rapport DBO/DCO (Beltrán et al., 1999 ; Cossu et al., 2008).

Bien que ces méthodes reposent sur le principe de la respiration aérobie, elles permettent également d'évaluer la production potentielle de méthane en conditions anaérobies. Certaines études ont, en effet, montré des corrélations intéressantes entre l'indice de respiration et la production de méthane en condition anaérobie. Des corrélations de $R^2 = 0,80$ et $R^2 = 0,94$ ont été obtenues entre l'Indice de Respiration (RI_{4+} , l'oxygène cumulé consommé en 4 jours), et l'indice de production de biogaz, GS_{21} , (Generation Gas Sum), représentant le cumul de biogaz mesuré après 21 jours d'incubation (Cossu et al., 2008 ; Ponsa et al., 2008) (Figure 7).

Ces méthodes aérobies offrent la possibilité de prédire le potentiel méthane plus rapidement que le test BMP, et d'être également utilisables pour n'importe quel type de déchet.

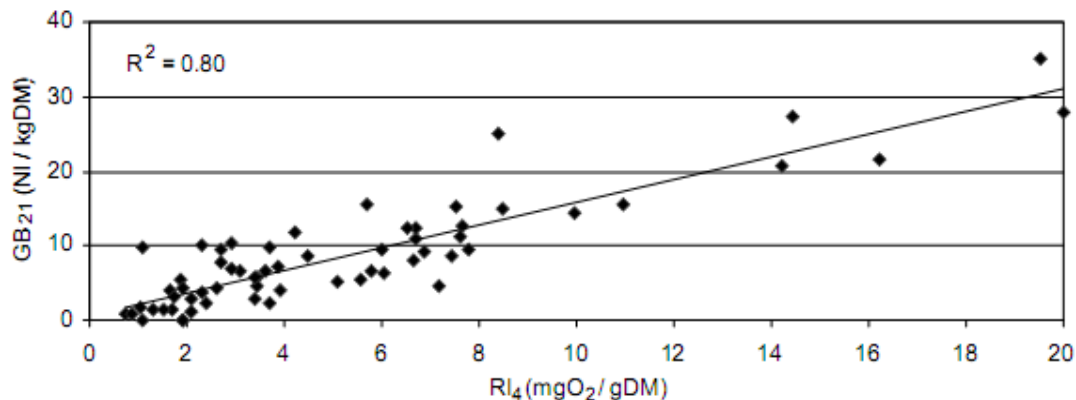


Figure 7 : Corrélation entre l'indice aérobic RI_4 et l'indice de production de biogaz (GB_{21}) (Cossu et al., 2008).

Cependant, la faible durée du test suggère que seules les molécules facilement et rapidement assimilables sont dégradées. Les molécules complexes mais biodégradables, comme la cellulose, ne sont pas prises en compte dans les mesures (Wagland et al., 2009). De plus, comme le test repose sur des réactions biologiques, son résultat est toujours dépendant de l'activité de l'inoculum. Egalement, en test aérobic la croissance bactérienne est plus importante qu'en test anaérobic, le dioxyde de carbone alors dégagé n'est pas représentatif de la biodégradabilité de l'échantillon.

La réduction du temps d'analyse via des méthodes respirométriques aérobies est donc une voie alternative mais ces tests sont encore trop longs et consommateurs de mains d'œuvre. C'est pourquoi d'autres méthodes d'évaluation du BMP, indirectes, ont été recherchées : elles passent par la caractérisation chimique ou biochimique de la matière organique.

II.4. Prédiction de la valeur BMP basée sur l'étude de la matière organique des déchets

Alors que les méthodes décrites précédemment tentent à la fois de mimer le processus de dégradation de la matière et de l'accélérer, l'approche suivante consiste à chercher le lien entre le potentiel méthane et la composition chimique des déchets. Trois approches seront utilisées selon qu'elles utilisent:

- des tests d'activité enzymatique, qui peuvent être en lien avec la matière organique présente dans le déchet,
- la composition élémentaire (C, H, O, N) du déchet,
- la composition moléculaire (glucides, protéines, fibres...)

II.4.1. Les tests enzymatiques

Les réactions biochimiques de la méthanogénèse sont toutes réalisées par des enzymes, intra ou extra-cellulaires. La mesure de l'activité d'enzymes spécifiques à la méthanogénèse permettrait une meilleure connaissance de la matière organique présente dans le déchet. Cette approche a déjà été utilisée pour évaluer la dynamique du processus de compostage (Pelaez et al., 2004). Les enzymes utilisées pour l'étude étaient des protéases, cellulases, amylases, invertases, phosphatases et des déshydrogénases. Le pool d'enzyme à utiliser est bien sûr en lien avec la complexité du déchet à analyser. Rodriguez et al., (2005) mettent en œuvre une méthode enzymatique afin de prédire la biodégradabilité d'ordures ménagères. Ils utilisent des enzymes cellulolitiques et hemicellulolitiques afin de corréliser les résultats des tests enzymatiques aux valeurs BMP de 3 lots d'échantillons provenant de 3 sites d'enfouissements différents en Belgique (37 échantillons au total). Les résultats de corrélations (R^2) montrent l'intérêt de cette méthode (Figure 8).

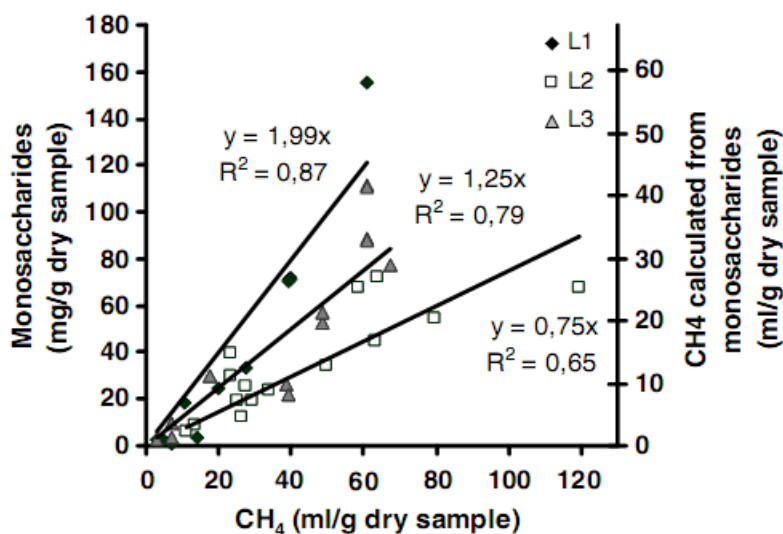


Figure 8 : Corrélation entre la quantité de sucre libérée par hydrolyse enzymatique et les tests BMP (Rodriguez et al. 2005).

Cependant, les auteurs nuancent leurs résultats sur plusieurs points. Les enzymes du test semblent avoir moins dégradé la matière organique que lors du processus anaérobie. Le pool d'enzymes provenant des micro-organismes est indéniablement plus complet que le test enzymatique. Une partie du méthane produit provient également de source carbonée autre que les sucres, tels que les protéines et les lipides. Une partie de la cellulose peut aussi ne pas être directement accessible aux enzymes, car confinée dans des particules plus grosses. Or les micro-organismes ont un pool d'enzymes plus important, ce qui permet d'hydrolyser une plus grande partie de la matière. Pour des déchets simples n'ayant qu'un type de molécule majoritaire, le pool d'enzyme peut être facilement constitué. Pour des déchets complexes, comme les ordures ménagères, afin d'être représentatif de l'ensemble de la matière organique méthanisable, le pool d'enzyme doit être également plus complexe, donc plus fastidieux à élaborer. Qui plus est, le test enzymatique précédant dure 48h, avec un pool d'enzyme simple. Dans les cas plus complexes, la durée des tests peut être plus longue et donc nécessiter davantage de mise en œuvre.

II.4.2. Analyse de la composition élémentaire

En 1933, Buswell (Buswell et Neave, 1930) développent la formule de Buswell (3) pour le calcul de la valeur théorique de méthane ($B_{o,th}$) produit par digestion anaérobie du déchet, à partir de sa composition élémentaire C, H, O, N.

$$C_n H_a O_b N_c + \left(n - \frac{a}{4} - \frac{b}{2} + \frac{3c}{4}\right) H_2 O \rightarrow \left(\frac{n}{2} + \frac{a}{8} - \frac{b}{4} - \frac{3c}{8}\right) CH_4 + \left(\frac{n}{2} - \frac{a}{8} + \frac{b}{4} + \frac{3c}{8}\right) CO_2 + c.NH_3$$

$$B_{o,th} = \frac{22.4 \cdot \left(\frac{n}{2} + \frac{a}{8} - \frac{b}{4} - \frac{3c}{8}\right)}{12n + a + 16b + 14c} L CH_{4,STP} \cdot g^{-1} MV \quad (3)$$

L'équation chimique représente la réaction de méthanisation, avec à gauche, la matière organique composée de carbone, hydrogène, oxygène et azote, qui réagit avec l'eau pour former à droite, du méthane, du dioxyde de carbone et de l'ammoniac. La valeur théorique de méthane représente la quantité maximale théorique de méthane qu'un déchet pourrait produire en digestion anaérobie, si toute la matière qu'il contient était biodégradable et convertie en méthane.

La mesure de la composition élémentaire est très rapide, ne nécessitant que quelques heures. Les quantités de carbone et d'hydrogène peuvent être déterminées par combustion. L'azote est déterminé par une minéralisation totale et dosé par la méthode de Kjeldahl. L'oxygène est mesuré par pyrolyse et analyse du gaz CO. L'autre inconvénient de la méthode, en dehors de la mesure de l'oxygène, est qu'elle ne différencie pas la matière biodégradable de la matière non biodégradable. Or la matière biodégradable uniquement est utilisée par les micro-organismes pour produire du méthane. De plus, une partie de cette matière biodégradable est utilisée pour la croissance bactérienne, et n'est donc pas convertie en méthane (Hansen, 2005). La valeur théorique de méthane surestime donc toujours la valeur BMP expérimentale ($B_{o,th} > BMP$) (Davidsson et al., 2007).

La meilleure façon pour différencier la matière organique biodégradable de celle non-biodégradable est de mesurer la composition chimique (glucides, protéines, fibres, etc.).

II.4.3. Analyse de la composition chimique (glucides, protéines, fibres, etc.)

II.4.3.1. Le potentiel méthane des composants de la matière organique

La matière organique est fractionnée en composés facilement biodégradables (sucres, protéines et lipides) et en composés difficilement biodégradables (lignine, acide humique et ses dérivés) (Batstone et al., 2002 ; Eleazer et al., 1997 ; Neves et al., 2008 ; Lehtomäki et al., 2008 ; Lissens et al., 2004).

La matière organique biodégradable est divisée en différentes catégories. Les sucres simples que sont les monosaccharides (glucose, fructose ou galactose) et les disaccharides (saccharose ou lactose), les acides aminés et les acides gras volatils. Ces composés n'ont pas besoin d'une étape d'hydrolyse pour être dégradés. Ils sont directement assimilables. Les molécules complexes, comme les polysaccharides (amidon, cellulose, hémicellulose), les protéines et les acides gras longues chaînes (AGLC) doivent être hydrolysées en monomères avant d'être assimilés. La formule de Buswell donne la valeur théorique maximale de méthane pour des composés les plus couramment rencontrés. Les sucres ont le plus faible rendement de méthane avec $415 \text{ ml CH}_4 \cdot \text{g}^{-1}$

MV. Puis viennent les protéines, avec une valeur de BMP théorique de 496 ml CH₄.g⁻¹ MV. Cependant, l'hydrolyse des protéines engendre une solubilisation d'ion NH₄⁺, augmentant ainsi le pH, ce qui est néfaste voire toxique pour les bactéries méthanogènes (Hansen et al., 1998). La détermination du taux de protéines, ou d'azote total, des déchets permettrait d'identifier le risque de toxicité par NH₄⁺, avant l'utilisation du déchet. Et enfin, les lipides et leurs longues chaînes carbonées ont le plus fort rendement théorique avec 1014 ml CH₄.g⁻¹ MV. Cependant, les lipides, notamment les acides gras longues chaînes, peuvent également être sources de problèmes (Angelidaki et al., 1992). A forte concentration, ils peuvent avoir un effet inhibiteur en endommageant la membrane bactérienne (Cavaleiro et al., 2008). Les AGLC peuvent également former une barrière autour de la matière organique, empêchant l'accès aux micro-organismes (Neves et al., 2008 ; Mata-Alvarez (2003)).

Comme explicité précédemment, il existe également de la matière organique difficilement biodégradable sous conditions anaérobies. Les formes les plus couramment rencontrées dans la littérature sont la lignine, les acides humiques et ses dérivés. Par exemple, l'utilisation de déchets d'orge, comme substrat, composé à 69% de fibre ne permet d'obtenir que 20 ml CH₄.g⁻¹ MV (Neves et al., 2006 (a et b)).

II.4.3.2. Modèle de prédiction de la biodégradabilité à une ou deux variables.

Chandler et al. (1980) a mis en évidence l'effet négatif de la lignine sur la biodégradabilité des déchets. Plusieurs types de déchets dont de lisier, des journaux, de la paille, des feuilles et d'autres produits issus de l'agriculture, ont été utilisés pour développer le modèle liant la quantité de lignine et la biodégradabilité. Le modèle proposé par Chandler est le suivant (4) :

$$B = 0.83 - (0.028) \times X_1 \quad (4)$$

Avec B , la fraction biodégradable de la matière volatile ($0 < B < 1$) et X_1 ($0 < X_1 < 20\%$), le pourcentage de lignine (exprimé en % de MV). Le coefficient de corrélation du modèle proposé est de 0.94. Le coefficient négatif (- 0,028) indique que la lignine est non biodégradable. Le modèle montre également que pour chaque augmentation du pourcentage de lignine, la biodégradabilité de la matière volatile du déchet diminue de 3%. Enfin, selon l'équation ci-

dessus, il est possible de démontrer que seulement 83% de la matière volatile peut être convertie en méthane, les 17% restants étant convertis en biomasse et en métabolite.

Eleazer et al., (1997) propose une relation entre (cellulose + hémicellulose) et quantité de méthane produit. Une corrélation positive est démontrée avec plusieurs types de déchets (ordures ménagères, journaux, papiers, nourritures, feuilles, graines, herbes) (Figure 9).

D'après cette relation, plus la quantité de cellulose et d'hémicellulose est élevée, plus la production de méthane sera importante. La plupart des points suivent cette relation. Seul le point représentant la nourriture, « food », est en dehors de la relation. Sa valeur BMP est supérieure à celle prédite par le modèle, sûrement dû à la présence de lipides dans l'échantillon, qui augmentent de façon importante la valeur BMP.

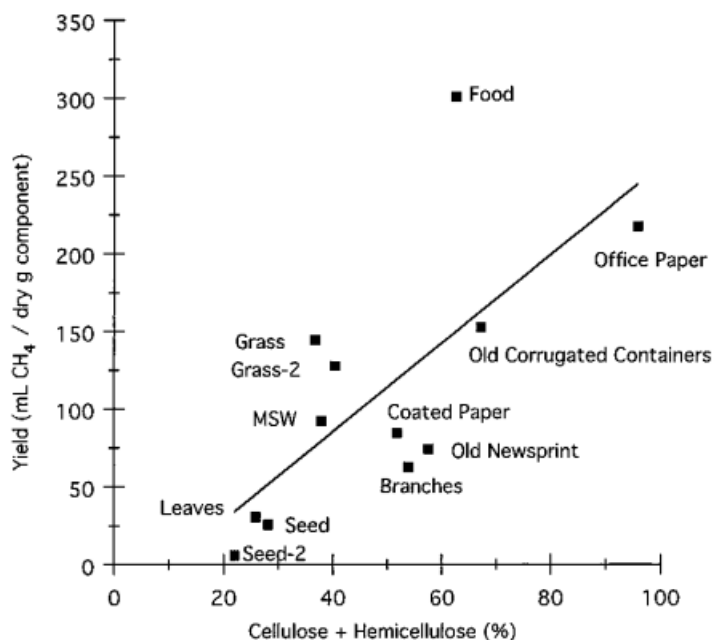


Figure 9 : Relation entre le rendement en méthane et le pourcentage en glucides complexes (Cellulose + Hémicellulose) (Eleazer et. al, 1997).

Buffière et al., (2006) propose une relation entre (lignine + cellulose) et la biodégradabilité comme définie en II.2.1. (Figure 10).

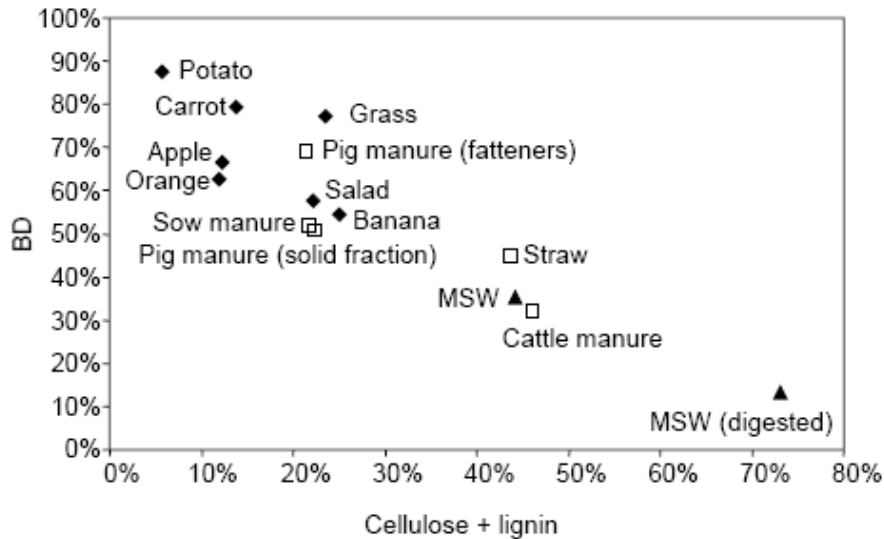


Figure 10 : Relation entre (cellulose + lignine) et la biodégradabilité de divers déchets non broyés (Buffière et al., 2006).

Dans cette relation, la cellulose est considérée comme non biodégradable. Cette relation est en contradiction avec la précédente, où la cellulose était considérée comme biodégradable. Avoir ajouté la cellulose, un composé communément admis comme étant biodégradable à un composé non biodégradable, la lignine est surprenant. Cependant des études récentes ont montré l'importance du broyage lors de l'obtention du potentiel méthane de déchets solides lignocellulosique. En effet, la lignine joue un rôle de protection dans le monde végétal. Composé non biodégradable, la lignine empêche les micro-organismes de pénétrer dans les cellules végétales. Après un broyage fin, la matière est alors plus facilement accessible.

Cependant ces modèles ne permettent pas de prédire la biodégradabilité ou le potentiel méthane de déchets n'ayant pas de cellulose ou de lignine, tout comme les déchets fortement lipidiques, tels que les déchets d'abattoirs ou d'industries laitières.

Certains modèles prennent en compte plus de 2 composés.

II.4.3.3. Modèle à plusieurs variables

Certains modèles de régression arrivent à prédire correctement le potentiel méthane de déchets, à l'aide de plusieurs paramètres. Nallathambi Gunaseelan (2007) établit plusieurs modèles de

régression sur des extraits de fruits et de légumes. Les modèles à un seul paramètre (glucides totaux) ou deux (glucides/fibres) ne donnent pas de bonnes régressions. Seuls des modèles à plusieurs paramètres conduisent à une bonne prédiction. Les meilleurs résultats sont obtenus en utilisant les composés suivants : glucides totaux, lignine, lignine/fibres extraites au détergent acide (ADF), le ratio Carbone total/ADF, azote total (NTK) et le résidu après combustion à 500°C. Ce modèle affiche un R^2 de 0,90 et en se restreignant à 2 types de substrats, il passe à un R^2 de 0,99. Le calcul de la racine carrée de la moyenne des écarts (entre la valeur expérimentale et prédite) au carré (Root Mean Square Error of Prediction, RMSEP) permet de juger de la qualité du modèle. La valeur de RMSEP obtenue est ici de 73 ml $\text{CH}_4 \cdot \text{g}^{-1}$ MV. Avec comme valeur moyenne du BMP des échantillons de 370 ml $\text{CH}_4 \cdot \text{g}^{-1}$ MV, cela fait près de 20% d'erreur relative. L'équation donnée par le modèle est la suivante (5) :

$$BMP = 0,45 + 0,35 \times C - 0,32 \times ADF - 0,18 \times L / ADF - 0,41 \times N - 3,4 \times A \quad (5)$$

Avec B_0 , la valeur de potentiel méthane donnée par le modèle, avec C, ADF, L, N et A les valeurs respectives des glucides totaux, fibres extraites aux détergent acide (Acid Detergent Fiber), la lignine, l'azote total et le résidu après combustion à 500°C.

Les seuls coefficients positifs sont ceux des glucides totaux. Les protéines ont un effet négatif sur la valeur BMP, peut-être dû à l'accumulation des ions NH_4^+ . Ces résultats sont à tempérer par le faible nombre d'échantillons utilisés (23 en étalonnage et 6 en validation) et par leur origine et catégorie très spécifiques. Bien que les fruits et légumes ne contiennent que peu de lipides, l'ajout de la concentration en lipides aurait pu expliquer une part de la variabilité des valeurs BMP.

Mottet et al., (2010) prédit également la valeur de potentiel méthane de boues de stations d'épuration à partir de la composition chimique (glucides, protéines, lipides, DCO/COT, carbone organique soluble) (Figure 11).

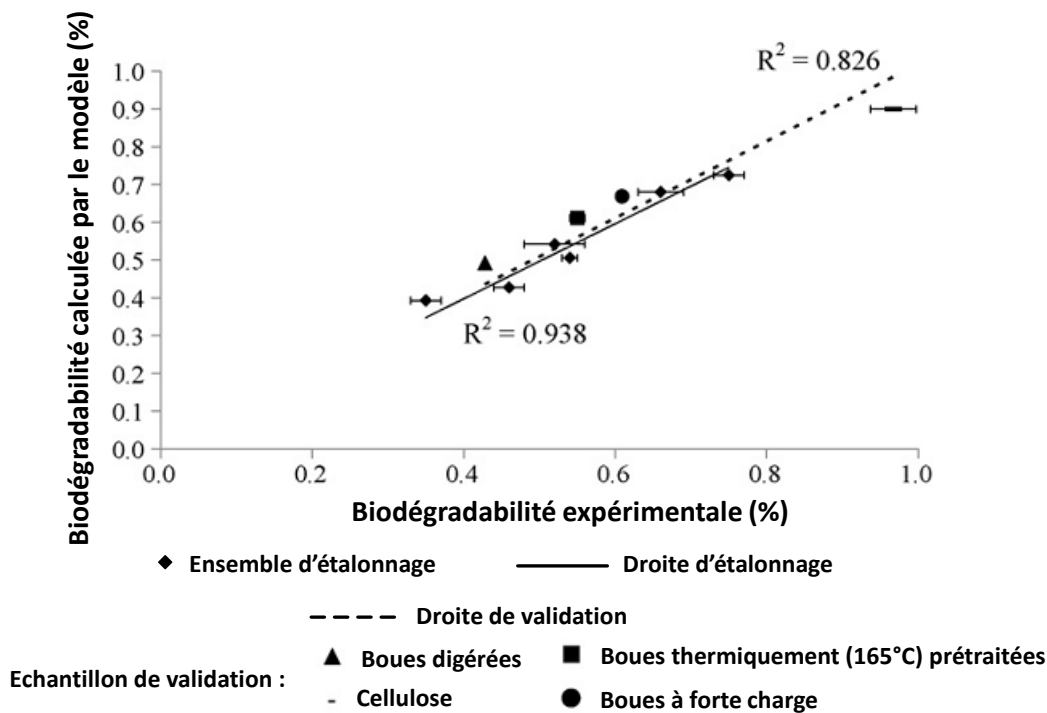


Figure 11 : Prédiction de la biodégradabilité de boues (Mottet et al., 2010).

D'autres modèles ont été établis en utilisant conjointement la composition de la matière organique et des paramètres associés à des tests respirométriques rapides.

Schievano et al., (2008) ont réalisé un modèle prédictif basé sur la quantité de matière volatile, le carbone organique total (COT), la phase soluble des déchets obtenue par la méthode de Van Soest (Van Soest et al., 1991) et le taux spécifique de consommation d'oxygène (la demande en oxygène pendant 20h, OD₂₀). Un R² de 0,92 est obtenu, sans validation du modèle. Certaines relations entre plusieurs variables indépendantes et la valeur BMP sont précisées. Certaines d'entre elles sont négatives, notamment avec l'azote total, la quantité d'ion ammonium, la lignine, et P₂O₅. Ces relations négatives indiquent des interférences possibles à de fortes concentrations de ces composés, diminuant ainsi la production de méthane en condition anaérobie. De plus les composés, NH₃, NTK et ADL semblent avoir une valeur seuil au-dessus de laquelle l'inhibition agit. L'estimation de ces composés sur les déchets donnerait une estimation du potentiel inhibiteur du déchet.

Des travaux plus récents de Schievano et al., (2009) tempèrent ces résultats par l'origine commune des déchets. Ces auteurs proposent un nouveau modèle étalonné à partir de 23 déchets

différents (ordures ménagères, déchets de fruit et légumes, déchets de viande et poissons, déchets de laiteries, boues, lisiers, fèces des volailles et lapins, des céréales, de l'huile) et le valident sur 12 autres déchets (légumes, extraits de levure, boues, lisiers, compost). Le modèle, basé uniquement sur le test respirométrique DO_{20} et la quantité de matière organique (MV), donne les meilleurs résultats ($R^2 = 0,88$), ce qui finalement, revient à corrélérer le test BMP à un test respirométrique plus court. Dans ce cas, le RMSEP et le biais sont respectivement de 134 et de 42 $ml\ CH_4.g^{-1}\ MV$ et un RPD (Ecart type des échantillons/RMSEP) de 2,7. Ce RMSEP semble élevé, sûrement du fait de l'hétérogénéité importante entre les différents échantillons.

Deux modèles théoriques prédisant la valeur BMP ont été comparés au test BMP sur des fractions fermentescibles d'ordures ménagères (Davidsson et al., 2007). Un premier modèle a été obtenu par la méthode de Buswell (Buswell et Neave, 1930) utilisant la composition élémentaire (C, H, O, N) (partie II.4.2). Le second est obtenu par la méthode utilisant la composition chimique (lipides, protéines, glucides). La figure 12 montre les résultats expérimentaux. Le BMP théorique, basé sur la composition élémentaire, surestime toujours les valeurs par rapport aux valeurs obtenues par la méthode conventionnelle.

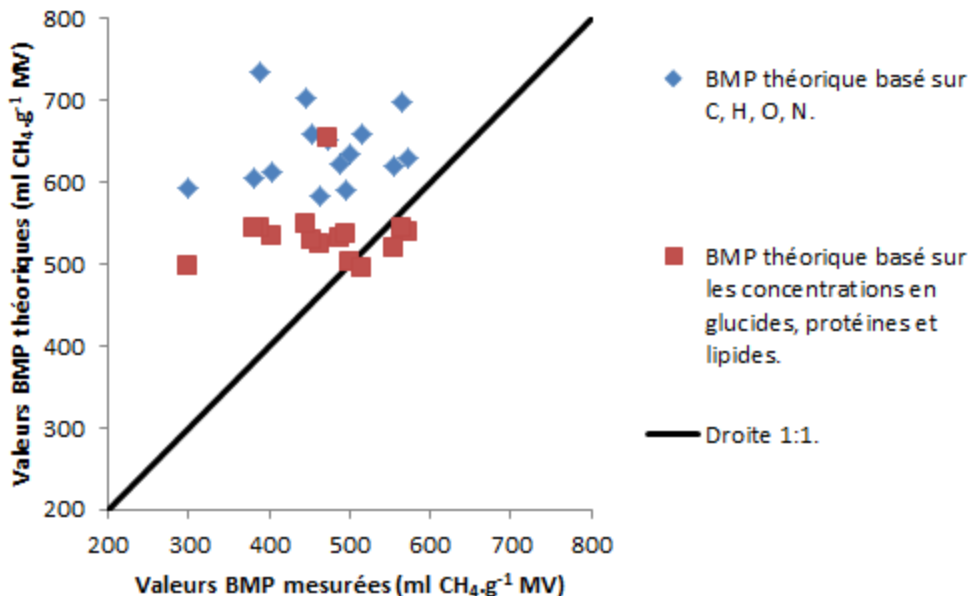


Figure 12 : Régression entre les valeurs BMP mesurées et les valeurs BMP théoriques obtenues par la méthode (C, H, O, N) et la méthode (glucides, protéines, lipides) pour plusieurs fractions fermentescibles d'ordures ménagères. (Davidson et al., 2007)

II.4.4. Conclusion

L'hypothèse du lien entre la biodégradabilité et la composition chimique conduit à la proposition de modèles validés et prédisant le potentiel méthane. Ces méthodes permettent une prédiction du potentiel méthane sans réalisation de fermentation et une meilleure compréhension de l'origine de la biodégradabilité des déchets. Cependant, pour utiliser ces modèles, il est nécessaire au préalable de mettre en œuvre des techniques de caractérisation de la matière organique afin d'obtenir toutes les valeurs d'entrées des modèles. Ces méthodes ont recours à l'emploi de produits dangereux (mesure de la DCO, utilisation d'acides...) et certaines d'entre elles, notamment la détermination des fibres par la méthode de Van Soest, peuvent être longues.

Des méthodes analytiques rapides peuvent être utilisées afin de réduire le temps d'analyse. Nécessitant généralement peu d'échantillon et peu ou pas de préparation, les méthodes spectroscopiques non destructives semblent les plus appropriées pour une caractérisation rapide de la valeur BMP d'échantillons complexes.

II.5. Utilisation de méthodes non-destructives : les méthodes spectrophotométriques

La spectroscopie est la science qui étudie les interactions entre les photons et la matière. Les photons sont caractérisés par leurs longueurs d'onde. Les longueurs d'onde les plus utilisées dans l'étude de la matière organique sont l'UV-visible, le proche infrarouge (PIR) et le moyen infrarouge (MIR).

II.5.1. La spectrophotométrie UV-Visible

II.5.1.1. Principe

L'UV-visible s'étend de 200 nm à environ 800 nm, avec la séparation entre l'UV et le visible autour de 400 nm. L'UV-visible est le siège des interactions électroniques. Ce sont les électrons de valence (positionnés sur les couches externes des atomes) qui absorbent les photons

correspondant au niveau d'énergie de l'UV-visible. Cette méthode est utilisée pour l'étude des molécules insaturées, telles que les molécules aromatiques (les lignines), les lipides complexes et les protéines avec des acides aminés aromatiques. De nombreuses études ont montré l'intérêt de la spectroscopie UV-visible pour appréhender la matière organique dissoute. La mesure est réalisée par transmittance à travers une cuvette en plastique (350 – 800 nm) ou en verre (250 – 800 nm) ou en quartz (200 – 800 nm). L'étude de composés solides par spectroscopie UV-visible passe forcément par une étape de solubilisation dans un solvant.

II.5.1.2. Application de la spectrophotométrie UV-Visible à la caractérisation de la matière organique dans des matrices complexes

Dans l'étude des déchets, la spectroscopie UV-visible a été utilisée pour la caractérisation d'eaux usées de fabrication d'huile avant et après dégradation aérobie (El Hajjouji et al., 2007). Ces auteurs attribuent le pic à 290 nm aux composés aromatiques générés par la dégradation de la lignine. La spectroscopie UV a également été utilisée pour le suivi du processus de compostage avec les mesures d'absorbances dans la plage 250 – 300 nm, caractéristiques de la phase d'humification des groupements aromatiques (Domeizel et al., 2004). Plusieurs ratios de longueurs d'onde sont utilisés pour le suivi du compostage. Le Tableau 3 en fournit quelques exemples.

Longueurs d'onde (nm)	Intérêt de la zone spectrale
260 – 280	Identification des fractions lignine et des quinones.
460 – 480	Reflète la matière organique au début de la phase d'humification
600 – 670	Indice de la matière humifiée
Rapports utilisés	$Q_{2/6} = 280 / 664$; $Q_{4/6} = 472 / 664$; $Q_{2/4} = 280 / 472$

Tableau 3 : Longueurs d'onde d'intérêt dans l'UV/Visible (Zbytniewski et Buszewski, 2005).

Ces ratios ont également été proposés afin d'évaluer la maturité du compost des déchets issus de la production d'huile (Sellami et al., 2008), ou de compost de boues de station d'épuration (Zbytniewski et Buszewski, 2005).

La déconvolution spectrale est également très utilisée en spectroscopie UV. La déconvolution consiste à reproduire un spectre d'échantillon inconnu à partir d'une combinaison linéaire de spectres de références (7). Les spectres de références peuvent être des spectres de la même molécule à différentes concentrations, ou des solutions types.

$$Sa(\lambda) = \sum_{i=1}^p a_i REF_i(\lambda) \pm r \quad (7)$$

avec, $Sa(\lambda)$, l'absorbance à la longueur d'onde λ de l'échantillon inconnu, $REF_i(\lambda)$, l'absorbance à la longueur d'onde λ des spectres de références i , a_i , les p coefficients associés aux spectres de références et r , l'erreur.

Domeizel et al. (2004) se sont basés sur la déconvolution de spectre UV pour établir des indices de maturité du compost, avec comme spectres de référence des spectres d'acide humique, d'acide fulvique et une fraction non humifiée. El Khorassani et al. (1999) a appliqué la déconvolution de spectre UV obtenus sur les eaux usées pour la détermination des nitrates, de la DCO, du carbone organique total et de la matière dissoute.

II.5.1.3. Spectrophotométrie UV/ Visible et mesure de la biodégradabilité et de la traitabilité

Comme expliqué au paragraphe II.3, le rapport DBO/DCO peut être utilisé comme un indice de la biodégradabilité (Beltrán et al., 1999 ; Cossu et Raga, 2008). L'estimation de la DCO a d'abord été effectuée à une longueur d'onde, 254 nm (Mrkva, 1983) puis à deux longueurs d'onde, 254 et 350 nm, afin d'atténuer l'effet de diffusion présent dans les milieux turbides (Matsché et Stumwhorer, 1996). Avec les progrès de l'instrumentation, l'acquisition de spectres est maintenant courante et rapide (quelques secondes). De plus, l'approche monolongueur d'onde réduit l'information sur la composition de l'échantillon comparée à l'exploitation du spectre complet et d'autre part, la matrice de l'eau usée peut aussi modifier la longueur d'onde du pic d'absorption (effet de matrice). L'utilisation de spectres complets a l'avantage de résoudre ces contraintes. Ainsi, pour estimer la DCO dans les eaux usées, une approche semi-déterministe basée sur une déconvolution du spectre UV et un modèle de corrélation a été validée sur un grand

nombre d'échantillons (Tsoumanis et al., 2010). Plus récemment, Fogelman et al. (2006) utilise le spectre UV (190 – 350 nm) et la mesure de la turbidité du milieu comme variables d'entrée dans un modèle à réseau de neurones pour prédire la DCO des eaux usées. Le modèle est cependant limité à une gamme restreinte de turbidités comprises entre 0 et 150 NTU (Unité de Turbidité Néphélométrique). Quant à la DBO₅, elle a été estimée dans des effluents et le lisier à 280 nm (Brookman, 1997).

Par ailleurs, la forme du spectre UV permet d'établir un indice jugeant de la capacité de l'échantillon à être biodégradé en se basant sur le SF (Shape Factor) (Muret et al., 2000) (8).

$$SF = -D(\lambda) / A(\lambda) \times 10^3 \quad (8)$$

Avec $A(\lambda)$, l'absorbance à la longueur d'onde λ , $D(\lambda)$, la valeur de la dérivée seconde à la longueur d'onde λ .

Au regard de la valeur de SF, 3 classes ont pu être identifiées (Figure 13) permettant de prévoir leur biodégradabilité au regard de l'allure des spectres UV.

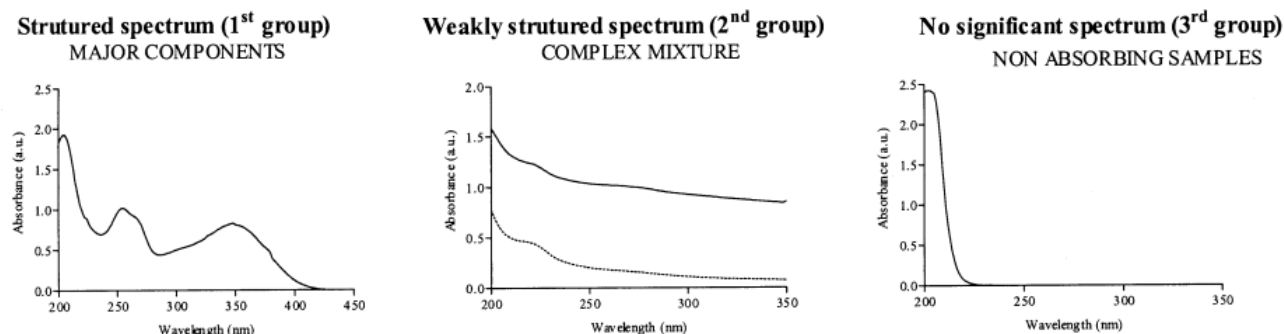


Figure 13 : Classification des différentes formes de spectres UV (Muret et al., 2000).

Le premier groupe (SF>3), est composé de spectres avec des pics bien identifiables, qui par comparaison à des spectres de références peuvent aider à l'identification.

Le second groupe (0,3<SF<3), est composé de spectre UV monotone signe de mélanges complexes, ou présentant des particules en suspensions.

Le troisième groupe (SF<0,3) correspond à des composés qui n'absorbent pas en UV ou de solution fortement minérale.

II.5.2. *La spectrophotométrie de fluorescence*

Lors de l'absorption de photons, les molécules passent de l'état initial à un état excité. Cet état excité est instable. Le retour à l'état initial peut se réaliser par ré-émission d'un photon, c'est la fluorescence.

Classiquement, les mesures de fluorescences se font sur les matières dissoutes, mais de la matière solide peut également être analysée directement par fluorescence (Garrimela Purna et al., 2005 ; Moreira et al., 2004). Seules les molécules ayant des insaturations (cycles aromatiques ou les doubles liaisons conjuguées) telles que les protéines via certains acides aminés (phénylalanine, tryptophane, tyrosine), les acides humiques et les lignines sont analysables par fluorescence. Les molécules telles que les glucides, lipides et acides gras volatiles ne peuvent pas être détectées directement par fluorescence.

Cependant des travaux ont montré l'intérêt de la fluorescence pour la détermination des AGV et de l'acide lactique par réaction de fixation de ces composés avec le N-(1-naphtyl) ethylenediamine (Robert-Peillard et al., 2009).

En lien avec la biodégradabilité, une relation linéaire a été établie entre la DBO₅ d'eaux usées et la fluorescence à 340 nm (excitation à 280 nm) (Reynolds et Ahmad, 1997). La fluorescence a également été mise en œuvre afin de suivre l'évolution de compost industriel et domestique (Miiki et al., 1997).

II.5.3. *La spectrophotométrie moyen-infrarouge : MIR*

II.5.3.1. Principe

Le moyen infrarouge s'étend de 2 500 nm à 25 000 nm (400 – 4000 cm⁻¹). Ces longueurs d'onde correspondent à des niveaux d'énergie de vibration ou de rotation des liaisons moléculaires. La mesure MIR d'échantillons liquides ou pâteux est généralement effectuée par « réflectance » en utilisant une cellule ATR (Attenuated Total Reflection ou Réflexion Totale Atténuée) (Figure 13). La mesure se fait également par transmission, à travers soit une cuvette en quartz pour les liquides, soit une pastille formée par le mélange de l'échantillon et de KBr pour les échantillons solides - qui doivent être au préalable séchés soit par lyophilisation en 24 à 48h ou

dans un four à 60°C pendant 48h puis broyés finement. Hormis ce temps de préparation, la mesure MIR se fait en quelques minutes.

II.5.3.2. Application au traitement des déchets

Dans le cadre du traitement des déchets, la spectroscopie MIR a déjà été utilisée pour diverses applications telles que la caractérisation de la matière organique ou le suivi de compostage. Le ratio des bandes d'absorbances 1650/2930 cm^{-1} (carbone aromatique/carbone aliphatique) a été proposé afin d'établir un indice décrivant le processus de compostage de déchets ménagers (Castaldi et al., 2005). Le suivi de l'indice met en évidence la diminution des molécules aliphatiques facilement biodégradables et en conséquence l'augmentation relative des molécules aromatiques telle que la lignine difficilement biodégradable (Castaldi et al., 2005). La « disparition » de certaines bandes d'absorbance est significative de la stabilisation du compost (Smidt et al., 2002 ; Smidt et Schwanninger, 2005). Par exemple, la bande à 2925 cm^{-1} attribuée aux groupements méthyles sur une chaîne aliphatique et la bande à 1240 cm^{-1} attribuée aux vibrations du groupement C-O de la fonction amide II, diminuent, voire disparaissent, lorsque le compost est arrivé à maturité. Différentes conditions de compostage (traitements, matières premières, stockage) peuvent être comparées par le suivi de l'absorbance à 2925 cm^{-1} informant de l'étape du process et de sa cinétique (Smidt et al., 2002). La spectroscopie MIR a également été employée afin de prédire la quantité d'acide humique ($r^2 = 0,86$; SEP = 2,8% de matière sèche) et l'indice de respiration ($r^2 = 0,92$; SEP = 3,5 mg.g^{-1} MS de compost) (Meissl et al., 2007).

En milieu liquide, la spectroscopie MIR peut être utilisée pour le suivi de procédés biologiques. Steyer et al. (2001 ; 2002) a utilisé un spectrophotomètre MIR en ligne afin de mesurer en continu certains paramètres de la digestion anaérobie (DCO, COT, AGV, alcalinité totale et partielle).

II.5.4. *La spectrophotométrie proche infrarouge : SPIR*

II.5.4.1. Principe et application

La spectroscopie proche infrarouge couvre la gamme 800 – 2500 nm. Ce domaine est le siège des harmoniques et des bandes de combinaison des modes de vibrations des molécules (Pasquini, 2003). La SPIR est couramment utilisée dans de nombreuses applications afin de classer des échantillons (Roussel et al., 2003) ou prédire un constituant particulier (protéines, glucides) (Osborne et Fearn, 1986) ou un paramètre comme la tendreté d'une viande (Prieto et al., 2009). La SPIR est également utilisée pour le suivi de procédé en milieu liquide. La production des AGV et de la biomasse lors de la production d'hydrogène en digestion anaérobie peut être suivie par SPIR (Lesteur et al, 2008). De même, la SPIR est appliquée par Jacobi et al. (2009) pour suivre la production de méthane et par Garrido-Vidal et al. (2004) pour le suivi de la fermentation acétique. Le suivi de procédé en milieu solide, comme la prédiction de la durée de compostage a déjà été réalisé par SPIR (Albrecht et al., 2008).

La spectroscopie proche infrarouge a l'avantage d'être non destructive et rapide. Pour des mesures de routine, elle peut s'avérer peu chère. En milieu liquide, la mesure s'effectue soit par transmission à travers une cuvette en quartz, soit par transflection - qui couple à la fois la transmission et la réflexion - par l'utilisation d'une coupelle munie d'un fond qui réfléchit la lumière infrarouge et double ainsi le trajet optique (Bertrand et al., 2006) (Figure 14). Pour les milieux solides, la mesure s'effectue presque exclusivement par réflexion diffuse. Le rayon lumineux pénètre de quelques dizaines de millimètres à quelques centimètres en fonction de la nature de l'échantillon. De plus la pénétration est inversement proportionnelle à la valeur de longueur d'onde. Certains appareils permettent également de réaliser des mesures en transmission diffuse, sur de fines surfaces de matières telle qu'avec des poudres, des comprimés ou des gels. Un autre intérêt de la SPIR est l'utilisation de fibres optiques, permettant de réaliser des mesures déportées ou en ligne lors de suivi de fermentation ou sans préparation d'échantillon (Pasquini, 2003 ; Lesteur et al., 2008 ; Hongqiang et al., 2008).

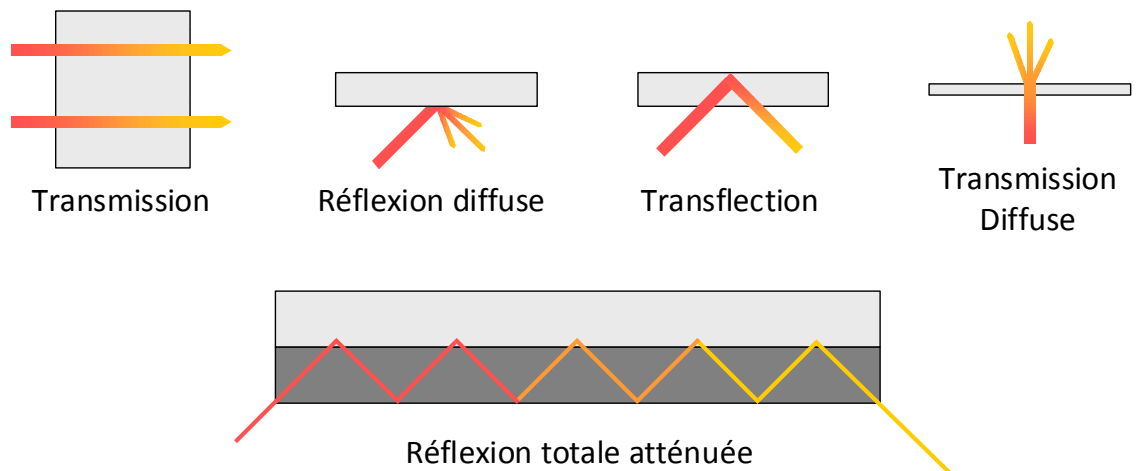


Figure 14 : Les différents principes de la mesure dans l'infrarouge

II.5.4.2. Utilisation de la SPIR pour prédire la biodégradabilité

La SPIR pourrait être utilisée de deux façons afin de prédire la biodégradabilité anaérobie.

La première façon d'utiliser la spectroscopie proche infrarouge pour prédire le BMP serait de déterminer la composition chimique des déchets (glucides, protéines, fibres, etc.) et d'en déduire la valeur BMP par un modèle de régression, comme décrit en II.4.3. La concentration, des composés jouant un rôle dans la biodégradabilité des déchets, a déjà été prédite dans de nombreuses matrices différentes (Bertrand et al., 2006). Par exemple, Sanderson et al. (1996) et Bruno-Soares et al. (1998) ont réalisé des modèles sur différentes plantes (avoine, orge, triticale, blé, ray-grass, sorgho, panic, canne à sucre, maïs, et divers écorces de bois) afin de prédire certains paramètres comme la matière volatile, les protéines, les fibres totales, les fibres extraites par la méthode Van Soest (NDF, ADF, ADL), le carbone total, l'azote total, le glucose, le galactose et d'autres sucres. Thuriès et al. (2005) ont construit un étalonnage robuste basé sur l'utilisation d'échantillons de natures et d'origines diverses : peau de raisin, pépins de raisin, café, graines et coques de chocolat, pulpe d'olive, riz, orge, tournesol ou soja. Les fibres (cellulose, hemicellulose, lignine) sont mieux prédites que la composition élémentaire (azote et carbone total). La SPIR a également été utilisée pour prédire certains paramètres des composts (Albrecht et al., 2008 ; Hermann et al., 2009 ; Vergnoux et al., 2009) tels que le pourcentage d'humidité, le pH, le carbone et l'azote organique, les acides humiques et fulviques, les fibres, l'indice de

stabilité biologique. Les paramètres tels que la DCO et la DBO5 ont été prédits à l'aide de la SPIR. La DCO des particules d'eaux usées a été mesurée avec de bons résultats en réflectance après filtration : $r^2 = 0.96$ et un SEP de $19 \text{ mg O}_2\cdot\text{L}^{-1}$ pour une gamme de valeurs allant de 75 à $300 \text{ mg O}_2\cdot\text{L}^{-1}$ (Sousa et al., 2007). La DBO₅ a également été mesurée par SPIR dans les eaux usées par transmission après filtration, SEP = $27,6 \text{ mg}\cdot\text{L}^{-1}$ pour une gamme de valeur comprise entre 35 et $275 \text{ mg}\cdot\text{L}^{-1}$ (Stephens et al., 2002). Cependant, l'inconvénient de prédire la composition chimique par la SPIR afin de l'utiliser dans un modèle pour en déduire la valeur de BMP serait le cumul des incertitudes.

La seconde façon d'obtenir rapidement la biodégradabilité est de la prédire directement par SPIR en utilisant les spectres des échantillons et un étalonnage dédié. Certains travaux relatent cette approche, pour l'étude de l'âge des composts, la décomposition de la litière végétale, la digestibilité des aliments pour animaux et la production de gaz suite à la digestion animale.

Albrecht et al. (2008) et Vergnoux et al. (2009) ont utilisé la SPIR afin de prédire l'âge des composts, ce qui se rapproche de la mesure de la biodégradabilité car au cours du compostage, la matière facilement biodégradable diminue, laissant place aux éléments moins biodégradables.

En sciences du sol, la spectroscopie SPIR a été utilisée pour prédire la décomposition de la litière végétale. Les paramètres prédits par SPIR sont plus spécialement, la perte de poids de la litière (Litter Mass Losses – LML) et le paramètre cinétique k de l'équation de décomposition (9) (Figure 15) (Gillon et al., 1999).

$$LML = e^{-kt} \quad (9)$$

En zoologie, de nombreuses études portent sur la prédiction de la digestibilité des aliments destinés aux animaux d'élevages. Adesogan et al., (1998) a étalonné des modèles prédictifs de la digestibilité *in vivo* de blé entier, à l'aide de différentes variables comme la digestibilité *in vitro*, la composition chimique ou les spectres infrarouges.

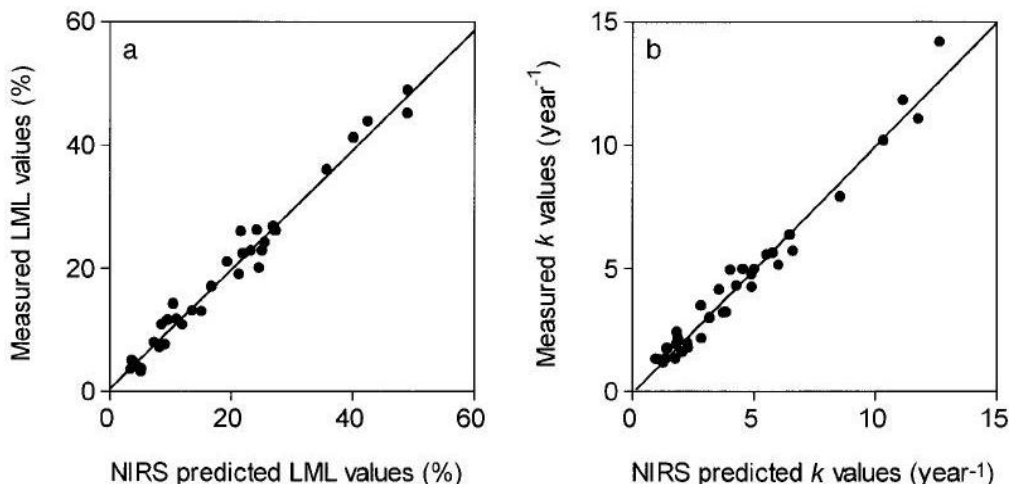


Figure 15: Relation entre les valeurs prédites par SPIR et les valeurs mesurées pour les données d'étalonnage pour (a) la perte de poids de la litière végétale après une semaine d'incubation et (b), la constante cinétique k du modèle exponentiel (Gillon et al., 1999).

Le meilleur modèle est obtenu avec la SPIR ($R^2 = 0,87$ et erreur Standard d'étalonnage de $13,0 \text{ g.kg}^{-1}$ pour une gamme de valeurs $558\text{-}708 \text{ g.kg}^{-1}$), mais il n'a pas été validé par un set de données indépendant. De nombreuses études montrent l'intérêt d'utiliser la SPIR pour prédire les paramètres de digestibilité de fourrages (De Boever et al., 1996 ; Adesogan et al., 1998 ; Lovett et al., 2004). Park et al. (1998) a même établi plusieurs modèles d'étalonnage afin de prédire divers paramètres de composition chimique (fibres, divers alcools et acides, azote total, sucres solubles) et la digestibilité en utilisant des spectres réalisés sur de l'ensilage frais, non séché. Les résultats de prédiction de la digestibilité sont satisfaisants: $\text{SECV} = 25,3 \text{ g MS.kg}^{-1} \text{ MS}$ ($r^2 = 0,9$) ; $\text{SEP} = 23,3 \text{ g MS.kg}^{-1} \text{ MS}$ ($r^2 = 0.85$) et un RPD de 2,72.

Toujours en zoologie, des études ont montré les potentialités de la SPIR pour la prédiction du volume de gaz produit (GP) *in vitro* après 6, 12, 24, 36 et 48h de fermentation d'herbe à gazon (*Kikuyu*) (Herrero et al., 1997). La prédiction sera plus tard améliorée en utilisant une base d'étalonnage homogène. Seule la valeur de l'asymptote (A), *i.e.* le volume de gaz cumulé dans l'équation (10), a été correctement prédite. Les autres paramètres, c et L , ne sont pas correctement prédits. D'après les auteurs, cela provient du faible accord entre la SPIR et les paramètres utilisés dans l'exponentielle de la formule cinétique.

$$GP = A[1 - e^{-c(t-L)}] \quad (10)$$

Cependant, une autre étude indique la possibilité de prédire le paramètre c de façon acceptable (Andrès et al., 2005). L'étude montre également que les prédictions par SPIR sont meilleures que celles prenant la composition chimique comme variables (Tableau 4). Les auteurs émettent l'hypothèse que le spectre complet mesuré par la SPIR permet de prendre en compte davantage de paramètres, telles que les interactions au sein de la matière organique comparé aux méthodes classiques.

Y	X (Variables)	R ²	SEP
GP24	NDF, ash, ADL	0.789	6.757
	NIR (p = 3)	0.889	3.877
GP96	Ash, ADL, LIG, CP	0.767	3.910
	NIR (p=6)	0.827	2.416
A	Ash, ADL, LIG, CP	0.819	3.716
	NIR (p=5)	0.801	2.711
c	NDF	0.906	0.0043
	NIR (p = 4)	0.941	0.0038
L	-	-	-
	NIR (p = 3)	0.597	0.288

Tableau 4 : Prédiction des paramètres cinétiques de la production de gaz cumulé par la composition chimique et la spectroscopie proche infrarouge (Andrès et al., 2005).

Comme précisé dans le tableau 4, le paramètre L n'est jamais correctement prédit que ce soit par la composition chimique ou par SPIR. Cela résulte du fait que des valeurs cinétiques ne dépendent pas seulement de la composition chimique des échantillons, mais aussi de l'inoculum bactérien. Si les inocula utilisés pour les différents échantillons ne sont pas identiques, c'est une variable supplémentaire à prendre en compte pour les valeurs cinétiques.

II.5.5. Conclusion

Les techniques spectroscopiques ont déjà été abondamment utilisées sur des échantillons liquides ou solides dans diverses applications. Chaque méthode présente des avantages et des inconvénients suivant la nature du matériau, de l'information recherchée, de la préparation nécessaire et donc du temps d'analyse nécessaire. De nombreux paramètres peuvent être prédits,

tels que la composition chimique, liée à la biodégradabilité, ou directement certains indices (DBO₅, digestibilité ou production de gaz de la digestion animale) proches de la définition de la biodégradabilité anaérobie et du test BMP de déchet.

II.6. Conclusion : choix d'une méthode rapide

Dans un but d'optimisation de la production de méthane, la connaissance rapide de la biodégradabilité anaérobie d'un déchet est primordiale. Actuellement la biodégradabilité d'un déchet est principalement déterminée par le test BMP, nécessitant environ 30 jours d'analyse, ce qui est incompatible avec une gestion en temps réel du procédé.

Des méthodes ont déjà été mises au point afin de réduire le temps d'analyse, notamment par :

- des tests aérobies plus courts,
- modélisation du lien entre biodégradabilité et qualité et quantité de matière organique contenue dans le déchet.

Certaines méthodes spectrophotométriques (UV ; fluorescence ; PIR ; MIR) ont été utilisées pour la caractérisation de la matière organique, voir pour déterminer des indices proches de la biodégradabilité.

La méthode la plus prometteuse est certainement la spectroscopie proche infrarouge, à la fois d'un point de vue théorique et pratique. Dans d'autres domaines d'application, elle a déjà été mise en œuvre pour caractériser la matière organique, prédire des stades de décomposition (compost, litière), les paramètres de digestibilité de fourrages et la production totale de gaz par les ruminants.

La spectroscopie proche infrarouge a de nombreux avantages :

- elle réalise une analyse à la fois qualitative et quantitative de la matière organique,
- c'est une méthode rapide,
- elle nécessite peu ou pas de préparation d'échantillon,
- c'est une méthode répétable,
- il existe des techniques robustes pour l'analyse en ligne,

- l'analyse peut être déportée par l'utilisation de fibres optiques.

De plus, les informations fournies par le PIR et le MIR sont redondantes et de nombreuses applications ont déjà été réalisées en PIR qui sont proches de notre thématique (mesure de la digestibilité), favorisant le choix du PIR sur le MIR. Egalement, le MIR est plus sensible aux conditions de préparation des échantillons. Concernant la spectrophotométrie UV, de par ses inconvénients, et en particuliers son manque de sensibilité aux glucides, ne sera pas plus approfondie dans le cadre de cette thèse. De plus d'un point de vue pratique, l'UV nécessite obligatoirement de mettre en solution les échantillons, ce qui augmente la durée de préparation et peut induire des erreurs de manipulation. Par conséquent, le travail présenté dans le cadre de cette thèse sera principalement axé sur l'application de la SPIR

Par conséquent, la partie qui suit, a pour objet d'étudier le potentiel de la spectroscopie proche infrarouge pour la prédiction du potentiel méthane, dans un contexte simplifié, par l'utilisation d'un seul type de déchets solides, les ordures ménagères.

III.

*Faisabilité de la spectroscopie proche
infrarouge pour la prédiction du potentiel
méthane de déchets solides*

Dans ce chapitre est abordée la prédiction par spectroscopie proche infrarouge (SPIR) de la valeur BMP d'un **ensemble homogène** de déchets solides issus d'une usine de tri-compostage.

Ce chapitre correspond largement à la traduction de l'article :

M. Lesteur, E. Latrille, V. Bellon Maurel, J.M. Roger, C. Gonzalez, G. Junqua, J.P. Steyer

First step towards a fast analytical method for the determination of biochemical methane potential of solid wastes by near infrared spectroscopy

Accepté dans Bioresource Technology en septembre 2010

III.1. Introduction

Les ordures ménagères sont composées de différentes catégories de déchets (ADEME, 2009) dont 2/3 ont un potentiel de valorisation organique et pourraient ainsi être utilisées en digestion anaérobie. Leur composition moyenne est de 25% de déchets putrescibles, 15% de papiers, 7% de cartons, 9% de textiles (tissus, coton, essuie-mains) et 7% de fines particules. Cette multitude de catégories, au niveau de leur composition, fait des déchets ménagers une source de matière organique intéressante pour la production de méthane, mais cela augmente d'autant plus la complexité et la variabilité de ces déchets. L'optimisation de la production de biogaz nécessite de connaître, sans délais trop importants, le potentiel méthane des déchets avant leur utilisation afin d'éviter un stockage des déchets. La valeur de potentiel méthane de déchet est actuellement déterminée par le test BMP, reposant sur une fermentation (Angelidaki et al., 2009). Mais il a été montré par l'utilisation de modèle de prédiction que cette valeur dépend exclusivement de la quantité et de la qualité de la matière organique (Eleazer et al., 1997 ; Buffière et al. (2006) ; Nallathambi Gunaseelan (2007) ; Schievano et al. (2008), Mottet et al., 2010). La spectroscopie proche infrarouge se révèle donc un outil potentiellement intéressant pour obtenir plus rapidement cette valeur de potentiel méthane, en mettant à profit sa capacité à analyser globalement la matière organique de matrice diverses. L'objectif de l'étude présentée ci-dessous est de démontrer les potentialités offertes par la SPIR pour prévoir le BMP dans le cas de fraction fermentescible d'ordures ménagères (FFOM) issus d'usines de traitement.

III.2. Matériel et méthodes

III.2.1. Les échantillons utilisés

Ces travaux ont été réalisés à l'aide de 102 échantillons, incluant 68 déchets ménagers (DM) extraits d'un site unique de tri-compostage, 7 DM provenant d'un site différent, 9 déchets dits "simples" dont : deux types de carton, riz, biscotte, pomme de terre, papier, un mélange de légumes, pain, un déchet obtenu par le mélange de certains des déchets « simples » précédents et 18 déchets lignocellulosiques. Les déchets issus de l'usine de tri-compostage proviennent d'une chaîne de transformation schématisée Figure 16.

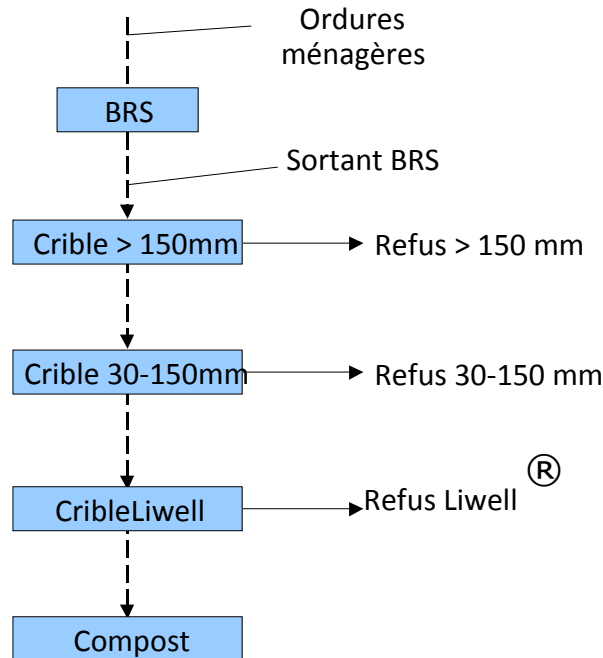


Figure 16 : Présentation des différentes étapes du process de tri-compostage.
BRS : Bioréacteur stabilisateur.

Le BRS est un long tuyau rotatif, à l'intérieur duquel une segmentation de la matière est réalisée. Il permet également de casser les verres en fines particules. Les déchets y séjournent environ 2 jours et y subissent un compostage accéléré.

Les échantillons sont issus de différents points de cette chaîne: ordures ménagères brutes, déchets à la sortie du BRS, déchets issus des différents refus de crible et compost final.

Ces échantillons sont ensuite divisés en 3 ensembles, comme présenté Tableau 5.

	Déchets ménagers #1	Déchets simples	Déchets lignocellulosiques	Déchets ménagers divers	Total
Ensemble d'étalonnage	45*	5 ^a	1 ^a		51
Ensemble de test	23*				23
Ensemble de validation indépendante		4 ⁺	17 ^b	7 ⁺	28
Total	68	9	18	7	102

Tableau 5 : Répartition des différents déchets dans chaque ensemble.

a: tous scannés le même jour ; b: tous scannés le même jour ; *: scanné sur 4 jours différents ; +: scans indépendants

L'ensemble d'étalonnage (51 échantillons) a permis de réaliser l'étalonnage du modèle de prédiction de la valeur BMP. Un ensemble d'étalonnage « réduit » ne comprenant pas les déchets simples a également été utilisé. L'ensemble de test (23 échantillons) a servi à effectuer un premier test du modèle. Ces 23 échantillons ont été choisis de la façon suivante : les 68 DM ont été classés par ordre croissant de valeur BMP et un échantillon sur trois a été pris pour construire l'ensemble de test (Figure 17).

L'ensemble de validation indépendant (28 échantillons) a été utilisé pour valider le modèle sur des échantillons très différents (différentes origines, différentes natures, différentes dates d'analyse spectrale). Pour intégrer d'éventuels problèmes de reproductibilité liés au spectrophotomètre, l'ensemble des déchets ont été scannés à des temps différents.

Afin d'améliorer la robustesse du modèle avant de le valider sur l'ensemble de validation indépendant, les ensembles d'étalonnage et de test ont été rassemblés pour construire un modèle final plus robuste suivant les recommandations de Dardenne (2010).

Tous les échantillons ont été lyophilisés et broyés. Le broyeur utilisé était un MF 10 basic IKA WERKE avec un tamis de 2 mm. Le broyage a permis d'améliorer l'homogénéité des matériaux et donc l'échantillonnage.

III.2.2. Les tests BMP

Les tests BMP ont été effectués en double. Pour chaque échantillon, 1,35g de matière sèche de chaque échantillon a été placé en condition mésophile (35°C) dans des fioles de 0,6L. Une boue anaérobie a été utilisée comme inoculum. Pendant 35 jours, l'accumulation de méthane a été mesurée tous les 2 jours à l'aide d'un Micro-GC (Chromatographie en phase gazeuse) Varian CP4900. La valeur BMP correspond alors à la quantité de gaz cumulée produite après 35 jours d'incubation. La méthode est décrite en détail dans la littérature (Angelidaki et al., 2009 ; Hansen et al., 2004).

III.2.3. La spectroscopie proche infrarouge

Les déchets solides, broyés et lyophilisés, ont été analysés en réflectance (R), sur la gamme de nombre d'onde du proche infrarouge 10000 - 4000 cm^{-1} , avec une résolution de 4 cm^{-1} . Les unités de la gamme ont ensuite été transformées en longueurs d'onde, 1000-2500 nm, avec un intervalle de 1 nm. Le spectrophotomètre est un BUCHI NIRFlex N-500, équipé du module « solide », avec l'accessoire boîte de Pétri. L'absorbance ($\log(1/R)$) a ensuite été calculée à partir des spectres en réflectance. Chaque échantillon a été divisé en trois sous-échantillons « scannés » indépendamment afin d'obtenir des triplicatas de spectres par échantillon. Les triplicatas ont servi à évaluer l'homogénéité (ou l'hétérogénéité) des mesures spectrales. Chaque spectre est la moyenne de 100 scans, afin de réduire le bruit de fond lié à la mesure spectrale. Le temps nécessaire pour effectuer ces 100 scans fut d'environ 2 minutes. Enfin, 10 spectres d'un ensemble de 3 échantillons différents ont été acquis et utilisés pour évaluer la répétabilité du spectrophotomètre.

III.2.4. Traitement des données et paramètres statistiques

L'analyse des données a été réalisée à l'aide de la PLS Toolbox d'Eigenvector Research, INC, utilisée avec le logiciel de programmation Matlab (© Mathworks). La méthode PLS a été utilisée pour construire le modèle.

III.2.4.1. Prétraitements chimiométriques et choix du modèle

Les prétraitements chimiométriques suivants ont été testés afin de réduire les effets de diffusion et de décalage de la ligne de base induits par la nature particulière des déchets (Zeaiter et al., 2005) : Standard Normal Variate (SNV) (Barnes et al., 1989), Detrend (Barnes et al., 1989) et une sélection de variables. Le modèle final d'étalonnage a été choisi en prenant en compte (i) la valeur de l'Erreur Standard de Validation Croisée (Standard Error of Cross-Validation, SECV) calculée sur l'ensemble d'étalonnage, (ii) le nombre de variables latentes nécessaires au modèle (Bertrand et al., 2006).

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

avec \hat{y}_i , la valeur de BMP prédite de l'échantillon i de l'ensemble d'étalonnage et y_i , la valeur de BMP mesurée de l'échantillon i de l'ensemble d'étalonnage, n_e , le nombre d'échantillons dans l'ensemble d'étalonnage.

La validation croisée dite « leave-one-out », par lot de triplicat a été utilisée afin d'évaluer le nombre optimal de variables latentes (Bertrand et al., 2006). Son fonctionnement est le suivant : un échantillon est retiré de l'ensemble d'étalonnage puis un étalonnage est alors créé avec les échantillons restants. Cet étalonnage est utilisé pour prédire l'échantillon indépendant écarté. Cette procédure est répétée pour chacun des échantillons. Le SECV est alors calculé.

Une Analyse en Composante Principale (ACP) a été mise en œuvre. Le principe de l'ACP, comme outil d'analyse en spectroscopie, est de remplacer les variables d'origines (absorbances), par de nouvelles variables (composantes principales), moins nombreuses. Les composantes principales sont des combinaisons linéaires des variables d'origines construites dans le but d'avoir une variance maximale et d'être orthogonales deux à deux. Les premières composantes principales contiennent la quasi-totalité de l'information. En général, pas plus de 20 composantes principales sont utilisées lors d'une ACP. La combinaison linéaire se réalise à l'aide de vecteurs propres, qui ont les mêmes dimensions qu'un spectre et peuvent être interprétés comme un spectre.

Une représentation schématique de l'ACP est présentée Figure 17.

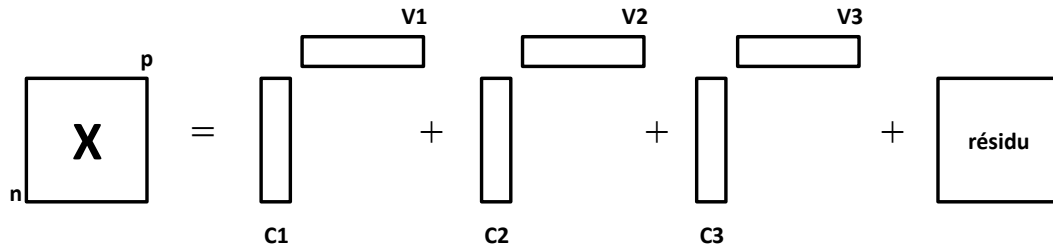


Figure 17 : Représentation schématique de l'ACP

X : matrice de spectres de dimension (n x p), Cn : Composante principale de taille (n x 1), Vn : vecteurs propres de dimension (1 x p).

III.2.4.2. Test du modèle

Le modèle établi est testé dans un premier temps avec l'ensemble de test, puis validé avec l'ensemble de validation indépendante. La qualité du modèle est estimée par l'erreur quadratique moyenne de prédiction (Root Mean Square Error of Prediction, RMSEP_t et RMSEP_{iv}) et du Biais (Biais_t et Biais_{iv}) (estimant l'erreur systématique liée à l'ensemble de test ou de validation indépendante),

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

avec \hat{y}_i , la valeur de prédiction de l'échantillon i de l'ensemble de test ou de l'ensemble de validation indépendante, y_i , la valeur BMP mesurée de l'échantillon i de l'ensemble de test ou de l'ensemble de validation indépendante, n , le nombre d'échantillons soit de l'ensemble de test (n_t), ou de l'ensemble de validation indépendante (n_{iv}).

III.2.4.3. Calcul des écarts types de répétabilité et de reproductibilité

Les **écarts type de répétabilité** (ETr) des deux méthodes ont été calculés selon les équations suivantes :

$$ETr_{PIR} = \sqrt{\frac{1}{n_t} \sum \text{Variance}(\text{triplicata d'échantillon } n)} \quad ETr_{BMP_{test}} = \sqrt{\frac{1}{n_d} \sum \text{Variance}(\text{duplicata d'échantillon } n)}$$

avec n_t , le nombre de répétitions d'échantillons (trois dans le cas des spectres) et n_d le nombre de répétitions d'échantillons (deux dans le cas des tests BMP).

L'écart-type de reproductibilité du test BMP ($ETR_{testBMP}$) a été calculé afin de comparer le test BMP avec la mesure PIR. L'erreur de reproductibilité du test BMP a été calculée avec 13 répétitions de test BMP avec de l'éthanol comme substrat (pureté : 96% ; densité : 0,81). L'éthanol est couramment utilisé comme étalon positif dans le test BMP. Les mesures ont été faites par deux personnes différentes, à des temps différents et en utilisant des inocula différents. $ETR_{testBMP}$ est défini comme étant l'écart type des 13 répétitions faites sur l'éthanol.

L'écart-type de reproductibilité étant calculé sur de l'éthanol, l'erreur d'échantillonnage est nulle, pour cela nous considérerons que l'écart-type de reproductibilité du test BMP ainsi calculé ne considère que le test BMP en lui-même. Dans un contexte d'échantillonnage, cette erreur de reproductibilité serait plus importante.

L'écart-type de reproductibilité du modèle NIR (ETR_{NIR}) a également été évalué. Un échantillon (échantillon n°44 : 231 ml $CH_4.g^{-1}$ MV) a été « scanné » 18 fois, par deux personnes différentes, pendant plusieurs jours, avec une prise d'échantillon différente pour chaque mesure. L'écart-type de reproductibilité du modèle proche infrarouge est alors défini comme étant l'écart type des 18 prédictions sur cet échantillon.

III.2.4.4. Calcul de l'incertitude des deux méthodes

III.2.4.4.1. Incertitude du test BMP

Le calcul de l'incertitude peut se faire soit à l'aide de l'écart type de reproductibilité, soit de l'écart type de répétabilité. Afin de prendre en compte l'erreur d'échantillonnage, l'écart-type de répétabilité sera utilisé.

Avec l'hypothèse d'une distribution des erreurs suivant la loi normale, les limites d'incertitudes pour une nouvelle mesure sont obtenues en multipliant l' $ETR_{testBMP}$ par un facteur d'élargissement, le t de Student ($t_{1-\alpha, dl}$) où $1-\alpha$ est le niveau de confiance (ici choisi égal à 0,95) et dl , le degré de liberté (le nombre de répétition : 2), le t de Student est donc $t_{95\%, 13} = 2.16$.

Les limites de l'intervalle de confiance du test BMP sont alors $ETR_{testBMP} \times t_{95\%, 13}$ (Widmer-Girod et al. 2004), avec $ETR_{testBMP}$, l'écart type de reproductibilité du test BMP, $t_{95\%, 13}$, le t de Student (2.16).

III.2.4.4.2. Incertitude du modèle proche infrarouge

Les limites de l'intervalle de confiance d'une mesure PIR ont également été calculées, suivant la même méthode que précédemment. Un facteur d'élargissement, $t_{1-\alpha, N-dl}$ de Student est calculé avec un niveau de confiance $1-\alpha=0,95$, N le nombre d'échantillons de l'ensemble d'étalonnage et dl , le nombre de degrés de liberté. Cependant, dans le cas d'une PLS, dl n'est pas connu directement et se situe entre le nombre de variables latentes et le nombre de longueurs d'onde (Zhang et al., 2009).

Van der Voet (1999) propose de calculer un pseudo degré de liberté, pdf tel que le présente la formule suivante,

$$pdf = n \cdot \left(1 - \frac{SEC}{SECV}\right)$$

avec n , le nombre d'échantillon dans l'ensemble d'étalonnage (51), SEC , l'erreur standard d'étalonnage (Standard Error of Calibration) et le $SECV$ (précédemment calculé).

Le SEC est calculé de la même façon que le $RMSEP$, mais avec l'ensemble d'étalonnage,

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

avec \hat{y}_i , la valeur de prédiction de l'échantillon i de l'ensemble d'étalonnage, y_i , la valeur BMP mesurée de l'échantillon i de l'ensemble d'étalonnage, n , le nombre d'échantillons de l'ensemble d'étalonnage.

La différence entre le SEC et le $SECV$ réside dans le principe de la validation croisée, où un échantillon est extrait pour être prédit avec le modèle construit à l'aide des échantillons restant. Pour le calcul du SEC , tous les échantillons sont inclus dans le modèle pour prédire les mêmes échantillons un à un.

Dans notre cas le pdf s'élève à 11. N étant égal à 51, $N-pdf$ est de 40. Le t -Student est alors égal à $t_{95\%,40} = 1.96$. C'est valeur est très proche de la valeur communément admis dans la communauté de la spectroscopie proche infrarouge de 2. Dans le cas de la PLS, c'est le $SECV$ qui est utilisé pour le calcul des limites d'incertitudes.

La limite d'incertitude de la méthode SPIR est alors $SECV \times 1.96$

Cela signifie donc que la valeur vraie d'un nouvel échantillon se trouvera dans l'intervalle de confiance autour de la valeur prédite Y_p , *i.e.* $[Y_p \pm SECV \times 1.96]$ avec une probabilité de 0,95.

III.3. Résultats et discussion

III.3.1. Choix du meilleur prétraitement chimiométrique

III.3.1.1. Présentation des valeurs BMP des échantillons

Les moyennes, min/max et écarts types des ensembles d'étalonnage et de validation sont présentés Tableau 6.

	Gamme	Moyenne (Ecart type)
Ensemble d'étalonnage (n = 51)	23 – 400	234 (66)
Ensemble de test (n = 23)	87 – 322	227 (57)
Ensemble de validation indépendante (n = 28)	69 - 380	189 (80)

Tableau 6 : Gamme, moyenne et écart type des valeurs BMP des ensembles d'étalonnage, de test et de validation indépendante (toutes les valeurs sont en $\text{ml.CH}_4.\text{g}^{-1} \text{ MV}$).

Les valeurs BMP des échantillons testés sont présentées Figure 17.

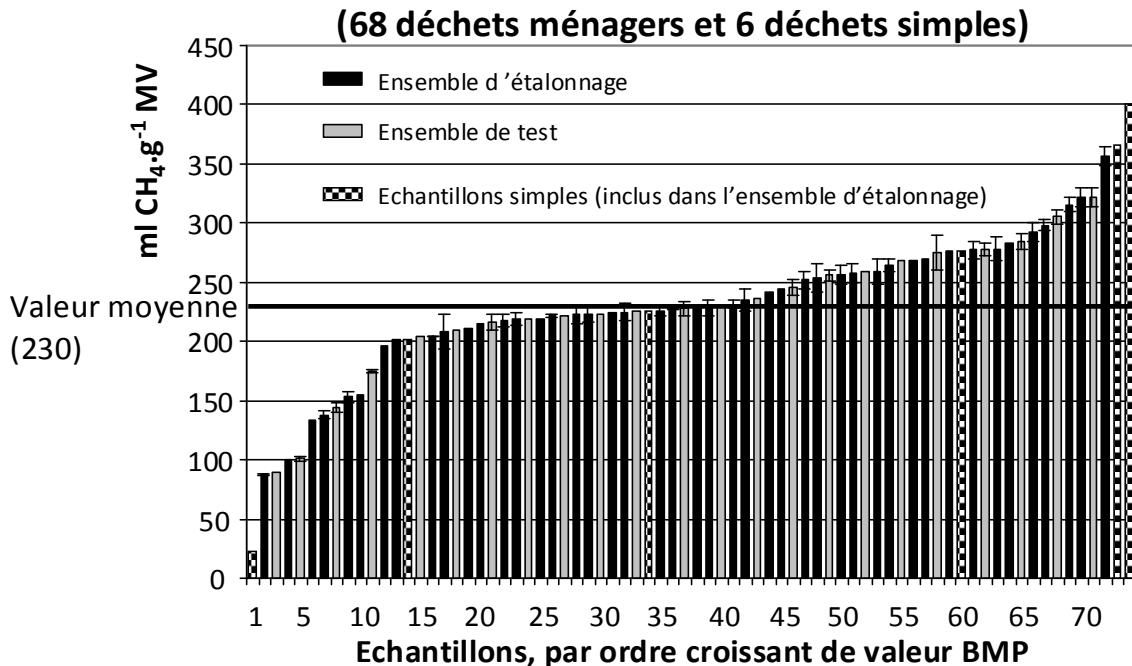


Figure 17 : Ensemble des valeurs BMP des 74 échantillons utilisés pour l'étalonnage et le test.

La gamme de valeur est étendue avec un minimum à 23 et un maximum à 400 ml CH₄.g⁻¹ MV, mais la plupart des valeurs sont dans la gamme 200-300 ml CH₄.g⁻¹ MV. Afin d'éviter une extrapolation du modèle, les valeurs BMP de l'ensemble de test ont été incluses dans la gamme de l'ensemble d'étalonnage.

III.3.1.2. Observation des spectres en absorbance

La Figure 18 présente une sélection de spectres des déchets ménagers (20 échantillons pris le long de la gamme des valeurs BMP) et des déchets « simples » lyophilisés.

L'observation des spectres des ordures ménagères (Figure 18a) montre une homogénéité dans leur forme, mais sans pics clairement visibles. Au contraire, les spectres des déchets simples - à l'exception des cartons-, présentent des pics mieux définis. Cette différence provient essentiellement d'effets multiplicatifs différents.

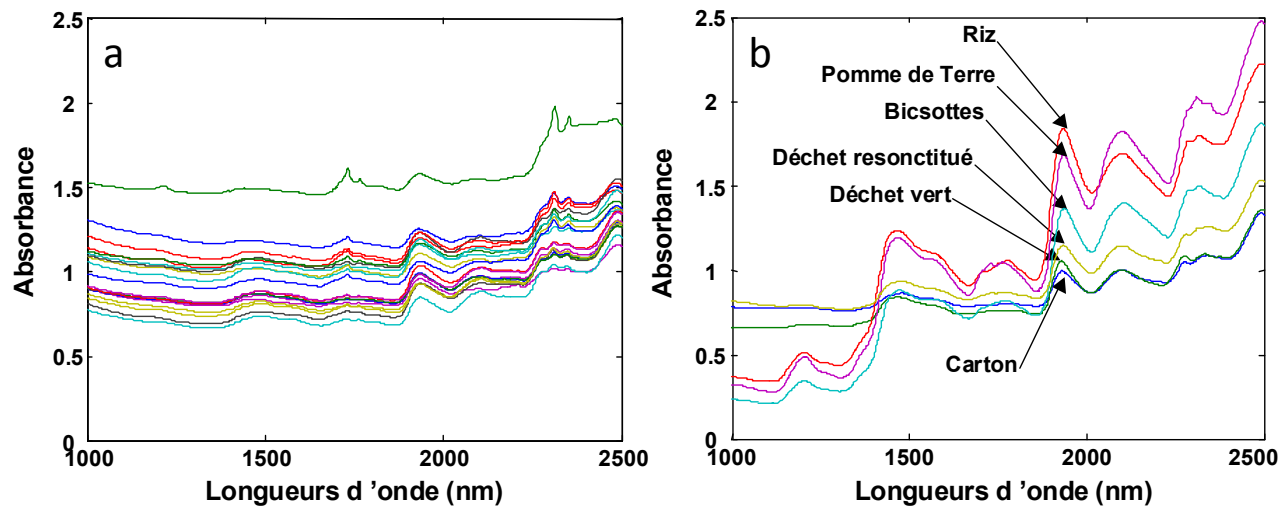


Figure 18 : Spectres des ordures ménagères (a) et des déchets simples (b) en absorbance.

Les variations de ligne de base proviennent essentiellement des effets de diffusion induits par les interactions entre le rayonnement proche infrarouge et les particules solides des échantillons (Zeaiter et al., 2005 ; Bertrand et al., 2006). En spectroscopie proche infrarouge sur des milieux diffusants, on rencontre 2 lignes de bases successives sur le spectre (Bellon Maurel et Dubois, 2009), qui correspondent à l'évolution de type « courbe en V » du coefficient de diffusion : la première montrant une diminution du coefficient jusqu'à un minimum situé aux alentours de

1300 nm, puis la seconde, montrant un accroissement du coefficient de diffusion à partir de 1300 nm jusqu'à la fin du domaine de proche infrarouge.

La loi de Beer Lambert est un modèle généralement accepté pour traduire la relation entre la concentration d'un composé et l'absorbance. La modification de l'absorbance suite aux effets de diffusion peut être traduite dans l'expression de la loi de Beer Lambert comme illustré Figure 19.

Deux effets de la diffusion sont possibles (Bertrand et al., 2006), (Figure 19b), l'augmentation du trajet optique ℓ , qui conduit à un effet multiplicatif indépendant de la longueur d'onde et (Figure 19c), la perte de photons, dépendant de la longueur d'onde, qui augmente la valeur de la ligne de base par un effet additif. Ces effets de diffusion sont fonction de la taille des particules solides analysées, plus grande est la taille des particules, plus faible est la diffusion. Il faut donc s'attendre à une diminution du niveau et de l'amplitude des spectres d'absorbance lorsque la taille des particules augmente (Bertrand et al., 2006). Les déchets ménagers ont un aspect plus fibreux et cotonneux que particulaire, ce qui implique une taille globalement plus importante que de la poudre, et donc une diffusion moins forte, pouvant expliquer la faible présence de pics clairement visibles par rapport aux déchets simples.

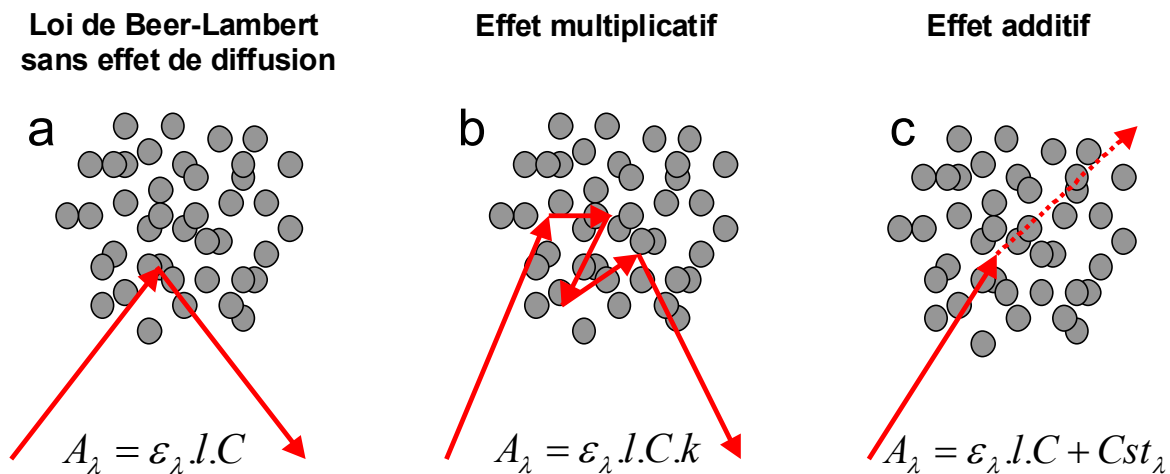


Figure 19 : Illustration des différentes interactions possibles entre le rayonnement proche infrarouge et les particules solides, induisant des effets multiplicatifs et additifs.

a : sans phénomène de diffusion. b : phénomène de diffusion induisant des effets multiplicatifs (augmentation de la longueur du trajet optique, ℓ). c : phénomène de diffusion induisant des effets additifs (augmentation de l'absorbance provoquée par la perte de photons, qui ne seront alors pas détectés par le capteur photosensible). A_λ : absorbance dépendante de la longueur d'onde (sans unité) ; ε : coefficient d'extinction molaire ($L \cdot mol^{-1} \cdot cm^{-1}$)

dépendant de la longueur d'onde ; ℓ : longueur du trajet optique (cm) ; C : concentration (mol.L^{-1}) ; k : constante multiplicative indépendante de la longueur d'onde ; $C_{st\lambda}$: constante additive dépendante de la longueur d'onde.

III.3.1.3. Observation des spectres après application des prétraitements

Afin de pallier ces effets de diffusion et décalages de ligne de base, des prétraitements chimiométriques ont été utilisés. L'utilisation des prétraitements SNV et Detrend ont diminué les effets de diffusion et ont favorisé l'homogénéisation de l'ensemble des spectres. Peu de zones apparaissent alors différentes entre les deux ensembles de déchets (Figure 20). Les pics visibles vers 1730/1765 nm et 2310/2350 nm, indiqués par les flèches, dans les déchets ménagers (Figure 20a) sont moins visibles pour les déchets simples (Figure 20b). Dans ces régions de longueurs d'onde, c'est le groupement $-\text{CH}_2$ qui absorbe. Ce groupement est souvent rencontré dans les lipides et les composés plastiques. Or dans les déchets simples, il y a peu de lipides et pas du tout de composés plastiques. Le pic à 2100 nm est également plus prononcé pour les déchets simples, signe d'une plus grande quantité de glucides.

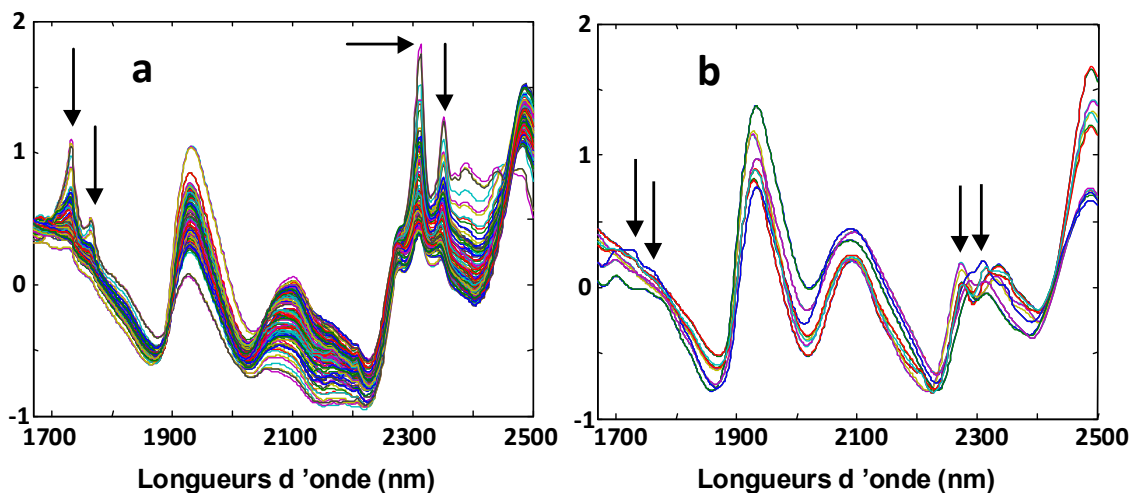


Figure 20 : Spectres des ordures ménagères (a) et des déchets simples (b) après application des prétraitements SNV et Detrend.

Les flèches indiquent les pics qui différencient distinctement les déchets ménagers des déchets simples.

L'observation des composantes principales de l'ACP (Figure 21) permet de visualiser des différences dans l'ensemble des spectres avant et après les prétraitements.

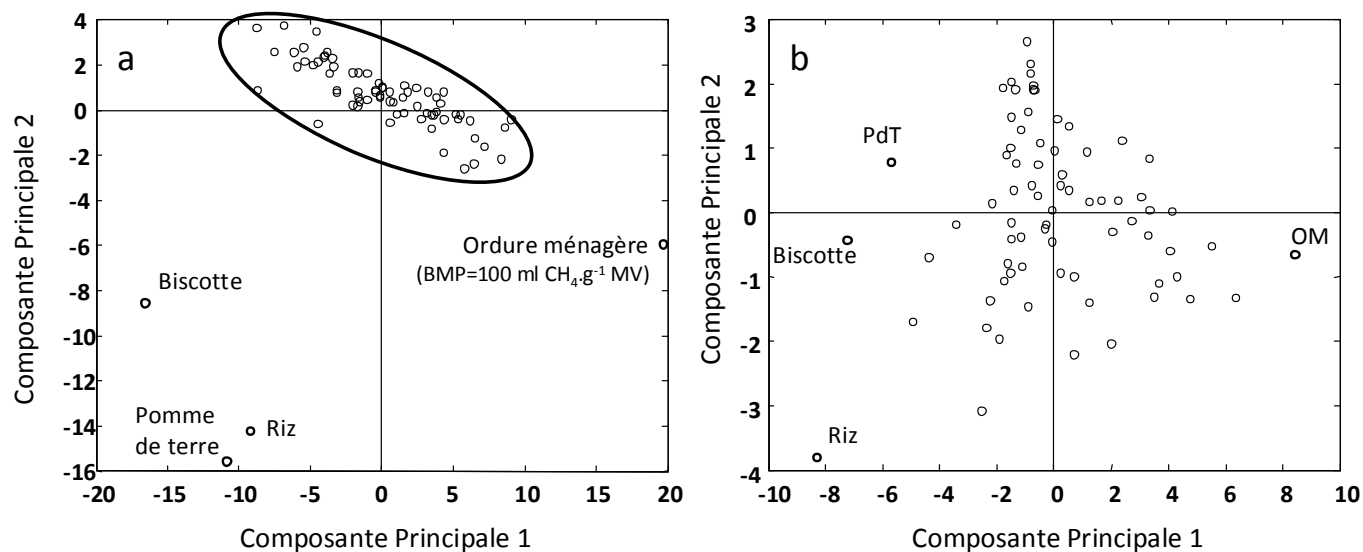


Figure 21 : Analyse en composante principale sur l'ensemble des spectres avant (a) et après prétraitements (b).

Ceux qui ressortent du lot sont -parmi les déchets simples- le riz, la biscotte et la pomme de terre (Figure 21a). Un des déchets ménagers ressort également du lot de 68 échantillons. Il est facilement identifiable à la Figure 18a, avec des valeurs d'absorbance globalement plus élevées, supérieures à 1,5 unité d'absorbance.

Autre point intéressant, la position des spectres sur l'axe 1 de la Figure 21a relate bien l'effet de diffusion et les différences observées sur les spectres Figure 18. Comme déjà explicité, les valeurs d'absorbances des spectres bruts sont liées à la diffusion : plus la diffusion est élevée, plus les valeurs d'absorbances seront fortes. On observe que les 3 déchets simples, en bas à gauche de la Figure 21a, sont les seuls à avoir des valeurs d'absorbances inférieures à 0,5 (Figure 17b). A l'opposé, l'échantillon d'ordures ménagères complètement à droite (Figure 21a) est le seul à avoir des valeurs d'absorbances supérieures à 1,5 (Figure 18a). Cette relation est exprimée Figure 22, où un Rr^2 de 0,97 est obtenu entre les valeurs de la première composante principale et l'absorbance à 1191 nm. Cette relation reste constante sur toute la gamme 1000 – 1400 nm.

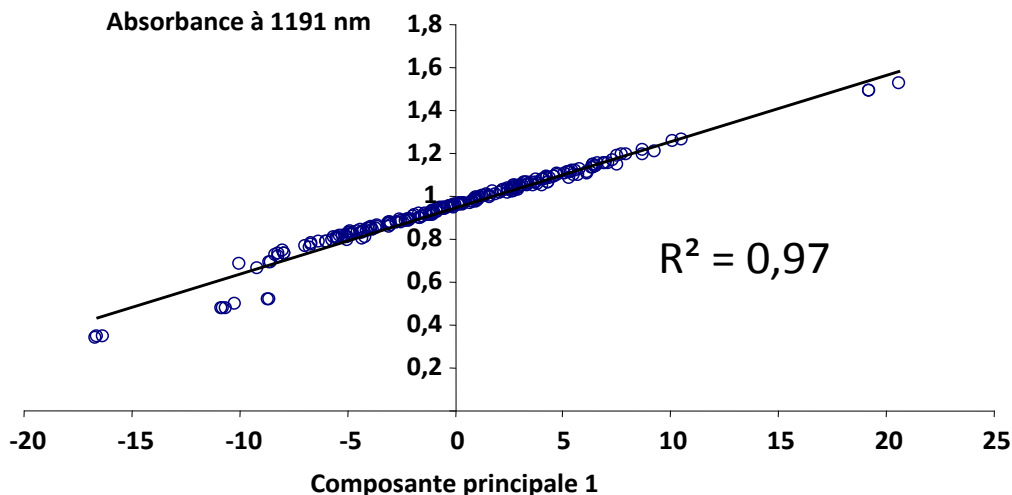


Figure 22 : Relation entre les valeurs de la première composante principale et l'absorbance à 1191 nm.

Cela montre que la première composante principale exprime fortement les phénomènes de diffusion. L'utilisation des prétraitements a permis de diminuer les effets de diffusion et de restreindre la variabilité entre les spectres à la composition chimique (Figure 21b).

III.3.2. Discussion autour du modèle sélectionné

III.3.2.1. Modèle construit sur le lot d'étalonnage de 51 échantillons

Le meilleur modèle a donc été obtenu avec les prétraitements SNV et Detrend appliqués sur une sélection de variables (1668 – 2500 nm) (Tableau 7). Le modèle sans prétraitement (noté Abs) et avec uniquement la sélection de variables ou celui avec dérivée première donnent des résultats proches et également satisfaisants. L'utilisation des prétraitements a également permis d'augmenter la robustesse des modèle PLS en diminuant notamment le nombre de variables latentes.

A	Avec déchets simples	Facteurs PLS ^a	R ² étalonnage	SECV mL CH ₄ .g ⁻¹ MV	R ² validation	RMSEP _t mL CH ₄ .g ⁻¹ MV	Biais mL CH ₄ .g ⁻¹ MV	RPD (SD/RMSEP _t)
Variables 1000-2500 nm	Abs	10	0.58	44	0.59	37	1	1.8
	SNV	8	0.68	38	0.74	29	0	2.3
	SNV – Dt	7	0.69	37	0.69	32	2	2.1
	SNV-(45,3,1)	6	0.65	39	0.66	33	3	2
	SNV-(13,3,2)	7	0.49	48	0.56	38	0	1.7
Variables 1668-2500 nm	Abs	9	0.77	31	0.68	32	3	2.1
	SNV	7	0.77	32	0.71	30	3	2.2
	SNV – Dt	7	0.79	31	0.76	28	4	2.4
	SNV-(17,3,1)	6	0.66	39	0.68	32	0	2.1
	SNV-(23,3,2)	6	0.62	41	0.51	40	1	1.7

B	Sans déchets simples	Facteurs PLS ^a	R ² étalonnage	SECV mL CH ₄ .g ⁻¹ MV	R ² validation	RMSEP _t mL CH ₄ .g ⁻¹ MV	Biais mL CH ₄ .g ⁻¹ MV	RPD (SD/RMSEP _t)
Variables 1000-2500 nm	Abs	8	0.77	26	0.62	38	2	1.4
	SNV	10	0.79	25	0.77	27	1	2
	SNV – Dt	8	0.77	26	0.78	27	1	2
	SNV-(51,3,1)	9	0.8	24	0.77	28	1	1.9
	SNV-(19,3,2)	8	0.7	30	0.72	30	1	1.8
Variables 1668-2500 nm	Abs	8	0.77	26	0.7	32	3	1.7
	SNV	8	0.76	26	0.76	28	4	1.9
	SNV – Dt	9	0.79	25	0.7	31	1	1.7
	SNV-(51,3,1)	9	0.8	25	0.78	27	3	2
	SNV-(29,3,2)	9	0.72	29	0.74	29	1	1.9

Tableau 7 : Résultats des différents étalonnages et validation testées.

a: nombre de variables latente utilisées; Abs: absorbance; Dt: Detrend; (45,3,1) sont les paramètres utilisés pour les dérivées avec respectivement la taille de la fenêtre, le degré du polynôme et l'ordre de la dérivée.

La Figure 23 présente la prédiction des ensembles d'étalonnage et de test par le modèle sélectionné (sélection de variables 1668-2500 nm, SNV et Detrend). Les valeurs de SECV et RMSEP_t sont respectivement de 31 et 28 ml CH₄.g⁻¹ MV pour les ensembles d'étalonnage et de test avec $r^2_{\text{étalonnage}} = 0,79$ et $r^2_{\text{test}} = 0,76$. Pour le modèle incluant les déchets "simples" (Tableau

7A), les résultats en étalonnage sont meilleurs avec que sans sélection de variables. Cela montre l'intérêt de la sélection de variables, aidant à réduire l'effet matrice entre les échantillons d'ordures ménagères et les déchets « simples ». En ne prenant que les déchets ménagers dans les modèles (Tableau 7B), les résultats en étalonnage diffèrent peu avec ou sans sélection de variables. La sélection de variables semble améliorer la robustesse des modèles établis par la diminution de la variabilité entre échantillons.

Certains des déchets "simples" (déchets verts, déchet reconstitué, pomme de terre), indiqués par les flèches (Figure 23), sont moins bien prédits. Ceci est dû au fait qu'il n'y a qu'un seul échantillon par type de déchet simple. De plus, le déchet vert et la pomme de terre sont aux extrêmes des valeurs BMP. Les échantillons extrêmes sont souvent mal prédit, car très différents des échantillons moyens. Lorsque ces déchets simples sont enlevés pour la validation croisée, l'erreur peut être importante et dégrade la valeur du SECV. Le modèle, majoritairement basé sur les déchets ménagers, est donc plus juste pour la prédiction de déchets ménagers que de déchets « simples ».

Sur la Figure 23, La ligne représente la droite de relation 1 :1. L'ensemble d'étalonnage est représenté par des cercles et l'ensemble de validation par des carrés. Pour chaque point, les barres horizontales sont les écarts-types de répétabilité du test BMP pour les duplicatas et les barres verticales sont les écarts-types de répétabilité de la mesure des triplicatas. La courbe en pointillé représente l'effet « Dunne » pour lequel les faibles valeurs sont surestimées et les fortes valeurs sont sous-estimées. Cet effet se produit car le modèle est « calé » sur la moyenne et donc la précision est meilleure pour des valeurs proches de la moyenne. Cet effet augmente lorsque, pour le paramètre prédit, l'ensemble d'étalonnage a une distribution gaussienne (Williams et Norris, 2001). Bien que le lot d'étalonnage ne soit pas formellement gaussien, il montre une distribution assez symétrique avec de nombreux échantillons proches de la valeur moyenne (230 ml CH₄.g-1 MV), ce qui le rend sensible à l'effet Dunne (afin d'éviter l'effet Dunne, il est donc recommandé de disposer d'un ensemble d'étalonnage distribué de façon uniforme).

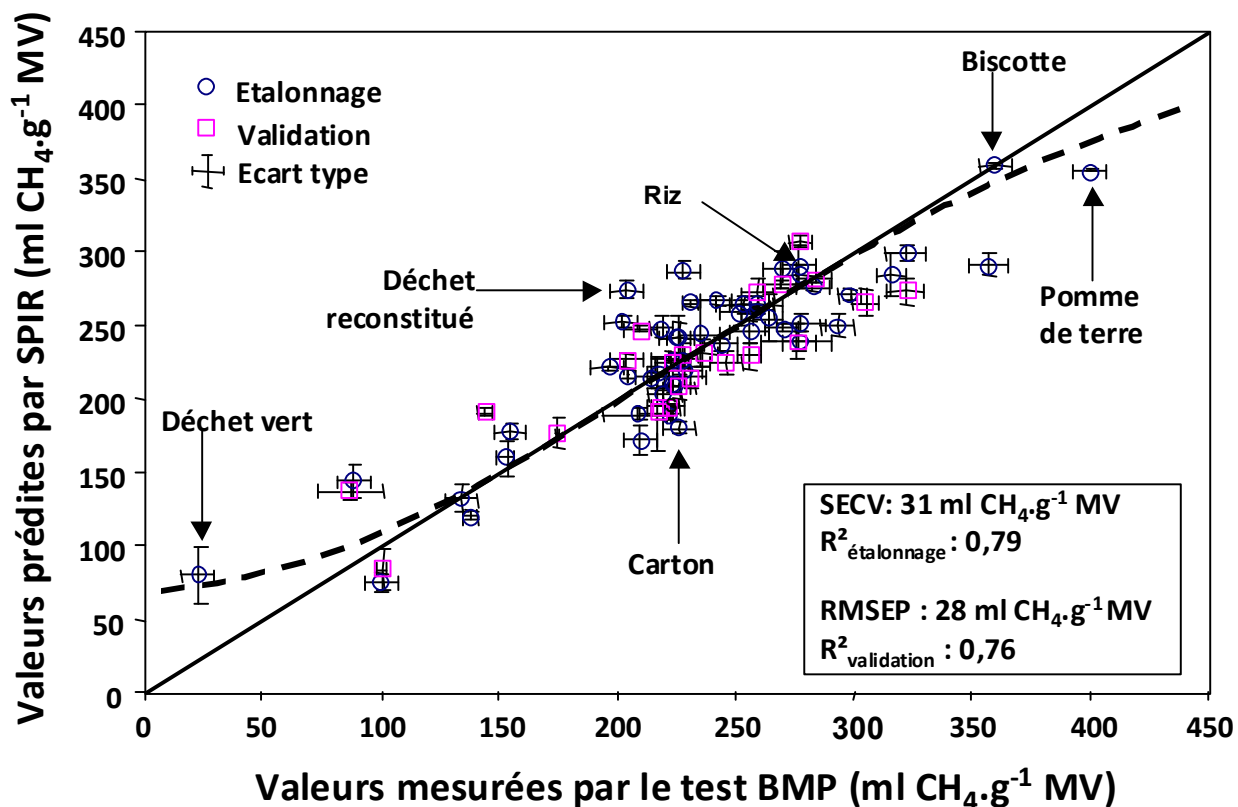


Figure 23 : Valeurs BMP prédites versus valeurs BMP mesurées.

III.3.2.2. Modèle construit avec le lot d'étalonnage construit sans déchets simples.

La régression PLS a également été réalisée sans tenir compte des déchets simples dans la base d'étalonnage (Tableau 7B). Les résultats d'étalonnage sont meilleurs, dû à l'homogénéité de la base d'étalonnage. Pour les résultats de validation, la différence avec les résultats d'étalonnage est très faible. Cela montre que les modèles construits avec ou sans les déchets simples sont identiques.

Afin d'améliorer les résultats d'étalonnage, des prétraitements chimiométriques ont été appliqués aux spectres bruts pour supprimer les effets de diffusion. Le meilleur modèle a été obtenu après la sélection de variables (1668 – 2500 nm) et de l'utilisation des prétraitements SNV et Detrend. Le modèle est cependant basé essentiellement sur des déchets ménagers. La validation de ce modèle par un ensemble d'échantillons indépendants est nécessaire afin de confirmer ces résultats.

III.3.3. Validation avec les échantillons indépendants

Une validation a été effectuée sur des échantillons indépendants, différents des ensembles d'étalonnage et de test (différentes origines, différentes natures...). Les résultats de la prédiction de ces échantillons indépendants sont présentés Figure 24.

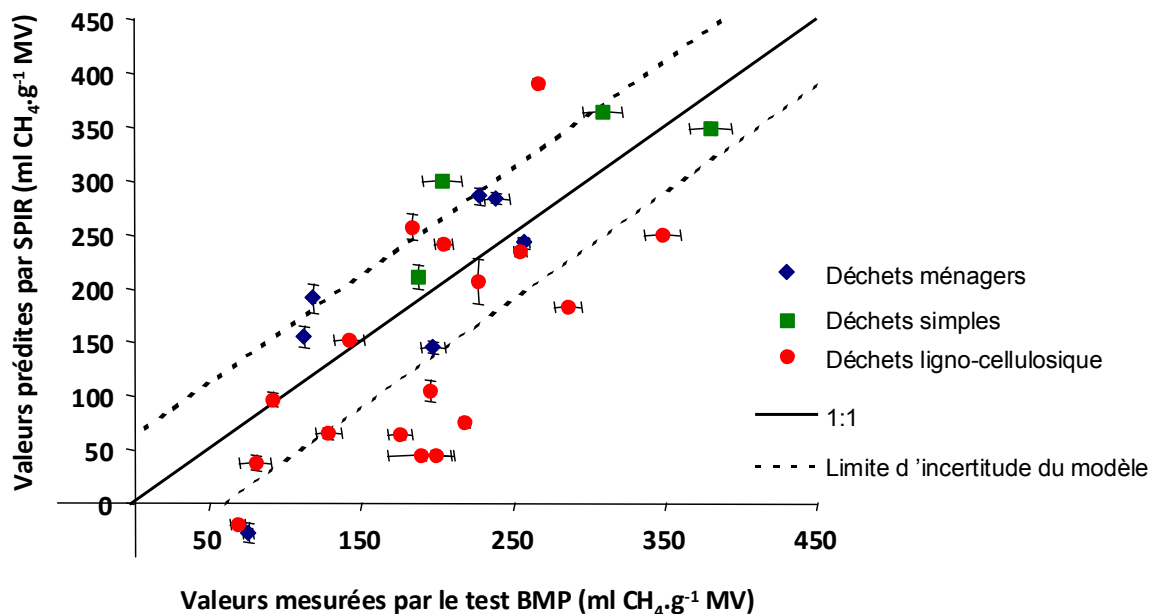


Figure 24 : Régression linéaire de l'ensemble de validation externe

Le $RMSEP_{iv}$ et le R_{iv}^2 sont respectivement égaux à $80 \text{ ml CH}_4 \cdot \text{g}^{-1}$ ($\text{biais}_{iv} = 23 \text{ ml CH}_4 \cdot \text{g}^{-1} \text{ MV}$) et $0,53$. Le $RMSEP_{iv}$ est important mais proche de la limite d'incertitude du modèle ($61 \text{ ml CH}_4 \cdot \text{g}^{-1} \text{ MV}$). Environ 50% des échantillons sont inclus dans les limites d'incertitude. Cinq déchets ménagers (sur 6 au total) et 3 déchets simples (sur 4 au total) sont bien prédits. Mais seulement 3 (sur 17 au total) des déchets ligno-cellulosiques sont correctement prédits. Ainsi, le modèle établi est satisfaisant pour prédire la valeur BMP d'ordures ménagères. Les déchets ligno-cellulosiques sont mal représentés par le lot d'étalonnage. En effet, ils sont trop différents au niveau de leurs compositions et donc au niveau spectral par rapport aux déchets ménagers qui composent l'essentiel de l'ensemble d'étalonnage.

Ce point constitue une limite de la méthode. La matrice déchet étant très complexe, et pouvant avoir une hétérogénéité importante, l'élaboration d'un modèle représentant cette variabilité et

pouvant prédire correctement le BMP est complexe. Deux voies de réponse semblent se dessiner. La première voie serait d'utiliser autant de modèle de prédiction que de types de déchet (*i.e.* même genre d'échantillon mais d'origines différentes), comme par exemple une base dédiée à la prédiction du potentiel méthane de déchets carnés. L'avantage d'une telle méthode est d'obtenir, à priori, des prédictions précises. La seconde voie, est d'utiliser un seul modèle d'étalonnage « global », à partir d'échantillons hétérogènes (différents types d'échantillons, différentes origines). L'intérêt de cette dernière approche est la simplicité d'utilisation, avec un seul modèle, robuste, mais à priori moins précis. Ces deux voies seront discutées en V.

III.3.4. Comparaison des résultats avec la méthode basée sur la caractérisation chimique

Une des méthodes les plus étudiées afin de réduire le temps pour prédire la valeur BMP, est l'utilisation de la caractérisation chimique des déchets et l'utilisation de modèle de prédiction basé sur cette caractérisation. En 2007, Nallathambi Gunaseelan se base sur des déchets de fruits et légumes pour construire un modèle d'étalonnage homogène de prédiction de la valeur BMP. La validation est réalisée sur le même type de déchet. Il obtient des valeurs de RMSEP = 91 ml CH₄.g⁻¹ MV (biais = 54 ml CH₄.g⁻¹ MV). En 2009, Nallathambi Gunaseelan (2009) utilise un nouveau modèle, basé sur du Jatropha, et le valide sur des échantillons de mûriers et sur des échantillons d'études antérieures (Nallathambi Gunaseelan, 2007 ; Buffière et al., 2006), donc différent du Jatropha. Il obtient de meilleurs résultats, RMSEP = 66 ml CH₄.g⁻¹ MV (biais = 5 ml CH₄.g⁻¹ MV), notamment et contrairement au modèle précédent (Nallathambi Gunaseelan, 2007) du fait de la prise en compte de la teneur en lipide dans le modèle. Egalement, Schievano et al. (2009) réalise un modèle à partir d'échantillons hétérogènes (épluchures de légumes, boues, déchets liquides, déchets ménagers) et obtient des résultats moins probants, RMSEP = 140 ml CH₄.g⁻¹ MV (biais = -42 ml CH₄.g⁻¹ MV). Les bons résultats de Nallathambi Gunaseelan (2009) sont intéressants et montrent l'intérêt de la méthode. Cependant, les résultats obtenus par Schievano et al. (2009) dus certainement à l'importante variabilité de la base d'étalonnage. L'utilisation d'une base de déchets hétérogènes n'est pas supposée augmenter la précision de la prédiction mais seulement la robustesse. L'intérêt de la spectroscopie proche infrarouge réside essentiellement dans le gain de temps. Elle a également l'intérêt d'étudier globalement la matière

organique de différentes matrices en une analyse, alors que les méthodes biochimiques n'analysent que les molécules les plus importantes. Les dosages les plus utilisés ne sont pas forcément adaptés à tous les types de molécule. Par exemple, la méthode à l'Anthrone pour le dosage des glucides est sensible à moins de types de glucides que la méthode au phénol, alors qu'elle est plus utilisée dans la caractérisation des déchets (Piccolo et al., 1996 ; Feller et al., 1991). La méthode à l'Anthrone sous-estime la concentration de nombreux glucides tels que le galactose, le mannose, le xylose, l'arabinose (Feller et al., 1991). La même conclusion peut être trouvée dans la littérature pour la détermination des protéines entre la méthode BCA (Acide Bincichoninique) et la méthode de Lowry (Ras et al., 2008). La méthode au BCA surestime la concentration alors que celle de Lowry la sous-estime (Ras et al., 2008). De plus, l'erreur de prédiction pour un modèle basé sur la caractérisation chimique prend en compte de nombreuses sources d'erreurs. Ces erreurs vont provenir des différents dosages, de l'échantillonnage et de la préparation de l'échantillon pour le dosage. L'erreur dans la préparation va principalement venir de la phase d'extraction nécessaire au dosage, ceux-ci se réalisant toujours en solution. La spectroscopie proche infrarouge a l'avantage de ne nécessiter que peu de préparation, où seule l'étape d'échantillonnage peut induire une erreur de préparation.

III.3.5. Discussion sur la faisabilité de remplacer le test BMP par les mesures PIR

III.3.5.1. Performances métrologiques

Cette étude métrologie se basera essentiellement sur deux critères, la justesse et l'exactitude. Etant la méthode de référence, la mesure du BMP par le test biochimique est considérée comme "juste" suivant la définition métrologique de la justesse *i.e.* "étroitesse de l'accord entre la moyenne d'un nombre infini de valeurs mesurées répétées et une valeur de référence" (VIM, 2009). Dans notre cas, la SPIR a démontré sa justesse avec un biais très faible (4 ml CH₄.g⁻¹ MV). Il faut cependant relativiser ce résultat par le fait que les échantillons proviennent d'un même ensemble initial. Le biais, important obtenu avec l'ensemble de validation indépendant en atteste, biais_{vi} = 23 ml CH₄.g⁻¹ MV. Le RMSEP_t (28 ml CH₄.g⁻¹ MV) est très satisfaisant, surtout si nous prenons en compte la précision de la méthode de référence (33 ml CH₄.g⁻¹ MV), *i.e.*

l'écart type calculé sur une série répétée d'expérimentations (ici l'éthanol), comme préconisé par Faber (2004). Comme la précision de la méthode de référence ($SDR_{testBMP}$) est au alentour de 32 ml $CH_4.g^{-1}$ MV, nous pouvons conclure que la méthode PIR est aussi bonne que la méthode de référence.

Le $RMSEP_{iv}$ est comparativement élevé par rapport au SECV (31 ml $CH_4.g^{-1}$ MV). Cela est expliqué par le fait que l'ensemble de validation externe contient des échantillons, majoritairement des déchets lignocelulosique, très différents de ceux de l'ensemble d'étalonnage.

Les valeurs d'écarts types de répétabilité (E_{tr}) sont similaires, respectivement 7 et 10 ml $CH_4.g^{-1}$ MV pour le test BMP et la mesure SPIR. Ce résultat est plutôt satisfaisant puisque la mesure PIR prend en compte plusieurs sources d'erreur, tels que l'homogénéité des échantillons, la qualité du modèle, la répétabilité du spectrophotomètre PIR et la répétabilité de la mesure de référence, *i.e.* le test BMP.

Pour évaluer l'exactitude des mesures, *i.e.* "étroitesse de l'accord entre une valeur mesurée et une valeur vraie d'un mesurande" (VIM, 2009), nous pouvons comparer les intervalles de confiance obtenus avec les deux méthodes. Les intervalles de confiance du test BMP et de la mesure spectrale sont respectivement 69 ml $CH_4.g^{-1}$ MV and 61 ml $CH_4.g^{-1}$ MV. Cela démontre que la SPIR prédit avec une bonne limite d'incertitude la valeur BMP et cela plus rapidement que le test biochimique (quelques heures).

Le Tableau 8 résume les résultats du modèle. La moyenne des valeurs prédites de l'ensemble de test (223 ml $CH_4.g^{-1}$ MV) est proche de celle des valeurs de références (227 ml $CH_4.g^{-1}$ MV). La prédiction ne présente pas de biais, mais ce dernier résultat était attendu puisque l'ensemble de validation et d'étalonnage proviennent du même ensemble d'échantillons. Egalement, comme attendu, l'erreur de répétabilité du spectrophotomètre PIR est négligeable avec un écart-type autour de 1,7 ml $CH_4.g^{-1}$ MV.

	Prédiction par la SPIR	Mesure par le test BMP
Moyenne ($ml\ CH_4.g^{-1}\ MV$)	227	223
Ecart type de répétabilité (ETr) ($ml\ CH_4.g^{-1}\ MV$)	9	7 (calculé sur les ordures ménagères)
Ecart type de reproductibilité (ETR) ($ml\ CH_4.g^{-1}\ MV$)	8	32 (calculé sur l'éthanol)
Intervalle de confiance pour une nouvelle mesure BMP ($ml\ CH_4.g^{-1}\ MV$)	61 ($SECV \times 1.96$)	69 ($SDR_{BMP_{test}} \times 2.16$)
Ecart type de répétabilité du spectrophotomètre ($ml\ CH_4.g^{-1}\ MV$)	1.7	

Tableau 8 : Comparaison des performances entre le modèle SPIR et le test BMP.

Le rapport (Erreur de prédiction / Ecart type des échantillons d'étalonnage) (Ratio of Performance to Deviation RPD) a été calculé afin d'évaluer et comparer les différents étalonnages. Un RPD de 2,36 est obtenu avec le modèle sélectionné. Ce ratio est plutôt satisfaisant, (Williams et Sobering, 1993) mais justifie le besoin d'améliorer le modèle d'étalonnage en augmentant le nombre d'échantillons dans la base d'étalonnage (Williams et Norris, 2001). L'extension du modèle à d'autres types de déchets incluant les spectres PIR et les valeurs BMP associées permettrait de concevoir un modèle global applicable quel que soit le déchet. Cette méthode est applicable à tous types de déchets : déchets ménagers, boues de station d'épuration, déchets solides d'industries agro-alimentaire ou de déchets agricoles.

L'ensemble de ces résultats est prometteur. En prenant en compte la variabilité du test BMP (*i.e.* l'activité de l'inoculum et la complexité de la matrice des déchets ménagers), les résultats obtenus sont très satisfaisants d'un point de vue pratique. En effet, l'intérêt majeur de la méthode SPIR est sa simplicité de mise en œuvre et sa rapidité d'analyse. Le gain de temps par rapport à la méthode classique est très important, on passe d'une analyse classique effectuée après plus de 20 jours d'attente à une analyse SPIR en 2 jours. La durée de stockage est donc considérablement réduite.

III.3.5.2. Perspectives de mise en pratique

Dans notre étude, afin d'avoir des valeurs BMP représentatives, il a été nécessaire de broyer les échantillons pour augmenter leur homogénéité. Une lyophilisation a été réalisée afin d'obtenir des échantillons secs, nécessaire au broyage. L'eau ayant une absorbance très forte dans le proche

infrarouge, l'utilisation d'échantillons secs permet également d'obtenir un meilleur signal spectral. Malheureusement, la lyophilisation est relativement longue (plus de 24 h), et pourrait, en analyse de routine, être remplacée par un séchage à 60°C ou 90°C, utilisé plusieurs heures jusqu'à un taux de matière sèche acceptable.

L'étape du broyage pourrait également être supprimée. En utilisant une méthode de mesure appropriée, ne nécessitant pas de préparation d'échantillon, il pourrait être possible de remplacer la mesure de triplicata de spectres après broyage par de nombreuses (plusieurs dizaines) mesures du déchet. En réalisant plusieurs mesures spectrales du même échantillon, à différents endroits, par l'utilisation de fibre optique par exemple, il devrait être possible d'obtenir un ensemble représentatif du déchet analysé. Une autre solution serait d'utiliser des bancs de mesure, où l'on pourrait placer une quantité importante du déchet. Il serait alors analysé sur une grande surface et le spectre obtenu serait ainsi représentatif de l'ensemble. Enfin, l'imagerie Hyperspectrale, où chaque pixel est composé d'un spectre pourrait être également une solution afin d'avoir une image globale du déchet (Bellon Maurel et Dubois, 2009).

III.4. Conclusion

L'utilisation de la spectroscopie proche infrarouge est proposée pour la prédiction du potentiel méthane de déchets solides issus d'une usine de tri-compostage. La valeur BMP est bien prédite, avec un RPD de 2,36 et un RMSEP de 28 ml CH₄.g⁻¹ MV. L'erreur de répétabilité de la mesure PIR est proche de celle du test BMP, respectivement 9 et 7 ml CH₄.g⁻¹ MV. La limite d'incertitude de la prédiction par PIR est du même ordre de grandeur que le test biochimique, 61 et 65 ml CH₄.g⁻¹ MV respectivement.

L'aspect le plus novateur dans l'approche proposée pour la prédiction du BMP est principalement la réduction notable du temps d'analyse, passant de plusieurs semaines pour le test biochimique à moins de deux jours pour la spectroscopie proche infrarouge (incluant la préparation de l'échantillon). Cette avancée permettra une évaluation plus rapide de la biodégradabilité du déchet et une augmentation de la production de méthane dans les digesteurs anaérobies industriels : une connaissance rapide du potentiel méthane du déchet permettra d'améliorer les performances de production en co-digestion par de meilleurs mélanges. De plus, la gestion du

stock de déchets se fera plus rapidement et plus facilement. Le risque de dégrader un déchet faiblement biodégradable et donc de perdre de la biomasse sera diminué.

*IV.
Etude conjointe des b coefficients PLS et
des longueurs d'onde d'intérêt pour les
modèles de prédiction du BMP*

Dans ce chapitre, il sera question d'étudier les coefficients PLS obtenus par le modèle établi précédemment. Cette étude permettra de vérifier la pertinence du modèle. En effet, en comparant l'attribution spectrale des *b* coefficients PLS aux données de la littérature relatives à la fois à la composition chimique des déchets ménagers et aux pouvoirs méthanogènes des différents composés, le modèle pourra être validé d'un point de vue chimique. En parallèle, des informations sur la composition chimique des déchets étudiés pourront être obtenues.

Ce chapitre est basé sur l'article:

M. Lesteur, V. Bellon Maurel, J.M. Roger, E. Latrille, C. Gonzalez, G. Junqua, J.P. Steyer
Parallel study of the PLS *b* coefficients and the near infrared wavelengths from the PLS regression of the methane potential of municipal solid wastes. (en préparation)

IV.1. Introduction

Comme on l'a vu en introduction, le potentiel méthane dépend à la fois de la matière organique biodégradable et non biodégradable présente dans le déchet. Or la fraction fermentescible des déchets ménagers est complexe et composée de différentes catégories (résidus de cuisines, papiers, carton, plastiques, déchets vert...). La spectroscopie proche infrarouge analyse rapidement et de façon globale (quantitative et qualitative) la matière organique de matrices diverses. La caractérisation de la biodégradabilité des déchets ou la connaissance des composés qui sont ou non dégradés pendant la digestion anaérobie peuvent être déterminées par l'étude conjointe des spectres et des *b* coefficients du modèle obtenus par la régression PLS.

Depuis le développement de la spectrométrie proche infrarouge, de nombreux chercheurs ont créé des bases de spectres caractéristiques ou ont tenté d'élucider les bandes d'absorption présentes en NIR. Williams et Norris (2001) fournissent les plages de longueurs d'onde des harmoniques et des combinaisons des liaisons chimiques qui absorbent dans le proche infrarouge : une base de 328 spectres de molécules purs ou de produits alimentaires est proposée, aidant à l'observation de leurs pics majeurs. Osborne et Fearn (1986) énumèrent pour de nombreuses longueurs d'onde l'élément prédominant qui absorbe. Quelques auteurs utilisent les *b* coefficients PLS ou directement les spectres afin d'identifier les composés de la matière organique analysée. Kays et al. (2009) utilisent les *b* coefficients PLS afin de montrer les longueurs d'onde les plus pertinentes (2250 et 2076 nm) pour la prédiction des fibres solubles et insolubles de céréales.

Ben-Dor et al. (1997) utilisent les spectres d'absorbance et de dérivée seconde afin d'identifier potentiellement les composés majeurs rencontrés lors du processus de compostage. Ils assignent certaines longueurs d'onde à la cellulose (2068 nm), la lignine (1930 nm), les protéines (2193 nm), l'amidon (1582 nm), des tanins (2142 nm), l'acide humique (1929 nm), les lipides (2309 nm), le glucane et les pectines (2111 nm). Baillères et al., (2002) énumèrent les longueurs d'onde principales rencontrées dans les spectres proche infrarouge de l'eucalyptus et leurs attribuent des caractéristiques moléculaires.

L'objectif de ce chapitre est de montrer comment l'analyse conjointe des *b* coefficients PLS et des spectres apporte des informations complémentaires sur le processus (la biodégradabilité des molécules) et facilite l'identification des molécules qui se dégraderont.

Pour cela, les vecteurs de *b*-coefficients seront comparés aux bandes d'absorption des composés majeurs des déchets, mais également aux variables sélectionnées par des algorithmes de sélection de variables. Une bonne concordance entre les zones d'intérêt retenues par ces différentes approches (*b* coefficients, bandes d'absorption, variables sélectionnées) confirmerait la qualité du modèle.

Les algorithmes de sélection de variables peuvent également être mis en œuvre, par exemple les algorithmes génétiques (Leardi et al., 1998 ; Leardi, 2000), ou des méthodes plus récentes comme l'algorithme CovSel (Roger et al., 2010). Ils nous aideront dans l'identification des longueurs d'onde d'intérêt. En effet, ces algorithmes reposent sur le principe de sélectionner les variables les plus utiles aux modèles ou dont la covariance avec la variable à prédire est maximale. Ainsi, ces algorithmes vont sélectionner les longueurs d'onde les plus pertinentes expliquant le mieux le potentiel méthane des déchets. Ils permettront de valider les conclusions tirées de l'analyse conjointe des *b* coefficients PLS et des spectres.

Notre démarche consistera donc d'abord à étudier en parallèle le vecteur de *b* coefficients PLS avec la dérivée seconde des spectres. Ainsi les longueurs d'onde d'intérêt seront mises en évidence. L'utilisation d'une table de corrélation permettra également d'identifier des corrélations entre les longueurs d'onde et ainsi aider à leur identification. Dans un second temps,

les algorithmes génétiques et CovSel seront utilisés pour confirmer les choix des longueurs d'onde d'intérêt.

IV.2. Théorie

IV.2.1. Les b coefficients de la PLS

La régression PLS est une méthode de régression linéaire multiple qui consiste à trouver des variables latentes T (également appelés *scores*), qui sont combinaisons linéaires à la fois des variables prédictives X (les spectres proche infrarouge) et des variables à prédire Y (ici les valeurs BMP) (Bertrand et al, 2006).

$$X = T_k P_k + R_k \quad \text{et} \quad y = T_k q_k + f_k$$

avec : X , la matrice des variables prédictives (les spectres proche infrarouge)

T_k , la matrice des *scores*

P_k , la matrice des *loadings* (poids) associé à la prédiction de X

R_k , la matrice des résidus associée à la prédiction de X

Y , le vecteur de la variable à prédire (les valeurs BMP)

q_k , le vecteur des *loadings* associé à la prédiction de y

f_k , le vecteur de résidus associé la prédiction de y

L'algorithme de la PLS travaille en boucle. Après la détermination de la première variable latente, une erreur résiduelle reste présente. On calcule à partir de cette erreur, une nouvelle variable latente. On atteint donc une seconde dimension dans la PLS. L'algorithme continue ainsi jusqu'à une dimension définie par l'utilisateur.

Les b coefficients PLS permettent de réaliser la régression linéaire sur les X afin de prédire les y :

$$\hat{y} = b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_n x_n \quad (\text{données centrées})$$

Avec : \hat{y} , la valeur prédite

b_n , la valeur du coefficient PLS à la longueur d'onde n .

x_n , la valeur du spectre à la longueur d'onde n .

Les b coefficients sont obtenus par combinaison linéaire entre les différents loadings (q et P) et le vecteur unitaire (W) du produit entre la matrice X et le vecteur y .

$$b = W(P^T W)^{-1} q \quad \text{avec} \quad W = \frac{X^T y}{\sqrt{(X^T y)^T X^T y}}$$

Le vecteur des b coefficients PLS a la même dimension qu'un spectre proche infrarouge, égale au nombre de longueurs d'onde utilisées dans le modèle. Les valeurs du vecteur peuvent être positives ou négatives et être interprétées telles quelles. L'interprétation se fait souvent en parallèle de l'observation de la moyenne des spectres. Les règles d'interprétation sont alors les suivantes :

- on ne regarde les valeurs des b coefficients PLS que lorsqu'elles sont associées à un pic dans les spectres d'absorbance.
- une valeur de b coefficient positive signifie que le composé qui absorbe à ces longueurs d'onde a une influence positive sur la variable prédite.
- une valeur de b coefficient négative signifie que le composé qui absorbe à ces longueurs d'onde a une influence négative sur la variable prédite.

Donc, une valeur positive dans les b coefficients, signifie que le composé qui absorbe à ces longueurs d'onde, contribue à augmenter la valeur de potentiel méthane. Le raisonnement inverse est valable pour une valeur de b coefficient négative.

Cependant, l'analyse des b coefficients PLS dans le domaine de la spectroscopie proche infrarouge, reste une démarche délicate. En effet, la spectroscopie proche infrarouge est le domaine des harmoniques et des bandes de combinaisons en relation avec les vibrations des molécules. Comparativement au moyen infrarouge, les pics sont moins bien définis et se

superposent entre molécules. Il en résulte une difficulté à définir avec précision un pic d'absorbance à une molécule particulière.

IV.2.2. Les algorithmes génétiques

Le principe des algorithmes génétiques (AG) est d'analyser l'ensemble des variables, afin d'en extraire une des meilleures combinaisons, c'est à dire celle qui maximise (ou minimise) un critère de qualité. Les AG conduisent souvent à une amélioration des résultats de prédiction, comparativement à ceux obtenus par les modèles comprenant l'ensemble des variables. L'algorithme génétique fonctionne en s'appuyant sur le paradigme de l'évolution, qui tend, dans une population donnée, à optimiser un critère de qualité lors du renouvellement des générations. Chaque individu de cette population est une combinaison de longueurs d'onde (représentée par un chromosome). A chaque génération, les gènes se croisent ou mutent, c'est-à-dire que les combinaisons de longueurs d'onde évoluent et seuls les gènes les plus performants sont conservés, provoquant l'amélioration globale de la population. Les AG fonctionnent donc par itération comme le décrivent Bertrand et al. (2006). Le fonctionnement des AG consiste à créer initialement une matrice C de $(k \times p)$ chromosomes, avec p représentant le nombre de longueurs d'onde. Un chromosome présente alors le même nombre de variables qu'un spectre. Via un filtre binaire, des valeurs (0 ou 1) sont affectées aux variables des chromosomes de façon aléatoire. 0 signifie que la variables n'est pas sélectionnée et 1, qu'elle est sélectionnée dans le chromosome.

Cette étape n'est effectuée qu'une seule fois au début du processus. Les étapes suivantes sont répétées en boucle :

- **Détermination de la valeur d'adaptation de chaque chromosome**

Avec chaque chromosome, est effectué un modèle de prédiction des y . L'erreur résiduelle de la prédiction des y est calculée, c'est la valeur d'adaptation

- **Sélection des deux meilleurs chromosomes**

Les deux chromosomes ayant la plus faible variance entre les valeurs prédites et les valeurs vraies sont sélectionnés. Ces deux chromosomes deviennent des chromosomes « parents ».

- **Etape de reproduction**

Deux nouveaux chromosomes « enfants » sont créés par croisement « crossing-over » des deux chromosomes parents. Le crossing-over consiste, tout comme en biologie, à croiser les éléments de deux chromosomes. Cette étape est réalisée en espérant que les « enfants » aient un potentiel de prédiction supérieur à ceux de leurs « parents ».

- **Etape de mutation**

La mutation consiste à inverser de façon aléatoire la valeur binaire (de 1 à 0 ou inversement) d'un élément de chromosome.

- **Actualisation de la population de chromosome**

Les deux chromosomes aux valeurs d'adaptations les plus faibles sont remplacés par les deux chromosomes enfants.

La boucle est ainsi répétée jusqu'à un nombre d'itérations défini ou jusqu'à la stabilisation des valeurs d'adaptation.

IV.2.3. L'algorithme CovSel

L'algorithme CovSel (Roger et al., 2010) a été mis en œuvre afin de répondre aux problèmes de sélection de variables, lors d'une régression linéaire à multi-réponses. Il peut cependant être utilisé pour une régression linéaire à une seule réponse. L'algorithme CovSel fonctionne également par itération. Les étapes principales de l'algorithme reposent sur 2 principes :

- sélectionner la variable i , parmi l'ensemble des variables des spectres, qui présente la plus grande covariance avec la variable y à prédire. Cette variable est un vecteur de n lignes, n représentant le nombre d'échantillons,
- projeter les spectres X et les réponses Y (qui sont des matrices $(n \times p)$ et $(n \times q)$) orthogonalement à ce vecteur, dans \mathbb{R}^n , de manière à ôter l'information portée par cette variable.

Cette boucle est ainsi répétée un nombre de fois égal au nombre maximal de variables défini par l'utilisateur. Le nombre de variables optimal peut ensuite être choisi par validation croisée de manière similaires à une PLS.

IV.3. Matériel et méthodes

Cette étude a été réalisée à l'aide des 74 échantillons (51 échantillons de l'ensemble d'étalonnage et 23 échantillons de l'ensemble de test) ayant permis de construire le modèle précédent.

L'analyse des données a été effectuée à l'aide de la PLS Toolbox d'Eigenvector Research, INC, utilisée avec le logiciel de programmation Matlab (© Mathworks). La méthode PLS a été utilisée pour construire le modèle. Le modèle étudié résulte donc de l'utilisation de la sélection de variables (1668 – 2500 nm) et des prétraitements SNV et Detrend.

Les traitements de sélection de variables par AG et par CovSel ont été appliqués. La source Matlab de l'algorithme génétique utilisé, est mise à disposition par Riccardo Leardi, sur le site Internet : <http://www.models.life.ku.dk/GAPLS>. La source Matlab de l'algorithme CovSel a été fournie par son concepteur, J.M. Roger (Cemagref ITAP, Montpellier).

Pour les AG, Leardi conseille dans sa documentation, de réduire au préalable le nombre de variables dans les spectres afin d'éviter un sur-ajustement du modèle. Pour cela, les moyennes sur 6 variables contiguës ont été réalisées afin de diminuer la taille des spectres de 1500 variables à 250 variables. Le nombre de chromosomes a été fixé à 30, soit C (30 x 250).

Le nombre d'itérations, la probabilité de crossing-over et la probabilité de mutation sont restées à leurs valeurs par défaut, soit respectivement 100, 50% et 1%.

L'AG ayant un fonctionnement stochastique, plusieurs répétitions de l'AG ne donneront pas obligatoirement les mêmes résultats. Afin de juger de la répétabilité de la méthode, Leardi conseille de renouveler l'AG un minimum de 5 fois par jeu de données. Nous avons réalisé 20 répétitions successives de l'AG.

Deux sélections de variables sont fournies (Figure 25).

Le meilleur modèle possible correspond au modèle ayant le meilleur SECV, nous l'appellerons « le modèle AG1 ». Une seconde sélection fournit le modèle avec le plus petit nombre de variables sélectionnées et dont la valeur de SECV n'est pas significativement différente de la plus faible valeur de SECV au terme d'un test de Fisher ($p < 0,1$). Ce modèle, dans un contexte de sélection de variables est suggéré, nous appellerons ce dernier modèle « le modèle AG2 ».

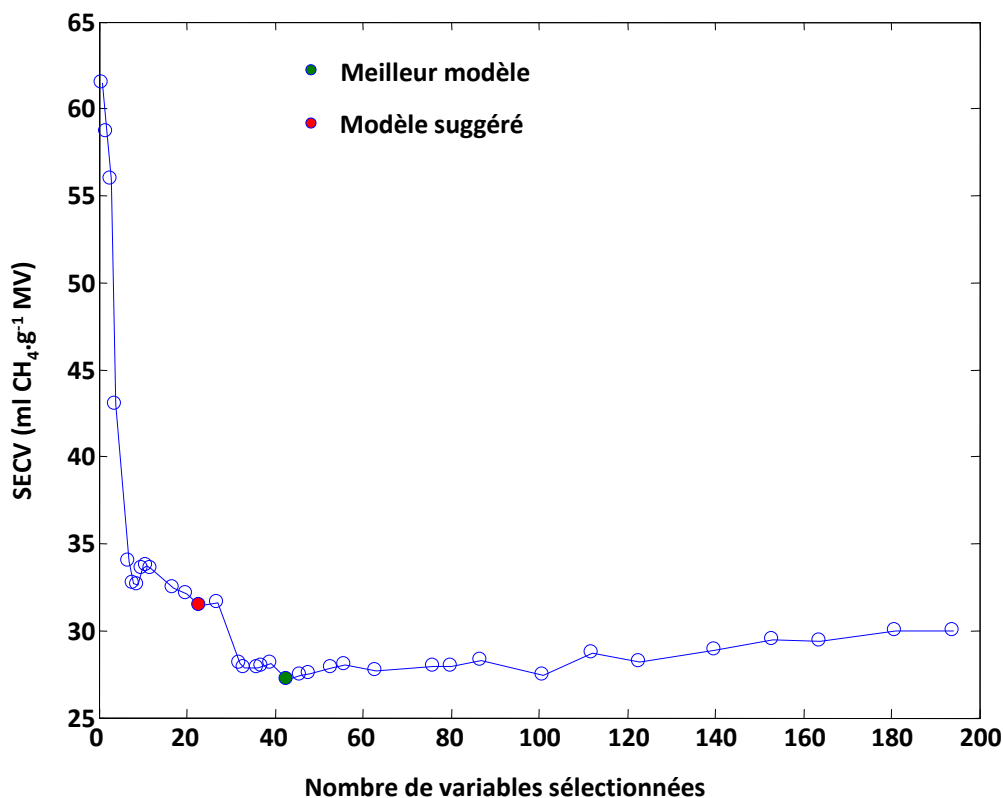


Figure 25 : Exemple de représentation de l'évolution des valeurs de SECV en fonction du nombre de variables sélectionnées dans le modèle lors d'une répétition. Le nombre de variable aux modèles AG1 et AG2 sont représentés.

Pour l'algorithme CovSel, aucune réduction de variables n'a été réalisée au préalable.

IV.4. Résultats et discussion

IV.4.1. Observation des b coefficients PLS

L'ensemble des spectres utilisés pour l'étude, ainsi que la moyenne et les écarts types de ces spectres sont présentés Figure 26. Les prétraitements, réduction de gamme à 1668 – 2500 nm,

SNV et Detrend ont été appliqués à l'ensemble des spectres. L'observation de la Figure 26b permet d'identifier les zones où les variations d'intensités spectrales ont lieu majoritairement, à savoir autour de 1930 nm, vers 2080 – 2100 nm, et enfin vers 2300 - 2480 nm.

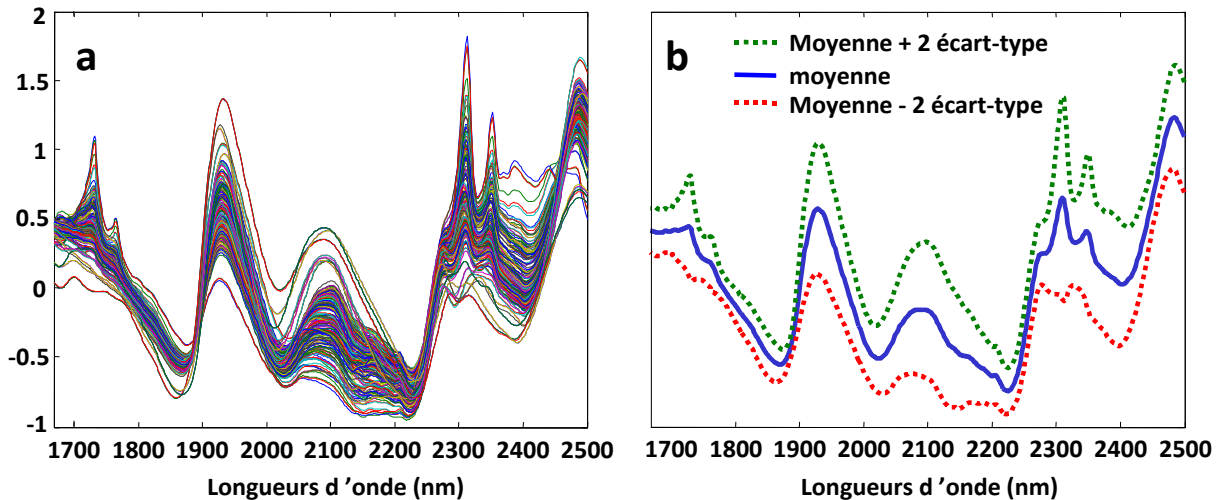


Figure 26 : Spectres proche infrarouge des déchets (a), spectre moyen et +/- 2 x écarts types (b).

Les b coefficients PLS correspondant aux 8 premières variables latentes et la dérivée seconde (taille de la fenêtre : 11 ; degré du polynôme : 3 ; dérivée : 2) du spectre moyen sont présentés Figure 27. L'intérêt de la dérivée seconde est qu'elle permet de déconvoluer le signal et donc de mettre les bandes d'absorption en évidence. Son inconvénient est qu'elle inverse les spectres, c'est-à-dire qu'un pic dans le spectre en absorbance se présente alors en dérivée seconde comme un « creux », ou « vallée ». Pour faciliter l'observation des pics d'intérêt, nous utiliserons l'opposée de la dérivée seconde (multiplication par -1) qui est montrée Figure 27. Un pic sur l'opposée de la dérivée seconde correspond donc à un pic d'absorbance du spectre original ($\log 1/R$).

IV.4.1.1. Les b coefficients PLS positifs

Intéressons-nous d'abord aux valeurs de b coefficients PLS positifs. C'est à dire ayant une influence en faveur d'un fort potentiel méthane.

La valeur la plus forte des b coefficients PLS se situe à 2308 nm. C'est la liaison méthyle, $C-H$, qui est souvent attribuée à cette longueur d'onde (Williams et Norris, 2001 ; Osborne et Fearn,

1986). La liaison méthyle est majoritairement présente dans les composés lipidiques, tels que les acides gras aliphatiques. Les lipides représentent une part non négligeable des déchets ménagers (Tableau 9). Et compte tenu du fort potentiel méthane théorique (1014 ml CH₄.g⁻¹ MV) des composés lipidiques, leur présence contribue à la production de méthane.

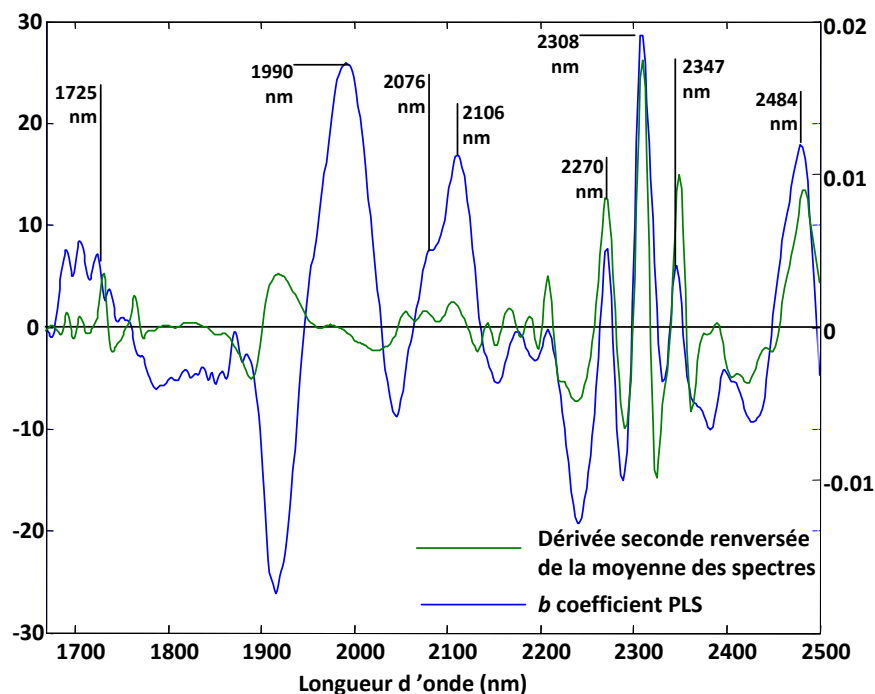


Figure 27 : Représentation des *b* coefficients PLS, de la dérivée seconde renversée de la moyenne des spectres et indication des longueurs d’onde correspondantes aux *b* coefficients PLS positifs.

Matière sèche (%)	Matière volatile (% MS)	Lipides (% MV)	Protéines (% MV)	Fibre cellulose (% MV)	Amidon (% MV)	Glucides simples (% MV)
29	86	17	18	19	17	7

Tableau 9 : Pourcentage des composés majeurs des déchets ménagers (Davidsson et al., 2007).

Pour mettre en évidence les corrélations entre différentes zones spectrales, nous utilisons une carte des corrélations (R²) (Figure 28). Elle montre clairement l’effet de contiguïté des longueurs d’onde (le long de la diagonale) et, la redondance d’informations contenue dans un spectre.

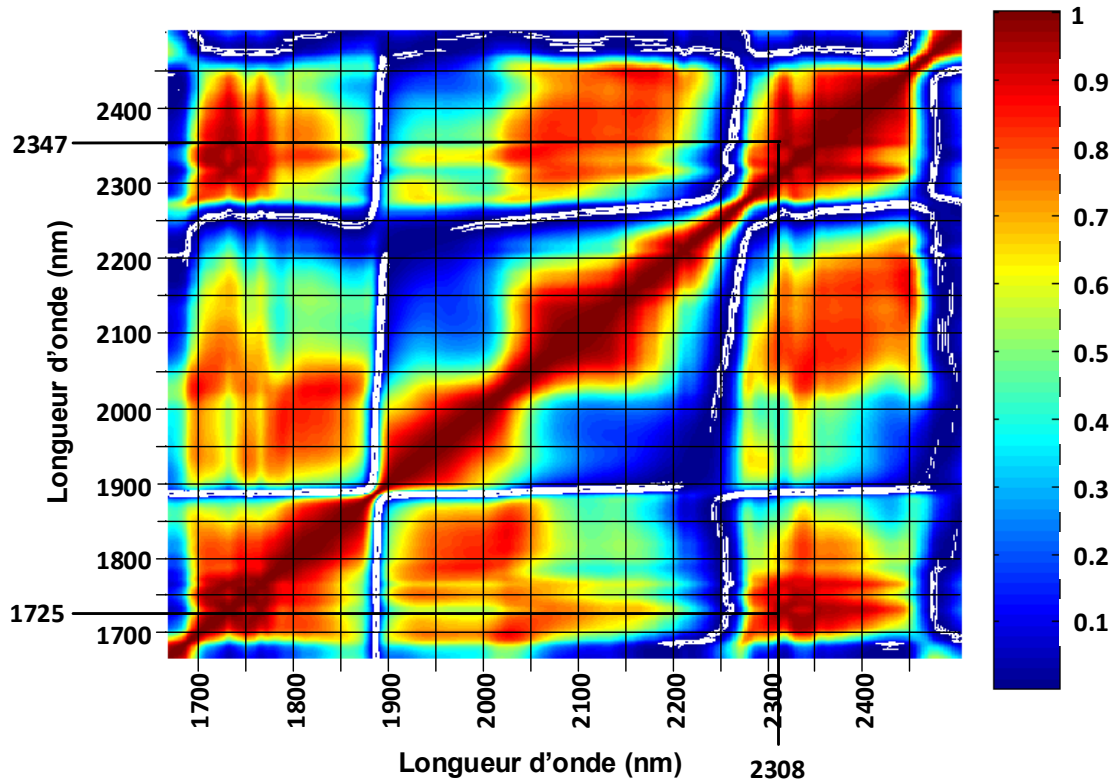


Figure 28 : Table de corrélation (R^2) entre les longueurs d'onde des spectres prétraités (SNV et Detrend sur la gamme 1668 – 2500 nm).

Cette carte met en évidence que l'absorbance à 2308 nm est fortement corrélée à trois zones distinctes. La première est située autour de 1725 nm, la seconde autour de 2347 nm et enfin la troisième entre 2050 nm et 2100 nm.

La zone autour de 1725 nm est la plus fortement corrélée avec 2308nm ($r^2 > 0,9$). Cette zone est également attribuée à la liaison $C-H$ couramment rencontrée au sein des lipides (Williams et Norris, 2001 ; Osborne et Fearn, 1986).

Le pic à 2347 nm est identifié comme étant également le siège de l'étirement de la liaison $C-H$ présent chez les lipides.

La 3ème zone corrélée à 2308 nm, entre 2050 nm et 2100 nm, est plus difficilement attribuable. Première harmonique de la combinaison $C-O/O-H$ et seconde harmonique de l'étirement de la double liaison $-C=C-$ (Williams et Norris, 2001), elle conforte l'hypothèse de l'absorbance des lipides à ces longueurs d'onde.

Cependant, à 2080 nm, le pic est également attribué à la liaison alcool ($O-H$) de la cellulose par Williams et Norris (2001) et Osborne et Fearn (1986). Ces derniers auteurs attribuent les pics à 2073 nm et 2318 nm respectivement, au glucose et au couple saccharose et amidon.

De même, la longueur d'onde à 1725 nm, attribuée à la cellulose par BenDor et al. (1997) et aux composés lipidiques/ glucidiques par Lebzein et al. (1997). Le pic situé à 2347 nm, est attribué à la cellulose par Ben-Dor et al., (1997) et Osborne et Fearn (1986).

On peut donc conclure que l'absorbance à 2308 nm est majoritairement due aux lipides mais les glucides semblent également y contribuer.

Le second pic positif dans les *b* coefficients se situe à 1990 nm et correspond au flanc d'un large pic autour de 1900 – 1960 nm, qui est la combinaison de 2 pics, l'un à 1910 nm (cf. section suivante) et un second vers 1930 nm. En absence d'eau, cette zone à 1930 nm est le lieu de l'absorbance de la fonction $O-H$ de l'amidon et de la cellulose (Williams et Norris, 2001). Ce « pic » des *b* coefficients est donc situé sur le flanc de ce large pic pour prendre en compte le pic à 1930 nm sans être influencé par le pic à 1910nm.

Le pic à 2484 nm (Figure 27) est identifié comme les combinaisons $C-H/C-C$ (Osborne et Fearn, 1986) ou $C-O/O-H$ des acides carboxyliques (Williams et Norris, 2001). Il est classiquement attribué à la cellulose (2480 nm) ou à l'amidon (2485 nm) (Williams et Norris, 2001 ; Osborne et Fearn, 1986 ; Jones et Schimleck, 2006).

Un large pic vers 2100 nm (Figure 27), est également visible. Il est la combinaison d'un premier pic à 2076 nm, proche des longueurs d'onde, 2080 nm et 2079 nm respectivement attribués au saccharose et à la cellulose (Williams et Norris, 2001 ; Osborne et Fearn, 1986 ; Jones et Schimleck, 2006). Le second pic, situé à 2106 nm peut également être attribué aux glucides (glucose, cellulose, amidon, hémicellulose) ($C-H, C-O, O-H$) (Williams et Norris, 2001 ; Osborne et Fearn 1986).

Enfin, le pic à 2270 nm, lieu de l'étirement de la liaison $C-O$ et de la combinaison $C-O/O-H$ des acides carboxyliques, est attribué aux glucides et plus spécifiquement à la cellulose (Williams

et Norris, 2001 ; Jones et al., 2006 ; Bailleres et al., 2002) ou à l'amidon (Osborne et Fearn, 1986) (Figure 29). Ce pic est également corrélé à la longueur d'onde à 1930 nm, attribuée à l'amidon/cellulose.

La plupart des *b* coefficient PLS positifs ont été attribués, aux lipides et aux glucides. Le pic le plus important dans les *b* coefficients PLS positifs est d'ailleurs attribué aux lipides mais avec une part de son absorbance également due aux glucides. Ces résultats sont en accord avec la littérature citant ces composés comme biodégradables et contribuant majoritairement au BMP. Les glucides contribuent principalement à la valeur BMP des déchets ménagers. De plus, une part de l'absorbance à 2308 nm serait également attribuée aux glucides.

Toute la matière organique n'est cependant pas biodégradable, par les micro-organismes en condition anaérobie. A l'opposé des *b* coefficients PLS positifs, attribués aux composés biodégradables, des *b* coefficients PLS négatifs devraient indiquer la présence de composés non biodégradables (plus ces composés sont présents, plus la valeur BMP sera faible).

IV.4.1.2. Les *b* coefficients PLS négatifs

Au sein des déchets solides, les molécules non biodégradables les plus rencontrées sont les différentes formes de lignines, les acides humiques, fulviques et les tanins. Globalement, ce sont des molécules complexes, composées en grande partie par des cycles aromatiques. Les matières plastiques non biodégradables, sont également présentes dans les déchets ménagers.

La vallée la plus importante dans les valeurs de *b* coefficients négatifs est située à 1910 nm (Figure 29). Elle est proche de la valeur du pic de l'eau (autour de 1940 nm), mais les échantillons ayant été lyophilisés, le pourcentage d'eau est très faible et ne peut expliquer cette vallée. Cette longueur d'onde est attribuée à la fonction alcool $O-H$, et à la double liaison $C=O$, dans les longues chaînes aliphatiques et les esters (Williams et Norris, 2001 ; Osborne et Fearn, 1986 ; Brinkmann et al., 2002). Elle a déjà été attribuée à la lignine par de nombreuses études (Brinkmann et al., 2002 ; Nousiainen et al., 2004 ; Wilman, et al., 2000, Ben-Dor et al., 1997).

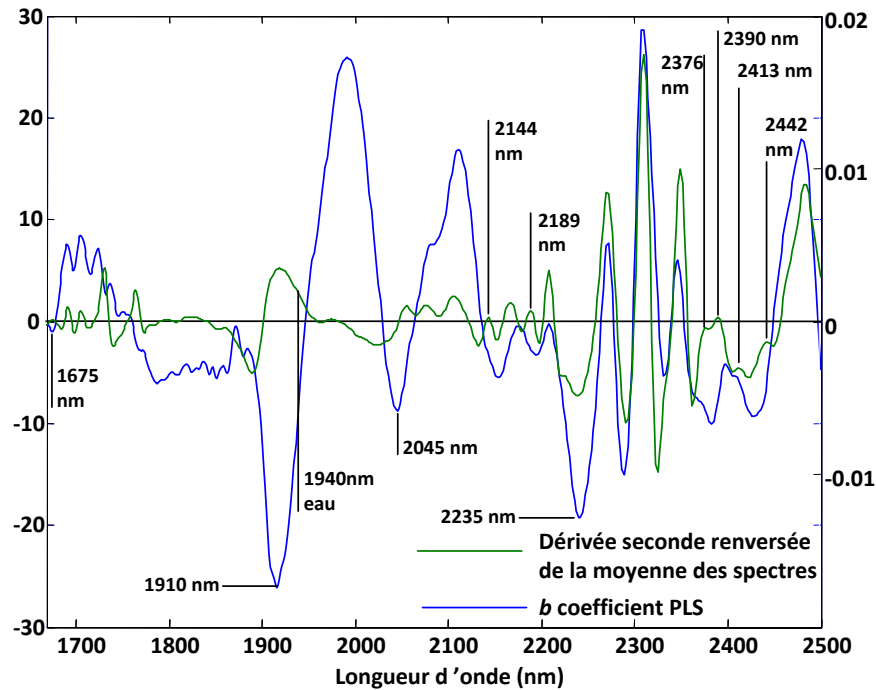


Figure 29 : Représentation des *b* coefficients PLS, de la dérivée seconde renversée de la moyenne des spectres et indication des longueurs d'onde correspondant aux *b* coefficients PLS négatifs.

La Figure 30 présente le spectre proche infrarouge de la lignine.

Le creux à 2390 nm, visible dans les *b* coefficients PLS, pourrait être également attribué à la lignine car il est répertorié par de nombreux auteurs (Brinkmann et al., 2002 ; Hodge et al., 2004 ; Kays et al., 2000). Il est notamment attribué par Williams et Norris (2001) au groupement phényl, fortement présent dans les lignines.

Mais les deux pics correspondant à ces deux creux ne sont pas corrélés (Figure 28). Deux explications sont possibles :

- Ce pic correspond effectivement à de la lignine mais, le terme « lignine » recouvre une famille de molécules très complexes. Il en existe de différentes formes, qui diffèrent de par les pourcentages et les types de monomères (Buranov et al., 2008), ce qui peut expliquer l'absence de corrélation entre les bandes que l'on attribue à la lignine.

- Ce pic à 2390nm est également relié à un autre composé (voir plus bas).

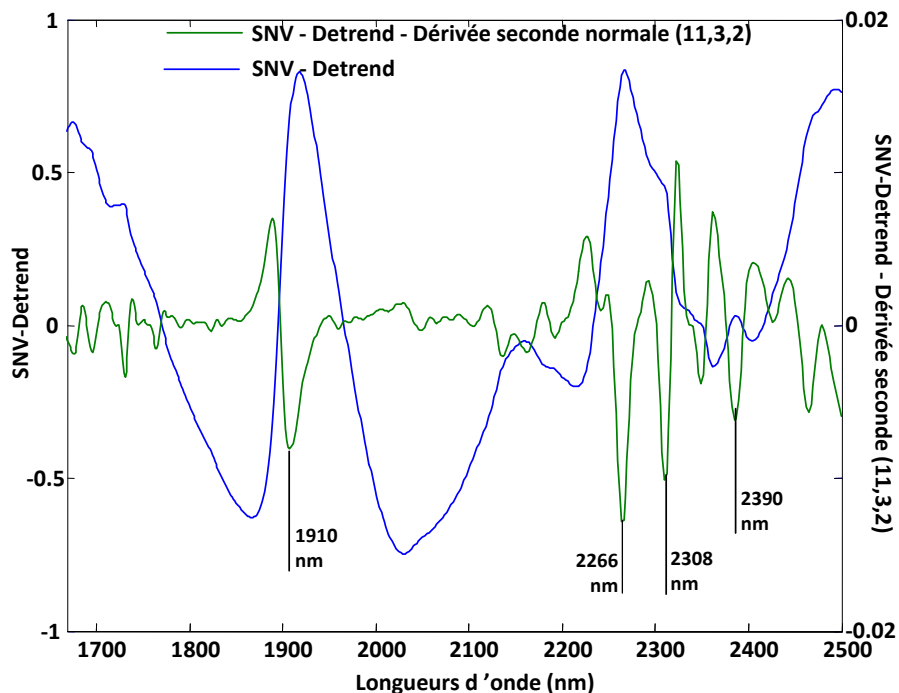


Figure 30 : Spectres prétraités, SNV-Detrend et SNV-Detrend-dérivée seconde (normale) (11,3,2) de la lignine.

A 2144 nm et 2189 nm (Figure 29), deux légères vallées peuvent être attribuées à la liaison $HC=CH$ (Osborne et Fearn, 1986). Cette liaison se rencontre dans les cycles benzéniques. La zone 2060 – 2150 nm est attribuée par Williams et Norris (2001) à la déformation du cycle benzénique. Egalement, la petite vallée à 1675 nm, étant attribuée à la déformation du cycle benzénique, est très souvent citée comme étant caractéristique de la lignine et plus généralement aux composés aromatiques (Williams et Norris, 2001 ; Osborne et Fearn, 1986 ; Jones et al., 2006 ; Bailleres et al., 2008 ; Brinkmann et al., 2002).

D'autres vallées sont rencontrées au niveau des *b* coefficients PLS négatifs : un massif entre 1780 et 1880 nm, 2045 nm, 2235 nm, un massif entre 2360 nm et 2460 nm (avec des vallées à 2376, 2413 nm et 2442 nm) (Figure 29).

Nous formulons l'hypothèse que les matières plastiques peuvent être responsables de ces vallées.

La Figure 31 présente les spectres proche infrarouge de divers composés plastiques pris dans les échantillons avant broyage.

On peut remarquer de nombreux pics dans la zone 2300 – 2500 nm, notamment à 2444 nm, qui pourraient expliquer l'origine du massif des *b* coefficients PLS négatifs obtenus dans cette zone. On retrouve, également entre 2144 nm et 2189 nm, des pics plus ou moins accentués selon le type de plastique. Cette zone est déjà attribuée aux cycles aromatiques, composés que l'on rencontre également dans les dérivés de plastiques.

La vallée à 2045 nm serait liée à l'absorbance dans les plastiques à 2055 nm.

On remarque également sur les spectres des plastiques, des pics plus ou moins prononcés autour de 1910 nm et 2390 nm. Ces pics étaient également présents dans les lignines. Il est possible qu'ils soient dus à l'un ou à l'autre de ces deux composés dans le spectre des déchets.

Un pic à 2308 nm, lieu de l'absorbance de *C-H*, est également très prononcé sur les plastiques. A cette longueur d'onde, les lipides absorbent également comme vu au 5.4.1.1. Cependant, les lipides et les plastiques ont des comportements opposés vis à vis de la biodégradabilité et du BMP. Les lipides sont dégradés et contribuent à la valeur BMP, c'est donc normal de retrouver un *b* coefficient positif à 2308 nm. Les plastiques qui absorbent également à cette même longueur d'onde (2308 nm), ne contribuent pas à la valeur BMP.

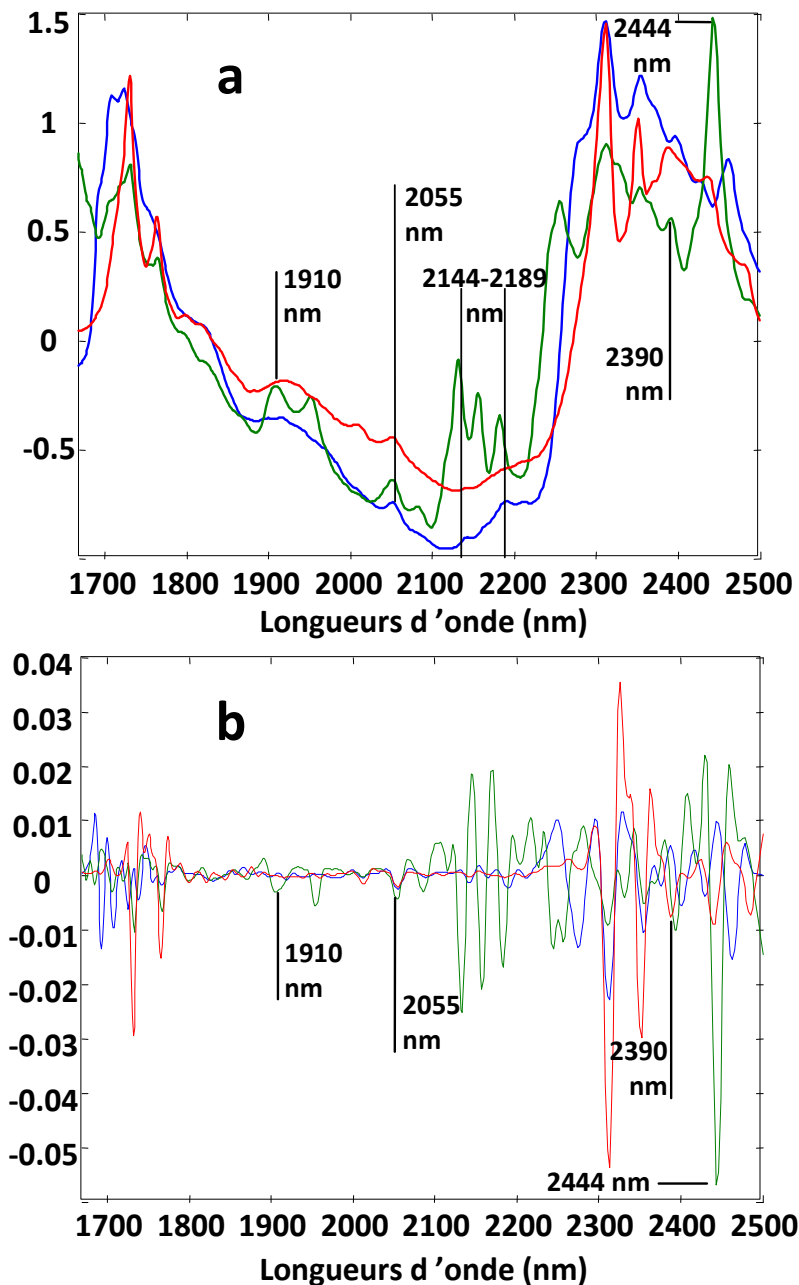


Figure 31 : Spectres de différents plastiques pris dans les échantillons.
 a : spectres SNV – Detrend des plastiques. b : dérivée seconde (11,3,2) des spectres SNV – Detrend des plastiques.

La relation entre l'absorbance à 2308 nm et la valeur BMP montre bien la présence de 2 composés aux comportements antagonistes (Figure 32). En effet, malgré les forts *b* coefficients positifs, le R^2 entre ces deux variables n'est que de 0,17. Du fait des *b* coefficients positifs à 2390 nm et de l'absorbance des plastiques à 2390 nm, la présence de ces derniers devrait donc se traduire par un accroissement du BMP. Pour corriger cette contribution positive des plastiques à

2308 nm, certaines zones, dans lesquelles les plastiques absorbent mais pas les lipides ou d'autres composés biodégradables, auront des *b* coefficients négatifs. Ces *b* coefficients négatifs apparaissent à 1780 nm– 1880 nm, 2140 – 2240 nm, et enfin dans la zone 2360-2460 nm où sont présents de nombreux pics d'absorbance des plastiques (Figure 33). La bande 2140-2240 nm englobe la vallée à 2235 nm, corrélée à 2440 nm, précédemment attribuée aux plastiques.

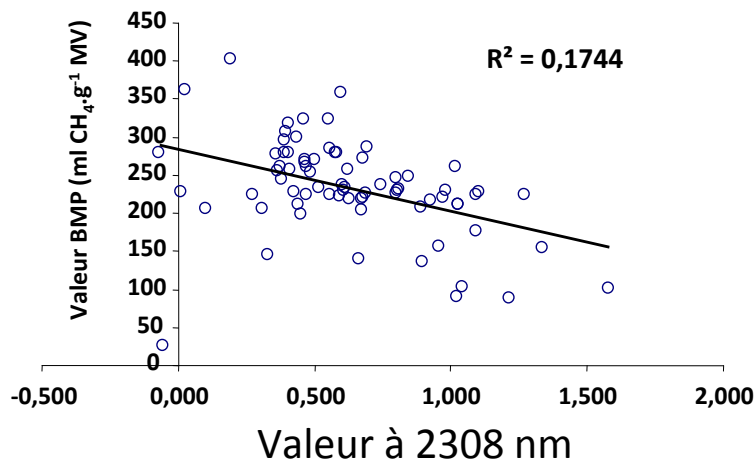


Figure 32 : Régression linéaire entre la valeur à 2308 nm et la valeur BMP.

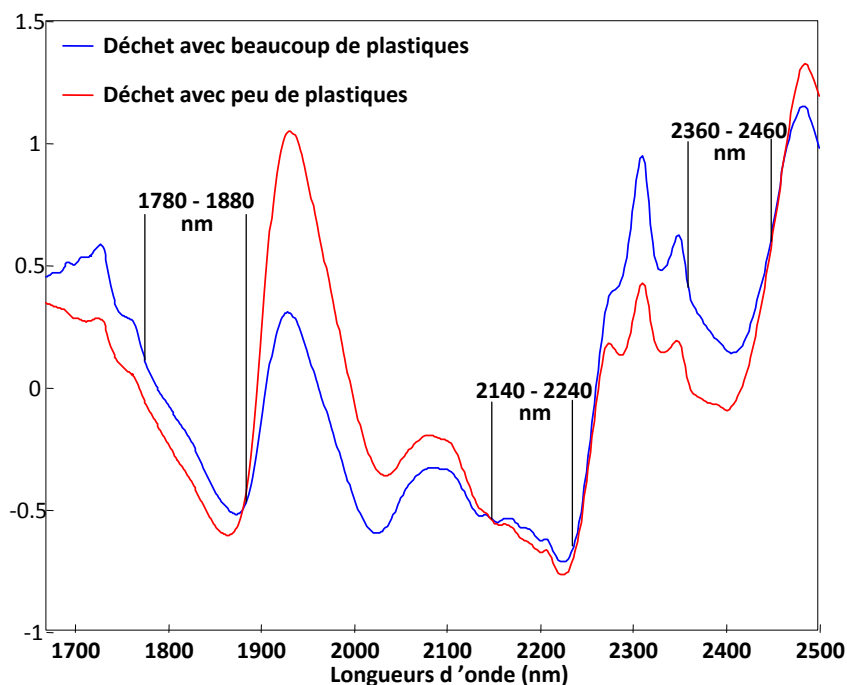


Figure 33 : Spectre d'un déchet avec beaucoup de plastiques (valeur BMP la plus faible) et spectre d'un déchet avec peu de plastiques (valeur BMP la plus forte), observation des zones de corrections de la contribution positive des plastiques à 2318 nm.

IV.4.2. Conclusion sur l'observation des *b* coefficients

La majorité des *b* coefficients positifs et négatifs ont été identifiés et expliqués, en accord avec la littérature. Les composés ayant une influence en faveur du BMP sont majoritairement les glucides (1930 nm, 2076 nm, 2106 nm, 2270 nm, 2484 nm et également dans une moindre mesure, 2308 nm) et les lipides (1725 nm, 2308 nm, 2347 nm). Les composés à *b* coefficients négatifs sont majoritairement les lignines (1910 nm, 2390 nm), les autres composés à cycles aromatiques tels que les acides humiques (1675 nm, 2144 nm, 2189 nm) et enfin les plastiques (1910 nm, 2045 nm, 2235 nm, 2390 nm et 2442 nm). Des corrections ont également lieu, aux longueurs d'onde (1780-1880 nm, 2140-2240 nm, 2360-2460 nm) pour compenser l'absorption du rayonnement par les lipides et les plastiques à la même longueur d'onde, 2308 nm.

Nous allons maintenant compléter cette approche par une étude utilisant la sélection de variables (par des algorithmes de sélection de variables tels AG et CovSel) pour identifier les composés ayant le plus d'impact dans la valeur BMP finale.

IV.4.3. Sélection de variables

Les résultats des sélections effectuées par 2 méthodes, les AG et l'algorithme CovSel, sont présentés.

IV.4.3.1. Variables sélectionnées par algorithme génétique

Les SECV obtenues par les différents modèles sont représentées Tableau 10.

	Sans AG	Meilleur modèle AG1	Second modèle AG2
Nombre de variables	500	87	28
Nombre de VL	8	6	9
SECV (ml CH₄.g⁻¹ MV)	31	27,6	31
R²	0,79	0,81	0,76

Tableau 10 Comparaison des résultats des modèles obtenus après sélection des variables par AG. VL : Variable latente

La Figure 34 présente les longueurs d'onde sélectionnées par les 20 répétitions de l'algorithme génétique, pour le modèle AG2 (Figure 34a) et AG1 (Figure 34b).

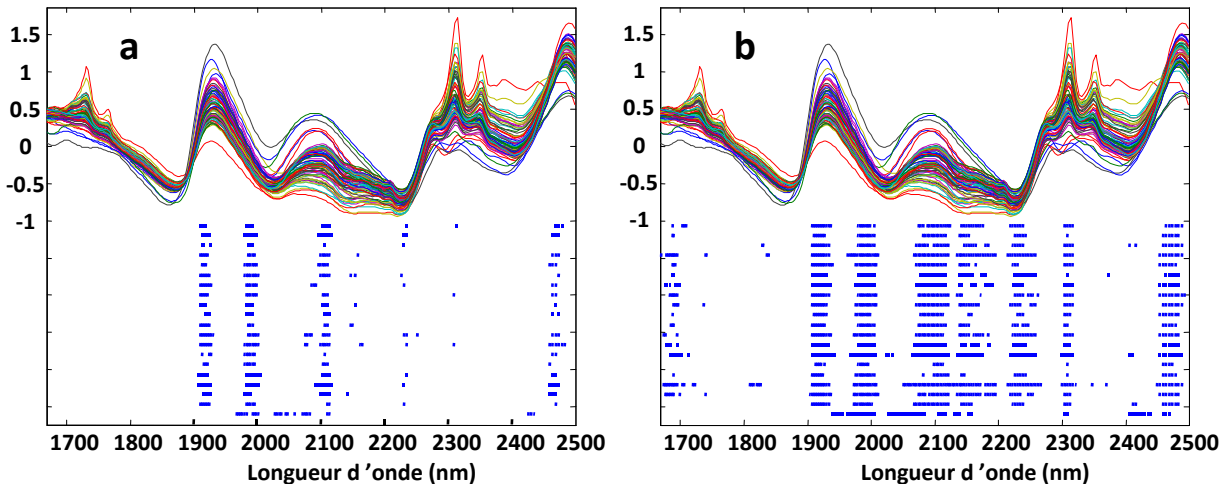


Figure 34 : Ensemble des spectres utilisés (sélection de variables (1668 – 2500 nm) - SNV – Detrend) et les variables sélectionnées par les 20 répétitions de l'algorithme génétique. a: variables sélectionnées correspondant au modèle AG2; b: variables sélectionnées par le modèle AG1.

Le modèle AG1 améliore les résultats d'étalonnage en prenant 87 variables. Le modèle AG2 est sensiblement identique que le modèle initial, avec seulement 28 variables. Il est intéressant de constater que les zones sélectionnées dans le modèle AG2 se retrouvent également dans le modèle AG1.

Nous allons commencer l'analyse par ce deuxième modèle dont la lecture est plus simple. Quatre zones sont régulièrement sélectionnées par l'algorithme pour le modèle AG2 (Figure 34a):

- La zone 1910 – 1930 nm, correspond à la zone la plus fortement négative au niveau des b coefficients négatifs (pic à 1910 nm) (Figure 29). Parmi l'ensemble des b coefficients négatifs (Figure 29), cette zone est la seule à être sélectionnée par l'AG. En effet, elle prend en compte l'ensemble des composés identifiés ayant un impact négatif sur le BMP, la lignine et les composés plastiques.
- La zone 1980 – 2000 nm, correspond à l'important pic positif des b coefficients, identifié comme étant la fonction $O-H$ de l'amidon ou cellulose en absence d'eau à la longueur d'onde 1930 nm (Osborne et Fearn, 1986 , Williams et Norris, 2001) (Figure 27). Cette zone est également corrélée (Figure 28) à la zone 1800 – 1850 nm,

où Osborne et Fearn (1986) et Williams et Norris (2001) indiquent une absorbance de la cellulose (1820 nm).

- La zone 2100 – 2115 nm correspond au pic des *b* coefficients PLS positifs à 2106 nm attribué aux glucides (glucose, cellulose, amidon, hémicellulose) (*C-H, C-O, O-H*) (Williams et Norris, 2001 ; Osborne et Fearn, 1986).
- A l'extrémité des spectres, la zone 2460 – 2476 nm a également été sélectionnée par l'algorithme génétique. Elle est contiguë à la bande à 2484 nm, attribuée aux glucides (amidon), qui est le lieu d'un pic positif très marqué des *b* coefficients.

L'analyse des longueurs d'onde sélectionnées met en évidence que les glucides totaux, cellulose et amidon principalement, (zone à 1980-2000 ; 2100-2215 nm ; 2460-2476 nm) contribuent principalement à la valeur BMP. Le modèle AG2, pour des résultats identiques en SECV, n'inclut pas la zone autour du pic à 2308 nm, malgré sa forte valeur dans les *b* coefficients positifs. Cela peut être expliqué par deux choses. D'une part, nous l'avons vu précédemment, la longueur d'onde à 2308 nm est une bande d'absorption à la fois des lipides et des plastiques qui ont des comportements opposés vis à vis du BMP ($R^2=0,17$ entre le pic et le BMP). D'autres part, la valeur BMP d'un déchet ménager est majoritairement due aux glucides (principalement cellulose et amidon), pas aux lipides.

Dans le modèle AG1, 4 zones supplémentaires ont été sélectionnées par l'algorithme, comparativement au « second modèle ». Les 4 zones supplémentaires sont 1680 – 1700 nm, 2140 – 2180 nm, 2220 – 2260 nm et enfin 2304 – 2320 nm (Figure 34b).

- La zone 1680 – 1700 nm correspond à un petit pic positif dans les *b* coefficients à 1690 nm. Un composé possible à cette longueur d'onde serait les protéines (Williams et Norris, 2001). De plus dans la table de corrélation, 1690 nm apparaît corrélé à 2030 nm et à 2265 nm, attribués aux protéines (Osborne et Fearn, 1986 ; Williams et Norris, 2001). 1690 nm est, en effet, le premier pic de la série de 3 pics autour de

1700 nm. Il est associé à un *b* coefficient PLS positif, signe de la biodégradabilité des protéines, en accord avec la littérature (Chen et al., 2008).

- La zone 2140 – 2180 nm inclut la vallée à 2144 nm, attribuée aux cycles benzéniques.
- La zone 2220 – 2260 nm représente la vallée à 2235 nm, attribuée également aux cycles benzéniques. Cette zone s'arrête juste avant le pic à 2270 nm, que l'on avait attribué à la cellulose.
- Enfin, la zone 2304 – 2320 nm inclut le pic à 2308 nm déjà attribué aux lipides et aux plastiques.

Dans le meilleur modèle AG, le pic à 2308 nm a été sélectionné, prenant en compte la fraction lipidique. La zone 2140-2180 nm permet alors de corriger l'absorbance des plastiques à 2308 nm. Enfin, la zone 1680 – 1700 nm a également été sélectionnée afin d'exprimer la part protéique des déchets.

Conclusion : les zones spectrales sélectionnées par AG concordent parfaitement avec les pics positifs ou négatifs des *b* coefficients et sont expliquées au niveau moléculaire, ce qui confirme la pertinence et la stabilité du modèle.

IV.4.3.2. Variables sélectionnées par CovSel

La Figure 35 présente les variables sélectionnées par l'algorithme CovSel.

Le nombre de variables sélectionnées est très inférieur au nombre proposé par l'algorithme génétique. Seulement huit variables ont été sélectionnées. L'étalonnage est légèrement meilleur que les précédents étalonnages: $SECV = 27,6 \text{ ml CH}_4 \cdot \text{g}^{-1} \text{ MV}$; $R^2 = 0,81$, au lieu de $SECV = 31 \text{ ml CH}_4 \cdot \text{g}^{-1} \text{ MV}$; $R^2 = 0,79$, avec la PLS réalisée sur les 74 échantillons en étalonnage. Ces résultats sont identiques à ceux du meilleur modèle AG1, avec seulement 8 variables sélectionnées contre 87 dans l'AG2.

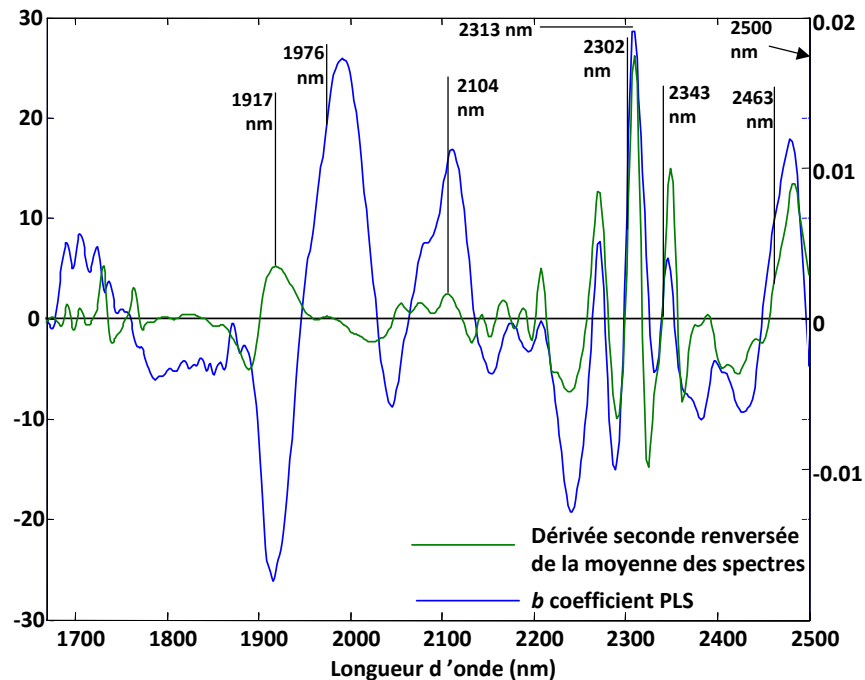


Figure 35 : Représentation des b coefficients PLS, de la dérivée seconde renversée de la moyenne des spectres et indication des longueurs d'onde sélectionnées par l'algorithme CovSel.

Certaines variables sélectionnées sont proches ou identiques de celles sélectionnées par le modèle AG1 (1917 ; 1976 ; 2104 ; 2302 ; 2313 ; 2463 nm). Seules les variables 2343 nm et 2500 nm n'ont pas été sélectionnées par l'AG.

Deux variables autour de 2310 nm ont été sélectionnées : 2302 et 2313 seraient attribuées aux lipides. La bande à 2343 nm est proche de 2347 nm, précédemment citée et attribuée aux lipides. La variable à 2500 nm est également sélectionnée, ce qui n'est pas forcément significatif car il semblerait que les longueurs d'onde extrêmes soient régulièrement sélectionnées par l'algorithme CovSel afin de corriger une ligne de base résiduelle (communication personnelle de J.M. Roger, concepteur de CovSel). Comme précédemment, la seule variable à b coefficients négatifs sélectionnée est à 1917 nm, attribuée majoritairement à la lignine, mais également aux plastiques. Enfin, la variable à 1976 nm est également sélectionnée par CovSel. Elle correspond au flanc du pic d'absorption à 1930 nm (pic des $O-H$ de l'amidon-cellulose) dans l'analyse des b coefficients et est contiguë à la zone 1980-2000 repérée par l'AG.

Une régression multiple (MLR) est faite à l'aide des variables sélectionnées par CovSel. Les coefficients sont présentés à la Figure 36.

Les plus fortes valeurs positives sont maintenant associées aux longueurs d'onde attribuées aux glucides (1976 nm, 2104 nm, 2463 nm). Cela confirme bien que ce sont les glucides qui ont la plus forte contribution envers le BMP. Les coefficients associés aux lipides, notamment autour de 2310 nm ont considérablement diminué par rapport aux b coefficients de la PLS. Cela permet de prendre en compte l'effet positif des lipides sur le BMP, mais en limitant la contribution des plastiques, qui absorbent également à cette longueur d'onde. Le coefficient associé à la longueur d'onde 2500 nm est également très faible, ne jouant qu'un rôle mineur (redressement de la ligne de base).

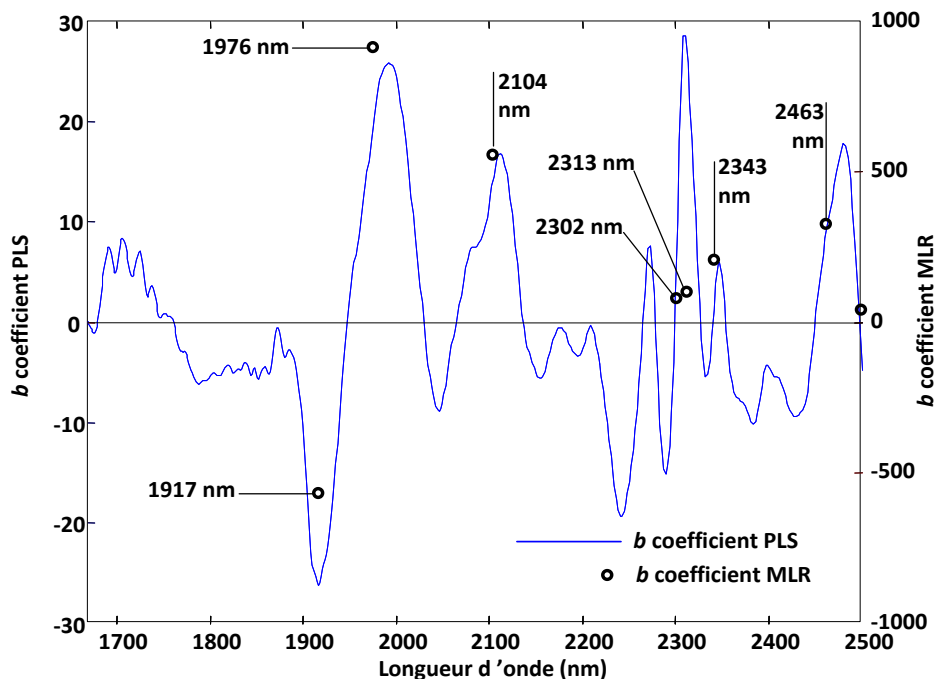


Figure 36 : Représentation des b coefficients PLS et des b coefficients associés à la régression multiple (MLR) faite à l'aide des variables sélectionnées par CovSel.

IV.5. Conclusion

L'objectif de cette étude, était d'identifier les b coefficients PLS les plus importants dans la régression PLS, précédemment mise en place, afin de caractériser le potentiel méthane des

déchets ménagers. L'étude des *b* coefficients PLS permet d'identifier les molécules ayant un impact « positif » (production de méthane), ou un impact « négatif » (pas de production de méthane). Nos résultats montrent, en accord avec la littérature, que les glucides (cellulose, amidon et glucides simples) sont les composés jouant le rôle majoritaire dans le BMP. Cependant, malgré leur plus faible concentration, les lipides ont également un rôle méthanogène essentiel.

L'utilisation d'algorithme de sélection de variables (algorithme génétique ou CovSel) nous a permis de valider les longueurs d'onde d'intérêt les plus facilement identifiables, et d'en déceler des nouvelles (1976 nm et 2463 nm). Une synthèse des longueurs d'onde, énoncées dans ce chapitre, est présentée Tableau 11. Certaines longueurs d'onde d'intérêt n'ont cependant pas été retenues par les algorithmes de sélection de variables. L'étude conjointe de l'analyse des *b* coefficients PLS et des résultats de sélection de variables permet de vérifier la pertinence du modèle. En effet, cela permet de s'assurer que les variables sélectionnées ont bien un sens chimique vis à vis de la variable à prédire.

Longueur d'onde (nm)	Liaison chimique en jeu	Composé probable	Pic b coefficient PLS	Sélection		Sélection Cov-Sel	Réf.
				AG 1	AG 2		
1675	HC = CH	Composés aromatiques	-				1-2-3-4-5
1690	CO - NH	Protéines liaison peptidique	-				1-2
1725	C - H	Lipides	+				1-2
1910	Combi C - H / C - C ou O - H ou C = O	Lignines plastiques	-			1917	1-2-5-6 7-8
1930							
1976	O - H	Amidon - cellulose	+				1-2
1990							
2060 à 2150	C - H	Déformation du cycle benzénique	-				1
2076	C - O ; O - H	Saccharose (2080 nm) Cellulose (2079 nm)	+				1-2-3
2106	C - H C - O O - H	Glucides (glucose ; cellulose ; amidon ; hémicellulose	+				1-2
2144	HC = CH C - H HC = CH	Cycle benzénique Plastiques Cycle benzénique Plastiques	-				2 1 2
2189	C - H	Déformation du squelette carboné	-				1
2230	C - H	Déformation du cycle benzénique	-				1
2265	CO - NH	Protéines	+				1
2270	C - O Combi C - O / O - H	Cellulose Amidon	+				1-3 4-2
2300	N - H	Protéines	+				1
2308	C - H	Lipides ; Plastiques ; glucides	+			2302 2313	1-2
2345	N - H	Protéines	+				1
2347	C - H	Lipides	+				1-2
2390	C - H ; C - C	Lignines et plastiques	-				5-9-10
2444	C(CH ₃) ₃ Combi	Plastiques Cellulose	-				1 2
2484	C - H / C - C Combi C - H / C - C	Cellulose Amidon	+			2463	1 - 3

Tableau 11 : Attribution chimique des longueurs d'onde d'intérêt du modèle de prédiction de la valeur BMP d'ordures ménagères.

1 : Williams et Norris (2001) ; 2 : Osborne et Fearn (1986) ; 3 : Jones et al., (2006) ; 4 : Baillères et al., (2002) ; 5 : Brinkmann et al. (2002) ; 6 : Nousiainen et al. (2004) ; 7 : Wilman et al. (2000) ; 8 : Ben-Dor et al. (1997) ; 9 : Hodge et al.(2004) ; 10 : Kays et al. (2000).

V.

*Etude de la robustesse du modèle vis à vis
de la nature des échantillons*

Ce chapitre est dédié à l'étude de la robustesse de la méthode vis à vis de l'origine/nature des déchets, pour répondre à l'objectif d'un modèle opérationnel en conditions industrielles.

V.1. Introduction

Le modèle, établi précédemment sur les déchets ménagers, montre le potentiel de la spectroscopie proche infrarouge, pour la prédiction du potentiel méthane de déchets solides. Cependant, ce modèle est basé sur des ordures ménagères, provenant d'un site unique. Malgré le fait que ces échantillons soient issus de divers niveaux d'un processus de tri par criblage, ils ont la même origine et ne couvrent pas de manière exhaustive la diversité des déchets que l'on peut rencontrer en entrée de digesteurs anaérobies industriels. Il serait donc intéressant, de voir comment se comporte le modèle établi précédemment, avec des échantillons d'origines et de natures différentes de la base d'étalonnage.

Les objectifs de cette partie sont d'étudier la robustesse du modèle précédemment construit, vis à vis de la nature et de l'origine des déchets.

Dans un premier temps, notre démarche sera de tester le modèle, construit précédemment, avec un nouvel ensemble d'échantillons, d'origines et de natures diverses. Dans un second temps, une analyse multivariée sera réalisée, afin de comprendre l'origine de la variabilité entre les spectres. Enfin, différentes solutions au problème de la robustesse seront avancées et l'une d'elle sera validée.

V.2. Matériel et méthodes

V.2.1. *Les échantillons*

Pour cette partie, deux ensembles d'échantillons ont été utilisés.

Le premier ensemble, correspond aux 74 échantillons utilisés dans les parties précédentes. Ils sont composés de 68 déchets d'ordures ménagères, issus d'un site unique de tri-compostage et de 6 déchets simples. Ces 74 échantillons, ont été lyophilisés et broyés. Ils composent la base

d'étalonnage. Le second ensemble, comprend, 100 nouveaux échantillons d'origines et de natures diverses, comme indiqué dans le Tableau 12.

Catégories	Nombre	Exemples de déchets
Ordures ménagères	11	OM brutes d'origines diverses, digestats, déchet reconstitué, refus d'affinage
Boues	14	Boues de STEP ; boue grasseuse (chargée en lipides)
Déchets d'origine végétale	22	Paille de riz, tige de <i>miscanthus</i> , feuille de maïs, feuille d'oliviers, topinambour, sorgho, eucalyptus, écorces de jardin, pomme de terre, gomme arabique
Déchets d'origine animale	7	Déchet d'abattoir, steak, carcasse, poisson, mollusque, coquillage, fémur de poulet
Déchets d'élevages agricoles	9	Fumier (bovin, caprin équin), mélange de fumiers, lisiers, fientes
Algues (micro et macro)	14	Sargassum, <i>Palmaria Palmata</i> , <i>Chlorella</i> , <i>Laminaria</i>
Pailles	16	Mêmes échantillons de pailles, broyées à différentes tailles
Déchets « simples », assimilés à des déchets d'industrie agro-alimentaire (matières transformées) et à des déchets industriels	7	Carton, papier journaux, pain, coquillage, déchets de cantine, yaourt, jus de cuisson

Tableau 12 : Différentes catégories de déchets et nombre d'échantillons correspondants.

Ils sont indépendants deux à deux et proviennent tous de sites différents, comparativement à l'ensemble d'étalonnage. Ce second ensemble présente des déchets d'origine végétale (différentes parties de plantes cultivées), d'origine animale (déchets carnés, poissons broyés), des déchets d'élevages agricoles (lisiers, fumiers), des boues de stations d'épuration, des ordures ménagères d'origines diverses, des algues (micro et macro-algues) et des déchets dit « simples » (pain, carton, papier, yaourt, etc.).

Ces 100 échantillons peuvent être globalement considérés comme indépendants. Ils sont tous issus de sources différentes (*i.e.* d'entreprises différentes). Ils sont de natures différentes (ordures

ménagères, algues, déchets agricoles...). De plus, ils ont été analysés par spectrophotométrie PIR à des dates différentes, afin d'intégrer d'éventuels problèmes de reproductibilité liés au spectrophotomètre (les mesures ont été faites sur 17 jours, sur 2 années).

Les moyennes, écarts-types et les minimums et maximums des deux ensembles sont présentés Tableau 13.

	Moyenne (Ecart type) (ml CH ₄ .g ⁻¹ MV)	Minimum – maximum (ml CH ₄ .g ⁻¹ MV)
74 échantillons d'étalonnage	232 (63)	23 - 400
100 échantillons de test	250 (135)	13 – 766

Tableau 13 : Moyennes, écarts-types, minimums et maximum des deux ensembles d'échantillons.

V.2.2. *Traitement des données et paramètres statistiques*

L'analyse des données a été effectuée à l'aide de la PLS Toolbox d'Eigenvector Research, INC, utilisée avec le logiciel de programmation Matlab (© Mathworks).

V.2.2.1. Test du modèle avec les nouveaux échantillons

Le modèle est celui établi précédemment sur la base des 74 échantillons de déchets. Les prétraitements, sélection de gamme (1668 – 2500 nm), SNV et Detrend, ont également été appliqués à l'ensemble des 100 échantillons indépendants.

V.2.2.2. Analyse en Composante Principale (ACP)

V.2.2.3. Construction et validation d'un modèle global

Deux solutions peuvent être imaginées pour résoudre les problèmes de robustesse vis-à-vis de la variabilité des échantillons :

- Utiliser l'ensemble des échantillons disponibles pour construire un modèle global plus robuste,
- utiliser l'ensemble des échantillons dans un modèle local utilisant le principe de « plus proches voisins ».

La première solution a été mise en œuvre en utilisant les 174 échantillons. Ils ont donc été séparés en deux groupes : un groupe d'étalonnage (117 échantillons) et un groupe de test (57 échantillons). Le Tableau 14 présente les valeurs moyennes, écarts types, min et max des ensembles d'étalonnage et de validation.

Ensemble d'étalonnage (ml CH ₄ .g ⁻¹ MV)		Ensemble de validation (ml CH ₄ .g ⁻¹ MV)	
Moyenne (Ecart type)	Min – Max	Moyenne (Ecart type)	Min – Max
244 (116)	13 – 766	235 – 97	23 – 507

Tableau 14 : Moyenne, écart type, minimum et maximum des ensembles d'étalonnage et de validation.

Plusieurs modèles de régression PLS ont été testés avec différentes combinaisons de prétraitements. Les prétraitements utilisés sont, SNV (Barnes et al., 1989), Detrend (Barnes et al., 1989), la dérivée première et seconde de Savitsky et Golay (Savitsky et al., 1964), et des plages spectrales.

V.3. Résultats et discussion

V.3.1. Prédiction des nouveaux échantillons à l'aide du modèle précédent

La prédiction de l'ensemble des 100 nouveaux échantillons a été réalisée en prenant comme ensemble d'étalonnage l'ensemble des 74 échantillons d'ordures ménagères. Les résultats de la prédiction sont présentés Figure 38.

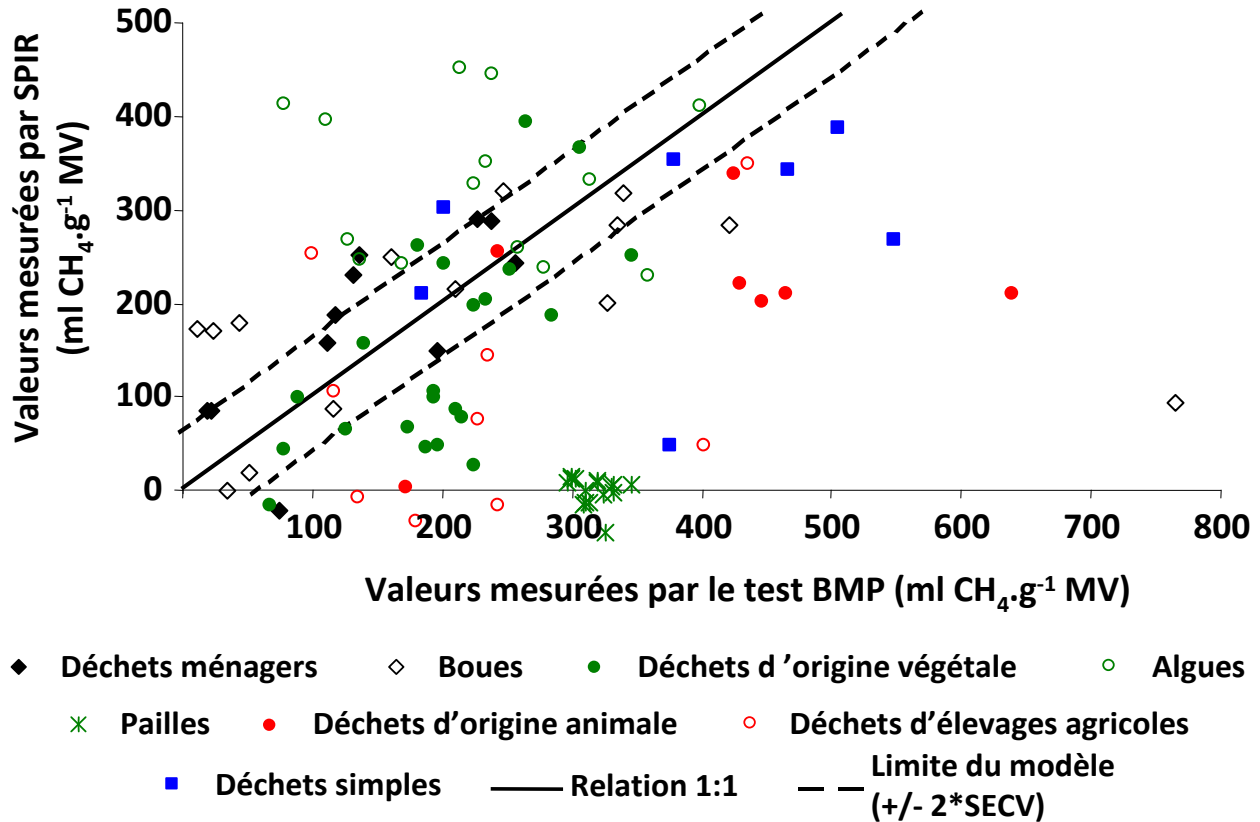


Figure 38 : Prédiction des 100 nouveaux échantillons.

Le RMSEP est de 177 ml CH₄.g⁻¹ MV avec un R² = 0,02. Ces résultats sont très loin de la limite du modèle qui était de 61 ml CH₄.g⁻¹ MV (1,96*SECV).

Seulement 30 déchets sur les 100 sont compris dans les limites d'incertitudes (+/- 61 ml CH₄.g⁻¹ MV) du modèle. Près de 73% des déchets ménagers sont inclus dans ces limites d'incertitudes du modèle (RMSEP_{déchets ménagers} = 65 ml CH₄.g⁻¹ MV). Cela montre que le modèle utilisé est correct pour des déchets ménagers.

Certains autres déchets, sont largement hors des limites du modèle, comme par exemple, l'ensemble des échantillons de paille, la quasi-totalité des échantillons de déchets d'origines animales, les algues, nombreux déchets d'origines végétales, certaines boues. Onze déchets sont hors de la gamme du modèle d'étalonnage (23 - 400 ml CH₄.g⁻¹ MV), forçant le modèle à extrapoler, ce qui explique certaines mauvaises prédictions (déchet animal, déchet de cantine, boue grasseuse). Même sans ces quelques échantillons, le RMSEP reste inchangé avec 168 ml CH₄.g⁻¹ MV.

La bonne prédiction des déchets ménagers et les mauvaises prédictions pour les autres catégories de déchets montre bien le manque de robustesse du modèle vis à vis de la nature des déchets. Celle-ci était prévisible – mais demandait à être confirmée- étant donnée la diversité réduite des échantillons d'étalonnage.

Une analyse multi-variée a été réalisée afin d'analyser le manque de robustesse du modèle.

V.3.2. Analyse multi-variée des échantillons

V.3.2.1. Projection des échantillons de test sur l'espace engendré par les deux premières composantes PLS.

Les 100 échantillons ont été projetés dans l'espace des 2 premières composantes PLS du modèle de régression basé sur les déchets ménagers (Figure 39). Ces deux premières composantes expliquent à elles seules près de 97% de la variance des déchets ménagers.

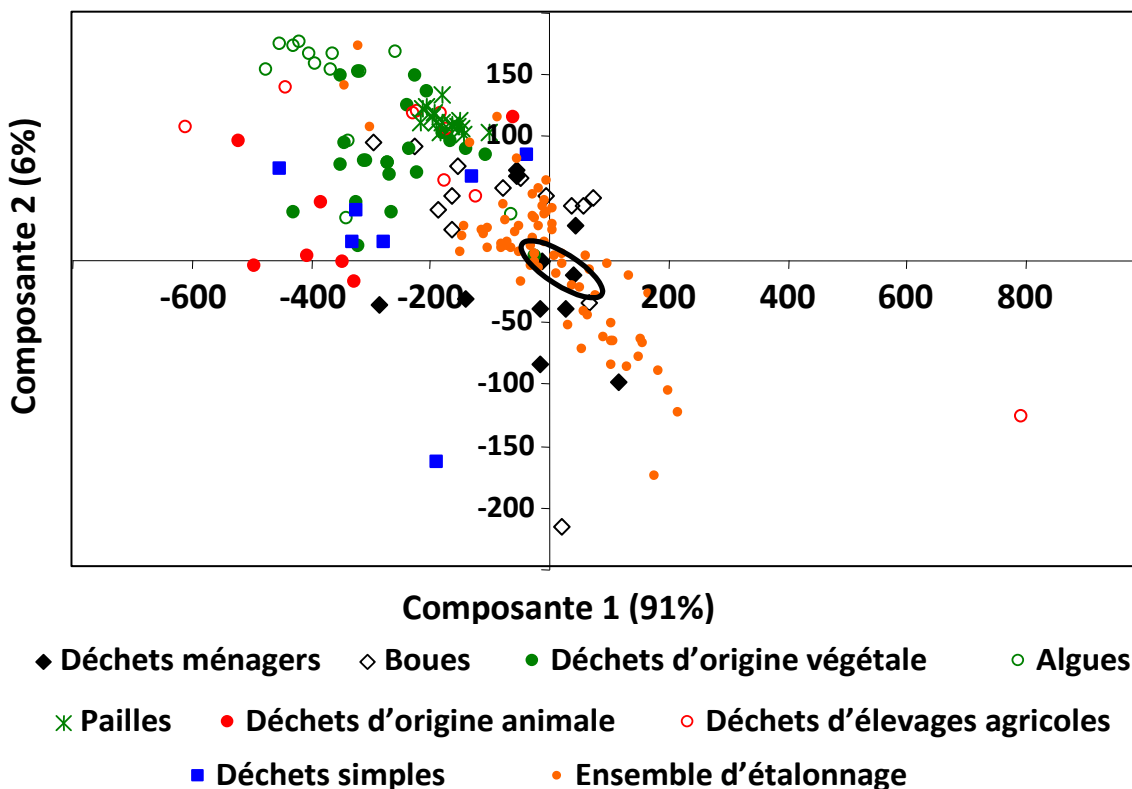


Figure 39 : Représentation de deux premières composantes PLS du modèle sur les ordures ménagères et projection des 100 nouveaux échantillons sur ces deux composantes.

Les échantillons d'étalonnage apparaissent clairement comme un nuage de point distinct et séparé des autres échantillons. La plupart des nouveaux échantillons de test ne sont pas inclus dans ce nuage de point, expliquant en partie leurs mauvaises prédictions. Seuls les nouveaux déchets ménagers sont inclus dans le nuage de point. Globalement, les échantillons inclus dans le nuage de point sont bien prédits, mais ce n'est pas une condition suffisante. En effet, les deux échantillons situés au centre du graphique et entourés d'une ellipse noire ne sont pas correctement prédits.

Nous avons donc essayé de prédire la variabilité entre les différents échantillons par une ACP.

V.3.2.2. Réalisation d'une analyse ACP sur l'ensemble des échantillons d'étalonnage et de test

Afin d'analyser la variabilité au sein des 174 échantillons prétraités (sélection de variable (1668 – 2500 nm, SNV et Detrend), une ACP a été réalisée. La Figure 40 représente les deux premières composantes, qui expriment ensemble 76% de la variance totale des échantillons.

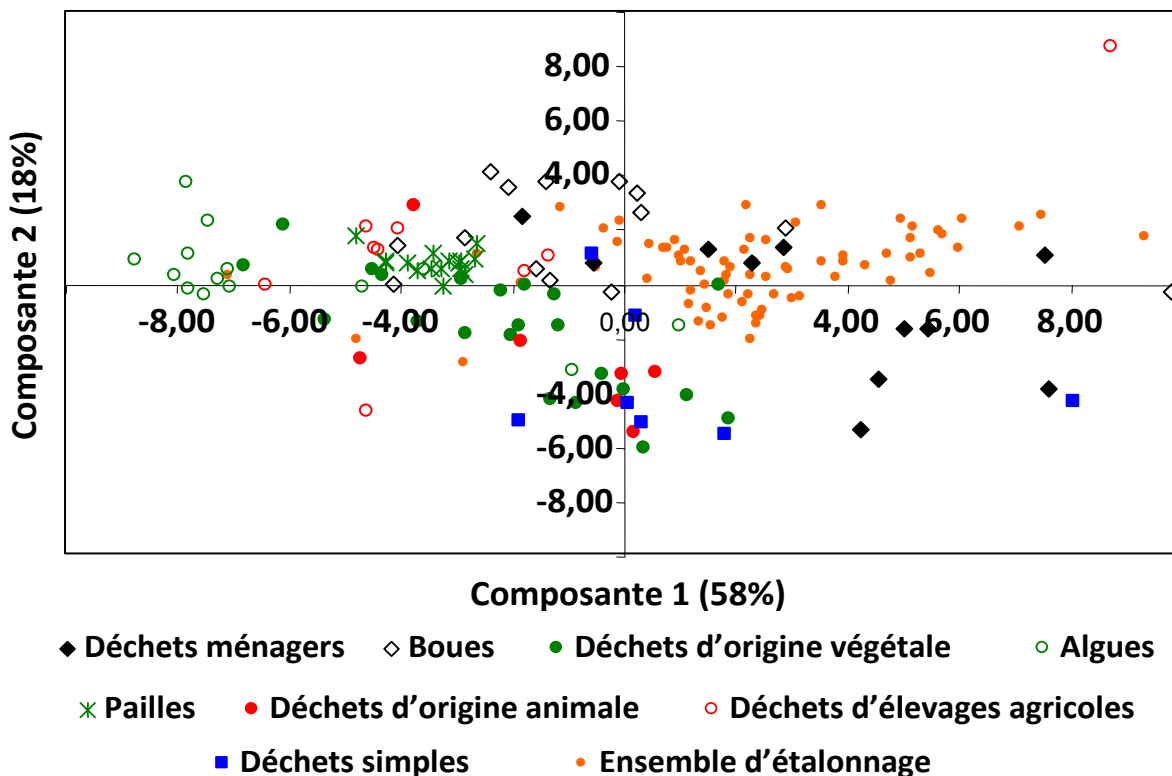


Figure 40 : Première et seconde composante principale de l'ACP effectuée sur l'ensemble des échantillons.

Globalement, les 74 échantillons ayant servi à l'étalonnage sont séparés des 100 échantillons de test. Selon la première composante, on voit clairement que l'ensemble d'étalonnage n'est pas représentatif de la variabilité au sein des échantillons de test. A noter que les nouveaux échantillons d'ordures ménagères sont en majorité du même côté du graphique que l'ensemble d'étalonnage. L'axe 1 discrimine presque les ordures ménagères des autres échantillons.

Le premier vecteur propre est représenté Figure 41.

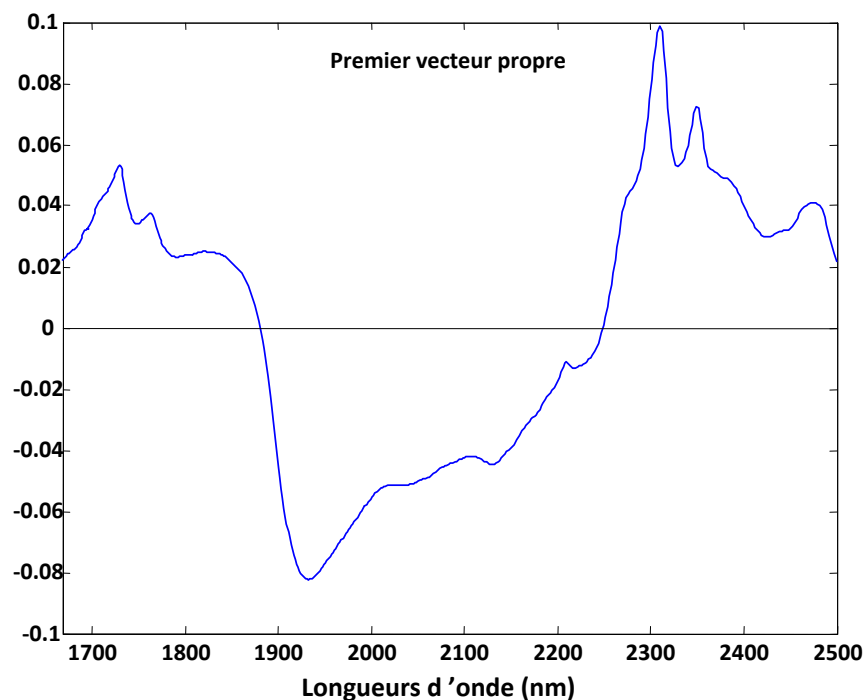


Figure 41 : Premier vecteur propre de l'ACP.

Ce premier vecteur propre semble opposer l'absorbance à 2308 nm à la large vallée débutant vers 1900 nm et finissant pas avant 2200 nm. Comme vu dans la précédente partie, l'absorbance à 2308 nm correspond à la liaison $C-H$, contenue principalement dans les lipides, les glucides et les plastiques. Au sein de la large vallée du premier vecteur propre, est inclut la vallée à 1930 nm qui représente l'absorbance de la fonction alcool $O-H$ de l'amidon et de la cellulose, mais également le pic à 2100 nm (glucides) et tous les pics intermédiaires. Un groupe d'algues est situé aux valeurs extrêmes négatives de la première composante. Certaines algues peuvent être

très riches en glucides (Sialve et al., 2009). La figure 42a, présente les spectres proche infrarouge de ce groupe d'algue.

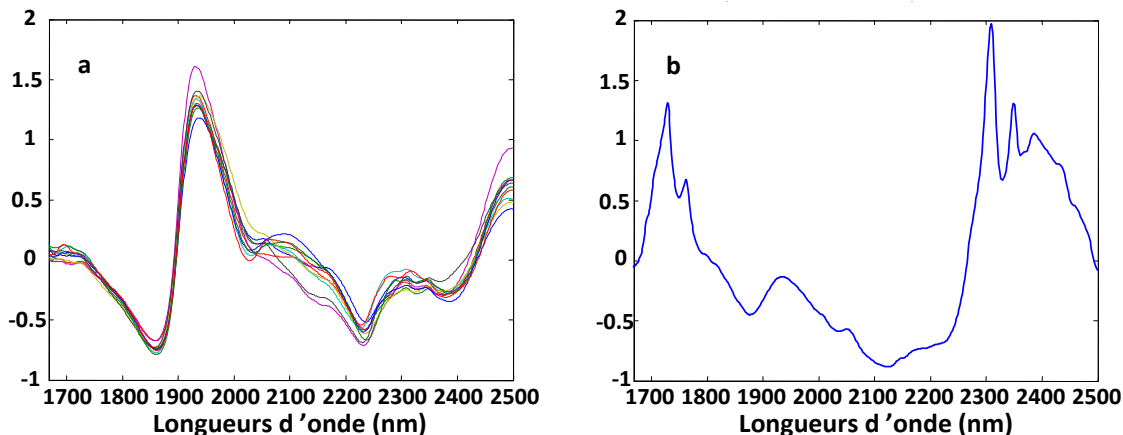


Figure 42 : Spectre proche infrarouge d'algues (a), présentant un pic important à 1930 nm et d'une boue grasseuse (b), présentant des pics d'absorbance de la liaison $C-H$.

On voit clairement le pic à 1930 nm et l'absence de pic à 2308 nm. Le pic à 2100 nm est également visible dans les algues, ce qui montre la présence de glucides. Il est ensuite plus difficile d'attribuer le reste de la vallée, après 1930 nm.

Aux valeurs positives extrêmes de la première composante principale sont présents des composés ayant beaucoup de liaison $C-H$, donc par exemple fortement concentrés en lipides comme la boue grasseuse (Figure 42b) ou contenant beaucoup de matériaux plastiques (l'échantillon ayant les plus fortes quantités de plastiques dans l'ensemble d'étalonnage est situé à l'extrême droite de l'axe 1). Contenant majoritairement des fibres (cellulose, amidon) et peu de lipides, les déchets d'origine végétale sont situés normalement du côté gauche de la première composante principale (Figure 40). La première composante oppose donc les deux composés ayant majoritairement un impact favorable au BMP, les lipides et les glucides.

Le second vecteur propre est présenté Figure 43.

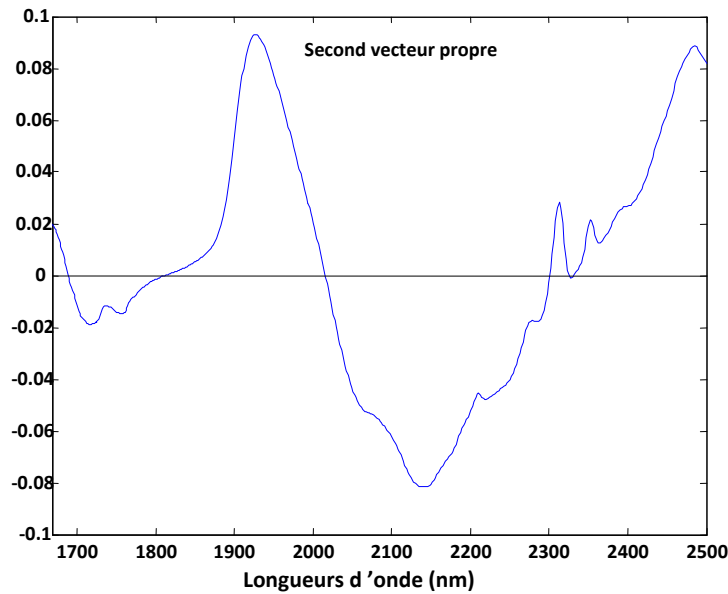


Figure 43 : Second vecteur propre de l'ACP.

L'information contenue dans ce second vecteur propre est plus difficilement identifiable. Une nouvelle fois, le pic à 1930 nm explique une part de la variabilité sur la seconde composante principale. Egalement visible, le pic à 2485 nm, précédemment attribué à la combinaison $C-H/C-C$ de l'amidon et de la cellulose. Ils sont tous deux opposés à la vallée située à 2140 nm, attribuée aux cycles aromatiques.

La troisième composante est plus intéressante. Elle sépare les échantillons en fonction des fibres et des composés protéo-lipidiques (Figure 44). En effet, la majorité des échantillons d'origine végétale et les pailles est située dans les valeurs négatives de cette composante. Les deux échantillons « simples » à gauche sont du papier et du carton, composés à teneur importante en fibres. Les déchets ménagers également constitués de carton et de papiers sont également majoritaires à gauche. Les échantillons à droite de la troisième composante principale sont des échantillons qui auraient de fortes concentrations en protéines et en lipides (déchets carnés, jus de cuisson, boue graisseuse). Le troisième vecteur propre (Figure 45) oppose les pics des fibres (cellulose et amidon : 2100 nm, 2485 nm et les composés à concentrations importantes en protéines et en lipides (2000 nm et 2308 nm).

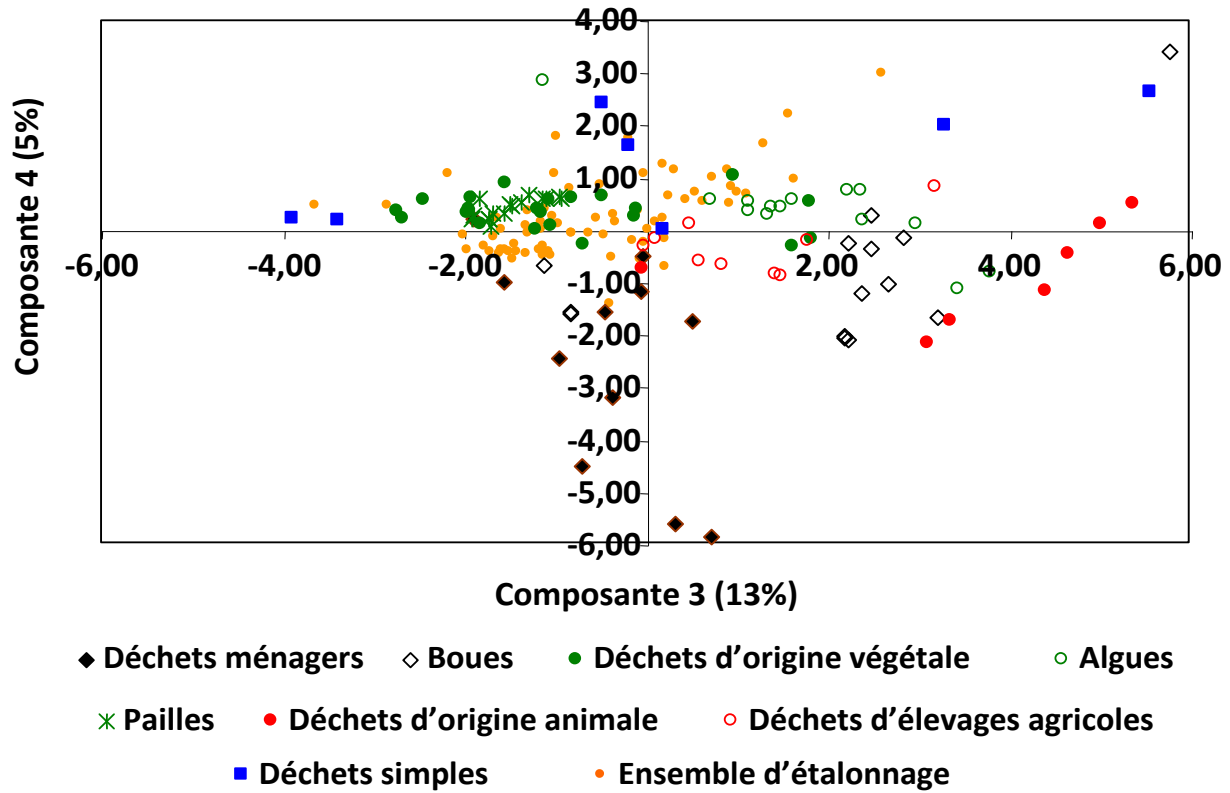


Figure 44 : Troisième et quatrième composantes principales de l'ACP sur les 174 échantillons de déchets.

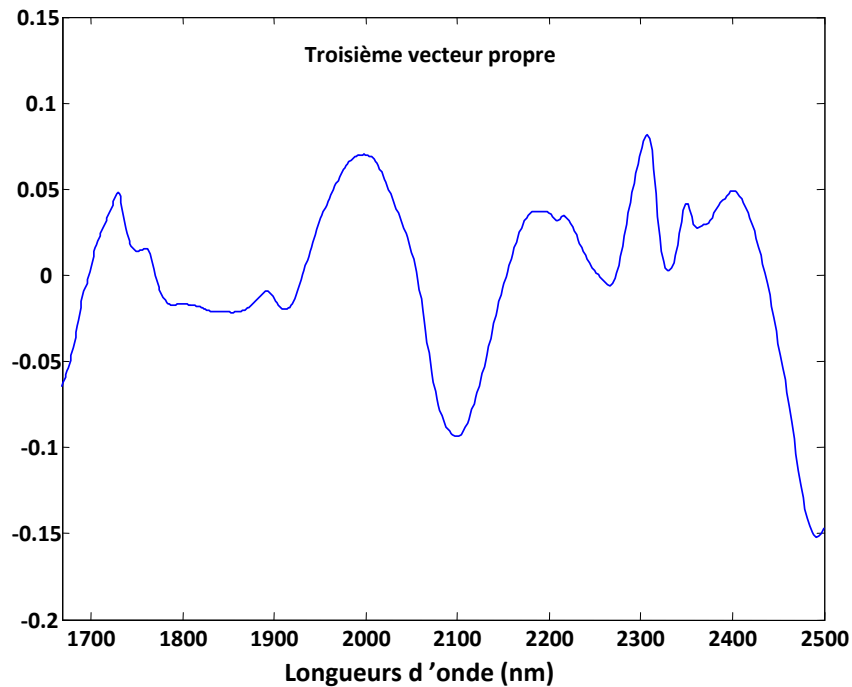


Figure 45 : Troisième vecteur propre de l'ACP.

V.3.2.3. Conclusion de l'analyse ACP

L'analyse ACP de l'ensemble des échantillons permet de confirmer que les échantillons de l'ensemble de test sont trop différents de ceux de l'ensemble d'étalonnage pour être correctement prédits.

Sur l'axe 1 de l'ACP (58% de la variance totale), les échantillons d'étalonnage sont placés symétriquement aux échantillons de test par rapport au 0. La première composante semble opposer principalement, l'absorbance des liaisons attribuées à la cellulose et l'amidon (massif entre 1930 nm et 2100nm) à l'absorbance de la liaison C-H, à 2308 nm, attribuée principalement aux lipides et plastiques des ordures ménagères.

L'axe 2 (18% de la variance totale) n'a pu être déterminé avec certitude, il opposerait également l'absorbance à 1930 nm à l'absorbance à 2140 nm, attribué aux cycles benzéniques.

Enfin, l'axe 3 (13% de la variance totale), opposerait également les composés fibreux (cellulose et amidon) absorbant à 2100 nm et 2490 nm, aux composés protéiques et lipidiques (2000 nm et 2308 nm). Les échantillons protéo-lipidiques ne sont pas du tout représentés dans l'ensemble d'étalonnage et peu de pic de protéines ont en effet été identifiés dans l'ensemble d'étalonnage.

L'ensemble d'étalonnage n'a pas une variabilité suffisante pour espérer prédire correctement certains types de déchets. Le modèle n'est donc pas assez robuste vis à vis de la nature des échantillons. C'est pourquoi nous avons envisagé la construction d'un modèle global.

V.3.3. *Modèle global de régression PLS*

Le Tableau 15 présente les résultats d'étalonnage et de validation pour les différents prétraitements.

	Var. Latente	SECV ^a	SEC ^a	R ² etal.	RMSEP ^a	Biais ^a	R ² test.
Abs	8	69	59	0,65	71	16	0,5
SNV	19	56	38	0,79	90	16	0,45
SNV(1645 – 2500)	7	89	72	0,43	89	13	0,3
SNV – Dt	18	58	39	0,77	84	18	0,47
Sav.(7,2,1)	11	73	50	0,62	81	9	0,47
Sav.(7,3,2)	9	70	45	0,64	73	1	0,49
SNV – Sav.(31,2,1)	18	58	51	0,75	87	21	0,47
SNV – Sav.(11,3,2)	8	65	49	0,7	70	8	0,52
SNV(1645-2500) Sav.(13,3,2)	9	63	47	0,71	71	10	0,53

Tableau 15 : Résultat d'étalonnage et de prédiction des modèles « globaux ».

a : exprimé en ml CH₄.g⁻¹ MV ; Var : Variable ; Dt : Detrend ; Sav : Dérivée de Savitsky et Golay. Les nombres entre parenthèses sont respectivement, la fenêtre de dérivation, le degré du polynome et l'ordre de la dérivée.

Le meilleur modèle en cross-validation est obtenu après les prétraitements SNV et la dérivée seconde (11,3,2) (SECV : 65 ml CH₄.g⁻¹ MV, R²_{et} : 0.7). Ce modèle obtient également le meilleur résultat en validation (RMSEP : 70 ml.CH₄.g⁻¹ MV, R²_{val} = 0.52). Les résultats de la régression PLS sont présentés Figure 46.

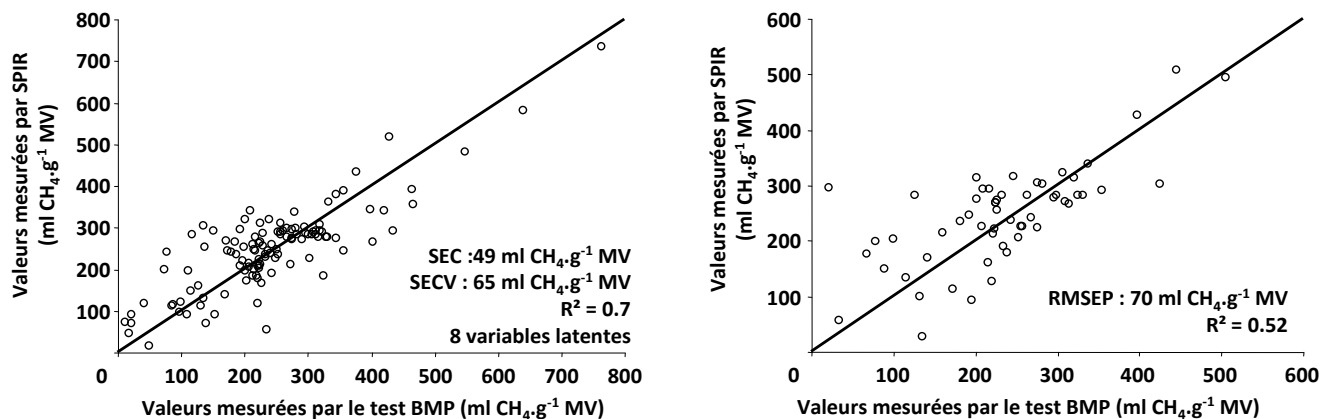


Figure 46 : Régression PLS sur les ensembles d'étalonnage et de validation.

Les performances du modèle sans prétraitement ne sont pas très éloignées du meilleur modèle. L'utilisation des prétraitements a amélioré légèrement les résultats en diminuant la variabilité due aux effets de diffusion et surtout en diminuant le biais. La variabilité des spectres après prétraitement est, en théorie, uniquement due à la composition chimique.

Les modèles avec uniquement SNV et Detrend nécessitent trop de variables latentes. D'après Bertrand et al. (2006), le nombre de dimension PLS conditionne le rapport entre la précision et la robustesse. Plus le nombre de variable latente est grand, meilleure est la précision, mais plus faible est la robustesse et inversement.

Le meilleur modèle en cross-validation, a donc été obtenu lors de l'utilisation d'une taille de fenêtre restreinte dans la dérivée seconde (SNV – Savgol(11,3,2)), ce qui peut engendrer du bruit au niveau des spectres et donc également au niveau des b coefficients.

Comparativement au chapitre 3, les erreurs obtenues ici (RMSEP = 70 ml.CH₄.g⁻¹ MV et $R^2 = 0,56$), sont deux fois et demi plus importantes puisque nous avons un RMSEP de 29 ml.CH₄.g⁻¹ MV ($R^2 = 0,76$), sur des échantillons homogènes. L'augmentation de l'hétérogénéité a diminué la précision du modèle global PLS (Bertrand et al., 2006). Cette diminution de la précision due à l'accroissement de l'hétérogénéité des échantillons n'est pas inhérente à la spectrophotométrie PIR : elle était également observée pour un modèle de prédiction de BMP basée sur la caractérisation chimique des échantillons (Gunaseelan (2007 ; 2009) ; Schievano et al.(2009)).

Ces nouveaux résultats montrent que le modèle ainsi créé n'est pas satisfaisant pour répondre correctement à une bonne prédiction du potentiel méthane.

Le manque de précision de ce modèle est dû au problème majeur de ce type de modèle PLS en mode global. Son utilisation émet l'hypothèse d'une linéarité entre les variables prédictives (les absorbances des spectres PIR) et la variable à prédire (le BMP). Cette linéarité existe également pour l'ensemble des composés de la matière organique. Or, le BMP est obtenu par un processus biologique complexe, qui n'est sûrement pas un processus complètement linéaire. De plus, des phénomènes d'interactions ou d'inhibitions peuvent se produire lors de fortes concentration en protéines ou en lipides. L'hydrolyse, de quantités importantes de protéines, engendre l'accumulation de NH₄⁺ inhibant la phase de méthanogène par augmentation du pH (Hansen et al., 1998). Les fortes quantités de lipides peuvent créer une barrière entre les micro-organismes et le substrat, jusqu'à endommager la membrane bactérienne (Mata-Alvarez, 2003 ; Cavaleiro et al., 2008). Des méthodes non linéaires ou locales seraient peut être mieux appropriées.

Les plus prometteuses sont certainement les méthodes locales, telles que LWR (Locally Weighted Regression) (Naes et al. 1990), LOCAL (Shenk et al., 1997). Elles sont de plus en plus utilisées, notamment en analyse des sols, où, la problématique de l'hétérogénéité des échantillons de la base d'étalonnage est également présente (Brown et al., 2006). Ces méthodes utilisent la

notion de noyau (*Kernel* en anglais) et se base sur l'hypothèse que, dans un espace de dimension n , des spectres proches dans cet espace auraient des valeurs y semblables.

Ces méthodes reposent sur le principe des plus proches voisins : un nouvel échantillon sera prédit à partir d'un ensemble d'échantillons situés à moins d'une distance d de lui. Une valeur de d élevée sélectionnera des échantillons très éloignés, et pouvant n'avoir que peu de rapport avec l'échantillon inconnu. Une petite valeur de d ne prendra en compte que peu d'échantillons et pourra donner une trop grande importance à chaque échantillon (Bertrand et al. 2006). Un poids peut être attribué à chaque échantillon en fonction de sa distance à l'échantillon inconnu.

Zamoras-Rojas et al. (2010) comparent trois méthodes locales pour la prédiction de la composition d'aliments pour animaux par spectrophotométrie PIR. Ils utilisent une base de départ de 6018 échantillons hétérogènes pour l'ensemble d'étalonnage et 1505 autres pour l'ensemble de test. Ils montrent que suivant la méthode, 5, 100 ou 300 échantillons sont sélectionnés en tant que proches voisins pour obtenir 100% de bonne prédiction des échantillons de test.

Un effort est donc à fournir pour élaborer une base d'échantillon plus importante afin qu'une variabilité importante soit atteinte. Le nombre d'échantillons d'une telle base doit être déterminé, en fonction de la variabilité que l'on peut attendre sur ce genre d'échantillon. Cette condition étant remplie, des méthodes de modélisations locales pourront être utilisées.

V.4. Conclusion

Dans cette partie, nous avons dans un premier temps utilisé le modèle précédemment établi pour prédire un ensemble de 100 échantillons différents de ceux d'étalonnage quant à leurs natures et de leurs origines. Les résultats montrent que la majorité des échantillons ne se trouvent pas dans les limites d'incertitudes du modèle ($\pm 2 \text{ SECV} = \pm 61 \text{ ml CH}_4 \cdot \text{g}^{-1} \text{ MV}$). Nous en avons conclu que le modèle n'était pas robuste vis à vis de la nature des échantillons. Dans un second temps, nous avons, à l'aide d'une analyse multi-variée, montré l'origine de la variabilité entre les deux ensembles d'échantillons. Il en ressort principalement que l'ensemble d'étalonnage n'est pas assez représentatif de la variabilité qui existe au sein de l'ensemble de test. En effet, certains déchets, tels que les déchets végétaux, ne présentent que peu de lipides, ou à l'inverse, tels que les déchets carnés, ne sont composés que de lipides et protéines et ne présentent pas les pics

caractéristiques des fibres (cellulose, lignine) et sur l'amidon. Or le modèle est principalement basé sur ces composés. Une solution à ce problème de robustesse a été envisagée : la construction d'un modèle global. Plusieurs combinaisons de prétraitement ont été testées pour construire ce modèle. Ce modèle donne des résultats deux fois moins performants en SECV que le modèle sur échantillons homogènes, mais les performances sont équivalentes à celles des prédictions du BMP à l'aide de la caractérisation chimique. Un RMSEP de 70 ml CH₄.g⁻¹ MV, avec un R² de 0,52 ont été obtenus. Ces résultats sont loin d'être satisfaisants. La construction d'un modèle basé sur un ensemble d'échantillons hétérogènes vis à vis de leurs natures et de leurs origines a accru la robustesse mais en contrepartie a diminué la précision, et ceci en deçà de ce qui est acceptable au niveau industriel. Une autre solution serait donc de construire des modèles locaux mais ceux-ci nécessitent des lots d'étalonnage encore plus exhaustifs.

VI.
Conclusion et perspectives

La digestion anaérobie répond pleinement aux problèmes actuels de développement durable. Elle permet de traiter les déchets, voire d'en tirer un profit à travers la production d'énergie. Cela amène de nombreux industriels et collectivités à développer des installations de méthanisation. L'optimisation de la production de méthane est un facteur clé à prendre en compte pour la rentabilité de l'exploitation. Une voie d'optimisation est la sélection, en amont, des déchets à potentiel méthane important, cette sélection devant se faire le plus rapidement possible, afin de ne pas avoir à gérer de stock de déchets.

La problématique de la thèse a donc consisté à élaborer une méthode de caractérisation rapide de la matière organique de déchets solides afin de prédire plus rapidement et plus facilement leur potentiel méthane.

Notre démarche s'est donc articulée en quatre points.

Le premier objectif de cette thèse a été d'identifier une méthode rapide d'analyse de la matière organique permettant d'en extraire l'information nécessaire à l'estimation du potentiel méthane. Ce premier objectif a été mené au travers d'une étude bibliographique. Le test BMP, permettant la mesure du potentiel méthane, a été expliqué et ses avantages et inconvénients discutés. Il en ressort qu'il n'est pas adapté à une optimisation et à une prise de décision rapide quant au choix des déchets à méthaniser. Ainsi, des tests plus rapides ont été étudiés, tels que des tests respirométriques aérobies ou des tests enzymatiques. Le lien entre la valeur BMP et la matière organique des déchets a également été abordé : les nombreuses études faites sur le sujet montrent que la valeur BMP peut être prédite à partir de la connaissance de la « formulation » de la matière organique des déchets (soit en connaissant les composés principaux en présence, soit à partir de la composition élémentaire). Les méthodes de caractérisation chimique des déchets sont cependant fastidieuses à mettre en œuvre. Enfin, les méthodes spectroscopiques (UV/NIR/MIR) ont été envisagées. Elles ont l'avantage d'analyser rapidement et globalement la matière organique. A la lumière de cette étude bibliographique, il apparaît que la spectroscopie proche infrarouge (SPIR) possède le potentiel théorique le plus adapté pour la prédiction du potentiel méthane de déchets solides. En effet, la SPIR nécessite peu de préparation de l'échantillon pour analyser rapidement et globalement toute la matière organique (glucides, lipides, protéines, fibres, matières organiques inertes). La possibilité qu'offre cette gamme de longueur d'onde

(1000 – 2500 nm) de réaliser la mesure par fibre optique est un atout supplémentaire pour une utilisation industrielle. Ce travail bibliographique a fait l'objet d'une publication sous forme d'une synthèse des « méthodes alternatives pour la mesure de la biodégradabilité » dans *Process Biochemistry* (Lesteur et al., 2010).

Le second objectif a été alors de montrer la faisabilité de la prédiction du potentiel méthane par spectroscopie proche infrarouge. Pour cette première approche, nous nous sommes placés dans des conditions simplifiées, avec l'utilisation d'un seul type de déchet, des échantillons broyés et lyophilisés d'ordures ménagères. Un modèle d'étalonnage a été créé et validé, sur un ensemble homogène d'échantillons de déchets ménagers. Les résultats obtenus (RMSEP = 29 ml CH₄.g⁻¹ MV – pour une moyenne de 234 ml CH₄.g⁻¹ MV - et un R² = 0,67) montrent clairement le potentiel de la SPIR pour la prédiction du potentiel méthane. Les incertitudes de la méthode SPIR et du test BMP ont été comparées. L'incertitude liée à la mesure SPIR est semblable à celle du test BMP, respectivement 61 ml CH₄.g⁻¹ MV et 69 ml CH₄.g⁻¹ MV. Une comparaison des précisions a été effectuée entre la méthode SPIR et celle basée sur la caractérisation chimique des déchets. Dans une situation identique d'un ensemble homogène de déchet, nos résultats sont meilleurs (RMSEP de 29 ml CH₄.g⁻¹ MV) que ceux de la méthode basée sur la caractérisation chimique (RMSEP de 91 ml CH₄.g⁻¹ MV) (Nallathambi Gunaseelan, 2007). La SPIR a l'avantage supplémentaire de pouvoir prédire le BMP en moins de temps. Nous avons également montré, qu'en restreignant la gamme de variables (1668 – 2500 nm), la variabilité inter-spectre diminue et les modèles en étalonnage en sont améliorés. Ce chapitre a également fait l'objet d'un article scientifique, étudiant les « premières étapes vers une prédiction rapide du potentiel méthane de déchets solides par spectroscopie proche infrarouge », publié dans *Bioresource Technology* (Lesteur et al. ; 2010).

Ce modèle a ensuite été étudié plus en détail, avec l'objectif de comprendre l'origine de la biodégradabilité des déchets étudiés. Les *b* coefficients PLS ont été analysés. Les *b* coefficients positifs, attribués aux composés biodégradables, correspondent majoritairement aux glucides, notamment la cellulose et l'amidon (1930 nm, 2076 nm, 2106 nm, 2270 nm, 2484 nm et dans une moindre mesure, 2308 nm). Nous avons également montré qu'au sein des ordures ménagères, les lipides (2308 nm et 2347 nm) jouent un rôle important sur la valeur BMP. Les *b* coefficients négatifs sont attribués aux composés non biodégradables, principalement la lignine

(principalement 1910 nm et 2380 nm) ou plus généralement les composés à nombreux cycles aromatiques (1675 nm, la bande 2060 – 2150 nm, 2144 nm, 2189 nm et 2230 nm) et les plastiques (1910 nm, 2144 nm, 2189 nm, 2308 nm et 2444 nm), dont on rencontre des résidus dans les ordures ménagères. Afin de compléter cette étude et de la valider les algorithmes de sélection de variables (algorithme génétique et CovSel) ont été utilisés. Les résultats obtenus confirment les conclusions de l'étude des b coefficients PLS. Ce sont bien les glucides qui ont le rôle majeur dans la valeur BMP. En effet, le modèle à base d'algorithmes génétiques à 28 variables dit « modèle suggéré AG2 » donne des performances identiques en étalonnage sans besoin d'inclure les longueurs d'onde relatives aux lipides (2308 nm et 2347 nm). De meilleurs résultats d'étalonnage sont obtenus avec CovSel. Cet algorithme a, quant à lui, sélectionné les longueurs d'onde des lipides. Mais il leur a attribué un coefficient positif significativement plus faible que dans les b coefficients PLS. Le modèle prend donc en compte les lipides, mais en limitant l'impact des plastiques. Cette partie fait l'objet d'un article en préparation.

Enfin, le quatrième objectif a été d'étudier la robustesse du modèle vis à vis de la nature des échantillons. Pour cela, notre démarche aura été de tester le modèle à l'aide d'échantillons de nature et d'origines différentes. Un ensemble de 100 échantillons a été construit à partir de déchets d'origines végétale, animale, agricole (lisier et fumiers, pailles), ménagère (d'origines différentes que ceux du modèle) et différentes espèces d'algues. Avec un RMSEP de 177 ml $\text{CH}_4 \cdot \text{g}^{-1}$ MV, cette validation montre le manque de robustesse du modèle vis à vis de la nature des échantillons. Seuls les échantillons de déchets ménagers sont, à 73%, correctement prédits. Afin de comprendre ce manque de robustesse au sein du modèle, une analyse ACP a été réalisée. Elle met en évidence que l'ensemble des 100 échantillons présentent une variabilité trop importante, comparativement à l'ensemble d'étalonnage. En effet, certains déchets possèdent des teneurs en lipides ou en protéines nettement plus élevées que l'ensemble d'étalonnage. De plus, ces mêmes échantillons ne présentent ni de composantes fibreuses (cellulose/lignine) ni d'amidon. Le modèle utilisé semble satisfaisant pour les déchets ménagers mais ne montre pas assez de variabilité pour prendre en compte d'autres types de déchets. Afin de répondre à ce problème de robustesse, une première solution fut testée. Un modèle exhaustif, dit « modèle global » a été mis en œuvre afin de répondre à ce problème de robustesse. Incluant 117 échantillons hétérogènes en étalonnage, il a été testé sur les 57 déchets restants. Ce modèle génère de meilleurs résultats (RMSEP = 70 ml $\text{CH}_4 \cdot \text{g}^{-1}$ MV et un $R^2 = 0,52$) que le précédent, mais reste insatisfaisant. En

effet, dans le cas d'un modèle PLS global, l'utilisation d'échantillons d'origines et de natures différentes accroît sa robustesse, mais diminue sa précision.

Ces travaux de thèse ont montré pour la première fois les potentialités importantes de la spectroscopie proche infrarouge pour la prédiction du potentiel méthane des déchets et ouvrent de futurs axes de recherches.

Le premier, incontestable, est de solutionner le problème lié à l'hétérogénéité quasi infinie de la matrice déchet. L'utilisation d'un modèle global n'est assurément pas la solution. Cependant quelques pistes ont pu être suggérées, notamment l'utilisation de modèles locaux. Ces modèles permettent de diminuer ce problème d'hétérogénéité des échantillons. Cependant, pour le fonctionnement de ces modèles, une base très importante d'échantillons est nécessaire.

Un second axe serait, d'ajouter à la valeur BMP, la prise en compte d'autres critères de la caractérisation des déchets. En effet, si la connaissance de la valeur BMP d'un déchet est une condition nécessaire, elle n'est pas suffisante. La vitesse de dégradation, liée à la qualité des différents glucides et lipides, est un facteur essentiel de l'optimisation dynamique des méthaniseurs. De même, les effets de ces différents contributeurs au BMP ne sont pas linéaires. Un déchet trop riche en glucides simples, conduira à une production rapide d'acides gras volatils. Un risque d'acidification du digesteur par baisse du pH entraînera alors une baisse de la productivité en méthane. De même, une trop forte proportion de protéines pourrait engendrer un risque d'obtenir une concentration en ammoniacque inhibitrice de la production de méthane. Enfin, une proportion importante de lipides risque d'entraîner une accumulation de propionate, néfaste aux étapes de la digestion anaérobie. Une caractérisation fine des déchets permettrait d'anticiper ces événements.

Toujours dans le but de connaître la vitesse de dégradation de déchets solides, des recherches pourraient être menées sur l'utilisation de processus d'oxydation avancées, tel que l'ozonation ou la photo-oxydation. En effet, des études ont déjà montré l'intérêt de la photo-oxydation, comme outil d'étude de la vitesse de dégradation lors de tests aérobies (Castillo et al., 1999).

Un troisième axe pourrait donc être le suivi in-situ de la méthanisation. Plusieurs paramètres seraient intéressants à suivre. Notamment, la mesure en ligne par SPIR de la teneur en eau dans les digesteurs par voie sèche en améliorerait son contrôle et sa régulation. Egalement la mesure

in-situ des métabolites de la réaction, tels les acides gras volatiles, l'ammoniaque ou les acides gras longues chaînes, serait utile aux contrôles de l'optimisation de la production de méthane.

Références Bibliographiques

ADEME. (2009). "Campagne nationale de caractérisation des ordures ménagères, résultats année 2007". *Agence de l'Environnement et de la Maitrise de l'Energie*.

Adesogan, A.T., Owen, E. and Givens, D.I. (1998). "Prediction of the in vivo digestibility of whole crop wheat from in vitro digestibility, chemical composition, in situ rumen degradability, in vitro gas production and near infrared reflectance spectroscopy". *Animal Feed Science and Technology*. 74(3): 259-272.

Albrecht, R., Joffre, R., Gros, R., Le Petit, J., Terrom, G. and Perissol, C. (2008). "Efficiency of near-infrared reflectance spectroscopy to assess and predict the stage of transformation of organic matter in the composting process". *Bioresource Technology*. 99(2): 448-455.

Alvarez, J.A., Otero, L. and Lema, J.M. (In press). "A methodology for optimising feed composition for anaerobic co-digestion of agro-industrial wastes". *Bioresource Technology*. 101(4): 1153-1158.

Anderson, J. (2003). "The environmental benefits of water recycling and reuse". *Water Science and Technology*. 3(4): 1-10.

Andres, S., Calleja, A., Lopez, S., Gonzalez, J.S., Rodriguez, P.L. and Giraldez, F.J. (2005). "Prediction of gas production kinetic parameters of forages by chemical composition and near infrared reflectance spectroscopy". *Animal Feed Science and Technology*. 123: 487-499.

Angelidaki, I. and Ahring, B.K. (1992). "Effects of Free Long-Chain Fatty-Acids on Thermophilic Anaerobic-Digestion". *Applied Microbiology and Biotechnology*. 37(6): 808-812.

Angelidaki, I. and W. Sanders (2004). "Assessment of the anaerobic biodegradability of macropollutants." *Reviews in Environmental Science and Bio/Technology*. 3(2): 117-129.

Angelidaki, I., Alves, M., Bolzonella, D., Borzacconi, L., Campos, J.L., Guwy, A.J., Kalyuzhnyi, S., Jenicek, P. and van Lier, J.B. (2009). "Defining the biomethane potential (BMP) of solid organic wastes and energy crops: a proposed protocol for batch assays". *Water Science and Technology*. 59(5): 927-934.

Bailleres, H., Davrieux, F. and Pichavant, F.H. (2002). "Near infrared analysis as a tool for rapid screening of some major wood characteristics in a eucalyptus breeding program". *Annals of Forest Science*. 59(5-6): 479-490.

Barnes R., Dhanoa M., and Lister J. (1989). "Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra". *Applied Spectroscopy*. 43(5): 772-777.

- Barrena, R., d'Imporzano, G., Ponsá, S., Gea, T., Artola, A., Vázquez, F., Sánchez, A. and Adani, F.** (2009). "In search of a reliable technique for the determination of the biological stability of the organic matter in the mechanical-biological treated waste". *Journal of Hazardous Materials*. 162(2-3): 1065-1072.
- Batstone, D.J., Keller, J., Angelidaki, I., Kalyuzhnyi, S., Pavlostathis, S.G., Rozzi, A., Sanders, W., Siegrist, H. and Vavilin, V.A.** (2002). "The IWA anaerobic digestion model N°1 (ADM1)". *Water Science and Technology*. 45(10): 65-73.
- Bellon-Maurel, V., Orliac, O. and Christen, P.** (2003). "Sensors and measurements in solid state fermentation: a review". *Process Biochemistry*. 38(6): 881-896.
- Bellon-Maurel, V. and Dubois, J.** (2009) "Near-Infrared Hyperspectral Imaging in Food and Agricultural Science in Infrared and Raman Spectroscopic Imaging. In: Infrared and Raman Spectroscopic Imaging". Salzer, R. and Siesler, H.W., eds. pp. 259-292, Wiley-VCH, Allemagne.
- Beltrán, F.J., García-Araya, J.F., Frades, J., Álvarez, P. and Gimeno, O.** (1999). "Effects of single and combined ozonation with hydrogen peroxide or UV radiation on the chemical degradation and biodegradability of debittering table olive industrial wastewaters". *Water Research*. 33(3): 723-732.
- Ben-Dor, E., Inbar, Y. and Chen, Y.** (1997). "The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400-2500 nm) during a controlled decomposition process". *Remote Sensing of Environment*. 61(1): 1-15.
- Bertrand D., Dufour E.** (2006). "La spectroscopie infrarouge et ses applications analytiques". Editions Tech&Doc, Paris, 2nd édition, Lavoisier. Collection Sciences & Techniques Agroalimentaire, 660p.
- Brinkmann, K., Blaschke, L. and Polle, A.** (2002). "Comparison of different methods for lignin determination as a basis for calibration of near-infrared reflectance spectroscopy and implications of lignoproteins". *Journal of Chemical Ecology*. 28(12): 2483-2501.
- Brookman, S.K.E.** (1997). "Estimation of biochemical oxygen demand in slurry and effluents using ultra-violet spectrophotometry". *Water Research*. 31(2): 372-374.
- Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne Mays, M. and Reinsch, T.G.** (2006). "Global soil characterization with VNIR diffuse reflectance spectroscopy". *Geoderma*. 132(3-4): 273-290.
- Bruno-Soares, A.M., Murray, I., Paterson, R.M. and Abreu, J.M.F.** (1998). "Use of near infrared reflectance spectroscopy (NIRS) for the prediction of the chemical composition and nutritional attributes of green crop cereals". *Animal Feed Science and Technology*. 75(1): 15-25.
- Buffiere, P., Loisel, D., Bernet, N. and Delgenes, J.P.** (2006). "Towards new indicators for the prediction of solid waste anaerobic digestion properties". *Water Science and Technology*. 53(8): 233-241.

Buffiere, P., Mirquez, L.D., Steyer, J.P., Bernet, N., Delgenes, J.P. (2008). "Anaerobic digestion of solid wastes needs research to face an increasing industrial success". *International Journal of Chemical Reactor Engineering*. 6, A94.

Disponible en ligne : <http://www.bepress.com/ijcre/vol6/A94>

Buranov, A.U. and Mazza, G. (2008). "Lignin in straw of herbaceous crops". *Industrial Crops and Products*. 28(3): 237-259.

Buswell, E.G., Neave, S.L. (1930). "Laboratory studies of sludge digestion. Illinois Division of State Water Survey, Bulletin N°30.

Calmon, A., Silvestre, F., Bellon-Maurel, V., Roger, J.M. and Feuilloley, P. (1999). "Modelling easily biodegradability of materials in liquid medium-relationship between structure and biodegradability". *Journal of Environmental Polymer Degradation*. 7(3): 135-144.

Castaldi, P., Alberti, G., Merella, R. and Melis, P. (2005). "Study of the organic matter evolution during municipal solid waste composting aimed at identifying suitable parameters for the evaluation of compost maturity". *Waste Management Special Issue - VARIREI 2003*. 25(2): 209-213.

Castillo, L., El Khorassani, H., Trebuchon, P. and Thomas, O. (1999). "UV treatability test for chemical and petrochemical wastewater". *Water Science and Technology*. 39(10-11): 17-23.

Cavaleiro, A.J., Pereira, M.A. and Alves, M. (2008). "Enhancement of methane production from long chain fatty acid based effluents". *Bioresource Technology*. 99(10): 4086-4095.

Chandler, J.A., Jewell, W.J., Gossett, J.M., Vansoest, P.J. and Robertson, J.B. (1980). "Predicting Methane Fermentation Biodegradability". *Biotechnology and Bioengineering*. 22 : 93-107.

Chen, T.-H. and Hashimoto, A.G. (1996). "Effects of pH and substrate:inoculum ratio on batch methane fermentation". *Bioresource Technology*. 56(2-3): 179-186.

Chen, Y., Cheng, J.J. and Creamer, K.S. (2008). "Inhibition of anaerobic digestion process: A review". *Bioresource Technology*. 99(10): 4044-4064.

Cho, J.K., Park, S.C. and Chang, H.N. (1995). "Biochemical methane potential and solid state anaerobic digestion of Korean food wastes". *Bioresource Technology*. 52(3): 245-253.

Chynoweth, D.P., Turick, C.E., Owens, J.M., Jerger, D.E. and Peck, M.W. (1993). "Biochemical methane potential of biomass and waste feedstocks". *Biomass and Bioenergy*. 5(1): 95-111.

Cossu, R. and Raga, R. (2008). "Test methods for assessing the biological stability of biodegradable waste". *Waste Management*. 28(2): 381-388.

Dardenne, P. (2010). "Some considerations about NIR spectroscopy: Closing speech at NIR-2009". *NIR news*. 21(1): 8–14.

Damelon, C., Thiénard, H. (2005). "Les énergies renouvelables en France 1970 - 2003". pp. 53, *Observatoire de L'Energie*.

Davidsson, A., Gruvberger, C., Christensen, T.H., Hansen, T.L. and Jansen, J.I.C. (2007). "Methane yield in source-sorted organic fraction of municipal solid waste". *Waste Management*. 27(3): 406-414.

De Boever, J.L., Cottyn, B.G., De Brabander, D.L., Vanacker, J.M. and Boucqué, C.V. (1996). "Prediction of the feeding value of grass silages by chemical parameters, in vitro digestibility and near-infrared reflectance spectroscopy". *Animal Feed Science and Technology*. 60(1-2): 103-115.

Domeizel, M., Khalil, A. and Prudent, P. (2004). "UV spectroscopy: a tool for monitoring humification and for proposing an index of the maturity of compost". *Bioresource Technology*. 94(2): 177-184.

Eleazer, W.E., Odle, W.S., Wang, Y.-S. and Barlaz, M.A. (1997). "Biodegradability of Municipal Solid Waste Components in Laboratory-Scale Landfills". *Environmental Science & Technology* 31(3): 911-917.

El Hajjouji, H., Fakharedine, N., Ait Baddi, G., Winterton, P., Bailly, J.R., Revel, J.C. and Hafidi, M. (2007). "Treatment of olive mill waste-water by aerobic biodegradation: An analytical study using gel permeation chromatography, ultraviolet-visible and Fourier transform infrared spectroscopy". *Bioresource Technology*. 98(18): 3513-3520.

El Khorassani, H., Trebuchon, P., Bitar, H. and Thomas, O. (1999). "A simple UV spectrophotometric procedure for the survey of industrial sewage system". *Water Science and Technology*. 39(10-11): 77-82.

EurObserv'ER (2009). "Etat des énergies renouvelables en Europe". pp.171. <http://www.eurobserv-er.org/>

Faber, N.M., Schreutelkamp, F.H. and Vedder, H.W. (2004). "Estimation of prediction uncertainty for a multivariate calibration model". *Spectroscopy Europe* 16, 17-20

Feller, C., Jeanson, P., Giummelly, P. and Bonaly, P. (1991). "Comparaison de différentes méthodes d'hydrolyse acide en vue du dosage des glucides totaux dans les sols." *Science du Sol*. 29(1): 13-22.

Fogelman, S., Zhao, H.J. and Blumenstein, M. (2006). "A rapid analytical method for predicting the oxygen demand of wastewater". *Analytical and Bioanalytical Chemistry*. 386(6): 1773-1779.

Garimella Purna, S.K., Prow, L.A. and Metzger, L.E. (2005). "Utilization of front-face fluorescence spectroscopy for analysis of process cheese functionality." *Journal of Dairy Science*. 88 : 470-477.

Garrido-Vidal, D., Esteban-Diez, I., Perez-del-Notario, N., Gonzalez-Saiz, J.M. and Pizarro, C. (2004). "On-line monitoring of kinetic and sensory parameters in acetic fermentation by near infrared spectroscopy". *Journal of near Infrared Spectroscopy*. 12(1): 15-27.

Gillon, D., Joffre, R. and Ibrahima, A. (1999). "Can litter decomposability be predicted by near infrared reflectance spectroscopy ?" *Ecology*. 80(1): 175-186.

Hansen, K.H., Angelidaki, I. and Ahring, B.K. (1998). "Anaerobic digestion of swin manure: inhibition by ammonia". *Water Research*. 32(1): 5-12.

Hansen, T.L., Schmidt, J.E., Angelidaki, I., Marca, E., Jansen, J.I.C., Mosbaek, H. and Christensen, T.H. (2004). "Method for determination of methane potentials of solid organic waste". *Waste Management*. 24(4): 393-400.

Hansen, T.L. (2005). "Quantification of environmental effects from anaerobic treatment of source-sorted organic house hold wastes". p.42. Th. Technical University of Denmark. Lyngby.

Herrmann, S., Mayer, J., Michel, K. and Ludwig, B. (2009). "Predictive capacity of visible-near infrared spectroscopy for quality parameter assessment of compost". *Journal of near Infrared Spectroscopy*. 17(5): 289-301.

He, P.J. (2010). "Anaerobic digestion: An intriguing long history in China". *Waste Management*. 30(4): 549-550.

Herrero, M., Murray, I., Fawcett, R.H. and Dent, J.B. (1996). "Prediction of the in vitro gas production and chemical composition of kikuyu grass by near-infrared reflectance spectroscopy". *Animal Feed Science and Technology*. 60(1-2): 51-67.

Herrero, M., Jessop, N.S., Fawcett, R.H., Murray, I. and Dent, J.B. (1997). "Prediction of the in vitro gas production dynamics of kikuyu grass by near-infrared reflectance spectroscopy using spectrally-structured sample populations". *Animal Feed Science and Technology*. 69(1-3): 281-287.

Hodge, G.R. and Woodbridge, W.C. (2004). "Use of near infrared spectroscopy to predict lignin content in tropical and sub-tropical pines". *Journal of near Infrared Spectroscopy*. 12(6): 381-390.

Hongqiang, L. and Hongzhang, C. (2008). "Near-infrared spectroscopy with a fiber-optic probe for state variables determination in solid-state fermentation". *Process Biochemistry*. 43(5): 511-516.

- Jacobi, H.F., Moschner, C.R. and Hartung, E.** (2009). "Use of near infrared spectroscopy in monitoring of volatile fatty acids in anaerobic digestion". *Water Science and Technology*. 60(2): 339-346.
- Jones, P.D., Schimleck, L.R., Peter, G.F., Daniels, R.F. and Clark, A.** (2006). "Nondestructive estimation of wood chemical composition of sections of radial wood strips by diffuse reflectance near infrared spectroscopy". *Wood Science and Technology*. 40(8): 709-720.
- Kays, S.E. and Barton, F.E.** (2002). "Near-Infrared Analysis of Soluble and Insoluble Dietary Fiber Fractions of Cereal Food Products". *Journal of Agricultural and Food Chemistry*. 50(10): 3024-3029.
- Knol, W., Vandermost, M.M. and Waart, J.D.** (1978). "Biogas Production by Anaerobic Digestion of Fruit and Vegetable Waste - Preliminary-Study". *Journal of the Science of Food and Agriculture*. 29(9): 822-830.
- Lasaridi, K.E. and Stentiford, E.I.** (1998). "A simple respirometric technique for assessing compost stability". *Water Research*. 32(12): 3717-3723.
- Leardi, R. and Gonzalez, A.L.** (1998). "Genetic algorithms applied to feature selection in PLS regression: how and when to use them". *Chemometrics and Intelligent Laboratory Systems*. 41(2): 195-207.
- Leardi, R.** (2000). "Application of genetic algorithm-PLS for feature selection in spectral data sets". *Journal of Chemometrics*. 14(5-6): 643-655.
- Lebzein, P., Paul, Chr.** (1997). "Use of near infrared spectroscopy for the estimation of the microbial portion of non-ammonia-nitrogen in the duodenum of dairy cows". *Animal Feed Science Technology*. 68: 225-233.
- Lehtomäki, A., Huttunen, S., Lehtinen, T.M. and Rintala, J.A.** (2008). "Anaerobic digestion of grass silage in batch leach bed processes for methane production". *Bioresource Technology*. 99(8): 3267-3278.
- Lesteur, M., Latrille, E., Roger, J. M., Bellon-Maurel, V., Gonzalez, C., Junqua, G., and Steyer, J.-P.** (2008). "A chemometric approach to minimise the diffusive effect of the biomass when using near infrared spectroscopic measurements for the monitoring of bioprocesses". 11th conference on Chemometrics in Analytical Chemistry - CAC2008, Montpellier, France, 2008/06/30-2008/07/04, 2:65-70.
- Lesteur, M., Bellon-Maurel, V., Gonzalez, C., Latrille, E., Roger, J.M., Junqua, G. and Steyer, J.P.** (2010). "Alternative methods for determining anaerobic biodegradability: A review". *Process Biochemistry*. 45(4): 431-440.

Lesteur, M., Latrille, E., Maurel, V.B., Roger, J.M., Gonzalez, C., Junqua, G. and Steyer, J.P. "First step towards a fast analytical method for the determination of biochemical methane potential of solid wastes by near infrared spectroscopy". *Bioresource Technology*. 102(3): 2280-2288.

Lissens, G., Thomsen, A.B., De Baere, L., Verstraete, W. and Ahring, B.K. (2004). "Thermal wet oxidation improves anaerobic biodegradability of raw and digested biowaste". *Environmental Science & Technology*. 38(12): 3418-3424.

Lovett, D.K., Deaville, E.R., Mould, F., Givens, D.I. and Owen, E. (2004). "Using near infrared reflectance spectroscopy (NIRS) to predict the biological parameters of maize silage". *Animal Feed Science and Technology*. 115(1-2): 179-187.

Mata-Alvarez, J., Macé, S. and Llabrés, P. (2000). "Anaerobic digestion of organic solid wastes. An overview of research achievements and perspectives". *Bioresource Technology*. 74(1): 3-16.

Mata-Alvarez, J., ed. (2003) *Biomethanization of the Organic Fraction of Municipal Solid Wastes*, pp. 323, IWA, London.

Matsche, N. and Stumwohrer, K. (1996). "UV absorption as control-parameter for biological treatment plants". *Water Science and Technology*. 33(12): 211-218.

McKay, G. (2002). "Dioxin characterisation, formation and minimisation during municipal solid waste (MSW) incineration: review". *Chemical Engineering Journal*. 86(3): 343-368.

Meissl, K., Smidt, E. and Schwanninger, M. (2007). "Prediction of humic acid content and respiration activity of biogenic waste by means of Fourier transform infrared (FTIR) spectra and partial least squares regression (PLS-R) models". *Talanta*. 72(2): 791-799.

Miikki, V., Senesi, N. and Hanninen, K. (1997). "Characterization of humic material formed by composting of domestic and industrial biowastes : Part 2 spectroscopic evaluation of humic acid structures". *Chemosphere*. 34(8): 1639-1651.

Moller, H.B., Sommer, S.G. and Ahring, B.K. (2004). "Methane productivity of manure, straw and solid fractions of manure". *Biomass and Bioenergy*. 26(5): 485-495.

Moletta, R. (2008) *La méthanisation*, pp. 532, Editions Tec & Doc Lavoisier.

Moreira, A.B., Dias, I.L.T., Neto, G.O., Zagatto, E.A.G. and Kubota, L.T. (2004). "Solid-phase fluorescence spectroscopy for the determination of acetylsalicylic acid in powdered pharmaceutical samples". *Analytica Chimica Acta*. 523(1): 49-52.

Mottet, A., Francois, E., Latrille, E., Steyer, J.P., Deleris, S., Vedrenne, F. and Carrere, H. (2010). "Estimating anaerobic biodegradability indicators for waste activated sludge". *Chemical Engineering Journal*. 160(2): 488-496.

Muret, C., Pouet, M.F., Touraud, E. and Thomas, O. (2000). "From UV spectra to degradability of industrial wastewater/definition and use of a "shape factor"". *Water Science and Technology*. 42(5-6): 47-53.

Mrkva, M. (1983). "Evaluation of Correlations between Absorbance at 254nm and Cod of River Waters". *Water Research*. 17(2): 231-235.

Naes, T., Isaksson, T., Kowalski, B. (1990). "Locally weighted regression and scatter correction for near-infrared reflectance data". *Analytical Chemistry*. 62(7) :664–673

Nallathambi Gunaseelan, V. (1995). "Effect of inoculum/substrate ratio and pretreatments on methane yield from Parthenium". *Biomass and Bioenergy*. 8(1): 39-44.

Nallathambi Gunaseelan, V. (2004). "Biochemical methane potential of fruits and vegetable solid waste feedstocks". *Biomass and Bioenergy*. 26(4): 389-399.

Nallathambi Gunaseelan, V. (2007). "Regression models of ultimate methane yields of fruits and vegetable solid wastes, sorghum and napiergrass on chemical composition". *Bioresource Technology*. 98(6): 1270-1277.

Nallathambi Gunaseelan, V. (2009). "Predicting ultimate methane yields of *Jatropha curcus* and *Morus indica* from their chemical composition". *Bioresource Technology*. 100(13): 3426-3429.

Neves, L., Oliveira, R. and Alves, M.M. (2006). "Anaerobic co-digestion of coffee waste and sewage sludge". *Waste Management*. 26(2): 176-181.

Neves, L., Ribeiro, R., Oliveira, R. and Alves, M.M. (2006). "Enhancement of methane production from barley waste". *Biomass and Bioenergy*. 30 : 599-603.

Neves, L., Gonçalo, E., Oliveira, R. and Alves, M.M. (2008). "Influence of composition on the biomethanation potential of restaurant waste at mesophilic temperatures". *Waste Management*. 28(6): 965-972.

Nousiainen, J., Ahvenjarvi, S., Rinne, M., Hellamaki, M. and Huhtanen, P. (2004). "Prediction of indigestible cell wall fraction of grass silage by near infrared reflectance spectroscopy". *Animal Feed Science and Technology*. 115(3-4): 295-311.

Observatoire de l'Énergie (2005). Les énergies renouvelables en France 1970-2003.

OCDE - Organisation de Coopération et de Développement Economique. (2008). Politiques de l'environnement : quelles combinaisons d'instruments ? p.262. ISBN : 9789264018426

Osborne, B.G. and Fearn, T. (1986) Near Infrared Spectroscopy in Food Analysis, pp. 200, Longman Scientific & Technical, Harlow, UK.

- Park, R.S., Agnew, R.E., Gordon, F.J. and Steen, R.W.J.** (1998). "The use of near infrared reflectance spectroscopy (NIRS) on undried samples of grass silage to predict chemical composition and digestibility parameters". *Animal Feed Science and Technology*. 72(1-2): 155-167.
- Palmowski, L.M. and Müller, J.A.** (2000). "Influence of the size reduction of organic waste on their anaerobic digestion". *Water Science and Technology*. 41(3): 155-162.
- Pasquini, C.** (2003). "Near Infrared Spectroscopy: fundamentals, practical aspects and analytical applications". *Journal of the Brazilian Chemical Society*. 14(2): 198-219.
- Pelaez, C., Mejia, A. and Planas, A.** (2004). "Development of a solid phase kinetic assay for determination of enzyme activities during composting". *Process Biochemistry*. 39(8): 971-975.
- Piccolo, A., Zena, A. and Conte, P.** (1996). "A comparison of acid hydrolyses for the determination of carbohydrate content in soils". *Communications in Soil Science and Plant Analysis*. 27(15-17): 2909-2915.
- Ponsa, S., Gea, T., Alerm, L., Cerezo, J. and Sanchez, A.** (2008). "Comparison of aerobic and anaerobic stability indices through a MSW biological treatment process". *Waste Management*. 28(12): 2735-2742.
- Prieto, N., Rohe, R., Lavín, P., Batten, G. and Andrés, S.** (2009). "Application of near infrared reflectance spectroscopy to predict meat and meat products quality: A review". *Meat Science*. 83(2): 175-186.
- Raposo, F., Banks, C.J., Siegert, I., Heaven, S. and Borja, R.** (2006). "Influence of inoculum to substrate ratio on the biochemical methane potential of maize in batch tests". *Process Biochemistry*. 41(6): 1444-1450.
- Raposo, F., de la Rubia, M.A., Borja, R., Alaiz, M., Beltran, J., Cavinato, C., Clinckspoor, M., Demirer, G., Diamadopoulos, E., Helmreich, B., Jenicek, P., Marti, N., Mendez, R., Noguero, J., Pereira, F., Picard, S. and Torrijos, M.** (2009). "An interlaboratory study as useful tool for proficiency testing of chemical oxygen demand measurements using solid substrates and liquid samples with high suspended solid content". *Talanta*. 80(1): 329-337.
- Ras, M., Girbal-Neuhauser, E., Paul, E., Spérandio, M. and Lefebvre, D.** (2008). "Protein extraction from activated sludge: An analytical approach". *Water Research*. 42(8-9): 1867-1878.
- Reynolds, D.M. and Ahmad, S.R.** (1997). "Rapid and direct determination of wastewater BOD values using a fluorescence technique". *Water Research*. 31(8): 2012-2018.
- Robert-Peillard, F., Palacio-Barco, E., Dudal, Y., Coulomb, B. and Boudenne, J.L.** (2009). "Alternative Spectrofluorimetric Determination of Short-Chain Volatile Fatty Acids in Aqueous Samples". *Analytical Chemistry*. 81(8): 3063-3070.

- Rodriguez, C., Hiligsmann, S., Ongena, M., Charlier, R. and Thonart, P.** (2005). "Development of an enzymatic assay for the determination of cellulose bioavailability in municipal solid waste". *Biodegradation*. 16(5): 415-422.
- Roger, J.M., Palagos, B., Bertrand, D. and Fernandez-Ahumada, E.** "CovSel: Variable selection for highly multivariate and multi-response calibration. Application to IR spectroscopy". *Chemometrics and Intelligent Laboratory Systems*. In Press, Accepted Manuscript.
- Roussel, S., Bellon-Maurel, V., Roger, J.-M. and Grenier, P.** (2003). "Authenticating white grape must variety with classification models based on aroma sensors, FT-IR and UV spectrometry". *Journal of Food Engineering*. 60(4): 407-419.
- Sánchez, A.** (2009). "Test methods to aid in the evaluation of the diversion of biodegradable municipal waste (BMW) from landfill". *Waste Management*. 29(8): 2306-2307.
- Sanderson, M.A., Agblevor, F., Collins, M. and Johnson, D.K.** (1996). "Compositional analysis of biomass feedstocks by near infrared reflectance spectroscopy". *Biomass & Bioenergy*. 11(5): 365-370.
- Savitzky, A., Golay, M. J. E.** (1964). "Smoothing and Differentiation of Data by Simplified Least Squares Procedures". *Analytical Chemistry*. 36(8) :1627–1639.
- Schievano, A., Pognani, M., D'Imporzano, G. and Adani, F.** (2008). "Predicting anaerobic biogasification potential of ingestates and digestates of a full-scale biogas plant using chemical and biological parameters". *Bioresource Technology*. 99(17): 8112-8117.
- Schievano, A., Scaglia, B., D'Imporzano, G., Malagutti, L., Gozzi, A. and Adani, F.** (2009). "Prediction of biogas potentials using quick laboratory analyses: Upgrading previous models for application to heterogeneous organic matrices". *Bioresource Technology*. 100(23): 5777-5782.
- Sellami, F., Hachicha, S., Chtourou, M., Medhioub, K. and Ammar, E.** (2008). "Maturity assessment of composted olive mill wastes using UV spectra and humification parameters". *Bioresource Technology*. 99(15): 6900-6907.
- Shenk, J.S., Westerhaus, M.O.** (1997). " Investigation of a LOCAL calibration procedure for near infrared instruments". *Journal of Near Infrared Spectroscopy*. 5(4) : 223–232
- Sialve, B., Bernet, N. and Bernard, O.** "Anaerobic digestion of microalgae as a necessary step to make microalgal biodiesel sustainable". *Biotechnology Advances*. 27(4): 409-416.
- Smidt, E., Lechner, P., Schwanninger, M., Haberhauer, G. and Gerzabek, M.H.** (2002). "Characterization of waste organic matter by FT-IR spectroscopy: Application in waste science". *Applied Spectroscopy*. 56(9): 1170-1175.
- Smidt, E. and Schwanninger, M.** (2005). "Characterization of waste materials using FTIR spectroscopy: Process monitoring and quality assessment". *Spectroscopy Letters*. 38(3): 247-270.

Sousa, A.C., Lucio, M., Bezerra, O.F., Marcone, G.P.S., Pereira, A.F.C., Dantas, E.O., Fragoso, W.D., Araujo, M.C.U. and Galvao, R.K.H. (2007). "A method for determination of COD in a domestic wastewater treatment plant by using near-infrared reflectance spectrometry of seston". *Analytica Chimica Acta*. 588(2): 231-236.

Stephens, A.B. and Walker, P.N. (2002). "Near-infrared spectroscopy as a tool for real-time determination of BOD5 for single-source samples". *Transactions of the Asae*. 45(2): 451-458.

Steyer, J.-P., Bouvier, J.-C., Conte, T., Gras, P., Harmand, J. and Delgenès, J.-P. (2001) « On-line measurements of COD, TOC, VFA, total and partial alkalinity in anaerobic digestion processes using infra-red spectrometry ». In: *9th world congress Anaerobic Digestion 2001 : anaerobic conversion for sustainability. Part I.* (Van Velsen, A.F.M. and Verstraete, W.H., eds.), pp. 305-314. IWA, Netherlands Association for Water Management,, Antwerpen, Belgium, September 2-6

Steyer, J.P., Bouvier, J.C., Conte, T., Gras, P. and Soubie, P. (2002). "Evaluation of a four year experience with a fully instrumented anaerobic digestion process". *Water Science and Technology*. 45(4-5): 495-502.

Thuries, L., Bastianelli, D., Davrieux, F., Bonnal, L., Oliver, R., Pansu, M. and Feller, C. (2005). "Prediction by near infrared spectroscopy of the composition of plant raw materials from the organic fertiliser industry and of crop residues from tropical agrosystems". *Journal of near Infrared Spectroscopy*. 13(4): 187-199.

Tsoumanis, C.M., Giokas, D.L. and Vlessidis, A.G. (2010). "Monitoring and classification of wastewater quality using supervised pattern recognition techniques and deterministic resolution of molecular absorption spectra based on multiwavelength UV spectra deconvolution". *Talanta*. 82(2): 575-581.

Turick, C.E., Peck, M.W., Chynoweth, D.P., Jerger, D.E., White, E.H., Zsuffa, L. and Andy Kennedy, W. (1991). "Methane fermentation of woody biomass". *Bioresource Technology*. 37(2): 141-147.

Van der Voet, H. (1999). "Pseudo-degrees of freedom for complex predictive models: The example of partial least squares". *Journal of Chemometrics*. 13(3-4): 195-208.

Van Soest, P.J., Robertson, J.B. and Lewis, B.A. (1991). "Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition." *Journal of Dairy Science*. 74(10): 3583-3597.

Vergnoux, A., Guiliano, M., Le Dreau, Y., Kister, J., Dupuy, N. and Doumenq, P. (2009). "Monitoring of the evolution of an industrial compost and prediction of some compost properties by NIR spectroscopy". *Science of The Total Environment*. 407(7) : 2390-2403.

VIM. Vocabulaire International de Métrologie. (2006) "Concepts fondamentaux et généraux et termes associés". 3^e édition.

- Wagland, S.T., Tyrrel, S.F., Godley, A.R. and Smith, R.** (2009). "Test methods to aid in the evaluation of the diversion of biodegradable municipal waste (BMW) from landfill". *Waste Management*. 29(3): 1218-1226.
- Ward, A.J., Hobbs, P.J., Holliman, P.J. and Jones, D.L.** (2008). "Optimisation of the anaerobic digestion of agricultural resources". *Bioresource Technology*. 99(17): 7928-7940.
- Wey, M.-Y., Ou, W.-Y., Liu, Z.-S., Tseng, H.-H., Yang, W.-Y. and Chiang, B.-C.** (2001). "Pollutants in incineration flue gas". *Journal of Hazardous Materials*. 82(3): 247-262.
- Widmer-Girod, C., Staub, C.** (2004). "Measurement uncertainty : to be define in forensic toxicology". *Annales de Toxicologie Analytique*. 16(3): 215-219.
- Williams, P. and Norris, K.** (2001) "Near-Infrared Technology in the Agricultural and Food Industries". 2nd Edition, pp. 312, Amer Assn of Cereal Chemists, St. Paul, Minnesota, USA.
- Williams, P.C. and Sobering, D.C.** (1993). "Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds". *Journal of Near Infrared Spectroscopy*. 1(1): 25-32.
- Wilman, D., Field, M., Lister, S.J. and Givens, D.I.** (2000). "The use of near infrared spectroscopy to investigate the composition of silages and the rate and extent of cell-wall degradation". *Animal Feed Science and Technology*. 88(3-4): 139-151.
- Zamora-Rojas, E., Garrido-Varo, A., Van den Berg, F., Guerrero-Ginel, J.E. and Pérez-Marín, D.C.** (2010). "Evaluation of a new local modelling approach for large and heterogeneous NIRS data sets". *Chemometrics and Intelligent Laboratory Systems*. 101(2): 87-94.
- Zhang, L. and Garcia-Munoz, S.** (2009). "A comparison of different methods to estimate prediction uncertainty using Partial Least Squares (PLS): A practitioner's perspective". *Chemometrics and Intelligent Laboratory Systems*. 97(2): 152-158.
- Zeaiter, M., Roger, J.-M. and Bellon-Maurel, V.** (2005). "Robustness of models developed by multivariate calibration. Part II: The influence of pre-processing methods". *TrAC Trends in Analytical Chemistry*. 24(5): 437-445.
- Zbytniewski, R. and Buszewski, B.** (2005). "Characterization of natural organic matter (NOM) derived from sewage sludge compost. Part 1: chemical and spectroscopic properties". *Bioresource Technology*. 96(4): 471-478.

Liste des Figures

Figure 1 : Production d'électricité (GWh) et de chaleur (ktep) à partir du biogaz (France) (Observatoire de l'énergie, 2005).....	5
Figure 2 : Les étapes de la digestion anaérobie (Mata-Alvarez, 2003).	7
Figure 3 : Représentation de résultats obtenus lors du test BMP. Les chiffres indiquent la valeur BMP.....	8
Figure 4 : Exemple de résultats de l'étude réalisée entre 26 laboratoires, afin de comparer des résultats de dosage de la DCO.	16
Figure 5 : Mise en œuvre des fioles à plasma couramment utilisées pour le test BMP (Angelidaki et al., 2009).	18
Figure 6 : Production de méthane cumulé de divers échantillons (Cho et al., 1995).	18
Figure 7 : Corrélation entre l'indice aérobie RI_4 et l'indice de production de biogaz (GB_{21}) (Cossu et al., 2008)....	23
Figure 8 : Corrélation entre la quantité de sucre libérée par hydrolyse enzymatique et les tests BMP (Rodriguez et al. 2005).	24
Figure 9 : Relation entre le rendement en méthane et le pourcentage en glucides complexes (Cellulose + Hémicellulose) (Eleazer et. al, 1997).....	28
Figure 10 : Relation entre (cellulose + lignine) et la biodégradabilité de divers déchets non broyés (Buffière et al., 2006).	29
Figure 11 : Prédiction de la biodégradabilité de boues (Mottet et al., 2010).	31
Figure 12 : Régression entre les valeurs BMP mesurées et les valeurs BMP théoriques obtenues par la méthode (C, H, O, N) et la méthode (glucides, protéines, lipides) pour plusieurs fractions fermentescibles d'ordures ménagères. (Davidson et al., 2007).....	32
Figure 13 : Classification des différentes formes de spectres UV (Muret et al., 2000).....	36
Figure 14 : Les différents principes de la mesure dans l'infrarouge	40
Figure 15 : Relation entre les valeurs prédites par SPIR et les valeurs mesurées pour les données d'étalonnage pour (a) la perte de poids de la litière végétale après une semaine d'incubation et (b), la constante cinétique k du modèle exponentiel (Gillon et al., 1999).	42
Figure 16 : Présentation des différentes étapes du process de tri-compostage.....	49
Figure 17 : Représentation schématique de l'ACP	53
Figure 17 : Ensemble des valeurs BMP des 74 échantillons utilisés pour l'étalonnage et le test.	56
Figure 18 : Spectres des ordures ménagères (a) et des déchets simples (b) en absorbance.	57
Figure 19 : Illustration des différentes interactions possibles entre le rayonnement proche infrarouge et les particules solides, induisant des effets multiplicatifs et additifs.	58
Figure 20 : Spectres des ordures ménagères (a) et des déchets simples (b) après application des prétraitements SNV et Detrend.....	59
Figure 21 : Analyse en composante principale sur l'ensemble des spectres avant (a) et après prétraitements (b). ...	60
Figure 22 : Relation entre les valeurs de la première composante principale et l'absorbance à 1191 nm.	61
Figure 23 : Valeurs BMP prédites versus valeurs BMP mesurées.	64

Figure 24 : Régression linéaire de l'ensemble de validation externe.....	65
Figure 25 : Exemple de représentation de l'évolution des valeurs de SECV en fonction du nombre de variables sélectionnées dans le modèle lors d'une répétition. Le nombre de variable aux modèles AG1 et AG2 sont représentés.	81
Figure 26 : Spectres proche infrarouge des déchets (a), spectre moyen et +/- 2 x écarts types (b).....	82
Figure 27 : Représentation des b coefficients PLS, de la dérivée seconde renversée de la moyenne des spectres et indication des longueurs d'onde correspondantes aux b coefficients PLS positifs.....	83
Figure 28 : Table de corrélation (R^2) entre les longueurs d'onde des spectres prétraités (SNV et Detrend sur la gamme 1668 – 2500 nm).....	84
Figure 29 : Représentation des b coefficients PLS, de la dérivée seconde renversée de la moyenne des spectres et indication des longueurs d'onde correspondant aux b coefficients PLS négatifs.	87
Figure 30 : Spectres prétraités, SNV-Detrend et SNV-Detrend-dérivée seconde (normale) (11,3,2) de la lignine. ...	88
Figure 31 : Spectres de différents plastiques pris dans les échantillons.	90
Figure 32 : Régression linéaire entre la valeur à 2308 nm et la valeur BMP.....	91
Figure 33 : Spectre d'un déchet avec beaucoup de plastiques (valeur BMP la plus faible) et spectre d'un déchet avec peu de plastiques (valeur BMP la plus forte), observation des zones de corrections de la contribution positive des plastiques à 2318 nm.....	91
Figure 34 : Ensemble des spectres utilisés (sélection de variables (1668 – 2500 nm) - SNV – Detrend) et les variables sélectionnées par les 20 répétitions de l'algorithme génétique.....	93
Figure 35 : Représentation des b coefficients PLS, de la dérivée seconde renversée de la moyenne des spectres et indication des longueurs d'onde sélectionnées par l'algorithme CovSel.	96
Figure 36 : Représentation des b coefficients PLS et des b coefficients associés à la régression multiple (MLR) faite à l'aide des variables sélectionnées par CovSel.	97
Figure 38 : Prédiction des 100 nouveaux échantillons.	106
Figure 39 : Représentation de deux premières composantes PLS du modèle sur les ordures ménagères et projection des 100 nouveaux échantillons sur ces deux composantes.	107
Figure 40 : Première et seconde composante principale de l'ACP effectuée sur l'ensemble des échantillons.....	108
Figure 41 : Premier vecteur propre de l'ACP.....	109
Figure 42 : Spectre proche infrarouge d'algues (a), présentant un pic important à 1930 nm et d'une boue grasseuse (b), présentant des pics d'absorbance de la liaison C – H	110
Figure 43 : Second vecteur propre de l'ACP.	111
Figure 44 : Troisième et quatrième composantes principales de l'ACP sur les 174 échantillons de déchets.....	112
Figure 45 : Troisième vecteur propre de l'ACP.....	112
Figure 46 : Régression PLS sur les ensembles d'étalonnage et de validation.	114

Liste des Tableaux

Tableau 1 : Production primaire de biogaz dans différents pays de l'Union Européenne en 2007 (en ktep) (EurObserv'ER, 2009).....	6
Tableau 2 : Comparaison des différents déchets avec leur valeur de potentiel méthane associée.....	21
^a Valeurs calculées à partir des données de références des articles.	21
Tableau 3 : Longueurs d'onde d'intérêt dans l'UV/Visible (Zbytniewski et Buszewski, 2005).....	34
Tableau 4 : Prédiction des paramètres cinétiques de la production de gaz cumulé par la composition chimique et la spectroscopie proche infrarouge (Andrès et al., 2005).....	43
Tableau 5 : Répartition des différents déchets dans chaque ensemble.....	50
Tableau 6 : Gamme, moyenne et écart type des valeurs BMP des ensembles d'étalonnage, de test et de validation indépendante (toutes les valeurs sont en ml.CH ₄ .g ⁻¹ MV).....	56
Tableau 7 : Résultats des différents étalonnages et validation testées.	62
Tableau 8 : Comparaison des performances entre le modèle SPIR et le test BMP.	69
Tableau 9 : Pourcentage des composés majeurs des déchets ménagers (Davidsson et al., 2007).....	83
Tableau 10 Comparaison des résultats des modèles obtenus après sélection des variables par AG. VL : Variable latente.....	92
Tableau 11 : Attribution chimique des longueurs d'onde d'intérêt du modèle de prédiction de la valeur BMP d'ordures ménagères.....	99
Tableau 12 : Différentes catégories de déchets et nombre d'échantillons correspondants.....	103
Tableau 13 : Moyennes, écarts-types, minimums et maximum des deux ensembles d'échantillons.	104
Tableau 14 : Moyenne, écart type, minimum et maximum des ensembles d'étalonnage et de validation.	105
Tableau 15 : Résultat d'étalonnage et de prédiction des modèles « globaux ».	114