

COMMUNAUTÉ FRANÇAISE DE BELGIQUE

UNIVERSITÉ DE LIÈGE – GEMBLoux AGRO-BIO TECH

**ÉVALUATION DES
POTENTIALITÉS DE L'IMAGERIE
HYPERSPÉCTRALE PROCHE
INFRAROUGE ET DE LA
CHIMIOMÉTRIE DANS L'ÉTUDE
DE SYSTÈMES RACINAIRES**

Damien EYLENBOSCH

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

Promoteur : Professeur Bernard Bodson
Année civile : 2018

Illustration de la page de couverture: prédiction de la nature des pixels d'une image hyperspectrale proche infrarouge acquise sur un échantillon contenant des racines de froment, des résidus de cultures, des particules de sol et le tamis utilisé pour le transport de l'échantillon. Une couleur est assignée à chaque classe spectrale des modèles de discrimination utilisés.

Copyright. Cette œuvre est sous licence Creative Commons. Vous êtes libre de reproduire, de modifier, de distribuer et de communiquer cette création au public selon les conditions suivantes :

- paternité (BY) : vous devez citer le nom de l'auteur original de la manière indiquée par l'auteur de l'œuvre ou le titulaire des droits qui vous confère cette autorisation (mais pas d'une manière qui suggérerait qu'ils vous soutiennent ou approuvent votre utilisation de l'œuvre) ;

- pas d'utilisation commerciale (NC) : vous n'avez pas le droit d'utiliser cette création à des fins commerciales ;

- partage des conditions initiales à l'identique (SA) : si vous modifiez, transformez ou adaptez cette création, vous n'avez le droit de distribuer la création qui en résulte que sous un contrat identique à celui-ci.

À chaque réutilisation ou distribution de cette création, vous devez faire apparaître clairement au public les conditions contractuelles de sa mise à disposition. Chacune de ces conditions peut être levée si vous obtenez l'autorisation du titulaire des droits sur cette œuvre. Rien dans ce contrat ne diminue ou ne restreint le droit moral de l'auteur.

*A celles et ceux qui ont cru en ma capacité à réaliser cette thèse et qui m'ont
encouragé dans mes recherches.*

EYLENBOSCH Damien (2018). Évaluation des potentialités de l'imagerie hyperspectrale proche infrarouge et de la chimiométrie dans l'étude de systèmes racinaires (Thèse de doctorat). Université de Liège, Gembloux Agro-Bio Tech, Gembloux, Belgique, 227 p., 8 tabl., 68 fig.

Résumé

Les racines jouent un rôle vital dans le développement des plantes. Premier organe à se développer lors de l'installation des plantes, elles permettent leur ancrage et le prélèvement d'eau et de nutriments dans le sol. Les racines permettent aussi des relations entre plantes, de la même espèce ou non, et avec des microorganismes du sol. Enfin, les racines, par leurs exsudats et lors de leur décomposition, apportent des quantités importantes de carbone dans les sols. Leur étude est donc essentielle face aux défis agronomiques et climatiques actuels: diminution des ressources en eau, en fertilisants, compétition interspécifique, cultures associées, stockage de carbone dans les sols. Cette étude est d'autant plus intéressante lorsqu'elle est menée dans les conditions naturelles de croissance des plantes.

Toutes les études racinaires menées dans les conditions naturelles font face au même problème : la présence du sol qui empêche une observation directe des racines. Les études menées sur les racines sont donc toutes basées sur une série d'estimations et de compromis entre la qualité des observations et le nombre d'observations pouvant être réalisées.

Parmi les techniques d'étude des racines, celle basée sur le prélèvement de carottes de sol semble être la plus utilisée. Cette méthode, dont la mise en place est relativement simple, permet d'acquérir rapidement, et en grand nombre, des échantillons de sol contenant des racines. Ce nombre élevé de prélèvements permet de répéter les mesures dans le temps et l'espace ce qui permet d'étudier la dynamique de développement racinaire et la distribution racinaire dans le sol. Cette technique est dès lors souvent utilisée comme référence.

Pour pouvoir observer, mesurer et quantifier les racines prélevées avec cette méthode, les carottes de sol doivent être lavées pour en extraire les racines. Ces racines doivent ensuite être séparées manuellement des autres éléments présents au sein des carottes de sol : racines mortes ou d'autres espèces, résidus de cultures, particules de sol, macroorganismes... Ces étapes de lavage et de tri prennent beaucoup de temps, sont fastidieuses, dépendent de l'habileté de l'opérateur et sont donc les principaux freins à la répétition des observations. De plus, certains éléments présents au sein des échantillons lavés sont difficilement séparables sur base de leur aspect visuel.

Afin de faciliter, d'objectiver et de réduire le temps nécessaire à l'étape de tri, une nouvelle méthode basée sur l'utilisation combinée de l'imagerie hyperspectrale proche infrarouge et de la chimiométrie a été proposée. L'imagerie hyperspectrale proche infrarouge est une technique d'analyse chimique qui, en combinant les techniques d'imagerie et de spectroscopie, permet l'acquisition d'un spectre proche infrarouge pour chaque pixel d'une image. Ce spectre est une signature liée à la

nature physico-chimique de l'élément sur lequel il a été acquis. La chimiométrie est quant à elle la science qui permet d'extraire l'information pertinente des spectres.

Cette méthode a d'abord été utilisée pour discriminer des racines, des particules de sol et des résidus de cultures extraits par lavage de carottes de sol. Les modèles chimiométriques mis au point ont permis de classer, selon des classes prédéfinies, chaque pixels des images hyperspectrales acquises sur les échantillons de sol et dès lors de connaître la composition des échantillons étudiés. En liant un nombre de pixels de racines à une masse de racines, il a ensuite été possible de prédire la masse de racine présente au sein d'un échantillon sans devoir réaliser le tri manuel et la pesée de celui-ci. Sur base de leurs signatures spectrales, cette méthode a également permis de discriminer et de quantifier des racines de différentes espèces au sein d'échantillons de sol et d'étudier la composition de ces racines.

Dans le cadre des recherches menées durant cette thèse, l'imagerie hyperspectrale proche infrarouge a été utilisée pour étudier le développement racinaire du froment d'hiver selon différentes modalités de travail du sol, pour quantifier les racines de cultures de froment et de pois d'hiver cultivées en association et pour mesurer des teneurs en leghémoglobine, une protéine associée à la fixation de l'azote atmosphérique par les légumineuses, dans des nodosités de pois.

Les recherches ont également eu pour objectif d'étudier la qualité de différents modèles de discrimination, de caractériser la stabilité du système d'acquisition des images hyperspectrales et d'identifier les sources d'erreurs lorsque cette méthode était utilisée pour quantifier des masses de racines. Si ces recherches ont montré les potentialités de l'imagerie hyperspectrale proche infrarouge dans l'étude des racines (suppression des étapes de tri et de pesées, possibilité de quantifier des racines de natures différentes et de composés au sein des racines), elles ont également mis en évidence plusieurs limites à cette méthode. Ces limites dans la quantification précise des masses de racines sont dues à l'instabilité du système d'acquisition, à la résolution spatiale des images, au recouvrement des racines et des résidus de cultures lors de l'acquisition des images et à la difficulté de classer certains spectres. De plus, les modèles mis au point ne sont valables que pour discriminer les classes de spectres pour lesquels ils ont été calibrés et ne permettent donc pas de classer des éléments de natures différentes tels que les racines d'autres cultures.

Avec ses avantages et ses limites, l'imagerie hyperspectrale proche infrarouge couplée à la chimiométrie peut donc être utilisée pour faire des estimations grossières de masses racinaires, ce qui permet d'étudier des distributions racinaires, mais ne permet actuellement pas de réaliser des mesures de masses précises. Elle semble également être particulièrement intéressante pour étudier des teneurs en leghémoglobine à l'échelle de nodosités individuelles et pour étudier la distribution de ces teneurs sur des systèmes racinaires de légumineuses.

EYLENBOSCH Damien (2018). Evaluation of the potentialities of near infrared hyperspectral imaging and chemometrics in the study of root systems (PhD thesis). University of Liège, Gembloux Agro-Bio Tech, Gembloux, Belgium, 227 p., 8 tabl., 68 fig.

Abstract

Roots play a vital role in plant development. First organ appearing at plant installation, they provide plant anchorage and absorb water and nutrients in the surrounding soil. Roots play also a role in relations between plants, from the same species or not, and with soil microorganisms. Furthermore, by their exudates and their decay, roots bring large amount of carbon in soils. Their study is therefore essential in order to face the present challenges in agronomy and with climate changes: decrease of water and nutrient resources, interspecific competition, intercropping and carbon storage in soils. This study is even more interesting when it is conducted in natural plant growing conditions.

All root studies conducted in natural conditions face to the same problem: the soil presence which hinder roots and do not allow direct observations. Root studies are therefore based on estimations and a compromise between observation quality and number of feasible observations.

Among existing techniques for root study, soil coring seems to be the most used. This technique, which is easy to use, allows rapid acquisitions of large number of soil samples containing roots. This ability to acquire a large number of samples allow the sampling replication in time and space and allow therefore to study root development dynamic and root distribution in soils. This technique is therefore often used as a reference.

In order to observe, measure and quantify roots acquired with this method, soil cores have to be washed to extract roots from soil. These roots have then to be separated from the other elements present in the cores: dead roots, roots from other species, crop residues, soil particles, macro organisms... These washing and sorting steps are time consuming and tedious, depend on operator ability and are therefore the main limitation in observation replication. Moreover, some elements extracted from soil cores are hardly separable on their visual aspect.

In order to facilitate, objectify and reduce the time needed for the manual sorting step, a new method, based on near infrared hyperspectral imaging and chemometrics, was proposed. Near infrared hyperspectral imaging is a chemical analysis technique combining imagery and spectroscopy and allowing the acquisition of one single near infrared spectrum for each pixel of an image. This spectrum is a signature linked to the physicochemical nature of the pixel on which it is acquired. Chemometrics is the science allowing the extraction of relevant information from spectra.

This method was first used to discriminate roots, soil particles and crop residues extracted by washing soil cores. The developed chemometric models allowed the classification, in predefined classes, of each pixel of hyperspectral images acquired

on soil samples. The composition of studied samples was therefore known. By linking the number of pixels classified as root to a root mass, it was then possible to predict the root mass in a sample without having to sort it manually and to weight it. Based on their spectral signature, this method also allowed discrimination and quantification of roots from different species and study on root composition.

As part of the research conducted in this thesis, near infrared hyperspectral imaging was used to study root development of winter wheat cultivated after contrasted tillage practices, to quantify roots from winter wheat and pea crops cultivated together and to measure leghaemoglobin content, a protein involved in nitrogen fixation by legumes, in pea nodules.

Research also aimed to study the quality of several discriminant models, to characterize hyperspectral image acquisition system stability and to identify error sources when this method was used to quantify root masses. Although these research showed the potentialities of near infrared hyperspectral imaging in root studies (sorting and weighting steps suppression, opportunities in quantification of roots from several species and of root compounds), they also highlighted the limits of this method. These limits in precise root mass quantification are due to image acquisition system instability, image resolution, overlapping of roots and crop residues during image acquisition and difficulties to classify some spectra. Moreover, developed models are only usable in discrimination of spectral classes for which they were built. They are therefore not useable on new elements like roots from other crops.

With their advantages and limitations, near infrared hyperspectral imaging and chemometrics can therefore be used to estimate roughly root masses but is currently not usable for precise measurements. They seem to be particularly interesting to study leghaemoglobin content in single nodules and to study the distribution of these contents on legume root systems.

Remerciements

Je tiens tout d'abord à remercier le Professeur Bernard Bodson, mon promoteur, qui m'a proposé à la fin de mes études de réaliser cette thèse et d'être son assistant. Merci pour ces six années de collaboration, pour tout ce que vous m'avez appris et pour la confiance que vous m'avez accordée.

Je tiens ensuite à remercier Vincent Baeten qui a eu l'idée de ce travail. Merci pour la confiance que tu m'as accordée, pour tous tes encouragements et pour m'avoir donné accès aux laboratoires de ton unité.

Un tout grand merci à Juan Antonio Fernández Pierna qui m'a encadré de près durant cette thèse. Merci pour ton aide précieuse dans les analyses chimométriques, mais surtout, pour tes nombreuses relectures en profondeur et cela malgré le fait que je t'envoyais souvent les textes en dernières minutes.

Merci à Nicaise Kayoka, mon professeur d'imagerie hyperspectrale. Tu m'as appris à utiliser la caméra hyperspectrale proche infrarouge, cette "vieille dame capricieuse", et nos longues discussions m'ont permis d'avancer dans cette thèse.

Merci à Damien Vincke, pour l'amitié qui est née au cours des conférences auxquelles nous avons participé ensemble et pour ton aide précieuse dans la création des scripts Matlab qui m'ont permis d'extraire un maximum d'informations de mes données spectrales.

Merci à Quentin Arnould qui allumait la caméra tôt le matin lorsque je devais faire mes acquisitions et à toute l'équipe de l'Unité Qualité des produits (CRA-W) qui m'a si gentiment accueilli.

Merci aux membres de mon comité de thèse, les Professeurs Jean-Pierre Destain, Gilles Colinet et Patrick du Jardin, qui m'ont encadré et conseillé durant cette thèse.

Merci au Professeur Xavier Draye qui a suivi de loin mes travaux et a accepté de participer à l'évaluation de ma thèse. Merci pour ta relecture attentive du document final et tous tes commentaires. Merci également de m'avoir permis de participer à deux formations très intéressantes sur l'utilisation de l'imagerie dans l'étude des plantes et le phénotypage.

Merci à Fred Colinet pour son amitié et pour m'avoir donné accès au serveur ayant permis la réalisation de mes calculs. Merci aussi pour ton aide lorsque le serveur ou le vnc plantait.

Parmi mes collègues de l'Unité de Phytotechnie tempérée, je tiens tout d'abord à remercier Benjamin Dumont pour sa disponibilité et son aide précieuse dans le traitement des données et la rédaction des articles scientifiques. Je te remercie également pour les longues discussions que nous avons eues au cours de ces quelques années où nous avons travaillé ensemble et pour la bonne ambiance que tu mettais dans l'équipe.

Un tout grand merci à Rodrigo Meza, pour ton amitié et les longues journées passées ensemble aux champs et lors de la préparation des Livres Blancs, à Marie-Pierre Hiel, pour les bons moments passés ensemble en conférences et pour ton aide

dans la création des scripts R, à Rémi Delcroix et Loïc Dewaele pour leur aide dans le prélèvement et le lavage des échantillons de racines, à Micheline Joly et Carine Demazy pour leur gentillesse, leur aide pour les démarches administratives et le café qui était prêt tous les jours. Merci à tous mes autres collègues de l'Unité de Phytotechnie tempérée avec qui j'ai vécu de nombreux beaux moments, entre autres lors des moissons. Je remercie donc Joseph Cap, Jean-Marc Dricot, Olivier Decelle, Laurent Quertinmont, Jules Latour, Bruno Monfort, Jean-Pierre Denis, Christine Cartrysse, Jérôme Pierreux, Patrick Gustin, Jean-Christophe Pector, Alexandre Detournay, Mélissa Lonneux et Rémy Blanchart. Merci également à Tanguy Manise et Donat Regaert avec qui j'ai partagé mon bureau.

Merci à tous les étudiants qui ont travaillé avec moi dans le cadre de cette thèse: Guillaume Fraipont, Adrien Boufflette, Louis Dumont de Chassart, Ariel Doudet, Florent Druart, Laëtitia Machiels, Audrey Sougnez, Marine Torfs, François Godin, Xavier Bordet, Affi Attikora et Florian Boucheron. Sans vous, beaucoup de travaux n'auraient pas été possibles et la quantité de résultats obtenue serait bien maigre.

Je n'en serais pas arrivé là sans l'aide et l'affection de mes parents et de ma famille. Un immense merci Papa et Maman pour m'avoir permis de réaliser mes études dans les meilleures conditions. Elles m'ont permis d'obtenir ce travail épanouissant. Mes frères n'ont pas toujours bien compris l'objectif de ma thèse mais ça ne les a pas empêchés de m'encourager. Merci Eloy et Benoît! Merci également à Marie-Claire et à Jan, ma tante et mon oncle, et à Jean-Louis, le papa d'Anaïs, pour leur soutien et leurs encouragements. Ils savent à quel point une thèse demande souvent de la persévérance.

Merci à mes amies, mes amis, et particulièrement mes coloc's, qui m'ont si souvent demandé des nouvelles de ma thèse et qui m'ont encouragé dans ce travail de longue durée.

Un immense merci enfin à Anaïs, ma fiancée. Vouloir se marier et présenter une thèse le même mois n'était certainement pas la meilleure idée... Mais tu as su prendre en main les préparatifs pendant que je me concentrais sur la finalisation de cette thèse. Merci pour ton amour, ton aide, ton soutien, tes encouragements, tes attentions et ta compréhension, même lorsque je passais une journée de congé avec ma thèse plutôt qu'avec toi. Le projet que nous construisons ensemble est le plus beau qui soit pour une vie!

Table des matières

Évaluation des potentialités de l'imagerie hyperspectrale proche infrarouge et de la chimométrie dans l'étude de systèmes racinaires

Résumé	VII
Abstract	IX
Remerciements	XI
Table des matières	XIII
Liste des figures.....	XIX
Liste des tableaux	XXIII
Liste des équations	XXV
Liste des abréviations	XXVII
Avant-propos.....	1
Définitions et contexte.....	1
Entendons-nous bien!	3
Contexte de la thèse et collaborations	5
Chapitre I: Revue bibliographique	9
1. L'étude des racines	11
1.1. Le système racinaire	11
1.2. Les représentations des systèmes racinaires	11
1.3. Dynamique de développement racinaire	14
1.4. Le rôle des racines	15
1.5. Facteurs influençant le développement racinaire	16
1.6. L'importance de mesurer les racines.....	21
1.7. Les paramètres mesurés sur les racines	22
1.8. L'étude et la quantification des racines aux champs	27
1.9. Utilisation de l'imagerie dans l'étude des systèmes racinaires	36
1.10. Méthodes de lavage et de séparation de racines et de matières organiques étrangères	37
1.11. Méthodes de discrimination et de quantification de racines de différentes espèces.....	40
1.12. Méthodes de séparation de racines vivantes et mortes	41

1.13.	Les limites de l'étude des systèmes racinaires.....	42
2.	Spectroscopie et imagerie hyperspectrale proche infrarouge	44
2.1.	La spectroscopie proche infrarouge	44
2.2.	L'imagerie hyperspectrale proche infrarouge.....	45
2.3.	Les principes sous-jacents de l'imagerie hyperspectrale	47
2.4.	Acquisition des images hyperspectrales	48
2.5.	Chimiométrie et analyse des images hyperspectrales proche infrarouge.....	51
2.6.	Avantages et limites de l'imagerie hyperspectrale proche infrarouge.....	60
3.	Utilisations de la spectroscopie dans l'étude racinaire	63
3.1.	La spectroscopie de réflectance proche infrarouge dans l'étude racinaire	63
3.2.	La spectroscopie en moyen infrarouge dans l'étude racinaire	66
3.3.	La spectroscopie de fluorescence dans l'étude racinaire	67
3.4.	Contraintes dans l'utilisation de la spectroscopie pour l'étude de systèmes racinaires	67
Chapitre II: Objectifs de la thèse		69
Chapitre III: Matériels et méthodes		73
1.	Introduction	75
2.	Préparation des échantillons	76
2.1.	Prélèvement des racines.....	76
2.2.	Lavage des échantillons de sol et récupération des racines	76
3.	Acquisition des images hyperspectrales et des spectres proche infrarouge.....	81
3.1.	Le système d'acquisition	81
3.2.	L'acquisition des images hyperspectrales.....	86
3.3.	L'extraction des spectres proche infrarouge.....	89
3.4.	Traitement des spectres.....	89
3.5.	Sélection des sets de calibration et de validation.....	90
4.	Construction des modèles de discrimination et de quantification	92
4.1.	Prétraitement des spectres.....	92
4.2.	Modèles de discrimination.....	92
4.3.	Construction de l'arbre de classification dichotomique.....	92

4.4.	Prédictions sur images	93
4.5.	Modèles de quantification	93
5.	Validation	95
Chapitre IV: Suivi de la stabilité du système d'acquisition des images hyperspectrales proche infrarouge.....		97
1.	Introduction	99
2.	Matériels et méthodes.....	100
3.	Résultats et discussion	101
3.1.	Suivi de l'absorbance des échantillons de référence en fonction de la fréquence de calibration.....	101
3.2.	Suivi de l'évolution des prédictions sur le standard paille en fonction de la fréquence de calibration.....	104
3.3.	Relation entre l'absorbance minimale de la céramique blanche et le pourcentage de pixels de paille correctement prédits	106
3.4.	Prise en compte de la variabilité spectrale dans le set de calibration des modèles de discrimination.....	107
4.	Conclusions et perspectives.....	108
Chapitre V: Utilisation de l'imagerie hyperspectrale proche infrarouge et de la chimiométrie pour la discrimination de racines et de résidus de cultures extraits d'échantillons de sol.....		109
Introduction au Chapitre V		111
NIR hyperspectral imaging spectroscopy and chemometrics for the discrimination of roots and crop residues extracted from soil samples		112
1.	Summary.....	112
2.	Introduction	113
3.	Materials and Method.....	115
4.	Results and discussion	118
5.	Conclusions	125
6.	Acknowledgements	126
7.	References	127
Conclusions du Chapitre V.....		130
Chapitre VI: Utilisation de l'imagerie hyperspectrale proche infrarouge et de la chimiométrie pour quantifier la masse de racines extraites de carottes de sol		131
Introduction au Chapitre VI.....		133

1. Utilisation de l'imagerie hyperspectrale proche infrarouge combinée aux outils de la chimiométrie dans l'étude de systèmes racinaires.....	134
1.1. Résumé	134
1.2. Introduction	135
1.3. Matériels et méthodes.....	137
1.4. Résultats et discussion.....	140
1.5. Discussion générale et conclusion	149
1.6. Références bibliographiques.....	151
2. Validation de la régression permettant la prédiction de la masse de racines de froment	154
2.1. Matériels et méthodes.....	154
2.2. Résultats et discussion.....	155
2.3. Conclusions	158
3. Utilisation de l'imagerie hyperspectrale et de la chimiométrie en recherche agronomique.....	159
3.1. Introduction	159
3.2. Matériels et méthodes.....	159
3.3. Résultats et discussion.....	161
3.4. Conclusions	166
Conclusions du Chapitre VI	167
Chapitre VII: Quantification des teneurs en leghémoglobine au sein de nodosités de pois protéagineux (<i>Pisum sativum</i> L.) sur base de l'imagerie hyperspectrale proche infrarouge et de la chimiométrie	169
Introduction au Chapitre VII	171
Quantification of leghaemoglobin content in pea nodules based on near-infrared hyperspectral imaging spectroscopy and chemometrics.....	172
1. Abstract.....	172
2. Introduction	174
3. Materials and method	175
3.1. Nodule production	175
3.2. Leghaemoglobin quantification	176
3.3. NIR spectra acquisition and preprocessing.....	177
3.4. Linking NIR hyperspectral and multispectral imaging to leghaemoglobin content.....	178

3.5.	Statistical analysis	178
4.	Results and discussion	179
4.1.	Pea aerial biomass and nodule development	179
4.2.	PLS regression.....	180
4.3.	Validation on single nodules	182
4.4.	Multiple linear regression model applied to a restricted number of wavelengths	183
5.	Conclusion.....	185
6.	Acknowledgements	186
7.	References	186
	Conclusions au Chapitre VII et perspectives.....	189
	Chapitre VIII: Discussion et conclusions	191
	Discussion générale.....	193
1.	Opportunités offertes par l'imagerie hyperspectrale proche infrarouge couplée à la chimiométrie dans l'étude de systèmes racinaires	193
1.1.	Discrimination et quantification de racines et de résidus de cultures extraits de carottes de sol.....	194
1.2.	Discrimination et quantification de racines de différentes espèces	195
1.3.	Estimation de la composition chimique de systèmes racinaires ..	196
2.	Limites et contraintes liées à l'utilisation de l'imagerie hyperspectrale proche infrarouge et de la chimiométrie dans l'étude de systèmes racinaires. 197	
2.1.	Contraintes liées au traitement des échantillons.....	197
2.2.	Contraintes rencontrées lors de l'acquisition des images hyperspectrales	198
2.3.	Contraintes liées à la construction des modèles chimiométriques.....	201
2.4.	Limites des modèles chimiométriques	202
2.5.	Limites de la quantification racinaire basée sur un nombre de pixels	204
3.	Perspectives dans l'utilisation de l'imagerie hyperspectrale proche infrarouge et de la chimiométrie dans l'étude de systèmes racinaires.....	205
3.1.	Discrimination et quantification de racines vivantes et mortes ...	205
3.2.	Mesures de paramètres de l'architecture de systèmes racinaires.	206
3.3.	Couplage à de la reconnaissance de formes	206

3.4.	Identification des longueurs d'onde d'intérêt.....	207
3.5.	Amélioration d'autres méthodes de quantification racinaire	208
3.6.	Amélioration de la qualité des images acquises	209
	Conclusions générale.....	211
	Références bibliographiques	213

Liste des figures

Figure 1: Diversité de systèmes racinaires.	13
Figure 2: Prélèvement de carottes de sol.	29
Figure 3: La méthode <i>core break</i>	33
Figure 4: Le minirhizotron.	35
Figure 5: Les différentes zones du spectre lumineux.	45
Figure 6: Groupes fonctionnels du spectre proche infrarouge.....	48
Figure 7: Différents systèmes d'acquisition d'images hyperspectrales.....	50
Figure 8: Lissage par l'algorithme de Savitsky-Golay.....	53
Figure 9: Spectres proche infrarouges avant et après traitement avec la méthode Detrend.	54
Figure 10 : Représentation géométrique de la régression PLS.....	56
Figure 11 : Représentation géométrique de l'algorithme <i>Support Vector Machine</i> (SVM).....	57
Figure 12: Algorithme Duplex.	60
Figure 13: Zones d'ombre et réflexion spéculaire sur des images selon la configuration du système d'acquisition et la forme des échantillons	62
Figure 14: Schéma général présentant les étapes de la quantification des racines, de l'acquisition d'échantillons de racines permettant la calibration et la validation des modèles de discrimination jusqu'à l'application de ces modèles sur de nouveaux échantillons de racines.....	75
Figure 15: Vérin hydraulique et sonde permettant le prélèvement de carottes de sol.	77
Figure 16: Vue générale du système de lavage des racines.....	78
Figure 17: Echantillon de sol après lavage et séchage.	79
Figure 18: Tri de racines et de résidus de cultures présents au sein d'un échantillon de sol.....	80
Figure 19: Système d'acquisition d'images hyperspectrales proche infrarouge ...	81
Figure 20: Etalon vitesse.....	84
Figure 21: Prédiction d'une image hyperspectrale acquise sur un échantillon de grains de blé contenant six ergots.....	85
Figure 22: Standard paille.....	86
Figure 23: Instabilité du système d'acquisition perçu sur la céramique blanche.	87
Figure 24: Comparaison de masses de racine réelles et prédites.....	88
Figure 25: Spectres aberrants acquis sur un échantillon de racines de froment... ..	90
Figure 26: Sélection de 1000 spectres avec l'algorithme Duplex.	91
Figure 27: Arbre de classification dichotomique.....	93
Figure 28: Échantillon de racine contenant des résidus de cultures, des particules de terre et des cailloux.....	94
Figure 29: Visualisation de l'instabilité du système d'acquisition lors de la prédiction de la nature des pixels sur deux image hyperspectrale proche infrarouge acquise sur un même échantillon de référence ne contenant que des pailles..	100

Figure 30: Évolution de la valeur minimale de l'absorbance moyenne d'une céramique blanche au cours d'une journée d'acquisition et selon la fréquence de calibration de la caméra hyperspectrale..... 102

Figure 31: Evolution au cours des acquisitions de la valeur d'absorbance pour quatre longueurs d'onde des spectres moyens acquis sur le standard paille et pour trois fréquences de calibration du système d'acquisition 103

Figure 32: Valeurs du coefficient de détermination (r^2) de la relation établie entre la valeur minimale de l'absorbance moyenne de la céramique blanche et la valeur d'absorbance de quatre longueurs d'onde du spectre moyen des pailles de l'échantillon de référence. 103

Figure 33: Évolution du pourcentage de pixels de pailles correctement prédits par les modèles de discrimination au cours des acquisitions lorsque la calibration du système d'acquisition est réalisée en début de journée et après la 30^e image..... 104

Figure 34: Évolution du pourcentage de pixels de pailles correctement prédits par les modèles de discrimination au cours des acquisitions lorsque la calibration du système d'acquisition est réalisée toutes les deux heures..... 105

Figure 35: Évolution du pourcentage de pixels de pailles correctement prédits par les modèles de discrimination au cours des acquisitions lorsque la calibration du système d'acquisition est réalisée toutes les 30 minutes. 105

Figure 36: Relation entre la valeur minimale de l'absorbance moyenne de la céramique blanche et le pourcentage de pixels de paille correctement prédits sur les images hyperspectrales du standard paille..... 106

Figure 37: Comparaison des résultats des prédictions obtenues avec deux modèles de discrimination sur l'ensemble des images hyperspectrales acquises sur le standard paille au cours du suivi de la stabilité du système d'acquisition 107

Figure 38: Picture of the sieve with the different elements extracted from soil cores..... 116

Figure 39: Near infrared (NIR) hyperspectral imaging system 117

Figure 40: Dichotomous classification tree. 119

Figure 41: Picture of crop residues and predictions of NIR hyperspectral images of crop residues with SVM and PLS-DA dichotomous classification trees. 122

Figure 42: Picture of roots and predictions of NIR hyperspectral images of roots with SVM and PLS-DA dichotomous classification trees..... 123

Figure 43: Mean spectra of roots and crop residues of winter wheat after second derivative or Standard Normal Variate (SNV) transformation..... 125

Figure 44: SVM prediction on an NIR hyperspectral image acquired on a sample containing wheat roots, crop residues, small stones and clots and the sieve 125

Figure 45: Système d'acquisition d' images hyperspectrales proche infrarouge 138

Figure 46: Racines de froment, racines de pois, nodosités, fragments de pailles, particules de terre et cailloux extraits d'un échantillon de sol..... 139

Figure 47: Relation linéaire reliant la masse d'échantillons composés uniquement de racines de froment et le nombre de pixels classé comme étant des racines de froment sur les images hyperspectrales. 143

Figure 48: Densités racinaires (g/dm^3) d'une culture de froment d'hiver calculée dans trois horizons de sol. Effet du travail du sol..... 144

Figure 49: Spectres moyens de racines de pois et de froment, de résidus de culture de froment et de nodosités de pois acquis en proche infrarouge..	147
Figure 50 : Densités racinaires (g/dm^3) de cultures de froment d'hiver cultivé seul ou en association et de pois d'hiver cultivé seul ou en association.....	148
Figure 51: Relation entre les masses de racines mesurées et prédites.....	155
Figure 52: Distribution de la différence relative entre la masse racinaire prédite et la masse racinaire réelle en fonction de la masse réelle des échantillons de racines.....	156
Figure 53: Distribution de la densité de masse racinaire prédite (g/dm^3) d'une culture de froment d'hiver conduite en labour conventionnel ou travail du sol réduit avec ou sans restitution au sol des résidus de la culture précédente. Saison 2011-2012.....	162
Figure 54: Distribution de la densité de masse racinaire prédite (g/dm^3) d'une culture de froment d'hiver conduite en labour conventionnel ou travail du sol réduit avec ou sans restitution au sol des résidus de la culture précédente. Saison 2015-2016.....	163
Figure 55: Évolution des densités de masse racinaire prédite (g/dm^3) d'une culture du froment d'hiver entre quatre dates de prélèvement au cours de la saison culturale 2011-2012 et sur trois horizons de sol.....	164
Figure 56: Évolution des densités de masse racinaire prédite (g/dm^3) d'une culture de froment d'hiver entre deux dates de prélèvement au cours de la saison culturale 2015-2016 et sur six horizons de sol.....	165
Figure 57: Évolution des densités de masse de résidus de cultures prédites (g/dm^3) entre quatre dates de prélèvement au cours de l'année culturale 2011-2012 et sur trois horizons de sol.....	166
Figure 58A and B: Average near-infrared (NIR) spectra of pea nodules used for PLS model calibration before and after preprocessing.....	177
Figure 59: Leghaemoglobin content (mg leghaemoglobin $\cdot\text{g}^{-1}$ fresh nodule) measured with the cyanmethemoglobin method at each harvest time.....	180
Figure 60: Calibration results of the partial least square (PLS) regression model calibrated on leghaemoglobin content measured with cyanmethemoglobin method and predicted leghaemoglobin content of nodules on the basis of their near-infrared (NIR) spectra.....	181
Figure 61: Cross-validation and validation results of the partial least square (PLS) regression model.....	182
Figure 62: Predicted values of leghaemoglobin on individual nodules.....	183
Figure 63: Most relevant wavelengths selected using the graphical representation of coefficients against the corresponding wavelength.....	184
Figure 64: Identification of the seven wavelengths used to perform multiple linear regression analysis.....	185
Figure 65: Prédiction des teneurs en leghémoglobine au sein de nodosités présentes sur un système racinaire de pois protéagineux.....	190
Figure 66: Effet de la répartition de racines sur la quantité de pixels de racines pouvant être enregistrée.....	198
Figure 67: Prédiction d'une image hyperspectrale proche infrarouge acquise sur un échantillon de racines.....	203

Figure 68: Représentation schématique de l'acquisition de spectres sur des racines de différentes épaisseurs..... 205

Liste des tableaux

Tableau 1: Results obtained during calibration, cross-validation and validation steps of the discrimination models constructed respectively with support vector machine (SVM) or partial least squares discriminant analysis (PLS-DA).....	120
Tableau 2: Number and percentage of pixels predicted as background, soil, crop residues or roots on NIR hyperspectral images acquired on samples containing only one spectral class.	121
Tableau 3: Validation sur spectres proche infrarouge des trois modèles permettant la discrimination des racines du froment, des résidus de la culture précédente, des particules de sol restantes après lavage de l'échantillon et du tamis permettant la manipulation de l'échantillon.	141
Tableau 4: Validation des modèles de discrimination permettant la discrimination des résidus de la culture précédente et des racines de froment.....	142
Tableau 5: Validation sur spectres proche infrarouge des trois modèles permettant la discrimination de racines de froment et de pois, de nodosités, de résidus de la culture de froment, de particules de sol restantes après lavage de l'échantillon et du tamis permettant la manipulation de l'échantillon..	145
Tableau 6: Validation des modèles de discrimination permettant la discrimination de racines de froment et de pois..	145
Tableau 7: Protocole de prélèvements des échantillons de racines dans l'essai SOLRESIDUS.....	160
Tableau 8: Number of nodules, aerial dry biomass production and nodule leghaemoglobin content measured per growing pot at the four harvest dates.....	179

Liste des équations

Équation 1: Relation entre le nombre d'onde et la longueur d'onde.	44
Équation 2: Transformation de l'intensité brute et non corrigée du signal perçu au niveau de la camera hyperspectrale en réflectance relative.....	82
Équation 3: Relation entre la réflectance (R) et l'absorbance (A).....	83
Équation 4: Quantification de la masse de racines de froment (g) en fonction du nombre de pixels classés dans cette classe spectrale par les modèles d'analyse discriminante calibrés dans le cadre de l'étude ($r^2 = 0,98$).....	142
Équation 5: Quantification de la masse de racines de froment (g) en fonction du nombre de pixels classés dans cette classe spectrale par les modèles d'analyse discriminante calibrés dans le cadre de l'étude ($r^2 = 0,89$).....	146
Équation 6: Quantification de la masse de racines de pois (g) en fonction du nombre de pixels classés dans cette classe spectrale par les modèles d'analyse discriminante calibrés dans le cadre de l'étude ($r^2 = 0,76$).....	147

Liste des abréviations

CRA-W: Centre wallon de Recherches agronomiques

GxABT: Gembloux Agro-Bio Tech, Université de Liège

NIR: *Near Infrared*, proche infrarouge

PCA: *Principal Component Analysis*, analyse en composante principale

PLS: *Partial Least Square*, moindres carrés partiels

PLS-DA: *Partial Least Square Discriminant Analysis*, analyse discriminante par les moindres carrés partiels

ppp : points (ou pixels) par pouce

SVM: *Support Vector Machines*

R: Réflectance

RMSE: *Root Mean Square Error*, erreur quadratique moyenne

RMSEC: *Root Mean Square Error of Calibration*, erreur quadratique moyenne de calibration

RMSECV: *Root Mean Square Error of Cross Validation*, erreur quadratique moyenne de validation croisée

RMSEP: *Root Mean Square Error of Prediction*, erreur quadratique moyenne de prédiction

RPD: *Ratio of Prediction to Deviation*, rapport de l'écart-type des données sur l'erreur quadratique moyenne de prédiction

r^2 : coefficient de détermination

AVANT-PROPOS

Définitions et contexte

"Ce que l'on conçoit bien s'énonce clairement et les mots pour le dire viennent aisément"

Nicolas Boileau (1636-1711)

Entendons-nous bien!

En guise d'introduction à cette thèse, il me semble nécessaire de définir les termes qui en composent le titre de manière à ce que le lecteur puisse facilement comprendre ce qui suit.

Tout d'abord, il est important de définir ce qu'est l'**imagerie hyperspectrale proche infrarouge**. Cette technique d'analyse chimique est la combinaison de la **spectroscopie**, une technique d'analyse basée sur la relation existant entre la lumière réfléchie par la surface d'un objet et la composition chimique et physique de cet objet, et l'**imagerie** qui donne une représentation spatiale de l'objet. Cette technique permet donc d'analyser la composition d'un objet et de donner la répartition spatiale de sa composition. Dans ce sens, cette technique pourrait être comparée à la photographie qui donne une représentation d'un ou plusieurs objets et dont les couleurs dépendent de la composition de la surface de ces objets. Mais contrairement à la photographie classique qui n'est basée que sur trois couleurs (le rouge, le vert et le bleu qui par leurs combinaisons donnent toutes les couleurs visibles), cette technique d'analyse permet d'enregistrer de très nombreuses informations dans la lumière (plusieurs dizaines voire centaines). Elle est donc définie comme étant **hyperspectrale**, car elle permet d'enregistrer un grand nombre d'informations qui, mises ensemble, constituent un **spectre**, la représentation physique de la lumière réfléchie par la surface d'un objet. Un spectre est en quelque sorte la signature lumineuse de la composition chimique et physique de la surface de l'objet qui est étudié. Enfin, le titre de la thèse précise que cette technique est **proche infrarouge**, car elle ne se base pas sur la lumière visible, mais sur les composantes de la lumière qui ne sont pas visibles par l'œil humain et qui se situent dans la région proche de la lumière rouge.

Le titre de la thèse précise également que cette technique d'analyse a été utilisée avec la **chimiométrie** qui est la science qui permet de traiter les données acquises par analyse chimique. Elle permet entre autre de relier le profil des spectres acquis à la surface de l'objet étudié par spectroscopie à la composition de cet objet. La chimiométrie va donc en quelque sorte fournir le décodeur qui permet de connaître la composition d'un objet sur base de la lumière qu'il renvoie.

Ces techniques d'analyse et de traitement de données ont été utilisées pour étudier des **racines**, c'est-à-dire les organes cachés des plantes qui leur permettent de s'ancrer dans le sol et d'y puiser ce dont elles ont besoin pour se nourrir (eau et éléments nutritifs). Au sein du sol, ces racines peuvent avoir des tailles et des âges différents, elles ont une distribution temporelle et spatiale qui évolue et elles peuvent être associées à d'autres êtres vivants. Cet ensemble organisé et complexe est donc défini comme étant un **système racinaire**. L'étude de ces systèmes racinaires, particulièrement des plantes cultivées, est importante afin d'avoir une meilleure compréhension de la manière dont les plantes se nourrissent. D'autres raisons, telles que le transfert de carbone et d'azote atmosphérique vers les sols via les racines, rendent les études sur les systèmes racinaires très intéressantes, notamment dans le

contexte actuel de réchauffement climatique et de diminution des ressources nutritives pour les cultures.

Les racines se développant dans le sol, elles sont donc cachées ce qui rend leur étude compliquée. De nombreuses méthodes ont été proposées pour néanmoins étudier les racines, mais ces méthodes prennent souvent beaucoup de temps à cause des étapes de lavage et de tri des racines qui sont nécessaires avant de pouvoir réaliser des mesures sur celles-ci. Ces étapes de tri sont par exemple le tri des racines et d'autres éléments présents dans le sol ou le tri de racines de différentes espèces. Pour faciliter ce tri, l'utilisation de l'imagerie hyperspectrale proche infrarouge, combinée à la chimiométrie, a été proposée. Cette technique n'ayant jamais été utilisée auparavant pour cette application, nous avons étudié dans quelle mesure cette nouvelle utilisation était possible. Nous avons donc réalisé une **évaluation de ses potentialités** en cherchant ses avantages et ses limites par rapport à d'autres techniques déjà utilisées.

En résumé, cette thèse de doctorat s'est donc intéressée à la manière dont une technique d'analyse, qui donne à la fois des informations sur la composition d'un objet et la répartition de cette composition à la surface de cet objet, pouvait être utilisée en recherche agronomique pour tenter de mieux comprendre la partie cachée de la plante qui lui permet de s'ancrer dans le sol et de se nourrir. Cette étude s'est focalisée sur deux plantes cultivées: le froment et le pois protéagineux.

Contexte de la thèse et collaborations

L'essai SOLRESIDUS

La thèse de doctorat décrite dans le présent document a été entamée dans le cadre de l'essai pluridisciplinaire et de longue durée SOLRESIDUS mis en place par Gembloux Agro-Bio Tech, Université de Liège (Belgique) en collaboration avec le Centre wallon de Recherches agronomiques (CRA-W) sur la parcelle "Grand-Bordia" de la Ferme expérimentale de Gembloux Agro-Bio Tech (Gembloux, Belgique). Cet essai, mis en place en 2008, a pour objectif d'étudier les effets du travail du sol et de la gestion des résidus de cultures sur le développement de la culture en place et l'évolution du sol (structure, composition chimique, populations d'organismes vivants, émissions de gaz à effet de serre...). De nombreuses mesures ont depuis été réalisées sur cet essai, certaines ayant fait l'objet de travaux de fin d'études et de thèses de doctorat. Des descriptions détaillées de l'essai ont été réalisées par Degrunne et al. (2016) et Hiel et al. (2018).

Dans le cadre de cet essai, Françoise Vancutsem et Jérôme Pierreux¹ faisaient le suivi de la dégradation des résidus de cultures au cours du temps. Leurs mesures se basaient sur le prélèvement de volumes de sol et l'extraction de ces résidus de cultures par lavage à l'eau. D'autres éléments, tels que les racines des cultures, étaient extraits en même temps des échantillons de sol. Un tri des résidus de cultures et de ces autres éléments extraits des échantillons de sol devait donc être réalisé avant de pouvoir peser les résidus de cultures et les quantifier. Ce tri se faisait de manière manuelle. Vu la taille et le nombre d'éléments à trier, cette étape prenait beaucoup de temps, était fastidieuse et était entachée d'une imprécision due à la ressemblance de certains éléments.

L'utilisation de l'imagerie hyperspectrale

Afin de faciliter cette étape de tri, l'idée a été émise avec Vincent Baeten² d'utiliser une technique d'imagerie utilisant la spectroscopie proche infrarouge pour séparer les résidus de cultures et les racines sur base de leurs propriétés spectrales. Un test préliminaire avait été réalisé en avril 2012 et avait donné des résultats encourageants.

C'est sur base de ces travaux préliminaires et des échantillons qui avaient déjà été prélevés dans l'essai SOLRESIDUS, lavés et triés que j'ai entamé ma thèse de doctorat le 17 septembre 2012 avec le Professeur Bernard Bodson comme promoteur et encadré de près par Vincent Baeten et Juan Antonio Fernández Pierna². Mon comité de thèse était également constitué des Professeurs Gilles Colinet, Jean-Pierre Destain et Patrick du Jardin.

¹ Unité de Phytotechnie tempérée, Gembloux Agro-Bio Tech, Université de Liège, Belgique.

² Unité Qualité des produits, Département Valorisation des productions, Centre wallon de Recherches agronomiques, Gembloux, Belgique

La thèse était intitulée initialement " Détection et quantification des racines et des résidus de cultures présents dans le sol à l'aide de l'imagerie proche infrarouge" et était orientée sur la culture du froment d'hiver. Mais les recherches qui ont été menées dans le cadre de cette thèse se sont finalement, au fil des découvertes, des rencontres, des discussions et des opportunités, intéressées à l'utilisation de l'imagerie hyperspectrale proche infrarouge dans l'étude des systèmes racinaires dans un sens plus large et se sont peu intéressées aux résidus de cultures. La mise au point de modèles permettant la discrimination de racines et de résidus de culture de froment d'hiver est restée la base du travail de recherche, mais celui-ci s'est aussi élargi aux racines de la culture de pois protéagineux. Les racines de cette culture ont été étudiées en culture pure ou associée avec du froment d'hiver³ et une méthode de quantification de la leghémoglobine présente dans les nodosités se situant sur les racines de cette culture, et permettant la fixation symbiotique de l'azote atmosphérique, a été mise au point. Des essais ont également été menés sur les racines d'une culture de maïs et sur les résidus d'une culture de féverole, mais ceux-ci n'ont jamais été publiés. Enfin, une étude sur la stabilité du système d'acquisition des images hyperspectrales a été réalisée afin de s'assurer de la bonne qualité des acquisitions de données.

Un travail qui est le fruit de collaborations

Cette thèse a pu aboutir grâce aux collaborations avec plusieurs collègues, tout particulièrement Nicaise Kayoka Mukendi et Damien Vincke travaillant au CRA-W ainsi que Marie-Pierre Hiel et Benjamin Dumont de l'Unité de Phytotechnie tempérée (GxABT), qui m'ont apporté une aide très précieuse dans la construction des scripts permettant l'analyse des données acquises. Je tiens à les remercier chaleureusement!

Les résultats obtenus dans ce travail de recherche sont également le fruit de collaborations avec de nombreux étudiants qui, dans le cadre de travaux de fin d'études ou de travaux de groupes, ont participé au développement et à l'évaluation des modèles de discrimination et de quantification mis au point durant cette thèse. Guillaume Fraipont a ainsi utilisé les modèles permettant la discrimination et la quantification des racines et des résidus de culture de froment dans le cadre de son TFE en 2015⁴. Il a appliqué ces modèles sur des échantillons prélevés durant la saison culturale 2011-2012 dans une culture de froment d'hiver implantée sur l'essai SOLRESIDUS. C'est dans le cadre de son travail que l'instabilité du système d'acquisition des images hyperspectrales a été mise en évidence. Il a ensuite réalisé des prélèvements d'échantillons de racines dans le même essai ainsi que dans un

³ Cette recherche a été menée dans les essais conduits par Jérôme Pierreux dans le cadre du projet "Produire durablement des graines riches en protéines en optimisant la conduite de la culture associée de pois protéagineux d'hiver et de froment d'hiver" financé par le Service Public de Wallonie, DGO3.

⁴ Fraipont G. "Détection et quantification des racines et des résidus de cultures de froment d'hiver (*Triticum aestivum* L.) par imagerie hyperspectrale proche infrarouge". Travail de fin d'études. Gembloux Agro-Bio Tech, Université de Liège. Année académique 2014-2015.

essai sur la fertilisation azotée du froment d'hiver durant la saison culturale 2015-2016 dans le cadre de sa thèse. Les quantités de racines présentes dans ces échantillons prélevés dans l'essai SOLRESIDUS ont ensuite été déterminées dans le cadre du travail de groupe mené par Affi Jean-Paul Attikora, Florian Boucheron, Xavier Bordet et François Godin durant l'année académique 2017-2018⁵. Les racines présentes dans les échantillons prélevés dans l'essai sur la fertilisation azotée ont quant à elles été quantifiées par Adrien Boufflette dans le cadre de son TFE en 2018⁶. Adrien Boufflette a également prélevé et quantifié des racines de froment d'hiver dans un essai sur la fertilisation azotée mené en 2017-2018. Dans le cadre de son travail, de nouveaux modèles de discrimination ont dû être calibrés pour tenir compte d'autres résidus de cultures présents dans les échantillons.

Pour leur travail de groupe réalisé durant l'année académique 2016-2017; Ariel Doudet, Florent Druart, Laëtitia Machiels, Audrez Sougnez et Marine Torfs ont produit les nodosités et ont réalisé les mesures qui ont permis de mettre au point le modèle de quantification des teneurs en leghémoglobine sur base de spectres proche infrarouge⁷. Enfin, Louis Dumont de Chassart a appliqué les modèles de discrimination des racines de froment d'hiver et de pois d'hiver et le modèle de quantification de la leghémoglobine sur des échantillons qu'il a prélevés dans le cadre de son TFE en 2018⁸.

C'est donc grâce aux travaux réalisés par ces étudiants que les modèles mis au point ont pu être améliorés et appliqués sur de nombreux échantillons. Merci à eux!

⁵ Godin F., Bordet X., Attikora A.J.P., Boucheron F. "Etude de l'effet de la profondeur du travail du sol sur le développement du système racinaire du froment d'hiver". Projet de groupe. Gembloux Agro-Bio Tech, Université de Liège. Année académique 2017-2018.

⁶ Boufflette A. "Etude de l'influence de la fertilisation azotée sur le développement du système racinaire du froment d'hiver (*Triticum aestivum* L.) par imagerie hyperspectrale proche infrarouge" Travail de fin d'études. Gembloux Agro-Bio Tech, Université de Liège. Année académique 2017-2018.

⁷ Doudet A., Druart F., Machiels L., Sougnez A., Torfs M. "Étude des potentialités de l'imagerie hyperspectrale proche infrarouge pour la mesure de la teneur en leghémoglobine au sein des nodosités de pois protéagineux (*Pisum sativum* L.)". Projet de groupe. Gembloux Agro-Bio Tech, Université de Liège. Année académique 2016-2017.

⁸ Dumont de Chassart L. "Evaluation de l'activité fixatrice d'azote de nodosités de pois protéagineux par imagerie hyperspectrale proche infrarouge" Travail de fin d'études. Gembloux Agro-Bio Tech, Université de Liège. Année académique 2017-2018.

CHAPITRE I

Revue bibliographique

1. L'étude des racines

Les racines constituent la partie cachée de la plante, "*the hidden half*" comme le dit très justement le titre du recueil édité par Eshel et Beeckman en 2013. Au cours des dernières années, elles ont été le sujet de très nombreuses études, car leurs rôles dans le développement des plantes et dans le stockage de carbone dans les sols doit de plus en plus être pris en compte. Or, bien que leurs rôles soient de mieux en mieux connus, l'étude de leur développement et des interactions auxquelles elles sont sujettes dans le sol reste difficile par la présence même du sol qui limite leur observation. De nombreuses techniques d'observation ont donc été mises au point et continuent d'être développées afin de pouvoir les étudier. Cette première section de la revue bibliographique définira le système racinaire et ses représentations. Elle présentera ensuite la dynamique de développement des racines, leurs rôles, les facteurs qui influencent leur développement, l'importance de les étudier, quelques paramètres couramment mesurés, plusieurs méthodes utilisables dans leur étude et enfin les limites rencontrées dans leur étude.

1.1. Le système racinaire

Le **système racinaire** d'une plante est constitué d'une structure arborescente de racines de différents ordres topologiques (primaires et latérales d'ordres supérieurs). Elles ont des origines, des dimensions, des comportements de croissance et des âges différents qui déterminent en partie leurs fonctions. De ce fait, leur distribution spatiale ainsi que la distribution de leur activité dans le sol évoluent au cours de la croissance de la plante.

L'architecture du système racinaire dépend de l'espèce, des conditions pédoclimatiques, des pratiques agronomiques, de la compétition pour les ressources et de nombreux autres facteurs. Par leur effet sur la distribution des racines et des ressources, ces facteurs influencent la manière dont le système racinaire remplit ses fonctions. Les liens entre l'architecture racinaire et les facteurs qui l'influencent sont néanmoins encore peu explorés.

1.2. Les représentations des systèmes racinaires

La représentation spatiale du système racinaire peut être décrite soit en s'intéressant à l'organisation des racines au sein d'une architecture racinaire, soit en étudiant la distribution des racines dans les différents horizons du sol.

L'**architecture du système racinaire** est la configuration spatiale du système racinaire d'une plante dans le sol (de Dorlodot et al. 2007). Elle fait référence au déploiement géométrique de ses axes racinaires (Lynch 1995). Elle permet de décrire la **forme** et la **structure** du système racinaire (Hodge et al. 2009; de Dorlodot et al. 2007). L'âge et le type de racines d'une plante constituent également l'architecture du système racinaire (Zhu et al. 2011).

La **distribution des racines** fait référence à la présence des racines (plutôt qu'à leur orientation) à un endroit du sol (Lynch 1995).

Les études sur l'architecture racinaire portent soit sur l'ensemble d'un système racinaire, soit sur un sous-ensemble important du système racinaire d'une plante individuelle. Par contre, les mesures réalisées sur la distribution racinaire en agriculture et dans les communautés végétales naturelles portent sur les racines de plusieurs plantes ou de plusieurs espèces (Lynch 1995).

Quelques systèmes racinaires avec des architectures contrastées sont représentés à la Figure 1.

1.2.1. La forme du système racinaire

La **forme du système racinaire** est définie par la localisation des racines dans l'espace et la manière dont les racines occupent le sol. La profondeur d'enracinement, l'extension latérale des racines, la longueur des racines, la densité racinaire, la ramification et la surface totale de racines sont des exemples de paramètres qui permettent de décrire cette forme (Kuijken et al. 2015; Hodge et al. 2009).

1.2.2. La structure du système racinaire

La **structure du système racinaire** est définie par les constituants du système racinaire, tels que les racines et les segments de racines, et par leurs relations, telles que la connexion entre les racines et leur ordre (Hodge et al. 2009).

Cette structure est définie par la génétique de la plante considérée (de Dorlodot et al. 2007). A titre d'illustration des différences pouvant exister entre espèces au niveau de la structure du système racinaire, prenons les cultures du froment et du pois protéagineux. Le système racinaire du froment (*Triticum aestivum* L.) est fasciculé. Les racines sont nombreuses et peu ramifiées. Il est constitué de deux systèmes racinaires distincts qui se développent successivement (Gate 1995; Moule 1971). Le système racinaire primaire (séminal) est actif de la levée jusqu'au début du stade tallage de la culture. Le système racinaire secondaire se met en place lors de l'apparition des talles. Les racines de ce deuxième système racinaire apparaissent à l'aisselle des talles et se ramifient. Le système racinaire du pois protéagineux (*Pisum sativum* L.) est quant à lui de type pivotant. Il est composé d'une racine principale (pivot) et de racines latérales primaires et secondaires (Munier-Jolain et al. 2005).

1.2.3. La distribution racinaire

La **distribution racinaire** décrit la manière dont les racines se répartissent au sein du sol. Elle est souvent exprimée en termes de longueur, de masse ou de surface racinaire par horizon de sol. La profondeur de ces horizons de sol est variable d'une étude à l'autre. Elle peut également être exprimée en terme de distance par rapport à la tige ou de position entre des plantes voisines (Lynch 1995). Cette distribution racinaire est fonction de l'espèce étudiée (Fan et al. 2016), du génotype, de l'âge des plantes et du milieu dans lequel elle se développe. Les racines ne sont donc pas réparties uniformément dans le sol et de nombreuses études ont montré que les racines se répartissent dans le sol de telle manière que leur longueur et leur masse décroît de manière exponentielle avec la profondeur du sol (Gregory 2006). Leurs distributions spatiales et temporelles déterminent de manière importante

l'accessibilité à l'eau et aux nutriments pour la plante (Plaza-Bonilla et al. 2014; Gregory 2006).

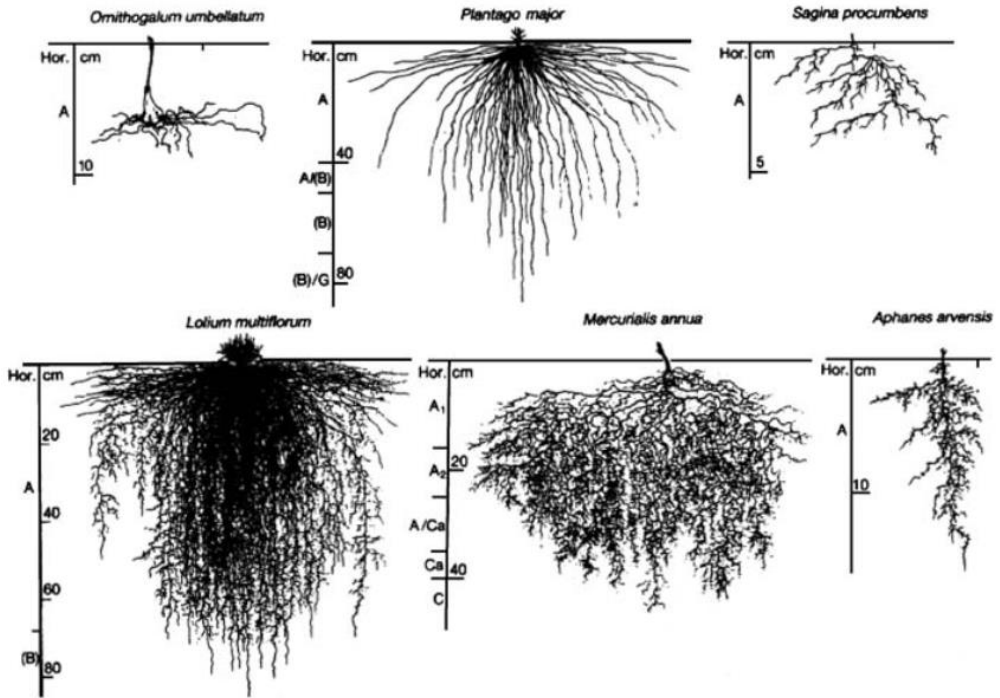


Figure 1: Diversité de systèmes racinaires. Dans le haut de la figure, les systèmes racinaires sont peu ramifiés, contrairement aux systèmes racinaires représentés dans le bas de la figure. La dominance d'un axe principal unique augmente de la gauche vers la droite de la figure. Illustration tirée d'un article de Hodge et al. (2009).

Dans une compilation de données effectuée pour 11 espèces cultivées en régions tempérées, Fan et al. (2016) ont mis en évidence chez les monocotylédones une relation linéaire légèrement décroissante entre la profondeur de prélèvement et la distribution du ratio de la masse sur la longueur racinaire. Pour les dicotylédones, cette relation décroissante avait une allure exponentielle. Les racines les plus épaisses étaient donc majoritairement dans les horizons supérieurs du sol et cela était d'autant plus observé chez les dicotylédones. De manière générale, les racines des céréales et des légumineuses étaient réparties de manière homogène dans le profil de sol alors que les racines du colza étaient observées majoritairement dans l'horizon supérieur du sol. Pour l'ensemble des données analysées dans cette étude, la profondeur estimée à laquelle 50 % des racines étaient accumulées variait entre 8 et 20 cm de profondeur (pour une profondeur maximale d'enracinement variant de 78 à 177 cm).

Des études menées sur l'enracinement de cultures de froment d'hiver (*Triticum aestivum* L.) (Qin et al. 2004; Wilhelm et al. 1982) et de maïs (*Zea mays* L.) (Cai et al. 2014; Wang et al. 2014) ont montré que la majorité du système racinaire était concentrée dans l'horizon supérieur du sol. Selon les études, plus de 60 % du système racinaire mesuré sur 80 à 120 cm de profondeur était situé entre 20 et 40 cm de profondeur.

Chez le pois protéagineux, la profondeur maximale d'enracinement varie de 50 à 80 cm en fonction du génotype et des conditions pédoclimatiques. Le système racinaire des légumineuses est donc peu développé par rapport au système racinaire des céréales. Tout au long du développement de la culture, la majorité des racines (jusqu'à 90 %) se situent dans l'horizon superficiel. A maturité, à biomasse aérienne égale pour les céréales et le pois, la longueur des racines des légumineuses atteint seulement la moitié de celle des céréales (Munier-Jolain et al. 2005).

1.3. Dynamique de développement racinaire

De la levée jusqu'à la mort d'une plante, le système racinaire suit une dynamique de croissance puis de sénescence. Cette dynamique est influencée par l'espèce considérée et les conditions de croissance de la plante.

En froment d'hiver, Wang et al. (2014) ont mesuré une augmentation rapide de la densité de masse racinaire lors de la reprise de la végétation en sortie d'hiver. Cette densité racinaire était maximale lors de la floraison et déclinait lentement jusqu'à la maturité de la culture. Dans une comparaison la biomasse racinaire entre le tallage et la floraison de deux cultures de céréales d'hiver, le froment d'hiver et l'orge d'hiver (*Hordeum vulgare* L.), Plaza-Bonilla et al. (2014) ont observé une densité de surface racinaire plus importante à la floraison des deux cultures. Kätterer et al. (1993) ont observé que la biomasse racinaire du froment atteignait son plus haut niveau en juin, lors de l'émergence des épis. Cette biomasse racinaire augmentait jusqu'à la sortie des épis et l'anthèse, puis elle diminuait jusqu'à la récolte de la culture. La réduction de biomasse racinaire la plus importante était observée dans les 10 premiers cm du sol. Huang et al. (2012) ont aussi observé que la biomasse racinaire du froment était la plus importante à la floraison de la culture et qu'elle diminuait par la suite. Plusieurs autres études ont aussi montré qu'après la floraison, il y a généralement une décomposition des racines les plus fines, et une réduction de la longueur, de la surface et de la masse racinaire (Plaza-Bonilla et al. 2014). Muñoz-Romero et al. (2010) rapportent aussi que la croissance du système racinaire du froment décroît ou cesse après l'anthèse.

Selon une étude de Guan et al. (2014), c'est également à la floraison que le développement racinaire en termes de biomasse est le plus important en maïs. Dans cette étude, la biomasse racinaire avait été mesurée jusqu'à un mètre de profondeur. Pour cette culture, la croissance racinaire a lieu de la levée jusqu'à la floraison. Elle est assez lente au début du cycle de la culture et s'accélère à partir de la montaison, moment où l'augmentation de la masse des parties aérienne de la plante devient également rapide (Tardieu & Manichon 1987).

Chez le pois protéagineux (*Pisums sativum* L.), la mise en place du système racinaire se fait en trois phases (Munier-Jolain et al. 2005):

- Une élongation rapide de la racine principale qui s'accompagne d'un développement horizontal et superficiel des racines latérales de premier ordre durant la période végétative des parties aériennes de la plante;
- Une forte augmentation de la masse racinaire due au développement des racines latérales de second ordre durant la floraison et la formation des graines. C'est au cours de cette phase que la profondeur maximale d'enracinement est atteinte;
- Un ralentissement du développement racinaire et l'émission des dernières racines latérales au cours du remplissage des graines.

Selon les études, la croissance racinaire du pois protéagineux s'arrête après la floraison ou à maturité physiologique de la culture (Munier-Jolain et al. 2005).

Au niveau des racines de pois, l'apparition des nodosités et la fixation symbiotique suivent également des dynamiques précises (Tricot-Pellerin et al. 1994). Les premières nodosités s'installent sur le pivot et sont visibles dès le stade deux feuilles de la culture. Elles s'installent ensuite aussi sur les racines latérales de premier ordre durant toute la phase végétative. La biomasse totale de nodosité atteint son maximum un peu avant la floraison. La fixation symbiotique augmente quant à elle jusqu'au début de la floraison, stagne durant celle-ci et chute au début du remplissage des grains.

Selon Munier-Jolain et al. (2005), il semblerait que les légumineuses aient une stratégie opportuniste. Dans un premiers temps, elles mettent en place un système racinaire peu développé. Par la suite, elles ont la capacité d'accroître leur masse racinaire en fonction des variations de la disponibilité en eau du sol. Chez les céréales par contre, le système racinaire est toujours bien développé, quelles que soient les conditions environnementales.

1.4. Le rôle des racines

Les racines sont le lien vital entre les ressources souterraines et le développement des plantes (Plaza-Bonilla et al. 2014). Leurs rôles principaux sont d'ancrer et de stabiliser la plante dans le sol ainsi que d'assimiler l'eau et les nutriments (Hodge et al. 2009; Atkinson 2000), mais elles ont aussi de nombreux autres rôles associés. En effet, les racines jouent le rôle d'organe de réserve, elles sont le site de synthèse d'hormones importantes nécessaires au développement de la plante, elles sont le lieu d'interactions de la plante avec des plantes de la même espèce ou d'autres espèces et avec des microorganismes. De plus elles jouent un rôle important dans l'apport de carbone dans les sols ainsi que dans leur stabilisation (Bardgett et al. 2014; Plaza-Bonilla et al. 2014; Kätterer et al. 2011; Zhu et al. 2011; Atkinson 2000). Les racines ne sont donc pas de simples organes de prélèvement d'eau et de nutriments au sein d'un sol qui ne servirait que de "frigo", mais bien des organes qui vont s'adapter à leur milieu de croissance, qui vont permettre l'interaction de la plante avec ce milieu complexe, riche et vivant qu'est le sol, au sein duquel la plante pourra aussi interagir avec d'autres plantes, et qui vont pouvoir le modifier. En effet, les racines, et les

organismes qui leur sont associés, peuvent avoir un impact majeur sur la structure du sol. Par exemple, les racines et des hyphes mycorhiziens sont impliqués dans la création d'agrégats de sol résistants à l'eau qui permettent la stabilisation des sols et la création de réserves en nutriments (Atkinson 2000). De plus, l'apport de matière organique dans les sols via les racines influence des propriétés importantes des sols telles que la capacité d'échange cationique. Une synthèse réalisée par Atkinson (2000) développe plus largement ces rôles des racines et l'importance de leur étude.

1.5. Facteurs influençant le développement racinaire

La configuration spatiale du système racinaire des plantes varie fortement en fonction de l'espèce, de la composition du sol et particulièrement de la disponibilité en eau et en nutriments (Hodge et al. 2009). Au cours de l'évolution, les racines ont évolué de telle manière à être extrêmement adaptées à leur environnement local. Leur croissance, leur morphologie et leur physiologie sont intimement liées au génotype de la plante, aux propriétés du milieu dans lequel elles se développent et à l'interaction des deux (Downie et al. 2014; Zhong et al. 2009). L'extension du système racinaire dans le temps et l'espace est gouvernée par des règles de développement génétiquement déterminées et modulées par l'action des stimuli biotiques et abiotiques de l'environnement. Ces stimuli peuvent entre autres être induits par l'action de l'homme sur les plantes, notamment par les pratiques agronomiques appliquées sur les cultures.

1.5.1. Le facteur génétique

La génétique de la plante définit l'architecture de son système racinaire. Elle détermine aussi la biomasse racinaire qui peut être produite, le renouvellement des racines, leur distribution et leur profondeur maximale d'enracinement (Fan et al. 2016). Au sein d'une même espèce, des différences en termes d'architecture racinaire sont observées entre génotypes (Hochholdinger 2009). De nombreux espoirs sont mis actuellement dans ce facteur de variabilité pour minimiser l'impact négatif du changement climatique et de la raréfaction des ressources en eau et en nutriments sur les productions agricoles. De nombreuses études sont dès lors menées pour identifier les gènes qui permettraient de modifier la distribution racinaire dans le sol de manière à améliorer l'absorption d'eau et de nutriments (de Dorlodot et al. 2007).

1.5.2. Les facteurs abiotiques

Les principaux facteurs abiotiques qui influencent le développement racinaire sont la structure et la composition du sol. Ces facteurs peuvent affecter l'élongation racinaire et impacter le taux de prélèvement d'eau et de nutriments par les racines (Downie et al. 2014). Il peut s'agir de la disponibilité en eau et en nutriments, de l'aération du sol, de sa densité, de la présence de macro pores, de la taille des agrégats, de son pH, de sa salinité et de sa teneur en élément tel que l'aluminium par exemple (Hodge et al. 2009; Gregory 2006). Si la plante subit des stress abiotiques au niveau de ses racines, le rendement final de la culture est impacté (Downie et al. 2014).

La structure du sol

Les conditions physiques du sol ont des effets importants à la fois sur la manière dont le système racinaire peut s'étendre dans le sol et sur les transferts d'eau, de gaz et de nutriments vers et à partir des racines (Downie et al. 2014; Gregory et al. 2013). Lorsque la densité du sol augmente, le contact entre le sol et les racines augmente ce qui permet un meilleur approvisionnement en eau et en nutriments de la plante. Si la densité du sol est élevée, la longueur des racines diminue et leur extension radiale augmente (Tracy et al. 2012). Dans un sol ayant une haute densité ou une forte teneur en eau, le taux d'élongation des racines et le nombre de racines latérales peuvent être réduits. Il en résulte alors aussi une diminution de la croissance aérienne de la plante (Downie et al. 2014). Ce type de développement racinaire s'observe dans des sols compactés.

Si par contre la densité du sol diminue et si le contact entre le sol et les racines diminue, les quantités d'eau et de nutriments prélevées par unité de longueur racinaire diminuent. Un sol avec de larges pores n'est donc pas favorable à la croissance racinaire; les racines préfèrent un réseau de pores étroits (Gregory et al. 2013).

La composition chimique du sol

Lorsqu'une plante pousse dans des conditions limitantes en termes de nutriments, elle va augmenter l'allocation de la biomasse dans ses racines et son développement racinaire sera modifié (Zhu et al. 2011). La disponibilité en nutriments tels que le phosphore peut par exemple causer des modifications dans l'architecture du système racinaire et l'anatomie des racines (Downie et al. 2014). Mais selon les ressources en nutriments, la stratégie mise en place par la plante peut être différente. Rasmussen et al. (2015) expliquent que pour des nutriments immobiles tels que les phosphates, la prolifération locale de racines permet une meilleure exploitation de la ressource. Si les nutriments sont mobiles, tels que les nitrates, le rôle d'une prolifération locale des racines est moins clair. Elle pourrait permettre à une plante isolée d'avoir un avantage compétitif vis-à-vis des plantes voisines, mais cela n'améliore pas le prélèvement général d'azote par la culture (Robinson et al. 1999). Pour le prélèvement d'azote, il semblerait donc qu'un enracinement plus profond serait plus intéressant, car celui-ci donne accès à la culture à un plus grand volume de sol et aux ressources qu'il contient (Rasmussen et al. 2015).

Selon une synthèse réalisée par Han et al. (2015), les nutriments présents dans la couche de sol située en dessous de l'horizon labouré pourraient contribuer jusqu'à deux tiers des besoins totaux de la plante. Cette contribution dépendrait des nutriments et des cultures considérés. Des variations ont été observées entre les études prise en compte dans cette synthèse. La proportion d'azote prélevé par le froment d'hiver en dessous de l'horizon labouré était de 18 à 38 %. En froment de printemps, les proportions des prélèvements totaux de potassium et de phosphore étaient à cette profondeur de 34 % et de 30 à 85 % respectivement.

La teneur en eau du sol

La manière dont la plante va développer son système racinaire dépend des ressources auxquelles elle a accès au cours de son développement. La teneur en eau

du sol est un des facteurs qui influence fortement la densité et la profondeur racinaire: une teneur en eau modérée dans les sols peut significativement augmenter la densité de longueur racinaire dans le sol (Wang et al. 2014). L'effet de la teneur en eau du sol, plus particulièrement l'effet d'un déficit en eau, peut cependant avoir des effets différents selon le stade de la culture et l'espèce considérée.

Des mesures réalisées en froment d'hiver par Muñoz-Romero et al. (2010) ont montré qu'en conditions sèches, le développement racinaire était lié au moment où de l'eau était disponible par rapport au stade de développement de la plante. Ainsi, un apport d'eau suffisamment important au tallage de la culture permettait le développement le plus important du système racinaire par la production d'un grand nombre de talles et l'initiation d'un grand nombre de racines. Les plantes étaient ensuite plus à même de supporter une période sèche. La teneur en eau du sol au moment de la plantation et la pluviométrie enregistrée avant et après ce stade de développement avaient beaucoup moins d'effet sur le développement racinaire.

Benjamin et Nielsen (2006) ont de leur côté étudié l'effet d'un déficit en eau sur la distribution racinaire de trois cultures de légumineuses: le pois chiche (*Cicer arietinum* L.), le pois (*Pisum sativum* L.) et le soja (*Glycine max* L. Merr.). Ils ont montré dans leur étude que la distribution racinaire du soja n'était pas influencée par un déficit hydrique. Au contraire, les deux autres cultures étaient fortement impactées et développaient leur système racinaire plus en profondeur en condition de stress.

1.5.3. Les facteurs biotiques

L'architecture du système racinaire et la croissance des racines sont également influencées par des facteurs biotiques tels que des microorganismes et des macroorganismes pathogènes et saprophytes ainsi que par l'association symbiotique avec des mycorhizes arbusculaires et des bactéries promotrices de croissance (Downie et al. 2014). Ces dernières peuvent par exemple modifier l'architecture racinaire en augmentant la ramification racinaire et en favorisant le développement des poils absorbants (Vacheron et al. 2013).

Si l'environnement de la plante influence le développement de son système racinaire, l'inverse est vrai aussi et le système racinaire de la plante va à son tour modifier son environnement. Les modifications du sol environnant le système racinaire d'une plante sont dues à la croissance et à la respiration racinaire, aux rhizodépôts, à l'émission de composés volatils ainsi qu'au prélèvement d'eau et de nutriments. Ces modifications sont d'ordre biologique, chimique et physique et peuvent être observées sur une distance variable par rapport aux racines si bien que la rhizosphère, c'est-à-dire le volume de sol environnant les racines et affecté par leur présence, peut avoir des dimensions différentes dans le temps et l'espace selon le processus considéré (Gregory 2006).

De nombreuses recherches ont démontré que des composés émis par les racines jouent le rôle de messagers dans la communication et l'interaction des racines entre elles, entre des racines et des microorganismes et entre des racines et la faune (Delory et al. 2016; Gregory 2006). Les interactions souterraines entre les plantes

influencent leur croissance, leur survie et les propriétés de leurs racines (Atkinson 2000).

La compétition interspécifique pour les ressources du sol et les interactions directes entre plantes via leurs racines influencent également le développement racinaire des plantes (Rewald et al. 2012). Dans une synthèse des interactions facilitatrices jouées par les racines dans des cultures en associations⁹, Hauggaard-Nielsen et Jensen (2005) citent une série de mécanismes par lesquelles les plantes, via leurs racines, peuvent améliorer des conditions environnementales hostiles et accroître la disponibilité en ressources pour d'autres espèces. Ils listent dans cet article une série d'interactions facilitatrices potentielles directes et indirectes. Ces interactions directes sont par exemple le transfert d'azote fixé par symbiose par des légumineuses cultivées en associations avec des plantes n'ayant pas cette capacité de fixation, la libération de phosphore provenant de la matière organique grâce à des enzymes extracellulaires ou la dissolution de phosphore inorganique fixé dans les sols grâce à un abaissement du pH par des légumineuses. Les exemples d'interactions facilitatrices indirectes qu'ils citent sont l'élimination de compétiteurs potentiels via des isothiocyanates produites par des crucifères et la stimulation de microorganismes bénéfiques du sol tels que les champignons mycorhiziens arbusculaires qui pourraient lier entre elles des plantes de différentes espèces et permettre le transfert de nutriments entre les individus d'une communauté végétale.

1.5.4. Les facteurs anthropiques

L'homme peut aussi influencer le développement racinaire des plantes. Ce point décrira les moyens d'action que l'homme a sur le développement racinaire et sera focalisé sur les pratiques agricoles.

Le travail du sol

Selon la manière dont il est réalisé, le type de sol et les conditions climatiques rencontrées, le travail du sol peut fortement affecter les propriétés du sol. Sa réalisation affecte la structure du sol, ses propriétés mécaniques, sa porosité ainsi que sa teneur en eau et en oxygène (Guan et al. 2014; Martínez et al. 2008). Cette modification des teneurs en eau du sol s'observe particulièrement dans l'horizon proche de la surface du sol et affecte son utilisation par les plantes (Muñoz-Romero et al. (2010).

La manière dont le travail du sol influence le développement racinaire semble être fortement liée au type de sol et aux conditions climatiques rencontrées par la culture. Donner une conclusion générale sur la manière dont un type de travail du sol influence à lui seul le développement racinaire n'est donc pas possible (Martínez et al. 2008).

Cependant, certaines études ont montré que l'impact du travail du sol sur la distribution racinaire était évident dans l'horizon de sol affecté par le travail du sol :

⁹ La culture en association est définie comme la croissance d'au moins deux cultures de manière simultanée sur une même surface. Elle est basée sur la gestion des interactions entre les plantes cultivées pour maximiser leur croissance et leur productivité (Hauggaard-Nielsen & Jensen 2005).

dans un système sans travail du sol, il y a une accumulation des racines dans les cinq premiers cm du sol par rapport à un système où le labour conventionnel est utilisé et l'opposé est vrai dans les horizons plus profonds. L'impact d'un système sans travail du sol sur le développement racinaire dépend cependant du temps entre la mise en place de ce système et la prise de mesures (Qin et al. 2004).

La fertilisation minérale et l'irrigation des cultures

L'interaction et la complémentarité entre l'eau et les nutriments sont les facteurs principaux qui affectent la production agricole : l'eau augmente la disponibilité des nutriments alors que les nutriments augmentent l'efficacité d'utilisation de l'eau par la plante (Wang et al. 2014). L'effet d'une fertilisation sur le développement du système racinaire peut donc dépendre des teneurs en eau du sol (Wang et al. 2014). De plus le prélèvement des nutriments dépend des nutriments disponibles à proximité des racines et du développement du système racinaire. Comme l'explique Atkinson (2000), l'importance du système racinaire dans le prélèvement des nutriments est donc variable. Lorsque les nutriments sont présents en abondance et que les conditions du sol sont telles que ces ressources sont facilement assimilables par les plantes, la longueur des racines et les caractéristiques du système racinaire sont peu importantes. Si la disponibilité en nutriments diminue, les propriétés du système racinaire deviennent très importantes.

De nombreuses études ont été menées sur la **fertilisation azotée**. Celles-ci montrent des effets qui peuvent varier selon la dose apportée, le type de sol, l'année d'étude et le climat. Des études menées sur la relation entre la croissance racinaire et la disponibilité en azote ont montré une tendance générale à la prolifération des racines dans les zones du sol riches en azote et une suppression de la croissance racinaire dans les zones à faible teneur en azote (Rasmussen et al. 2015). A contrario, une faible disponibilité en azote dans un sol améliore la biomasse racinaire et une haute teneur en azote dans un sol réduit cette biomasse. En régions sèches, la fertilisation améliore la croissance de la jeune plante et favorise la croissance rapide des racines en profondeur ce qui lui permet d'absorber l'eau accumulée en profondeur durant l'interculture. Cependant, une fertilisation excessive supprime la croissance racinaire dans les horizons profonds du sol et réduit donc l'utilisation de l'azote (Wang et al. 2014).

L'**irrigation** est une autre technique agricole qui peut influencer le développement du système racinaire. Plusieurs études ont montré que l'irrigation limitait la profondeur d'enracinement des cultures (Kätterer et al. 1993) mais un apport d'eau à un stade précis d'une culture peut augmenter significativement la densité de masse racinaire (Wang et al. 2014).

La rotation des cultures

La succession des cultures sur une même parcelle peut également influencer le développement racinaire de la culture étudiée. Han et al. (2015) ont ainsi montré qu'un froment d'hiver cultivé après une culture de chicorée fourragère (*Cichorium intybus* L.) bénéficiait de la porosité créée par cette culture plus que s'il était cultivé après une culture de fétuque élevée (*Festuca arundinacea* Schreb.). Les pores créés par les racines pivotantes de la chicorée constituaient un chemin d'accès rapide pour

les racines du froment vers les horizons profonds du sol permettant ainsi une meilleure exploration du sol et donc une amélioration de l'absorption d'eau et de nutriments dans ces horizons profonds en début de culture.

L'association de cultures

L'association de cultures, telle que l'association de céréales et de légumineuses, entraîne aussi des modifications du développement racinaire. La compétition des racines aurait un effet plus important dans le développement de la plante que la compétition au niveau aérien (Li et al. 2006). Dans une étude réalisée en rhizotrons, Ramirez-Garcia et al. (2015) ont mis en évidence que dans les horizons les plus profonds (80-120 cm), l'orge (*Hordeum vulgare* L.) cultivée en association avec de la vesce (*Vicia sativa* L.) produisait plus de racines. Lorsque la longueur racinaire était exprimée par plante, l'orge atteignait des valeurs deux fois plus élevées en association qu'en culture pure. Dans une étude réalisée sur une autre association, Li et al. (2006) ont montré que la culture associée de froment de printemps (*Triticum aestivum* L.) et de maïs (*Zea mays* L.), par alternance de six rangs de blé et de deux rangs de maïs sur une largeur de 1.5 m¹⁰, entraînait une augmentation de la densité de longueur racinaire pour les deux cultures par rapport aux cultures pures. L'association de ces cultures entraînait également une modification de la distribution racinaire horizontale des cultures par rapport aux cultures pures. Les essais qu'ils ont menés sur la culture en association du maïs avec la féverole n'ont pas montré d'effets marqués au niveau du maïs par rapport à la culture pure, mais ont montré une augmentation de la densité de longueur racinaire pour la féverole entre 0 et 30 cm de profondeur lorsqu'elle était cultivée en association.

La sélection variétale

L'agriculture va devoir faire face à la diminution de l'accessibilité des ressources en nutriments et à l'évolution des systèmes agricoles (Atkinson 2000). En effet, en limitant leur impact sur l'environnement, les systèmes agricoles cultivent de manière plus extensive et utilisent les nutriments de manière plus raisonnée et sous des formes plus complexes. Face à ces évolutions, la capacité des plantes à développer leur système racinaire pour utiliser les nutriments du sol est une voie de développement qui fait l'objet de nombreuses recherches.

Ce nouveau critère de sélection variétale a été étudié par plusieurs auteurs au cours des dernières années. Wasson et al. (2014) ont par exemple montré que pour des froments de printemps, il y avait des différences significatives de taux de pénétration racinaire entre les génotypes.

1.6.L'importance de mesurer les racines

Selon Price (2016), la recherche sur les racines s'intensifie car il est de plus en plus évident qu'une meilleure compréhension de leur fonctionnement et des facteurs génétiques déterminant leur développement offre des stratégies importantes pour

¹⁰ De telles associations de cultures en bandes étroites ne sont pas pratiquées dans nos régions, mais semblent être des pratiques utilisées depuis longtemps et répandues dans la région au nord-ouest de la Chine (Li et al. 2006).

faire face aux défis globaux causés par l'accroissement de la population, le changement climatique et la sécurité alimentaire.

Les caractéristiques racinaires ont une influence forte sur les processus écosystémiques tels que les cycles du carbone et de l'azote (Bardgett et al. 2014). Selon Kätterer et al. (2011), le carbone provenant des racines contribue plus fortement dans le stockage de carbone relativement stable dans les sols que le carbone provenant de résidus de cultures (à quantités de racines et de résidus égales). Une partie de ce carbone apporté aux sols par les racines provient du renouvellement des cellules racinaires et des exsudats racinaires. Les quantités de carbone injectées dans les sols par ces deux mécanismes seraient l'équivalent de 65 à 100 % de la biomasse racinaire mesurable. Gregory (2006) signale de son côté qu'au cours de la saison culturale, la quantité de carbone transférée dans le sol serait deux fois plus importante que la quantité de racines laissée dans le sol au moment de la récolte de la culture.

En ce qui concerne la relation entre l'importance d'étudier les racines et la sécurité alimentaire, il est maintenant largement reconnu que l'architecture du système racinaire est une composante essentielle de la productivité des écosystèmes agronomiques et naturels (Lobet et al. 2011). Cette architecture influence également l'efficacité avec laquelle la plante s'ancre dans le sol et assure la stabilité du sol environnant (Downie et al. 2014). Des modifications au niveau de l'architecture racinaire des plantes pourraient donc permettre des progrès de productivité au niveau des cultures, particulièrement pour celles qui sont cultivées dans des conditions limitantes (Zhu et al. 2011).

Cette importance de l'architecture racinaire dans la productivité des plantes est liée à la distribution non homogène des ressources du sol. Le déploiement spatial des racines dans le sol détermine donc dans une large mesure l'accessibilité des plantes à ces ressources (Lynch 1995). Des racines nombreuses et profondes peuvent par exemple améliorer le prélèvement d'azote dans les horizons profonds du sol et réduire ainsi les pertes de nitrates dans l'environnement sous forme de lixiviation (Rasmussen et al. 2015).

La compréhension des processus impliqués dans l'exploration du sol et dans l'acquisition des nutriments par les racines pourrait donc permettre le développement de nouvelles stratégies de production durable via une meilleure efficacité de l'utilisation de l'eau et des nutriments, le dépassement des contraintes du sol et l'amélioration de la séquestration du carbone (Downie et al. 2014). Cela permettra d'utiliser ces ressources d'une manière raisonnée, d'éliminer le gaspillage et les effets environnementaux négatifs qui peuvent être liés à ces pratiques agricoles et ainsi réduire les coûts qui y sont associés (Wang et al. 2014; Atkinson 2000).

1.7. Les paramètres mesurés sur les racines

Il existe de nombreuses manières d'estimer le développement racinaire et d'en donner une représentation. Certains paramètres sont régulièrement rencontrés dans la littérature sur le sujet, d'autres sont plus anecdotiques. Dans une synthèse bibliographique réalisée en 2000 dans "*Root methods: a handbook*", Atkinson

définit plus de 20 paramètres clés des systèmes racinaires. Certains de ces paramètres ont été repris dans cette revue bibliographique. Ils ont été choisis pour leur importance dans le développement des cultures, leur fréquence d'utilisation ou leurs liens, actuels ou potentiels, avec la méthode d'étude des systèmes racinaires basée sur l'imagerie hyperspectrale proche infrarouge. Par commodité, les paramètres présentés sont triés par ordre alphabétique. Pour faciliter le lien avec la littérature anglaise, la traduction des paramètres est à chaque fois indiquée.

1.7.1. Diamètre racinaire (*root diameter*)

Le diamètre racinaire est un paramètre important de la morphologie des racines. Il reflète l'influence et la réponse des racines d'une plante aux conditions physiques du sol (Cai et al. 2014). Selon leur épaisseur, les racines peuvent avoir un effet différent sur les sols. Durant leur croissance, les racines écartent les particules de sol, créant ainsi des zones de sol avec une plus forte densité adjacentes aux racines. Cependant, de fines racines peuvent aussi réduire la masse volumique du sol en augmentant la porosité du sol (Bardgett et al. 2014). Ce paramètre peut également donner des indications sur le potentiel de prélèvement des nutriments dans le sol, de développement de mycorhizes, de régulation du stress hydrique et de croissance des plantes (Qin et al. 2004; Atkinson 2000). Enfin, il donne une indication sur la capacité du système racinaire à stabiliser un sol: un système racinaire dense et avec des racines fines stabilise mieux un sol qu'un système racinaire avec des racines épaisses (Bardgett et al. 2014).

Ce paramètre est soit mesuré directement, soit calculé sur base de la longueur et du volume racinaire en supposant que les racines sont cylindriques (Atkinson 2000). Il est exprimé en millimètres. La mesure directe de la longueur est réalisée en répartissant les racines sur une grille et en réalisant les mesures avec un binoculaire. Les logiciels d'analyse d'image ont maintenant fortement simplifié la mesure de la longueur et du diamètre racinaire et permettent d'affiner les mesures (Han et al. 2015; Plaza-Bonilla et al. 2014; Qin et al. 2004; Himmelbauer et al. 2004; Oliveira et al. 2000).

Qin et al. (2004) ont observé que les diamètres racinaires du froment d'hiver étaient influencés par la position du prélèvement (dans le rang de la culture ou dans l'interrang) et par la profondeur à laquelle les racines étaient prélevées. Les plus grosses racines étaient observées dans le rang de la culture à une profondeur de sol comprise entre 0 et 5 cm. Amato et Pardo (1994) ont également observé qu'au sein d'une culture de froment et de féveroles, le diamètre mesuré sur des racines prélevées dans le rang de la culture était plus important que celui mesuré entre les rangs.

À l'échelle de la plante, les racines ayant un large diamètre influencent fortement la biomasse racinaire. Les racines avec un petit diamètre comptent pour leur part beaucoup plus dans la surface du système racinaire et sont le lieu d'échange entre la plante et le sol pour le prélèvement d'eau et de nutriments (Muñoz-Romero, et al.

2010). On comprend donc pourquoi le diamètre racinaire est un des paramètres les plus importants pour la modélisation¹¹ de la rhizosphère (Himmelbauer et al. 2004).

1.7.2. Longueur racinaire (*root length*)

La longueur de toutes les racines du système racinaire permet d'estimer le potentiel d'absorption de nutriments et d'eau dans le sol par la plante (Bengough et al. 2000). Ce sont donc des paramètres ciblés dans l'amélioration des cultures (Wasson et al. 2014; Watt et al. 2013). C'est également un indicateur du potentiel de l'activité des microorganismes du sol, spécialement les champignons mycorhiziens arbusculaires, et du fonctionnement microbien, par exemple dans le catabolisme du phosphore (Atkinson 2000).

Ce paramètre peut être déterminé par la méthode d'intersection de lignes qui consiste à estimer la longueur de racine en comptant le nombre d'intersections entre des racines placées aléatoirement sur une surface plane et une grille à mailles carrées (Martínez et al. 2008) ou avec des lignes droites positionnées aléatoirement sur une surface rectangulaire (Newman 1966). La méthode de Newman (1966) est celle qui a été la plus utilisée pour estimer la longueur racinaire (Plaza-Bonilla et al. 2014). Cette méthode a été revue par Tennant (1975).

Ces mesures de la longueur racinaire peuvent être réalisées au sein de monolithes, de carottes de sol, dans des rhizotrons, des minirhizotrons ou sur des profils de sols (Atkinson 2000).

Sur base de la longueur racinaire, deux autres paramètres peuvent être calculés. Il s'agit de la densité de longueur racinaire et de la longueur racinaire spécifique.

La **densité de longueur racinaire** (*root length density*) est exprimée en unité de longueur racinaire par unité de volume de sol ($m.m^{-3}$ ou $cm.cm^{-3}$). Ce paramètre indique la longueur de racine présente dans un volume de sol jusqu'à une certaine profondeur ou dans un horizon donné (Gregory 2006; Atkinson 2000). Il est donc souvent utilisé pour décrire la distribution racinaire (Qin et al. 2004) et reflète l'accessibilité des plantes à l'eau et aux nutriments du sol ainsi que les éventuelles limitations dans leur exploitation (Atkinson 2000). C'est un paramètre clé dans de nombreux modèles sur le prélèvement d'eau (Pierret et al. 2003).

De manière générale, dans des cultures semées en rang, la densité de longueur racinaire est plus élevée dans les lignes de culture que dans les interlignes (Qin et al. 2004).

La **longueur racinaire spécifique** (*specific root length*) est égale à la longueur racinaire divisée par la masse sèche des racines (Atkinson 2000). Ce paramètre est donc exprimé en unité de longueur par unité de masse. Il permet d'évaluer la part de racines fines au sein du système racinaire. Plus la valeur de ce rapport est élevée, plus le système racinaire est ramifié (Wilhelm et al. 1982).

¹¹ La modélisation est la conception de modèles qui sont des "formulations simplifiées qui imitent les phénomènes du monde réel de telle sorte qu'ils nous permettent de comprendre des situations complexes et de faire des prévisions" (Odum cité par Gate, 1995).

Notons que la précision des mesures effectuées sur la longueur racinaire dépend fortement de la capacité à récupérer les racines fines de la plante (Amato & Pardo 1994).

1.7.3. Masse racinaire (*root mass* ou *root biomass*)

Selon Hirte et al. (2017), ce paramètre est un des plus mesurés dans les études se focalisant sur la réponse des plantes aux changements environnementaux. La biomasse racinaire est en effet fortement liée à l'énergie investie par la plante dans son système racinaire, c'est-à-dire à la quantité de carbone allouée à la partie souterraine de la plante (Bengough et al. 2000).

Elle est déterminée par la pesée de racines extraites de volumes de sol et puis séchées (Hirte et al. 2017; Plaza-Bonilla et al. 2014). Plusieurs auteurs font remarquer qu'il faut tenir compte du fait que cette masse comprend également les microorganismes associés au système racinaire et éventuellement des particules de sol qui n'auraient pas été éliminées lors du lavage des racines (Atkinson 2000; Oliveira et al. 2000). Selon les auteurs, les masses racinaires sont exprimées par unité de surface ou par unité de volume. Lorsqu'elle est exprimée par unité de volume, on parle de **densité de masse racinaire** (Wang et al. 2014).

Selon Amato et Pardo (1994), les données trouvées dans la littérature pour les biomasses racinaires mesurées sur des plantes de la même espèce sont plus semblables d'une étude à l'autre que pour les données de densité de longueur racinaire. Néanmoins, la méthodologie utilisée pour récupérer les racines peut varier fortement ce qui entraîne de larges différences dans la biomasse racinaire finalement mesurée (Hirte et al. 2017). Ces biais seront abordés dans la section dédiée au lavage et au tri des racines (point 1.10 de ce chapitre).

C'est ce paramètre que nous avons prédit sur base d'images prises en imagerie hyperspectrale proche infrarouge sur des échantillons de racines. La méthodologie utilisée sera décrite au Chapitre VI dédié à la quantification racinaire.

Sur base de la masse racinaire, il est possible de calculer un **rapport de la masse racinaire sur la masse aérienne** (*root-to-shoot ratio*). Ce paramètre est calculé en divisant la biomasse sèche souterraine par la biomasse sèche aérienne (Atkinson 2000). Il donne une représentation de la distribution de la matière sèche dans la plante. En froment d'hiver, Wang et al. (2014) ont montré que la valeur de ce paramètre diminue graduellement au cours de la croissance de la plante et qu'il atteint un minimum à la maturité. Le rapport inverse (*shoot-to-root ratio*) est également rencontré dans la littérature. Plaza-Bonilla et al. (2014) ont réalisé une synthèse de valeurs obtenues pour ce ratio par différents auteurs. Ils ont montré que la valeur de ce rapport dépend du stade de la culture auquel il est mesuré, de la culture considérée et du milieu de croissance de la culture. Dans une étude qu'ils ont menée sur des céréales d'hiver, ce ratio était fortement affecté par le site d'étude. Selon eux, cette variation illustre bien l'influence des conditions de sol et du climat sur l'allocation des ressources par la plante.

1.7.4. Profondeur d'enracinement (*root depth*)

La profondeur d'enracinement est la profondeur maximale du sol à laquelle des racines peuvent être observées (Atkinson 2000). Elle permet de prédire le potentiel de prélèvement des nutriments (Hodge et al. 2009) et conditionne le prélèvement d'eau par les plantes (Gregory 2006). Ce paramètre est également intéressant dans l'étude des apports de carbone dans les sols et il est lié à la stabilité des sols et à l'ancrage des plantes en profondeur (Atkinson 2000; Bardgett et al. 2014).

Cette profondeur est affectée par la génétique de la plante et les facteurs environnementaux tels que des obstacles physiques et chimiques (Gregory 2006; Hodge et al. 2009). Cette profondeur d'enracinement peut également être influencée par les conditions agroclimatiques de l'année de culture, le site d'étude et le cultivar étudié (Rasmussen et al. 2015; Wasson et al. 2014). Le contexte dans lequel le système racinaire a proliféré doit donc être pris en compte et cette relation entre la longueur racinaire et le prélèvement de nutriments est surtout observable en cas de concurrence interspécifique, pour l'azote par exemple (Hodge et al. 2009).

À titre d'exemples de l'influence de l'espèce étudiée sur la profondeur d'enracinement, des études menées en champs sur le froment d'hiver (*Triticum aestivum* L.) ont montré une croissance racinaire pouvant aller jusqu'à 1,5 – 2 m de profondeur (Rasmussen et al. 2015). Han et al. (2015) ont mesuré des profondeurs d'enracinement de 1,8 et 1,3 m pour une culture de chicorée fourragère (*Cichorium intybus* L.) et une culture de fétuque élevée (*Festuca arundinacea* Schreb.) respectivement. Chez le pois protéagineux (*Pisum sativum* L.), la profondeur maximale d'enracinement varie de 50 à 80 cm (Munier-Jolain et al. 2005).

Etant donné la faible densité racinaire en profondeur et l'hétérogénéité de la répartition racinaire (Maeght et al. 2013), ce paramètre est difficile à estimer. Il est probablement souvent sous-estimé.

Lorsque des mesures de profondeur d'enracinement sont réalisées régulièrement, il est possible de calculer un **taux d'enracinement en profondeur** (*root depth penetration rate*). Ces mesures régulières peuvent par exemple être réalisées au sein de minirhizotrons (voir point 1.8.4 de ce chapitre). Ce paramètre est exprimé en $\text{mm } ^\circ\text{C jours}^{-1}$ et est estimé sur base de la pente de la régression liant la profondeur maximale d'enracinement à l'accumulation de températures journalières moyennes avec pour référence une température basale de $0\text{ }^\circ\text{C}$. Dans une étude réalisée sur deux ans sur différents cultivars de froment d'hiver, Rasmussen et al. (2015) ont mesuré des taux moyens d'enracinement en profondeur compris entre $0,7$ à $1\text{ mm } ^\circ\text{C jour}^{-1}$.

1.7.5. Surface racinaire (*Surface area*)

La mesure de ce paramètre est intéressante car le prélèvement d'eau par les racines serait plus lié à leur surface qu'à leur longueur (Plaza-Bonilla et al. 2014).

En posant l'hypothèse que les racines sont circulaires, la surface racinaire est calculée en multipliant l'aire mesurée en deux dimensions sur les racines par π . Elle peut également être estimée sur base de la longueur et du diamètre racinaire (Oliveira et al. 2000). Enfin, ce paramètre peut être mesuré à l'aide de logiciels

d'analyse d'images (Cai et al. 2014; Plaza-Bonilla et al. 2014). Ce paramètre est fortement influencé par la quantité de racines fines prises en compte lors des mesures et donc par la qualité de la récupération des racines lors des prélèvements de sol.

Ce paramètre peut ensuite être utilisé pour calculer une **densité de surface racinaire** (*root surface density*). Celle-ci est calculée en divisant la surface racinaire par un volume de sol (Plaza-Bonilla et al. 2014).

1.8. L'étude et la quantification des racines aux champs

Les études sur le développement racinaire sont importantes pour identifier des pratiques de gestion du sol adaptées à une production agricole durable (Huang et al. 2012). Les études de systèmes racinaires de plantes cultivées réalisées aux champs donnent une représentation réelle de la croissance racinaire dans un contexte agricole. C'est également le contexte le plus pertinent pour réaliser du phénotypage¹² sur les racines (Price 2016).

Historiquement, l'organisation des systèmes racinaires a été observée aux champs en retirant la terre autour du système racinaire ou en prélevant un volume de sol dans le champ avec la plante puis en séparant les racines du sol et en les mesurant manuellement (Zhu et al. 2011). Les premières études sur les systèmes racinaires datent du 18^e siècle. Elles étaient réalisées en extrayant les racines du sol et en mesurant manuellement leur poids et leur longueur. L'écologiste J.E. Weaver (1884-1966) est un pionnier dans l'étude des systèmes racinaires par excavation en champs et par culture de plantes en pots pour observer leurs systèmes racinaires (Downie et al. 2014). Ces méthodes d'excavation sont encore toujours utilisées, car elles permettent d'observer simultanément un grand nombre de racines du système racinaire (Zhu et al. 2011).

1.8.1. La diversité des techniques

Le nombre de techniques développées pour étudier les racines aux champs est assez vaste. Le choix d'une ou l'autre technique dépend de la question de recherche qui est posée, du temps et des ressources disponibles (Pierret et al. 2003). Il est effectivement important de se rendre compte qu'il n'y a aucune méthode de mesure racinaire qui puisse être utilisée dans toutes les situations.

Les méthodes standards pour étudier l'abondance, la distribution et la dynamique des systèmes racinaires sont réalisées soit par la mesure destructive de matériel échantillonné, soit par l'observation directe *in situ* (Rewald et al. 2012). Au sein des techniques d'observation des racines, il est possible de distinguer celles qui sont utilisées pour observer les racines dans leur environnement naturel de croissance et celles qui ont été développées pour observer les racines dans des milieux de croissance contrôlés ou artificiels. Parmi les techniques utilisées pour l'étude en conditions naturelles de croissance, très souvent des terres agricoles, on peut citer l'excavation de monolithes, le prélèvement de carottes de sol, l'observation sur des

¹² Le phénotypage consiste à observer des plantes et les caractères qu'elles expriment dans leurs conditions de croissance.

profils de sol, l'acquisition d'images dans le sol par minirhizotron, l'utilisation d'ondes radio à haute fréquence ou l'imagerie en résistivité électrique.

La méthode basée sur les monolithes consiste à extraire les racines de larges volumes de sol (1000-5000 cm³) prélevés aux champs. Le prélèvement de carottes de sol permet de manipuler des volumes de sol réduits (< 1000 cm³) (Heeraman & Juma 1993). Les méthodes basées sur des ondes radio à haute fréquence (*ground-penetration radar*) ou l'imagerie en résistivité électrique sont des techniques à faible résolution qui permettent quant à elles d'estimer des biomasses racinaires de manière non invasive (Zhu et al. 2011).

Seules les méthodes basées sur le prélèvement de carottes de sol et les minirhizotrons, ainsi que leurs évolutions, seront présentées dans ce document d'une part parce qu'elles ont été utilisées dans le cadre de cette thèse (prélèvement de carottes de sol) et d'autre part parce qu'elles pourraient avantageusement être combinées à l'utilisation de l'imagerie hyperspectrale proche infrarouge (*core break method* et minirhizotrons).

Au niveau des techniques d'étude en conditions contrôlées, on peut citer l'utilisation de techniques d'imagerie en 2 D ou en 3 D telles que des scanners, des caméras, des tomographies aux rayons X sur des plantes cultivées en rhizotrons transparents, en hydroponie ou en aéroponie. Ces techniques ne seront pas développées dans ce chapitre. Pour plus d'informations sur ces méthodes, plusieurs revues bibliographiques leur ont été consacrées dont le livre de Böhm (1979) qui reste une référence dans le domaine, le livre "*Root method: a handbook*" édité plus récemment par Smit et al. (2000) et l'article rédigé par Downie et al. (2014) sur les défis et les opportunités de la quantification racinaire par imagerie et analyse d'images.

Selon Atkinson (2000), les principaux facteurs qui influencent le choix de la méthode sont la disponibilité des équipements et des installations, la culture et/ou le sol étudié ainsi que le type d'effet du système racinaire auquel on s'intéresse. Mais le critère principal dans ce choix est vraisemblablement, selon l'auteur, si les informations recueillies doivent refléter un changement dans le temps ou si elles doivent donner une information sur la distribution spatiale à un moment précis. Lorsqu'il s'agit d'un suivi au cours du temps, une méthode d'observation directe, telle que le minirhizotron, sera préconisée. Si par contre on s'intéresse plutôt à la distribution spatiale, une méthode basée sur un prélèvement direct sera plus approprié.

1.8.2. La quantification racinaire basée sur le prélèvement de carottes de sol

Description

La méthode de quantification racinaire basée sur le prélèvement de carottes de sol (*soil coring method*) est une méthode qui est simple, rapide, qui ne nécessite pas de compétences spéciales et qui donne une bonne représentation du développement racinaire dans les sols (Frasier et al. 2016; Pierret et al. 2003; Neill 1992). Cette méthode consiste à prélever un volume de sol avec une sonde. Ce prélèvement est réalisé soit manuellement, soit mécaniquement avec, par exemple, un vérin hydraulique monté sur un tracteur (Figure 2) (Oliveira et al. 2000).

Cette technique demande un grand nombre d'échantillons pour tenir compte de l'hétérogénéité de la distribution des racines (Oliveira et al. 2000). Cela est d'autant plus vrai pour l'étude des racines profondes (au-delà de 1 m de profondeur) qui sont moins nombreuses et réparties de manière plus hétérogène (Maeght et al. 2013). Les méthodes de quantification racinaire basées sur l'excavation demandent par conséquent beaucoup de travail, détruisent une grande partie de l'architecture racinaire, dont les racines fines, et elles ne permettent pas des observations répétées sur une même plante (Zhu et al. 2011). Un autre désavantage majeur de cette technique est que de nombreuses petites racines sont rompues et éliminées avec le sol lors du lavage ce qui entraîne des erreurs de quantification (Pierret et al. 2003; Livesley et al. 1999; Amato & Pardo 1994). Ce problème sera discuté dans la section dédiée aux méthodes de lavage (point 1.10 de ce chapitre).

L'utilisation de la méthode basée sur le carottage de sol mène donc souvent à une sous-estimation de la production nette de racines. Car, en plus des pertes de racines lors du lavage des échantillons et de l'hétérogénéité du système racinaire, d'autres sous-estimations peuvent être causées par la non-coïncidence des minimums et maximums de production saisonniers avec les dates de prélèvement des racines, par les pertes de racines au cours de la saison et par la croissance et la mortalité simultanées des racines durant un intervalle de prélèvement (Neill 1992). Enfin, cette méthode est difficilement utilisable dans les sols caillouteux, contenant un grand nombre de racines d'arbres ou dans des sols secs et argileux (Oliveira et al. 2000).



Figure 2: Prélèvement de carottes de sol à l'aide d'une sonde enfoncée dans le sol par un vérin hydraulique.

Le nombre de carottes de sol, le diamètre des sondes utilisées et la position du prélèvement sont des paramètres importants à définir avant le prélèvement de racines. Ceux-ci déterminent la représentativité de l'échantillonnage par rapport à la variabilité spatiale du système racinaire rencontrée au champ (Koteen & Baldocchi

2013; Oliveira et al. 2000). Il n'existe malheureusement aucun consensus concernant ces paramètres. Ils varient donc fortement d'une étude à l'autre. Oliveira et al. (2000) rapportent néanmoins que le diamètre recommandé est de 7 cm et que les diamètres les plus utilisés sont compris entre 5 et 8 cm. Prélever des carottes plus larges et en plus grand nombre permet une meilleure caractérisation de la distribution spatiale et temporelle des racines, mais demande beaucoup plus de temps pour le prélèvement et le traitement des échantillons (Koteen & Baldocchi 2013). Un juste compromis entre la représentativité de l'échantillon et le temps nécessaire à son traitement doit donc être trouvé.

Détermination du nombre de carottes de sol à prélever

Selon les études, le nombre de carottes de sol prélevées par unité expérimentale varie entre trois et dix (Oliveira et al. 2000). Frasier et al. (2016) conseillent de répliquer les mesures au sein de chaque parcelle expérimentale de telle sorte à réduire la variabilité spatiale inhérente aux racines et de prendre en compte que cette variabilité augmente avec la profondeur de prélèvement. Notons que selon Van Noordwijk (1993), cité par Oliveira et al. (2000), l'erreur standard de 25 échantillons est seulement cinq fois plus faible que celle d'un seul échantillon. Afin de déterminer le nombre de répétitions nécessaires, Oliveira et al. (2000) conseillent d'étudier la variabilité des mesures au sein d'un certain nombre de carottes de sol avant de mener une campagne de prélèvement. C'est ce qui a été réalisé par Watt et al. (2013) afin de clarifier le nombre de carottes de sol nécessaire pour estimer de manière adéquate la profondeur d'enracinement et la densité de longueur racinaire. Ils ont donc prélevé 25 carottes de sol au sein d'une parcelle de 1,2 x 1,2 m² de froment d'hiver et ont étudié la manière dont évoluait la valeur moyenne des paramètres en fonction du nombre de prélèvements. Les prélèvements ont été réalisés au stade floraison de la culture jusqu'à 120 cm de profondeur avec une sonde de 4,1 cm de diamètre. Ils ont montré que, pour une parcelle expérimentale de cette taille, deux carottes de sol suffisaient pour avoir une estimation précise de la profondeur maximale d'enracinement, qu'il en fallait quatre pour estimer la densité de longueur racinaire et qu'il en fallait cinq pour compter le nombre de racines sur les faces horizontales de carottes de sol. Frasier et al. (2016) conseillent de leur côté de prendre quatre carottes de sol en quatre endroits différents dans une parcelle expérimentale pour tenir compte de la variabilité de la distribution racinaire.

Prise en compte de l'hétérogénéité spatiale des racines

Lorsque les échantillons de sol sont prélevés dans des cultures semées en rangs, plusieurs auteurs ont noté l'importance de prendre des échantillons aussi bien dans les lignes que dans les interlignes de la culture pour tenir compte de l'hétérogénéité spatiale des racines dans le sol et ainsi éviter de sous-estimer ou de surestimer la quantité de racines en ne prenant que des échantillons dans les interlignes ou les lignes respectivement (Frasier et al. 2016; Plaza-Bonilla et al. 2014; Li et al. 2006).

Dans une étude sur l'effet du travail du sol sur le développement racinaire de cultures de céréales d'hiver, Plaza-Bonilla et al. (2014) prenaient quatre carottes de sol par parcelle, la moitié dans la ligne de culture, l'autre moitié au centre de l'interligne. Ils faisaient attention à éviter le lieu de passage des machines lors de

leurs prélèvements. Cette remarque est valable quelle que soit la méthode d'observation racinaire utilisée. Lorsque l'interligne est important, plusieurs carottes de sol peuvent être prélevées. C'est ce qui a été fait par Li et al. (2006) pour étudier la distribution spatiale d'une culture de maïs associée soit avec du froment, soit avec de la vesce. Les carottes de sol étaient prélevées au centre, à 10 et à 20 cm du rang de maïs.

Oliveira et al. (2000) mettent également en garde sur les précautions à prendre lorsque des mesures prises à différents endroits sont extrapolées à l'échelle du champ car de sérieux biais pourraient être introduits si la valeur moyenne des échantillons est supposée représenter la valeur moyenne du champ. Frasier et al. (2016) ont proposé une série d'équations permettant de quantifier la biomasse racinaire totale (g.m^{-2}) sur base de la taille des carottes de sol, de leur espacement et de la taille de l'interligne et conseillent de calculer un pourcentage d'influence qui permet de tenir compte de masses de racines différentes dans la ligne et dans l'interligne de la culture. Ils signalent que ne tenir compte que de la biomasse racinaire présente dans l'interligne mène à une sous-estimation de la biomasse racinaire par unité de surface et qu'une simple moyenne des quantités prélevées dans la ligne et dans l'interligne entraîne une surestimation de cette biomasse racinaire.

Profondeur des prélèvements

La profondeur de prélèvement est un autre facteur important lors de l'étude racinaire basée sur le prélèvement de carottes de sol. Selon Oliveira et al. (2000), la profondeur de prélèvement devrait idéalement être égale à la profondeur maximale d'enracinement. Ils précisent néanmoins que les horizons les plus profonds sont difficiles à atteindre et que la variabilité y est plus grande. Ils préconisent donc de prélever au minimum jusqu'à 30 cm, ou jusqu'à la profondeur du labour, car les racines sont concentrées dans cette couche de sol.

Traitement des échantillons

Avec ce système de prélèvement, les racines doivent être séparées du sol. Les carottes de sol doivent donc être lavées à l'eau afin d'éliminer les particules de sol et extraire les racines. Un tri des racines, des particules de sol et des matières organiques étrangères¹³ doit ensuite être réalisé. Les différentes méthodes de lavage seront présentées au point 1.10 de ce chapitre. Les points 1.10.3 et 1.10.4 de ce chapitre seront quant à eux consacrés au tri selon qu'il est réalisé manuellement ou sur des images.

Applications

Le prélèvement de racines par carottage de sol est spécialement recommandé pour l'étude de la stratification des racines dans le sol et pour tenir compte de la variabilité spatiale de la distribution de racines fines (diamètre de moins de 2 mm) (Majdi 1996). Le carottage du sol est cité par de nombreux auteurs comme étant

¹³ Les matières organiques étrangères sont par exemple des résidus de la culture précédente, des racines de plantes adventices, des matières organiques enfouies dans le sol (fumiers, couverts d'intercultures) (Hirte et al. 2017).

une des méthodes les plus utilisées pour l'étude et la quantification racinaire aux champs (Maeght et al. 2013; Pierret et al. 2003).

Cette méthode permet d'étudier la longueur et la masse racinaire, ainsi que leur densité au sein d'un volume de sol, la morphologie des racines, leur distribution, leur renouvellement (Frasier et al. 2016; Wasson et al. 2014; Maeght et al. 2013; Oliveira et al. 2000). Elle permet souvent d'obtenir les meilleures informations quantitatives sur la biomasse du système racinaire et la densité de longueur racinaire. Dès lors, elle est fréquemment utilisée comme référence pour calibrer d'autres méthodes (Xu et al. 2016; Kusumo et al. 2010; Oliveira et al. 2000; Cheng et al. 1990). De plus, les prélèvements de racines par carottage peuvent être facilement répétés et perturbent peu les parcelles de culture (Oliveira et al. 2000). Cette méthode a donc été utilisée par de nombreux auteurs pour étudier, par exemple, les effets sur le développement racinaire du travail du sol (Guan et al. 2014; Plaza-Bonilla et al. 2014; Huang et al. 2012; Qin et al. 2004; Wilhelm et al. 1982), de l'irrigation ou de la fertilisation (Kätterer et al. 1993).

1.8.3. La méthode "core break"

Cette méthode est une évolution de la méthode de prélèvement de carotte de sol présentée ci-dessus. Elle a été développée afin de réduire le besoin en main-d'œuvre des méthodes de prélèvement de systèmes racinaire (Wasson et al. 2016) et le temps nécessaire pour acquérir des données sur les racines. Avec cette méthode, des carottes de sol sont également prélevées, mais les racines ne sont pas extraites par lavage. En effet, la méthode consiste à prélever des carottes de sol, à les casser en différents segments et à observer les racines directement sur les faces "cassées" des carottes de sol. Le fait que les carottes soient cassées en différents segments, et non coupées, permet la création d'un plan de fracture duquel dépassent des racines qui peuvent donc être comptées (Figure 3) (Wasson et al. 2016). Cette méthode rapide et simple permet de répliquer de nombreuses fois les mesures (Van Noordwijk et al. 2000).

En se basant sur le comptage de racines sur deux faces d'un volume de sol, cette méthode suppose que les racines sont orientées aléatoirement dans le sol (Drew & Saker 1980). Comme l'expliquent Wasson et al. (2016) et Wasson et al. (2014), cette méthode est basée sur la relation existant entre des vecteurs orientés aléatoirement au sein d'un volume et le nombre d'intersections de ces vecteurs avec un plan d'observation dans ce volume. Cependant, les racines ne sont clairement pas des vecteurs orientés aléatoirement et la relation entre les intersections et la longueur des racines est altérée par la ramification et le regroupement des racines, par l'influence du gravitropisme sur la direction de la croissance racinaire, par l'ordre de développement des racines et par la réponse du système racinaire à la structure du sol. La relation entre le nombre de racines comptées sur les faces horizontales des carottes de sol et la densité de longueur racinaire varie donc d'une espèce à l'autre et en fonction de l'environnement du sol. Si les racines sont préférentiellement orientées dans le plan vertical, les comptages de racines réalisés uniquement sur les faces horizontales des carottes de sol vont tendre à surestimer la longueur réelle de racines (Drew & Saker 1980). Cette relation entre le nombre de racines comptées et

leur longueur doit donc être établie pour chaque espèce, chaque site d'étude et chaque stade de la culture. En effet, lorsque Drew et Saker (1980) ont établi la relation entre les longueurs racinaires réelles et prédites par cette méthode, ils ont observé des différences statistiquement différentes entre les relations construites sur base d'échantillons acquis sous des cultures implantées par semis direct ou après labour. Ils ont également observé que les régressions évoluaient selon le stade de la culture à laquelle les observations étaient réalisées. De plus, Wasson et al. (2014) ont signalé qu'il était difficile d'estimer de manière précise et rapide la forte densité racinaire dans les 10 premiers cm du sol autour du collet de plantes de froment avec cette méthode et que la relation avec la densité de longueur racinaire établie sur des échantillons lavés était faible dans cet horizon de sol. Cette méthode était par contre tout à fait adaptée pour l'estimation de la profondeur maximale d'enracinement de différents génotypes de froment de printemps.

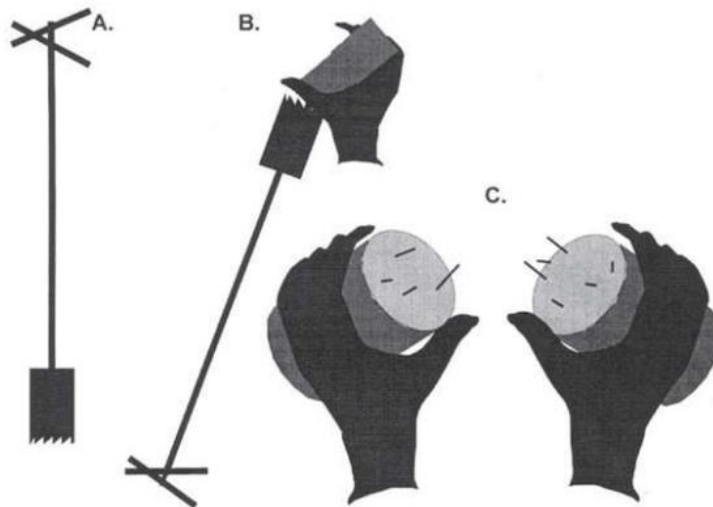


Figure 3: La méthode *core break* permet d'estimer une densité racinaire sur base d'un nombre de racines comptées à l'intersection d'une surface d'observation horizontale et plane. (A) carottage du sol, (B) récupération de la carotte de sol, (C) fraction des carottes et comptage des racines visibles sur les faces horizontales. Illustration extraite de l'article de Van Noordwijk et al. (2000)

Enfin, la qualité du comptage des racines peut être influencée par la couleur du sol; un sol de couleur claire offre moins de contraste avec les racines et rend donc leur détection plus difficile (Wasson et al. 2014). Cette méthode doit donc être calibrée empiriquement pour chaque espèce et chaque site d'étude.

Ayant étudié la relation qui existait entre le nombre de racines visibles sur les faces horizontales des carottes de sol et la longueur des racines présentes dans ces échantillons de sol, Drew et Saker (1980) ont conclu que cette méthode d'estimation

constituait une base convenable pour comparer rapidement la distribution relative des racines au sein du profil de sol en condition de plein champ. Elle ne permettait cependant pas de donner directement l'information sur la longueur ou la masse absolue des racines présentes dans le sol.

La mise au point de méthodes d'étude racinaire basées sur la combinaison de cette méthode avec la spectroscopie proche infrarouge ou en fluorescence sera présentée au point 3 de ce chapitre.

1.8.4. Le minirhizotron, une méthode de quantification racinaire basée sur l'observation directe et l'imagerie

Description

Un minirhizotron est un tube transparent enterré dans le sol à proximité de la plante dont on souhaite étudier le système racinaire (Zeng et al. 2006). Il peut être positionné horizontalement, verticalement ou avec un certain angle (Majdi 1996). Selon les études, la longueur, le diamètre et l'inclinaison des tubes sont variables (voir par exemple les études réalisées par Muñoz-Romero et al. (2010), Heeraman et Juma (1993) et Cheng et al. (1990)). Les racines qui poussent autour de la paroi extérieure du tube peuvent être observées à l'aide d'une caméra miniaturisée insérée à l'intérieur du tube (Zhu et al. 2011; Zeng et al. 2006) ou grâce à un miroir incliné et une lentille (Figure 4) (Taiz et al. 2015). Les observations sont réalisées à intervalles réguliers et les images sont ensuite analysées par ordinateur pour obtenir la longueur, le diamètre, l'état et la durée de vie des racines (Zeng et al. 2006). Il est ensuite possible de déterminer le taux d'élongation, la production, le renouvellement, la densité de longueur racinaire, la surface et la distribution racinaire à différentes profondeurs tout au long de la saison la croissance (Rewald et al. 2012; Zhu et al. 2011). Cette méthode a également été utilisée pour étudier la profondeur d'enracinement et l'effet de pathogènes sur des racines (Cheng et al. 1990) ou la décomposition de racines fines (Majdi 1996).

Cette méthode d'observation racinaire a été développée afin de s'affranchir de certaines limitations de l'étude racinaire basée sur des prélèvements de sol (Zhu et al. 2011). Elle est non destructive et permet d'étudier les systèmes racinaires *in situ* (Muñoz-Romero et al. 2010). Les minirhizotrons permettent des observations régulières avec une perturbation minimale du sol (Zeng et al. 2006). Avant le développement de logiciels d'analyse d'images, la procédure d'acquisition des données sur les images acquises en minirhizotrons était limitante car elle se faisait entièrement à la main. Elle était donc fastidieuse, coûteuse et prenait beaucoup de temps (Zeng et al. 2006; Jose et al. 2001). Plusieurs recherches ont donc été menées pour automatiser le traitement des images et l'acquisition des données (Zeng et al. 2006).

Avantages et contraintes

Koteen et Baldocchi (2013) font un état des lieux de l'utilisation des minirhizotrons. Selon eux, cette technique est préférable lorsque des échantillonnages destructifs et répétitifs doivent être évités dans de petites parcelles. Elle est également particulièrement adaptée pour l'estimation du renouvellement des racines et permet une visualisation directe des dynamiques racinaires *in situ*. Ils

précisent cependant qu'une sous-estimation des biomasses racinaires et des longueurs racinaires cumulées est souvent observée avec cette technique lorsqu'elle est comparée avec des estimations basées sur des prélèvements répétés au cours du temps, par carottage ou prélèvement de monolithes, bien que ces méthodes entraînent des pertes de racines lors du lavage. Ils rapportent de plus que cette technique est connue pour influencer l'environnement de croissance et l'architecture des racines à l'interface entre le sol et la surface du tube transparent. Enfin, ils expliquent que le traitement des images acquises au sein des minirhizotrons est laborieux et que la discrimination des racines vivantes et mortes est difficile jusqu'au moment où les racines disparaissent.

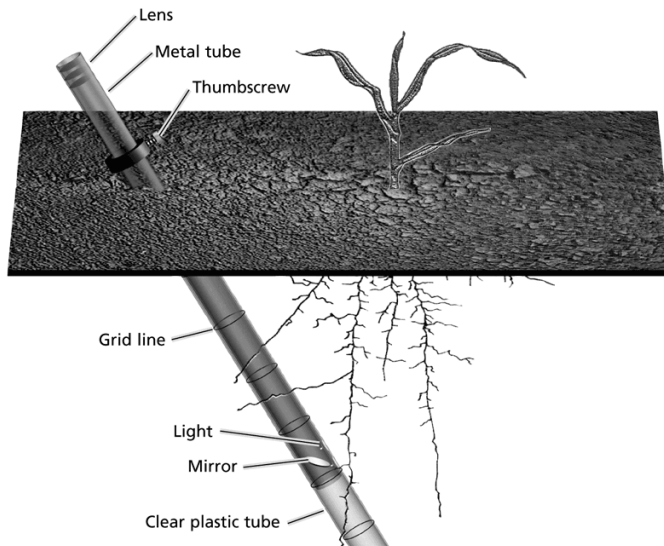


Figure 4: Un minirhizotron est un tube transparent enfoncé dans le sol. Il est installé à proximité de la plante dont les racines doivent être étudiées. Les observations des racines se développant le long de la paroi du tube sont réalisées à l'aide d'une petite caméra ou avec un miroir incliné et une lentille. Illustration tirée de la page web de Taiz et al. (2015).

Muñoz-Romero et al. (2010) listent également quelques contraintes à l'utilisation des minirhizotrons. Tout d'abord, cette méthode serait plus adaptée pour l'étude des racines fines. Ils notent également que les conditions du sol peuvent poser problème, que ce soit lors de l'installation des minirhizotrons, lorsque les sols sont trop durs ou au contraire lorsqu'ils manquent de structure, ou que ce soit par l'apparition de bavures sur la face extérieure du tube empêchant dès lors les observations ou par l'apparition de crevasses dans le sol qui entraîne une entrée de lumière dans les tubes. Enfin, ils notent que la plupart des études rapportent une sous-estimation de la longueur racinaire dans les horizons supérieurs du sol et une surestimation de la longueur racinaire en profondeur avec les minirhizotrons par rapport à la méthode basée sur le prélèvement de carottes de sol et le calcul de la densité de longueur

racinaire. La sous-estimation serait attribuée à un faible contact entre le sol et la surface du tube et une inhibition de la croissance racinaire due à une entrée de lumière dans l'horizon supérieur du sol. La surestimation observée en profondeur serait quant à elle due à une croissance préférentielle des racines le long du tube. Par exemple, dans une étude comparant les minirhizotrons avec le prélèvement de sol sous forme de carottes ou de monolithes, Heeraman et Juma (1993) ont obtenu une meilleure estimation de la densité de longueur racinaire avec les méthodes destructives qu'avec les minirhizotrons. Enfin, les études réalisées avec des minirhizotrons se concentrent majoritairement sur les jeunes racines, car celles-ci sont de couleur plus claire que le sol environnant ce qui facilite leur détection (Zeng et al. 2006).

Applications

Cette méthode a par exemple été utilisée par Muñoz-Romero et al. (2010) pour étudier l'effet du travail du sol sur la croissance racinaire (longueur et diamètre racinaire) et par Heeraman et Juma (1993) pour étudier la distribution racinaire de l'orge (*Hordeum vulgare* L.) et de la vesce (*Vicia faba* L.).

1.9. Utilisation de l'imagerie dans l'étude des systèmes racinaires

Les avancées dans les systèmes d'analyse d'images, obtenues sur des racines observées au sein de minirhizotrons ou extraites par lavage d'échantillons de sol lavé, ont mené au développement de techniques de mesure automatique des longueurs et des surfaces racinaires. Des systèmes d'analyse d'image dédiés à l'étude des racines existent, mais leur prix peut être rédhibitoire. Une alternative consiste à utiliser des logiciels *open source* et gratuits qui peuvent être personnalisés pour réaliser des mesures sur des racines (Plaza-Bonilla et al. 2014).

Comparativement aux méthodes traditionnelles de mesure de paramètres racinaires, les analyses d'images sont plus rapides, plus précises et moins sujettes aux erreurs humaines (Richner et al. 2000). La plupart des mesures sur racines sont réalisées *ex situ* en prenant des images de racines posées sur une surface plane et en les localisant ensuite (Downie et al. 2014). Avec ces techniques, il est également possible de mesurer de nouveaux paramètres morphologiques tels que la distribution de diamètres racinaires au lieu du diamètre moyen au sein d'un échantillon, des paramètres de ramification et des paramètres de formes mesurés sur des racines individuelles (Lobet et al. 2011; Richner et al. 2000).

Des informations complémentaires sur l'analyse d'image utilisée sur les racines et sur l'interprétation de ces données peuvent être trouvées dans les revues de la littérature réalisées par Downie et al. (2014) et Richner et al. (2000). Les techniques d'imagerie combinées à la spectroscopie seront quant à elles présentées au point 2 de ce chapitre.

1.10. Méthodes de lavage et de séparation de racines et de matières organiques étrangères

1.10.1. Le lavage des échantillons de sol

Lorsque les racines sont prélevées au sein d'un volume de sol et que ces racines doivent être récupérées, une étape de lavage est nécessaire.

Différentes techniques peuvent être utilisées pour éliminer le sol des échantillons et récupérer les racines. La plupart des systèmes sont basés sur un lavage à l'eau : système d'éluatriation hydropneumatique (Hirte et al. 2017; Qin et al. 2004; Oliveira et al. 2000; Heeraman & Juma 1993), lavage manuel avec un jet d'eau à faible pression sur un tamis (Plaza-Bonilla et al. 2014), mise en suspension des particules de sol dans un courant d'eau et récupération des racines sur un tamis (McKell et al. 1961), immersion répétée d'échantillons de sol dans un volume d'eau grâce à un tambour rotatif (Benjamin & Nielsen 2004; Brown & Thilenius 1976) ou agitation horizontale d'échantillons de sol dans de l'eau (Fehrenbacher & Alexander 1955) par exemple. Les échantillons peuvent également être tamisés à sec si le sol est sableux, mais cela ne s'applique qu'à des racines d'arbre de plus de 2 mm de diamètre (Oliveira et al. 2000).

Afin de faciliter la désagrégation de la terre présente avec les racines, différents agents dispersifs peuvent être utilisés et sont rapportés dans la littérature : l'hexaméthaphosphate de sodium (Plaza-Bonilla et al. 2014; Martínez et al. 2008), le métaphosphate de calcium (Wilhelm et al. 1982) et le pyrophosphate de sodium (Drew & Saker 1980) par exemple. Une synthèse d'agents dispersifs utilisés par différents auteurs dans des études sur le système racinaire a été réalisée par Oliveira et al. (2000).

Plusieurs sources d'erreurs peuvent avoir lieu durant le lavage et le tri des échantillons. Celles-ci ont été synthétisées par Oliveira et al. (2000). Il s'agit de la perte de racines fines, de la perte de matière organique, du changement de composition chimique des racines, de la présence de débris et de particules de sol avec les racines, de la présence de racines mortes, des différences de traitement des échantillons entre différents opérateurs et du temps investi dans le traitement des échantillons qui peut être variable. L'étape de lavage fragmente aussi la majorité des racines (Koteen & Baldocchi 2013) ce qui peut augmenter la perte de racines lors de cette étape. Seuls les effets de la taille du tamis sur la perte de racines fines et de l'étape de tri des éléments présents au sein des échantillons seront développés ici.

1.10.2. La taille des mailles du tamis

La taille des mailles du tamis utilisé lors du lavage a une forte influence sur la quantité de racines récupérées. Amato et Pardo (1994) ont réalisé une étude sur les pertes de longueur et de biomasse racinaire lors du lavage des échantillons en fonction de la taille des mailles du tamis. Cette étude avait été réalisée sur des racines de froment et de féveroles. Ils ont mesuré que les racines collectées sur un tamis avec des mailles de 2 mm représentaient 55 % du poids et seulement 10 % de la longueur des racines récupérées sur un tamis ayant des mailles de 0,2 mm. Lorsque la comparaison était réalisée entre des tamis ayant des mailles de 1 mm et

0,2 mm, 75 % du poids et 34 % de la longueur étaient retenus sur le tamis avec les plus grandes mailles. Les auteurs concluent que l'utilisation de tamis avec des mailles plus grandes que le diamètre des racines fines n'est pas recommandée pour des études sur des phénomènes liés à la surface racinaire dans lesquelles la quantification de la longueur racinaire est nécessaire. C'est par contre envisageable pour des comparaisons grossières de masses racinaires et de distributions spatiales. Malgré l'importance de la taille des mailles du tamis utilisé sur la mesure de la quantité et de la taille des racines, cette information est rarement reportée dans les études ce qui rend difficile la comparaison entre les études (Amato & Pardo 1994).

Dans une étude menée sur des racines de maïs, Livesley et al. (1999) ont mesuré que 80 % de la biomasse racinaire totale échantillonnée dans l'horizon de sol compris entre 0 et 15 cm étaient récoltées sur des tamis avec des mailles de 1 et 2 mm (mesures cumulées) par rapport à un tamis de 0,25 mm, mais que seulement 60 % des longueurs de racines étaient récupérées avec des tamis de cette taille. Ils conseillent dès lors d'utiliser un tamis avec des mailles de 0.25 mm pour la mesure précise de longueur de racines fines et concluent qu'un tamis avec des mailles de 0.5 mm suffit pour des mesures de biomasse racinaire.

1.10.3. Le retrait des matières organiques étrangères

Quelle que soit la technique de lavage utilisée, des matières organiques autres que des racines sont récupérées avec les racines. Ces matières organiques étrangères sont des racines mortes de la culture précédente, des racines d'adventices, ainsi que des résidus de plantes, d'amendements enfouis dans le sol ou de la faune du sol (Hirte et al. 2017; Benjamin & Nielsen 2004). Ces matières organiques étrangères sont régulièrement regroupées sous le terme "débris". Leur quantité dans les échantillons de racines dépend des pratiques agronomiques (apports de fumier, enfouissement de matières organiques, qualité du désherbage...), de la manière dont est réalisé l'échantillonnage (profondeur, position par rapport aux rangs de la culture) et des conditions de décomposition de ces débris (Hirte et al. 2017).

En utilisant les différences isotopiques pouvant apparaître entre des racines de maïs et des matières organiques étrangères présentes au sein d'échantillons de racines, Hirte et al. (2017) ont mesuré que seulement 60 % de la biomasse de leurs échantillons étaient réellement des racines de maïs de la culture en place. Si elles ne sont pas retirées des échantillons de racines, ces matières organiques constituent donc une source d'erreur majeure lors de la quantification de la biomasse racinaire et peuvent entraîner de fortes surestimations (Hirte et al. 2017). Elles peuvent également poser problème lors de la mesure de paramètres tels que la surface racinaire sur des images (Benjamin & Nielsen 2004). Il faut donc les éliminer des échantillons avant de pouvoir faire des mesures sur les racines.

Ce tri peut être réalisé sur des échantillons frais (Plaza-Bonilla et al. 2014; Heeraman & Juma 1993) ou séchés sur base de critères tels que la forme, la couleur, l'élasticité ou la structure de la surface des éléments présents dans l'échantillon (Hirte et al. 2017).

Dans tout le processus d'étude des racines, cette étape de tri est celle qui limite le plus le débit de traitement des échantillons. Elle peut prendre de quelques minutes à

quelques jours pour chaque échantillon et elle limite donc la taille et le nombre d'échantillons de sol pouvant être traités. Koteen et Baldocchi (2013) et Benjamin et Nielsen (2004) ont rapporté des temps de tri pouvant aller jusqu'à plus de 20 heures par échantillon. Leurs échantillons avaient été extraits de volumes de terre allant jusqu'à entre 2400 et 3980 cm³ respectivement.

Koteen et Baldocchi (2013) expliquent qu'afin de réduire la charge de travail par échantillon, et qu'étant donné que la quantité de racines récupérée diminue au cours du temps lors du tri d'un échantillon, certains auteurs fixent régulièrement une limite sur les racines qu'ils récupèrent. Cette limite est souvent basée sur la longueur des racines. Lorsque la longueur des racines restantes est inférieure à la limite fixée, le tri est arrêté et ils considèrent alors que les racines restantes contribuent faiblement à la biomasse totale de racines. Cependant, selon les auteurs de l'article, cette supposition est souvent incorrecte étant donné que les courts fragments de racines peuvent constituer une part importante des échantillons et elle peut donc mener à une sous-estimation de la quantité de racines présente au sein des échantillons de sol. Le choix de cette limite est également souvent réalisé de manière subjective dans la pratique et peut varier selon l'échantillon et l'opérateur. Il faut cependant remarquer que l'utilisation de cette limite dans le tri des échantillons a une influence différente selon le paramètre étudié et qu'elle impactera plus la mesure de la longueur racinaire que celle de la masse.

De manière à réduire le temps nécessaire au tri des racines et des résidus, pouvoir accroître la taille et le nombre d'échantillons sans augmenter exagérément le temps nécessaire à leur traitement, améliorer la précision des estimations et ne plus devoir fixer une limite en dessous de laquelle les racines ne sont plus comptabilisées, Koteen et Baldocchi (2013) ont imaginé une méthode de tri réalisée sur un sous-échantillon. Dans cette méthode, les racines sont encore extraites des échantillons de sol par lavage et récupération sur des tamis. Un premier tri est ensuite réalisé sur l'ensemble de l'échantillon pour les racines ayant plus d'un centimètre de long et pour les débris de grande taille. Les racines et les débris de petite taille sont ensuite randomisés et leur tri est effectué uniquement sur un sous échantillon. Avec cette méthode, les auteurs ont montré que le tri prenait plus de dix fois moins de temps qu'avec la méthode de tri classique et que la masse de racines qu'ils mesuraient différait de $3 \pm 0,6$ % par rapport à un tri méticuleux des échantillons.

Après l'étape de tri, les racines peuvent être pesées, leurs longueurs peuvent être analysées, éventuellement avec des systèmes d'analyse d'image, et des analyses chimiques peuvent être réalisées (Frasier et al. 2016).

1.10.4. L'utilisation de l'imagerie pour séparer des racines et d'autres éléments présents au sein d'échantillons

Afin de s'affranchir de l'étape de tri manuel des racines et des matières organiques étrangères ou limiter le temps qui y est consacré, de nouvelles techniques utilisent des scanners et des logiciels de traitement d'images pour étudier les racines extraites de carottes de sol.

Dowdy et al. (1998) ont mis au point un algorithme de traitement d'image qui permet d'éliminer la majeure partie du tri manuel. Cet algorithme est basé sur le

rapport entre la longueur et le diamètre de chaque objet de l'image. Lorsqu'un objet présente un rapport longueur/diamètre inférieur à une certaine valeur de référence, il n'est pas considéré comme étant une racine et il est automatiquement rejeté par l'algorithme. Dans le cas de racines de maïs et des débris qui y étaient associés, la meilleure valeur de référence de ce rapport était de 15/1. Les auteurs ont considéré que cet algorithme permettait de diminuer de 50 % le temps nécessaire à la détermination précise de la longueur racinaire totale. Wasson et al. (2016) ont fait la séparation de racines et de débris sur des images sur base de l'épaisseur de ces deux éléments, les débris étant plus épais que les racines.

Sur des images d'échantillons de racines acquises en couleur avec un scanner à plat, Plaza-Bonilla et al. (2014) ont séparé les racines vivantes des autres éléments présents dans l'échantillon (racines âgées, particules de sable, certains débris) ainsi que l'arrière-plan de l'image par seuillage sur base des intensités de rouge, vert et bleu des différents éléments. Seuls les débris les plus fins étaient éliminés avec le logiciel d'analyse d'images. Un premier tri des racines, des particules de sol et des débris récupérés sur le tamis lors du lavage avait été réalisé en immergeant ce tamis dans un bac contenant 10 cm d'eau afin de récupérer les racines à la surface de l'eau en évitant les débris les plus larges.

Richner et al. (2000) mettent cependant en garde par rapport à la discrimination des débris sur des images: bien qu'un rapport de la longueur sur la largeur puisse exclure la majorité des objets étrangers avant de réaliser des mesures morphologiques sur des racines, un nombre inconnu de ces objets sera considéré comme étant des racines et un nombre inconnu de racines sera exclu à tort. Ces techniques donnent donc uniquement de bons résultats lorsque les échantillons sont faiblement contaminés par des débris. Benjamin et Nielsen (2004) ont observé que lorsque les débris ne dépassaient pas 15 % de la surface de l'image, il n'y avait pas de différence entre les surfaces racinaires mesurées sur des images acquises sur des échantillons dont les débris organiques avaient été préalablement retirés ou non. Ces techniques sont plus performantes pour déterminer des longueurs que des masses racinaires et elles ne permettent pas de faire la distinction entre les racines vivantes et mortes (Koteen & Baldocchi 2013).

1.11. Méthodes de discrimination et de quantification de racines de différentes espèces

La plupart des échantillons de sol et des images acquises au sein de minirhizotrons contiennent des racines provenant de plantes de plusieurs espèces. L'identification des racines des différentes espèces permet de les quantifier séparément et de mener des études sur la compétition entre espèces par exemple (Rewald et al. 2012). La compréhension des interactions compétitives au sein de communautés de plantes et de la manière dont ces interactions sont affectées par le changement de ressources demande une évaluation précise de la biomasse souterraine de ces espèces dans la communauté. La proportion de biomasses racinaires de chaque espèce est liée à la capacité de la plante à accéder aux ressources en eau et en nutriments du sol (Roumet et al. 2006). Or, les racines montrent beaucoup moins de caractéristiques

distinctives que les parties aériennes des plantes (Rewald et al. 2012) et l'observation de paramètres simples tels que le diamètre des racines ou l'épaisseur du cortex ne suffit parfois pas à les discriminer (Ramirez-Garcia et al. 2015). La quantification de la biomasse racinaire de différentes espèces en mélange est donc difficile étant donné les problèmes associés à l'identification et à la séparation des racines de chaque espèce (Roumet et al. 2006) et leur identification fait appel à différentes techniques d'observation. Lorsque les espèces ont des morphologies racinaires similaires, la séparation des racines sur base d'une analyse visuelle n'est pas réalisable vu à quel point elle est fastidieuse, chronophage, dépendante de la personne qui fait le tri et sujette à des erreurs d'identification (Roumet et al. 2006).

Rewald et al. (2012) ont réalisé une synthèse des techniques d'identification des taxons au sein de mélanges de racines de différentes espèces. Ces techniques sont basées sur l'observation de critères anatomiques et morphologiques, l'utilisation de colorants, des analyses chimiques et biochimiques, des identifications basées sur l'ADN ainsi que sur des techniques basées sur la fluorescence et la spectroscopie infrarouge. Cette dernière technique fait l'objet d'une discussion au point 3 de ce chapitre. L'utilisation de radio-isotopes appliqués sur la partie aérienne des plantes et détectés par autoradiographie dans des tranches de sol ou au sein de racines lavées est aussi possible (Hauggaard-Nielsen & Jensen 2005). Dans une étude réalisée en 2006, Roumet et al. ont utilisé les concentrations en alcanes et en alcools présentes dans les racines pour faire la discrimination de différentes espèces. Li et al. (2006) ont utilisé la couleur, la texture et le schéma d'enracinement pour différencier des racines de froment, de maïs et de féveroles.

Mais l'identification des taxons des espèces ne donne qu'une partie de l'information nécessaire à la compréhension de la complexité des interactions des systèmes racinaires dans le sol. En effet, le fonctionnement des systèmes racinaires est fortement lié à la densité des racines. Il est donc nécessaire de pouvoir également quantifier ces racines. Parmi les méthodes d'identification des taxons, une partie seulement permet de faire également de la quantification (avec ou sans tri manuel) (Rewald et al. 2012).

Les méthodes permettant la quantification de la biomasse racinaire de différentes espèces étant moins développées, la compréhension de l'allocation des ressources des plantes dans leur système racinaire et de la compétition entre les racines de différentes espèces est en conséquence limitée (Roumet et al. 2006).

1.12. Méthodes de séparation de racines vivantes et mortes

La distinction entre les racines mortes et vivantes au sein d'un échantillon peut être difficile. Pour faire leur distinction, des colorants vitaux tels le rouge Congo ou le bleu de trypan peuvent être utilisés (Hirte et al. 2017; Plaza-Bonilla et al. 2014). L'identification des racines mortes et vivantes peut également être réalisée sur base de leur souplesse, leur fragilité, leur couleur et leur élasticité (Lei & Bauhus 2010; Picon-Cochard et al. 2009; Oliveira et al. 2000). Une méthode basée sur l'utilisation de la spectroscopie a également été mise au point pour discriminer des racines en

fonction de leur vitalité (Picon-Cochard et al. 2009). Ce sujet sera développé au point 3.1.1 de ce chapitre.

1.13. Les limites de l'étude des systèmes racinaires

Durant bien longtemps, les racines n'ont fait l'objet que de très peu d'études. Dans une revue de la littérature réalisée en 2006, Gregory signale que les articles sur les racines et leurs activités n'apparaissent dans les revues dédiées à l'étude du sol qu'à partir des années 1960. De même, jusqu'à très récemment, le phénotypage des plantes s'est focalisé sur des caractéristiques de la partie aérienne des plantes, telles que le rendement, la vigueur de la plante, la qualité des productions, la résistance aux maladies et aux ravageurs, et s'est très peu intéressé aux racines. Cela s'explique facilement par la difficulté technique d'observer les caractéristiques des racines qui se trouvent dans le sol, d'autant plus si on souhaite le faire de manière non destructive (Kuijken et al. 2015). Or, afin d'améliorer génétiquement la structure racinaire des plantes dans le futur, il est nécessaire d'identifier les gènes qui déterminent l'architecture des systèmes racinaires (Zhong et al. 2009). Mais, en plus du génotype de la plante, les facteurs biotiques et abiotiques de l'environnement des racines influencent aussi la morphologie des systèmes racinaires, leur croissance et finalement le rendement des cultures. Comme le signalent Downie et al. (2014), le défi pour les chercheurs intéressés par le phénotypage des systèmes racinaires n'est donc pas juste la mesure des racines et du lien existant entre le phénotype racinaire et le génotype de la plante, mais aussi la compréhension de la manière dont la croissance des racines est influencée par l'environnement. Selon ces auteurs, dans un monde idéal, le phénotypage des racines serait réalisé par l'acquisition *in situ*, à pas de temps régulier, d'images de racines se développant dans un sol non perturbé en serre ou en plein champ. Les systèmes d'analyse d'images seraient développés non seulement pour enregistrer la forme du système racinaire à un moment donné, mais aussi pour fournir des informations sur les mécanismes de croissance racinaire et la réponse génétique ou physiologique de la plante face aux conditions de croissance qu'elle rencontre au cours du temps. Cela serait lié aux informations acquises sur l'hétérogénéité biologique et physique de l'environnement du sol. Hélas, les contraintes à l'observation dans les sols sont telles que pour pouvoir faire l'acquisition d'images sur des racines vivantes, les scientifiques doivent souvent trouver un compromis entre les conditions de croissance des racines et la qualité des données acquises.

Il est en effet difficile de caractériser correctement les systèmes racinaires car le traitement des racines extraites du sol, par les méthodes traditionnelles telles que le carottage du sol et les monolithes, demande des ressources en temps, en personnel et financières considérables (Koteen & Baldocchi 2013). Les contraintes principales aux études réalisées en conditions naturelles sont la présence du sol qui empêche la visualisation du système racinaire *in situ*, et les difficultés d'extraction des racines qui peuvent former un réseau complexe et étendu dans le sol (Zhu et al. 2011; Muñoz-Romero et al. 2010). De plus, lorsque les racines ont été prélevées et traitées, les résultats obtenus peuvent être peu représentatifs étant donné l'hétérogénéité spatiale des racines dans le sol et la vitesse du renouvellement des racines fines. Une

quantification des racines et de leur développement nécessite donc des échantillonnages fréquents dans le temps et l'espace. Le prélèvement de larges volumes de sol est préférable, mais le traitement des échantillons, si laborieux, est contraignant (Price 2016; Koteen & Baldocchi 2013). Enfin, parmi le grand nombre de méthodes qui ont été proposées pour la caractérisation et la quantification des racines, la plupart sont sujettes à la subjectivité de l'opérateur, destructives et lentes (Downie et al. 2014; Plaza-Bonilla et al. 2014). Un effort constant est donc réalisé dans les études sur le système racinaire pour réduire le temps, le travail et le prix des techniques traditionnellement utilisées et principalement pour surmonter l'inconvénient de devoir réaliser des échantillonnages destructifs pour pouvoir accéder aux racines (Muñoz-Romero et al. 2010). Car en plus du fait que ces échantillonnages destructifs empêchent de faire des mesures répétées et de réaliser des observations simultanées sur des phénomènes liés, ils entraînent des conséquences importantes sur la qualité des données collectées lors de chaque prélèvement : le nombre d'échantillons est souvent inférieur à ce qui est nécessaire pour mettre en évidence des différences significatives et des études à haute résolution ne sont pas possibles (Muñoz-Romero et al. 2010).

Vu les difficultés d'étudier les caractéristiques de l'architecture du système racinaire, cette architecture racinaire est donc estimée par différentes mesures lorsque les plantes se développent dans leur environnement naturel, soit observée dans des milieux artificiels transparents ou permettant une récupération facile de l'ensemble des racines tout en les gardant reliées entre elles (Kuijken et al. 2015).

Des estimations sont également fréquemment réalisées pour les biomasses racinaires. Les données sur les biomasses racinaires de cultures étant rares, elles sont en effet souvent estimées sur base de la biomasse aérienne grâce à des coefficients obtenus dans la littérature et liant la biomasse aérienne à la biomasse racinaire (Frasier et al. 2016; Plaza-Bonilla et al. 2014). Or, dans le cadre de modèles permettant d'établir des projections sur des stocks de carbone, des données fiables sont nécessaires et une simple estimation de la biomasse racinaire sur base de la biomasse aérienne mesurée n'est pas suffisamment crédible au vu de la forte variabilité des données (Frasier et al. 2016). Le développement de nouveaux systèmes de phénotypes est donc nécessaire pour pouvoir observer et mesurer les caractéristiques complexes du système racinaire (Zhu et al. 2011).

2. Spectroscopie et imagerie hyperspectrale proche infrarouge

2.1. La spectroscopie proche infrarouge

La **spectroscopie** est définie par le dictionnaire Larousse comme étant "l'étude des spectres des rayonnements électromagnétiques émis ou absorbés par une substance". Les **spectres** étant un "ensemble de radiations monochromatiques résultant de la décomposition [...] d'un rayonnement complexe". La spectroscopie peut également être définie comme l'étude de la lumière en fonction de la longueur d'onde qui a été émise, réfléchi ou diffusée par un solide, un liquide ou un gaz (Schaepman 2009).

Les régions du spectre sont exprimées soit en termes de longueurs d'onde (λ , nm) soit en nombres d'ondes ($\bar{\nu}$, cm^{-1}), le nombre d'onde étant proportionnel à l'inverse de la longueur d'onde (Équation 1) (Türker-Kaya & Huck 2017; Rewald et al. 2012).

Équation 1: Relation entre le nombre d'onde (exprimé en centimètres⁻¹) et la longueur d'onde (exprimée en nanomètres).

$$\bar{\nu} = 10^7 / \lambda$$

L'**infrarouge** est le rayonnement compris entre la lumière visible et les microondes dans le spectre électromagnétique (Rewald et al. 2012). La région de l'infrarouge est elle-même divisée en proche infrarouge (13.500-4000 cm^{-1} , 780-2500 nm), infrarouge moyen (4000-400 cm^{-1} , 2500-25.10³ nm) et infrarouge lointain (400-10 cm^{-1} , 25.10³-1.10⁶ nm) (Türker-Kaya & Huck 2017; Abbas et al. 2012).

La région du proche infrarouge est la région de l'infrarouge la plus énergétique et la plus proche de la région visible du spectre électromagnétique (Agelet & Hurburgh 2010). La spectroscopie est donc parfois étendue à la gamme visible du spectre électromagnétique. Il s'agit alors de **spectroscopie visible-proche infrarouge**. Les longueurs d'onde concernées sont alors comprises entre 400 et 2500 nm (Figure 5) (Picon-Cochard et al. 2009).

La **spectroscopie proche infrarouge** est une méthode d'analyse standard largement utilisée en laboratoire et en industrie pour quantifier ou identifier des substances inconnues. Le principe de la spectroscopie infrarouge est d'irradier un échantillon et d'enregistrer les valeurs d'absorbance du spectre qui en résulte. Les valeurs d'absorbance enregistrées pour les différentes longueurs d'onde du spectre sont déterminées par la composition chimique de l'échantillon (Rewald et al. 2012). Elle se caractérise donc par l'acquisition d'un spectre proche infrarouge qui est représentatif de l'échantillon et qui peut être considéré comme étant la signature spectrale de cet échantillon. Cette signature spectrale est unique pour un matériau étant donné qu'elle est basée sur les propriétés physiques et chimiques de ce matériau (Ravikanth et al. 2015; Dale et al. 2013).

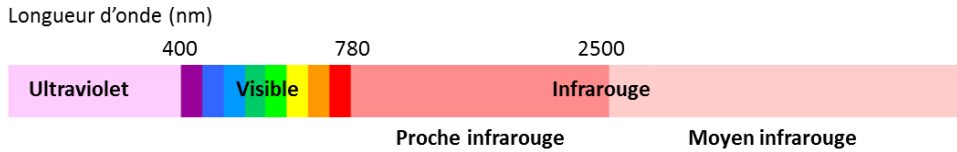


Figure 5: Les différentes zones du spectre lumineux. Illustration inspirée de IGN (2017).

L'aptitude des composants d'une matière organique, les tissus végétaux par exemple, à interagir avec la lumière infrarouge fait de la spectroscopie infrarouge un outil utile pour les analyses quantitatives et qualitatives avec de nombreux avantages (Türker-Kaya & Huck 2017). Ces avantages sont sa facilité d'utilisation, sa reproductibilité, sa capacité à analyser rapidement et simultanément différents composants, le faible coût par analyse et sa nature non invasive, non destructive et non polluante étant donné qu'elle n'utilise pas ou peu de réactifs chimiques (Dale et al. 2013; Vermeulen et al. 2013; Türker-Kaya & Huck 2017). Elle permet de plus la détermination et la quantification des éléments majeurs d'un échantillon avec seulement une préparation minimale de celui-ci. Enfin, cette technique permet des acquisitions de données en ligne et directement sur le terrain (Dale et al. 2013; Dale et al. 2012).

La spectroscopie proche infrarouge est communément utilisée depuis plus de 40 ans pour déterminer la composition nutritionnelle dans le domaine agricole et alimentaire et pour faire de la discrimination au sein de mélanges de fourrages complexes (Roumet et al. 2006). Au début des années 1960, Karl Norris fut le pionnier du développement des analyses en spectroscopie proche infrarouge. Cette méthode fut ensuite reconnue comme méthode d'analyse des protéines du blé au début des années 1980. Le nombre d'applications s'est ensuite multiplié dans les secteurs agricole, pharmaceutique, médical, artistique, alimentaire et dans les sciences des matériaux par exemple (Agelet & Hurburgh 2010).

La spectroscopie infrarouge classique ne peut cependant donner qu'un spectre moyen d'un échantillon et ne donne donc aucune information sur la distribution de la composition chimique de cet échantillon (Türker-Kaya & Huck 2017).

2.2.L'imagerie hyperspectrale proche infrarouge

L'**imagerie hyperspectrale proche infrarouge** est la combinaison des technologies de la spectroscopie proche infrarouge et de l'imagerie (Dale et al. 2013; Abbas et al. 2012; Fernández Pierna et al. 2004).

Si on se réfère à la définition du Larousse, l'**imagerie** est la technique permettant d'obtenir des images à partir de différents types de rayonnements, une **image** étant la représentation d'un objet matériel donnée par un système optique, par la photographie ou un système apparenté par exemple.

L'imagerie permet d'acquérir plusieurs informations pour chaque unité de surface (appelée pixel) d'un échantillon, typiquement l'intensité du bleu, du rouge et du vert

lorsqu'il s'agit d'imagerie en couleur. L'**imagerie spectrale**, en combinant l'imagerie et spectroscopie, permet donc d'obtenir les informations spatiales et spectrales d'un objet (Gowen et al. 2007; Fernández Pierna et al. 2006) en faisant l'acquisition d'un spectre pour chaque pixel de cet objet. Selon le nombre de longueurs d'onde acquises pour chaque spectre, il s'agira d'**imagerie multispectrale** (moins de 20 longueurs d'onde) ou **hyperspectrale** (plus de 20 longueurs d'onde) (Ravikanth et al. 2015). L'imagerie hyperspectrale peut être utilisée pour obtenir des spectres et des informations spatiales dans les régions du spectre électromagnétique correspondant à l'ultraviolet, le visible, le proche infrarouge ou le moyen infrarouge (Türker-Kaya & Huck 2017; Huang et al. 2014).

L'imagerie hyperspectrale permet l'acquisition simultanée d'un grand nombre de données spatiales et spectrales sur une surface étendue (Huang et al. 2014; Dale et al. 2013). Elle fournit des analyses quantitatives et qualitatives avec la distribution spatiale des composants (Dale et al. 2013).

L'imagerie hyperspectrale a été initialement développée pour les applications en télédétection (Huang et al. 2014). Elle a ensuite été utilisée pour des applications dans des domaines aussi divers que l'agriculture, l'agroalimentaire, l'astronomie, la pharmacie et la médecine, où, grâce à ses nombreux avantages, elle gagne rapidement du terrain (Baeten et al. 2015; Dale et al. 2013; Fernández Pierna et al. 2012; Fernández Pierna et al. 2009; Naganathan et al. 2008; Gowen et al. 2007). De nombreuses revues bibliographiques y ont été consacrées. Dans une revue de la littérature sur les applications de l'imagerie hyperspectrale moyen et proche infrarouge dans l'analyse des tissus végétaux, Türker-Kaya & Huck (2017) présentent l'utilisation de cette technique pour des applications telles que la discrimination d'échantillons de différentes plantes (différenciation de fruits et semences de différentes variétés), la détermination de la distribution spatiale de l'azote au sein des différents organes de plantes, la mesure de l'humidité, de molécules solubles, de l'acidité, de la chlorophylle, de caroténoïdes au sein de légumes et de fruits, la détection de blessures sur des fruits, l'analyse de la fermeté de fruits, l'évaluation de la texture de l'endosperme de grains de maïs ou de l'état de développement de graines mises en germination. Une revue de Mahesh et al. (2015) s'est focalisée sur les utilisations de l'imagerie hyperspectrale pour classer et surveiller la qualité de productions agricoles et horticoles. Une autre revue réalisée par Manley (2014) synthétise des applications de la spectroscopie proche infrarouge et de l'imagerie hyperspectrale sur des produits alimentaires (denrées agricoles, viandes, poissons, produits laitiers, fruits et légumes, produits céréaliers, vins, bières) et non alimentaires (bois, sols, médicaments) ainsi que des applications médicales. Dale et al. (2013) ont également synthétisé les applications de l'imagerie hyperspectrale dans les produits agricoles et agroalimentaires et se sont intéressés au contrôle de la qualité et de la sécurité de ces produits. Quelques années plus tôt, Baeten et Dardenne (2007) réalisaient une revue de la littérature sur les applications de l'imagerie proche infrarouge dans la surveillance des produits agricoles destinés à l'alimentation humaine et animale.

2.3. Les principes sous-jacents de l'imagerie hyperspectrale

La spectroscopie infrarouge est basée sur l'analyse de la lumière infrarouge interagissant avec les molécules. Elle suit donc la loi de Beer-Lamberts qui démontre la relation linéaire entre les propriétés chimiques de l'échantillon et la lumière incidente (Ravikanth et al. 2015; Abbas et al. 2012).

Dans un article de 2014 publié dans *Sensors*, Huang et ses collaborateurs ont synthétisé les principes sous-jacents de l'imagerie hyperspectrale. Ils expliquent que si on se positionne à l'échelle moléculaire, la matière émet et absorbe en permanence de l'énergie en abaissant et en augmentant son niveau d'énergie moléculaire. Les longueurs d'onde auxquelles les molécules absorbent, réfléchissent et transmettent les radiations électromagnétiques sont caractéristiques de leur structure. Les ondes électromagnétiques comprennent notamment les radiations dans l'ultraviolet, la lumière visible, le proche infrarouge, le moyen infrarouge et le moyen infrarouge lointain. À chacune de ces régions du spectre électromagnétique est lié un type particulier de transition atomique ou moléculaire en fonction de la fréquence étudiée. Il y a au sein des matériaux biologiques des liaisons moléculaires et des forces qui maintiennent les molécules entre elles. L'eau, les hydrates de carbone et les graisses sont riches en liaisons O-H ou C-H. Les composés organiques et les dérivés du pétrole sont riches en liaisons C-H ou N-H. Lorsqu'un échantillon est exposé à la lumière, les ondes électromagnétiques sont transmises au travers de celui-ci, l'énergie des ondes électromagnétiques réfléchies change à cause des vibrations et des torsions des liaisons chimiques telles que les liaisons O-H, N-H et C-H. C'est ce qui permet à la spectroscopie de fournir des empreintes caractéristiques et détaillées d'échantillons de matières sur base de l'observation des modifications au niveau de l'énergie moléculaire.

Dans la région du proche infrarouge, on rencontre deux types de vibrations : les harmoniques (*overtones*) et les combinaisons. Les harmoniques sont des vibrations des liaisons au sein des groupes fonctionnels C-H, N-H, O-H, S-H et C=O. Le second type de vibrations est dû aux combinaisons de vibrations fondamentales de ces groupes fonctionnels. Bien que le nombre d'harmoniques possible dans un groupe fonctionnel d'une molécule soit limité, un large nombre de combinaisons est observé. Il en résulte donc une répétition et une superposition des informations chimiques au sein des spectres proche infrarouge (Türker-Kaya & Huck 2017; Agelet & Hurburgh 2010). Ces répétitions sont illustrées à la Figure 6.

Toutes les molécules contenant des atomes d'hydrogène ont un spectre qui est mesurable dans le proche infrarouge. Il y a donc une large gamme de molécules qui peuvent être analysées dans cette gamme du spectre électromagnétique (Türker-Kaya & Huck 2017).

À une échelle plus large, la lumière atteignant l'échantillon est réfléchi, diffractée ou transmise. Étant donné qu'une partie de la lumière pénètre à l'intérieur de l'échantillon, les valeurs d'absorbances mesurées pour les différentes longueurs d'onde dépendent de l'état chimique et physique du matériel étudié. Les images hyperspectrales obtenues peuvent donc apporter de l'information sur la constitution chimique et les propriétés physiques de l'échantillon. Les spectres acquis sur des

échantillons différents peuvent être comparés. Les similitudes entre spectres indiquent des similitudes de composition chimique et de caractéristiques physiques (Huang et al. 2014). Par exemple, la dispersion de la lumière au niveau d'une surface étant fortement dépendante de la densité du produit et de sa structure interne, le profil de dispersion de la lumière peut donner des indications sur les propriétés du produit telles que la texture ou la dureté (Gowen et al. 2007). C'est dans ce sens que Williams et al. (2009) ont utilisé l'imagerie hyperspectrale pour classer des grains de maïs selon leur dureté.

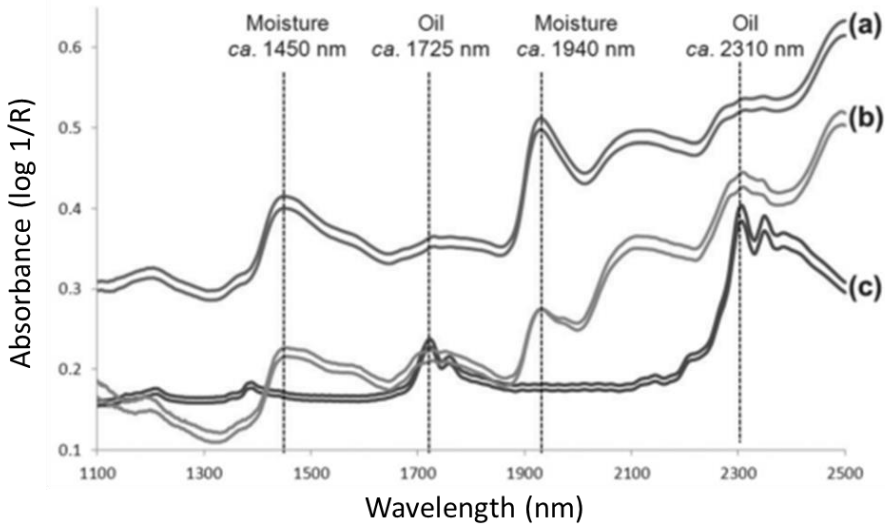


Figure 6: Les informations des groupes fonctionnels se répètent au sein du spectre proche infrarouge comme l'illustre cet exemple avec les liaisons moléculaires O-H (humidité) et C-H (huile) sur des spectres acquis sur (a) des plantes de thé, (b) du poivre noir moulu et (c) de l'huile d'olive. Illustration extraite d'un article de Manley (2014).

2.4. Acquisition des images hyperspectrales

2.4.1. La configuration du système d'acquisition d'images hyperspectrales

Les systèmes d'acquisition d'images hyperspectrales sont composés de quatre éléments principaux : une source de lumière, une unité d'imagerie, un plateau pour déposer l'échantillon et un ordinateur avec un logiciel de contrôle (Pu et al. 2015). Au sein de l'unité d'imagerie se trouvent un dispositif qui permet de décomposer la lumière en différentes longueurs d'onde, un détecteur de photons qui enregistre le signal et un objectif qui permet de faire un grossissement de l'échantillon (Türker-Kaya & Huck 2017; Pu et al. 2015; Bertrand & Baeten 2006).

Selon la configuration du système, les mesures sont réalisées en mode transmission, réflexion ou interaction (Figure 7) (Wu & Sun 2013). En

transmission, l'échantillon est illuminé par la source lumineuse et le détecteur est placé du côté opposé de l'échantillon. L'échantillon doit donc être partiellement transparent (Türker-Kaya & Huck 2017). Pour les mesures en **réflexion**, le détecteur est placé du même côté de l'échantillon que la source lumineuse dans une conformation telle que la réflexion spéculaire ne soit pas enregistrée. Le signal mesuré est donc celui qui est réfléchi par l'échantillon (Türker-Kaya & Huck 2017). En réflexion, les caractéristiques externes telles que la taille, la forme, la couleur, la texture de surface et les défauts extérieurs sont généralement détectées (Wu & Sun 2013). Pour les mesures en **interaction**, la source lumineuse et le détecteur sont également mis du même côté de l'échantillon, mais ils sont positionnés parallèlement l'un à l'autre. Cette configuration permet de réaliser des mesures plus en profondeur dans l'échantillon et est moins affectée par la réflexion spéculaire (Wu & Sun 2013). La majorité des recherches publiées en imagerie hyperspectrale auraient été réalisées en mode réflexion (Gowen et al. 2007).

Les lampes halogènes sont couramment utilisées dans les systèmes d'imageries hyperspectrales fonctionnant en mode réflexion ou en mode transmission, car elles fournissent une illumination de l'échantillon dans une large gamme de longueur d'onde dans les régions visibles et proche infrarouge (Pu et al. 2015). Ces lampes produisent une lumière dont le spectre est régulier et continu, mais elles ont les désavantages d'avoir une durée de vie courte, d'émettre beaucoup de chaleur, d'entraîner des variations au sein des spectres à cause de la température, d'avoir une production lumineuse instable à cause des fluctuations de voltage et d'être sensibles aux vibrations (Wu & Sun 2013). La forte production de chaleur de ces lampes peut aussi perturber la stabilité des échantillons (Wang & Paliwal 2007).

Des lampes LED, des lasers et des sources de lumière accordables (*tunable light sources*) peuvent également être utilisés en imagerie hyperspectrale. Ces systèmes d'illumination, ainsi que les appareils permettant la décomposition de la lumière et la transformation des photons en électrons et *in fine* en signal ne sont pas décrits dans ce chapitre. Des informations complémentaires à ces sujets peuvent être obtenues dans l'article de Wu et Sun (2013). Notons seulement que ces sources lumineuses ne produisent pas ou peu de chaleur (Wu & Sun 2013; Wang & Paliwal 2007).

2.4.2. Les méthodes d'acquisition des images hyperspectrales proche infrarouge

Une image hyperspectrale consiste en une série d'images d'un même objet acquises à différentes longueurs d'onde du spectre proche infrarouge. Il en résulte une structure en trois dimensions qui est appelée "*hypercube*" (Huang et al. 2014). Dans ce cube, l'échantillon est compartimenté en petites surfaces, les pixels, pour lesquelles un spectre complet est systématiquement enregistré (Türker-Kaya & Huck 2017).

L'hypercube est composé de vecteurs de pixels, représentant la composition et l'apparence de l'échantillon analysé (Huang et al. 2014). Cet hypercube est représenté sous forme d'une matrice à trois dimensions contenant les données spatiales en deux dimensions (x lignes * y colonnes) acquises à chaque longueur

d'onde du spectre proche infrarouge (λ longueurs d'onde sur l'axe z) (Türker-Kaya & Huck 2017; Ravikanth et al. 2015). Une quantité énorme d'information est donc acquise pour chaque image (Türker-Kaya & Huck 2017).

Cet hypercube permet de représenter la répartition de la composition chimique d'un échantillon sur une image étant donné que les régions de cette image ayant des compositions chimiques similaires ont des propriétés spectrales similaires (Mahesh et al. 2015; He et al. 2014; ElMasry et al. 2012; Kamruzzaman et al. 2012; Menesatti et al. 2009; Elmasry et al. 2007).

L'hypercube d'une image hyperspectrale peut généralement être construit de quatre manières : par acquisition de toutes les longueurs d'onde point par point ou par lignes de la largeur de l'image, par l'acquisition de l'entièreté de la surface de l'image à une longueur d'onde à la fois ou par l'acquisition de toute la surface de l'image à de toutes les longueurs d'onde en une seule fois (Figure 7) (Huang et al. 2014; Wu & Sun 2013).

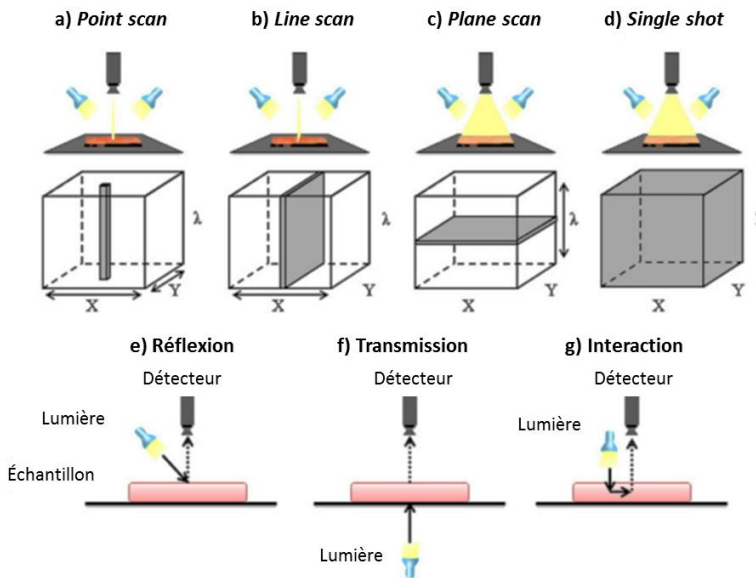


Figure 7: Les différents systèmes d'acquisition d'images hyperspectrales. L'hypercube de données des images hyperspectrales comporte deux dimensions spatiales (x et y) et une dimension spectrale (λ). Il peut être construit par acquisition (a) point par point, (b) par ligne,

(c) par plan ou (d) en une seule fois. Les directions des acquisitions de données sont indiquées par les flèches et les données acquises à chaque prise de mesure sont indiquées par la surface grise. Les spectres qui composent l'hypercube sont acquis en mode (e) réflexion, (f) transmission ou (g) interaction. Illustration extraite de l'article de Wu & Sun (2013) et modifiée.

Dans le système d'acquisition point par point (système d'acquisition *whiskbroom* ou *point scan*) un point unique est mesuré à l'échelle d'un pixel pour acquérir le spectre de ce point. Les autres points de l'échantillon sont ensuite acquis en déplaçant soit le détecteur, soit l'échantillon (Wu & Sun 2013).

Lorsque l'hypercube est construit par lignes de pixels (système d'acquisition *pushbroom* ou *line-scan*), seule la largeur de l'image est définie. Sa longueur est variable et est définie selon les besoins. Ce système d'acquisition est le plus adapté pour de la télédétection, des systèmes avec un tapis roulant ou pour des mesures en temps réel lors d'un processus (Fernández Pierna et al. 2012; Gowen et al. 2007).

Les méthodes *point scan* et *line scan* sont considérées comme étant des méthodes d'acquisition spatiales, car l'hypercube est construit en déplaçant les mesures dans la dimension spatiale de l'image. Avec ces systèmes d'acquisition, l'entièreté du spectre est acquise directement pour chaque pixel de l'image.

La méthode consistant à faire l'acquisition d'une surface (*plane scan*), est une méthode d'acquisition spectrale. En effet, dans les systèmes permettant l'acquisition d'une surface entière, le champ de vue est fixe et l'acquisition se fait à une longueur d'onde à la fois (Wu & Sun 2013). Dans le système d'acquisition en plan, la caméra est positionnée parallèlement à la surface de l'échantillon et une image d'une taille définie par les dimensions du détecteur est acquise. Les images hyperspectrales sont obtenues en modulant les radiations réfléchies atteignant la caméra via l'utilisation d'un filtre positionné avant la caméra (Fernández Pierna et al. 2012). L'hypercube est donc créé dans ce cas-ci en superposant des images acquises à différentes longueurs d'onde.

Un dernier système (*single shot* ou *snapshot*), permettant l'acquisition de l'entièreté de l'hypercube en une seule fois, est mentionné par Wu et Sun (2013). Hagen & Kudenov (2013) y ont consacré une revue de la littérature.

2.5.Chimométrie et analyse des images hyperspectrales proche infrarouge

La **chimométrie** est la science qui permet d'extraire des informations pertinentes de données acquises par analyses chimiques en utilisant des procédures mathématiques et statistiques ainsi que des outils informatiques (Dale et al. 2012; Fernández Pierna et al. 2012; Agelet & Hurburgh 2010; Wold 1995). Cette science permet également de concevoir des procédures expérimentales adaptées (Fernández Pierna et al. 2012).

La chimométrie rend possible la gestion de spectres proche infrarouge qui sont composés de bandes larges et se recouvrant, qui sont très sensibles aux caractéristiques physiques des échantillons et qui contiennent des informations redondantes (Agelet & Hurburgh 2010). Ces caractéristiques des spectres acquis dans le proche infrarouge ont été illustrées à la Figure 6, au point 2.3 de ce chapitre.

Pour extraire de l'information chimique d'un spectre électromagnétique, différentes étapes sont nécessaires (Türker-Kaya & Huck 2017). La première étape consiste à réaliser un prétraitement des spectres. Cela permet d'améliorer le rapport

signal-sur-bruit des spectres et donc d'améliorer l'information contenue dans les spectres en réduisant l'impact des informations non pertinentes dues, par exemple, à la dispersion de la lumière sur l'échantillon. Les spectres ainsi traités peuvent ensuite être utilisés pour faire de la classification ou de la quantification.

2.5.1. Correction des images hyperspectrales

Le détecteur du système d'acquisition peut avoir des défauts de fabrication et se dégrader au cours de son utilisation. Un certain nombre de pixels défectueux (*bad pixels*), dont des pixels morts (*dead pixels*), apparaissent alors sur les images (Amigo 2010; Naganathan et al. 2008; Burger & Geladi 2006). Le système d'acquisition doit donc également être calibré de telle manière à détecter et corriger ces pixels défectueux. La détection des pixels défectueux peut se faire en comparant l'absorbance de chaque pixel avec l'absorbance médiane des pixels voisins (Caporaso et al. 2018)

Après l'acquisition des images hyperspectrales et avant d'en extraire l'information utile, l'utilisation de seuils et de masques permet de retirer de l'hypercube de données les informations redondantes du support d'acquisition des images (Gowen et al. 2007) et les informations provenant de pixels défectueux et non corrigée par le système d'acquisition (Manley 2014).

2.5.2. Prétraitement des spectres

Plusieurs phénomènes indésirables lors de l'acquisition des spectres peuvent poser problème lors de leur analyse. Il s'agit de phénomènes tels que la diffraction de la lumière, la taille des particules de l'échantillon, des variations morphologiques de l'échantillon et des défauts au niveau du détecteur du système d'acquisition (Ravikanth et al. 2015; Manley 2014). Ces phénomènes peuvent conduire à des effets multiplicatifs ou additifs sur les valeurs d'absorbance des spectres (Wise et al. 2006). Les spectres acquis ne respectent alors plus la loi de Beer-Lambert et la relation linéaire existant entre la nature chimique de l'échantillon et la lumière qu'il réfléchit (Ravikanth et al. 2015). Les techniques de traitement des spectres permettent d'éliminer ces variations qui ne sont pas dues à la nature chimique des échantillons et fournissent des données qui sont plus appropriées pour le développement de modèles. Ces traitements n'affectent en rien l'information chimique contenue dans les spectres.

Il existe plusieurs méthodes de traitement des spectres : mise à l'échelle, lissage, normalisation, dérivée... Pour choisir le prétraitement qui sera utilisé, il n'existe aucune règle générale. Le traitement optimal pour des spectres donnés dépend en effet du mode d'acquisition (réflexion ou transmission), des caractéristiques de l'échantillon, de la configuration de l'instrument, des applications et de l'objectif final (calibration ou discrimination). Le choix du prétraitement devra donc se baser sur un processus d'essais-erreurs guidé par l'expérience (Agelet & Hurburgh 2010) et ce prétraitement ne devra être utilisé que s'il aide réellement à améliorer les performances des modèles (Wu & Sun 2013). Seules les méthodes de prétraitement qui ont été utilisées dans le cadre de cette thèse sont décrites ici.

La **mise à l'échelle automatique** (*autoscale*) est une méthode de prétraitement très utilisée. Elle est réalisée en calculant la moyenne centrée des données avant de diviser chaque variable par l'écart-type obtenu pour cette variable. Cela revient dans le cadre de spectre à centrer les spectres sur une absorbance moyenne égale à zéro et à diviser ensuite la valeur obtenue à chaque longueur d'onde par l'écart-type observé pour cette longueur d'onde de manière à ce qu'elle ait finalement un écart-type égal à un (Wise et al. 2006). Ce traitement permet de donner la même importance à chaque longueur d'onde lors de la construction des modèles (Agelet & Hurburgh 2010).

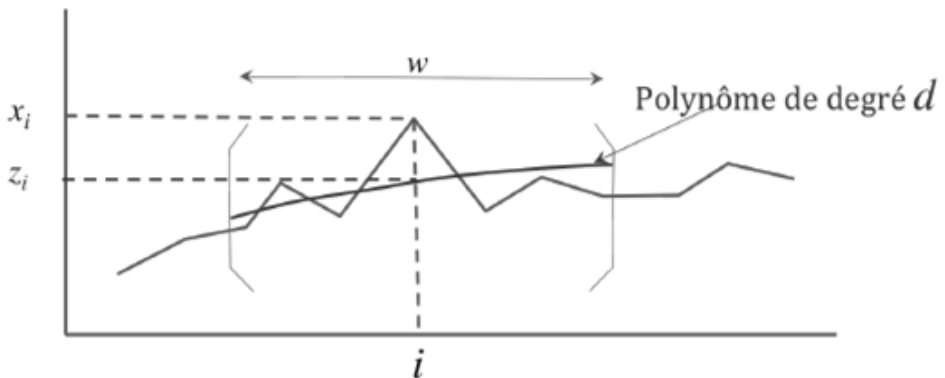


Figure 8: Lissage par l'algorithme de Savitzky-Golay. En chaque point i du spectre, la valeur brute x_i est remplacée par la valeur z_i d'un polynôme de degré d ajusté sur une fenêtre de largeur w autour du point i . Illustration extraite de (Roger & Ecarnot 2016)

Le **lissage** (*smoothing*) est un filtre passe-bas utilisé pour retirer le bruit¹⁴ de haute fréquence des spectres. Ce traitement suppose que les valeurs d'absorbances mesurées à des longueurs d'onde proches sont liées entre elles et contiennent des informations similaires qui peuvent être moyennées de sorte à réduire le bruit sans perdre significativement d'information au niveau du signal. L'algorithme de Savitzky-Golay est souvent utilisé pour lisser les spectres (Wise et al. 2006). Cet algorithme permet de réaliser un lissage polynomial appliqué sur une fenêtre mobile. En chaque point du spectre, la valeur brute est remplacée par la valeur calculée pour ce point par un polynôme ajusté sur une fenêtre autour de ce point (Figure 8). L'ordre du polynôme et la taille de la fenêtre (largeur du filtre) doivent être définis avant l'application de l'algorithme. Plus la largeur de la fenêtre est grande et plus l'ordre du polynôme est faible, plus le spectre sera lissé (Roger & Ecarnot 2016; Wise et al. 2006).

¹⁴ Le **bruit** est la variabilité qui accompagne les mesures et qui est due à l'imprécision inhérente aux mesures. Les instruments de mesure ne sont en effet pas parfaitement stables ce qui entraîne une source de variabilité au sein des mesures qui n'est pas due à ce qui est mesuré. Cette variabilité peut être due à une variation de la pression atmosphérique, de l'humidité relative, de la température ou de la tension électrique par exemple (Wold 1995).

Le traitement **SNV** (*Standard Normal Variate*) permet de retirer les interférences multiplicatives et l'effet de la diffraction générée par les particules de l'échantillon (Ravikanth et al. 2015). C'est une méthode de normalisation pondérée. La normalisation n'est donc pas réalisée de la même manière pour toutes les variables. L'algorithme SNV calcule l'écart-type de toutes les variables regroupées pour un échantillon donné. L'échantillon entier est ensuite normalisé par cette valeur, ce qui donne à l'échantillon un écart-type unitaire. Cet algorithme soustrait également la valeur moyenne individuelle de chaque spectre (Wise et al. 2006).

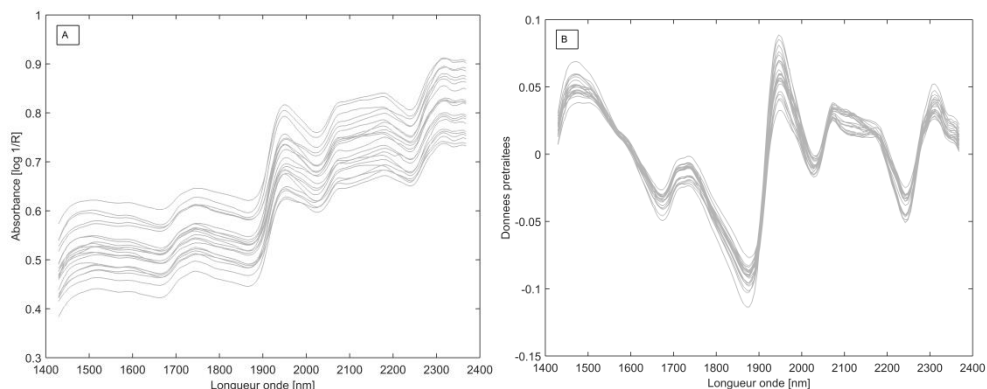


Figure 9: Spectres proche infrarouges, acquis sur des nodosités sèches de pois protéagineux, avant (A) et après (B) traitement avec la méthode Detrend.

La méthode **Detrend** est utilisée pour retirer du spectre un effet de décalage constant, linéaire ou courbe (Wise et al. 2006). Elle permet de réduire la ligne de base des spectres affectés par ce type de décalage. A titre d'exemple, la Figure 9 présente des spectres moyens acquis sur des nodosités de pois protéagineux. Une tendance croissante affecte clairement ces spectres lorsqu'aucun prétraitement des spectres n'est réalisé. Ce phénomène n'est pas dû à la composition chimique des échantillons mais à la diffusion de la lumière dont l'intensité augmente avec la longueur d'onde (dans cette gamme de longueur d'onde) (Roger & Ecarnot 2016). La méthode Detrend ajuste une fonction polynomiale sur l'ensemble du spectre et soustrait ensuite cette fonction aux valeurs du spectre. L'ordre de la fonction polynomiale doit être défini avant l'application de l'algorithme. C'est un traitement simple, mais qui ne fonctionne bien que lorsque la plus large part du signal est due à un effet de fond et n'est pas fortement influencé par des processus chimiques ou physiques d'intérêt (Wise et al. 2006).

2.5.3. Modèles non supervisés

Les **modèles non supervisés** analysent les caractéristiques des données sans connaissance préalable de classes qui permettraient de regrouper ces données (Wu & Sun 2013). Ils sont utilisés pour réaliser des analyses exploratoires des données et déterminer des relations possibles entre les échantillons (Manley 2014).

L'**analyse en composante principale (PCA)**, par exemple, crée des composantes principales qui expliquent le maximum de variabilité des données sans tenir compte de la nature des échantillons sur lesquels les données ont été acquises (Naganathan et al. 2008). Ces composantes principales sont des combinaisons de variables qui permettent de décrire les tendances majeures observables au sein des données (Wise et al. 2006). Elles sont non corrélées et organisée de telle sorte que les premières composantes englobent la majorité de la variabilité présente au sein de l'ensemble des données d'origine (Kemsley 1996). Elles permettent de compresser des données au sein desquelles il y a un large nombre de variables interconnectées et d'en extraire de l'information (Wise et al. 2006; Kemsley 1996).

2.5.4. Modèles supervisés

Les modèles supervisés sont construits sur base de données de référence appartenant à des classes prédéfinies (Wu & Sun 2013). Ces modèles supervisés sont utilisés pour faire de la discrimination et/ou de la quantification.

Modèles de quantification

Les modèles chimiométriques de quantification permettent de prédire la teneur d'un composé sur base du profil d'un spectre. Ces modèles permettent la résolution de problèmes de régression dont le but est de prédire la valeur d'une ou plusieurs variables dépendantes (Y) au moyen d'un certain nombre de variables indépendantes ou de prédictions (X) (Wold et al. 2001). Les variables dépendantes (Y), aussi appelées variables de réponse ou prédites, pourraient par exemple être la composition d'un échantillon, les variables de prédictions (X) étant le profil des spectres acquis sur cet échantillon.

L'**algorithme PLS** (*Partial Least Square* ou *Projection to Latent Structures* selon Wold et al. (2001)) est une méthode linéaire multivariée qui est souvent utilisée pour sa simplicité, sa rapidité et ses bonnes performances (Fernández Pierna et al. 2011). Selon Lebot (2012), la régression au sens des moindres carrés partiels est la technique de régression la plus utilisée pour développer des modèles prédictifs en proche infrarouge.

Cet algorithme permet de réduire de larges ensembles de données brutes en un petit nombre de facteurs orthogonaux (non corrélés) de telle manière à minimiser la somme des carrés des erreurs des valeurs qui doivent être prédites (Roumet et al. 2006; Barker & Rayens 2003). Lorsqu'il est utilisé pour des cas de régression, il tente de trouver des facteurs qui saisissent la variance des variables de prédiction (X) et qui permettent de corréler ces variables avec les variables prédites (Y) (Wise et al. 2006). Il est alors utilisé comme technique de réduction de la dimensionnalité des données en définissant de nouvelles composantes, les variables latentes, qui expliquent aussi bien la variation des variables de prédiction et prédites (Agelet & Hurburgh 2010; Naganathan et al. 2008).

Au sein de cet algorithme, chacun des paramètres du modèle est estimé itérativement comme étant la pente d'une régression bivariée simple (moindre carré) entre la matrice ligne ou une colonne de la variable prédite Y et le vecteur de paramètres de la variable prédictive X. La régression est définie comme étant

"partielle" car le vecteur X est considéré comme étant fixe dans l'estimation (Wold et al. 2001). Une interprétation géométrique du modèle de régression PLS est illustrée à la Figure 10.

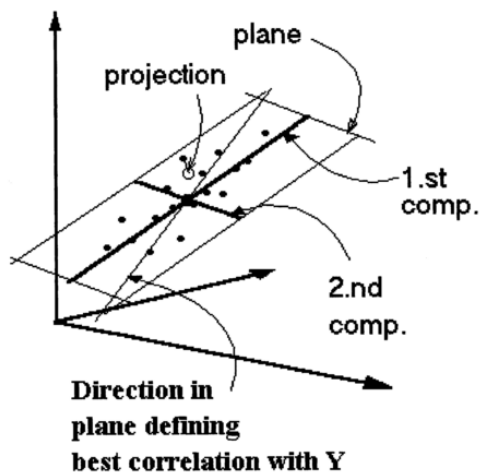


Figure 10 : Représentation géométrique de la régression PLS utilisée avec des spectres. La matrice des variables de prédiction (X) peut être représentée sous forme de points (un par observation) dans un espace multidimensionnel où chaque longueur d'onde du spectre définit un axe de coordonnées. Le modèle de régression PLS va permettre de trouver un hyperplan, défini par un nombre réduit de composantes (ou variables latentes), sur lequel la projection des variables de prédiction (X) sera fortement corrélée avec les variables prédites (Y). Illustration extraite de l'article de Wold et al. (2001) et modifiée.

La sélection du bon nombre de variables latentes est une des étapes critiques dans la construction d'un modèle précis et qui permettra d'obtenir de bonnes prédictions (Kamruzzaman et al. 2012). La sélection d'un nombre trop important ou trop peu important de variables latentes peut limiter les capacités du modèle. Nous verrons au point 2.5.5 comment la validation croisée permet de sélectionner le bon nombre de variables latentes.

Modèles de discrimination

Il existe plusieurs modèles chimiométriques permettant de discriminer des éléments sur base du profil de leur spectre.

Parmi ces modèles de discrimination, deux ont été utilisés dans le cadre de nos recherches. Il s'agit des analyses discriminantes basées sur les algorithmes PLS-DA et SVM.

L'**algorithme PLS** présenté comme méthode de quantification peut également être utilisé pour réaliser de la discrimination. Il est alors nommé PLS-DA (*Partial Least Square Discriminant Analysis*). Dans ce cas, il s'agit d'une régression PLS classique dans laquelle la variable prédite est une catégorie. Les valeurs de la variable prédite

(Y) sont donc remplacées par des variables nominales auxquelles sont rattachées les variables de prédiction (X). Cette technique d'analyse ne permet pas d'obtenir une réponse autre que celle définie par les catégories (ou classes) (Menesatti et al. 2009).

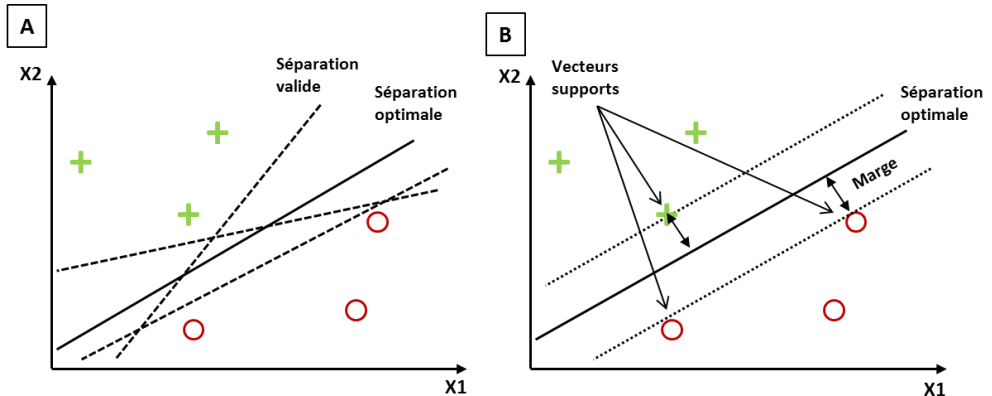


Figure 11 : Représentation géométrique de l'algorithme *Support Vector Machine* (SVM). À gauche (A), quatre limites possibles pour séparer deux groupes de données dans un espace à deux dimensions (X_1 , X_2) sont représentées. Parmi les limites valides (lignes en tirets), seule une limite est optimale (ligne pleine). À droite (B), l'algorithme SVM va permettre de déterminer cette limite optimale (ligne pleine) qui est en tout point à la distance maximale des deux groupes de données. Cette distance maximale est déterminée grâce aux vecteurs supports et à la marge (comprise entre la ligne pleine et les lignes en pointillés). Illustration adaptée de la présentation de Bousquet (2001).

L'algorithme SVM (*Support Vector Machine*) est un outil de la chimiométrie qui permet de modéliser des relations non linéaires entre des données en projetant, de manière linéaire, les données de l'espace original dans un nouvel espace de plus grande dimension où la séparation entre les données devient linéaire grâce à l'utilisation d'une fonction noyau. Cette fonction peut être de plusieurs types: linéaire, Gaussienne à base radiale (RBF), polynomiale, sigmoïdale... (Vermeulen et al. 2012; Fernández Pierna et al. 2004).

L'objectif de l'algorithme SVM est de dégager, des données qui doivent être discriminées, une fonction qui décrit la limite de décision, ou l'hyperplan, qui sépare de manière optimale deux groupes de données (Fernández Pierna et al. 2006). L'optimisation d'un modèle construit avec l'algorithme SVM consiste à déterminer l'hyperplan qui se trouve en tout point à la distance maximale (marge) des deux groupes de données qu'il sépare et qui réduit les erreurs de prédiction. Cette optimisation se base sur les données les plus proches des deux groupes (vecteurs supports) (Ravikanth et al. 2015; Fernández Pierna et al. 2012; Zhang 2010; Zhong et al. 2009; Fernández Pierna et al. 2006; Fernández Pierna et al. 2004; Cornuéjols 2002).

La Figure 11 illustre, dans un espace à deux dimensions, quatre limites possibles entre deux groupes de données. La limite qui sera finalement choisie est celle qui se trouve à la plus grande distance des deux groupes en se basant sur les données les plus proches des deux groupes. La détermination de cette limite de décision permet par la suite de déterminer si une nouvelle donnée, de nature inconnue, appartient à l'un ou l'autre groupe de données.

En recherche agronomique, l'algorithme SVM a par exemple été utilisé pour détecter ou classer des composés ou des substances indésirables au sein d'aliments destinés à l'alimentation humaine ou animale (Ravikanth et al. 2015; Vermeulen et al. 2012; Fernández Pierna et al. 2006; Fernández Pierna et al. 2004), pour étudier la complexité racinaire du maïs (Zhong et al. 2009) et pour estimer la densité racinaire de maïs au sein de carottes de sol (Xu et al. 2017).

2.5.5. Évaluation des modèles

L'étape d'évaluation des modèles calibrés, appelée validation, permet d'estimer les performances des modèles lorsqu'ils sont appliqués sur de nouvelles données et de calculer une erreur moyenne de prédiction (Wu & Sun 2013; Roumet et al. 2006; Faber & Kowalski 1997). Elle permet d'évaluer dans quelle mesure le modèle est optimiste. C'est-à-dire dans quelle mesure un modèle qui prédit bien les caractéristiques de l'échantillon à partir duquel il est établi est capable de prédire ces caractéristiques sur de nouveaux échantillons (Dramaix-Wilmet 2008; Steyerberg et al. 2001). Cette évaluation peut être réalisée de deux manières : soit sur les données du set de calibration (validation interne), soit sur de nouvelles données (validation externe). Dans le cas de la validation interne, il existe plusieurs approches.

Une première approche de la **validation interne** consiste à diviser aléatoirement le jeu de données initial en deux sous-ensembles (*data splitting*). Un de ces sous-ensembles servira à construire le modèle et l'autre sera utilisé pour tester ses performances. Avec cette approche, les performances du modèle sont déterminées sur des données similaires mais indépendantes (Steyerberg et al. 2001).

L'algorithme Duplex est un cas particulier de la division des données. Cet algorithme mis au point par R.W. Kennard et décrit par Ronald Snee (1977) est recommandé pour diviser des données en un jeu de données de calibration et un autre de validation lorsqu'il n'y a pas de variable évidente à utiliser comme base pour fractionner les données (le temps par exemple). Il fonctionne selon le principe suivant : l'algorithme commence par sélectionner les deux données les plus éloignées en termes de distance euclidienne¹⁵ et les rassemble au sein d'un même jeu de données. Ces données seront utilisées pour calibrer les modèles. Les deux données suivantes qui sont les plus éloignées sont mises dans un second jeu de données qui sera utilisé pour tester les modèles. La procédure se poursuit de manière à inclure successivement dans chaque jeu de données la donnée qui en est la plus éloignée. Le résultat est une séparation des données en deux groupes qui contiennent le maximum de diversité, couvrent approximativement la même variabilité et ont des

¹⁵ La distance euclidienne est la distance géométrique entre deux points dans un espace multidimensionnel.

propriétés statistiques similaires. Si le jeu de donnée initial contient des valeurs extrêmes, celles-ci seront reprises aussi bien dans le jeu de donnée de calibration que dans celui servant à la validation des modèles. La Figure 12, extraite de l'article de Snee (1977), illustre la sélection de ces deux jeux de donnée au sein d'un espace à deux dimensions.

Dans une seconde approche de la validation interne, appelée **validation croisée** (*cross validation*), certaines données du set de calibration sont gardées à part. Les modèles sont alors construits en ne tenant pas compte de ces données et ils sont ensuite testés sur ces données. La validation croisée a deux objectifs: (i) obtenir rapidement une estimation des performances des modèles calibrés lorsqu'ils sont appliqués sur de nouvelles données et (ii) estimer le nombre optimal de variables à utiliser dans les modèles pour obtenir les meilleures prédictions et éviter le sur-apprentissage¹⁶. Il s'agit par exemple de choisir le nombre de variables latentes pour un modèle PLS. Le choix du nombre de variables utilisées pour calibrer le modèle est basé sur la valeur de l'erreur standard obtenue lors de la validation croisée (RMSECV): le modèle choisi est celui qui obtient la plus faible valeur d'erreur standard de validation croisée avec le plus faible nombre de variables (Eigenvector 2016; Roumet et al. 2006).

Plusieurs procédures de validation croisée peuvent être utilisées. Celles-ci diffèrent par leur manière de sélectionner les données: de manière aléatoire ou de manière systématique, seules ou par groupes. La validation croisée est généralement réalisée en multipliant les tests et en utilisant successivement les données disponibles pour calibrer et valider les modèles. Les résultats des tests sont ensuite moyennés pour donner une estimation des performances du modèle. A titre d'exemple, si n est le nombre de division du jeu de données initial, les jeux de données utilisés pour tester les modèles avec la méthode *venetian blind* sont construit en sélectionnant de manière systématique chaque $n^{\text{ème}}$ objet du jeu de données, en commençant à l'objet 1 et en allant jusque n . Lorsque le nombre de données n 'est pas trop élevé (moins de 20), les tests sont répétés sur chaque donnée séparément jusqu'à ce que l'entièreté des données disponibles pour la calibration ait été utilisée pour la validation. Il s'agit de la méthode *leave-one-out* qui est la méthode de validation croisée la plus extrême (Eigenvector 2016; Steyerberg et al. 2001; Faber & Kowalski 1997).

La validation croisée, particulièrement lorsqu'elle est réalisée sur des données avec peu de variabilité, peut mener à des résultats excessivement optimistes par rapport à ceux qui sont obtenus lorsque les modèles sont appliqués sur de nouveaux échantillons. La meilleure manière de démontrer la puissance d'un modèle est donc de valider ce modèle sur des données indépendantes de celles utilisées pour la calibration (Xu et al. 2016).

¹⁶ Le sur-apprentissage (*overfitting*) apparaît lorsqu'un nombre excessif de variables est inclus dans le modèle lors de la calibration. Dans ce cas, une partie du bruit est également modélisée et la calibration devient trop liée aux données utilisées. Ce phénomène peut mener à une baisse de précision des modèles lorsqu'ils sont utilisés pour prédire de nouvelles données (Agelet & Hurburgh 2010).

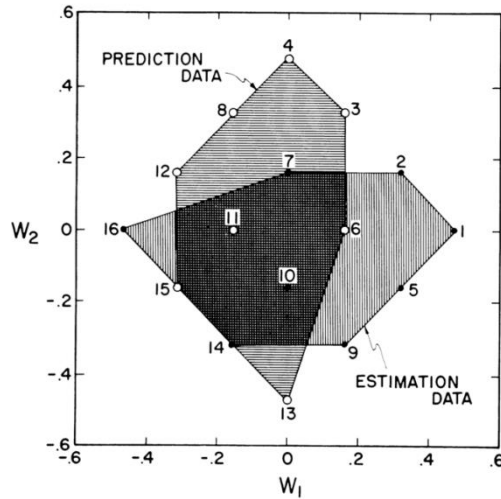


Figure 12: Algorithme Duplex appliqué sur des données au sein d'un espace à deux dimensions. Les données les plus éloignées sont successivement sélectionnées et rassemblées au sein de deux jeux de données : un pour calibrer le modèle (de classification ou de régression, par exemple), l'autre pour tester ce modèle. Illustration tirée d'un article de Snee (1977).

Dans le cas de la **validation externe**, de nouvelles données sont utilisées pour tester les performances des modèles calibrés. Cette méthode est la plus sûre car elle évalue mieux les capacités des modèles à réaliser des prédictions sur de nouvelles données, mais elle demande d'avoir accès à de plus larges ensembles de données que celles utilisées pour la construction des modèles (Faber & Kowalski 1997).

2.6. Avantages et désavantages de l'imagerie hyperspectrale proche infrarouge

2.6.1. Les avantages

En plus des qualités de la spectroscopie classique déjà cité précédemment (point 2.1 de ce chapitre), l'imagerie hyperspectrale permet de collecter des informations spatiales et spectrales détaillées et cela en une large quantité. Elle surpasse donc l'imagerie RGB (*Red Green Blue*, c.-à-d. l'imagerie utilisant les trois couleurs visibles), la spectroscopie proche infrarouge et l'imagerie multispectrale (Huang et al. 2014). Contrairement à la spectroscopie classique, elle permet d'étudier la distribution de constituants multiples au sein d'échantillons et elle n'en donne pas seulement une composition moyenne. Elle permet aussi de déterminer des régions d'intérêt après l'acquisition des images et d'en extraire les spectres. Le spectre d'un pixel peut être considéré comme étant la signature spectrale d'une région particulière et être enregistré dans une librairie spectrale si besoin. Grâce à la richesse des informations spatiales et spectrales recueillies avec cette technique, il est possible de

détecter et de discriminer des objets même s'ils ont la même couleur, la même morphologie (Wu & Sun 2013) pour autant que leurs spectres soient suffisamment différents.

2.6.2. Les désavantages

Le coût

Selon Türker-Kaya & Huck (2017), un des principaux désavantages de l'imagerie hyperspectrale est son coût. La collecte des données nécessite des détecteurs sensibles et leur analyse ne peut se faire que grâce à des ordinateurs avec de fortes capacités de calcul. La rapidité du traitement des données est donc un défi important avec cette technologie (Vermeulen et al. 2012) car la quantité importante de données acquises limite la vitesse d'acquisition et de traitement des données (Wu & Sun 2013). De plus, une grande capacité de stockage de données est nécessaire vu la taille des images hyperspectrales (Türker-Kaya & Huck 2017).

La qualité des données de référence

De la même manière que la spectroscopie classique, l'imagerie hyperspectrale est une technique d'analyse indirecte. Des échantillons de référence sont donc nécessaires. En principe, il faudrait connaître la valeur de référence à l'échelle du pixel pour pouvoir construire des modèles de quantification sur des images hyperspectrales. Cependant, obtenir une valeur de référence à cette échelle n'est pas faisable avec les méthodes d'analyses traditionnelles. Lorsque l'imagerie hyperspectrale proche infrarouge est utilisée à des fins de quantification, les données utilisées comme référence sont souvent obtenue sur une zone plus large et sont donc des valeurs moyennes (Manley 2014).

La complexité des données acquises

Par rapport à l'imagerie conventionnelle et à la spectroscopie classique utilisées séparément, un autre désavantage de l'imagerie hyperspectrale est la complexité des données acquises. Les données acquises par imagerie hyperspectrale peuvent en effet être affectées par des variations au niveau du système d'acquisition et par des facteurs externes. Elles contiennent dès lors du bruit. Des prétraitements sont donc nécessaires avant de pouvoir extraire l'information utile des images (Ravikanth et al. 2015).

La présence d'ombres sur les images hyperspectrale est un des facteurs affectant les données acquises (Wu et al. 2012; Bochow et al. 2011; Rodionova et al. 2005; Lawrence et al. 2004). Elles sont dues à une illumination non homogène de l'échantillon (Boldrini et al. 2012). Comme l'illustre la Figure 13, l'importance des zones d'ombre sur les images hyperspectrales dépend de la configuration du système d'acquisition. Les systèmes d'acquisition dans lesquels l'illumination est réalisée de manière diffuse permettent de réduire au maximum les zones d'ombre sur l'échantillon et le support d'acquisition des images.

Si le système d'acquisition utilisé crée des zones d'ombre sur les images hyperspectrales, il est possible de les retirer par traitement de ces images (Zhong et al. 2009). Ce traitement peut être réalisé grâce à des masques appliqués sur les images suite à la détermination d'un seuil au-delà duquel les pixels de l'image sont

classés comme étant de l'ombre. C'est ce qui fut par exemple utilisé par Bochow et al. (2011) sur des images acquises par télédétection et par Wu et al. (2012) et Zhong et al. (2009) sur des images hyperspectrales acquises en laboratoire. Les ombres peuvent aussi être prédites par les modèles de discrimination en incluant leurs spectres dans le set de calibration. Les zones ombragées sont alors classées soit avec les éléments étudiés (Lawrence et al. 2004), soit éliminées de l'image. L'algorithme SVM utilisé dans nos recherches pour ses bonnes performances en classification a plusieurs fois montré sa capacité à classer ces zones d'ombre (Tolt et al. 2011; Hearst et al. 1998).

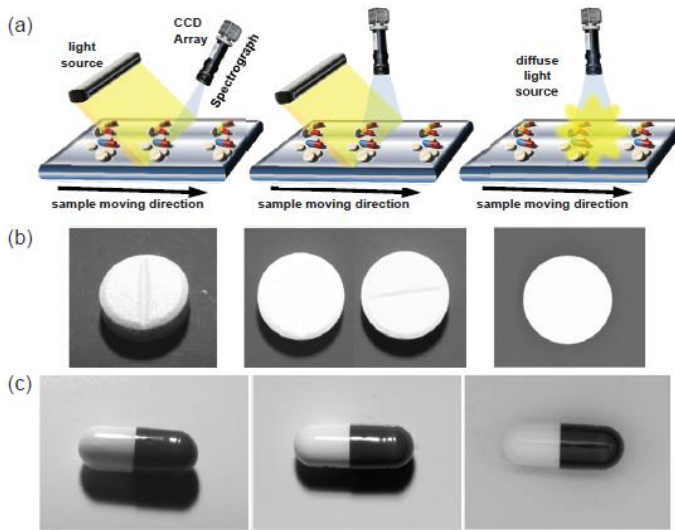


Figure 13: Importance des zones d'ombre et de la réflexion spéculaire sur des images selon la configuration du système d'acquisition et la forme des échantillons. (a) différents dispositifs optiques pour la mesure de la lumière spéculaire et de la lumière réfléchie diffuse, (b) un échantillon plane sous différents dispositifs d'éclairage et de détection ou avec une source d'éclairage diffus, (c) les mêmes dispositifs optiques appliqués à un objet de forme cylindrique avec une réflexion spéculaire potentiellement forte. Illustration extraite de Boldrini et al. (2012).

Un autre facteur affectant les données acquises est la forme de l'échantillon. Si la surface d'acquisition des pixels n'est pas plate, la lumière incidente va être réfléchi différemment. Le profil du spectre est donc modifié et cela peut mener à de mauvaises classifications. De tels effets ont été mis en évidence par plusieurs auteurs en bordure de leurs échantillons (M. Kamruzzaman et al. 2012; Williams et al. 2009).

Notons enfin que le profil final d'un spectre ne dépend pas uniquement de l'élément sur lequel un pixel est acquis, mais aussi de la nature des pixels voisins (Amigo 2010).

L'interprétation des données

Les modèles utilisés pour analyser les spectres sont souvent développés sur des constituants purs. Lorsqu'un échantillon contient différents éléments, les spectres acquis au niveau d'un pixel sont souvent le mélange de spectres des différents constituants et ils peuvent être acquis sur d'autres constituants présents au sein de l'échantillon, mais qui ne sont pas pris en compte par les modèles utilisés (Vermeulen et al. 2012). Si les éléments étudiés ont une taille importante, ces pixels se trouveront majoritairement en bordure des éléments. Si par contre l'élément étudié est de très petite taille, qu'il est plus petit qu'un pixel, son spectre sera inévitablement compris dans le spectre du pixel avec le spectre d'autres constituants (Amigo 2010). Ce phénomène est connu sous le nom de "pixels mixtes" et posent problème lors de l'interprétation des données extraites des images hyperspectrales.

L'interprétation des données est également difficile lors de la présence sur les images hyperspectrales de corps étrangers, dont les spectres ne sont pas pris en compte dans les modèles.

3. Utilisations de la spectroscopie dans l'étude racinaire

Dans le cadre de l'étude des racines, plusieurs techniques spectroscopiques ont été utilisées que ce soit pour la détermination de taxons, la quantification de biomasse racinaire ou la détermination de l'état de racines (vivantes ou mortes). Un aperçu de différentes utilisations de la spectroscopie dans l'étude de systèmes racinaires est donné dans les paragraphes qui suivent. La plupart de ces utilisations ont été synthétisées par Rewald et Meinen (2013) dans une révision de la littérature sur le sujet.

3.1. La spectroscopie de réflectance proche infrarouge dans l'étude racinaire

La spectroscopie de réflectance proche infrarouge est un outil rapide et rentable dans l'analyse de constituants et elle est considérée comme un outil utile pour des applications en écologie (Rewald et al. 2012). Dans l'étude des systèmes racinaires, elle permet de discriminer et de quantifier des racines de différentes espèces et de différents états (mortes ou vivantes). Elle permet aussi d'estimer des densités racinaires au sein de carottes de sol. Enfin, combinée à l'imagerie, elle permet de discriminer les racines du sol et des racines saines de racines infectées par des nématodes.

3.1.1. Discrimination d'espèces, de racines vivantes et mortes, et mesures de quantités relatives de racines

La discrimination et la quantification de la composition spécifique au sein de mélanges de racines sont des prérequis fondamentaux dans l'étude des interactions souterraines des plantes (Lei & Bauhus 2010). L'utilisation de la spectroscopie proche infrarouge dans le tri de racines de différentes espèces s'est montré être une

alternative prometteuse au tri manuel. Ce tri est basé sur l'identification visuelle des espèces et de la vitalité des racines fines sur base de critères morphologiques tels que la couleur, le diamètre ou l'ordre de ramification par exemple. Il prend beaucoup de temps et sa la qualité peut être influencée par l'opérateur (Rewald et al. 2012; Lei & Bauhus 2010).

La spectroscopie fut pour la première fois utilisée dans l'étude racinaire par Rumbaugh et al. (1988) pour prédire la proportion de biomasse racinaire de luzerne (*Medicago sativa* L.) au sein de mélanges binaires avec des racines de quatre espèces de graminées. Les mélanges préparés avec les racines de luzerne et de graminées contenaient des proportions de racines allant de 0 à 100 % de chaque espèce. Une forte corrélation a été obtenue entre les valeurs prédites sur base des spectres proche infrarouge acquis sur les mélanges et les valeurs réelles des proportions de racines.

Roumet et al. (2006) ont utilisé la spectroscopie proche infrarouge pour prédire des proportions de racines au sein de mélanges contenant des racines de deux ou trois espèces obtenues sur des plantes cultivées en serres ou récoltées dans des prairies. Ici encore, la proportion de racines de chaque espèce fut bien prédite ($r^2 = 0,97$ à $0,99$). Dans cette étude, la capacité de discrimination de racines sur base de leurs spectres proche infrarouge a été évaluée aussi bien sur des espèces apparentées (différentes espèces de poacées pérennes) que sur des espèces appartenant à deux familles botaniques différentes (Asteracées et Poacées).

La spectroscopie en proche infrarouge a également été utilisée par Picon-Cochard et al. (2009) afin de prédire le pourcentage de racines vivantes et mortes au sein d'échantillons de racines de cinq graminées cultivées séparément en prairies.

Lei et Bauhus (2010) ont quant à eux utilisé la spectroscopie proche infrarouge pour mesurer la proportion de racines au sein de mélanges contenant des racines fines d'arbres et de graminées. Les mélanges contenaient jusqu'à cinq classes de racines.

Il est important de noter que dans toutes ces études, ce sont des quantités relatives de racines qui ont été déterminées, sous forme de proportions au sein de mélanges réalisés de manière artificielle, et non des quantités absolues. Dans tous les cas, les acquisitions étaient réalisées en laboratoire sur des échantillons de racines lavées et broyées.

3.1.2. Quantification de racines et étude de leur distribution

La méthode mise au point par Kusumo et al. (2009) a permis pour la première fois de mesurer une densité racinaire directement dans un volume de sol sans devoir extraire les racines du sol. Ils ont pour cela mis au point une sonde permettant d'acquérir des spectres en mode réflexion dans le visible et le proche infrarouge sur une tranche de sol. Un modèle de régression a ensuite été calibré afin de relier l'allure du spectre infrarouge à la densité racinaire mesurée par la méthode de référence qui consiste à laver, sécher et peser les racines ($r^2 = 0,85$). Cette étude avait été réalisée sur des racines de raygrass italien (*Lolium multiflorum* Lam.) obtenues dans des cultures réalisées en pots et en serres.

Cette méthode a ensuite été testée sur des échantillons de sol prélevés en prairies (Kusumo et al. 2010). La sonde permettant l'acquisition des spectres visible-proche infrarouge a cette fois-ci été utilisée sur la surface horizontale de carottes de sol (80 mm de diamètre). Les mesures ont été réalisées à des profondeurs de sol de 15, 30 et 60 mm dans deux sites différents. Les auteurs ont montré que la qualité des prédictions de la densité racinaire était influencée par la nature du sol (r^2 obtenus lors de la validation croisée variant entre 0,75 et 0,83 selon la nature du sol) et que les régressions mises au point sur des échantillons obtenus en serre n'étaient pas applicables sur les carottes de sol prélevées en prairies. Une calibration séparée devait donc être réalisée pour chaque site d'étude et chaque condition de culture.

Dans une troisième étude menée par les mêmes auteurs, la méthode a été appliquée dans une culture de maïs (Kusumo et al. 2011). Dans cette étude, la spectroscopie visible-proche infrarouge a été utilisée pour prédire la densité racinaire du maïs dans des carottes de sol prélevées à différentes distances du rang de la culture et à différentes profondeurs. La densité racinaire du maïs déterminée sur base du spectre visible-proche infrarouge acquis sur des carottes de sol était modérément bien prédite ($r^2 = 0,75$ lors de la validation croisée).

Xu et al. (2016) ont utilisé la spectroscopie visible et proche infrarouge d'une manière similaire pour déterminer la densité racinaire du riz (*Oryza sativa* L.) dans des carottes de sol. Dans cette étude, ils ont également testé trois méthodes de régression multivariées pour prédire la densité racinaire. Il s'agissait des méthodes de régression PCR (*Principal component regression*), PLSR (*Partial least square regression*) et SVMR (*Support vector machine regression*). Parmi ces méthodes, la régression SVM a donné les meilleurs résultats et permettait de prédire la densité racinaire avec un r^2 de 0.88. Une prédiction de la densité racinaire encore meilleure a été obtenue en utilisant un nombre réduit de longueurs d'onde (Xu et al. 2017).

3.1.3. Discrimination de racines et du sol

Afin de s'affranchir de l'étape de lavage et pouvoir réaliser des mesures de manière non destructive, plusieurs auteurs ont utilisé l'imagerie hyperspectrale proche infrarouge étendue au spectre visible sur des systèmes racinaires obtenus dans des rhizotrons. Dans une étude menée sur des boutures de peuplier hybride (*Populus deltoides* x *Populus euramericana*) cultivées en conditions contrôlées, Nakaji et al. (2008) ont démontré que l'imagerie spectrale visible-proche infrarouge utilisée dans la gamme spectrale comprise entre 480 et 972 nm permettait de classer des racines vivantes en quatre classes d'âges et de discriminer des racines mortes, du terreau et du sol. Bodner et al. (2017) ont également testé les potentialités de l'imagerie hyperspectrale visible-proche infrarouge pour discriminer des racines de betterave sucrière (*Beta vulgaris* L.) et du sol.

3.1.4. Détection de kystes sur des racines

L'imagerie hyperspectrale a aussi été utilisée sur des racines de betteraves pour y détecter la présence de kystes causés par des nématodes (*Heterodera schachtii*). Dans le cas de cette étude menée par Fernández Pierna et al. (2012), l'imagerie hyperspectrale proche infrarouge a été utilisée pour détecter les kystes sur les

racines, les compter et classer des plants de betteraves en fonction de leur résistance aux nématodes.

3.2. La spectroscopie en moyen infrarouge dans l'étude racinaire

La spectroscopie en moyen infrarouge à transformée de Fourier et réflectance totale atténuée permet l'acquisition de donnée en surface des échantillons (quelques μm) alors que la spectroscopie en proche infrarouge permet des analyses légèrement plus en profondeur (de 27 à 180 μm de profondeur selon une étude de Clarke et al. (2002)). De plus, les analyses réalisées par moyen infrarouge se basent sur une surface d'analyse faible par rapport à la spectroscopie proche infrarouge. Le volume d'échantillon analysé est donc beaucoup plus faible en spectroscopie moyen que proche infrarouge. Cependant, comparée à la spectroscopie en proche infrarouge, la spectroscopie en moyen infrarouge offre des spectres beaucoup plus caractéristiques pour l'identification de substances spécifiques (Meinen & Rauber 2015; Rewald et al. 2012).

Cette technique a été utilisée par Naumann et al. (2010) afin de discriminer des fragments de racines de pois (*Pisum sativum* L.) et d'avoine (*Avena sativa* L.) obtenues dans différentes conditions de culture. Une discrimination parfaite a été obtenue entre les racines des deux espèces.

L'utilisation de la spectroscopie en moyen infrarouge a également été rapportée par Rewald et al. (2012) sur des fragments de racines de betterave (*Beta vulgaris* spp), de colza (*Brassica napus* L.), de froment (*Triticum aestivum* L.) et de maïs (*Zea mays* L.) obtenus sur des cultures en serre. Une analyse typologique sur les spectres acquis a montré une séparation en quatre groupes distincts correspondant aux quatre espèces. Cette analyse a montré que les différences interspécifiques étaient plus importantes que l'hétérogénéité intraspécifique.

Meinen et Rauber (2015) ont testé les potentialités de cette même technique de mesure pour discriminer les racines de cultures et d'adventices qui leur sont associées. Leurs essais ont porté sur des racines de maïs et de panic pied-de-coq (*Echinochloa crus-galli* L.), d'orge (*Hordeum vulgare* L.) et de folle avoine (*Avena fatua* L.), de froment d'hiver et de vulpin des champs (*Alopecurus myosuroides* Huds.) ainsi que sur des racines de betteraves sucrières et de chénopode blanc (*Chenopodium album* L.). Les spectres étaient acquis soit sur des segments de racines fraîches ou sèches, soit sur des racines sèches et broyées. Une séparation claire des différentes espèces ne fut pas possible sur les racines fraîches, mais elle fut parfaite sur les racines séchées et broyées. Sur les racines séchées et non broyées, la qualité de la discrimination dépendait des espèces et des segments de racines (base, milieu ou extrémité) considérés. Ces auteurs ont également démontré la possibilité de quantifier la proportion de racines au sein de mélanges de racines broyées de maïs, folle avoine et panic pied-de-coq.

3.3. La spectroscopie de fluorescence dans l'étude racinaire

Contrairement aux techniques de spectroscopie utilisant l'infrarouge (proche ou moyen), la spectroscopie en fluorescence utilise les régions du spectre électromagnétique comprises dans le visible et l'ultraviolet.

Wasson et al. (2016) ont développé un système portable d'imagerie en fluorescence qui permet le comptage automatique de racines à la surface de carottes de sol et remplace donc le comptage visuel nécessaire dans la méthode *core break*. Leurs résultats ont montré que la corrélation était meilleure entre la densité de longueur racinaire déterminée au sein de carottes de sol lavées (méthode de référence) et les prédictions réalisées sur base des mesures acquises en fluorescence qu'avec des mesures obtenues par comptage visuel par un opérateur sur des carottes de sol cassées ($r^2 = 0,68$ au lieu de $0,57$). Comparé à la méthode basée sur le comptage visuel, ce système permettait aussi de traiter un plus grand nombre d'échantillons par heure (4,3 carottes/heure/personne au lieu de 3,7). Selon les auteurs de l'étude, ce système pourrait permettre de faire du phénotypage racinaire directement au champ. Il pourrait aussi permettre d'acquérir des informations sur la santé racinaire ou d'autres paramètres liés aux racines.

3.4. Contraintes dans l'utilisation de la spectroscopie pour l'étude de systèmes racinaires

Pour la plupart des analyses en spectroscopie infrarouge, les racines doivent être extraites du sol. La préparation des échantillons se limite ensuite souvent à un séchage et un faible broyage est parfois réalisé. L'étape de séchage des échantillons est nécessaire avant l'acquisition des échantillons car l'eau a une absorbance forte et elle peut donc masquer la présence d'autres composants. Bien qu'une étude ait montré que des racines fraîches pouvaient également être distinguées sur base de leurs spectres, leur distinction est meilleure lorsqu'elles sont séchées et broyées (Rewald et al. 2012). Plusieurs travaux ont cependant montré qu'il était possible d'étudier des racines fraîches sans les extraire de leur substrat (Xu et al. 2016; Kusumo et al. 2010; Kusumo et al. 2011; Nakaji et al. 2008).

Selon Rewald et al. (2012), la nécessité d'avoir à sa disposition des séries de calibration pour déterminer les ratios d'espèces au sein des échantillons est la principale limitation dans l'utilisation de la spectroscopie infrarouge dans l'étude de systèmes racinaires. La création de mélanges artificiels de racines demande de cultiver ces racines séparément pour s'assurer qu'elles appartiennent bien à l'espèce cible. Cette étape est laborieuse et prend du temps, d'autant plus que la composition chimique des racines peut évoluer au cours de son développement et être influencée par son environnement.

CHAPITRE II

Objectifs de la thèse

Objectifs de la thèse

Cette thèse de doctorat a poursuivi plusieurs objectifs afin d'évaluer les potentialités de l'imagerie hyperspectrale proche infrarouge et de la chimiométrie dans l'étude de systèmes racinaires.

Lorsque les systèmes sont étudiés par carottage de sol, les échantillons doivent être lavés pour éliminer les particules de sol. Restent alors les racines et divers éléments de matières organiques, dont des résidus de cultures, qui doivent être triés avant de pouvoir réaliser des mesures sur les racines. Jusqu'à présent, ce tri est réalisé manuellement. Il prend beaucoup de temps, il est fastidieux et il dépend de la subjectivité de l'opérateur qui le réalise. Si des racines de différentes espèces sont présentes au sein d'un même échantillon, ce tri est d'autant plus difficile.

Le premier objectif de cette thèse était d'évaluer dans quelle mesure l'imagerie hyperspectrale proche infrarouge et la chimiométrie permettaient de discriminer les spectres des éléments extraits d'échantillons de sol et offraient une opportunité pour pouvoir se passer du tri manuel de ces éléments. Cette évaluation a d'abord été réalisée sur des racines de froment d'hiver et des résidus de la même culture. Une deuxième évaluation a été réalisée sur des résidus de froment, des racines de froment et de pois d'hiver et des nodosités extraits d'échantillons de sol prélevés sous des cultures associées.

Un second objectif était de déterminer quel traitement des spectres permettait d'obtenir la meilleure discrimination des éléments présents sur les images hyperspectrales. Nous avons donc comparé différents prétraitements de spectres et différents modèles de discrimination.

Le troisième objectif de nos travaux était d'évaluer la possibilité d'estimer la masse de racine présente au sein d'un échantillon sur base du nombre de pixels classés comme étant des racines sur l'image hyperspectrale acquise sur cet échantillon. Cette évaluation a également été réalisée pour estimer la masse de résidus de cultures. Nous avons ensuite testé cette méthode de quantification dans plusieurs essais agronomiques.

Suite à l'observation, au cours du temps, de différences entre les prédictions de la nature des pixels d'images hyperspectrales acquises sur un échantillon de référence, un autre objectif de la thèse a été de caractériser l'instabilité du système d'acquisition des images hyperspectrales et d'y apporter des solutions.

Enfin, le dernier objectif de cette thèse a été d'évaluer les potentialités de l'imagerie hyperspectrale proche infrarouge et de la chimiométrie pour la quantification de la leghémoglobine au sein de nodosités individuelles de pois protéagineux afin de remplacer la méthode actuellement utilisée qui prend beaucoup de temps, nécessite l'emploi de réactifs toxiques et ne permet pas de faire des mesures sur des nodosités individuelles.

CHAPITRE III

Matériels et méthodes

1. Introduction

Afin d'avoir une bonne vue d'ensemble du chemin du traitement des données entre le prélèvement de carottes de sol contenant les racines et la quantification de ces racines, la Figure 14 reprend de manière schématique toutes les étapes nécessaires à la construction des modèles permettant la détection de racines sur une image hyperspectrale prise en proche infrarouge, la quantification de ces racines et l'application des modèles de discrimination et de quantification sur de nouveaux échantillons.

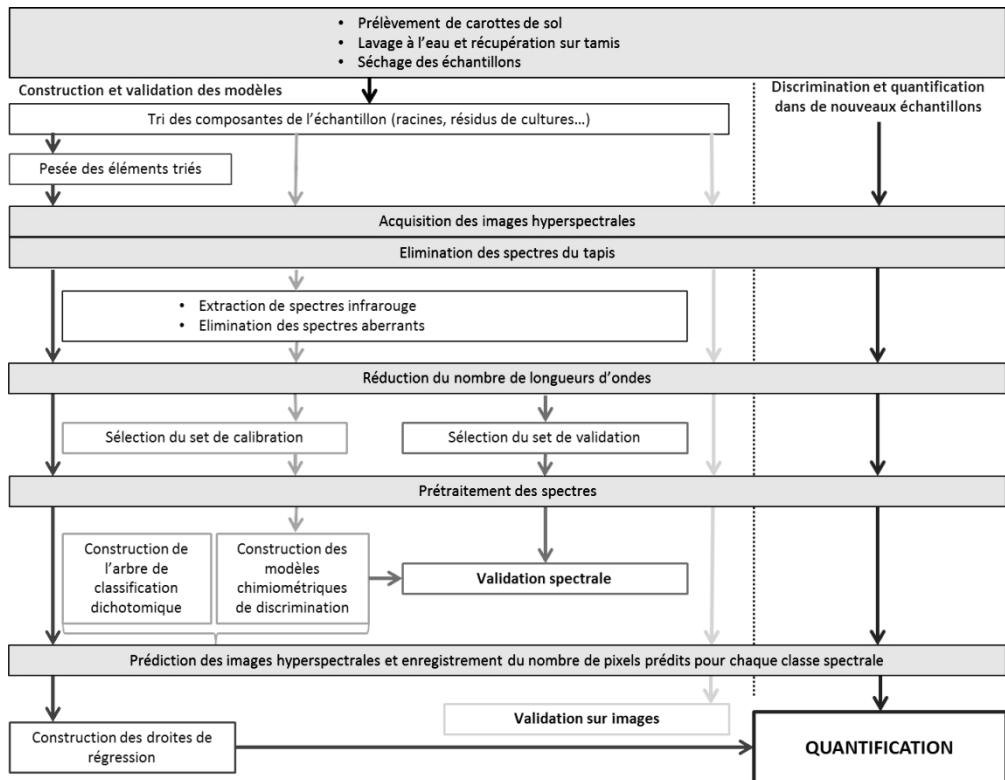


Figure 14: Schéma général présentant les étapes de la quantification des racines, de l'acquisition d'échantillons de racines permettant la calibration et la validation des modèles de discrimination jusqu'à l'application de ces modèles sur de nouveaux échantillons de racines.

Le schéma général de la quantification de racines contenues dans des échantillons de sol par imagerie hyperspectrale proche infrarouge consiste (i) à laver ces échantillons de sol pour en extraire les racines, (ii) à sécher ces racines, (iii) à

acquérir les spectres proche infrarouge des échantillons par imagerie hyperspectrale, (iv) à extraire et prétraiter le signal spectral de ces images, (v) à détecter le nombre de pixels de racines présents sur les images hyperspectrales et (vi) à relier ce nombre de pixels de racines à une masse de racines.

La mise au point des modèles permettant cette quantification nécessite l'extraction de spectres proche infrarouge sur des échantillons de référence et la calibration de modèles chimiométriques de discrimination permettant de prédire la nature de chaque pixel des images hyperspectrales acquises sur les échantillons de racines après leur lavage et leur séchage. Elle demande aussi la pesée d'échantillons de racines afin de calculer l'équation de la régression linéaire qui permet de relier un nombre de pixels classés comme étant des racines sur les images hyperspectrales à une masse de racine.

Enfin, afin de vérifier la validité des modèles, ceux-ci sont testés sur des spectres différents de ceux utilisés pour construire les modèles et sur des images.

2. Préparation des échantillons

2.1. Prélèvement des racines

Les échantillons de racines ont été obtenus par prélèvement de carottes de sol dans les conditions réelles de culture. Lorsque la végétation était fortement développée et rendait l'accès au sol difficile, elle était coupée à ras du sol avant de réaliser les prélèvements de racines.

Les prélèvements ont été réalisés à l'aide d'une sonde de 30 mm de diamètre sur une profondeur de 30 à 90 cm selon les cas d'études. Cette sonde était enfoncée dans le sol à l'aide d'un vérin hydraulique monté à l'avant d'un tracteur (Figure 15). Le système utilisé permettait de déplacer le vérin de gauche à droite ce qui permettait de prélever plusieurs carottes de sol l'une à côté de l'autre sans devoir déplacer le tracteur. Elles pouvaient ainsi être facilement prélevées dans la ligne de culture ou dans l'interligne selon les besoins. Selon les études que nous avons menées, deux ou cinq carottes de sol étaient prélevées dans la ligne de culture et l'interligne pour chaque parcelle expérimentale. Lors des prélèvements, les carottes de sols étaient divisées en section de 10 à 20 cm de long selon les horizons de sol étudiés. Elles étaient ensuite mises dans des sacs en plastique et conservées à 4 °C jusqu'à un traitement ultérieur.

2.2. Lavage des échantillons de sol et récupération des racines

2.2.1. Mise en suspension

Afin de faciliter la désagrégation de la terre contenue dans les carottes de sol et en extraire plus facilement les racines, les carottes de sol ont été immergées durant 48 h dans une solution contenant 50 g de sel de cuisine (NaCl) par litre d'eau. Pour faciliter la manutention des échantillons, chaque échantillon a été mis à tremper dans la solution dans un seau avec un couvercle. Les seaux ont été vigoureusement secoués à la main après l'ajout de la solution salée aux carottes de sol et juste avant

l'étape de lavage. Notons que le sel n'absorbe pas dans le proche infrarouge (Manley 2014). Ce traitement n'affecte donc pas les spectres proche infrarouge des éléments lavés.



Figure 15: Vérin hydraulique et sonde permettant le prélèvement de carottes de sol.

2.2.2. Lavage

Après trempage dans la solution salée et désagrégation de la terre, les carottes de sol ont été transférées dans un système de lavage construit au sein de l'Unité de Phytotechnie tempérée (GxABT).

Ce système, similaire à celui décrit par McKell et al. (1961), est constitué de plusieurs unités de lavage indépendantes (Figure 16). Chaque unité de lavage consiste en un seau à la base duquel deux entrées d'eau créent un mouvement d'eau ascendant qui agite le contenu du seau et entraîne les éléments mis en suspension vers un tube qui est situé au milieu du seau et qui est terminé par un tamis à fines mailles (1,6 mm, type moustiquaire). L'eau entraîne et élimine donc les fines particules de terre et les éléments plus gros sont retenus par le tamis. Parmi ces éléments se trouvent les racines, mais également des résidus de la culture précédente, des particules de sol d'une taille supérieure aux mailles du tamis et des macroorganismes du sol.

Lorsque l'eau circulant dans le seau devenait claire, les éléments récupérés sur le tamis situé à l'extrémité du tube central étaient transférés sur un tamis de plus

grande taille et mis dans un sachet micro-perforé destiné au séchage en étuve. Les éléments encore présents au fond du seau étaient également transférés sur un tamis, aspergés à l'eau afin de retirer la terre résiduelle et les racines restantes étaient combinées avec le reste de l'échantillon avant la mise à l'étuve.

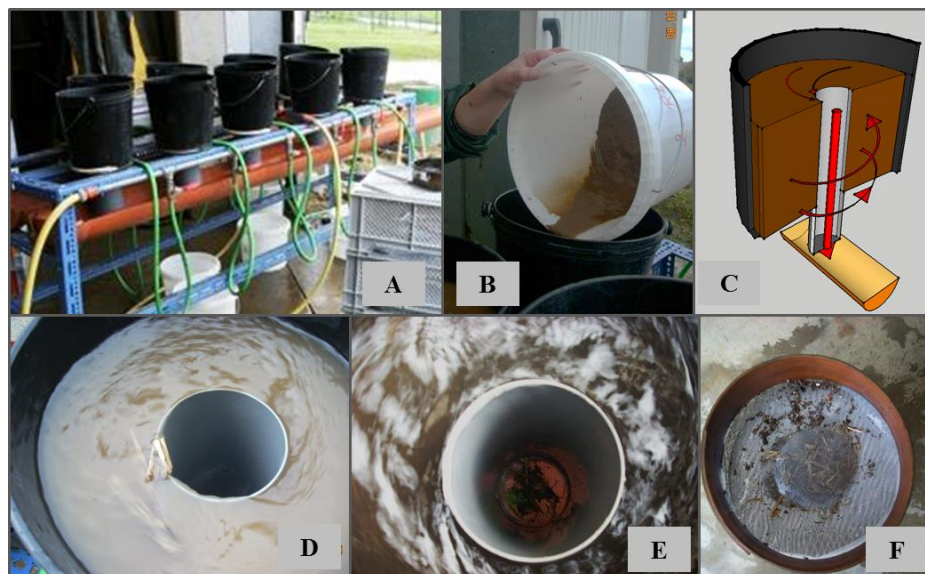


Figure 16: A : Vue générale du système de lavage des racines. Chaque seau est relié à sa base à une arrivée d'eau indépendante. L'eau entre en deux points au fond du seau et crée un mouvement d'eau ascendant. L'eau boueuse est évacuée par un tube situé au centre du seau et s'écoule dans une gouttière; B : Échantillon de sol transféré dans un seau après une immersion de 48 h dans une solution salée (50 g NaCl.l^{-1}); C : Mouvement de l'eau et évacuation; D : Les éléments en suspension tels que les racines et les résidus de cultures sont emportés par le courant d'eau vers le tube au centre du seau; E : Le tube central est terminé par un tamis qui retient tous les éléments plus grands que la taille des mailles (1,6 mm); F : Les racines et les résidus sont récupérés sur un tamis. Illustration adaptée du poster de Vancutsem et al. (2012).

2.2.3. Séchage

Le séchage des échantillons a été réalisé dans une étuve à $60 \text{ }^\circ\text{C}$ durant environ 48 h jusqu'à atteindre une masse constante. L'objectif du séchage était double : (i) extraire des échantillons l'eau libre qui pose problème en spectroscopie, car elle réagit fortement avec la lumière, masquant d'autres informations qui pourraient être intéressantes; (ii) permettre une bonne conservation des échantillons qui peuvent donc par la suite être maintenus dans les conditions ambiantes du laboratoire sans risque de se dégrader (Oliveira et al. 2000).

2.2.4. Tri et pesée des racines

Lors de l'étape de lavage, tous les éléments ayant une taille supérieure à la taille des mailles du tamis utilisé ont été récupérés. Ces éléments comprenaient des racines vivantes, la base des tiges de plantes qui se trouve sous la surface du sol et un ensemble de matières organiques mortes telles que des résidus de la culture précédente, des graines d'adventices, des racines et des bases de tige de la culture précédente de même que des racines mortes de la culture en place lors du prélèvement. Les Figures 17 et 18 illustrent la diversité des éléments extraits des échantillons de sol.



Figure 17: Echantillon de sol après lavage et séchage. Cet échantillon contient des racines, des résidus de la culture précédente, de petits cailloux et des particules de sol. L'échantillon est posé sur le tamis ayant servi à sa récupération après le lavage et comme support lors du séchage à l'étuve.

Un des objectifs de l'étude menée sur l'utilisation de l'imagerie hyperspectrale proche infrarouge dans l'étude des systèmes racinaires était de pouvoir faire la discrimination des racines et des autres éléments présents au sein des échantillons (résidus de cultures, cailloux, petites mottes de terre non éliminées par le lavage) sans devoir passer par un tri manuel. Pour atteindre cet objectif, les modèles de discrimination permettant cela devaient être préalablement calibrés sur des spectres de référence pour chaque classe. Le tri d'un certain nombre d'échantillons a donc été nécessaire (Figure 18). Ce tri manuel a été réalisé avec l'aide d'une pince fine (type pince à épiler) et était basé sur l'aspect visuel (couleur et forme) des éléments présents dans l'échantillon, les résidus de culture étant généralement plus foncés et plus larges que les racines. Le contenu des échantillons était réparti en trois classes : les racines, les résidus de la culture précédente et le sol (particules de terre et petits cailloux). La distinction des racines mortes et vivantes n'étant pas facile, elles ont été regroupées.

Le choix des échantillons triés, la recherche de variabilité au sein des échantillons et la rigueur avec laquelle le tri est réalisé déterminent la qualité des modèles qui sont calibrés sur des échantillons de référence (Lebot 2012). Cette étape de tri a donc été menée de manière à ne prendre en compte que des éléments bien identifiés, mais très variés.



Figure 18: Tri de racines et de résidus de cultures présents au sein d'un échantillon de sol. Au plus le tri avance, au plus la taille des éléments et leur ressemblance rendent leur tri difficile.

Plusieurs échantillons triés ont été pesés afin de calibrer et valider les modèles permettant la quantification des racines. Dans le cadre d'études s'intéressant à la quantification de résidus de cultures, cette étape de pesée a également été réalisée pour les résidus.

Les étapes de tri et de pesée des échantillons ont uniquement été nécessaires dans le cadre de la calibration et de la validation des modèles de discrimination et de quantification. L'objectif du travail réalisé dans ce cadre de cette thèse était de s'en passer par la suite grâce aux données fournies par l'imagerie hyperspectrale proche infrarouge.

3. Acquisition des images hyperspectrales et des spectres proche infrarouge

3.1. Le système d'acquisition

3.1.1. Description du système d'acquisition des images hyperspectrales proche infrarouge

Le système d'acquisition qui a été utilisé dans le cadre de cette thèse (Figure 19) faisait l'acquisition des images par balayage linéaire (système *line scan* aussi appelé *push-broom*). L'acquisition était réalisée par lignes successives de 320 pixels de large. Pour chaque pixel, 32 mesures ont été réalisées et un spectre moyen a été enregistré ce qui a permis d'améliorer le rapport signal-sur-bruit de la mesure. Ces spectres ont été acquis dans la gamme de longueurs d'onde comprises entre 1118 et 2424 nm avec une résolution spectrale de 6,3 nm. Au total, ce sont donc des spectres proche infrarouge constitués de 209 longueurs d'onde qui ont été acquis.

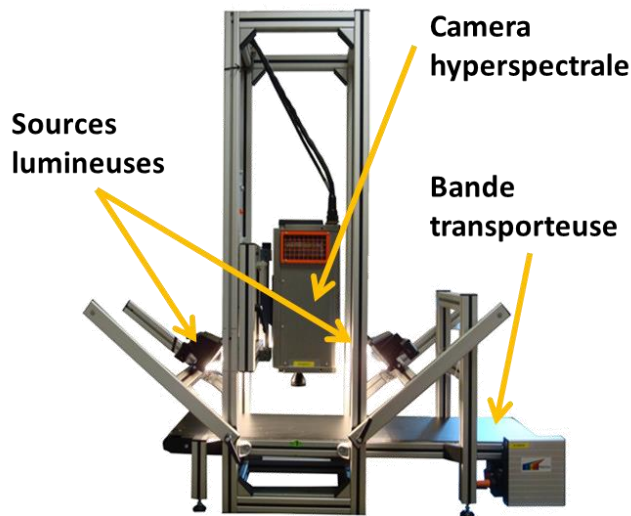


Figure 19: Système d'acquisition des images hyperspectrales proche infrarouge (BurgerMetrics SIA, Riga, Lettonie) installé au Centre wallon de Recherches agronomiques (Gembloux, Belgique). L'échantillon qui est éclairé par deux spots halogènes est posé sur une bande transporteuse et passe sous la caméra.

Ce système d'acquisition, mis au point par la société lettonne BurgerMetrics (Riga, Lettonie), était installé au Centre wallon de Recherches agronomiques (Département Valorisation des productions, Unité Qualité des produits). Il a été

décrit en détails par Vermeulen et al. (2012). Il était constitué d'une **caméra** SWIR XEVA CL 2.5 320 TE4 de la société XENICS nv (Leuven, Belgique) qui contenait un **spectrographe** ImSpector N25E fourni par la société SPECIM Ltd. (Oulu, Finlande). Le détecteur du spectrographe était de type mercure-cadmium-telluride (MCT) et sa température était stabilisée. Un **objectif**, positionné à l'entrée de la caméra, permettait de faire la mise au point de l'image. Le grossissement et le positionnement de l'objectif déterminaient la largeur de l'image. Dans le cadre de nos mesures, la largeur des images était de 10 cm. Le **système d'éclairage** de l'échantillon était constitué de deux lampes halogènes de 120 watts positionnées de part et d'autre de l'échantillon, dans le sens de son déplacement. Le support des échantillons lors de l'acquisition des images hyperspectrales était une **bande transporteuse** qui permettait le déplacement de l'échantillon sous la caméra. Sa vitesse était adaptable. Le **logiciel** HyperProVB (BurgerMetrics SIA, Riga, Lettonie) a été utilisé pour commander l'acquisition des images et le déplacement de la bande transporteuse.

3.1.2. Calibration de la réflectance

Toutes les images ont été acquises en mode réflexion. C'est donc l'intensité lumineuse réfléchi par l'échantillon qui a été mesurée. Cette intensité lumineuse a ensuite été transformée en valeur de réflectance. Pour cela, le système d'acquisition d'images hyperspectrales a été calibré en mesurant l'intensité lumineuse acquise en bloquant l'entrée du signal de la scène analysée au sein du système d'acquisition (image noire) et l'intensité lumineuse acquise sur une référence blanche ayant une réflectance élevée et uniforme qui réfléchit 99 % de la lumière (image blanche). Une céramique blanche¹⁷ a été utilisée à cet effet. Sur base de ces intensités de références, la valeur corrigée de la réflectance (réflectance relative) a pu être calculée pour chaque pixel de l'échantillon dont le signal réfléchi était mesuré (Équation 2). Ces mesures de références ont été réalisées préalablement à l'acquisition des images hyperspectrales sur les échantillons étudiés et pour chaque longueur d'onde du spectre enregistré.

Équation 2: Transformation de l'intensité brute et non corrigée du signal perçu au niveau de la caméra hyperspectrale en réflectance relative sur base de l'intensité lumineuse mesurée en obstruant l'objectif de la caméra (image noire) et sur base de l'intensité lumineuse mesurée sur une surface hautement réfléchissante (image blanche).

$$\text{Réflectance relative} = \frac{\text{intensité brute et non corrigée} - \text{intensité image noire}}{\text{intensité image blanche} - \text{intensité image noire}}$$

¹⁷ Nous avons observé que la céramique devait être à température ambiante pour réaliser la calibration. Si elle s'était échauffée suite à une exposition trop longue sous les lampes halogènes, la calibration semblait être de moins bonne qualité.

Cette calibration du système d'acquisition a été réalisée de telle sorte à compenser le décalage dû au *dark current*¹⁸, à la dérive de la température du système d'éclairage et au manque d'uniformité spatiale de l'éclairage (Huang et al. 2014; Vermeulen et al. 2012).

Les effets de la fréquence de calibration ont été étudiés lors d'un suivi de la stabilité du système d'acquisition des images hyperspectrales. Ce suivi régulier de la valeur de la réflectance mesurée sur différents supports a permis de mettre en évidence la nécessité de ne pas prendre des images directement après l'allumage du système d'acquisition et de calibrer régulièrement le système d'acquisition des images hyperspectrales. Le détail de ce suivi est présenté au Chapitre IV. Suite à cette étude, un délai d'au moins 1h30 était observé entre l'allumage de la caméra et l'acquisition de la première image et les calibrations se faisaient de manière régulière : la référence blanche était acquise au moins une fois par heure et la référence noire était acquise automatiquement avant chaque image.

La valeur de la réflectance a ensuite été transformée en absorbance grâce à l'Équation 3 :

Équation 3: Relation entre la réflectance (R) et l'absorbance (A).

$$A = \log(1/R)$$

3.1.3. Détection et retrait des pixels défectueux

Le détecteur du système d'acquisition peut avoir des défauts de fabrication et se dégrader au cours de son utilisation ce qui entraîne l'apparition de pixels défectueux sur les images hyperspectrales. Le système d'acquisition a donc été calibré de telle manière à détecter et corriger ces pixels défectueux. Cette calibration était réalisée avec le logiciel d'acquisition des images (HyperProVB, BurgerMetrics SIA, Riga, Lettonie) en même temps que la calibration de la réflectance.

3.1.4. Mise au point de la netteté de l'image

La mise au point de la netteté de l'image a été réalisée sur une latte. L'objectif placé à l'entrée de la caméra était réglé de telle sorte à obtenir une image nette des graduations de la latte. Une image hyperspectrale de la latte était acquise et la netteté de l'image était évaluée sur base d'une prédiction en nuances de gris de l'image obtenue avec le logiciel de traitement des images HyperSee (BurgerMetrics SIA, Riga, Lettonie). Cette étape était particulièrement importante pour l'acquisition des racines qui sont des éléments très fins.

3.1.5. Réglage de la vitesse d'avancement de la bande transporteuse

Afin d'avoir toujours des pixels de forme carrée, la vitesse d'avancement de la bande transporteuse a été adaptée en début de chaque journée d'acquisition. Cette

¹⁸ Le *dark current* correspond au signal enregistré par le détecteur en l'absence d'éclairage lumineux.

vitesse variait légèrement d'un jour à l'autre en fonction du nombre de pixels défectueux détectés lors de la calibration du système d'acquisition.

Le réglage de la vitesse a été réalisé en faisant l'acquisition de l'image d'un cercle imprimé sur une feuille de papier et dont le rapport de la longueur (mesurée dans le sens d'avancement de la bande transporteuse) sur la largeur (mesurée perpendiculairement au sens d'avancement de la bande transporteuse) a été calculé (Figure 20). La vitesse d'avancement de la bande transporteuse retenue était celle qui permettait d'obtenir un rapport longueur/largeur du cercle le plus proche de un.

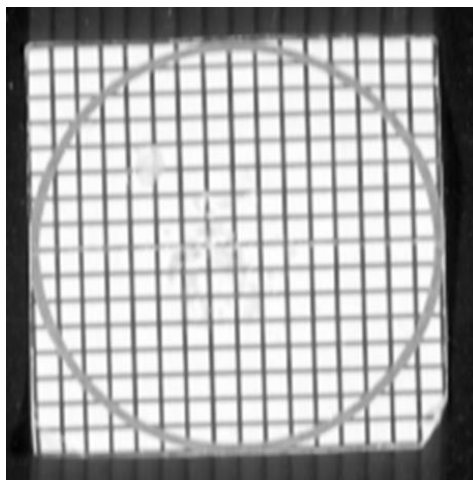


Figure 20: Etalon permettant de régler la vitesse d'avancement de la bande transporteuse du système d'acquisition des images hyperspectrales proche infrarouges. La forme du cercle permet de déterminer si la bande transporteuse a la bonne vitesse.

3.1.6. Vérification et suivi de la qualité du signal enregistré

Afin de déterminer la qualité du signal enregistré par le système d'acquisition, deux échantillons ont été utilisés comme références.

Le premier, illustré à la Figure 21, était un échantillon de grains de froment contenant de l'ergot (sclérotas de *Claviceps purpurea* Tul.). Cet échantillon de référence a également été utilisé par Vermeulen et al. (2012) pour suivre la stabilité du système d'acquisition. Le second était un échantillon contenant des résidus de cultures constitués de sept fragments de paille de froment d'hiver (Figure 22). Le suivi du nombre de pixels correctement prédits comme étant de l'ergot et des pailles par des algorithmes d'analyse discriminante permettait d'évaluer la qualité du signal enregistré. Une image hyperspectrale de ces deux échantillons était acquise au moins une fois en début de chaque journée d'acquisition. Si la qualité du signal était incertaine, d'autres images de ces références étaient acquises au cours de la journée et analysées. Lorsque la qualité du signal était faible, une nouvelle calibration du

système d'acquisition (calibration de l'absorbance et détection des pixels défectueux) était réalisée.

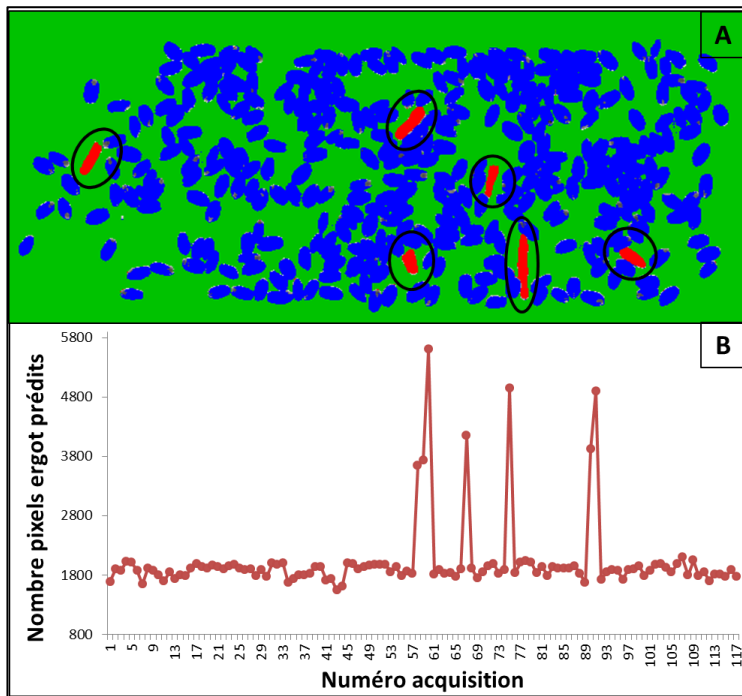


Figure 21: Prédiction d'une image hyperspectrale acquise sur un échantillon de grains de blé (en bleu sur l'image prédite) contenant six ergots (en rouge et entourés sur l'image prédite).

La qualité de la prédiction, jugée (A) visuellement et (B) sur base du nombre de pixels prédits comme étant de l'ergot, permet d'estimer la qualité des acquisitions des images hyperspectrales. Sur cette figure, une surestimation du nombre de pixels d'ergots a été observée à sept reprises.

Un autre critère simple d'évaluation de la qualité des acquisitions consistait à prendre une image hyperspectrale de la céramique blanche utilisée pour la calibration de la réflectance relative et d'observer l'aspect de cette image lorsqu'elle était ouverte avec le logiciel de traitement d'image HyperSee. La présence de bandes parallèles sur l'image était liée à des spectres très bruyants (Figure 23).



Figure 22: Standard paille. Échantillon de référence constitué de sept fragments de pailles de froment d'hiver et utilisé pour suivre la stabilité des mesures réalisées avec le système d'acquisition d'images hyperspectrales proches infrarouges et estimer la qualité de ces images.

3.2.L'acquisition des images hyperspectrales

Les images hyperspectrales proche infrarouge ont été acquises en plaçant les échantillons sur la bande transporteuse passant sous la caméra. Dans le cadre de la calibration et de la validation des modèles permettant la discrimination de racines et d'autres éléments présents au sein des échantillons de sol ainsi que leur quantification, ce sont les échantillons séchés et triés qui ont été passés sous la caméra hyperspectrale. Par la suite, ce sont les échantillons séchés et non triés qui ont été analysés.

Les échantillons étaient disposés sur la bande transporteuse de telle manière à ne pas dépasser de la zone d'acquisition de la caméra (c.-à-d. une bande d'une largeur de 10 cm au centre de la bande transporteuse) et afin de limiter tant que possible le recouvrement des éléments entre eux. En effet, le nombre de pixels de racines détectés sur une image dépend de la manière dont l'échantillon est disposé sur la bande transporteuse du système d'acquisition. Les erreurs de quantification des racines dues à leur recouvrement seront présentées dans la discussion générale (Chapitre VIII).

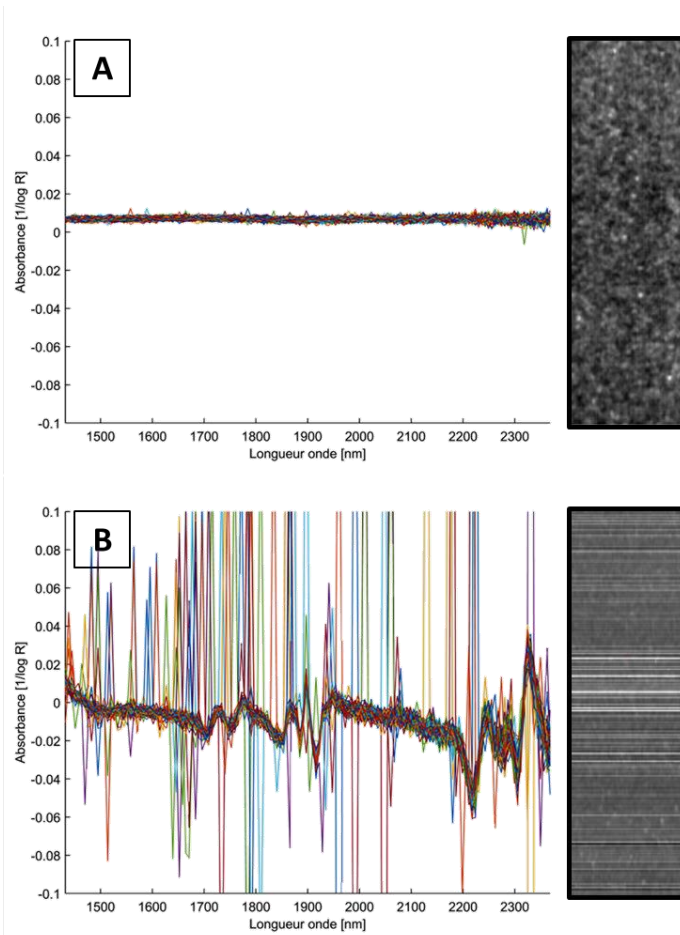


Figure 23: Aspect du profil spectral et prédiction en nuances de gris de l'image hyperspectrale de la céramique blanche lorsque le système d'acquisition est stable (A) ou instable (B). L'instabilité du système d'acquisition crée de fortes variations d'absorbance et du bruit au niveau des spectres. Ce bruit se matérialise sous forme de bandes parallèles au sens d'avancement de la bande transporteuse sur les prédictions des images hyperspectrales de la céramique blanche. Les deux figures des profils spectraux sont à la même échelle.

Pour l'acquisition des spectres destinés à la calibration des modèles de discrimination, les échantillons ont été acquis en les posant directement sur la bande transporteuse. Pour les autres échantillons, au sein desquels la quantité de racines a été mesurée, ils furent dans un premier temps acquis avec le tamis sur lequel ils avaient été récupérés et séchés. L'acquisition conjointe de l'échantillon et du tamis permettait une manipulation plus aisée des échantillons. Néanmoins, la présence du tamis lors de l'acquisition des images posait deux problèmes lors de la quantification des racines. D'une part, le tamis masquait une série de petites racines se trouvant sur

la bande transporteuse et sous le tamis. Ces racines de petite taille ou fragmentées passaient sous le tamis lors de l'étalage des échantillons juste avant leur acquisition¹⁹. D'autre part, une confusion entre les racines et le tamis entraînait une surestimation de la masse de racines présentes dans l'échantillon. Une analyse²⁰ réalisée sur trois échantillons acquis avec et sans le tamis a clairement mis cette surestimation en évidence (Figure 24). Il a donc été décidé à partir de février 2018 de ne plus faire l'acquisition des échantillons avec le tamis et de transférer l'entièreté de l'échantillon sur la bande transporteuse avant l'acquisition des images hyperspectrales.

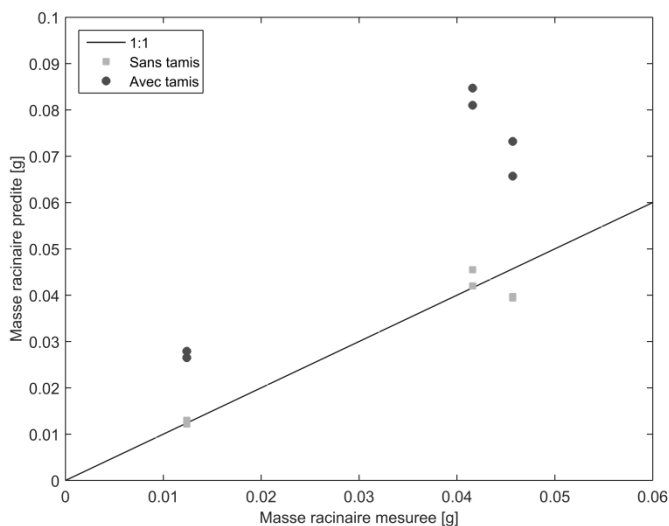


Figure 24: Comparaison des masses de racine (g) réelles et prédites sur des images hyperspectrales. La comparaison est basée sur trois échantillons de racines posés (points noirs) ou non (carrés gris) sur un tamis lors de l'acquisition des images hyperspectrales. Quatre images hyperspectrales ont été acquises pour chaque échantillon (deux avec le tamis et deux sans le tamis). Cette figure met en évidence la surestimation de la masse de racines prédite sur les images hyperspectrales due à la présence du tamis et à la confusion des spectres du tamis et des racines.

¹⁹ Une estimation du nombre de pixels de racines non détectés suite à la présence du tamis a été réalisée dans le cadre du travail de fin d'études de Guillaume Fraipont (année académique 2014-2015, GxABT). Sur base des images acquises sur trois échantillons de racines, nous avons pu déterminer que les racines positionnées sous le tamis et cachées par celui-ci représentaient 5 à 7 % des pixels de racines.

²⁰ Analyse réalisée dans le cadre du travail de fin d'études d'Adrien Boufflette (année académique 2017-2018, GxABT)

3.3. L'extraction des spectres proche infrarouge

L'extraction de spectres proche infrarouges d'images hyperspectrales permet d'obtenir les spectres nécessaires à la calibration et à la validation des modèles chimiométriques. Les spectres extraits étant par la suite utilisés comme référence pour construire et tester les modèles, une attention toute particulière doit être portée dans l'identification des pixels desquels les spectres sont extraits. Il faut en effet extraire des spectres acquis sur des éléments de nature parfaitement connue (des racines par exemple) et éviter de prendre des spectres dans des zones de bordure où la nature des spectres est plus incertaine. Afin de réduire les risques de confusion, les images ayant servi à l'extraction des spectres ne contenaient que des éléments triés et de nature connue. Par exemple, les spectres de racines et résidus de culture de froment ont été extraits sur des images séparées, acquises sur des échantillons ne contenant que des racines ou des résidus de cultures préalablement triés. Des spectres de la bande transporteuse et des zones d'ombre ont été extraits séparément sur plusieurs images de racines et de résidus de cultures.

L'extraction des spectres proche infrarouge des images hyperspectrales a été réalisée avec le logiciel HyperSee (BurgerMetrics SIA, Riga, Lettonie). Les bornes de l'histogramme d'intensité de l'absorbance qui est fourni par le logiciel ont été positionnées de manière à délimiter sur l'image des zones d'intérêt dont les spectres ont été extraits.

Pour la suite des analyses, les spectres ont été enregistrés sous forme de matrices comprenant les différents spectres en lignes et les valeurs d'absorbance pour chaque longueur d'onde en colonnes.

3.4. Traitement des spectres

Cette étape de traitement doit être réalisée pour nettoyer ou préparer les spectres en vue de les analyser (Burger & Geladi 2005). Elle a été réalisée sur les matrices de spectres grâce à une procédure développée avec le logiciel Matlab (The MathWorks Inc., Natick, MA, USA). Le prétraitement des spectres et la calibration des modèles chimiométriques ont été réalisés avec la PLS Toolbox 7.8 (Eigenvector Research Inc., Wenatchee, WA, USA).

3.4.1. Retrait des spectres de la bande transporteuse et aberrants

Avant de pouvoir utiliser les spectres pour construire des modèles chimiométriques, il est important d'éliminer les spectres aberrants des matrices de spectres utilisées comme référence. Ces spectres aberrants peuvent être dus à un pixel mort qui n'aurait pas été détecté ou un spectre acquis sur la bande transporteuse par exemple (Figure 25). Dans nos travaux, ces spectres ont été retirés des matrices grâce à un algorithme qui retirait tous les spectres dont valeur d'absorbance était supérieure à 1,5 à la longueur d'onde 1740 nm ou qui contenaient des valeurs d'absorbances plus de huit fois supérieures à la moyenne des différences d'absorbances entre deux longueurs d'onde successives.

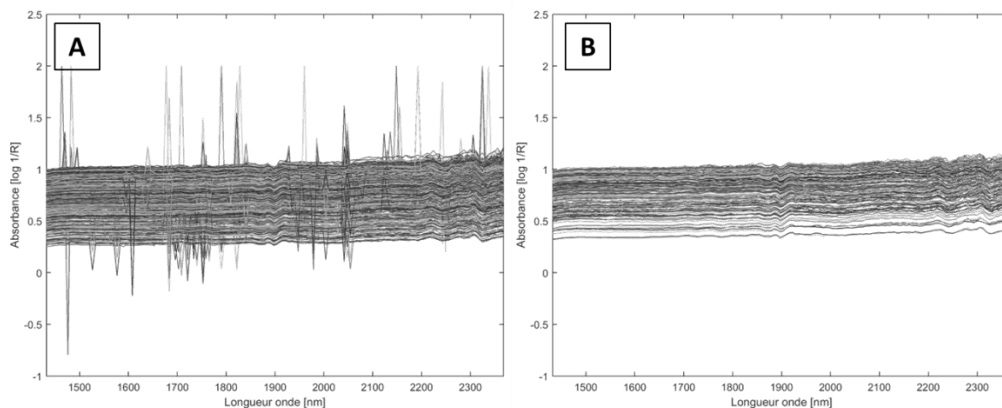


Figure 25: Spectres acquis sur un échantillon de racines de froment. Parmi ces spectres, certains sont aberrants (Figure A). Il s'agit des spectres ayant une valeur d'absorbance soudainement plus élevée. Ces spectres aberrants sont retirés du jeu de données avant la construction des modèles chimiométriques (Figure B).

3.4.2. Réduction du nombre de longueur d'onde des spectres

Les spectres acquis par la caméra hyperspectrale utilisée dans le cadre de cette étude étaient bruyants au début et à la fin du spectre. Les valeurs d'absorbance des longueurs d'onde inférieures à 1432 nm et supérieures à 2368 nm ont donc été éliminées et seules 150 longueurs d'onde restantes sur les 209 acquises initialement ont été utilisées pour la suite des analyses.

3.5. Sélection des sets de calibration et de validation

Selon Agelet et Hurburgh (2010), un set de calibration idéal devrait couvrir l'entièreté des caractéristiques chimiques, physiques et spectrales des échantillons devant être analysés de manière à construire des modèles robustes. Sachant que des études ont montré que le spectre proche infrarouge de racines pouvait être influencé par la date de prélèvement (Lei & Bauhus 2010) et par le milieu de culture (Naumann et al. 2010), les spectres utilisés pour calibrer les modèles de discrimination ont été extraits d'images acquises sur des échantillons très variables en termes de date de prélèvement, de profondeur de prélèvement, d'état de développement et de dégradation. Des spectres provenant d'images hyperspectrales acquises à différentes dates ont également été utilisés afin de tenir compte de l'instabilité du système d'acquisition des images hyperspectrales, instabilité qui peut être perçue sur des échantillons identiques acquis à différentes heures d'une journée d'acquisition ou lors de différentes journées. C'est pourquoi des spectres extraits d'images du standard paille acquis à différents moments ont été inclus dans les sets de calibration et de validation.

Pour la séparation des spectres en deux sets de données, l'un pour la construction des modèles, l'autre pour leur validation, l'algorithme Duplex a été utilisé. Cet

l'algorithme a aussi permis de réduire le nombre de spectres tout en maintenant leur variabilité. L'utilisation de tous ces spectres extraits des images hyperspectrales n'était en effet pas justifiée car ils contenaient des informations redondantes et le temps de calibration des modèles aurait été très long.

Dans le cas de nos travaux, les spectres ont été sélectionnés en nombre égal au sein de la matrice de spectre créée pour chaque image et non au sein d'une matrice regroupant l'ensemble des spectres acquis pour chaque classe de spectre. Cela a été fait afin de garantir la sélection de la plus grande diversité de spectres comme l'illustre la Figure 26.

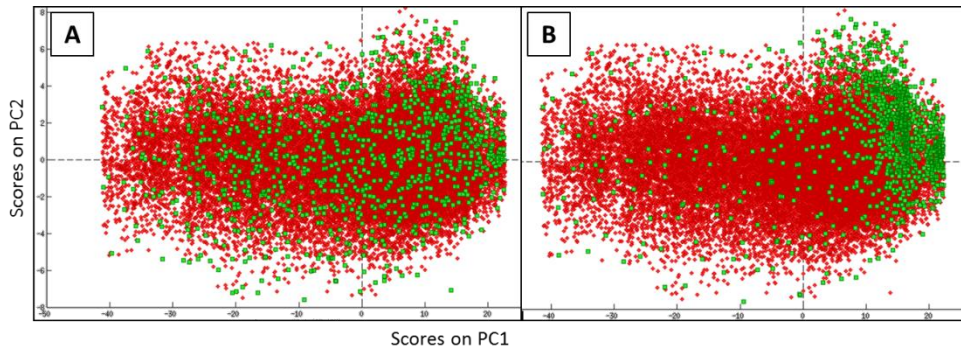


Figure 26: Sélection de 1000 spectres (points verts) réalisée en appliquant l'algorithme Duplex soit (A) sur différentes matrices de spectres et en sélectionnant un nombre égal de spectres au sein de chaque matrice, soit (B) sur une seule matrice rassemblant tous les spectres. L'analyse en composantes principales (PCA) réalisée sur les spectres avant (points rouges) et après la sélection duplex (points verts) montre une meilleure prise en compte de la variabilité des spectres par l'algorithme Duplex lorsque les spectres ne sont pas sélectionnés en une seule fois au sein d'une même matrice. Les axes des graphiques reprennent les résultats obtenus pour la première et la seconde composante principale (PC) de l'analyse, expliquant respectivement 97 % et 2 % de la variabilité des données.

Cette étape de sélection de spectres a été réalisée en deux temps. Dans un premier temps, un certain nombre de spectres ont été sélectionnés pour créer le set de calibration. Vu le nombre de spectres acquis, la sélection n'a pas été réalisée à raison de 50 % des spectres pour le set de calibration et 50 % pour le set de validation. Le nombre de spectres désiré pour le set de calibration a été déterminé et l'algorithme Duplex a été utilisé de telle sorte à sélectionner exactement ce nombre de spectres. Au sein des spectres restants, une seconde sélection a été réalisée pour le set de validation. Cette fois encore, un nombre précis de spectres a été sélectionné. Certains spectres acquis n'ont donc été utilisés ni pour la calibration ni pour la validation. Ce choix a été posé de manière à réduire le temps nécessaire à la calibration et à la validation des modèles.

4. Construction des modèles de discrimination et de quantification

4.1. Prétraitement des spectres

La première étape dans la construction des modèles chimiométriques consiste à choisir le prétraitement qui sera appliqué sur les spectres pour éliminer les variations à haute fréquence ainsi que les effets additifs et multiplicatifs qui sont présents dans les spectres, mais ne sont pas liés à la nature chimique des échantillons.

Le choix du (des) prétraitement(s) est basé sur un processus d'essais-erreurs avec plusieurs prétraitements. Ceux qui ont été utilisés dans le cadre de nos travaux sont la mise à l'échelle automatique (*autoscale*), le lissage avec l'algorithme de Savitzky-Golay ainsi que les traitements SNV et Detrend. Ces prétraitements ont été appliqués sur les jeux de données comprenant uniquement les spectres sélectionnés par la méthode Duplex, dont les spectres aberrants avaient été éliminés et le nombre de longueur d'onde réduit.

4.2. Modèles de discrimination

Les modèles de discrimination (ou classification) permettent le classement de spectres selon la nature des éléments sur lesquels ils ont été acquis. Parmi ces modèles, nous avons utilisé deux algorithmes de classification supervisés différents : l'algorithme PLS-DA et l'algorithme SVM.

Ces deux algorithmes ont été testés pour discriminer des racines de froment, des résidus de cultures, des particules de terre et le tamis sur lequel étaient récupérés les échantillons après leur lavage. Les résultats obtenus ont été publiés dans le *Journal of Chemometrics* et sont présentés au Chapitre V.

L'algorithme SVM ayant donné les meilleurs résultats pour la discrimination de ces éléments, il a été retenu pour les autres travaux menés dans le cadre de cette thèse, dont un travail de groupe²¹ et trois travaux de fin d'études²². Certains résultats de ces travaux sont présentés au Chapitre VI.

4.3. Construction de l'arbre de classification dichotomique

Dans le cas de l'utilisation de modèles de discrimination appliqués sur des échantillons contenant plus de deux classes spectrales, plusieurs modèles de discrimination ont été créés et agencés au sein d'un arbre de classification dichotomique. La méthodologie utilisée était inspirée de ce qui avait déjà été fait pour détecter différents composants au sein d'aliments pour bétail (Fernández Pierna

²¹ Travail réalisé par quatre étudiants de Gembloux Agro-Bio Tech, Université de Liège (année académique 2017-2018): A.J.P. Attikora, F. Boucheron, X. Bordet et F. Godin.

²² Travaux de fin d'études réalisés par des étudiants de Gembloux Agro-Bio Tech, Université de Liège : Guillaume Fraipont (Année académique 2014-2015), Adrien Boufflette et Louis Dumont de Chassart (Année académique 2017-2018)

et al. 2006) ou pour détecter des substances indésirables dans des aliments (Fernández Pierna et al. 2012).

À chaque étage de l'arbre, les spectres étaient classés en deux classes spectrales par un algorithme de classification binaire. Au sein de l'arbre illustré à la Figure 27, trois modèles de discrimination permettaient de classer successivement les spectres de l'arrière-plan, de particules de sol, de racines et de résidus de cultures extraits d'une image hyperspectrale.

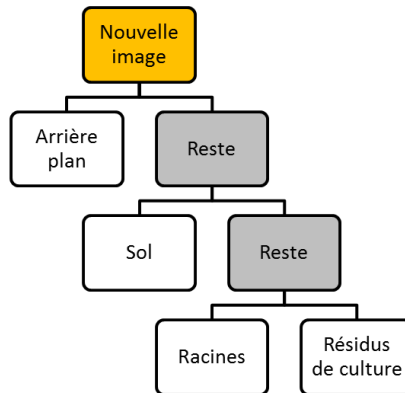


Figure 27: Arbre de classification dichotomique. Dans cet exemple, trois modèles de discrimination organisés en cascade permettent de classer successivement les spectres de l'arrière-plan, de particules de sol, de racines et de résidus de cultures extraits d'une image hyperspectrale.

4.4. Prédications sur images

En utilisant les modèles chimiométriques de discrimination et l'arbre de classification dichotomique, le spectre proche infrarouge de chaque pixel de l'image hyperspectrale a pu être classé. En assignant une couleur à chaque classe spectrale, il a ensuite été possible de reconstruire l'image en faisant apparaître les éléments selon leur nature. Chaque élément pouvait donc être visualisé sur l'image reconstruite. Un exemple de prédiction est donné à la Figure 28 pour un échantillon contenant des racines de froment d'hiver, des résidus de la culture précédente (également du froment d'hiver) et des particules de sol. Cet échantillon était posé sur un tamis lui-même posé sur la bande transporteuse du système d'acquisition.

4.5. Modèles de quantification

4.5.1. Quantification sur base d'un nombre de pixels

Sur base des prédictions des images hyperspectrales et du nombre de pixels classés en tant que racines pour un échantillon, il a été possible de quantifier la masse de racines au sein de cet échantillon. Ce type de relation entre un nombre de pixels prédits comme étant un élément sur une image hyperspectrale et la masse de cet

élément, dans notre cas la masse de racines, avait déjà été utilisée par Vermeulen et al. (2012) pour quantifier la masse d'ergot au sein d'échantillons de céréales.

Pour construire cette relation linéaire entre le nombre de pixels de racines et la masse de racines dans l'échantillon, des images hyperspectrales ont été acquises sur des échantillons de masse connue ne contenant que des racines. Le nombre de pixels prédits par les modèles de discrimination comme étant des racines sur ces images a ensuite été mis en relation avec la masse des échantillons.

La construction de droites de régression permettant la quantification de racines de froment d'hiver et de pois d'hiver ainsi que de résidus de culture est décrite au Chapitre VI. Des exemples d'utilisation y sont également présentés.

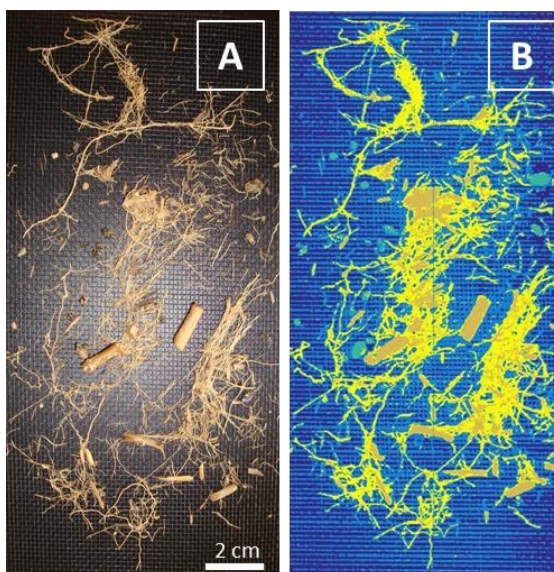


Figure 28: Échantillon de racine contenant des résidus de cultures, des particules de terre et des cailloux. L'échantillon est déposé sur un tamis. L'image de gauche (A) est une photographie de l'échantillon. L'image de droite (B) est une image hyperspectrale dont la nature du spectre extrait de chaque pixel a été prédite en utilisant des modèles chimiométriques de discrimination. Sur base de ces prédictions, une couleur a été assignée à chaque classe spectrale : les racines sont en jaune, les résidus de cultures en ocre, les éléments du sol en turquoise et l'arrière-plan en bleu.

4.5.2. Quantification sur base d'un profil spectral

Les modèles chimiométriques de quantification permettent de prédire la teneur d'un composé sur base de l'allure d'un spectre. La construction de tels modèles est basée sur des valeurs de références pour la teneur en un composé et sur des spectres acquis sur les mêmes échantillons (ou des échantillons apparentés). La teneur du composé étudié dans les échantillons utilisés comme références peut être déterminée par chimie analytique, ou par toute autre méthode de référence, ou être synthétisée

ce qui permet de créer des échantillons contenant des valeurs connues du composé (Cougnon et al. 2013).

L'algorithme de quantification PLS a été utilisé pour estimer la teneur en leghémoglobine, une molécule intervenant dans la fixation symbiotique de l'azote atmosphérique par les légumineuses, au sein de nodosités de pois (*Pisum sativum* L.). Les résultats de cette étude, menée en partie dans le cadre d'un travail de groupe d'étudiants de deuxième master²³, ont été publiés dans le *Journal of Spectral Imaging* et sont présentés au Chapitre VII. Dans le cadre de cette étude, la valeur de référence avait été déterminée par spectrophotométrie et par comparaison avec des solutions contenant des teneurs connues en hémoglobine humaine.

5. Validation

L'étape de validation est importante de manière à tester la qualité des modèles construits. Cette étape consiste à appliquer un modèle chimiométrique (qu'il soit de discrimination ou de quantification) sur des échantillons ayant des propriétés connues et d'observer comment ces propriétés sont prédites par le modèle. Elle peut être réalisée soit sur des échantillons autres que ceux utilisés pour la construction des modèles, soit sur les mêmes échantillons que ceux utilisés pour construire le modèle. C'est le cas de la validation croisée.

Tous les modèles développés dans cette étude ont été validés. Selon la nature des modèles, la validation a été réalisée de manières différentes. Pour les modèles de discrimination, ceux-ci ont d'abord été validés de manière croisée sur les spectres ayant servi à leur construction (validation interne). Cette validation a été réalisée automatiquement par le logiciel PLS Toolbox utilisé pour construire ces modèles.

Une seconde validation a été réalisée sur une sélection de spectres différents de ceux utilisés pour la construction des modèles grâce au set de validation sélectionné par l'algorithme Duplex. Les spectres utilisés pour cette seconde validation provenaient des mêmes images hyperspectrales que ceux utilisés pour la calibration des modèles.

Enfin, une fois combinés dans l'arbre de classification dichotomique, les modèles chimiométriques ont été appliqués sur des images ne contenant que des éléments appartenant à une classe spectrale unique. Le pourcentage de pixels correctement prédit a ensuite été calculé. Cette approche avait aussi été utilisée par Fernández Pierna et al. (2004). Cette troisième validation a été réalisée sur les images utilisées pour calibrer les modèles et sur de nouvelles images.

Les méthodes de validation et les résultats obtenus pour la discrimination des racines, des résidus de cultures, des particules de sol et du tamis sont présentés au Chapitre V

²³ Doudet Ariel, Druart Florent, Machiels Laëtitia, Sougnez Audrey, Torfs Marine. "Etude des potentialités de l'imagerie hyperspectrale proche infrarouge pour la mesure de la teneur en leghémoglobine au sein de nodosités de pois protéagineux (*Pisum sativum* L.)". Gembloux Agro-Bio Tech, Université de Liège. Année académique 2016-2017.

Les modèles de quantification, qu'ils soient basés sur un nombre de pixels détectés comme appartenant à une classe spectrale ou sur l'allure d'un spectre, ont également été validés. Ces validations sont présentées aux Chapitres VI et VII respectivement.

CHAPITRE IV

Suivi de la stabilité du système d'acquisition des images hyperspectrales proche infrarouge

"Accurate results require accurate spectra"

Burger et Geladi (2005)

1. Introduction

Un instrument de mesure tel qu'un spectrographe n'est jamais parfaitement stable. Il enregistre donc des valeurs légèrement différentes d'une acquisition à l'autre réalisée sur un même objet. Ces variations sont dues aux nombreux facteurs qui influencent l'instrument de mesure tels que la pression atmosphérique, l'humidité et la température ambiante et le voltage des circuits électriques par exemple (Wold 1995). D'autres facteurs tels que le système d'éclairage utilisé peuvent entraîner des variations au niveau des spectres. Les lampes halogènes, par exemple, ont le désavantage de créer des variations au sein du profil des spectres à cause de la température qu'elles produisent et d'avoir une production lumineuse instable à cause des fluctuations de voltage (Wu & Sun 2013).

Afin d'assurer la pertinence des données enregistrées avec un système d'acquisition d'imagerie hyperspectrale proche infrarouge, des efforts de calibration réguliers et continus sont donc nécessaires. Car, comme le soulignent Burger et Geladi (2005), "*accurate results require accurate spectra*". Cela demande donc un plan robuste de calibration du système d'acquisition de manière à assurer une continuité de mesure en mesure et de jour en jour dans les données acquises (Eylenbosch et al. 2016). Ce n'est que sous cette condition que l'imagerie hyperspectrale proche infrarouge peut être considérée comme une méthode rapide et fiable pour quantifier et discriminer différents éléments au sein d'échantillons complexes.

Au cours du suivi de la qualité des prédictions de la nature des pixels des images hyperspectrales acquises sur un échantillon de référence dans le cadre d'une étude sur la quantification de racines²⁴, une instabilité du système d'acquisition a été observée. Selon le jour et l'heure auxquels les images étaient acquises sur cet échantillon de référence constitué uniquement de pailles de froment, la quantité de pixels prédits comme étant des pailles par les modèles de discrimination était fortement variable (Figure 29). Le modèle de discrimination utilisé permettait de faire la discrimination entre des spectres de pailles et de racines et, pour 82 % des images, moins de 75 % des pixels étaient correctement prédits comme étant des pailles. Il avait été observé lors de cette étude que les images semblaient être mieux prédites après la calibration du système d'acquisition (calibration de la réflectance et retrait des pixels défectueux).

Selon Burger et Geladi (2005), trois situations peuvent rendre invalides les modèles construits : (i) des modifications chimiques ou physiques au sein des échantillons tels que la texture ou la taille des particules, (ii) des modifications du système d'acquisition des images hyperspectrales causées par à une instabilité inhérente, le vieillissement des composants ou l'usage de différents systèmes d'acquisition ou (iii) des modifications de l'environnement telles que l'humidité ou

²⁴ La dérive du système d'acquisition d'images hyperspectrales a été observée au cours du Travail de fin d'études de Guillaume Fraipont (année académique 2014-2015, Gembloux Agro-Bio Tech, Université de Liège).

la température. Les modifications spectrales qui en découlent peuvent être observées soit sur l'axe horizontal des spectres avec un décalage des longueurs d'onde ou un élargissement des pics, soit sur l'axe vertical avec un décalage de la ligne de base des spectres ou des différences d'absorbance.

Les variations observées lors des prédictions sur l'échantillon de référence nous semblant plutôt être dues à une instabilité inhérente du système d'acquisition, une étude a donc été réalisée afin de caractériser cette instabilité et tenter d'y trouver une solution.

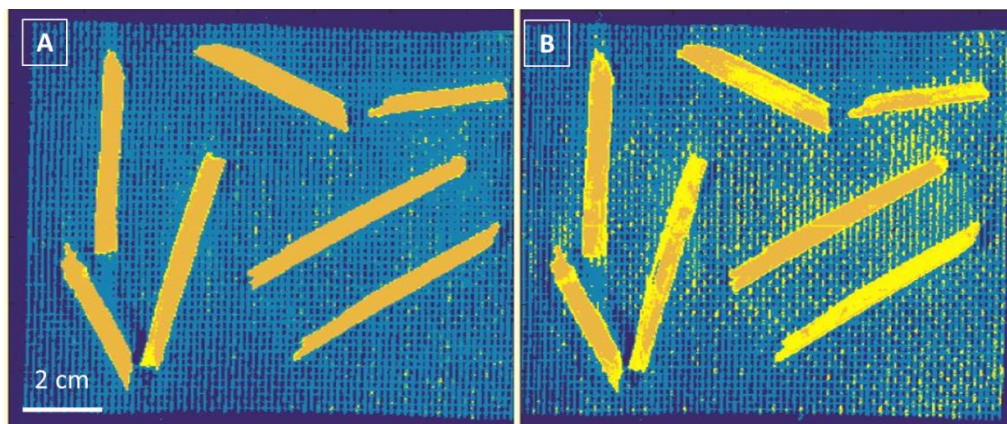


Figure 29: Prédications de la nature des pixels sur deux images hyperspectrales proche infrarouge acquises sur un même échantillon de référence ne contenant que des pailles. Les prédictions ont été réalisées grâce à l'arbre de classification dichotomique constitué de modèles de discrimination SVM (*Support Vector Machine*) permettant de classer les spectres de l'arrière-plan (bleu), des particules de sol (absentes sur ces images), des racines (jaune) et des résidus de cultures (ocre). Sur cet échantillon ne contenant que des pailles, les pixels en jaune sont des pixels mal classés. La comparaison des deux prédictions met clairement en évidence une classification variable des pixels selon le moment où l'image a été acquise. Sur l'image de gauche (A), 90 % des pixels de paille sont bien classés alors que sur l'image de droite (B), seulement 39 % de ces pixels sont bien classés.

2. Matériels et méthodes

Le suivi de la stabilité du système d'acquisition d'images hyperspectrales proche infrarouge a été réalisé sur base d'un suivi régulier du signal spectral acquis sur deux références et en faisant varier la fréquence de calibration du système d'acquisition au cours de neuf journée de mesure.

La première référence était une céramique blanche utilisée pour faire la calibration du système d'acquisition et pour calculer la réflectance relative. Sur base d'acquisitions répétées d'images hyperspectrales sur cette surface de référence, nous avons réalisé le suivi de l'évolution de son profil spectral moyen. Plus précisément,

nous avons suivi l'évolution de la valeur d'absorbance minimale de ce spectre moyen.

La seconde référence était l'échantillon de paille (standard paille) sur lequel l'instabilité du système d'acquisition avait été observée au cours de nos travaux. Sur cet échantillon de référence, l'évolution du pourcentage de pixels correctement prédit comme étant de la paille a été suivie. Ce suivi a été réalisé en appliquant sur les images acquises l'arbre de classification dichotomique permettant de discriminer l'arrière-plan, les particules de sol, les racines et les résidus de cultures. Ce pourcentage de pixels correctement prédits a été calculé en divisant le nombre de pixels prédits comme étant des pailles par la somme des pixels prédits comme étant du sol, des racines ou des pailles (résidus de cultures).

L'évolution des valeurs d'absorbance de quatre longueurs d'onde (1458, 1941, 2092 et 2350 nm) a également été suivie pour le standard paille. Ces longueurs d'onde ont été choisies, sur base d'un test de Fisher, pour leur importance dans la discrimination des racines et des résidus de cultures. Les valeurs d'absorbance de ces quatre longueurs d'ondes ont été obtenues sur le spectre moyen calculé sur base des trois répétitions des mesures. Un spectre moyen a été calculé, pour chaque fréquence de calibration du système d'acquisition, pour chaque acquisition de la journée.

Le suivi du profil spectral moyen de la céramique blanche et du pourcentage de pixels de pailles a été réalisé après trois fréquences de calibration du système d'acquisition : une calibration en début de journée et une seconde après 30 images, une calibration toutes les deux heures ou une calibration toutes les 30 minutes. Chaque fréquence de calibration était étudiée séparément au cours de journées entières. La fréquence d'acquisition des images hyperspectrales sur la céramique blanche et le standard paille était d'une image toutes les 10 minutes. Ce protocole a été répété trois fois au cours de trois semaines consécutives. L'ordre des fréquences de calibration au sein de chaque semaine était déterminé aléatoirement.

Quelle que soit la fréquence de calibration, la première calibration était réalisée 1h30 après l'allumage du système d'acquisition (caméra hyperspectrale et lampes). La première image était acquise 30 minutes après cette première calibration. Dans le cas d'une calibration toutes les 30 minutes, le système d'acquisition était donc calibré deux fois avant l'acquisition de la première image. Entre chaque journée d'acquisition, le système d'acquisition était mis à l'arrêt.

3. Résultats et discussion

3.1. Suivi de l'absorbance des échantillons de référence en fonction de la fréquence de calibration

Les résultats ont montré une décroissance systématique au cours du temps de la valeur minimale du spectre moyen de la céramique blanche, cette valeur devenant de plus en plus négative. Cette tendance était majoritairement observée en début de journée et de nouvelles calibrations permettaient de réduire cette chute et de stabiliser la valeur minimale de l'absorbance moyenne de la céramique (Figure 30).

Lorsque la calibration du système d'acquisition était réalisée une fois en début de journée d'acquisition et une seconde fois après 30 acquisitions, la valeur minimale de l'absorbance moyenne de la céramique blanche était d'abord négative puis se stabilisait autour d'une valeur d'absorbance proche de zéro après la seconde calibration.

Lorsque la calibration du système d'acquisition était réalisée toutes les deux heures, la décroissance de la valeur minimale du spectre moyen de la céramique blanche diminuait aussi fortement que dans le cas précédent après la première calibration et durant les deux premières heures de mesure. Elle se rapprochait aussi de zéro après la seconde calibration. Ces valeurs n'étaient cependant pas directement stabilisées et une nouvelle décroissance était observée entre la deuxième et la troisième calibration du système d'acquisition. Ce n'était qu'après cette troisième calibration, 5h30 après l'allumage de la caméra, que les valeurs enregistrées par le système d'acquisition se stabilisaient.

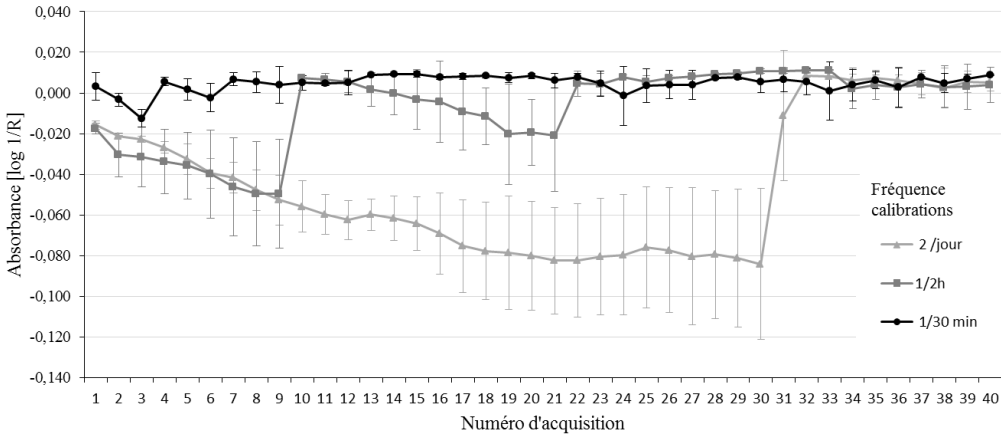


Figure 30: Évolution de la valeur minimale de l'absorbance moyenne d'une céramique blanche au cours d'une journée d'acquisition et selon la fréquence de calibration de la caméra hyperspectrale (une fois en début de journée et une fois après 30 images, toutes les deux heures ou toutes les 30 minutes). Plus la fréquence de calibration du système d'acquisition des images hyperspectrales est élevée, plus la valeur de l'absorbance moyenne de la céramique est stable. Les barres d'erreur représentent la variabilité observée au cours des trois journées de répétition des mesures pour chaque fréquence de calibration.

Un suivi des valeurs d'absorbance enregistrées pour quatre longueurs d'onde du spectre moyen des pailles de l'échantillon de référence a également été réalisé. Ce suivi a montré une évolution équivalente à celle observée pour la valeur d'absorbance minimale de la céramique pour chaque fréquence de calibration avec une amplitude de variation qui était fonction de la longueur d'onde considérée (Figure 31).

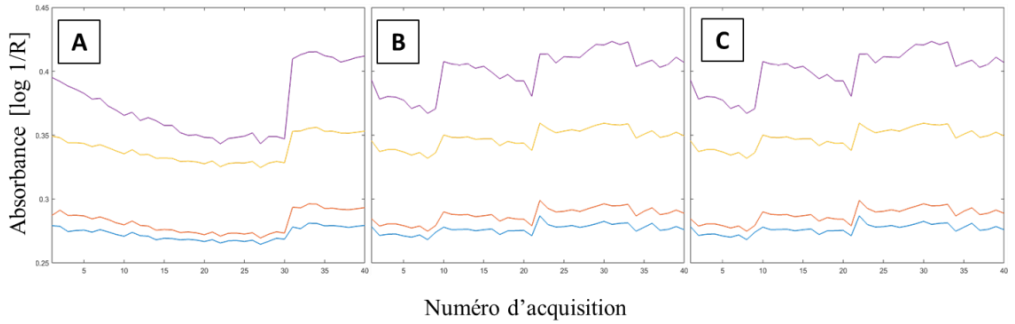


Figure 31: Evolution au cours des acquisitions de la valeur d'absorbance pour quatre longueurs d'onde des spectres moyens acquis sur le standard paille et pour trois fréquences de calibration du système d'acquisition: (A) deux fois par jour, (B) toutes les deux heures ou (C) toutes les 30 minutes. Ces quatre longueurs d'onde ont été choisies sur base d'un test de Fisher pour leur importance dans la discrimination des racines et des résidus de cultures (de bas en haut : 1458, 1941, 2092 et 2350 nm).

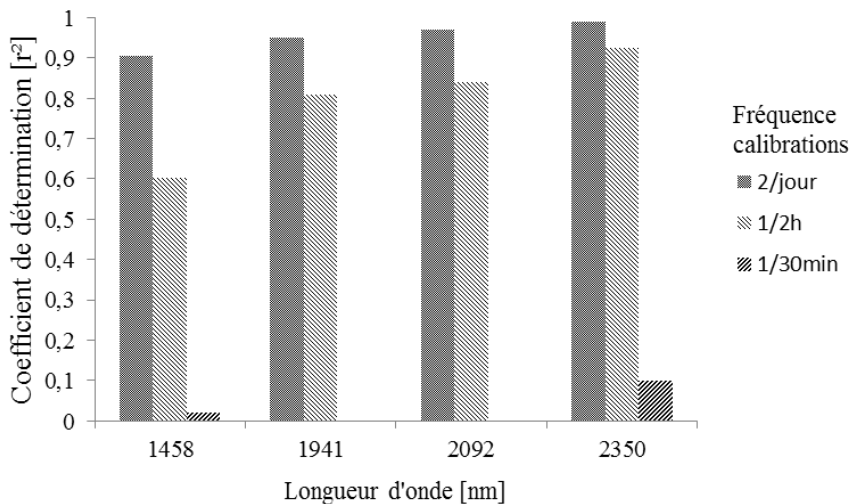


Figure 32: Valeurs du coefficient de détermination (r^2) de la relation établie entre la valeur minimale de l'absorbance moyenne de la céramique blanche et la valeur d'absorbance de quatre longueurs d'onde du spectre moyen des pailles de l'échantillon de référence. Le coefficient de détermination diminue lorsque la fréquence de calibration augmente.

L'étude de la relation liant la valeur de l'absorbance moyenne minimale de la céramique blanche aux valeurs d'absorbance mesurées sur le standard paille a montré une forte corrélation entre ces valeurs lorsque le système d'acquisition était calibré peu fréquemment (Figure 32). En effet, lorsque le système d'acquisition n'était calibré que deux fois sur la journée, le coefficient de détermination (r^2) était

supérieur à 0,9 pour les quatre longueurs d'onde considérées sur les pailles. Si la fréquence de calibration passait à une calibration toutes les deux heures, une corrélation était toujours observée entre la valeur minimale du spectre moyen de la céramique blanche et les valeurs d'absorbance obtenues sur les pailles de l'échantillon de référence ($r^2 > 0,6$). Lorsque le système d'acquisition était calibré très régulièrement, les variations observées au niveau des valeurs d'absorbances des pailles ne pouvaient plus être expliquées par les variations d'absorbance enregistrées au niveau de la céramique blanche ($r^2 < 0,1$).

3.2. Suivi de l'évolution des prédictions sur le standard paille en fonction de la fréquence de calibration

Avec la fréquence de calibration la moins intensive, le pourcentage de pixels correctement prédits comme étant des pailles suivait le même profil d'évolution que celui observé avec la valeur minimale du spectre moyen de la céramique blanche : les prédictions étaient bonnes pour les images acquises en début de journée, se dégradaient au fur et à mesure des acquisitions et se stabilisaient, autour d'une valeur nettement améliorée, dans deux cas sur trois, suite à la deuxième calibration réalisée après la 30^e acquisition (Figure 33).

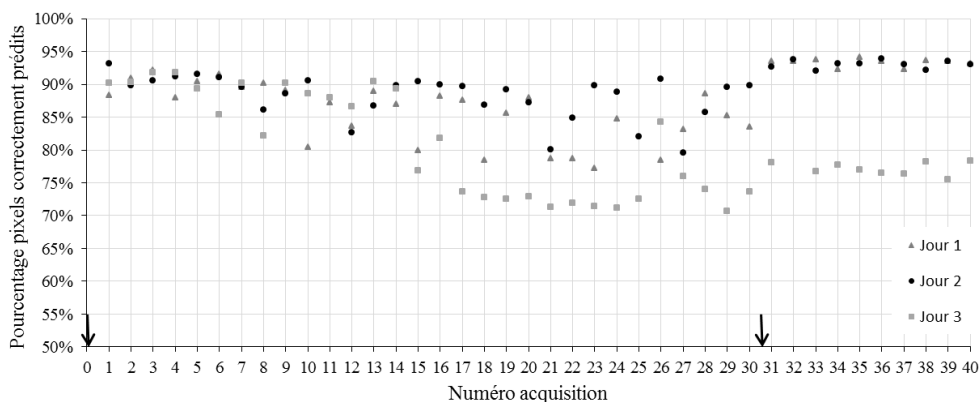


Figure 33: Évolution du pourcentage de pixels de pailles correctement prédits par les modèles de discrimination lorsque la calibration du système d'acquisition est réalisée en début de journée et après la 30^e image. Les acquisitions ont été répétées sur trois journées d'acquisition. Les flèches verticales indiquent quand ont été réalisées les calibrations du système d'acquisition.

Avec une calibration toutes les deux heures, le pourcentage de pixels correctement prédits était globalement plus élevé et plus stable qu'avec une calibration en début de journée et une en fin de journée, après la 30^e acquisition (Figure 34). Des résultats plus faibles ont cependant été observés pour une des trois journées d'acquisition avec cette fréquence de calibration sans qu'aucune explication n'ait pu être donnée.

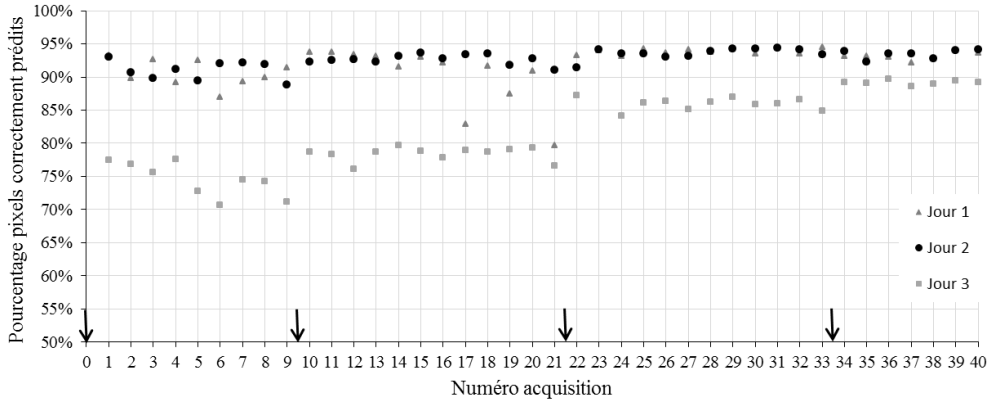


Figure 34: Évolution du pourcentage de pixels de pailles correctement prédits par les modèles de discrimination lorsque la calibration du système d’acquisition est réalisée toutes les deux heures. Les acquisitions ont été répétées sur trois journées d’acquisition. Les flèches verticales indiquent quand ont été réalisées les calibrations du système d’acquisition.

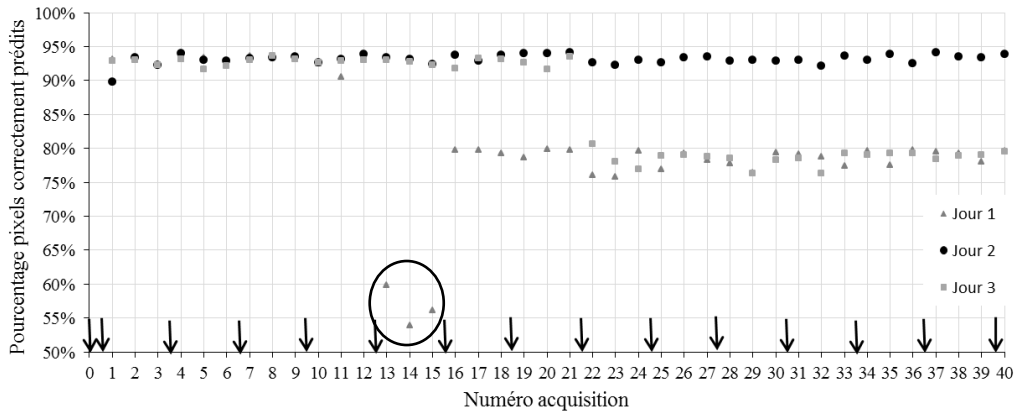


Figure 35: Évolution du pourcentage de pixels de pailles correctement prédits par les modèles de discrimination lorsque la calibration du système d’acquisition est réalisée toutes les 30 minutes. Les acquisitions ont été répétées sur trois journées d’acquisition. Les flèches verticales indiquent quand ont été réalisées les calibrations du système d’acquisition. Les points encerclés correspondent aux images pour lesquelles moins de 60 % des pixels de paille étaient correctement prédits.

Calibrer le système d’acquisition toutes les 30 minutes permettait de stabiliser rapidement les valeurs d’absorbance enregistrées sur la céramique blanche. Cette fréquence de calibration a aussi globalement permis de stabiliser les prédictions obtenues sur le standard paille bien qu’une chute du pourcentage de pixels correctement prédits ait été observée lors de deux journées d’acquisition (Figure 35).

À nouveau, aucune explication n'a pu être apportée à cette perte de qualité des prédictions. Notons que dans certains cas, moins de 60 % des pixels de paille étaient bien prédits (Figure 35, images 13, 14 et 15 de la première journée d'acquisition). Une nouvelle calibration du système a permis d'améliorer fortement le signal enregistré sur l'échantillon et les prédictions réalisées dessus.

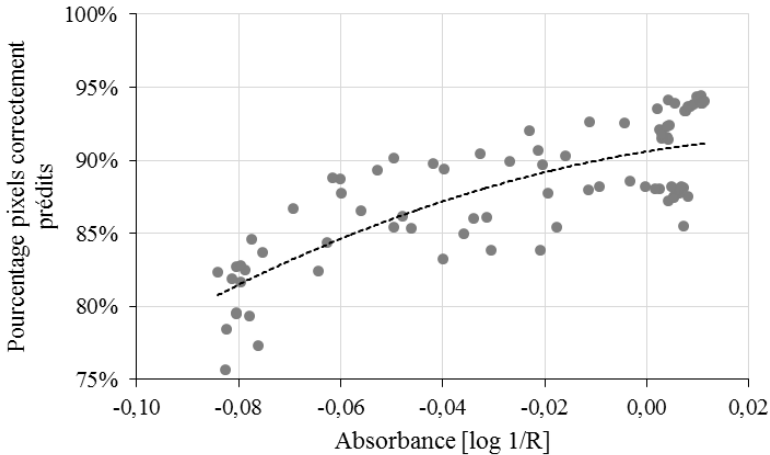


Figure 36: Relation entre la valeur minimale de l'absorbance moyenne de la céramique blanche et le pourcentage de pixels de paille correctement prédits sur les images hyperspectrales du standard paille lorsque le système d'acquisition est calibré deux fois par jour ou toutes les deux heures (valeurs combinées). L'équation de la relation polynomiale (ligne pointillée) permet d'estimer la qualité des prédictions sur base de la valeur minimale de l'absorbance moyenne de la céramique ($r^2 = 0.63$).

3.3. Relation entre l'absorbance minimale de la céramique blanche et le pourcentage de pixels de paille correctement prédits

Sur base de la valeur minimale de l'absorbance moyenne de la céramique blanche et du pourcentage de pixels de paille correctement prédits sur les images du standard paille, nous avons pu calculer la corrélation entre ces deux données. Les résultats ont montré que lorsque le système d'acquisition était instable, c'est-à-dire lorsque les calibrations n'étaient effectuées que deux fois par jour ou toutes les deux heures, la valeur minimale de l'absorbance moyenne de la céramique blanche permettait de donner une estimation de la qualité de la prédiction réalisée sur le standard paille. Au plus la valeur minimale de l'absorbance moyenne de la céramique blanche était négative, au plus la probabilité d'obtenir de bonnes prédictions sur le standard paille était faible. Les meilleures prédictions étaient obtenues lorsque la valeur minimale de l'absorbance moyenne de la céramique blanche était proche de zéro (Figure 36).

Lorsque le système d'acquisition était calibré toutes les 30 minutes, les variations enregistrées au niveau des prédictions du standard paille étaient plus importantes que celles observées sur le spectre moyen de la céramique blanche et aucun lien n'existait entre les données. Le suivi de l'absorbance minimale permettait donc d'estimer la stabilité du système d'acquisition sans pouvoir prédire la qualité des prédictions réalisées sur les images enregistrées.

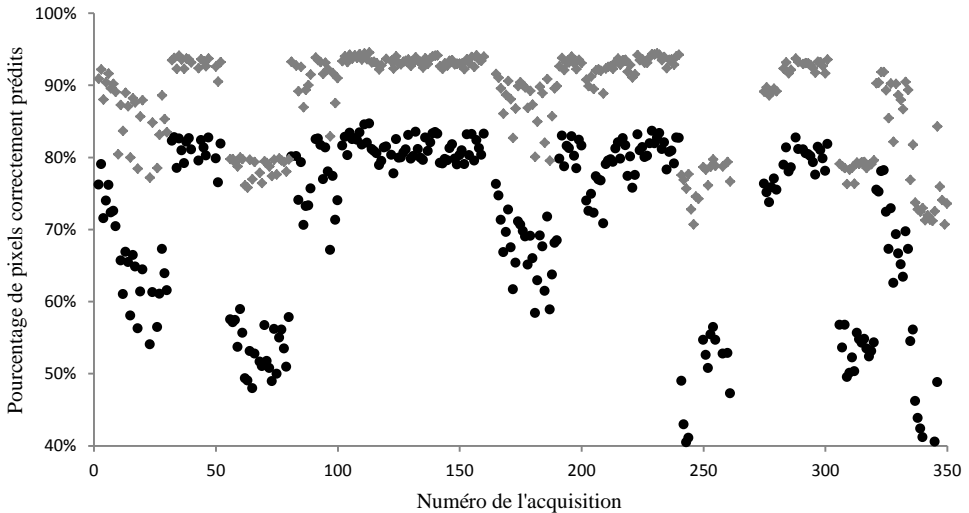


Figure 37: Comparaison des résultats des prédictions obtenues avec deux modèles de discrimination sur l'ensemble des images hyperspectrales acquises sur le standard paille au cours de la stabilité du système d'acquisition (neuf journées d'acquisitions). Le Modèle 1 (points noirs), contrairement au Modèle 2 (losanges gris), fut calibré sans tenir compte de l'instabilité du système d'acquisition des images hyperspectrales et des zones d'ombre présentes sur les images.

3.4.Prise en compte de la variabilité spectrale dans le set de calibration des modèles de discrimination

La prise en compte d'un maximum de variabilité spectrale au sein du set de calibration des modèles chimiométriques est importante afin d'obtenir des modèles robustes. Suite à l'observation de l'instabilité du système d'acquisition et de l'évolution des prédictions, des spectres extraits d'images hyperspectrales acquises sur le standard paille à plusieurs moments au cours des journées d'acquisition ont été ajoutés dans le set de calibration des modèles de discrimination. Lors de cette étude, les spectres des zones d'ombres en bordure des racines et des résidus de culture ont également été pris en compte et ont été classés avec l'arrière-plan. Le résultat de cette prise en compte d'une plus grande variabilité spectrale fut une meilleure prédiction des pixels de résidus de cultures.

La Figure 37 compare les résultats des prédictions obtenues avec deux modèles de discrimination SVM sur l'ensemble des images hyperspectrales acquises sur le standard paille au cours du suivi de l'instabilité du système d'acquisition. Le premier modèle (Modèle 1, points noirs), qui donne les résultats les plus variables, fut calibré sans tenir compte de l'instabilité du système d'acquisition des images hyperspectrales et des zones d'ombre présentes sur les images. Le second modèle de discrimination (Modèle 2, losanges gris) prenait en compte ces deux problèmes rencontrés lors de la prédiction des images hyperspectrales en incluant dans son set de calibration des spectres acquis sur de nombreuses images du standard paille et des spectres d'ombres. La prise en compte de l'instabilité a permis de diminuer la variabilité des prédictions alors que la prise en compte des zones d'ombre a permis d'augmenter la part de pixels de paille correctement prédits par le modèle de discrimination. Le pourcentage moyen (\pm écart-type) de pixels de paille correctement prédit sur le standard paille est ainsi passé de $69 \pm 14 \%$ à $88 \pm 8 \%$.

4. Conclusions et perspectives

Bien que cette étude n'ait pas pu apporter toutes les solutions pour stabiliser les valeurs obtenues lors de la prédiction de la nature des pixels d'images hyperspectrales acquises sur un échantillon de référence, elle fut néanmoins riche en enseignements.

Tout d'abord, elle a montré qu'une augmentation de la fréquence de calibration du système d'acquisition permettaient de stabiliser la valeur de l'absorbance acquise sur la céramique blanche utilisée pour calibrer le système d'acquisition et calculer la valeur de l'absorbance relative. Ces calibrations régulières permettaient également de stabiliser les valeurs d'absorbance et d'améliorer globalement la prédiction de la nature des pixels des images hyperspectrales acquises sur l'échantillon de référence (échantillon contenant uniquement des pailles de froment d'hiver).

Deuxièmement, les résultats obtenus montrent l'intérêt de suivre l'absorbance moyenne de la céramique blanche au cours de la journée d'acquisition. La valeur minimale de l'absorbance moyenne de la céramique blanche étant positivement corrélée aux valeurs obtenues lors de la prédiction de la nature des pixels de l'échantillon de référence, le spectre moyen de cette surface était un bon indicateur de l'état du système d'acquisition et de la qualité des mesures réalisées.

Cette étude a également mis en évidence la très forte variabilité des mesures acquises avec un même système d'acquisition. Un suivi des valeurs obtenues sur des échantillons de référence avec le système d'acquisition utilisé et une calibration régulière de celui-ci sont donc essentiels pour pouvoir assurer la répétabilité des mesures.

Enfin, cette étude a montré l'intérêt d'inclure des spectres acquis durant différentes journées d'acquisition dans le set de calibration des modèles de discrimination. Tenir compte des variations des spectres mesurés au cours du temps permet d'améliorer les prédictions des modèles de discrimination.

CHAPITRE V

Utilisation de l'imagerie hyperspectrale proche infrarouge et de la chimiométrie pour la discrimination de racines et de résidus de cultures extraits d'échantillons de sol

"Root coring has its obvious limitations, the most important being that [sample] processing is time-consuming and therefore limits the frequency of samplings during the growth period."

Hooshang Majdi, 1996

Introduction au Chapitre V

Ce chapitre reprend un article publié dans le *Journal of Chemometrics*. Cet article présente les résultats d'une étude ayant permis la mise au point d'une méthode de discrimination de spectres proche infrarouge de racines de froment d'hiver, de résidus de la même culture cultivée l'année précédente et de particules de sol.

L'objectif principal de cette étude était de déterminer dans quelle mesure l'imagerie hyperspectrale proche infrarouge, combinée aux outils de la chimiométrie, pouvait être utilisée pour discriminer des racines et des résidus de cultures extraits de carottes de sol. Si cette discrimination était possible, l'imagerie hyperspectrale permettrait de remplacer le tri manuel qui est jusqu'à présent nécessaire pour séparer les racines des matières organiques également présentes dans les échantillons de sol avant de pouvoir mesurer des paramètres sur les racines.

Un second objectif de cette étude a été de comparer deux modèles d'analyse discriminante (PLS-DA et SVM) pour discriminer les spectres acquis sur les éléments extraits de carottes de sol.

Parmi les matières organiques étrangères qui peuvent être trouvées au sein d'échantillons de sol avec les racines, nous nous sommes uniquement concentrés sur les résidus de cultures (*crop residues*). Ceux-ci constituaient, dans les échantillons analysés, les éléments organiques les plus rencontrés lors du tri manuel et les plus problématiques dans la quantification de la masse racinaire. Les particules de sol, également discriminées grâce aux modèles développés dans cette étude, étaient constituées de petites mottes de terre et de cailloux non éliminés lors du lavage des échantillons.

NIR hyperspectral imaging spectroscopy and chemometrics for the discrimination of roots and crop residues extracted from soil samples

D. Eylenbosch^a, B. Bodson^a, V. Baeten^b, J.A. Fernández Pierna^{b*}

^a Gembloux Agro-Bio Tech, University of Liège, AgroBioChem Department, Crop Science Unit, 2 Passage des déportés, 5030 Gembloux, Belgium.

^b Walloon Agricultural Research Centre, Valorisation of Agricultural Products Department, Food and Feed Quality Unit, Henseval Building, 24 Chaussée de Namur, 5030 Gembloux, Belgium.

*Corresponding author: Juan Antonio Fernández Pierna, j.fernandez@cra.wallonie.be

This is the peer reviewed version of the following article: NIR hyperspectral imaging spectroscopy and chemometrics for the discrimination of roots and crop residues extracted from soil samples, which has been published in final form at <http://onlinelibrary.wiley.com/doi/10.1002/cem.2982/full>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

1. Summary

Utilisation de l'imagerie hyperspectrale proche infrarouge et de la chimiométrie pour la discrimination de racines et de résidus de cultures extraits d'échantillons de sol

Les racines jouent un rôle majeur dans le développement des plantes. Leur étude dans les conditions naturelles de croissance est importante de manière à pouvoir identifier des pratiques de gestion des sols appropriées pour des productions agricoles durables. Le prélèvement de carottes de sol est une méthode couramment utilisée pour mesurer la production racinaire. Sa fréquence d'échantillonnage est limitée par l'étape de tri manuel qui est nécessaire après l'extraction des racines des échantillons de sol de manière à séparer les racines des autres éléments présents dans les carottes de sol tels que les résidus de cultures. Cette étape est chronophage, fastidieuse et sensible à l'habileté et à la subjectivité de l'opérateur. De manière à se passer de cette étape de tri manuel, réduire la confusion entre ces éléments et réduire le temps nécessaire à la quantification racinaire, une nouvelle procédure, basée sur l'imagerie hyperspectrale proche infrarouge et la chimiométrie, a été proposée. Elle a été testée pour discriminer des racines de froment d'hiver (*Triticum aestivum* L.), des résidus de cultures et des particules de sol. Deux algorithmes (SVM et PLS-DA) ont été comparés pour les analyses discriminantes. Les modèles construits avec ces deux algorithmes ont permis la discrimination des racines et des autres éléments mais les meilleurs résultats ont été atteints avec l'algorithme SVM. Les voies de validation des modèles, avec des spectres sélectionnés ou avec des images hyperspectrales, ont

fourni des informations de différent types mais complémentaires. Cette nouvelle procédure permettant la discrimination de racines est une première étape avant de pouvoir quantifier des racines dans des échantillons de sol sur base de l'imagerie hyperspectrale proche infrarouge. Les résultats montrent que cette méthodologie pourrait être un outil intéressant pour avoir une meilleure compréhension des effets du travail du sol ou de la fertilisation, par exemple, sur le développement des systèmes racinaires.

NIR hyperspectral imaging spectroscopy and chemometrics for the discrimination of roots and crop residues extracted from soil samples

Roots play a major role in plant development. Their study in field conditions is important to identify suitable soil management practices for sustainable crop productions. Soil coring, which is a common method in root production measurement, is limited in sampling frequency due to the hand-sorting step. This step, needed to sort roots from others elements extracted from soil cores like crop residues, is time consuming, tedious and vulnerable to operator ability and subjectivity. In order to get rid of the cumbersome hand-sorting step, avoid confusion between these elements and reduce the time needed to quantify roots, a new procedure, based on NIR hyperspectral imaging spectroscopy and chemometrics, has been proposed. It was tested to discriminate roots of winter wheat (*Triticum aestivum* L.) from crop residues and soil particles. Two algorithms (SVM and PLS-DA) have been compared for discrimination analysis. Models constructed with both algorithms allowed the discrimination of roots from other elements but the best results were reached with models based on SVM. The ways to validate models, with selected spectra or with hyperspectral images, provided different kinds of information but were complementary. This new procedure of root discrimination is a first step before root quantification in soil samples with NIR hyperspectral imaging. The results indicate that the methodology could be an interesting tool to improve the understanding of the effect of tillage or fertilization, for example, on root system development.

Keywords: NIR hyperspectral imaging, SVM, PLS-DA, classification, wheat root

2. Introduction

Roots play a major role in plant development. They are the link between belowground resources and aboveground growth¹ and therefore have a great influence on plant biomass production². Root morphology and physiology determine the capacity for nutrient uptake and water extraction by plants³. Their growth is determined by both the genetic traits of plants and the physical and chemical properties of soil, and shows high plasticity under different environmental conditions. The potential of plants to obtain water and mineral nutrients from the soil is primarily attributed to their capacity to develop extensive root systems².

Studies on root development are important for a better understanding of the interaction between crop root systems and the growing environment in order to identify suitable soil management practices for sustainable crop productions⁴. The time taken by roots to colonize the deep soil profile and the root system biomass production during the whole crop development cycle are therefore good indicators to interpret crop behaviour in various management modes and in a given soil-climate context. Nevertheless, roots are rarely directly evaluated in field experiments because soil limits accessibility for their observation^{1, 5}. Therefore, they need to be extracted from soil before any measurement can be made.

Soil coring is a commonly used method to sample roots in field experiments and measure their production⁶. This technique is not expensive and allows repeated measurements during crop growth at several soil depths. After soil coring, cores need to be washed to extract roots from soil. Other elements, mainly crop residues buried in crop soils, are also extracted during the washing step and need to be separated from roots, commonly by hand-sorting, before any analysis or quantification of roots^{1, 4, 5, 7, 8, 9, 10}. This sorting step is time-consuming, tedious and vulnerable to operator ability and subjectivity¹. It is therefore the most significant limitation of the soil coring method because it limits the frequency of sampling during crop growth¹¹.

Near Infrared Spectroscopy (NIRS) is used to identify and quantify components in agricultural products and is characterized by the acquisition of a distinctive spectral chemical profile, which can be considered as a fingerprint of the material. This method allows the acquisition of chemical information from samples with numerous advantages like ease of use and the possibility to quickly and simultaneously analyse several components without any reagents and without sample destruction. It is used to classify or even identify any given material without dependency on human subjectivity. NIRS being an indirect method of measurement needs to be coupled with chemometrics, a chemical discipline using mathematics and statistics to extract the relevant chemical information out of measured data^{12, 13, 14, 15, 16, 17}.

In the last decades, NIRS has been applied to study root systems and to detect differences in the chemical composition of roots. It was used to estimate root biomass proportions in different root mixtures^{18, 19}, to predict the percentage of dead versus living grass roots⁷ and to determine the taxa of herbaceous and woody species²⁰. NIRS was also used as a rapid method to predict pasture and maize root densities in soil cores, directly in the field, without separating soil and roots^{21, 22}.

Recently, new advances have permitted the coupling of NIR spectrometers with imaging technologies, thus creating new devices (Near Infrared Hyperspectral Imaging (NIR-HSI)) providing spectral and spatial information simultaneously^{16, 23}. With this method, thousands of spectra can be obtained for each sample, giving a complete picture of the distribution of chemical compounds at pixel level^{14, 24, 25}.

The aim of this work was to develop a new procedure based on NIR-HSI coupled with chemometrics, allowing a faster discrimination of elements extracted from soil cores in order to decrease the time needed to analyse samples and thereby allowing an increase in the number of soil samples that can be taken in a same time and space.

An increase of sampling throughout crop development will allow a better understanding of the root system's development. An increase of sampling in a same field will improve the knowledge of the spatial heterogeneity existing in field conditions and of the root system's development at several depths. Discrimination of roots and crop residues could lead to a better understanding of root system development in field conditions and its adaptation to its growth environment.

3. Materials and Method

Root and crop residue samples were collected by soil coring in a long-term trial on tillage and crop residue management. The experimental field was located on the experimental farm of Gembloux Agro-Bio Tech (University of Liège, Gembloux, Belgium). A complete description of the site was given by Degrune et al.²⁶. The trial was sown with winter wheat (*Triticum aestivum* L.) in 2010, 2011 and 2012. Soil samples were collected during the second and third year of this crop on several dates (November 2011, April and May 2012), corresponding to different development stages of the crop (first leaf development, tillering and last leaf deployment), or after the harvest (September 2013). After appropriate weeding of the field, all sampled roots and crop residues were considered as produced by the winter wheat crop. We considered as crop residues the above-ground parts of plants remaining in the field after harvest, mainly stubble, chopped straw and husks. These crop residues were mixed with the soil by ploughing (25 cm depth tillage) or reduced tillage (10 cm depth tillage). Soil samples were taken at a depth of 30 cm using a 30-mm diameter soil-coring tube driven into the soil by a tractor-mounted hydraulic push press. They were then divided into three soil horizons (0-10, 10-20 and 20-30 cm deep).

As in the common hand-sorting method, soil samples were washed with tap water in order to extract roots and crop residues. These elements were collected on a sieve (Figure 38). Small stones and soil aggregates were also present. The sieve containing all the elements was dried at 60 °C until a constant weight was reached. After drying, extracted elements were sorted by hand into three classes in order to calibrate and validate the discrimination models: roots, crop residues and soil (small stones and soil aggregates). After model development, this hand-sorting step is no longer necessary. Samples were then kept at ambient temperature and humidity until the acquisition of NIR hyperspectral images.

NIR hyperspectral images were acquired with a moving imager technique combining an NIR hyperspectral line scan, or push-broom imaging system, and a conveyor belt (BurgerMetrics SIA, Riga, Latvia) (Figure 39). The camera was an SWIR XEVA CL 2.5 320 TE4 camera (SPECIM Ltd., Oulu, Finland) using an ImSpector N25E spectrograph that includes a cooled, temperature-stabilized mercury-cadmium-telluride (MCT) detector (XENICS nv, Leuven, Belgium). Images were acquired in the 1118-2424 nm range with a 6,3 nm spectral resolution (i.e. 209 wavelengths) with a width of 320 pixels. For each pixel, 32 scans were co-added and the mean absorbance spectrum was saved. Illumination was achieved with two halogen lamps of 120 W each. The camera position was set up to give images 10 cm in width. Pixels resolution was of 0,31 mm width. The conveyor belt

speed was adapted in order to ensure the pixels were always square-shaped. Acquisitions were performed with HyperPro VB software (BurgerMetrics SIA, Riga, Latvia). The system was first calibrated with white ceramic (white reference) and by blocking the entrance of reflected light (dark reference). A dark reference was automatically performed before acquiring each image. A white reference was performed several times during acquisition day. Bad pixels were detected and removed at the same time.



Figure 38: Picture of the sieve with the different elements extracted from soil cores.

Samples were placed on the conveyor belt and care was taken so that the overlapping of constituents was reduced. For the calibration and the validation of our models, soil, crop residues and roots were sorted manually and images were acquired separately on each class of elements. Only perfectly washed samples of roots and crop residues were used so that soil was not present on these samples. During manual sorting, only perfectly identified elements were selected for model calibrations and validations. For crop residues, spectra from the inner and outer side of pieces of straw were acquired. Preliminary tests on spectra showed that spectral profiles of these two sides were different. It was therefore important to acquire both sides in order to take all the spectral variability into account for model calibrations. Images of the sieve were also acquired because, in the case of discrimination of the elements extracted together from soil cores, all elements stay on the sieve during the acquisition of NIR spectra in order to facilitate sample handling. NIR spectra of the sieve were therefore needed to calibrate discrimination models. In order to build the sieve library, a piece of sieve has been placed alone on the conveyor belt, in the same way than the other elements used for model calibrations and validations.

NIR spectra acquired on sample surface were extracted with HyperSee software (BurgerMetrics SIA, Riga, Latvia) from 24 NIR hyperspectral images of roots, 56

images of crop residues, 5 images of soil and 5 images of the sieve. One matrix of spectra was created for each image. These images were chosen in order to cover the highest possible spectral variability: images of roots and crop residues were acquired on samples collected at different times throughout crop growth and at different soil depths. Moreover, variability due to the system of acquisition was added in the calibration set by acquiring spectra from different samples at different dates and by collecting spectra from the same sample of crop residues used as a reference and acquired on 13 different dates. It was also important to take edge and shading effects into account: spectra were acquired on the whole sample surface and spectra from shadow areas created on the conveyor belt by roots and crop residues were extracted and saved separately. These shadow spectra did indeed have a specific spectral profile and were extracted from 14 NIR hyperspectral images. They were considered as background together with the sieve spectra.



Figure 39: Near infrared (NIR) hyperspectral imaging system installed at the Walloon Agricultural Research Center, Belgium (BurgerMetrics SIA, Riga, Latvia).

Spectra with spikes were removed from all images. These spikes are related to dead pixels that are due to a few elements in the detector responding with an unusually high or low value²⁷. All spectra with absorbance higher than 1,5 at 1740 nm, corresponding to conveyor belt absorbance, were also removed. Due to noise at the beginning and the end of spectra, only the 150 wavelengths between 1432 and 2368 nm were maintained. All this pre-processing was done on Matlab R2015a software (The MathWorks, Inc., Natick, MA, USA).

After the removal of spikes and conveyor belt spectra, data extracted from each image still contained between several hundred to several thousand of spectra. Such a large amount of spectra was useful to take the large sample spectral variability into account for the calibration of models but it contained redundancy and it increased

the time needed to calibrate models. In order to keep the spectral variability and reduce redundancy, the duplex algorithm described by Snee was used to select representative spectra for the calibration set²⁸. Images were gathered in groups (according to sample nature and acquisition date) and around 1000 spectra from each were selected. The duplex algorithm was applied to each image separately to select homogeneously spectra in order to build a spectral library. This method starts by selecting the two points furthest from each other and puts them both in a first set (training). Then the next two points furthest from each other are put in a second set (test), and the procedure is continued by alternatively placing pairs of points in the first or second set²⁹. As a result, data are divided into two groups which include the most diverse spectra, cover approximately the same spectral variability and have similar statistical properties. In total, around 9000 spectra were selected in a first test set which was used for the calibration of the discrimination models. A second selection was done with the duplex algorithm on the remaining spectra in order to create a second test set, the spectral validation set, which also contained around 9000 spectra.

In order to normalize spectra, an autoscale transformation was applied on selected spectra. Preliminary tests showed that this pre-process improved discrimination of our spectral classes and gave best results among classical data pre-treatments. Autoscale is a common pre-processing method which uses mean-centering followed by the division of each variable by its standard deviation³⁰. As a result, all variables of the dataset have a mean of zero and a standard deviation of one.

In this work, the two discrimination algorithms were the following supervised methods: Partial Least Square Discriminant Analysis (PLS-DA) and Support Vector Machines (SVM). PLS-DA is a dimensionality reduction technique that defines new latent variables such that they explain maximum variation of independent and dependent variables³¹. SVM is an algorithm designed to derive a function based on high-dimensional space that describes the hyperplane that optimally separates two classes of data. SVM tries to correctly classify the training data by maximizing the wide of the margin between the groups and by penalizing for errors in order to get a better generalization performance. This is done in a high dimensional space where the groups are linearly separable via a kernel function^{15, 32, 33}.

Pre-processing and model calibrations were done using the PLS Toolbox 7.8 software (Eigenvector Research, Inc., Wenatchee, WA, USA). Discrimination models were arranged in a dichotomous classification tree using Matlab R2015a software (The MathWorks, Inc., Natick, MA, USA). This classification method was chosen for its good performance in a previous study¹⁵.

4. Results and discussion

Spectra of background, soil, crop residues and roots were discriminated thanks to a dichotomous classification tree containing three nodes: each node was based on individual discrimination models by successively extracting background (sieve and shadows) and soil before discriminating crop residues from roots (Figure 40). The

aim of each model was to separate spectra of interest (considered as positive class) from the rest (negative class).

Parameter optimization was performed separately for each model. For PLS-DA models, the number of latent variables was chosen in order to reduce the cross-validation classification error average. A 10-group cross-validation was performed. For SVM models, the cost (penalizing parameter) and the gamma (radial basis function (RBF) kernel parameter) were optimized using the grid search procedure to minimize the misclassification fraction.

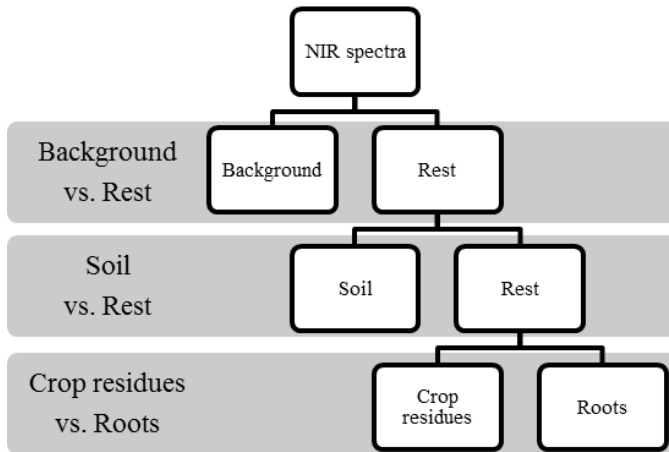


Figure 40: Dichotomous classification tree. Three successive discrimination models allow the separation of near infrared (NIR) spectra in four classes: background, soil, crop residues and roots.

Table 1 presents the results obtained with both algorithms during model calibration, cross-validation and validation. Calibration and cross-validation were obtained with spectra of the calibration set. For the validation, models were tested with spectra of the validation set. These results were expressed in terms of sensitivity and specificity. Sensitivity is the proportion of spectra that are actually positive and classified as positive for one class whereas specificity is the proportion of spectra that are actually negative and classified as negative for one class¹⁵. For instance, for the SVM model concerning the discrimination of background vs. the rest, a sensitivity of 97,1 % and a specificity of 99,8 % were obtained indicating that 97,1 % of the spectra from background were correctly classified as background and 99,8 % of the other spectra were correctly classified as spectra belonging to the rest (i.e. spectra of soil, crop residues and roots altogether).

Table 1: Results obtained during calibration, cross-validation and validation steps of the discrimination models constructed respectively with support vector machine (SVM) or partial least squares discriminant analysis (PLS-DA). Results are expressed in terms of model sensitivity and specificity.

		Calibration		Cross-Validation		Validation	
		Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Background vs. Rest	SVM	97,1 %	99,8 %	95,3 %	99,3 %	94,9 %	95,7 %
	PLS-DA	96,2 %	95,9 %	96,2 %	95,8 %	96,2 %	91,4 %
Soil vs. Rest	SVM	99,2 %	99,7 %	99,2 %	99,5 %	99,2 %	99,7 %
	PLS-DA	99,3 %	98,2 %	99,3 %	98,2 %	98,8 %	96,6 %
Crop residues vs. Root	SVM	97,1 %	92,5 %	94,0 %	85,2 %	94,1 %	85,9 %
	PLS-DA	88,0 %	86,5 %	87,7 %	86,0 %	88,5 %	89,5 %

All models allowed discrimination of the spectral classes. The best results were obtained with the soil vs. rest model: in general, up to 99 % of soil spectra were correctly classified. The background vs. rest model also provided good discrimination with up to 95 % of correctly classified spectra in general. For both models, the specificity was higher with SVM than with PLS-DA. The discrimination between crop residues and roots was more difficult: 14 % of root spectra used to validate models were classified as crop residues with SVM. For this third model, crop residues were always better predicted with SVM.

The validation of the three discrimination models applied separately on the spectra of the validation set confirmed the results obtained during the calibration of models and the ability of models to discriminate spectra from the different spectral classes. It gave a raw idea of model performance when applied to new spectra and was therefore very useful during model construction. However, it was observed that this validation of the model sometimes gave an inaccurate picture of the model's quality. In some cases, excellent results were obtained during model construction and validation but bad predictions were obtained when these models were used to predict the spectral nature of pixels on hyperspectral images. Moreover, validation done on each discriminant model separately did not take into account the sequence of models applied with the classification tree. The discrimination models were therefore also validated in a second way. Table 2 gives results obtained when the whole dichotomist classification trees, constructed respectively with SVM and PLS-DA algorithms, were applied to 264 new NIR hyperspectral images in order to predict the spectral nature of each pixel constituting these images. These images were acquired on samples containing only elements corresponding to one spectral class (sieve, soil, crop residues or roots). This second table presents the number of images

used for the four spectral classes, the number of pixels predicted for each class (in total and after removal of spectra predicted as background) as well as the percentage of pixels that were or were not correctly predicted in each group. These results do not take into account pixels having a high absorbance value which were removed during image pre-treatment and corresponding to conveyor belt. The results show that background was correctly predicted on images containing only the sieve (no elements extracted from soil cores): less than 1 % of pixels were detected as belonging to another spectral class with both algorithms and the SVM tree correctly predicting 99,9 % of background pixels. The sieve was not present on the other images and only shadow areas were thus considered as background. The amount of pixels predicted as shadow reached one to two thirds of pixels analysed by models. These pixels were present on the conveyor belt, along the edge of elements. The highest part of pixels predicted as shadow was reached on images of roots. These elements were thin and the ratio between shadow along the edge of roots and their surface was therefore higher.

Table 2: Number and percentage of pixels predicted as background (sieve and shadow areas), soil, crop residues or roots on NIR hyperspectral images acquired on samples containing only one spectral class.

Classification tree	SVM				PLS-DA			
Samples	Sieve	Soil	Residues	Roots	Sieve	Soil	Residues	Roots
Number of images	4	12	131	117	4	12	131	117
Total number of pixels	175.583	323.407	4.880.123	2.525.402	175.583	323.407	4.880.123	2.525.402
Pixels classified as Background	99,9 %	35,2 %	49,6 %	56,2 %	99,1 %	40,6 %	51,9 %	63,2 %
Remaining pixels after background removal								
Pixels classified as	Soil, residues or roots	209.572	2.460.102	1.107.303		192.168	2.345.244	928.890
	Soil (%)	99,3 %	1,8 %	0,8 %		97,5 %	2,2 %	5,0 %
	Crop residues (%)	0,2 %	80,8 %	8,3 %		0,0 %	81,7 %	7,3 %
	Roots (%)	0,5 %	17,4 %	90,8 %		2,5 %	16,0 %	87,7 %

Number of pixels and percentages were calculated by cumulating, for each spectral class, the results obtained on all hyperspectral images analysed for this class. PLS-DA indicates partial least squares discriminant analysis; SVM, support vector machine.

Images of soil were also well predicted. The SVM tree correctly predicted 99,3 % of pixels and the PLS-DA tree correctly predicted 97,5 % of pixels, with only slight confusion appearing with root spectra.

The weakest results were obtained on images of roots and crop residues with both classification trees. With the SVM tree, 17,4 % of pixels of crop residues were predicted as roots and 8,3 % of root pixels were predicted as crop residues. With the PLS-DA tree, 16 % of crop residue pixels were predicted as roots and part of the root pixels were predicted as crop residues (7,3 %) or as soil (5 %). These percentages did not take into account the number of pixels predicted as background by the first model of discrimination in the classification tree. Considering the results in absolute values, the number of pixels correctly classified as background, soil, crop residues or roots on the images of these four classes was always higher with the SVM tree than with the PLS-DA tree. The SVM tree correctly classified a higher proportion of pixels as background on images of the sieve and classified a lower proportion of pixels in this class on images of soil, crop residues and roots. The PLS-DA tree classified a higher proportion of pixels as roots on soil images and a higher proportion as soil on root images. The highest difference between SVM and PLS-DA trees was observed with images of roots: the SVM tree predicted 19 % more pixels as roots.

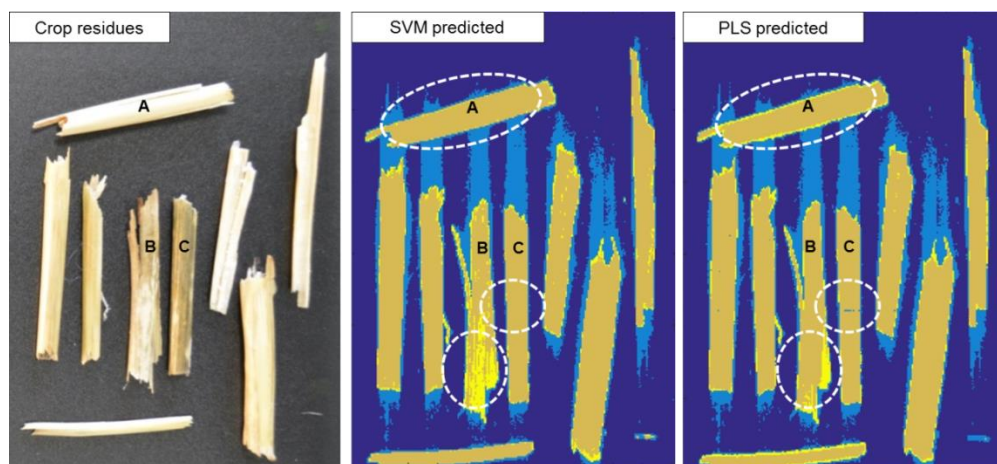


Figure 41: Picture of crop residues and predictions of NIR hyperspectral images of crop residues with SVM and PLS-DA dichotomous classification trees. Pixels classified as crop residues are in ochre. Pixels classified as roots are in yellow. Differences between predictions with SVM and PLS-DA are illustrated in dashed circles and can be observed (A) on the border of crop residues, (B) on the inner face of straw and (C) in terms of prediction of dead pixels.

This method of validating models was much more time-consuming in terms of prediction and analysis than the first way of validation but it gave interesting results:

it was possible to validate models on a very large number of spectra, therefore with a higher variability, and to quantify the proportion of spectra wrongly predicted in each spectral class separately. Furthermore, it was possible to identify regions of hyperspectral images where pixels were not well classified. Indeed, based on prediction results for each pixel, images were reconstructed by assigning one colour to each spectral class. The visual analysis of reconstructed images reveals how the different image regions were predicted and allows the identification of areas where spectra were wrongly classified. On images of crop residues, pixels located on the border of crop residues were often classified as roots (Figure 41 A). This mistake was lower when using SVM. In contrast, the SVM tree applied to images of the inner side of crop residues classified a higher number of pixels as roots in the central part of crop residues than PLS-DA (Figure 41 B). However, the analysis of pixel number predicted for each spectral class showed that, on 92,4 % of crop residue images, the number of pixels correctly classified as crop residues was higher with the SVM tree. Dead pixels were sometimes present on images. These pixels were generally better classified with the SVM tree (Figure 41 C). On images of roots, pixels classified as crop residues were mostly located in the central area of large roots. On these images, a large part of root pixels were classified as background (Figure 42). Considering all images of roots predicted with both models, a higher pixel number was classified as roots with the SVM tree on 99,1 % of root images.

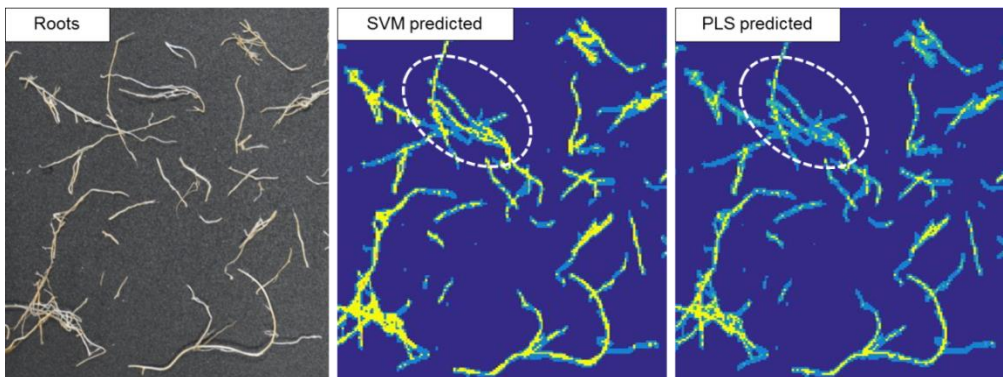


Figure 42: Picture of roots and predictions of NIR hyperspectral images of roots with SVM and PLS-DA dichotomous classification trees. The main differences between predictions with SVM and PLS-DA were a higher number of pixels classified as roots (in yellow) and a lower number of pixels classified as background (in royal blue) as illustrated in the dashed circle.

This confusion of roots and crop residues spectra was already observed during the calibration and the first validation of models. The analysis of predicted hyperspectral images allowed to highlight that this confusion appeared mainly on the border of crop residues. A large part of the spectra on the border of elements resulted from the chemical nature of several elements: the conveyor belt and elements extracted from

soil cores or, in the case of samples with mixed elements, different elements extracted from soil cores that overlapped on the conveyor belt during image acquisition. Vermeulen et al. reported that spectra in a pixel are often a mixture of the pure constituent spectra³⁴. Spectra acquired in the central part of crop residues were not affected by this phenomenon because crop residues are quite large elements and it was easy to acquire pure spectra on it. This confusion due to the low spatial resolution of the instrument and the mix of spectra was also well observed on predicted images of roots. The roots used in this study were sampled under a winter wheat crop and they were very thin: their diameter was mostly smaller than the pixel size. Much of the root spectra were therefore the result of mixed spectra of root and background resulting in confusion between these two spectral classes. Preliminary tests showed that model results were improved when these border areas were taken into account to construct models. Other elements of confusion, located in the central part of roots or crop residues, seem to be due to the shape of the elements and their interaction with light. Such edge and shaping effects were also observed and discussed by Manley et al.³⁵. All spectral confusions were therefore most likely linked to the way spectra were acquired. Figure 43 shows how similar were the spectra of root and crop residues used to calibrate the discriminant models. Furthermore, it seems that these confusions were not due to the chemical composition of roots and crop residues because chemical analysis demonstrated that their C/N content was different. In addition, they were perfectly discriminated when samples were ground (results not shown). Soon et Arshad also measured a higher C/N ratio in the straw of winter wheat than in their roots³⁶ and Kamruzzaman et al. observed that minced samples were much better classified, probably due to the homogeneous surface³⁷.

The comparison of models has indicated that models built with SVM better discriminated root and crop residue spectra both during model calibration and during prediction on hyperspectral images. These results confirmed the good performances of SVM previously observed and its superiority to solve complex problems^{15, 16, 32}. Spectral classes were better separated and areas of confusion like crop residue borders, thin roots or bad pixels were better predicted. It was observed on predicted images and it could explain the difference observed in Table 1 between the selectivity and the sensitivity of the SVM model discriminating crop residues and roots. Applied to the sample presented in Figure 38, SVM gave a good prediction of all the elements extracted from a soil core (Figure 44). The only limitation identified with SVM was the time needed to construct models, which was longer than with PLS-DA. Due to the large quantity of data, several hours were necessary to calibrate the SVM models whereas only a few seconds were needed for PLS-DA. However once the models are built the performance of SVM is larger in terms of sensitivity and robustness.

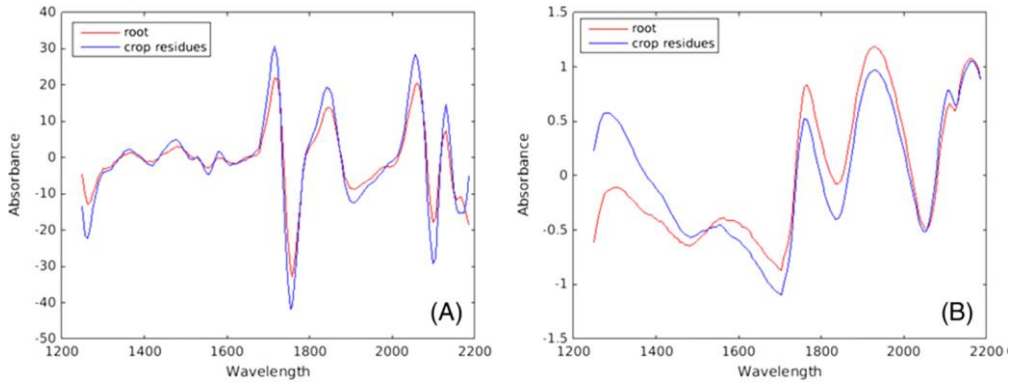


Figure 43: Mean spectra of roots and crop residues of winter wheat after (A) second derivative or (B) Standard Normal Variate (SNV) transformation. These mean spectra were calculated on the spectra used to calibrate the models. They were very similar, and the discrimination of root and crop residues was therefore challenging.

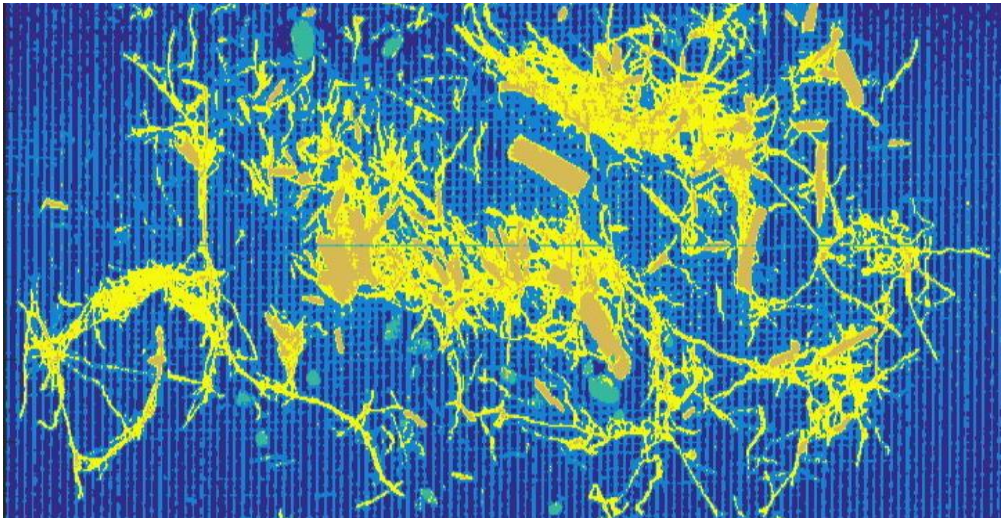


Figure 44: SVM prediction on an NIR hyperspectral image acquired on a sample containing wheat roots (in yellow), crop residues (in ochre), small stones and clots (in turquoise) and the sieve (in royal blue). The conveyor belt is in dark blue.

5. Conclusions

This work demonstrated the ability of NIR hyperspectral imaging spectroscopy and chemometrics to discriminate winter wheat roots, crop residues and soil particles extracted from soil cores. To reach the best discrimination, it was important to identify all sources of spectral variability on hyperspectral images. This

variability can be due to the nature of analysed elements, presence of shadow areas on images or variability induced by the acquisition system itself. In routine use, predictions would be improved if discrimination models were regularly updated with new spectra in order to take into account as much spectral variability as possible. The interest of validating discrimination models on NIR hyperspectral images was demonstrated. It allows the validation of models when they are combined in a dichotomous classification tree and the identification of areas of spectral confusion. This method of validation proved to be very useful to estimate a model's ability to discriminate elements extracted from soil cores.

The comparison of chemometric tools confirmed the good performances of SVM. Models built with this algorithm classified a higher proportion of spectra in the correct spectral class and were less sensitive in areas of spectral confusion than models built with PLS-DA.

Although discrimination between the spectra of roots and crop residues was not perfect, this new procedure, which is rapid and does not depend on operator subjectivity, offers new possibilities in studies on root systems. NIR hyperspectral imaging and chemometrics could be used to quantify roots extracted from soil cores sampled in field conditions in order to replace the time-consuming and tedious manual step currently required with the soil coring method after washing soil samples. Furthermore, this new procedure could also be of use to quantify crop residues in soils and study their degradation dynamics, an important step in soil carbon fixation. This root or crop residue quantification could be based on the link between the weight of roots or crop residues in a sample analysed with the hyperspectral camera and the number of pixels predicted as belonging to these spectral classes by the discriminant models. Coupled with the soil coring method, this new procedure could be used to study tillage and nitrogen fertilization effects on winter wheat root system development in field conditions.

6. Acknowledgements

The authors would like to thank Françoise Vancutsem and Jérôme Pierreux from the Crop Science Unit (GxABT) for supplying samples of roots and crop residues, as well as Remi Delcroix and Loïc Dewaele for their help in collecting these samples. They would also like to thank Guillaume Fraipont from the Crop Science Unit (GxABT) and Nicaise Kayoka Mukendi from the Valorisation of Agricultural Products Department (CRA-W) who helped with image acquisitions, as well as Frederic Collinet and the CAMI technical platform at the AgroBioChem Department (GxABT) for access to the computational server.

7. References

1. Plaza-Bonilla D, Álvaro-Fuentes J, Hansen NC, Lampurlanés J, Cantero-Martínez C. Winter cereal root growth and aboveground-belowground biomass ratios as affected by site and tillage system in dryland Mediterranean conditions. *Plant Soil*, 2014;**374**(12):925-939. doi:10.1007/s11104-013-1926-3.
2. Guan D, Al-Kaisi MM, Zhang Y, et al. Tillage practices affect biomass and grain yield through regulating root growth, root-bleeding sap and nutrients uptake in summer maize. *Field Crop Res.* 2014;**157**:89-97. doi:10.1016/j.fcr.2013.12.015.
3. Fan J, McConkey B, Wang H, Janzen H. Root distribution by depth for temperate agricultural crops. *Field Crop Res.* 2016;**189**:68-74. doi:10.1016/j.fcr.2016.02.013.
4. Huang G, Chai Q, Feng F, Yu A. Effects of different tillage systems on soil properties, root growth, grain yield, and water use efficiency of winter wheat (*Triticum aestivum* L.) in arid northwest China. *J Integr Agric.* 2012;**11**(8) : 1286-1296. doi:10.1016/S2095-3119(12)60125-7.
5. Cheng W, Coleman DC, Box JE. Root dynamics, production and distribution in agrosystems on the Georgia Piedmont using minirhizotrons. *J Appl Ecol.* 1990;**27**(2):592-604. doi:10.2307/2404304.
6. Jose S, Gillespie AR, Seifert JR, Pope PE. Comparison of minirhizotron and soil core methods for quantifying root biomass in a temperate alley cropping system. *Agroforestry Syst.* 2001;**52**(2):161-168. doi:10.1023/A:1010667921970.
7. Picon-Cochard C, Pilon R, Revaillet S, Jestin M, Dawson LA. Use of near-infrared reflectance spectroscopy to predict the percentage of dead versus living grass roots. *Plant Soil.* 2009;**317**(1-2):309-320. doi:10.1007/s11104-008-9810-2.
8. Izumi Y, Uchida K, Iijima M. Crop production in successive wheat-soybean rotation with no-tillage practice in relation to the root system development. *Plant Prod Sci.* 2004;**7**(3):329-336. doi:10.1626/pp.7.329.
9. Qin R, Stamp P, Richner W. Impact of tillage on root systems of winter wheat. *Agron J.* 2004;**96**(6):1523-1530. doi:10.2134/agronj2004.1523.
10. Kätterer T, Hansson AC, Andréon O. Wheat root biomass and nitrogen dynamics-effects of daily irrigation and fertilization. *Plant Soil.* 1993;**151**:21-30. doi:10.1016/j.agee.2011.02.029.
11. Majdi, H. Root sampling methods - application and limitations of the minirhizotron technique. *Plant Soil.* 1996;**185** : 255-258. doi:10.1007/BF02257530.
12. Baeten V, Rogez H, Fernández Pierna JA, Vermeulen P, Dardenne P. Vibrational spectroscopy methods for the rapid control of agro-food products. In : Nolle LML, Toldra F, eds. *Handbook of Food Analysis*. 3rd ed. CRC Press; 2015;**2**:591-614.
13. Abbas O, Dardenne P, Baeten V. Near-infrared, mid-infrared, and raman spectroscopy. In : Picó Y, ed. *Chemical Analysis of Food: Techniques and Applications*. Waltham, MA : Elsevier Inc; 2012:59-91. doi:10.1016/B978-0-12-384862-8.00003-0.
14. Dale LM, Thewis A, Boudry C, et al. Hyperspectral imaging applications in agriculture and agro-food product quality and safety control: a review. *Appl*

Spectrosc Rev, 2013;**48**(2):142-159. doi:10.1080/05704928.2012.705800.

15. Fernández Pierna JA, Vermeulen P, Amand O. NIR hyperspectral imaging spectroscopy and chemometrics for the detection of undesirable substances in food and feed. *Chemometr Intell Lab.* 2012;**117**:233-239. doi:10.1016/j.chemolab.2012.02.004.

16. Fernández Pierna JA, Baeten V, Michotte Renier A, Cogdill RP, Dardenne P. Combination of support vector machines (SVM) and near-infrared (NIR) imaging spectroscopy for the detection of meat and bone meal (MBM) in compound feeds. *J Chemometr.* 2004;**18**(7-8):341-349. doi:10.1002/cem.877.

17. Massart DL, Vandeginste BGM, Buydens LMC, De Jong S, Lewi PJ, Smeyers-Verbeke J. *Handbook of Chemometrics and Qualimetrics : Part A.* Amsterdam, Elsevier; 1997.

18. Roumet C, Picon-Cochard C, Dawson LA, et al. Quantifying species composition in root mixtures using two methods: near-infrared reflectance spectroscopy and plant wax markers. *New Phytol.* 2006;**170**(3):631-638. doi:10.1111/j.1469-8137.2006.01698.x.

19. Rumbaugh MD, Clark DH, Pendery BM. Determination of root mass ratios in alfalfa-grass mixtures using near infrared reflectance spectroscopy. *J Range Manage.* 1988;**41**(6) : 488-490. doi:10.2307/3899523.

20. Rewald B, Meinen C. Plant roots and spectroscopic methods - analyzing species, biomass and vitality. *Front Plant Sci.* 2013;**4**:393. doi:10.3389/fpls.2013.00393.

21. Kusumo BH, Hedley MJ, Tuohy MP, Hedley CB, Arnold G. Predicting soil carbon and nitrogen concentrations and pasture root densities from proximally sensed soil spectral reflectance. In : Viscarra Rossel R, McBratney A, Minasny B, eds. *Proximal Soil Sensing. Progress in Soil Science.* Springer, Dordrecht; 2010:177-190.

22. Kusumo BH, Hedley MJ, Hedley CB, Tuohy MP. Measuring carbon dynamics in field soils using soil spectral reflectance: prediction of maize root density, soil organic carbon and nitrogen content. *Plant Soil.* 2011;**338**(1-2):233-245. doi:10.1007/s11104-010-0501-4.

23. Dale LM, Thewis A, Rotar I, et al. Chemometric tools for NIRS and NIR hyperspectral imaging. *Bulletin UASVM Agriculture.* 2012;**69**(1):70-76. <http://orbi.ulg.ac.be/handle/2268/132855>. Accessed July 06, 2017.

24. Shahin MA, Symons SJ, Hatcher DW. Quantification of mildew damage in soft red winter wheat based on spectral characteristics of bulk samples: a comparison of visible-near-infrared imaging and near-infrared spectroscopy. *Food Bioprocess Tech.* 2014;**7**(1):224-234. doi:10.1007/s11947-012-1046-8.

25. Fernández Pierna JA, Baeten V, Dubois J, Burger J, Lewis EN, Dardenne P. NIR imaging - Theory and applications. In : Brown S, Tauler R, Walczak B, eds. *Comprehensive Chemometrics.* Oxford : Elsevier; 2009;**4**:173-196.

26. Degruene F, Theodorakopoulos N, Dufrière M, et al. No favorable effect of reduced tillage on microbial community diversity in a silty loam soil (Belgium). *Agr Ecosyst Environ.* 2016;**224**:12-21. doi:10.1016/j.agee.2016.03.017.

27. Burger J, Geladi P. Hyperspectral NIR image regression part I: Calibration and correction. *J Chemometr.* 2005;**19**:355-363. doi:10.1002/cem.986.

28. Snee RD. Validation of regression models: methods and examples. *Technometrics*. 1977;**19**(4):415-428. doi:10.2307/1267881.

29. Fernández Pierna J A, Lecler B, Conzen J P, Niemoeller A, Baeten V, Dardenne P. Comparison of various chemometric approaches for large near infrared spectroscopic data of feed and feed products. *Anal Chim Acta*. 2011;**705**(1-2):30-34. doi : 10.1016/j.aca.2011.03.023.

30. Wise BM, Gallagher NB, Bro R, Shaver JM, Windig W, Koch RS. PLS_Toolbox Version 4.0 for use with MATLAB TM. Wenatchee, WA: Eigenvector Research Inc; 2006. http://mitr.p.lodz.pl/raman/jsurmacki/pliki/zajecia/LMDiT/cw4i5/LMDiT_PLS_Manual_4.pdf Accessed July 06, 2017.

31. Konda Naganathan G, Grimes LM, Subbiah J, Calkins CR, Samal A, Meyer GE. Partial least squares analysis of near-infrared hyperspectral images for beef tenderness prediction. *Sens Instrumen Food Qual*. 2008;**2**(3):178-188. doi:10.1007/s11694-008-9051-3.

32. Zhang X. Support vector machines. In: Sammut C, Webb G, eds. *Encyclopedia of Machine Learning*. New York: Springer; 2010:941-946. doi:10.1007/978-0-387-77242-4.

33. Fernández Pierna JA, Baeten V, Dardenne P. Screening of compound feeds using NIR hyperspectral data. *Chemometr Intell Lab*. 2006;**84**(1-2):114-118. doi:10.1016/j.chemolab.2006.03.012.

34. Vermeulen P, Fernández Pierna JA, van Egmond HP, Dardenne P, Baeten V. Online detection and quantification of ergot bodies in cereals using near infrared hyperspectral imaging. *Food Addit Contam A*. 2012;**29**(2):232-240. doi:10.1080/19440049.2011.627573.

35. Manley M, du Toit G, Geladi P. Tracking diffusion of conditioning water in single wheat kernels of different hardnesses by near infrared hyperspectral imaging. *Anal Chim Acta*. 2011;**686**(1):64-75. doi:10.1016/j.aca.2010.11.042.

36. Soon YK, Arshad MA. Comparison of the decomposition and N and P mineralization of canola, pea and wheat residues. *Biol Fert Soils*. 2002;**36**:10-17. doi:10.1007/s00374-002-0518-9.

37. Kamruzzaman M, Barbin D, ElMasry G, Sun DW, Allen P. Potential of hyperspectral imaging and pattern recognition for categorization and authentication of red meat. *Innovative Food Sci Emerging Technol*. 2012;**16**:316-325. doi:10.1016/j.ifset.2012.07.007.

Citation : Eylenbosch D, Bodson B, Baeten V, Fernández Pierna JA. NIR hyperspectral imaging spectroscopy and chemometrics for the discrimination of roots and crop residues extracted from soil samples. *Journal of Chemometrics*. 2017;e2982. <https://doi.org/10.1002/cem.2982>

Conclusions du Chapitre V

Les résultats obtenus dans cette étude ont démontré qu'il était possible de discriminer des racines, des résidus de cultures et des particules de sols récupérés après lavage d'échantillons de sol sur base de leurs spectres acquis par imagerie hyperspectrale proche infrarouge. Cette discrimination n'était cependant pas parfaite et les spectres d'un certain nombre de pixels étaient mal prédits. Ces spectres provenaient principalement de pixels acquis en bordures des éléments de l'échantillon. Pour réduire les confusions lors des prédictions de la nature des spectres, la prise en compte des zones d'ombres présentes sur les images et de l'instabilité du système d'acquisition était nécessaire.

Néanmoins, malgré ces confusions, l'imagerie hyperspectrale ouvre des possibilités pour se passer de l'étape de tri manuel des échantillons qui était jusqu'à présent nécessaire dans les études sur les systèmes racinaires et constituait un frein important dans le traitement des échantillons.

La comparaison des performances des modèles de discrimination PLS-DA et SVM réalisée dans cette étude a aussi confirmé les meilleures performances du modèle SVM. Ce modèle classait un plus grand nombre de spectres dans la bonne classe spectrale.

CHAPITRE VI

**Utilisation de l'imagerie hyperspectrale
proche infrarouge et de la chimiométrie
pour quantifier la masse de racines
extraites de carottes de sol**

Introduction au Chapitre VI

Nous avons montré au chapitre précédent qu'il était possible de détecter des racines sur des images hyperspectrales grâce à l'analyse de leurs spectres proche infrarouge par des modèles de discrimination. L'objectif de ce chapitre est de présenter l'étape suivante dans l'étude de systèmes racinaires sur base de l'imagerie hyperspectrale proche infrarouge. Elle consiste à relier le nombre de pixels de racines détectés sur les images hyperspectrales à la masse de racine présente au sein des échantillons étudiés.

La construction de droites de régression permettant de relier des nombres de pixels de racines de froment d'hiver et de pois d'hiver avec des masses de racines de ces deux cultures a été présentée lors de la conférence *Phloème 2018, première biennale de l'innovation céréalière* qui s'est tenue à Paris les 24 et 25 janvier 2018. L'article repris à la première section de ce chapitre a été publié dans les actes du congrès. Cet article présente également deux exemples de quantifications racinaires basés sur l'utilisation de l'imagerie hyperspectrale proche infrarouge : une quantification de densités racinaires dans un essai de longue durée sur le travail du sol et la quantification séparée de racines de cultures de froment et de pois d'hiver cultivées en association et selon différentes fertilisations azotées.

La seconde et la troisième section de ce chapitre présentent la validation de cette méthode de quantification et son utilisation dans différents cas d'études agronomiques. La quantification de résidus de cultures au sein d'échantillons de sol par imagerie hyperspectrale proche infrarouge y est également présentée.

1. Utilisation de l'imagerie hyperspectrale proche infrarouge combinée aux outils de la chimométrie dans l'étude de systèmes racinaires

Damien Eylenbosch⁽¹⁾, Juan Antonio Fernández Pierna⁽²⁾, Vincent Baeten⁽²⁾, Bernard Bodson⁽¹⁾

⁽¹⁾ Phytotechnie tempérée, Département AgroBioChem, Gembloux Agro-Bio Tech, Université de Liège, Belgique.

⁽²⁾ Qualité des produits, Départements Valorisation des productions, Centre wallon de Recherches agronomiques, Belgique.

Avec la collaboration de Benjamin Dumont, Marie-Pierre Hiel, Guillaume Fraipont et Jérôme Pierreux.

1.1. Résumé

Utilisation de l'imagerie hyperspectrale proche infrarouge combinée aux outils de la chimométrie dans l'étude de systèmes racinaires

Lors d'études de systèmes racinaires sur base de prélèvements de carottes de sol, une étape préliminaire à la quantification racinaire consiste à éliminer manuellement les résidus de cultures et les particules de sol récupérés avec les racines lors du lavage des échantillons. En présence de racines de différentes espèces, cette étape s'accompagne d'un tri des racines selon différents critères. Pour faciliter et réduire le temps nécessaire à cette étape fastidieuse, la discrimination de ces éléments a été réalisée sur base d'images hyperspectrales prises en proche infrarouge. Les échantillons lavés et séchés ont été passés sous une caméra qui enregistre, pour chaque pixel de l'image, un spectre en proche infrarouge. Ces spectres sont liés à la nature physico-chimique des éléments présents sur l'image et ont été analysés à l'aide d'algorithmes d'analyses discriminantes qui permettent de les classer et de donner la nature de chaque pixel de l'image. Sur base du nombre de pixels classés comme étant des racines et grâce à la relation linéaire reliant ce nombre de pixels à la masse de racines présente dans l'échantillon, il fut finalement possible de quantifier des racines au sein d'échantillons de sol sans devoir réaliser un tri manuel préalable et les peser.

Cette méthode a été testée avec succès en froment et a permis de mettre très clairement en évidence l'effet du travail du sol sur le développement du système racinaire. Elle a également été testée sur une culture associée de froment et de pois afin de quantifier séparément les racines des deux espèces.

Use of near infrared hyperspectral imaging and chemometrics to study root systems

In studies on root systems using the soil coring method, roots, crop residues and soil particles extracted after washing of soil samples need to be manually sorted before root quantification. If different species are present in the sample, their discrimination is based on several criteria. In order to reduce the complexity and the time needed for this tedious sorting step, near infrared hyperspectral imaging was used to discriminate all these elements. Images of washed and dried samples were acquired with a camera saving, for each pixel, a near infrared spectrum. These spectra are linked to the physicochemical nature of the elements visible on the hyperspectral images and were analyzed with discriminant algorithms allowing their discrimination and to link each pixel spectrum to the nature of the elements. According to the number of pixels classified as roots and thanks to the linear relationship between the number of pixels and the root biomass in the sample, it was possible to quantify roots in soil samples without manual sorting and weighing.

This methodology was successfully tested on winter wheat roots and highlighted the tillage effect on root system development. It was also tested on wheat-pea intercropping in order to quantify separately the roots of each species.

1.2.Introduction

La compréhension des interactions entre le système racinaire des cultures et l'environnement dans lequel elles se développent est nécessaire afin d'identifier les pratiques de gestion du sol appropriées et augmenter les performances des cultures. La manière dont les racines se développent dans l'espace et le temps détermine le potentiel de la culture à absorber l'eau et les nutriments et, finalement, le rendement de la culture (Plaza-Bonilla et al. 2014; Huang et al. 2012; Qin et al. 2004; Cheng et al. 1990). Le développement du système racinaire dépend fortement de la saison culturale, de la teneur en eau et de la structure du sol ainsi que du stade de développement de la plante (Cai et al. 2014; Plaza-Bonilla et al. 2014; Morell et al. 2011). Le temps nécessaire aux racines pour coloniser le sol en profondeur et la production de biomasse au niveau du système racinaire tout au long du cycle de développement de la plante sont dès lors de bons indicateurs pour interpréter le comportement de la plante cultivée selon différents modes de gestion et dans un contexte de sol et de climat donné. Certains paramètres influençant le développement racinaire peuvent en effet être modifiés par les pratiques culturales. C'est entre autres le cas avec le travail du sol qui influence la densité du sol et sa teneur en eau (Strudley et al. 2008). C'est également le cas pour les éléments nutritifs disponibles pour la plante dont les quantités présentes dans le sol peuvent être modifiées par la fertilisation ou encore lors d'une culture en association avec une légumineuse. Cette deuxième pratique fait l'objet de nombreuses recherches afin de réduire la dépendance des cultures céréalières à l'azote de synthèse et donc aux ressources d'énergies fossiles.

Cependant, la mesure des effets de ces pratiques culturales sur le développement racinaire des cultures se fait rarement dans des essais en champs, car la présence du

sol limite les observations et l'extraction des racines demande dès lors beaucoup de travail et de temps (Plaza-Bonilla et al. 2014). Leurs effets sur le développement racinaire restent donc mal connus.

Parmi les quelques méthodes utilisables pour étudier les systèmes racinaires en conditions de plein champ, celle basée sur le prélèvement de carottes de sol est fréquemment utilisée. Cette méthode est peu coûteuse, perturbe peu la culture et permet de nombreux prélèvements, mais implique un lavage des carottes de sol et un tri des différents éléments présents au sein de l'échantillon de sol : racines, résidus de matières organiques, cailloux... Ce n'est qu'après ces étapes que les racines peuvent être analysées ou quantifiées (Plaza-Bonilla et al. 2014; Huang et al. 2012; Picon-Cochard et al. 2009; Izumi et al. 2004; Qin et al. 2004; Cheng et al. 1990). L'étape de tri est jusqu'à présent réalisée à la main ce qui la rend chronophage, pénible et sujette à l'habileté et la subjectivité de l'opérateur (Plaza-Bonilla et al. 2014). Selon Majdi (1996) cette étape constitue la limitation la plus importante de la méthode basée sur le prélèvement de carottes de sol, car elle réduit la fréquence des échantillonnages réalisables durant la croissance de la culture. Dans le cas d'études s'intéressant aux systèmes racinaires de plusieurs espèces cultivées en association, un tri des racines des différentes espèces est aussi nécessaire. Ce tri se base alors sur des critères morphologiques (couleur, texture, mode d'enracinement) si les espèces peuvent être distinguées sur base de tels critères ou par l'utilisation de colorants, de radio-isotopes, ou encore d'analyses d'ADN (Rewald et al. 2012; Li et al. 2006; Hauggaard-Nielsen & Jensen 2005).

Afin de réduire le temps et la pénibilité du tri des échantillons avant la quantification des racines, nous avons étudié les potentialités offertes par l'imagerie hyperspectrale proche infrarouge et les modèles chimiométriques pour la discrimination des différents éléments présents au sein de ces échantillons de sol ainsi que pour leur quantification.

L'imagerie hyperspectrale proche infrarouge résulte de la combinaison des technologies de l'imagerie et de la spectroscopie proche infrarouge. Utilisée seule, la spectroscopie proche infrarouge est une méthode d'analyse qui permet l'acquisition du spectre proche infrarouge propre à un échantillon. Ce spectre peut être considéré comme la signature spectrale de l'échantillon. Cette méthode permet donc l'acquisition d'informations chimiques et a de nombreux avantages tels que la facilité d'utilisation et la possibilité d'analyser rapidement et simultanément différents composants sans devoir utiliser de réactif et sans détruire l'échantillon (Dale et al. 2013; Fernández Pierna et al. 2012). L'analyse des spectres permet également de classer ou d'identifier des matériaux sans être influencé par la subjectivité humaine (Dale et al. 2013; Fernández Pierna et al. 2004).

Au cours des 30 dernières années, la spectroscopie a été utilisée de plusieurs façons pour étudier les systèmes racinaires. Elle fut utilisée pour estimer les proportions de biomasses racinaires au sein de mélanges de racines (Roumet et al. 2006; Rumbaugh et al. 1988), déterminer les taxons d'espèces herbagées ou forestières (Rewald & Meinen 2013), discriminer des racines d'avoine et de pois (Naumann et al. 2010) et pour prédire, au sein d'un échantillon de racines de graminées, le pourcentage de racines vivantes ou mortes (Picon-Cochard et al.

2009). Kusumo et al. (2011 et 2010) ont également utilisé la spectroscopie proche infrarouge, dans leur cas directement au champ, pour estimer rapidement la densité racinaire de maïs et de pâtures au sein de carottes de sol sans devoir séparer les racines du sol.

Toutes ces méthodes d'analyse spectroscopiques étaient basées sur l'étude du spectre moyen des échantillons analysés. Avec l'imagerie hyperspectrale, la combinaison de l'imagerie et de la spectroscopie permet d'acquérir simultanément des données spectrales et spatiales. Avec cette nouvelle technologie, chaque pixel d'une image hyperspectrale contient l'entièreté du spectre proche infrarouge ce qui permet de l'utiliser pour caractériser la composition des régions de l'image (Shahin et al. 2014). Cette technologie est souvent utilisée pour contrôler la composition, la qualité et la sécurité de produits agricoles (Dale et al. 2013; Wang & Paliwal 2007; Fernández Pierna et al. 2006).

Tout comme la spectroscopie classique, l'imagerie hyperspectrale proche infrarouge est une méthode indirecte. La calibration de modèles de discrimination ou de régression est donc nécessaire pour obtenir des analyses qualitatives ou quantitatives (Dale et al. 2013). La calibration de ces modèles est rendue possible grâce aux outils de la chimométrie qui permettent l'extraction des informations pertinentes des mesures effectuées par utilisation d'algorithmes mathématiques et statistiques (Massart et al., 1988 cité par Dale et al. 2012).

Dans le cadre de cet article, nous présentons deux cas d'études où l'imagerie hyperspectrale proche infrarouge a été utilisée pour quantifier le développement racinaire de cultures. Dans le premier cas, l'effet du travail du sol sur le développement du système racinaire du froment d'hiver (*Triticum aestivum* L.) a été étudié. Dans le second cas d'étude, les effets de la fumure azotée et de la culture en association du froment d'hiver et du pois protéagineux d'hiver (*Pisum sativum* L.) ont été étudiés sur le développement racinaire de ces deux cultures. Les effets du travail du sol et de l'association de cultures sont deux sujets qui soulèvent de nombreuses questions. Si le nombre d'études menées en conditions de plein champ sur les effets du travail du sol et de la fertilisation est en augmentation, la manière dont ces pratiques affectent le développement racinaire n'est pas encore parfaitement comprise (Rasmussen et al. 2015; Wang et al. 2014; Huang et al. 2012; Muñoz-Romero et al. 2010; Munkholm et al. 2008). Il en va de même pour l'effet de l'association de cultures (Li et al. 2006; Hauggaard-Nielsen & Jensen 2005).

1.3. Matériels et méthodes

1.3.1. Échantillonnage

Les échantillons de racines ont été prélevés dans des essais réalisés au champ à l'aide d'une sonde de 3 cm de diamètre enfoncée dans le sol à l'aide d'un vérin hydraulique monté à l'avant d'un tracteur. Les prélèvements ont été réalisés sur 30 cm de profondeur et ont été fractionnés en 3 horizons de 10 cm. Pour chaque échantillon de sol, 5 prélèvements étaient rassemblés. Les échantillons ont été mis dans des sacs en plastique et gardés à 4°C en attendant leur traitement.

1.3.2. Lavage des échantillons

Les échantillons de sol ont été lavés à l'eau afin d'en extraire les racines. Lors de ce lavage, d'autres éléments ont été extraits avec les racines tels que des résidus de la culture précédente ou de petits cailloux. Afin de faciliter la désagrégation du sol, les échantillons de sol ont été préalablement immergés durant 48 h dans une solution contenant 50 g NaCl par litre d'eau. Les éléments récupérés lors du lavage ont été étendus sur un grillage à fine maille (type moustiquaire) et séchés à l'étuve (60 °C durant 48 h).



Figure 45: Système d'acquisition des images hyperspectrales proche infrarouge (BurgerMetrics SIA, Riga, Latvia). L'échantillon qui est éclairé par deux spots halogènes est posé sur une bande transporteuse et passe sous la caméra.

1.3.3. Acquisition des images hyperspectrales

Les images hyperspectrales en proche infrarouge ont été acquises selon la même méthode que celle décrite par Vermeulen et al. (2012). La caméra hyperspectrale permettant l'acquisition de spectres proche infrarouge est positionnée au-dessus d'une bande transporteuse (BurgerMetrics SIA, Riga, Lettonie) (Figure 45). Le système projette un faisceau de lumière sur une matrice à plan focal à deux dimensions. Chaque image hyperspectrale est acquise à 209 longueurs d'onde (1118-2425 nm) avec une résolution spectrale d'environ 6,3 nm et sur 320 pixels de large. L'image hyperspectrale fournit un spectre pour chaque pixel, chaque spectre étant le résultat moyen de 32 mesures d'absorbance. Les images ont été acquises avec un objectif couvrant une plage de 10 cm de large sur la bande transporteuse. Les acquisitions ont été réalisées avec le logiciel HyperProVB (BurgerMetrics SIA, Riga, Lettonie). La vitesse d'avancement de la bande transporteuse a été adaptée de telle manière à avoir toujours des pixels de forme carrée. Afin de faciliter les manipulations, les échantillons ont été placés sous la caméra avec le grillage utilisé lors du séchage. Les échantillons ont été disposés de telle manière à ne pas dépasser

de la plage de mesure de la caméra. Ils ont été partiellement étendus afin de limiter autant que possible le recouvrement des éléments (Figure 46).



Figure 46: Racines de froment, racines de pois, nodosités, fragments de pailles, particules de terre et cailloux extraits d'un échantillon de sol, étendus sur le grillage de support et posés sur la bande transporteuse de la caméra hyperspectrale proche infrarouge. Les éléments sont disposés de telle manière à éviter tant que possible qu'ils se recouvrent. Un spectre proche infrarouge va être acquis pour chaque pixel de l'image.

1.3.4. Discrimination des spectres proche infrarouge

Les différents éléments présents sur les images hyperspectrales ont été regroupés en quatre classes spectrales. Dans le premier cas d'étude, ces classes étaient constituées des spectres i) du grillage, ii) des particules de sol et des cailloux restants après le lavage, iii) des résidus de la culture précédente et iv) des racines de froment. Dans le second cas d'études, une seule classe regroupait les spectres du grillage, des particules de sol, des cailloux et des résidus de la culture précédente. Les trois autres classes étaient constituées des spectres des racines de pois, des racines de froment et des nodosités. Ces différentes classes spectrales ont été discriminées grâce à un arbre de classification dichotomique comprenant trois algorithmes d'analyse discriminante selon la même méthode que celle utilisée par Fernández Pierna et al. (2012). Les algorithmes utilisés sont basés sur les *Support Vector Machines* (SVM). Ces algorithmes sont une méthode chimiométrique dont l'objectif est de trouver l'hyperplan maximisant la distance entre deux groupes prédéfinis de données en minimisant les erreurs de classification (Fernández Pierna et al., 2004; Cornuéjols, 2002).

Vu la grande variation de l'absorbance aux extrémités du spectre proche infrarouge acquis, seules les 150 longueurs d'onde comprises entre 1432 et 2368 nm ont été utilisées pour la calibration des algorithmes et la prédiction de la nature des pixels. Une élimination des spectres aberrants a également été réalisée avant la sélection des spectres utilisés pour la calibration. La calibration des modèles utilisés

dans les deux cas d'étude est basée sur une sélection d'environ 9000 et 6000 spectres respectivement. Ces spectres ont été sélectionnés de manière à contenir un maximum de variabilité pour chaque classe spectrale des modèles.

1.3.5. Quantification des racines

La quantification des racines présentes dans chaque échantillon de sol a été réalisée grâce à la relation linéaire qui existe entre le nombre de pixels classé par les modèles de discrimination comme étant des racines sur l'image hyperspectrale et la masse de racines présentes dans cet échantillon. Les équations utilisées dans les deux cas d'études ont été construites sur base d'images hyperspectrales d'échantillons de racines dont la masse avait été préalablement mesurée. Le nombre de pixels de racines présentes sur l'image hyperspectrale étant influencée par la manière dont l'échantillon est présenté devant la caméra et par le pourcentage de recouvrement des racines entre elles, au moins deux images ont été prises par échantillon utilisé pour la construction des équations. Entre chaque prise d'image, la disposition des racines sous la caméra était modifiée de sorte à avoir une densité de racine par cm² différente entre chaque image.

1.3.6. Traitement des données et des résultats

La calibration des modèles de discrimination a été réalisée à l'aide de la PLS Toolbox 7.8 (Eigenvector Research, Inc., Wenatchee, WA, USA). Le logiciel Matlab (The MathWorks, Inc., Natick, MA, USA) a servi pour le traitement des spectres proche infrarouge et pour la classification des pixels des images hyperspectrales. L'analyse statistique des résultats de la quantification de la densité racinaire a été réalisée grâce au logiciel R (R Core Team 2017). Le package "agricolae" a été utilisé pour les analyses de la variance (ANOVA) et les tests de Student-Newman-Keuls (Mendiburu & Simon 2015).

1.4. Résultats et discussion

1.4.1. Cas d'étude 1 : Quantification de racines de froment sous labour ou travail superficiel

Dans ce premier cas d'étude, l'objectif était de mesurer l'effet du travail du sol sur le développement du système racinaire du froment d'hiver. Les échantillons de sol ont été prélevés au sein de l'essai SOLRESIDUS mis en place en 2008 par Gembloux Agro-Bio Tech (Université de Liège, Belgique) et décrit par Degruene et al. (2016). Cet essai a pour objectif d'étudier les effets de différents modes de gestion des résidus de cultures : le labour conventionnel sur 25 cm de profondeur ou le travail superficiel réalisé sur une profondeur de 10 cm avec un outil à dents et à disques, avec ou sans restitution des résidus de cultures. La culture de froment d'hiver sous laquelle les prélèvements de sol ont été réalisés avait été implantée le 11 octobre 2011. Les prélèvements de sol ont été réalisés un mois après le semis, le 17 novembre 2011, et le 8 mai 2012 au stade phénologique "deux nœuds" de la culture (stade 32 sur l'échelle BBCH améliorée de Hack et al. (1992)). Pour chaque modalité de travail du sol, deux échantillons de sol ont été prélevés dans chacune des quatre répétitions de l'essai à raison d'un échantillon dans la ligne de culture et un

autre dans l'interligne. Chacun de ces échantillons était constitué de cinq carottes de sol.

Avant de quantifier les racines présentes dans les échantillons de sol, il fallait pouvoir discriminer ces racines des résidus de la culture précédente (du froment) et des particules de sol également présents dans les échantillons de sol. La validation des modèles permettant la discrimination de ces différentes classes a été réalisée sur des spectres indépendants de ceux utilisés pour la calibration des modèles. Les résultats obtenus sont présentés dans le Tableau 3 et sont exprimés en termes de sensibilité (pourcentage de spectres correctement classés dans la première classe du modèle de discrimination) et de spécificité (pourcentage de spectres correctement classés dans la seconde classe du modèle de discrimination) (Fernández Pierna et al. 2012). Une seconde validation a été réalisée en appliquant les modèles de discrimination sur des images de racines de froment (Tableau 4).

Tableau 3: Validation sur spectres proche infrarouge des trois modèles permettant la discrimination des racines du froment, des résidus de la culture précédente (froment), des particules de sol restantes après lavage de l'échantillon et du tamis permettant la manipulation de l'échantillon. Validation réalisée pour chaque modèle séparément sur une sélection de spectres différents de ceux utilisés pour la calibration des modèles. Les résultats sont exprimés en termes de sensibilité (pourcentage de spectres correctement classés dans la première classe du modèle de discrimination) et de spécificité (pourcentage de spectres correctement classés dans la seconde classe du modèle de discrimination).

Classes spectrales	Qualité des discriminations	
	Sensibilité	Spécificité
Tamis vs. Reste	94,9 %	95,7 %
Particules de sol vs. Reste	99,2 %	99,7 %
Résidus vs. Racines de froment	94,1 %	85,9 %

Les résultats obtenus lors de la validation de ces modèles ont montré une bonne classification des spectres dans chaque classe. Une confusion apparaît néanmoins entre les spectres de racines de froment et les résidus de cultures. Cette confusion apparaît principalement en bordure des résidus de cultures lors de l'application des modèles de discrimination sur les images hyperspectrales. Elle pourrait s'expliquer par le bruit présent au niveau des spectres de racines. Celles-ci ayant un diamètre bien souvent inférieur à la taille des pixels (0,3 mm), leurs spectres sont probablement influencés par l'arrière-plan de l'image, tout comme les pixels se trouvant en bordure des résidus de cultures.

Tableau 4: Validation des modèles de discrimination permettant la discrimination des résidus de la culture précédente (froment) et des racines de froment. Validation réalisée en appliquant la succession des trois modèles de discrimination sur des images hyperspectrales acquises sur des échantillons contenant uniquement des résidus de la culture de froment ou des racines de froment. Les résultats sont exprimés en termes de pourcentage moyen de pixels classés par les modèles de discrimination dans les différentes classes spectrales. Le pourcentage de pixels correctement classé est en caractère gras.

Image (nombre)	Classification des spectres (Moyenne ± écart-type)	
	Résidus	Racines de froment
Résidus (131)	81 ± 8 %	17 ± 6 %
Racines de froment (117)	8 ± 6 %	91 ± 5 %

Sur base du nombre de pixels prédits comme étant des racines de froment sur les images hyperspectrales acquises sur les échantillons, une quantification de la masse de racines présentes au sein de l'échantillon a pu être réalisée grâce à l'Équation 4. Cette relation, illustrée à la Figure 47, a été construite sur base du nombre de pixels prédits comme étant des racines de froment sur des images acquises sur des échantillons de masse connue. Cette équation est donc strictement liée aux modèles d'analyse discriminante calibrés dans le cadre de la présente étude. La masse de racines calculée a ensuite été divisée par le volume de terre échantillonné (0,353 dm³) afin d'exprimer les résultats en termes de densité de racine (g) par volume de sol (dm³).

Équation 4: Quantification de la masse de racines de froment (g) en fonction du nombre de pixels classés dans cette classe spectrale par les modèles d'analyse discriminante calibrés dans le cadre de l'étude ($r^2 = 0,98$).

$$\text{Masse racines} = \frac{\text{Nombre de pixels de racines}}{254435}$$

La Figure 48 présente les résultats obtenus suite à l'application des modèles de discrimination sur les images hyperspectrales acquises sur les échantillons prélevés dans l'essai sur le travail du sol et dont le nombre de pixels prédit comme étant des racines a été converti en termes de densités racinaires. Cette figure présente, pour chaque date de prélèvement et pour les deux intensités de travail du sol, la densité racinaire calculée sur les trois horizons de sol échantillonnés.

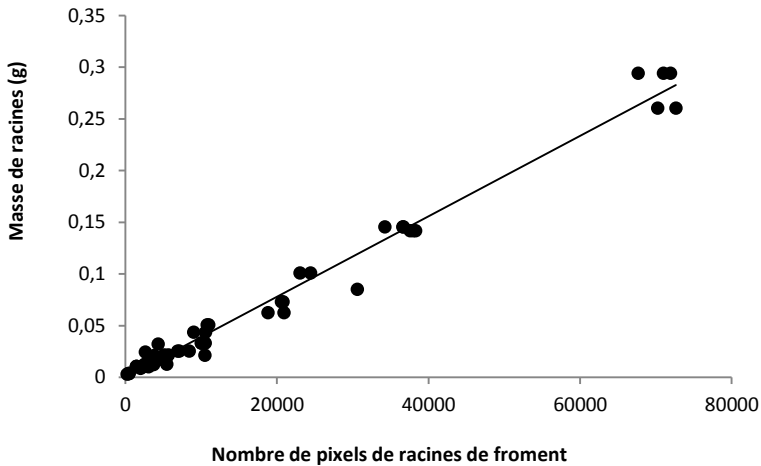


Figure 47: Relation linéaire reliant la masse (g) d'échantillons composés uniquement de racines de froment et le nombre de pixels classé, par les modèles de discrimination calibrés pour cette étude, comme étant des racines de froment sur les images hyperspectrales.

L'analyse statistique des résultats a montré une répartition homogène des racines sur l'entièreté du profil étudié (30 cm) pour le labour lors du premier prélèvement. Avec le travail superficiel, la densité racinaire un mois après le semis est plus élevée qu'en labour dans l'horizon supérieur et est moins élevée entre 10 et 30 cm de profondeur. Ces résultats montrent que selon la profondeur du travail du sol, le froment développe préférentiellement son système racinaire dans l'horizon ameubli et descend donc plus ou moins vite en profondeur. Six mois plus tard, lors du prélèvement réalisé au stade "deux nœuds" de la culture de froment, les densités racinaires sont équivalentes dans les deux horizons supérieurs, mais la densité racinaire entre 20 et 30 cm reste plus faible en travail superficiel. Il est néanmoins intéressant de remarquer qu'aux deux dates de prélèvement, la densité moyenne de racines sur 30 cm était statistiquement égale pour les deux modalités de travail du sol. Le travail du sol influencerait donc la distribution des racines dans le sol, mais pas la quantité totale produite. Cette affirmation devrait cependant être vérifiée sur l'entièreté du système racinaire du froment qui se développe jusqu'à une profondeur de 1,5-2 m (Rasmussen et al. 2015) ce qui n'a pas été fait dans le cadre de cette étude.

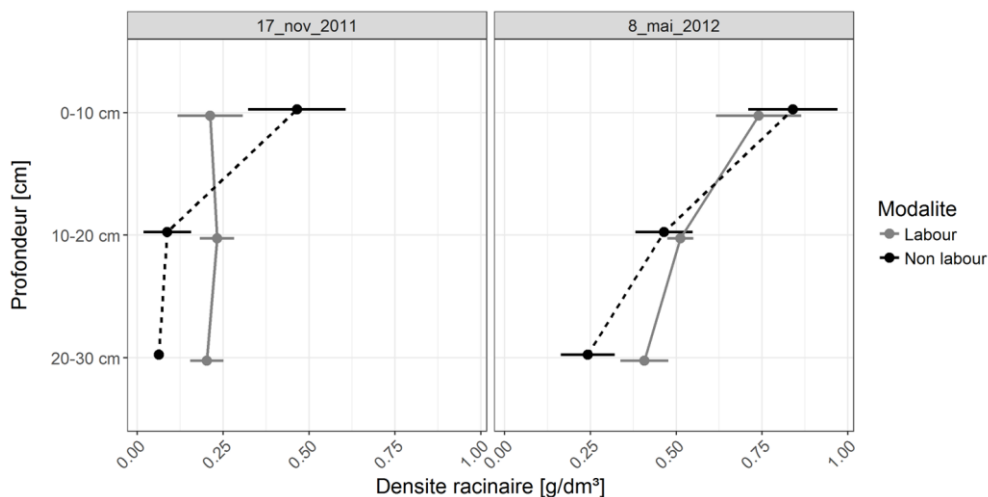


Figure 48: Densités racinaires (g/dm^3) d'une culture de froment d'hiver calculée dans trois horizons de sol (0-10, 10-20 et 20-30 cm) sur base d'images hyperspectrales prises en proche infrarouge sur des échantillons de sol lavés. Les échantillons ont été prélevés un mois après le semis de la culture (17 novembre 2011) et lorsque la culture était au stade phénologique "deux nœuds" (8 mai 2012) dans un essai comparant les effets du labour sur 25 cm de profondeur (trait continu) et un travail superficiel sur 10 cm de profondeur (non-labour; trait discontinu). La densité racinaire est statistiquement différente en labour et non-labour le 17 novembre dans les trois horizons de sol. Le 8 mai, une différence de densité racinaire fut observée entre les deux modes de travail du sol dans l'horizon de sol 20-30 cm. En labour, aucune différence de densité racinaire n'a été observée entre les trois horizons de sol lors du premier prélèvement.

1.4.2. Cas d'étude 2 : Quantification de racines de froment et de pois en culture associée

Dans ce second cas d'étude, les racines ont été prélevées le 28 juin 2016 dans un essai sur la fertilisation de cultures associées de froment d'hiver et de pois protéagineux d'hiver. Cet essai, réalisé à Gembloux (Belgique) dans un sol limoneux à drainage favorable après une culture de froment d'hiver, comprenait des parcelles de froment d'hiver et de pois protéagineux cultivés seuls ou en association. À cette date, la culture de froment était au stade "grain mi-laiteux" (BBCH 73) et les graines de pois étaient complètement formées (BBCH 79). La fumure avait été apportée en deux fractions aux stades phénologiques du froment "fin tallage-début redressement" (BBCH 30) et "dernière feuille" (BBCH 39). Les échantillons ont été prélevés au sein de quatre modalités de fumure : soit aucun apport d'azote (0-0), soit un apport d'azote de 40 kg/ha appliqué au stade "tallage-redressement" de la culture de froment suivi d'un second apport de 40 kg/ha (40-40), 80 kg/ha (40-80) ou 120 kg/ha (40-120) au stade "dernière feuille". Quatre échantillons de sol ont été prélevés pour chaque modalité : deux parcelles ont été échantillonnées à raison de

deux échantillons de sol par parcelle. Les échantillons ont été prélevés dans l'interligne de la culture.

Tableau 5: Validation sur spectres proche infrarouge des trois modèles permettant la discrimination de racines de froment et de pois, de nodosités, de résidus de la culture de froment, de particules de sol restantes après lavage de l'échantillon et du tamis permettant la manipulation de l'échantillon. Validation réalisée pour chaque modèle séparément sur une sélection de spectres différents de ceux utilisés pour la calibration des modèles. Les résultats sont exprimés en termes de sensibilité (pourcentage de spectres correctement classés dans la première classe du modèle de discrimination) et de spécificité (pourcentage de spectres correctement classés dans la seconde classe du modèle de discrimination).

Classes spectrales	Qualité des discriminations	
	Sensibilité	Spécificité
(Tamis + sol + résidus) vs. Reste	89 %	89 %
Racines de pois vs. Reste	78 %	83 %
Racines de froment vs. Nodosités	99 %	96 %

Tableau 6: Validation des modèles de discrimination permettant la discrimination de racines de froment et de pois. Validation réalisée en appliquant la succession des trois modèles de discrimination sur des images hyperspectrales acquises sur des échantillons contenant uniquement des racines de pois ou des racines de froment. Les résultats sont exprimés en termes de pourcentage moyen de pixels classés par les modèles de discrimination dans les différentes classes spectrales. Le pourcentage de pixels correctement classé est en caractère gras.

Images (nombre)	Classification des spectres (Moyenne \pm écart-type)		
	Racines pois	Racines froment	Nodosités
Racine pois (27)	75 \pm 9 %	24 \pm 9 %	1,1 \pm 0.4 %
Racine froment (32)	14 \pm 6 %	86 \pm 6 %	0 \pm 0 %

Les modèles de discrimination mis au point permettaient d'extraire successivement les spectres proche infrarouge rassemblés en quatre classes : (i) le tamis sur lequel était posé l'échantillon + les éléments du sol (terre et cailloux) + les résidus de la culture précédente (froment), (ii) les racines de pois, (iii) les racines de froment, (iv) les nodosités de pois. La qualité de la discrimination de ces quatre classes spectrales a été évaluée sur des spectres différents de ceux utilisés pour la calibration des modèles. La qualité des modèles a également été étudiée en appliquant les modèles sur des images ne contenant que des racines de pois ou de

froment d'hiver et en calculant le pourcentage moyen de pixels correctement classés. Les résultats de ces validations sont donnés dans les Tableaux 5 et 6.

Les résultats de la validation sur spectres montrent une excellente discrimination des racines de froment et des nodosités ce qui se confirme lors de la validation sur images. La discrimination des spectres de la première classe (tamis, particules de sol et résidus de la culture précédente) est bonne. Une analyse plus en profondeur des résultats a mis en évidence une excellente discrimination du tamis et des particules de sol et une certaine confusion entre les spectres de résidus et de racines des deux cultures. La confusion la plus importante se situe entre les racines de froment et de pois avec néanmoins un pourcentage de bonne classification proche de 80 % obtenu aussi bien lors de la validation sur spectres que lors de la validation sur images hyperspectrales.

Tout comme dans le cas de l'étude précédente, la confusion entre les résidus de la culture précédente et les racines, dans ce cas-ci des cultures de froment et de pois, apparait principalement en bordure des résidus. Elle peut donc à nouveau être expliquée par le fait que les spectres acquis dans cette zone sont influencés par l'arrière-plan²⁵ et ne correspondent donc pas exactement au profil des spectres utilisés pour la calibration des modèles. Il en va de même pour la confusion entre les racines des deux cultures qui, de par leurs tailles, ont des spectres fortement influencés par l'arrière-plan. L'analyse des spectres moyens de ces trois classes spectrales et des nodosités, calculés sur l'ensemble des spectres utilisés pour la calibration des modèles (1100 spectres de résidus de froment, 1100 de racines de pois, 1098 de racines de froment, 300 de nodosités) montre par ailleurs une parfaite discrimination de ces spectres moyens (Figure 49). La confusion des spectres des différentes classes n'est donc pas due à la nature chimique des classes, mais à la manière dont les spectres sont acquis et à la diversité des spectres présents au sein de chaque classe spectrale.

La conversion en masse du nombre de pixels classé comme étant des racines de froment ou de pois par les modèles d'analyse discriminante calibrés pour cette étude a été réalisée grâce aux Équations 5 et 6. La Figure 50 présente les densités racinaires calculées dans les horizons de sol échantillonnés pour les cultures de froment et de pois cultivées seules ou en association selon quatre niveaux de fumure azotée.

Équation 5: Quantification de la masse de racines de froment (g) en fonction du nombre de pixels classés dans cette classe spectrale par les modèles d'analyse discriminante calibrés dans le cadre de l'étude ($r^2 = 0,89$).

$$\text{Masse racines froment} = \frac{\text{Nombre pixels racines froment} - 6846}{233817}$$

²⁵ L'arrière-plan (*background*) est le support sur lequel sont posés les échantillons lors de l'acquisition des images hyperspectrales. L'arrière-plan contient aussi les zones d'ombres.

Équation 6: Quantification de la masse de racines de pois (g) en fonction du nombre de pixels classés dans cette classe spectrale par les modèles d'analyse discriminante calibrés dans le cadre de l'étude ($r^2 = 0,76$).

$$\text{Masse racines pois} = \frac{\text{Nombre pixels racines pois} + 3486}{426825}$$

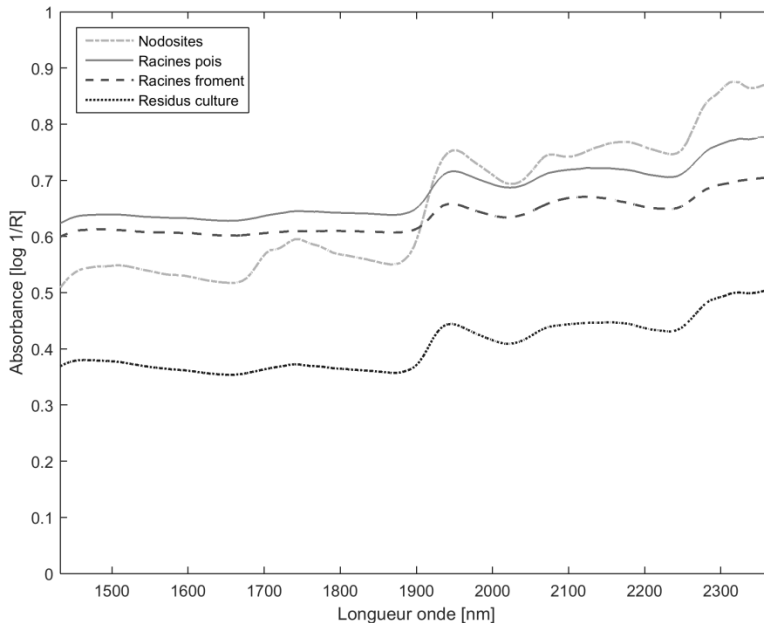


Figure 49: Spectres moyens de racines de pois et de froment, de résidus de culture de froment et de nodosités de pois acquis en proche infrarouge. Valeurs d'absorbances pour 150 longueurs d'onde comprises entre 1432 et 2368 nm.

Une première analyse de ces résultats fait apparaître une forte variabilité des densités calculées ce qui rend l'analyse des données difficile. Cette variabilité peut s'expliquer par différents facteurs liés aux méthodes d'échantillonnage, d'acquisition des images et de classification des spectres.

Concernant l'échantillonnage, un échantillon de sol était constitué dans ce second cas d'étude de cinq carottes de sol prélevées dans l'interligne de la culture. Sachant que la densité racinaire sous la ligne de semis peut être plus élevée (Li et al. 2006), le prélèvement de carottes de sol dans la ligne et l'interligne, comme dans le premier cas d'étude, aurait probablement permis de mieux caractériser la densité racinaire de la culture.

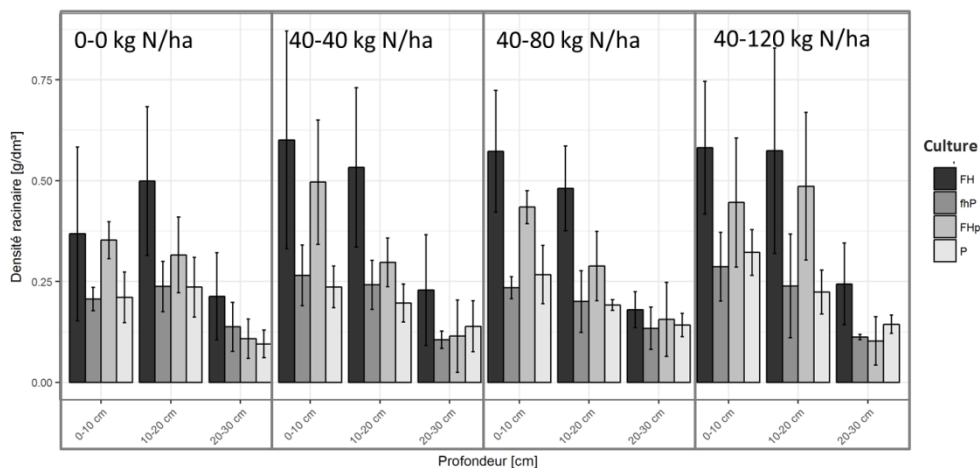


Figure 50 : Densités racinaires (g/dm^3) de cultures de froment d'hiver cultivé seul (FH) ou en association (FHp) et de pois d'hiver cultivé seul (P) ou en association (fhP). Ces densités ont été calculées sur base d'images hyperspectrales acquises en proche infrarouge sur des échantillons de sol lavés. Les échantillons ont été prélevés le 28 juin 2016 dans 3 horizons de sol (0-10, 10-20 et 20-30 cm) dans des cultures menées selon 4 fertilisations azotées contrastées : soit aucune fertilisation (0-0), soit un apport d'azote de 40 kg/ha appliqué au stade "tallage-redressement" de la culture de froment suivi d'un second apport de 40 kg/ha (40-40), 80 kg/ha (40-80) ou 120 kg/ha (40-120) au stade "dernière feuille".

Au niveau de l'acquisition des images, plus un échantillon contient des éléments, plus le risque de recouvrement entre eux est élevé. Un meilleur étalement de l'échantillon sur la bande transporteuse permettrait donc d'avoir une quantification plus précise du volume racinaire présent dans l'échantillon de sol, mais entraînerait une augmentation du temps nécessaire à son traitement. Un juste compromis doit donc être trouvé.

Enfin, pour la classification des spectres, la confusion observée sur les images hyperspectrales entre les spectres de racines de froment et de pois entraîne une sous-estimation de la quantité de racines dans les échantillons prélevés dans les cultures pures et un biais pour les échantillons contenant des racines des deux cultures. Ce biais n'est pas mesurable pour chaque image séparément, mais peut être estimé en moyenne sur les images acquises sur des échantillons purs tels que ce fut fait pour la validation sur images des modèles d'analyse discriminante. Nous avons alors calculé qu'en moyenne 24 % des pixels de racines de pois étaient classés comme étant du froment et que 14 % des pixels de racines de froment étaient classés comme étant des racines de pois.

Malgré une forte variabilité des valeurs calculées, les résultats obtenus donnent plusieurs informations intéressantes sur les densités racinaires de cet essai. Ils montrent clairement que la densité racinaire de pois est identique lorsque cette culture est cultivée seule ou en association, et cela malgré une densité de semis

adaptée (80 grains/m² en culture pure, 50 grains/m² en culture associée). L'analyse statistique n'a pas non plus montré de différence significative de densité racinaire entre le froment cultivé seul et en association. La densité du froment d'hiver ayant également été adaptée lors du semis en association (150 grains/m² au lieu de 300 grains/m² en culture pure), nous pouvons donc conclure à un meilleur développement des systèmes racinaires des deux cultures conduites en association. La densité racinaire totale de la culture en association, calculée pour chaque modalité de fertilisation sur 30 cm de profondeur, n'a néanmoins jamais surpassé significativement la densité racinaire calculée en froment pur. La densité racinaire de pois cultivé seul était par contre toujours inférieure à celles du froment ou de l'association des cultures.

Au niveau de la culture de froment, un effet significatif de la profondeur a été observé sur la densité racinaire. La densité racinaire calculée dans l'horizon supérieur du sol était plus élevée que celle calculée entre 20 et 30 cm de profondeur. Ceci est observable systématiquement pour le froment cultivé en association, quelle que soit la fumure azotée apportée. Dans le cas des fertilisations 0-0 et 40-120, la densité racinaire du froment associé était égale dans les deux horizons supérieurs alors que dans les deux autres modalités de fertilisation, la densité racinaire est statistiquement différente dans les trois horizons de sol. En froment pur, une différence de densité entre les horizons n'est observée que pour la modalité de fertilisation 40-80. Dans le cas du pois, cultivé seul ou en association, l'effet de la profondeur sur la densité racinaire est moins marqué mis à part pour le pois cultivé seul avec la fertilisation 40-120 pour lequel il y a une nette différence entre les trois horizons étudiés.

Aucun effet significatif de la fertilisation azotée n'a pu être mis en évidence sur la densité racinaire totale sur 30 cm de profondeur pour les deux cultures, qu'elles soient cultivées seules ou en association, bien qu'une tendance à l'augmentation des densités racinaires du froment soit observée lorsque la quantité d'azote apportée augmente.

1.5. Discussion générale et conclusion

La méthode de quantification de racines basée sur l'utilisation de l'imagerie hyperspectrale proche infrarouge a montré dans ces deux cas d'étude ses avantages et ses limites. Par rapport au tri manuel généralement utilisé pour séparer les racines des résidus de cultures, l'utilisation de l'imagerie hyperspectrale prend moins de temps ce qui permet de traiter un plus grand nombre d'échantillons sur un même laps de temps et donc de répéter plus facilement les mesures. De plus, elle est beaucoup moins fastidieuse pour l'opérateur et elle n'est pas influencée par sa subjectivité.

Pour le classement de racines d'espèces différentes, l'imagerie hyperspectrale proche infrarouge est moins onéreuse que l'utilisation d'isotopes (Rewald et al. 2012) et semble beaucoup moins laborieuse que la technique de microscopie permettant la discrimination des espèces sur base de l'analyse morphologique et

anatomique de la section transversale des racines utilisée par Ramirez-Garcia et al. (2015).

Par rapport à la spectroscopie proche infrarouge classique qui fut déjà utilisée à plusieurs reprises dans l'étude de systèmes racinaires (Meinen & Rauber 2015; Roumet et al. 2006; Rumbaugh et al. 1988), la combinaison avec l'imagerie permet de ne plus mesurer des valeurs relatives (telles que les proportions de racines de différentes espèces au sein d'un échantillon), mais de pouvoir calculer directement des masses de racines sur base de la relation existant entre la quantité de pixels présents sur l'image hyperspectrale et la masse de l'échantillon.

Dans les deux cas d'études présentés, cette nouvelle méthode de quantification a permis de mettre très clairement en évidence l'effet du travail du sol sur la densité racinaire du froment, l'absence d'effet de la fumure azotée sur la densité racinaire du pois cultivé seul ou en association et des différences significatives de densité racinaire du froment entre les trois horizons de sol étudiés. La bonne classification des spectres des résidus de cultures et de nodosités aurait également pu permettre de quantifier ces éléments dans ces deux cas d'études afin d'étudier l'effet du travail du sol sur la répartition et la dégradation des résidus de cultures ainsi que l'effet de la fumure azotée sur le développement des nodosités de pois.

Néanmoins, cette méthode de quantification de racines comporte encore un certain nombre de biais qui influencent le calcul de la quantité de racines présente dans un échantillon. Lors de l'acquisition des images, une partie de l'information peut être perdue par manque de séparation des éléments de l'échantillon (racines, résidus de cultures...). Avec les réglages de la caméra utilisés pour ces études, la largeur de la plage d'acquisition des spectres était de 10 cm. L'échantillon devait être dispersé sur cette largeur de telle manière à limiter le recouvrement des éléments. Lors de l'analyse des images, la confusion des spectres appartenant aux différentes classes des modèles induit un second biais dans la quantification des éléments de l'échantillon. Enfin, la quantité finale de racine calculée dépend aussi de la qualité de la relation linéaire permettant de convertir en termes de masse le nombre de pixels classés comme étant des racines. Il est donc nécessaire d'utiliser un grand nombre d'échantillons pour la calibration de cette relation.

Une meilleure répartition des éléments sur la bande transporteuse lors de l'acquisition des images, l'utilisation d'un plus grand nombre d'échantillons pour la calibration des modèles de discrimination et des droites de régression ainsi qu'une meilleure prise en compte des ombres et des effets de bordures, qui influencent la qualité de la discrimination des spectres, devraient permettre à l'avenir d'améliorer encore cette nouvelle méthode d'étude des systèmes racinaires.

Dans l'état actuel de son développement, cette méthode ne doit donc pas être utilisée en espérant avoir une quantification précise des racines présentes au sein d'un échantillon, mais comme une méthode permettant d'obtenir des valeurs relatives et permettant la comparaison de densités racinaires au sein de systèmes de culture ou d'horizons de sol différents et cela de manière plus rapide et moins fastidieuse que par la méthode de tri manuel et de pesée couramment utilisée avec le

prélèvement de carottes de sol dans les études sur le développement de systèmes racinaires en conditions de plein champ.

1.6. Références bibliographiques

Cai H. et al., 2014. Effect of subsoil tillage depth on nutrient accumulation, root distribution, and grain yield in spring maize. *Crop J*, 2(5), pp.297–307.

Cheng W., Coleman D.C. & Box J.E., 1990. Root dynamics, production and distribution in agrosystems on the Georgia Piedmont using minirhizotrons. *J Appl Ecol*, 27(2), pp.592–604.

Cornuéjols A., 2002. Une nouvelle méthode d'apprentissage : Les SVM. Séparateurs à vaste marge. *Bulletin de l'AFIA*, 51.

Dale L.M. et al., 2012. Chemometric tools for NIRS and NIR Hypespectral Imaging. *Bulletin UASVM Agriculture*, 69(1), pp.70–76.

Dale L.M. et al., 2013. Hyperspectral Imaging Applications in Agriculture and Agro-Food Product Quality and Safety Control: A Review. *Appl Spectrosc Rev*, 48(2), pp.142–159.

Degrune F. et al., 2016. No favorable effect of reduced tillage on microbial community diversity in a silty loam soil (Belgium). *Agr Ecosyst Environ*, 224, pp.12–21.

De Mendiburu F., Simon R., 2015 *Agricolae* - Ten years of an open source statistical tool for experiments in breeding, agriculture and biology. *PeerJ PrePrints*.

Fernández Pierna J.A. et al., 2004. Combination of support vector machines (SVM) and near-infrared (NIR) imaging spectroscopy for the detection of meat and bone meal (MBM) in compound feeds. *J Chemometr*, 18(7–8), pp.341–349.

Fernández Pierna J.A. et al., 2012. NIR hyperspectral imaging spectroscopy and chemometrics for the detection of undesirable substances in food and feed. *Chemometr Intell Lab*, 117, pp.233–239.

Fernández Pierna J.A., Baeten V. & Dardenne P., 2006. Screening of compound feeds using NIR hyperspectral data. *Chemometr Intell Lab*, 84 (1–2), pp.114–118.

Hack H. et al., 1992. Einheitliche Codierung der hänologischen Entwicklungsstadien mono- und dikotyler Pflanzen – Erweiterte BBCH-Skala, Allgemein. *Nachrichtenbl. Deut. Pflanzenschutzd.* 44, 265-270.

Hauggaard-Nielsen H. & Jensen E.S., 2005. Facilitative root interactions in intercrops. *Plant and Soil*, 274(1–2), pp.237–250.

Huang G. -b. et al., 2012. Effects of different tillage systems on soil properties, root growth, grain yield, and water use efficiency of winter wheat (*Triticum aestivum* L.) in arid northwest China. *J Integr Agric*, 11(8), pp.1286–1296.

Izumi Y., Uchida K. & Iijima M., 2004. Crop production in successive wheat-soybean rotation with no-tillage practice in relation to the root system development. *Plant Prod Sci*, 7(3), pp.329–336.

Kusumo B.H. et al., 2011. Measuring carbon dynamics in field soils using soil spectral reflectance: prediction of maize root density, soil organic carbon and nitrogen content. *Plant Soil*, 338(1–2), pp.233–245.

Kusumo B.H. et al., 2010. Predicting pasture root density from soil spectral reflectance: field measurement. *Eur J Soil Sci*, 61(1), pp.1–13.

Li L. et al., 2006. Root distribution and interactions between intercrop species. *Oecologia*, 147, pp.280–290.

Naumann A., Heine G. & Rauber R., 2010. Efficient discrimination of oat and pea roots by cluster analysis of Fourier transform infrared (FTIR) spectra. *Field Crops Research*, 119, pp.78–84.

Majdi H., 1996. Root sampling methods - application and limitations of the minirhizotron technique. *Plant Soil*, 185, pp.255–258.

Meinen C. & Rauber R., 2015. Root discrimination of closely related crop and weed species using FT MIR-ATR spectroscopy. *Frontiers in Plant Science*, 6(September), pp.1–13

Morell F.J. et al., 2011. Root growth of barley as affected by tillage systems and nitrogen fertilization in a semiarid mediterranean agroecosystem. *Agron J*, 103(4), pp.1270–1275.

Munkholm L.J., Hansen E.M. & Olesen J.E., 2008. The effect of tillage intensity on soil structure and winter wheat root/shoot growth. *Soil Use and Management*, 24(4), pp.392–400.

Muñoz-Romero V. et al., 2010. Monitoring wheat root development in a rainfed vertisol: Tillage effect. *European Journal of Agronomy*, 33(3), pp.182–187.

Picon-Cochard C. et al., 2009. Use of near-infrared reflectance spectroscopy to predict the percentage of dead versus living grass roots. *Plant Soil*, 317(1–2), pp.309–320.

Plaza-Bonilla D. et al., 2014. Winter cereal root growth and aboveground–belowground biomass ratios as affected by site and tillage system in dryland Mediterranean conditions. *Plant Soil*, 374(1–2), pp.925–939.

Qin R., Stamp, P. & Richner, W., 2004. Impact of tillage on root systems of winter wheat. *Agron J*, 96(6), pp.1523–1530.

Ramirez-Garcia J. et al., 2015. Intercropping effect on root growth and nitrogen uptake at different nitrogen levels. *Journal of Plant Ecology*, 8(4), pp.380–389.

Rasmussen I.S., Dresbøll D.B. & Thorup-Kristensen K., 2015. Winter wheat cultivars and nitrogen (N) fertilization-Effects on root growth, N uptake efficiency and N use efficiency. *European Journal of Agronomy*, 68, pp.38–49.

R Core Team, 2017. R : A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Rewald B. et al., 2012. Root taxa identification in plant mixtures - current techniques and future challenges. *Plant and Soil*, 359, pp.165–182.

Rewald B. & Meinen C., 2013. Plant roots and spectroscopic methods - analyzing species, biomass and vitality. *Front Plant Sci*, 4, p.393.

Roumet C. et al., 2006. Quantifying species composition in root mixtures using two methods: near-infrared reflectance spectroscopy and plant wax markers. *New Phytol*, 170(3), pp.631–638.

Rumbaugh M.D., Clarck D.H. & Pendery B.M., 1988. Determination of root mass ratios in alfalfa-grass mixtures using near infrared reflectance spectroscopy. *J Range Manage*, 41(6), pp.488–490.

Shahin M.A., Symons S.J. & Hatcher D.W., 2014. Quantification of mildew damage in soft red winter wheat based on spectral characteristics of bulk samples: a comparison of visible-near-infrared imaging and near-infrared spectroscopy. *Food Bioprocess Tech*, 7(1), pp.224–234.

Strudley M.W., Green T.R. & Ascough J.C., 2008. Tillage effects on soil hydraulic properties in space and time: State of the science. *Soil and Tillage Research*, 99(1), pp.4–48.

Vermeulen P. et al., 2012. Online detection and quantification of ergot bodies in cereals using near infrared hyperspectral imaging. *Food Addit Contam A*, 29(2), pp.232–240.

Wang W. & Paliwal J., 2007. Near-infrared spectroscopy and imaging in food quality and safety. *Sens Instrum Food Qual Saf*, 1(4), pp.193–207.

Wang C. et al., 2014. Effects of different irrigation and nitrogen regimes on root growth and its correlation with above-ground plant parts in high-yielding wheat under field conditions. *Field Crops Research*, 165, pp.138–149.

Citation : Eylenbosch, D. et al., 2018. Utilisation de l'imagerie hyperspectrale proche infrarouge combinée aux outils de la chimie dans l'étude de systèmes racinaires. In J. Mathieu & P. Gate, eds. *Phloèmes - Actes du congrès*. Paris, Arvalis - Institut du végétal, pp. 226–236.

2. Validation de la régression permettant la prédiction de la masse de racines de froment

2.1. Matériels et méthodes

Pour valider la droite de régression utilisée pour prédire la masse de racines de froment d'hiver sur base d'images hyperspectrales, des échantillons de racines prélevés dans l'essai SOLRESIDUS au cours de la saison culturale 2015-2016 ont été utilisés. Cette validation a été possible suite au tri manuel des racines et des résidus de cultures présents au sein d'une vingtaine d'échantillons et à la pesée des racines extraites de ces échantillons.

Les échantillons de racines de froment utilisés pour valider cette droite avaient une masse comprise entre 0,004 et 0,11 g. Cette gamme de variation des masses d'échantillons de racines était équivalente à celle de la majorité des échantillons utilisés pour calibrer la droite de régression précédemment utilisée. En effet, parmi les 30 échantillons utilisés pour établir l'équation de la droite de régression permettant de prédire la masse racinaire au sein des échantillons de sol de l'essai SOLRESIDUS (Equation 4 présentée à la section précédente), 28 échantillons avaient une masse comprise entre 0,009 et 0,146 g. Les deux échantillons restant étaient des échantillons combinés de racines. Cette combinaison d'échantillons avait été réalisée afin de s'assurer que la relation entre le nombre de pixels détectés comme étant des racines sur des images hyperspectrales et la masse de racines était toujours valable sur des échantillons contenant un plus grand nombre de racines. La masse de ces échantillons était de 0,26 et 0,29 g. Sans ces deux échantillons, le coefficient de détermination de la régression était de 0,96. Lorsque ces deux échantillons étaient utilisés pour calculer l'équation de la droite de régression, le coefficient de détermination était de 0,98. Dans les deux cas, le coefficient de détermination avait une valeur supérieure à 0,91 et la régression pouvait donc être considérée comme étant excellente (Saeys et al. 2005). La validation de la régression permettant de faire de la quantification racinaire a été réalisée sur la régression dont le domaine d'application était réduit.

Afin de tenir compte du nombre variable de pixels de racines détectés sur une image hyperspectrale selon la manière dont les échantillons étaient disposés sous la caméra lors de l'acquisition des images, la plupart des échantillons ont été acquis deux ou trois fois avec la caméra hyperspectrale. Entre chaque acquisition, la disposition de l'échantillon sur la bande transporteuse était modifiée. En pratique, un échantillon était de plus en plus étendu sur la bande transporteuse au cours des acquisitions. À une même masse de racines correspondait donc plusieurs quantités de pixels²⁶. Pour la construction de la régression, ce nombre de pixels était compris

²⁶ L'écart relatif entre le nombre minimum et maximum de pixels de racines prédits pour un même échantillon sur ses différentes images, exprimé par rapport au nombre moyen de pixels de racines prédits pour cet échantillon, a été calculé à l'échelle de l'ensemble des données. Il était compris entre 1 et 17 % (n = 41, moyenne = 5 %).

entre 1.359 et 38.158. Pour sa validation, ce nombre de pixels était compris entre 1.491 et 28.591.

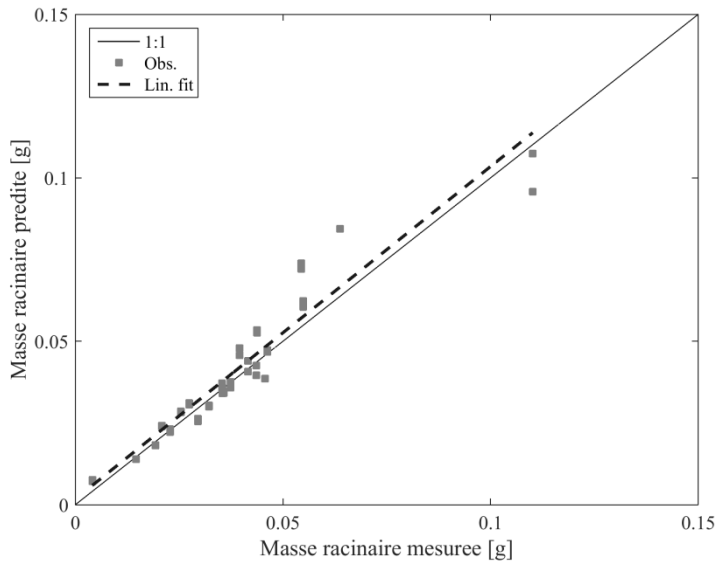


Figure 51: Relation entre les masses de racines mesurées et prédites (sur base de la régression linéaire reliant le nombre de pixels classés comme étant des racines sur les images hyperspectrales proche infrarouge et la masse de ces échantillons de racines). La régression utilisée pour prédire la masse des racines (ligne continue) a un r^2 de 0,96. La régression obtenue durant la validation (ligne discontinue) a un r^2 de 0,90 et une pente de 1,02. Pour la plupart des échantillons de racines, deux ou trois images hyperspectrales ont été acquises en modifiant la dispersion de l'échantillon sur la bande transporteuse entre les acquisitions ce qui explique un nombre de masses prédites supérieur au nombre de masses mesurées.

2.2. Résultats et discussion

2.2.1. Validation de la régression

Les résultats de la validation ont montré une bonne prédiction de la masse des nouveaux échantillons de racines. La régression linéaire liant les valeurs des masses réelles et prédites avait une pente de 1,02, un coefficient de détermination (r^2) de 0,90, une erreur quadratique moyenne de prédiction (RMSEP) de 0,0073 et un rapport de l'écart-type sur l'erreur quadratique moyenne (RPD) de 2,85 (Figure 51). Avec de telles valeurs obtenues pour le r^2 et le RPD, le modèle pouvait être considéré comme bon selon les critères rapportés par Saeys et al. (2005).

Au cours de cette validation, deux échantillons contenant des racines plus épaisses, obtenues dans l'horizon supérieur du sol (0-10 cm), ont également été utilisés pour tester la régression. Ces échantillons ont moins bien été prédits que les échantillons se trouvant dans la gamme de masses racinaires utilisée pour calibrer la régression. Leurs masses réelles étaient de 0,17 et 0,18 g et leurs masses prédites étaient de 0,14

et 0,10 g respectivement. Plusieurs raisons peuvent être évoquées pour expliquer la moins bonne prédiction de la masse de ces échantillons : (i) le fait que la régression ait été calibrée majoritairement sur des racines fines, (ii) un rapport surface/masse différent pour ces racines ou (iii) une modification du profil spectral de ces racines, dû à la forme de ces racines et à leur manière de réfléchir la lumière, et qui pourrait modifier leur reconnaissance par les modèles chimiométriques.

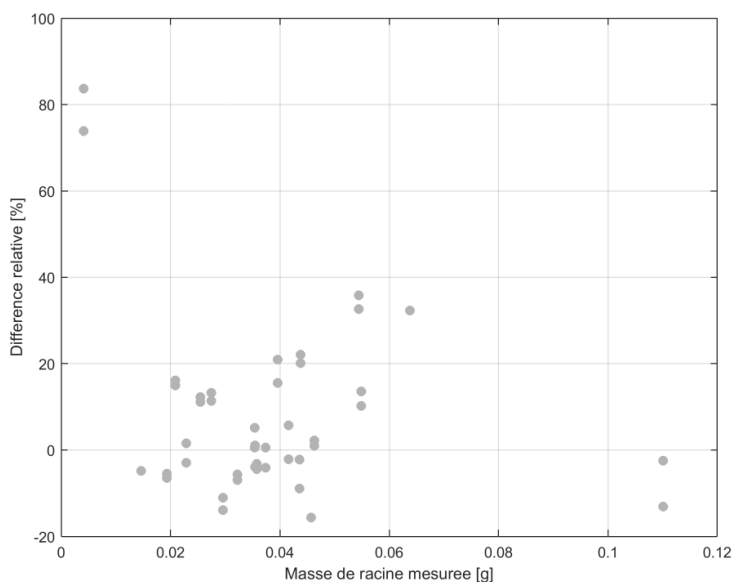


Figure 52: Distribution de la différence relative (%) entre la masse racinaire prédite et la masse racinaire réelle en fonction de la masse réelle des échantillons de racines (g). Les échantillons représentés dans cette figure ont été utilisés pour valider la régression linéaire permettant de relier un nombre de pixels classés comme étant des racines avec la masse réelle de ces racines.

Lors de cette validation, la différence relative entre la masse prédite et la masse réelle des échantillons a été calculée. La relation entre cette différence relative et la masse réelle des échantillons de racine est illustrée à la Figure 52. Cette figure montre une forte variabilité de l'erreur relative de prédiction. Une partie de cette variabilité est due aux répétitions des acquisitions d'images sur les échantillons avec une répartition différente des racines entre chaque acquisition. Pour les échantillons de faible masse, la masse racinaire prédite sur base du nombre de pixels de racines détectés sur les images hyperspectrales est la plupart du temps surestimée. Cette différence entre la masse prédite et la masse réelle de racines pourrait être expliquée par la taille des racines, la taille des pixels de l'image, la disposition des échantillons lors de l'acquisition des images et la précision des modèles de discrimination. Ces erreurs de quantification seront discutées au Chapitre VIII.

2.2.2. Comparaison avec d'autres méthodes de quantification racinaire

À notre connaissance, la masse racinaire n'a jamais été quantifiée sur des images hyperspectrales. Il est donc difficile de comparer nos résultats avec d'autres études. Cependant, plusieurs études ont comparé la masse racinaire mesurée dans des carottes de sol après lavage et récupération des racines sur un tamis avec des méthodes indirectes de quantification racinaire. Les résultats obtenus pour les droites de régression dans ces différentes études peuvent être comparés avec nos résultats.

Xu et al. (2016) et Kusumo et al. (2010) ont déterminé la densité racinaire dans des échantillons de sol en combinant la méthode de prélèvement de carottes de sol avec de la spectroscopie visible-proche infrarouge. Dans ces études, la quantité de racines présentes au sein des carottes de sol était directement estimée sur base de spectres acquis sur la surface horizontale des sections de carottes de sol. Les meilleurs coefficients de détermination qu'ils ont obtenus lors de la validation de la relation entre les densités racinaires mesurées avec la méthode classique, basée sur le lavage et la pesée des racines, et les densités prédites sur base de la réflectance spectrale du sol étaient de 0,88 (Xu et al. 2016) et 0,92 (Kusumo et al. 2010). Les valeurs de RMSEP qu'ils ont obtenues lors de la validation étaient respectivement 4,28 et de 2,32. Enfin, les valeurs de RPD obtenues étaient de 2,83 et de 2,25. Dans ces deux cas, les prédictions pouvaient être considérées comme bonnes (Saeys et al. 2005).

Wasson et al. (2016) ont développé un système similaire, mais basé cette fois sur de la spectroscopie en fluorescence, pour remplacer le comptage visuel nécessaire dans la méthode *core break*. Le coefficient de détermination de la corrélation qu'ils ont obtenue entre la densité de masse racinaire réelle et celle prédite sur base des mesures en fluorescence était de 0,68. Cette valeur indique que cette technique permettait de faire des prédictions quantitatives approximatives (Saeys et al. 2005).

Dans les trois cas d'étude présentés ci-dessus, la détermination de la densité racinaire était donc basée sur des données spectrales acquises sur des carottes de sol entières. Une autre comparaison peut être réalisée avec la corrélation entre la biomasse racinaire mesurée et celle prédite sur base d'images acquises dans des minirhizotrons. Dans ce cas, Jose et al. (2001) ont obtenu un r^2 de 0,67 sous une culture de maïs et un r^2 de 0,87 avec des racines de noyer (*Juglans nigra* L.). Dans ce cas-ci, la biomasse racinaire était donc déterminée sur base d'une surface de racine mesurée grâce à la réalisation d'observations au sein de minirhizotrons.

La méthode que nous proposons est en quelque sorte une combinaison de ces techniques : la spectroscopie est utilisée pour détecter les racines dans les échantillons de sol lavés et la surface de racines prédite sur les images hyperspectrales permet ensuite de quantifier ces racines.

Au vu des résultats obtenus dans ces différentes études prises à titre de comparaison, l'estimation de la masse racinaire sur base de l'imagerie hyperspectrale proche infrarouge semble avoir une précision comparable à d'autres méthodes indirectes et celle-ci peut être considérée comme bonne.

2.3. Conclusions

L'application de l'équation de la régression linéaire sur un nouvel ensemble d'échantillons de masses connues a montré que cette régression permettait de prédire correctement la masse de la majorité de ces échantillons sur base du nombre de pixels classés comme étant des racines sur les images hyperspectrales. En comparant les résultats obtenus dans notre étude avec d'autres études sur la quantification racinaire basée sur l'imagerie ou la spectroscopie, nous avons montré que nos prédictions étaient d'aussi bonne qualité.

Cette validation a néanmoins mis en évidence la difficulté de prédire correctement la masse de certains échantillons ayant une masse plus élevée et contenant des racines avec un diamètre plus important que celles utilisées pour établir la régression. Cette validation met donc en évidence que les résultats obtenus avec cette méthode de quantification racinaire doivent être interprétés avec prudence. Tout d'abord car des confusions apparaissent entre les spectres des différents éléments présents sur les images hyperspectrales. Ces confusions peuvent créer soit des sous-estimations, soit des surestimations des quantités de racines présentes au sein des échantillons. Ensuite parce que la régression permettant la transformation d'un nombre de pixels de racines détecté sur l'image hyperspectrale en masse racinaire ne permet pas de déterminer parfaitement la masse de certains échantillons. Ces erreurs de prédictions sont en parties dues aux conditions d'acquisition des images hyperspectrale et seront discutées au Chapitre VIII.

3. Utilisation de l'imagerie hyperspectrale et de la chimiométrie en recherche agronomique

3.1. Introduction

L'article repris dans la première section de ce chapitre présentait l'effet du travail du sol sur le développement racinaire du froment d'hiver. Les résultats présentés ont été donnés, à titre illustratif, pour deux dates de prélèvement au cours de la saison culturale 2011-2012 dans l'essai de longue durée SOLRESIDUS. Cette section-ci vise à fournir un plus grand nombre de résultats de manière à illustrer les potentialités de la combinaison de l'imagerie hyperspectrale à la méthode d'étude racinaire par carottage de sol en recherche agronomique. Elle s'intéressera plus spécifiquement à la quantification racinaire dans le but de caractériser la distribution spatiale et temporelle des densités racinaires de froment selon la profondeur du travail du sol et le mode de gestion des résidus de la culture précédente. Les résultats ayant été obtenus sur base de la prédiction de la nature des pixels d'images hyperspectrales proche infrarouge et sur base de la régression permettant de relier le nombre de pixels de racines détectés à une masse de racines, les données présentées dans cette section sont donc des **valeurs prédites** de densités de masses racinaires.

3.2. Matériels et méthodes

3.2.1. Description des modalités de l'essai SOLRESIDUS

L'essai de longue durée SOLRESIDUS, présenté dans l'avant-propos de cette thèse, est mené selon deux modalités de travail du sol et deux modalités de gestion des résidus de cultures. Les modalités de travail du sol et de gestion des résidus sont croisées si bien que l'essai comporte finalement quatre modalités d'étude. Ces modalités sont testées dans un essai en split-plot. Le travail du sol est soit un labour conventionnel réalisé sur environ 25 cm de profondeur, soit un travail du sol réduit réalisé sur environ 10 cm de profondeur. Les modalités de gestion des résidus de cultures sont l'exportation des résidus de la culture précédente ou la restitution au sol de ces résidus après broyage. Ces modalités et le matériel agricole utilisé pour le travail du sol ont été décrits en détails par Hiel et al. (2018). La culture précédant la culture de froment d'hiver étudiée en 2011-2012 sur l'essai SOLRESIDUS était également une culture de froment. En 2015-2016, la culture précédant la culture de froment était une culture de maïs grain.

3.2.2. Protocole de prélèvement des échantillons

Au cours de la saison culturale 2011-2012, quatre prélèvements de racines ont été réalisés aux stades levée, redressement, deux nœuds et récolte de la culture de froment. Seuls les prélèvements réalisés un mois après le semis (levée de la culture) et au stade phénologique "deux nœuds" de la culture ont été présentés dans la section précédente. L'ensemble des prélèvements réalisés au cours de cette saison a permis d'étudier la distribution temporelle et spatiale des racines de froment d'hiver dans trois horizons de sol (0-10, 10-20 et 20-30 cm de profondeur) et de mesurer l'effet de l'exportation des résidus de cultures sur le développement racinaire.

Au cours de la saison culturale 2015-2016, une seconde campagne de prélèvement de racines a été réalisée dans cet essai de longue durée. Durant cette saison, deux prélèvements de racines ont été réalisés : le premier prélèvement a été réalisé mi-mai, au stade redressement de la culture, et le second mi-juillet, lors du remplissage des grains. À ces deux dates, les prélèvements ont été réalisés à plus grande profondeur que durant la saison 2011-2012 : les carottes de sol étaient prélevées sur 90 cm de profondeurs et divisées en six horizons de sol (0-10, 10-20, 20-30, 30-50, 50-70 et 70-90 cm de profondeur).

Pour chaque prélèvement, un échantillon de sol a été prélevé au sein de chacune des quatre répétitions des modalités de l'essai. Ces échantillons étaient constitués de 4 ou 10 carottes de sol prélevées pour moitié dans les lignes et pour moitié dans les interlignes de la culture. Une synthèse du protocole de prélèvement est donnée au Tableau 7.

Tableau 7: Protocole de prélèvements des échantillons de racines dans l'essai SOLRESIDUS. Cet essai comporte deux modalités de travail du sol (conventionnel, sur 25 cm de profondeur (LC), et réduit, sur 10 cm de profondeur (LR)) et deux modalités de restitution des résidus de la culture précédente (incorporation au sol (IN) et exportation (OUT)). Les modalités de travail du sol et de restitution des résidus sont combinées pour donner au final quatre modalités expérimentales. Les prélèvements de racines ont été réalisés en nombre égal dans la ligne (L) et dans l'interligne (IL) de la culture de froment d'hiver.

Saisons culturales	Stades des cultures	Dates de prélèvements	Modalités	Horizons sol (cm)	Nombre de carottes
2011-2012	levée	17 novembre	LC-IN	0-10 10-20 20-30	5*L + 5* IL
	redressement	04 avril	LC-OUT		
	2 nœuds	08 mai	LR-IN		
	récolte	08 août	LR-OUT		
2015-2016	redressement	12 mai	LC-IN	0-10	2*L + 2* IL
			LC-OUT	10-20	
	remplissage	13 juillet	LR-IN	20-30	
			LR-OUT	30-50 50-70 70-90	

3.2.3. Traitement des échantillons et des résultats

Les échantillons de sol ont été traités exactement avec le même matériel et selon la même méthode que ce qui a été décrit au point 1.3 de ce chapitre. Il en va de même pour les analyses statistiques réalisées sur les résultats.

Pour la saison culturale 2011-2012, une prédiction de la masse de résidus de cultures présents au sein des échantillons de sol a également été réalisée. En effet, les modèles chimiométriques mis au point dans le cadre de cette thèse permettant la

discrimination de racines et de résidus de cultures, il était donc également possible de quantifier ces résidus sur les images hyperspectrales. Comme pour les racines, une régression linéaire entre un nombre de pixels prédits comme étant des résidus de cultures (dans ce cas-ci des pailles de froment d'hiver) et la masse de ces résidus a été établie ($r^2 = 0,97$). Cette relation a été utilisée pour faire la quantification des résidus de cultures dans les trois horizons de sol étudiés au cours de la saison 2011-2012 dans l'essai SOLRESIDUS.

3.3. Résultats et discussion

3.3.1. Effet du mode de gestion des résidus de cultures

Au cours des quatre prélèvements réalisés en 2011-2012, la densité racinaire mesurée dans le premier horizon du sol (0-10 cm) au mois de novembre, un mois après le semis, est la seule situation pour laquelle un effet statistiquement significatif du mode de gestion des résidus de cultures a été mis en évidence sur le développement du système racinaire (Figure 53). À cette date, la densité de masse racinaire prédite dans les modalités avec restitution des résidus au sol était plus élevée. L'interprétation de cette observation doit cependant être faite avec prudence. En effet, nous avons montré qu'une certaine confusion existe entre les racines et les résidus de cultures au niveau des modèles de discrimination utilisés et qu'un certain nombre de pixels de résidus de cultures peuvent dès lors être prédits comme étant des racines, principalement en bordure des résidus. La présence d'une plus grande quantité de résidus dans cette situation pourrait donc expliquer en partie cette valeur de densité racinaire prédite sur base des images hyperspectrales. Cet effet n'était d'ailleurs observable qu'en travail du sol réduit, lorsque les résidus de la culture précédente étaient concentrés dans les dix premiers cm du sol. En labour conventionnel, les résidus de cultures étaient répartis de manière homogène dans la couche de sol travaillée et la présence de ces résidus n'influçait pas la densité racinaire.

Pour l'ensemble des autres situations étudiées au cours de cette saison culturale (horizons et dates de prélèvement), aucun effet du mode de gestion des résidus de cultures n'a été mis en évidence sur la densité racinaire prédite. Au sein d'une même modalité de travail du sol, les densités de racine prédite dans les modalités avec ou sans restitution des résidus de la culture précédente étaient donc statistiquement identiques. C'est pourquoi ce facteur n'est plus considéré dans le reste de la présentation des résultats.

Pour les deux prélèvements réalisés au cours de l'année culturale 2015-2016, aucun effet statistiquement significatif du mode de gestion des résidus de cultures n'a été observé sur les densités racinaires prédites.

3.3.2. Effet de la profondeur du travail du sol

Pour la saison culturale 2011-2012, l'effet le plus marqué du travail du sol sur la densité de masse racinaire prédite au sein des trois horizons de sol étudiés a été observé en début de culture, lors du prélèvement réalisé mi-novembre 2011, un mois après le semis. À cette date, la densité racinaire prédite du froment d'hiver était plus importante dans le premier horizon du sol en travail du sol réduit qu'en labour et

cette tendance s'inversait dans les deux autres horizons de sol (Figure 53). Un effet marqué du travail du sol apparaissait donc à cette date en travail du sol réduit alors que le développement racinaire dans les trois horizons de sol était homogène en labour conventionnel.

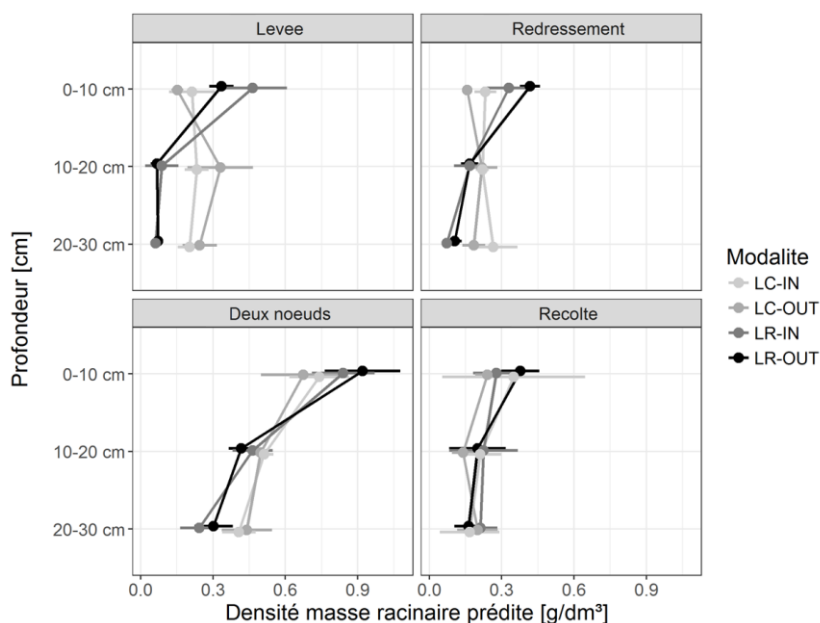


Figure 53: Distribution de la densité de masse racinaire prédite (g/dm^3) d'une culture de froment d'hiver conduite en labour conventionnel (LC) ou travail du sol réduit (LR) avec ou sans restitution au sol des résidus de la culture précédente (IN ou OUT). Les prélèvements de racines ont été réalisés à quatre stades culturaux au cours de la saison culturale 2011-2012 et sur trois horizons de sol.

Cet effet du travail du sol sur les densités racinaires prédites dans les trois horizons de sol a encore été observé lors des deux prélèvements suivants. Lors du prélèvement réalisé en avril, les tendances étaient identiques, mais les différences entre les deux modes de travail du sol étaient moins marquées. Au stade deux nœuds de la culture (prélèvement du mois de mai), la densité racinaire prédite en travail du sol réduit était toujours plus importante qu'en labour dans le premier horizon de sol et plus faible dans les deux horizons inférieurs. À cette date, contrairement aux prélèvements précédents, une différence de densité racinaire prédite entre les deux premiers horizons du sol apparaissait également en labour. La densité racinaire prédite était donc la plus élevée dans le premier horizon de sol, quelle que soit la modalité de travail du sol.

Lors du dernier prélèvement racinaire, juste avant la récolte de la culture, plus aucun effet du travail du sol n'était observable sur la densité racinaire prédite : les densités racinaires prédites en labour et travail réduit étaient identiques à toutes les profondeurs.

Lorsque les analyses statistiques ont été réalisées en combinant les densités racinaires prédites des trois horizons de sol, aucun effet du travail du sol n'a été observé, quel que soit le stade de développement de la culture. Le travail du sol modifiait donc la répartition du système racinaire dans les 30 premiers cm du sol, mais pas la densité racinaire prédite dans cet horizon de sol.

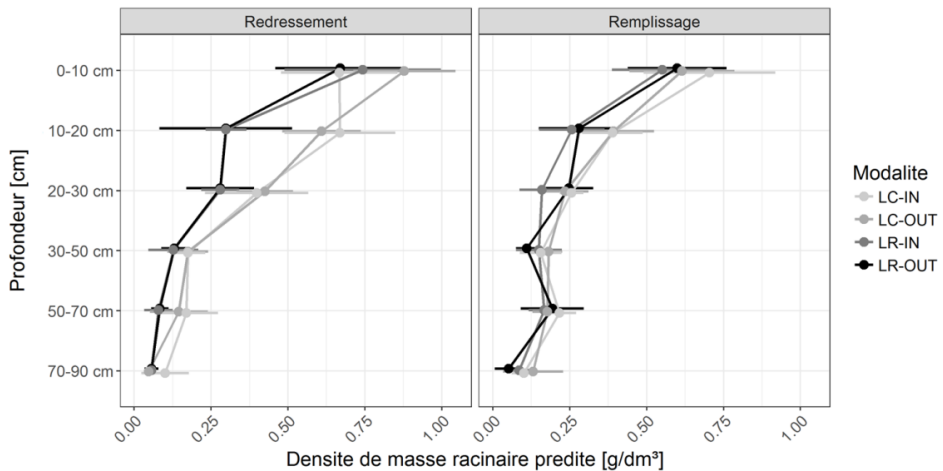


Figure 54: Distribution de la densité de masse racinaire prédite (g/dm^3) d'une culture de froment d'hiver conduite en labour conventionnel (LC) ou travail du sol réduit (LR) avec ou sans restitution au sol des résidus de la culture précédente (IN ou OUT). Les prélèvements de racines ont été réalisés à deux dates au cours de la saison culturale 2015-2016 (stades redressement de la culture et remplissage des grains) et sur six horizons de sol.

Pour l'année 2016, la seule différence entre les modes de travail du sol en termes de densités racinaires prédites par horizon de sol a été observée au stade redressement de la culture (Figure 54). Selon l'analyse statistique réalisée sur nos mesures, la densité racinaire prédite était, à ce stade de la culture, en moyenne plus élevée en labour conventionnel qu'en travail du sol réduit dans les horizons de sol compris entre 10-20, 20-30 et 50-70 cm de profondeur. Pour les autres horizons de sol étudiés et pour l'autre date de prélèvement de racines, aucune différence significative entre le labour et le travail du sol réduit n'a été mise en évidence lorsque les horizons étaient considérés séparément. Cependant, contrairement aux observations réalisées en 2011-2012, la densité de racine prédite dans l'entièreté du profil de sol étudié en 2016 était influencée par le travail du sol : aux deux dates de prélèvement, une plus grande densité de racine était prédite lorsque le sol était

labouré que lorsqu'il n'était travaillé que sur 10 cm de profondeur. Cette différence se marquait au stade redressement dans les horizons de sol travaillé (10-20 et 20-30 cm de profondeur), mais aussi plus en profondeur (50-70 cm de profondeur).

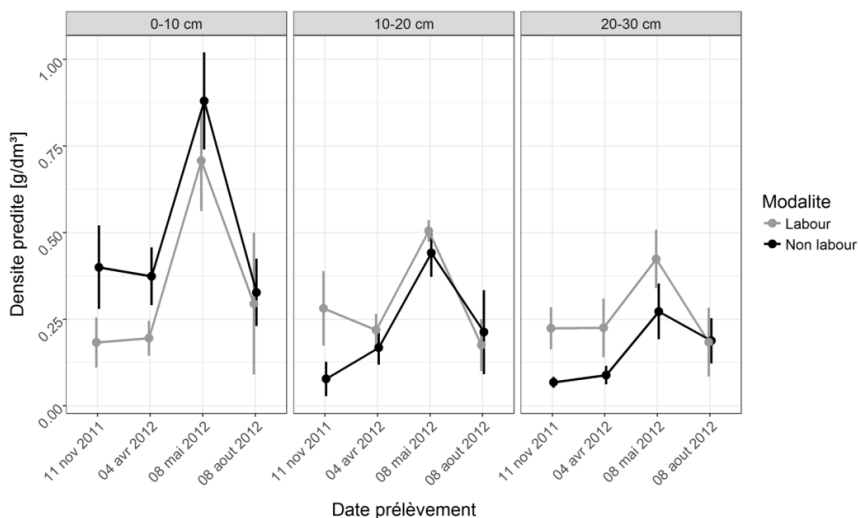


Figure 55: Évolution de la densité de masse racinaire prédite (g/dm^3) d'une culture de froment d'hiver entre quatre dates de prélèvement au cours de la saison culturale 2011-2012 et sur trois horizons de sol (0-10, 10-20 et 20-30 cm de profondeur). Les prélèvements ont été réalisés un mois après le semis, aux stades tallage et 2 nœuds de la culture et quelques jours avant la récolte. Les modalités de travail du sol et de gestion des résidus dans lesquels les racines ont été prélevées sont un labour conventionnel ou un travail du sol réduit (non labour).

3.3.3. Dynamique de développement racinaire

Sur base des prélèvements racinaires réalisés au cours de la saison culturale 2011-2012, les résultats obtenus avec l'imagerie hyperspectrale et les modèles chimiométriques ont mis en évidence une densité de masse racinaire prédite plus importante au début du mois de mai, au stade deux nœuds de la culture de froment d'hiver, qu'aux autres dates de prélèvements et quel que soit la modalité d'essai (Figure 55). À ce stade, la densité racinaire prédite était la plus importante dans le premier horizon du sol (0-10 cm) aussi bien en labour conventionnel qu'en travail réduit. Entre ce prélèvement et celui réalisé juste avant la récolte de la culture (08 août 2012), une diminution générale de la densité racinaire prédite a été observée. Une telle décroissance en fin de culture confirme ce qui avait été observé dans d'autres études (Plaza-Bonilla et al. 2014; Wang et al. 2014; Huang et al. 2012; Muñoz-Romero et al. 2010; Kätterer et al. 1993).

Aucun prélèvement n'a été réalisé au stade floraison au cours de cette saison de culture. Or, d'après les observations rapportées dans plusieurs études menées sur le froment d'hiver, le développement racinaire maximal aurait dû être mesuré à ce

stade de développement de la culture (Plaza-Bonilla et al. 2014; Wang et al. 2014; Huang et al. 2012; Kätterer et al. 1993). Nous pouvons donc raisonnablement poser l'hypothèse que les densités racinaires de la culture de froment d'hiver cultivée en 2011-2012 ont dû atteindre des valeurs plus élevées que celles prédites au stade deux nœuds de la culture.

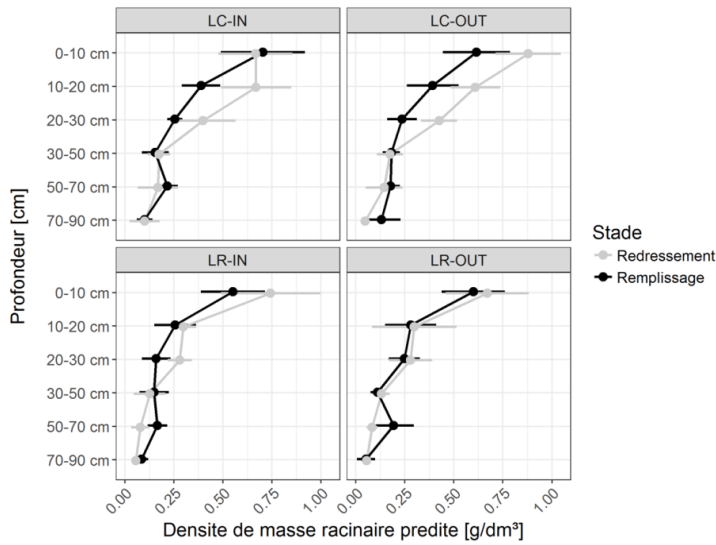


Figure 56: Évolution de la densité de masse racinaire prédite (g/dm^3) d'une culture de froment d'hiver entre deux dates de prélèvement au cours de la saison culturale 2015-2016 et sur six horizons de sol. Les dates de prélèvement correspondent aux stades redressement de la culture et remplissage des grains. Les modalités de travail du sol et de gestion des résidus de cultures dans lesquels les racines ont été prélevées sont un labour conventionnel (LC) ou un travail du sol réduit (LR) avec (IN) ou sans (OUT) restitution des résidus de la culture précédente.

Pour la seconde campagne de prélèvement réalisée en 2015-2016, les densités racinaires prédites lors du redressement de la culture et lors du remplissage des grains étaient statistiquement égales quel que soit la modalité étudiée. Seule une différence significative a été observée dans l'horizon de sol compris entre 10 et 20 cm de profondeur en labour avec restitution des résidus de la culture précédente. La densité racinaire prédite y avait diminué entre les stades redressement de la culture et remplissage du grain. Sans qu'aucune différence statistiquement significative ne puisse être mise en évidence, les densités prédites semblaient globalement plus élevées au stade redressement que lors du remplissage des grains (Figure 56).

3.3.4. Distribution spatiale des résidus de cultures et dynamique de dégradation

Les observations réalisées sur les résidus de cultures ont mis en évidence un effet prononcé du travail du sol sur leur répartition (Figure 57). En travail du sol réduit,

les densités prédites de masses de résidus de cultures étaient plus élevées dans le premier horizon de sol (0-10 cm de profondeur). En labour, ces densités de résidus de cultures étaient par contre équivalentes dans les trois horizons de sol. La dégradation des résidus de cultures a également pu être suivie : lors du dernier prélèvement, juste avant la récolte de la culture, la densité de résidus de la culture précédente prédite dans le sol était quasi nulle quel que soit le mode de gestion de ces résidus. La dégradation des résidus de cultures était donc importante dans la parcelle agricole échantillonnée, aussi bien en labour qu'en travail du sol réduit.

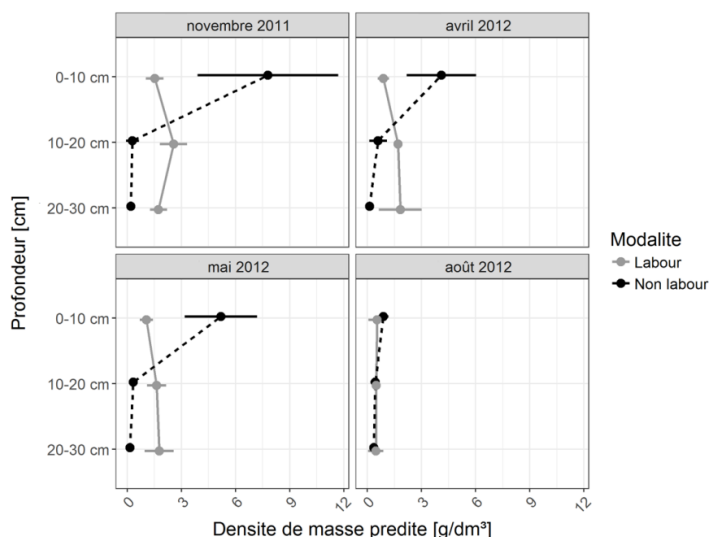


Figure 57: Évolution des densités de masses de résidus de cultures prédites (g/dm^3) entre quatre dates de prélèvement au cours de l'année culturale 2011-2012 et sur trois horizons de sol (0-10, 10-20 et 20-30 cm de profondeur). Les modalités de travail du sol dans lesquels les résidus de cultures ont été prélevés sont un labour conventionnel sur 25 cm de profondeur ou un travail du sol réduit sur 10 cm de profondeur (non labour). Cette figure présente le cas où les résidus de la culture précédente sont restitués au sol.

3.4. Conclusions

Les résultats obtenus sur les prélèvements racinaires réalisés au cours de deux saisons culturales au sein de l'essai SOLRESIDUS ont démontré que l'imagerie hyperspectrale proche infrarouge pouvait être utilisée dans des études agronomiques pour la prédiction de masses racinaires au sein d'échantillons de sol. Cette méthode a permis de mettre en évidence des différences entre les modalités de cultures, des horizons de sols et des dates de prélèvements et a donné des résultats plausibles par rapport à d'autres études.

Conclusions du Chapitre VI

Les études présentées dans ce chapitre ont permis de mettre en évidence plusieurs avantages et limites de l'imagerie hyperspectrale proche infrarouge utilisée dans l'étude de systèmes racinaires. Elles ont également permis de définir le domaine d'application de cette méthode de quantification.

Avantage:

Dans le chapitre précédent, nous avons montré que l'imagerie hyperspectrale et la chimiométrie permettait de discriminer des racines et des résidus de culture de froment d'hiver extraits d'échantillon de sol. Dans ce chapitre-ci, nous avons montré que la discrimination de nodosités, de racines de froment et de pois était également possible.

Grâce à la bonne relation existant entre un nombre de pixels détecté pour un élément sur une image hyperspectrale proche infrarouge et la masse de cet élément, nous avons montré qu'il était possible de prédire la masse de racines et de résidus de cultures sur des images hyperspectrales. Cette relation entre un nombre de pixels et une masse a été établie pour des racines de froment et de pois d'hiver ainsi que pour les résidus de culture de froment extraits d'échantillons de sol.

Par rapport à la méthode de référence qui consiste à trier manuellement les racines et les résidus de cultures et à les peser, cette méthode est plus rapide et plus facile d'utilisation. Ce constat est d'autant plus vrai lorsque des racines de différentes espèces sont présentes au sein d'un même échantillon. Le temps d'analyse d'un échantillon étant réduit, un plus grand nombre d'échantillons peut être traité sur un même laps de temps (par rapport à la méthode de référence) et les mesures peuvent donc être répétées plus facilement.

Limite:

Tout au long du traitement des échantillons, une série d'erreurs biaise la prédiction de la masse racinaire présente dans l'échantillon. Ces erreurs sont dues (i) à la manière dont l'échantillon est présenté sous la caméra hyperspectrale (recouvrement des éléments de l'échantillon), (ii) à la résolution de la caméra (beaucoup de racines ont une taille inférieure à la taille des pixels), (iii) à la confusion de certains spectres proche infrarouge par les modèles de discrimination et (iv) à la qualité de la relation entre un nombre de pixels de racines et la masse de racines.

Domaine d'application

Les études que nous avons menées avec l'imagerie hyperspectrale proche infrarouge pour estimer des masses racinaires dans le cadre de recherches agronomiques ont montré que cette méthode permet la comparaison de densités racinaires entre différents itinéraires techniques appliqués sur les cultures. Couplée à la méthode d'échantillonnage par carottage de sol, elle permet également d'étudier la stratification de systèmes racinaires au sein de profils de sol.

Sur base de ces résultats, nous remarquons que cette méthode ne doit donc pas être utilisée en espérant avoir une quantification précise des racines présentes au sein d'un échantillon, du moins dans l'état actuel de son développement. Mais cette méthode permet néanmoins de prédire des valeurs relatives et permet dès lors la

comparaison de densités racinaires au sein de systèmes de culture ou d'horizons de sol différents. Cette prédiction est plus rapide et moins fastidieuse que la quantification réalisée par la méthode de tri manuel et de pesée couramment utilisée avec le prélèvement de carottes de sol dans les études sur le développement de systèmes racinaires en conditions de pleins champs.

CHAPITRE VII

**Quantification des teneurs en
leghémoglobine au sein de nodosités de pois
protéagineux (*Pisum sativum* L.) sur base
de l'imagerie hyperspectrale proche
infrarouge et de la chimiométrie**

Introduction au Chapitre VII

Ce chapitre, publié dans le *Journal of Spectral Imaging*, présente la mise au point de la méthode permettant de faire la quantification de la leghémoglobine, une protéine impliquée dans la fixation symbiotique de l'azote atmosphérique par les légumineuses, au sein de nodosités de pois protéagineux par imagerie hyperspectrale proche infrarouge.

Cette étude a été imaginée après avoir mis en évidence les potentialités de l'imagerie hyperspectrale proche infrarouge et de la chimiométrie pour séparer les spectres proche infrarouge acquis sur des racines et des nodosités de pois protéagineux (résultats présentés dans le deuxième cas d'étude de l'article repris au Chapitre VI). Sur base de ces résultats, nous avons voulu savoir s'il était possible, sur une même image hyperspectrale, de discriminer les racines et les nodosités de pois protéagineux et de quantifier, au sein de chaque nodosité, une teneur en leghémoglobine. Cette protéine est en effet un excellent indicateur de l'activité fixatrice d'azote chez les légumineuses. Nous nous sommes également intéressés dans cette étude à la possibilité de quantifier cette leghémoglobine au sein des nodosités en se basant sur un nombre réduit de longueurs d'ondes de manière à réduire le temps de calcul nécessaire à la prédiction des images hyperspectrales et afin de pouvoir faire ces mesures avec des caméras multispectrales.

Cette étude a été menée sur des nodosités obtenues sur les systèmes racinaires de plantes de pois protéagineux cultivées en serre. Les teneurs en leghémoglobine des nodosités utilisées comme référence pour calibrer la régression PLS ont été mesurées par la méthode de la cyanmethémoglobine.

Quantification of leghaemoglobin content in pea nodules based on near-infrared hyperspectral imaging spectroscopy and chemometrics

Damien Eylenbosch^{a‡}, Benjamin Dumont^{b‡*}, Vincent Baeten^c, Bernard Bodson^a, Pierre Delaplace^d, Juan Antonio Fernández Pierna^c

^a AgrobioChem Department, Gembloux Agro-Bio Tech Faculty, University of Liège

^b AgrobioChem Department and Terra Teaching and Research Center, Gembloux Agro-Bio Tech Faculty, University of Liège

^c Walloon Agricultural Research Centre, Valorisation of Agricultural Products Department, Food and Feed Quality Unit

^d Terra Teaching and Research Center, Gembloux Agro-Bio Tech Faculty, University of Liège

*Corresponding author: Benjamin Dumont, benjamin.dumont@uliege.be

‡ These authors share senior authorship.

1. Abstract

Quantification des teneurs en leghémoglobine au sein de nodosités de pois protéagineux sur base de l'imagerie hyperspectrale proche infrarouge et de la chimométrie

La teneur en leghémoglobine au sein de nodosités est fortement liée à la quantité d'azote fixée par la symbiose entre légumineuses et rhizobiums. Elle est donc couramment mesurée de manière à évaluer les effets de paramètres favorisant la croissance tels que la fertilisation sur l'efficacité de la fixation symbiotique d'azote chez les légumineuses. La méthode faisant intervenir la cyanmethémoglobine est une méthode de référence pour la mesure de la concentration en leghémoglobine, mais cette méthode est chronophage, nécessite la réalisation précise et prudente d'opérations techniques et utilise du cyanure, un réactif toxique.

Une méthode alternative, plus rapide, plus facile et non destructive, basée sur l'imagerie hyperspectrale proche infrarouge, a été testée pour mesurer les teneurs en leghémoglobine au sein de nodosités séchées. Deux approches ont été évaluées: (i) une approche PLS appliquée sur l'entièreté du spectre acquis avec la caméra hyperspectrale et (ii) le potentiel de l'imagerie multispectrale qui a été testée suite à la présélection des longueurs d'ondes les plus pertinentes et à la construction d'un modèle de régression linéaire multiple. L'approche PLS a été testée sur des spectres moyens acquis d'une part sur des échantillons contenant plusieurs nodosités et acquis d'autre part séparément sur des nodosités individuelles. Les pois protéagineux (*Pisum sativum* L.) ont été cultivés en serre. Les nodosités ont été récoltées à quatre dates de manière à obtenir des variations au sein des teneurs en leghémoglobine.

Les teneurs en leghémoglobine mesurées avec la méthode de référence au sein de nodosités fraîches étaient comprises entre 1,4 et 4,2 mg de leghémoglobine par gramme de nodosité fraîche. Une régression PLS a été calibrée sur base des teneurs en leghémoglobine mesurées par la méthode de référence et les spectres moyens de nodosités sèches acquis avec une caméra hyperspectrale proche infrarouge. Sur un jeu de données prévu pour la validation, le modèle PLS a bien prédit les teneurs en leghémoglobine dans les échantillons de nodosités ($r^2 = 0,90$, RMSEP = 0,26). L'approche multispectrale a montré des performances similaires. Appliqué sur des nodosités individuelles, le modèle PLS a mis en évidence une large variabilité des teneurs en leghémoglobine au sein de nodosités récoltées sur la même plante. Ces résultats montrent que l'imagerie hyperspectrale proche infrarouge pourrait être utilisée pour mesurer rapidement et sans danger des teneurs en leghémoglobine au sein de nodosités de pois.

Quantification of leghaemoglobin content in pea nodules based on near-infrared hyperspectral imaging spectroscopy and chemometrics

Leghaemoglobin content in nodules is closely related with the amount of nitrogen fixed by the legume-rhizobium symbiosis. It is therefore commonly measured in order to assess the effect of growth-promoting parameters such as fertilization on the symbiotic nitrogen fixation efficiency of legumes. The cyanmethemoglobin method is a reference method in leghaemoglobin content quantification, but this method is time-consuming, requires accurate and careful technical operations and uses cyanide, a toxic reagent.

As a quicker, simpler and nondestructive alternative, a method based on near-infrared (NIR) hyperspectral imaging was tested to quantify leghaemoglobin in dried nodules. Two approaches were evaluated: (i) the PLS approach was applied to the full spectrum acquired with the hyperspectral device, and (ii) the potential of multispectral imaging was also tested through the preselection of the most relevant wavelengths and the building of a multiple linear regression model. The PLS approach was tested on mean spectra acquired from samples containing several nodules and acquired separately from individual nodules. Peas (*Pisum sativum* L.) were cultivated in the greenhouse. The nodules were harvested on four different dates in order to obtain variations in leghaemoglobin content.

The leghaemoglobin content measured with the cyanmethemoglobin method in fresh nodules ranged between 1,4 and 4,2 mg leghaemoglobin.g⁻¹ fresh nodule. A partial least square (PLS) regression model was calibrated on leghaemoglobin content measured with the reference method and mean NIR spectra of dried nodules acquired with a hyperspectral imaging device. On a validation dataset, the PLS model predicted well the leghaemoglobin content in nodule samples ($r^2 = 0,90$, RMSEP = 0,26). The multispectral approach showed similar performance. Applied to individual nodules, the PLS model highlighted a wide variability of leghaemoglobin content in nodules harvested from the same plant. These results show that NIR hyperspectral imaging could be used as a rapid and safe method to quantify leghaemoglobin in pea nodules.

Keywords: cyanmethemoglobin, hyperspectral, imaging, leghaemoglobin, multispectral, near-infrared, pea, PLS regression, quantification, spectroscopy

2. Introduction

Legume plants can fix atmospheric nitrogen thanks to a symbiosis with bacteria. The leghaemoglobin (Lb) protein is synthesized by the host plant within the cytoplasm of cells located in the nitrogen-fixing zone of nodules, the special organ of legumes that contains the nitrogen-fixing bacteria (Pourmoayyed 2016; Taiz & Zeiger 2002). Leghaemoglobin has a high affinity with oxygen and keeps the O₂ concentration in an optimal range to protect nitrogenase enzyme against oxygen damage and to supply bacterial respiration. The nitrogenase enzyme is contained in the nodule and transforms N₂ into NH₃, a form of nitrogen which is metabolically available for plants (Pourmoayyed 2016; Downie 2005). In the case of pea (*Pisum sativum* L.), the bacteria involved in this transformation is *Rhizobium leguminarosum* (Scherer et al. 2008; Kosmachevskaya & Topunov 2007).

Leghaemoglobin in root nodules is commonly quantified in order to assess the effects of plant growth-promoting parameters (like fertilization or UV exclusion of solar radiation) on symbiotic nitrogen fixation efficiency (Pourmoayyed 2016; Sharma & Sharma 2014; Chouhan et al. 2008). The concentration of leghaemoglobin is affected by several factors such as the age of the nodules, the content of nutrients such as S or N in the growing substrate, or plant health (Pourmoayyed 2016; Sharma & Sharma 2014; Orellana et al. 1978). Its concentration in root nodules is closely correlated with the amount of nitrogen fixed in the association between the plant and the bacteria (Kosmachevskaya & Topunov 2007; Wilson & Reisenauer 1963). Low leghaemoglobin content in nodules leads to low N₂ fixation due to higher O₂ concentration and inhibition of the nitrogenase enzyme (Pourmoayyed 2016; Scherer et al. 2008).

Common methods of leghaemoglobin quantification require prior extraction of leghaemoglobin from nodules before any possible quantification of its concentration. This extraction step is time-consuming, destructive and sensitive to operator ability and needs several chemical reagents. Moreover, several nodules are needed and leghaemoglobin quantification is, to the best of our knowledge, not possible in an individual nodule. Among the methods used to quantify leghaemoglobin, the most popular one is based on the cyanmethemoglobin method, also used in medicine to quantify hemoglobin in blood. As leghaemoglobin is structurally and chemically similar to human hemoglobin, the method was transposed to legume hemoglobin as described by Wilson & Reisenauer (1963). The principle of this method is the conversion of hemoglobin into cyanmethemoglobin by the addition of potassium cyanide and ferricyanide, whose absorbance is measured at 540 nm in a spectrophotometer against a standard solution (Nkrumah et al. 2011). This method is considered as a reference method because it is accurate, cheap, sensitive to small variations in hemoglobin concentration, reliable, reproducible and requires a minimum of specialized equipment and reagents (Srivastava et al. 2014; Nkrumah et al. 2011; Balasubramaniam & Malathi 1992;

Wilson & Reisenauer 1963). Nevertheless, the multiple operations required in this method make it time-consuming and thus unsuitable for large-scale field studies. Furthermore, it requires accurate and careful technical operations, which increase the risk of imprecision and errors (Nkrumah et al. 2011). Finally, this method uses cyanide, a toxic reagent (Srivastava et al. 2014).

In order to avoid the limitations of the cyanmethemoglobin method in leghaemoglobin extraction and content measurement, the use of near-infrared (NIR) hyperspectral imaging combined with chemometrics was studied as a new way to quantify leghaemoglobin in nodules. NIR hyperspectral imaging is the combination of NIR spectroscopy and imaging technologies. This technology provides spectral and spatial information simultaneously, *i.e.* thousands of spectra can be obtained for each sample and give a complete picture of the chemical compounds' distribution at pixel level. This technology requires the use of chemometrics, a chemical discipline using mathematics and statistics, in order to extract relevant information from the measurement procedures (Dale et al. 2012; Wold 1995).

The aim of the present study was to predict the leghaemoglobin concentration of nodules on the basis of their NIR spectra. Two approaches were evaluated. Firstly, the partial least squares (PLS) approach was applied to the full spectrum acquired with the hyperspectral device. Secondly, the potential of multispectral analysis was also tested through the preselection of the most responsive wavelengths and the application of a multiple linear regression model. Both procedures were tested on a standard calibration-validation approach.

3. Materials and method

3.1. Nodule production

To produce nodules that covered a large range of leghaemoglobin concentrations, a dedicated protocol was established taking into account that leghaemoglobin content in pea is affected by nitrogen fertilization and varies over time with plant growth (Sharma & Sharma 2014; Höflich et al. 2000; Sato et al. 1998; Orellana et al. 1978). Two types of fertilization (no fertilization and nitrogen fertilization) were applied at seeding time of pea plants. Under both fertilization treatments, nodules were harvested at four different dates during plant growth. Each one of the eight treatments (fertilization x harvest date) was replicated four times.

Nodules were produced in pots (20 cm in diameter, 15 cm high) in the greenhouses of Gembloux Agro-Bio Tech (University of Liège, Belgium). Five pea seeds were sowed in each pot containing field soil and sand (5 kg of substrate per pot with a mass ratio of $\frac{3}{4}$ field soil and $\frac{1}{4}$ sand). The field soil was collected in a field (loam type) where peas had been cultivated three months earlier to ensure the presence of rhizobium in the substrate. Seeds of the Dove variety were used. They were protected with WAKIL fungicide (10 % Cymoxanil, 5 % Fludioxonil and 17 % Metalaxyl-m). The fertilizer was applied in the top layer of pot soil at seeding time. Nitrate fertilization consisted of one input of 0,49 g ammonitrate per pot (equivalent

to 80 kg N.ha⁻¹). Plants were cultivated between October 12th and November 30th under a 24 h photoperiod (daylight and artificial light complement). Being exposed to aphids 35 days after seedling emergence, the plants were protected with Pirimor (50 % Pirimicarbe). The temperature was measured in the vicinity of the plants throughout their development. After germination, three plants were kept in each pot.

The first nodule harvest occurred 25 days after sowing (on November 6th). At this time, the plants had accumulated 500 heat units (basal temperature = 0 °C) and had produced seven leaves. The second harvest occurred when the plants had accumulated 630 heat units, the third after 760 heat units and the fourth after 820 heat units, 48 days after sowing. Approximately one week separated each harvest.

At each harvest date, four pots were randomly selected in each fertilization modality. In order to quantify plant development, the number of nodules present on each plant root system and the aerial dry biomass were measured. Roots were manually extracted from soil with tap water and nodules were separated from roots with tweezers. For each pot, a sample of 30 nodules was dried at 60 °C and then kept in ambient conditions until image acquisition. Other nodules were saved in falcon tubes, dipped in liquid nitrogen and kept at -80 °C. When the amount of nodules harvested on a plant was below 60, only 10 nodules were dried in order to save enough nodules for the chemical quantification.

3.2. Leghaemoglobin quantification

Leghaemoglobin content in frozen nodules was measured the day after each harvest. The assays were based on the cyanmethemoglobin method described in detail by Wilson and Reisenauer (1963). This method is based on spectrophotometric measurement and uses Drabkin's solution (Drabkin & Austin 1935). Briefly, Drabkin's solution was prepared with 52 mg KCN, 198 mg K₈Fe(CN)₆ and 1 g NaHCO₃ dissolved in water and made up to 1000 ml. To extract leghaemoglobin, frozen nodules were crushed in liquid nitrogen and 100 mg of crushed nodules was transferred to a 2 ml centrifuge tube with 0.6 ml Drabkin's solution. The mixture was centrifuged for 15 min at 500 x g and 4 °C. The supernatant was transferred to a 2 ml flask. The solid phase was extracted and centrifuged twice more. Supernatants were combined, made up to 2 ml with Drabkin's solution and centrifuged for 30 min at 20.000 x g and 4 °C. Absorbance of the cleared supernatant was read against Drabkin's solution in a 1.5 ml cell at 540 nm with a UV-visible spectrophotometer (UV-1650 PC, Shimadzu, Japan).

A reference curve was computed each day of measurement with a sample free of hemoglobin and five crystalline human hemoglobin (SIGMA-ALDRICH Cat nr H7379) dilutions: 0,05, 0,1, 0,2, 0,4 and 0,6 mg hemoglobin.ml⁻¹. Absorbance of each dilution was measured twice. Reference curves always had a regression coefficient superior to 0.99. Results were then expressed in milligrams of leghaemoglobin per gram of fresh nodules (mg.g⁻¹).

3.3. NIR spectra acquisition and preprocessing

NIR spectra of dried nodules were acquired with an NIR hyperspectral line scan camera combined with a conveyor belt (BurgerMetrics SIA, Riga, Latvia) installed at the Walloon Agricultural Research Centre (CRA-W, Belgium). The device is described in detail in Vermeulen et al. (2012) and Eylenbosch et al. (2017). Acquisitions were performed with HyperPro VB software (BurgerMetrics SIA, Riga, Latvia). Before image acquisition, the system was calibrated with a white ceramic plate (white reference) and by blocking the entrance of reflected light (dark reference). Bad pixels were detected and removed at the same time. A dark reference was automatically performed before the acquisition of each image. One hyperspectral image was acquired for each nodule sample (one sample per growing pot) and all nodules used for the study were scanned on the same day.

Nodule samples were put on a conveyor belt progressing at $1200 \mu\text{m}\cdot\text{s}^{-1}$ under the camera. For each pixel, one spectrum of 209 wavelengths was acquired in the 1118-2424 nm range. Pixel resolution was 0,31 mm wide. One mean spectrum representative of each nodule sample was then computed from each hyperspectral image with the HyperSee software (BurgerMetrics SIA, Riga, Latvia). A threshold was used to segment images and remove background prior mean spectra computation. The mean spectrum of each individual nodule was also computed.

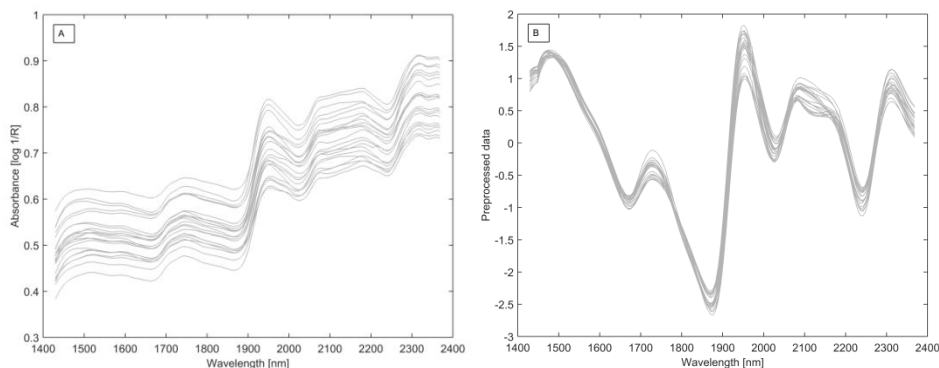


Figure 58A and B: Average near-infrared (NIR) spectra of pea nodules used for PLS model calibration before (A) and after (B) preprocessing (Savitzky-Golay smoothing (order 0, filter width 7), detrend and Standard Normal Variate (SNV) transformations).

In order to remove noisy regions at the beginning and the end of the NIR spectra, only the wavelengths between 1432 and 2368 nm were kept (Eylenbosch et al. 2017). The remaining noise and the increasing trend observed in the mean spectra were removed by use of Savitzky-Golay smoothing (order 0, filter width 7) and detrend transformations respectively. Spectra were also normalized with Standard Normal Variate (SNV) transformation (Figure 58).

3.4. Linking NIR hyperspectral and multispectral imaging to leghaemoglobin content

For each modality (harvest date x fertilization), the leghaemoglobin content and the average NIR spectrum from one nodule sample was kept for the validation of the regression model. This sample was randomly selected among the four replicates. The remaining data were used for model calibration (two samples were unusable due to handling errors). In total 22 mean spectra were used for calibration and 8 others for validation.

3.4.1. PLS regression model construction

Partial Least Square (PLS) regression was used to build a calibration model between the leghaemoglobin content and the NIR spectra of pea nodules. The number of latent variables was chosen in order to reduce the cross-validation classification error average. A 10-group venetian blind cross-validation was performed.

This work was performed using the PLS Toolbox 7.8 software (Eigenvector Research, Inc., Wenatchee, WA, USA) working with Matlab R2015a software (The Math Works, Inc., Natick, MA, USA).

3.4.2. Optimal wavelength selection and multiple linear regression fitting

Spectra acquired with hyperspectral imaging contain a large number of wavelengths. Most of the wavelengths are therefore likely to provide redundant information and therefore make little or no contribution to the predictions. Identification of wavelengths carrying the maximum spectral information would allow the simplification of the model and further technology transfer to more simple acquisition systems, allowing for the implementation of multispectral imaging.

The standard coefficients (β -coefficients) of the PLS regression were used to analyze and identify which wavelengths were the most significant in explaining the variations observed in leghaemoglobin content (Liu et al. 2013). As two proximal wavelengths might contain redundant information, rather than identifying the wavelength with the highest score (as a positive or negative value), it was decided to identify the peaks in the graphs of β -coefficients against wavelengths (cf Results section). The wavelengths for which peaks were identified were then used in a multiple linear regression model to predict leghaemoglobin content.

Algorithms used to select the wavelengths and create the multiple linear regression model were developed using Matlab R2015a software (The Math Works, Inc., Natick, MA, USA).

3.5. Statistical analysis

Statistical analyses of nodule numbers, leghaemoglobin content and dry aerial biomass were performed with the R software (R Core Team 2017). Analysis of variance (ANOVA) and pairwise comparisons with Student-Newman-Keuls tests were made using the agricolae package (Mendiburu & Simon 2015).

Three criteria were used to evaluate the quality of the models and their ability to predict leghaemoglobin content: root mean square error (RMSE), determination coefficient (r^2) and ratio of standard deviation to RMSE called ratio of prediction to deviation (RPD) (Xu et al. 2016; Saey et al. 2005). According to Saey et al. (2005), model predictions can be considered as good when the r^2 value is between 0,82 and 0,90. They are classified as good or excellent when the RPD values are between 2,5 and 3,0 or above 3,0 respectively.

4. Results and discussion

4.1. Pea aerial biomass and nodule development

At each harvest date, the number of nodules on each plant root system and the aerial biomass were measured. For the statistical analysis, these measurements were taken per pot (three plants). Both variables were significantly affected by accumulated heat units (p-value < 0,01 – Table 8). The ANOVA reported no significant impact of nitrogen fertilization (results not shown). Throughout the study, nodules of different ages and sizes were observed. An increase in nodule number was observed between the first and the second harvest; after which the number of nodules tended to stabilize. Aerial dry biomass increased with an exponential shape with accumulated heat units. At the end of the study, most developed plants had 13 leaves.

The leghaemoglobin content measured in the harvested nodules with the cyanmethemoglobin method was comprised between 1,4 and 4,2 mg.g⁻¹ (Table 8 and Figure 59). These concentrations were consistent with those reported in previous studies (Pourmoayyed 2016; Scherer et al. 2008). A broad range of leghaemoglobin content in nodules was therefore observed, as expected with the protocol established.

Table 8: Number of nodules, aerial dry biomass production (g) and nodule leghaemoglobin content (mg.g⁻¹ fresh nodule) measured per growing pot (3 plants) at the four harvest dates.

Mean and standard deviation are presented for each harvest date. N-fertilized and non-fertilized plants are pooled. Mean values followed by the same letter are not significantly different (Student-Newman-Keuls, $\alpha = 0,05$)

Accumulated heat units [°C days]	Nodules		Aerial dry biomass		Leghaemoglobin content	
	Mean [#pot ⁻¹]	SD [#pot ⁻¹]	Mean [g.pot ⁻¹]	SD [g.pot ⁻¹]	Mean [mg.g ⁻¹]	SD [mg.g ⁻¹]
500	95,6 (b)	27,4	0,79 (c)	0,22	1,63 (c)	0,25
630	179,7 (a)	89,9	1,22 (bc)	0,63	2,50 (b)	0,56
760	218,3 (a)	72,5	2,03 (b)	0,75	3,42 (a)	0,48
820	216,0 (a)	58,3	4,69 (a)	1,52	3,52 (a)	0,43

The analysis of leghaemoglobin content in pea nodules showed a highly significant effect (p -value $< 0,001$ – Table 8) of harvest time. An increase in leghaemoglobin content was observed between the first and third harvest. No statistical differences were reported between the last two harvest dates. Contrary to our expectations, there was no observable effect of nitrogen fertilization on leghaemoglobin content in nodules (result not shown). This absence of effect could be due to the high amount of nitrogen already present in the growth substrate, which was taken from a field where peas had been grown in the previous crop season. Furthermore, the growing conditions, *i.e.* higher temperature of the substrate due to the cultivation performed in pots within the greenhouse, have likely increased soil mineralization which also increased the nutrient content available for the plants. Figure 59 presents the leghaemoglobin content for the different harvest dates, gathering altogether fertilized and unfertilized datasets.

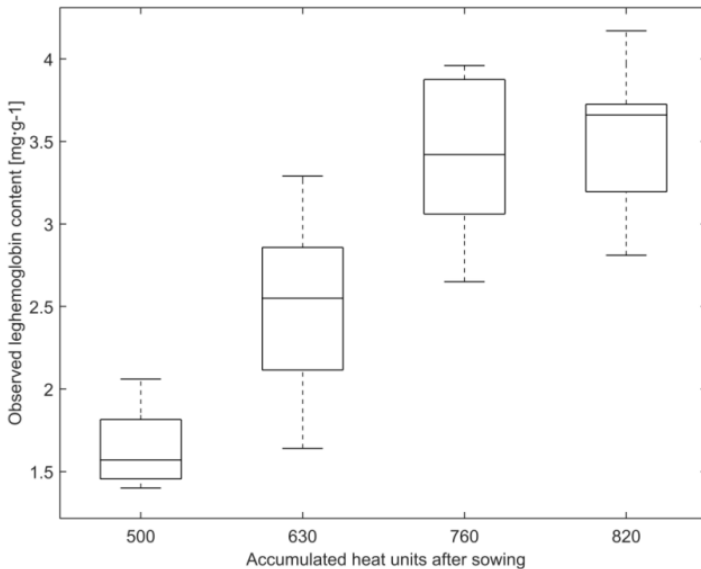


Figure 59: Leghaemoglobin content ($\text{mg leghaemoglobin} \cdot \text{g}^{-1}$ fresh nodule) measured with the cyanmethemoglobin method at each harvest time (expressed in heat units ($^{\circ}\text{C days}$) accumulated after plant sowing).

4.2. PLS regression

The main goal of this work was to investigate whether leghaemoglobin content could be predicted from average NIR spectra of pea nodules acquired by NIR hyperspectral imaging. The first attempt was made using a PLS regression to link leghaemoglobin content, measured with the cyanmethemoglobin method, and NIR spectra of the nodule. In the calibration phase, 22 nodules samples were used. Five latent variables were used to predict the leghaemoglobin content and explained

99,98 % of leghaemoglobin variance. The first latent variable explained 98,51 % of the variance.

The PLS regression showed a determination coefficient (r^2) of 0,87 and a root mean square error of calibration (RMSEC) of 0,32 (Figure 60). The cross validation of the PLS model gave an r^2 of 0,74 and a root mean square error (RMSECV) of 0,45. Lower values obtained with cross validation were probably due to the small number of samples.

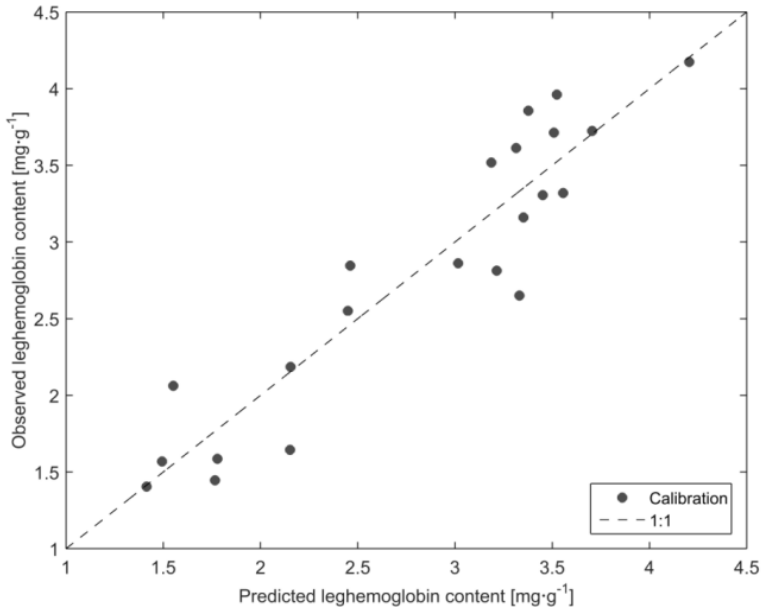


Figure 60: Calibration results of the partial least square (PLS) regression model calibrated on leghaemoglobin content measured with cyanmethemoglobin method and predicted leghaemoglobin content of nodules on the basis of their near-infrared (NIR) spectra. Results are expressed in mg leghaemoglobin.g⁻¹ fresh nodules. The regression was constructed on five latent variables. Leghaemoglobin was predicted with a root mean square error of calibration (RMSEC) of 0,32 and a determination coefficient (r^2) of 0,87.

The validation of the PLS model on eight samples gave an r^2 of 0,90 and a root mean square error of prediction (RMSEP) of 0,27 (Figure 61). The RPD was 3,42. Predictions were therefore classified as good (Saeys et al. 2005). The leghaemoglobin contents of the samples used for the validation were well predicted by the regression, and the predictions obtained from these samples were better than those obtained with the cross-validation: the r^2 was higher and the RMSEP was lower than the RMSECV. These samples were not completely independent from those used for the regression model calibration because they were obtained in the same conditions. However, the division of samples, or spectra in a dataset, acquired under the same conditions into calibration and validation sets is common in

chemometric studies to estimate the potential of a methodology (Xu et al. 2016; Dale et al. 2013; Fernández Pierna et al. 2012; Williams et al. 2009).

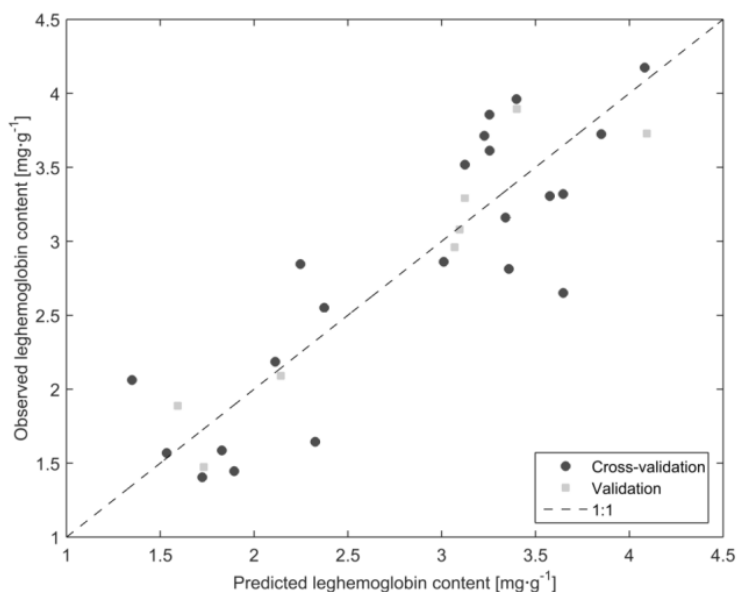


Figure 61: Cross-validation (dark grey circle) and validation (light grey square) results of the partial least square (PLS) regression model. Leghaemoglobin content was measured with cyanmethemoglobin method and predicted on the basis of nodule near-infrared (NIR) spectra. Results are expressed in mg leghaemoglobin.g⁻¹ fresh nodules. The regression was constructed on five latent variables. Leghaemoglobin was predicted with a root mean square error of cross-validation (RMSECV) of 0,45 and a determination coefficient (r^2) of 0,74. When the regression was applied to the spectra used for the validation, the root mean square error of prediction (RMSEP) was 0,27 and the r^2 was 0,90.

4.3. Validation on single nodules

As stated in the 'Materials and method' section, NIR hyperspectral imaging allows one spectrum per pixel to be acquired. The results obtained with the PLS model (PLS regression section) were based on the computation of an average spectrum representative of the whole sampling, *i.e.* a sample including all nodules harvested for each growing pot. To evaluate the variability within each sampling, which would make it possible to evaluate the leghaemoglobin content of individual nodules separately, we computed one mean spectrum per nodule within each image (using the information contained in all pixels related to each nodule of the image). The average pixel number for one nodule was 37 (ranging between 10 and 290 pixels per nodule). The PLS model was then applied to each mean spectrum representative of each individual nodule of the validation dataset. The results are presented in Figure 62.

The variability observed in leghaemoglobin content predicted from mean spectra tended to be higher where the measured content was greater, *i.e.* in nodule samples harvested from the oldest plants. These plants had nodules of different ages and diameters on their roots. Our observations were consistent with results reported for soybean by Sato et al.(1998) who showed that plants have nodules of different diameters and that the leghaemoglobin content increases with the nodules' diameter.

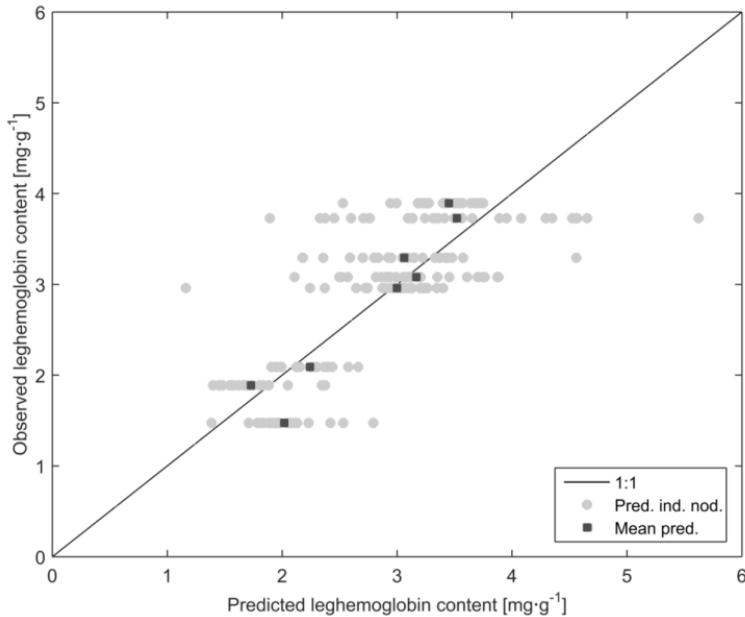


Figure 62: Predicted values of leghaemoglobin on individual nodules (mg leghaemoglobin.g⁻¹ fresh nodule). Predictions were made on each mean near-infrared spectrum computed for each nodule (light gray circles). Nodules were contained in height samples (validation dataset). The mean leghaemoglobin content (dark gray squares) computed on the whole sample was linked with the reference value of leghaemoglobin (observed leghaemoglobin content) measured for the sample with the cyanmethemoglobin method.

4.4. Multiple linear regression model applied to a restricted number of wavelengths

Figure 63 reports the β -coefficients corresponding to each wavelength, when the PLS analysis was performed (here, the Matlab function *plsregress* was used to retrieve the coefficients). The graph corresponded to the results obtained from plotting the output of the PLS regression using the first component of the PLS obtained with the *plsregress* function. The percentage of variance explained in this

case was about 71 %. As stated in the 'Materials and method' section, rather than identifying the highest coefficients, it was decided to identify the peaks within the β -coefficients graph.

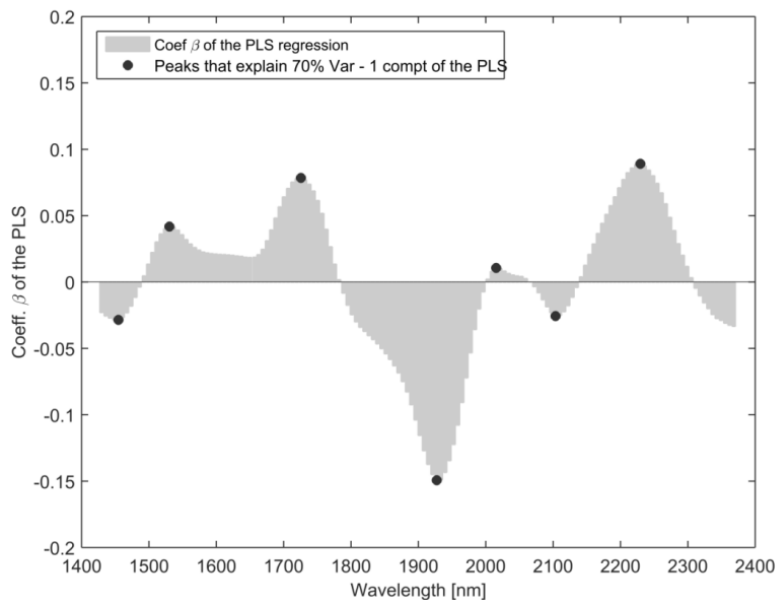


Figure 63: Most relevant wavelengths selected using the graphical representation of coefficients against the corresponding wavelength.

The wavelengths identified were: 1461, 1537, 1732, 1934, 2022, 2110 and 2236 nm (Figure 64). A multiple linear regression model was calibrated with the absorbance values measured at these wavelengths to predict leghaemoglobin content. The calibration displayed similar performances to those observed in the complete PLS approach, with an r^2 of 0,85 and an RMSEC of 0,35.

During the validation phase, the results again revealed similar performances, with an r^2 of 0,86 and an RMSEP of 0,33. Graphical representations were fairly similar to those reported when applying the complete PLS approach.

Not surprisingly, most of the identified wavelengths were linked to chemical structure containing carbon and nitrogen. These wavelengths are situated in the first overtone, the second overtone and the combination band regions of the NIR spectrum (Metrohm 2013). Key wavelengths identified at 1461, 1934 and 2110 nm are located in the regions of N-H bonds. The 1934 nm wavelength is also linked to moisture (Caporaso et al. 2018). Wavelengths identified at 1732 and 2236 nm are located in the C-H bond regions.

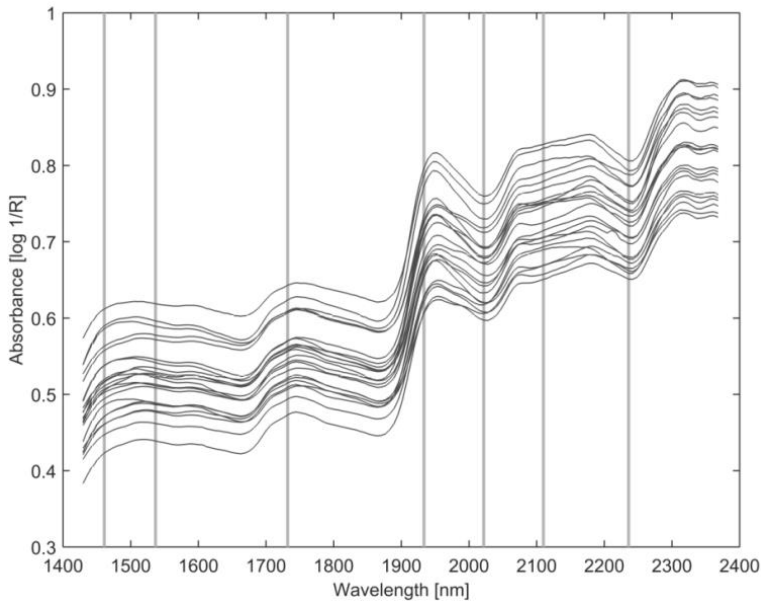


Figure 64: Identification (vertical gray lines) of the seven wavelengths (1461, 1537, 1732, 1934, 2022, 2110 and 2236 nm) that were used to perform multiple linear regression analysis. The wavelengths of interest were reported on the graphical representation of the pea nodule spectra used to calibrate the multiple linear regression model.

5. Conclusion

This paper has proposed a new method for leghaemoglobin quantification in dried pea nodules using NIR hyperspectral imaging combined with chemometrics as an alternative to the classical reference cyanmethemoglobin method. The PLS regression gave good results when the whole spectrum of pea nodules was used.

This PLS regression applied to the mean spectrum of individual nodules showed consistent results: nodules collected from the same plant had different leghaemoglobin content, due to the size and age of nodules, and their mean leghaemoglobin content as well as the variability of their leghaemoglobin content tended to increase with plant age.

Using the peaks identified in the β -coefficients graph resulting from the PLS analysis applied to the calibration dataset, seven wavelengths of interest were identified. A multiple linear regression model was calibrated on these wavelengths to predict leghaemoglobin content. This simpler model gave similar results on the validation dataset to those of the PLS regression applied to the whole spectrum.

Both proposed approaches showed good correlation with the reference method. However, they show some advantages over the reference approach: (i) they are more rapid; (ii) they are simpler (once the model is calibrated, the quantification does not

depend on operator ability); (iii) they do not involve the use of chemical and potentially dangerous reagents; (iv) the leghaemoglobin content of a single nodule can be predicted. Considering all these advantages and the first results acquired in this study, NIR hyperspectral imaging and multispectral analysis have demonstrated strong potential for evaluating the activity of pea nodules through the quantification of leghaemoglobin content. They appear suited to use within large-scale field studies in order to improve knowledge of nitrogen fixation and protein accumulation in legume crops.

6. Acknowledgements

The authors would like to thank Pete Iannetta (James Hutton Institute, Scotland), who gave us the inspiration for this work, Patrick du Jardin and Florence Paquet (Gembloux Agro-Bio Tech – University of Liège) for their help in this study and the access they provided to the Plant Biology laboratory, and the CAMI technical platform at the AgrobioChem Department (Gembloux Agro-Bio Tech – University of Liège) for access to the computational server. They would also like to thank Nikaise Kayoka Mukendi from the Valorization of Agricultural Products Department (CRA-W) who helped with NIR image acquisition and data preparation, as well as the students of Gembloux Agro-Bio Tech who collected the samples and performed the pre-analysis of the results: Ariel Doudet, Florent Druart, Laëtitia Machiels, Audrey Sougnez, Marine Torfs.

7. References

- Balasubramaniam & Malathi, A., 1992. Comparative study of hemoglobin estimated by Drabkin's and Sahli's methods. *J Postgrad Med*, 38(1), p.8.
- Caporaso, N., Whitworth, M.B. & Fisk, I.D., 2018. Protein content prediction in single wheat kernels using hyperspectral imaging. *Food Chem*, 240, pp.32–42.
- Chouhan, S. et al., 2008. Enhancement in leghemoglobin content of root nodules by exclusion of solar UV-A and UV-B radiation in soybean. *J Plant Biol*, 51(2), pp.132–138.
- Dale, L.M. et al., 2012. Chemometric tools for NIRS and NIR hyperspectral imaging. *Bulletin UASVM Agriculture*, 69(1), pp.70–76.
- Dale, L.M. et al., 2013. Discrimination of grassland species and their classification in botanical families by laboratory scale NIR hyperspectral imaging: preliminary results. *Talanta*, 116, pp.149–154.
- Downie, J.A., 2005. Legume haemoglobins: symbiotic nitrogen fixation needs bloody nodules. *Curr Biol*, 15(6), pp.196–198.
- Drabkin, D.L. & Austin, J.H., 1935. Preparation from washed blood cells; nitric oxide hemoglobin and sulfhemoglobin. *J Biol Chem*, 112, pp.51–65.

Eylenbosch, D. et al., 2017. NIR hyperspectral imaging spectroscopy and chemometrics for the discrimination of roots and crop residues extracted from soil samples. *J Chemom*, 32(1), pp.1–11.

Fernández Pierna, J.A. et al., 2012. NIR hyperspectral imaging spectroscopy and chemometrics for the detection of undesirable substances in food and feed. *Chemometr Intell Lab*, 117, pp.233–239.

Höflich, G. et al., 2000. Influence of agricultural crops and fertilization on microbial activity and microorganisms in the rhizosphere. *J Agron Crop Sci*, 184(1), pp.49–54.

Kosmachevskaya, O. V. & Topunov, A.F., 2007. Method of determination of the content of hemoglobin-like proteins in heterogenic mixtures. *Appl Biochem Micro*, 43(3), pp.313–319.

Liu, D., Sun, D.-W. & Zeng, X.-A., 2013. Recent advances in wavelength selection techniques for hyperspectral image processing in the food industry. *Food Bioprocess Tech*, 7(2), pp.307–323.

Mendiburu, F. & Simon, R., 2015. *Agricolae* -Ten years of an Open source Statistical tool for experiments in Breeding, agriculture and biology. *PeerJ Prepr*, 3, pp.1–17.

Metrohm, 2013. *NIR Spectroscopy: A guide to near-infrared spectroscopic analysis of industrial manufacturing processes*, Herisau, Switzerland: Metrohm AG.

Nkrumah, B. et al., 2011. Hemoglobin estimation by the HemoCue portable hemoglobin photometer in a resource poor setting. *BMC Clinic Pathol*, 11, pp.1–6.

Orellana, R.G., Fan, F. & Sloger, C., 1978. Tobacco ringspot virus and rhizobium interactions in soybean: impairment of leghemoglobin accumulation and nitrogen fixation. *Phytopathology*, 68, pp.577–582.

Pourmoayyed, P., 2016. The effect of sulfur deficiency in pea (*Pisum sativum*) on factors related to nitrogen fixation and on sulfate transporters. *Rheinischen Friedrich Wilhelms Universität Bonn*.

R Core Team, 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Saeyes, W., Mouazen, A.M. & Ramon, H., 2005. Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy. *Biosyst Eng*, 91(4), pp.393–402.

Sato, T. et al., 1998. Determination of leghemoglobin components and xylem sap composition by capillary electrophoresis in hypernodulation soybean mutants cultivated in the field. *Soil Sci Plant Nutr*, 44(4), pp.635–645.

Scherer, H.W. et al., 2008. Low levels of ferredoxin, ATP and leghemoglobin contribute to limited N₂ fixation of peas (*Pisum sativum* L.) and alfalfa (*Medicago sativa* L.) under S deficiency conditions. *Biol Fert Soils*, 44(7), pp.909–916.

Sharma, A. & Sharma, S., 2014. Effect of nitrogen and sulphur nutrition on nitrogen assimilating enzymes in soybean roots and nodules. *IJAEB*, 7(3), pp.471–480.

Srivastava, T. et al., 2014. Methods for hemoglobin estimation: a review of “What works”. *J Hematol Transfus*, 2(3), pp.2005–2006.

Taiz, L. & Zeiger, E., 2002. *Plant physiology*, third edition, Sunderland, Massachusetts: Sinauer Associates, Inc., Publishers.

Vermeulen, P. et al., 2012. Online detection and quantification of ergot bodies in cereals using near infrared hyperspectral imaging. *Food Addit Contam A*, 29(2), pp.232–240.

Williams, P. et al., 2009. Maize kernel hardness classification by near infrared (NIR) hyperspectral imaging and multivariate data analysis. *Analytica chimica acta*, 653, pp.121–130.

Wilson, D.O. & Reisenauer, H.M., 1963. Determination of leghemoglobin in legume nodules. *Anal Biochem*, 6, pp.27–30.

Wold, S., 1995. Chemometrics; what do we mean with it, and what do we want from it? *Chemometrics and Intelligent Laboratory Systems*, 30(1), pp.109–115.

Xu, S. et al., 2016. Determination of rice root density at the field level using visible and near-infrared reflectance spectroscopy. *Geoderma*, 267, pp.174–184.

Citation: D. Eylenbosch et al., “Quantification of leghaemoglobin content in pea nodules based on near infrared hyperspectral imaging spectroscopy and chemometrics”, *J. Spectral Imaging* 7, a9 (2018). <https://doi.org/10.1255/jsi.2018.a9>

Conclusions au Chapitre VII et perspectives

L'étude présentée dans ce chapitre a démontré qu'il était possible de prédire les teneurs en leghémoglobine au sein de nodosités individuelles de pois protéagineux sur base de spectres proche infrarouge extraits d'images hyperspectrales. L'application du modèle de régression PLS développé sur des échantillons de nodosités a permis de mettre en évidence la variation des teneurs en leghémoglobine au sein de nodosités d'une même plante. Ces variations étaient dues aux variations de la taille et de l'âge des nodosités. Cette méthode a aussi permis de mettre en évidence une augmentation de la concentration moyenne en leghémoglobine au cours du développement des plantes de pois. Par rapport à la méthode de référence utilisée pour la quantification de la leghémoglobine, cette nouvelle approche est plus rapide, plus simple, elle ne nécessite pas l'utilisation de réactifs chimiques toxiques et elle permet de faire de la quantification sur des nodosités individuelles. L'étude a également montré que sept longueurs d'ondes au sein du spectre proche infrarouge suffisaient pour prédire les teneurs en leghémoglobine.

Ayant montré la capacité de l'imagerie hyperspectrale proche infrarouge et des outils de la chimiométrie à prédire la teneur en leghémoglobine au sein de nodosités individuelles, l'étape suivante consistait à combiner la régression obtenue et le modèle de discrimination des racines et de nodosités de pois afin de quantifier la leghémoglobine sur des systèmes racinaires entiers, sans devoir détacher les nodosités des racines. Suite à la prédiction de la nature de chaque pixel d'images hyperspectrales acquises sur des systèmes racinaires de pois, un modèle de reconnaissance de forme nous a permis de détecter les nodosités. Un spectre moyen a été calculé pour chacune d'elle et la teneur en leghémoglobine moyenne a été prédite pour chaque nodosité présente sur l'image. Ces teneurs en leghémoglobine ont ensuite pu être affichées, selon un gradient de couleur, sur une image reconstruite. Il a ainsi été possible de visualiser sur une même image (i) les racines de pois et leur architecture, (ii) la distribution des nodosités sur les racines de pois, et (iii) la distribution des teneurs en leghémoglobine sur le système racinaire du pois (Figure 65). Sur base des résultats obtenus dans l'étude menée sur les racines prélevées sous une culture en association de froment et de pois d'hiver, nous pouvons imaginer que de telles prédictions pourraient également être réalisées sur des images plus complexes, comprenant par exemple aussi des racines d'autres cultures et des résidus de matières organiques.

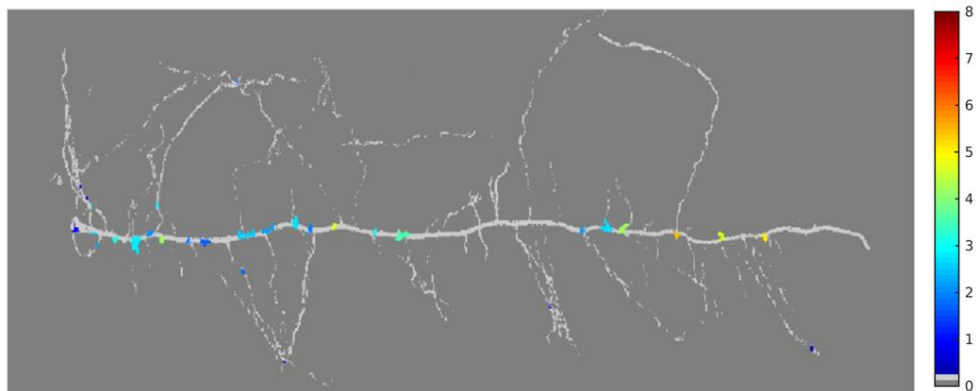


Figure 65: Prédiction des teneurs en leghémoglobine au sein de nodosités présentes sur un système racinaire de pois protéagineux (*Pisum sativum* L.) acquis par imagerie hyperspectrale proche infrarouge. Les teneurs en leghémoglobine, présentées selon une échelle de couleur, sont exprimées en mg de leghémoglobine par g de nodosité fraîche.

CHAPITRE VIII

Discussion et conclusions

"More energy may have been spent on developing technologies than on studying roots. Each method so far developed has serious shortcoming... No one technique, of the many so far developed, informs us about all aspects of root growth and structure."

Harper et al. (1991) cité par Atkinson (2000)

Discussion générale

Grâce à nos travaux, nous avons prouvé que l'imagerie hyperspectrale proche infrarouge et la chimiométrie peuvent avantageusement être utilisés dans l'étude de systèmes racinaires. Cette approche originale a notamment permis de trier les spectres de différents éléments extraits de carottes de sol et des racines de différentes espèces. Elle a aussi permis de quantifier, sur des images hyperspectrales, des masses de racines et des teneurs en composés organiques présents au sein de racines.

Cette discussion générale présente les opportunités et les perspectives offertes par l'imagerie hyperspectrale et la chimiométrie dans l'étude des systèmes racinaires. Elle présente également les limites et les contraintes rencontrées avec cette méthode dans la quantification de la masse racinaire. Ces limites sont principalement d'ordre technique et leur impact devrait donc être diminué avec l'évolution rapide des systèmes d'acquisition et de traitement d'images.

La revue bibliographique ayant déjà repris les avantages et les désavantages généraux de l'imagerie hyperspectrale proche infrarouge, nous ne présentons ici que nos considérations personnelles basées sur les résultats de nos travaux. Nous rappelons également dans cette discussion générale un certain nombre de points auxquels il faut être particulièrement attentif lors de l'utilisation de l'imagerie hyperspectrale dans l'étude de systèmes racinaires et nous proposerons certaines pistes d'amélioration.

1. Opportunités offertes par l'imagerie hyperspectrale proche infrarouge couplée à la chimiométrie dans l'étude de systèmes racinaires

L'évaluation que nous avons menée sur les potentialités de l'imagerie hyperspectrale proche infrarouge et la chimiométrie dans l'étude de systèmes racinaires nous a montré que cette méthode d'analyse permettait de se passer de l'étape de tri de racines et de résidus de cultures ou de tri de racines de différentes espèces qui était jusqu'à présent nécessaire lors de l'étude de racines prélevées par carottage de sol. Elle a également montré qu'elle pouvait remplacer une méthode d'analyse chimique laborieuse, utilisant un réactif toxique et peu adaptée à une bonne compréhension du fonctionnement du système racinaire du pois protéagineux.

Ces applications de l'imagerie hyperspectrales proche infrarouge et de la chimiométrie ont été utilisées dans des essais agronomiques. Elles ont permis de comparer l'influence de modalités de cultures sur le développement racinaire. Elles ont également permis de suivre l'évolution de la densité de masse racinaire, de leur distribution dans le sol et de leur composition au cours du temps.

1.1. Discrimination et quantification de racines et de résidus de cultures extraits de carottes de sol

L'étude racinaire par carottage de sol est souvent citée comme étant la méthode qui permet d'obtenir les meilleures informations quantitatives sur la biomasse du système racinaire et la densité de longueur racinaire. Elle est donc fréquemment utilisée comme méthode de référence (Xu et al. 2016; Kusumo et al. 2010; Oliveira et al. 2000; Cheng et al. 1990). Or, la fréquence d'échantillonnage de carottes de sol est limitée par le temps nécessaire au traitement des échantillons (Majdi 1996) et particulièrement par l'étape de tri manuel nécessaire après le lavage des échantillons de sol pour séparer les racines des autres éléments extraits de ces échantillons.

La possibilité de discriminer ces éléments sur des images hyperspectrales et l'existence d'une relation linéaire entre le nombre de pixels de racine et la masse de ces racines permet donc de se passer de ces étapes de tri manuel et de pesée des échantillons. En éliminant ces étapes, l'utilisation de l'imagerie hyperspectrale proche infrarouge, combinée aux outils de la chimométrie, permet d'analyser plus rapidement les échantillons. Elle permet donc de libérer du temps et d'augmenter potentiellement le nombre d'échantillons de racines prélevés. De plus, elle réduit la pénibilité du traitement des échantillons et est beaucoup moins sensible à la subjectivité de l'opérateur. Nous verrons qu'une certaine subjectivité reste néanmoins présente avec cette technique d'analyse au niveau du positionnement de l'échantillon sur le support d'acquisition des images hyperspectrales et au niveau de la sélection des spectres servant de référence pour la construction des modèles de discrimination.

Le potentiel de cette nouvelle méthode de quantification basée sur l'imagerie hyperspectrale proche infrarouge et de la chimométrie a été utilisé dans différentes études agronomiques au cours de cette étude pour traiter des échantillons de sol prélevés par carottage.

Tout d'abord, cette méthode a été utilisée pour mesurer les densités de masse racinaire au sein de carottes de sol prélevées au cours de deux saisons culturales dans un essai sur le travail du sol et la gestion des résidus de culture. Elle a permis de mettre en évidence un effet du travail du sol (labour ou travail réduit) sur la distribution racinaire du froment d'hiver dans l'horizon superficiel du sol (0-30 cm). Elle a également permis de mettre en évidence une évolution dans la quantité de racines produite par la culture au cours du temps. Enfin, elle a mis en évidence que la densité de masse racinaire était la plus importante dans l'horizon superficiel du sol et diminuait en profondeur.

Cette méthode de quantification a également été utilisée pour mesurer la quantité de résidus de cultures présente au sein de carottes de sol. Appliqués sur des échantillons prélevés à quatre dates, elle a permis de mettre en évidence la dégradation des résidus de cultures enfouis dans le sol au cours de la saison culturale.

1.2. Discrimination et quantification de racines de différentes espèces

En combinant imagerie et spectroscopie, l'imagerie hyperspectrale proche infrarouge permet d'analyser simultanément des éléments de natures différentes, de les discriminer et de les quantifier. La spectroscopie classique utilisée seule avait déjà montré par le passé son utilité pour discriminer et quantifier des racines de différentes espèces (Meinen & Rauber 2015; Rewald et al. 2012; Lei & Bauhus 2010; Naumann et al. 2010; Roumet et al. 2006; Rumbaugh et al. 1988) et des racines mortes et vivantes (Picon-Cochard et al. 2009).

Dans les recherches que nous avons menées, nous avons également démontré que l'imagerie hyperspectrale proche infrarouge permettait de discriminer des racines d'espèces différentes (froment et pois d'hiver) mais qu'elle permettait en plus de les quantifier sans devoir broyer ni peser les échantillons. En effet, les études réalisées précédemment avec la spectroscopie classique sur des mélanges de racines nécessitaient de broyer les échantillons et permettaient de déterminer uniquement les masses relatives des différentes espèces au sein des mélanges (Meinen & Rauber 2015; Roumet et al. 2006; Rumbaugh et al. 1988). Nous avons montré que l'imagerie hyperspectrale proche infrarouge permettait également de discriminer des nodosités et des racines de pois. En couplant cette discrimination à un logiciel de reconnaissance de formes, il a été possible de compter des nodosités sur des systèmes racinaires de pois.

La discrimination et la quantification de racines de froment et de pois d'hiver sur des images hyperspectrales a été utilisée dans un essai mené en plein champ sur les effets de l'association de cultures. La quantification séparée des racines des deux cultures conduites en association a permis de mettre en évidence que la densité de masse racinaire du pois protéagineux était identique quel que soit le mode de culture (seul ou associé) et cela malgré une adaptation de la densité de semis. Elle a aussi montré que le pois avait toujours une densité de masse racinaire plus faible que le froment et que cette densité de masse racinaire évoluait peu au sein du profil de sol, mis à part lorsqu'une forte fumure azotée (160 kg N/ha) était appliquée sur la culture. Enfin, cette étude a montré que la densité de masse racinaire de la culture en association était plus élevée pour la culture de froment pur et la culture en association que pour la culture de pois pure. L'association de cultures permettait donc une meilleure utilisation des ressources du sol qu'une culture pure de pois.

Par rapport aux méthodes d'analyses visuelles généralement utilisées pour trier les racines de différentes espèces, l'utilisation de l'imagerie hyperspectrale proche infrarouge est moins laborieuse et permet d'avoir un tri qui est moins influencé par la subjectivité de l'opérateur.

En comparaison avec d'autres méthodes de discrimination de racines de différentes espèces, l'imagerie hyperspectrale proche infrarouge est moins onéreuse que l'utilisation d'isotopes (Rewald et al. 2012) et semble beaucoup moins laborieuse que la méthode utilisée par Ramirez-Garcia et al. (2015) qui consistait à analyser par

microscopie la morphologie et l'anatomie de la section transversale de racines afin d'identifier leur appartenance à une espèce.

1.3. Estimation de la composition chimique de systèmes racinaires

Dans le cadre de nos travaux, nous avons montré que l'imagerie hyperspectrale proche infrarouge et la chimiométrie pouvaient être utilisées pour quantifier la leghémoglobine, une protéine essentielle dans la fixation symbiotique de l'azote par les légumineuses, au sein de nodosités de pois. Par rapport à la méthode de référence utilisée actuellement pour faire cette quantification, l'imagerie hyperspectrale offre plusieurs avantages et elle ouvre de nouvelles perspectives de recherche.

En effet, la méthode de référence nécessite la séparation des racines et des nodosités ainsi que l'extraction de la leghémoglobine de ces nodosités avant de pouvoir la quantifier. Cette extraction comporte différentes étapes qui prennent du temps et qui peuvent être influencées par l'habileté de l'opérateur. De plus, cette extraction nécessite une certaine masse de nodosité (elle ne peut donc pas être réalisée sur des nodosités individuelles) et utilise du cyanure, un réactif chimique connu pour sa toxicité. L'imagerie hyperspectrale ne nécessite quant à elle aucune extraction de la leghémoglobine. La seule étape de préparation des échantillons est un séchage en étuve. De plus, cette méthode d'analyse est simple d'utilisation une fois que les modèles chimiométriques sont mis au point et est donc peu sujette à l'habileté de l'opérateur. Enfin, l'imagerie hyperspectrale permet de faire des mesures sur des nodosités individuelles ce qui permet d'étudier séparément les nodosités d'un même système racinaire et d'avoir dès lors une représentation de la distribution des teneurs en leghémoglobine, et donc de leur activité, au sein de ce système racinaire.

Une estimation de la teneur en leghémoglobine au sein de nodosités de pois protéagineux sur base de l'imagerie hyperspectrale et des outils de la chimiométrie a été réalisée en 2018 dans des essais menés en champs²⁷. Elle a tout d'abord montré que les teneurs en leghémoglobine augmentaient dans les nodosités de pois d'hiver entre la reprise de la végétation de la culture, à la sortie de l'hiver, et la floraison de la culture. A ce stade de la culture, la concentration moyenne en leghémoglobine se stabilisait. Elle a également montré que les teneurs moyennes en leghémoglobine pouvaient être différentes selon la variété de pois considérée et que les nodosités présentes sur un même système racinaire n'avaient pas les mêmes teneurs en leghémoglobine.

En combinant la discrimination des racines et des nodosités de pois, la détection des nodosités par un logiciel de reconnaissance de formes et l'estimation de la teneur en leghémoglobine sur base du spectre proche infrarouge de ces nodosités, nous avons pu étudier la répartition des teneurs en leghémoglobine au sein des nodosités de systèmes racinaires de pois.

²⁷ Travail de fin d'études réalisé par Louis Dumont de Chassart. "Evaluation de l'activité fixatrice d'azote de nodosités de pois protéagineux par imagerie hyperspectrale proche infrarouge". Gembloux Agro-Bio Tech, Université de Liège. Année académique 2017-2018.

2. Limites et contraintes liées à l'utilisation de l'imagerie hyperspectrale proche infrarouge et de la chimiométrie dans l'étude de systèmes racinaires.

Tout au long du processus de quantification racinaire, une série d'erreurs peut biaiser le résultat final. Ces erreurs s'accumulent au cours des étapes de traitement des échantillons, de l'échantillonnage des racines au champ jusqu'à la prédiction de la masse de ces racines sur des images hyperspectrales.

Les premières erreurs de quantifications racinaires sont dues à l'hétérogénéité du système racinaire, aux méthodes de prélèvement et de traitement des échantillons de sol qui sont utilisées et aux pertes de racines lors de la manipulation des échantillons. Bien que les questions liées à l'échantillonnage et au traitement des échantillons de sol soient essentielles pour avoir une bonne représentation du développement des systèmes racinaires, cette discussion n'abordera pas ces questions. Elles ont été abordées dans la revue bibliographique et seule la question du nombre de carottes de sol à prélever a fait l'objet d'une analyse dans le cadre d'un travail de fin d'études lié à cette thèse²⁸. Les conclusions de ce travail sont que le prélèvement de quatre à six carottes de sol par parcelle expérimentale constitue le meilleur compromis entre la qualité des données acquises et le temps nécessaire pour acquérir ces données. Ce nombre de prélèvements permettrait de mettre en évidence des différences significatives entre des modalités ayant entre elles des différences relatives de l'ordre de 30 à 25% respectivement.

A partir de l'étape d'acquisition des images hyperspectrales, les erreurs de quantifications racinaires sont dues au recouvrement des échantillons, à la non détection des racines due à leur taille inférieure à la taille des pixels, à la présence de pixels mixtes sur les images, aux erreurs de classification par les modèles de discrimination et aux erreurs de quantification de la masse racinaire qui est basée sur un nombre de pixels de racines détectés sur les images hyperspectrales.

2.1. Contraintes liées au traitement des échantillons

Les échantillons de racines prélevés par carottage de sol dans les conditions de croissance des plantes doivent être lavés et séchés avant de pouvoir être analysés par imagerie hyperspectrale proche infrarouge. En ce sens, il n'y a pas d'évolution par rapport au traitement classique des carottes de sol.

L'étape de lavage est nécessaire pour extraire les racines du volume du sol prélevé. Sans cette étape, les racines ne peuvent pas être visualisées au sein des échantillons de sol, même si ceux-ci sont étalés en couche mince sur le support d'acquisition de la caméra hyperspectrale. Cette étape est généralement réalisée de manière manuelle.

²⁸ Travail de fin d'études réalisé par Adrien Boufflette. "Etude de l'influence de la fertilisation azotée sur le développement du système racinaire du froment d'hiver (*Triticum aestivum* L.) par imagerie hyperspectrale proche infrarouge". Gembloux Agro-Bio Tech, Université de Liège. Année académique 2017-2018.

Elle est chronophage, utilise de grandes quantités d'eau (plusieurs litres par échantillons) et peut mener à des pertes importantes de racines si les mailles du tamis utilisé pour récupérer l'échantillon sont trop grandes par rapport à l'épaisseur des racines (Livesley et al. 1999; Amato & Pardo 1994).

L'étape de séchage est quant à elle nécessaire pour retirer l'eau libre présente au sein des échantillons car l'eau a un signal fort dans le proche infrarouge et elle risque dès lors de masquer des informations utiles à la discrimination des spectres.

2.2. Contraintes rencontrées lors de l'acquisition des images hyperspectrales

2.2.1. Recouvrement au sein des échantillons

Le recouvrement des racines entre elles (et le recouvrement des racines par d'autres éléments de l'échantillon lorsque celui-ci ne contient pas que des racines) est le premier facteur qui biaise les mesures effectuées par imagerie hyperspectrale car il intervient avant même que les images soient acquises. Un recouvrement des éléments de l'échantillon lors de l'acquisition des images ne pourra d'aucune manière être corrigé et influence donc le reste du processus de quantification.

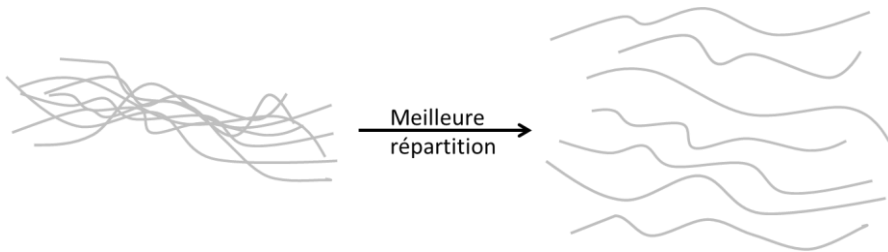


Figure 66: Effet de la répartition de racines sur la quantité de pixels de racines pouvant être enregistrée. Une meilleure répartition des racines entraîne une diminution de leur recouvrement, une augmentation de la surface pouvant être perçue et donc du nombre de pixels de racines détectables.

Cet effet de la répartition des échantillons est illustré à la Figure 66 avec des racines symbolisées par des lignes. Selon la manière dont les racines sont disposées, le nombre de pixels de racines pouvant être acquis est variable. Une répartition sans aucun recouvrement permettrait l'acquisition du plus grand nombre de pixels de racines. Néanmoins, vu la nature des échantillons analysés, le nombre d'éléments qu'ils contenaient et le fait que les racines et les résidus de cultures étaient souvent entremêlés et agglomérés, une répartition parfaite était en pratique impossible. Les échantillons étaient donc tant que possible étalés, mais un juste compromis a dû être trouvé entre le temps gagné par la quantification de racines par imagerie hyperspectrale par rapport à un tri manuel et le temps passé à étaler les échantillons.

Les points d'attention principaux lors de cette répartition étaient donc l'étalement des agglomérats de racines, l'écartement des résidus de cultures les plus larges et la régularité dans la manière d'étendre les échantillons sur la bande transporteuse d'acquisition en acquisition. Cette étape pouvait donc dépendre de la subjectivité de l'opérateur.

Ce problème de recouvrement n'est pas rencontré lorsque l'imagerie hyperspectrale proche infrarouge est utilisée sur des échantillons de grains, des poudres ou des denrées alimentaires (fruits, viande ...) qui de par leur taille peuvent être plus facilement répartis sur la bande transporteuse ou pour lesquels l'analyse ne se base pas sur un certain nombre de pixels visibles, mais plutôt sur l'allure du spectre acquis ou sur des ratios de pixels (Lim et al. 2016; Mishra et al. 2015; Fernández Pierna et al. 2014; Vermeulen et al. 2012). Dans le cadre de la méthode de quantification des racines mise au point dans cette étude, la masse de racines est par contre déterminée sur base du nombre de pixels classés comme étant des racines sur les images hyperspectrales. Le recouvrement de racines entre elles ou par des résidus de cultures entraîne donc une diminution du nombre de pixels visibles sur l'image et une diminution de la masse de racine prédite pour l'échantillon analysé.

2.2.2. Taille des racines et résolution de la caméra hyperspectrale

Comme l'ont mentionné Flavel et al. (2012) dans un article sur la quantification de racines de céréales par tomographie aux rayons X et à haute résolution, la limite pour observer la plus petite racine est fonction de la qualité de l'image (son rapport signal sur bruit) et de sa résolution. Ils ajoutent qu'il est couramment reconnu que le plus petit objet détectable sur une image doit avoir au moins la taille de deux pixels. Dans le cas de notre étude, beaucoup de racines avaient un diamètre inférieur à la taille du pixel (0,31 mm). La détection de ces racines très fines était dès lors difficile.

Une solution simple à ce problème serait d'utiliser un objectif plus puissant lors de l'acquisition des images. La résolution de l'image serait alors meilleure et les racines seraient plus facilement détectées sur les images hyperspectrales. Néanmoins, le grossissement de l'objectif détermine la taille de la plage d'acquisition de la caméra; si l'objectif est plus puissant, la plage d'acquisition est réduite et l'étalement des échantillons est d'autant plus difficile. Le choix de la résolution spatiale des images est donc un paramètre qui doit s'adapter à la taille des échantillons et de ce qui doit y être détecté (Amigo et al. 2015; Amigo 2010). Il faut espérer que l'évolution des systèmes d'acquisition des images hyperspectrales permettra d'améliorer la résolution des images sans réduire la taille de la plage d'acquisition.

2.2.3. Présence d'ombres sur le support d'acquisition des images

Des ombres peuvent être présentes sur les images hyperspectrales. Ces ombres sont dues à une illumination non homogène de l'échantillon. Leur importance sur les images hyperspectrales dépend de la configuration du système d'acquisition. Les systèmes d'acquisition dans lesquels l'illumination est réalisée de manière diffuse permettent de réduire au maximum les zones d'ombre sur l'échantillon et le support d'acquisition des images. Le système d'acquisition utilisé dans nos recherches n'était pas conçu avec un système d'éclairage diffus mais avec une source lumineuse de

part et d'autre de l'échantillon. Les zones d'ombres ont donc dû être prises en compte lors du traitement des images. Les spectres de ces zones d'ombre ont donc été inclus dans la classe spectrale de l'arrière-plan lors de la construction des modèles de discrimination.

2.2.4. Présence de pixels mixtes

Un certain nombre de pixels d'une image hyperspectrale peuvent être acquis en même temps sur des éléments de natures différentes. Ce sont les pixels mixtes. Le spectre de ces pixels est alors le résultat d'un mélange de spectres de ces éléments.

La prise en compte des spectres de ces pixels est importante. Si les spectres de ces pixels sont inclus dans les sets de calibration des modèles de discrimination, ils risquent de déformer ces modèles et de fausser les prédictions sur les images hyperspectrales. En effet, utiliser comme références, pour des classes qui doivent être bien séparées, des spectres contenant un mélange d'informations de ces classes mène inévitablement à une confusion des classes. Il faut donc bien identifier toutes les classes spectrales présentes au sein des images hyperspectrales (racines, résidus de cultures, ombres...), extraire des spectres correspondant à ces classes et n'utiliser que ces spectres pour la construction des modèles.

Dans le cas de notre étude, la plupart des racines étaient très fines et avaient une épaisseur inférieure à la taille d'un pixel. Les spectres acquis sur ces racines fines, mais également sur le bord de racines plus épaisses, contenaient donc l'information de la racine, mais aussi de l'arrière-plan. Pour calibrer nos modèles de discrimination, nous avons donc tant que possible pris les spectres de racines en dehors de ces zones de confusion et donc de préférence au centre de racines ayant une largeur de plusieurs pixels. Les spectres de l'arrière-plan ont également été pris en compte et ont constitué une classe à part dans les modèles de discrimination. Ces spectres ont été acquis dans les zones d'ombre des images hyperspectrales, le long des racines et des résidus de cultures. Cette classe spectrale contenait également les spectres du support d'acquisition des images hyperspectrales²⁹. Pour améliorer encore les prédictions, le choix de l'algorithme de discrimination utilisé pour classer les spectres a également été pris en compte. Comme nous l'avons montré lors de la comparaison des performances des modèles de discrimination, l'algorithme SVM prédisait mieux les pixels mixtes que l'algorithme PLS.

2.2.5. Instabilité du système d'acquisition

Le système d'acquisition utilisé dans cette étude montrait une certaine instabilité dans les valeurs d'absorbance qui étaient enregistrées pour un même échantillon au cours d'une journée d'acquisition. Sans prise en compte de cette instabilité, la

²⁹ Lors de l'analyse des images hyperspectrales, les spectres de la bande transporteuse de la caméra sont en principe éliminés par seuillage avant l'application des modèles de discrimination. Néanmoins, certains spectres de la bande transporteuse, notamment ceux qui sont acquis dans la zone de soudure de la bande, ne sont pas éliminés par ce seuillage. Des spectres représentatifs de cette zone de soudure ont donc été inclus dans les modèles de manière à être classés avec l'arrière-plan.

prédiction de la nature des pixels des images hyperspectrales acquises se dégradait au cours du temps.

Pour résoudre ce problème, deux mesures ont été prises dans notre étude: (i) le système d'acquisition a été calibré régulièrement (calibration de l'absorbance et détection des pixels défectueux) et (ii) des spectres acquis à différents moments au cours des journées d'acquisition ont été inclus dans le jeu de données servant de référence à la construction des modèles.

2.3. Contraintes liées à la construction des modèles chimiométriques

2.3.1. Acquisition des données de référence

L'imagerie hyperspectrale proche infrarouge est une méthode d'analyse indirecte. Pour pouvoir extraire l'information des échantillons analysés, des modèles chimiométriques sont utilisés. Ceux-ci sont construits sur des données de références pour pouvoir ensuite classer ou quantifier les éléments présents au sein des échantillons à analyser. L'identification de ces données de référence est importante car elle détermine le domaine d'application des modèles.

Comme nous l'avons vu au chapitre sur la quantification racinaire, les racines et les résidus de cultures ont des spectres différents. Cela s'observe clairement sur des spectres moyens de racines et de résidus de cultures (Figure 49, Chapitre VI) ou lorsque ces éléments sont broyés avant l'acquisition des images hyperspectrales. Cependant, lorsque les spectres sont enregistrés pour chaque pixel d'une image hyperspectrale acquise sur un échantillon brut, les variations au sein des spectres de racines et de résidus de cultures sont élevées et la discrimination est beaucoup plus difficile. Comme le notaient Rodionova et al. (2005), bien que les spectres médians de deux classes spectrales soient bien distincts, les spectres de ces classes prises dans leur ensemble peuvent ne pas être bien séparés. L'illumination non uniforme, les effets d'ombres et des impuretés au sein des échantillons sont plusieurs facteurs qui expliquent les variations observées au sein des spectres. Pour maximiser les chances de classer correctement un spectre il faut donc inclure une grande diversité de spectres pour chaque classe spectrale au sein des sets de calibration des modèles chimiométriques. Dans notre étude, le choix d'un grand nombre d'échantillons pour la calibration de chaque classe spectrale, la sélection des spectres du set de données sur base de leur variabilité et l'inclusion régulière de nouveaux spectres dans le jeu de donnée de référence des modèles a permis d'augmenter la robustesse des modèles chimiométriques. Comme nous l'avons vu dans le chapitre dédié au suivi de l'instabilité du système d'acquisition des images hyperspectrales, l'inclusion dans le set de calibration de spectres acquis sur un même échantillon, mais lors de différentes journées d'acquisition, ou à différentes heures d'une même journée d'acquisition, a permis d'améliorer la discrimination des racines et des résidus de cultures.

C'est à cette étape du développement des modèles que la subjectivité de l'opérateur peut avoir le plus d'impact. Dans le cas de la discrimination des racines et des résidus de cultures, cette subjectivité peut entraîner des erreurs à deux moments: lors

du tri des échantillons extraits de carottes de sol (il faut correctement identifier les racines et les résidus) et lors de la sélection des spectres de référence sur les images hyperspectrales (il faut sélectionner les zones d'intérêt sur les images en y incluant un maximum de variabilité pour chaque classe mais sans sélectionner des pixels mixtes qui pourraient par la suite déformer les modèles de discrimination).

2.3.2. Traitement des spectres

Un autre point important lors de la construction des modèles chimiométriques est le choix du prétraitement des spectres et du modèle. En effet, parmi les outils de la chimiométrie, il existe de nombreux prétraitements de spectres et plusieurs modèles de quantification ou de discrimination.

Pour choisir le prétraitement et le modèle qui doivent être utilisés, il n'existe aucune règle générale. Le traitement optimal pour des spectres donnés dépend en effet de plusieurs paramètres tels que les caractéristiques de l'échantillon et la configuration du système d'acquisition des images par exemple. Le choix du prétraitement doit donc se baser sur un processus d'essais-erreurs guidé par l'expérience de la personne qui traite les données. Il faut également noter qu'un prétraitement des données ne doit être utilisé que s'il aide réellement à améliorer les performances des modèles. Dans le cas de la discrimination des racines et des résidus de cultures, le prétraitement par défaut (*autoscale*) du logiciel utilisé (PLS Toolbox 7.8, Eigenvector Research Inc., Wenatchee, WA, USA) a permis d'obtenir les meilleurs résultats. Il en va de même pour le choix du modèle. Dans le cas de notre étude, c'est la comparaison de deux modèles de discrimination sur une large quantité de données qui a permis de conclure que le modèle de discrimination construit avec l'algorithme SVM convenait le mieux pour discriminer les spectres de racines et de résidus de cultures.

2.4.Limites des modèles chimiométriques

Un modèle est une vision simplifiée de la réalité. Il comporte dès lors une série d'erreurs qui peuvent, dans le cas de modèles de discrimination, mener à des confusions lors des prédictions. L'interprétation des données doit donc en tenir compte.

Dans le cadre de l'étude que nous avons menée sur la discrimination de racines et de résidus de cultures, nous avons montré qu'en moyenne 8,3 % des pixels de racines étaient prédits comme étant des résidus de cultures avec le modèle SVM (Chapitre V, Tableau 2). Cette valeur a été obtenue en appliquant l'arbre de classification dichotomique, permettant de classer des spectres de l'arrière-plan, des particules de sol, de racines et de résidus de cultures, sur des images acquises sur des échantillons contenant uniquement des racines. Les calculs réalisés pour obtenir ce résultat ne tenaient pas compte des pixels de racines prédits comme étant de l'arrière-plan (support de l'échantillon et zones d'ombre). Si on s'intéressait au pourcentage de pixels prédits comme étant de l'arrière-plan après le retrait des pixels du support d'acquisition des images, nous obtenions une valeur de 56 %. Les 44 % de pixels restant étaient ceux qui étaient prédits comme étant des racines, des résidus de cultures et des particules de sol. Ces pixels correspondaient à des zones d'ombre,

mais également à des racines trop petites pour être correctement détectées par le système d'acquisition (racines de taille inférieure à la taille des pixels). L'ensemble des pixels de racines prédits comme étant des résidus de cultures ou de l'arrière-plan échappaient donc à la quantification et entraînaient une sous-estimation des racines présentes au sein de l'échantillon. Les erreurs de prédictions sur des échantillons de racines sont illustrées à la Figure 67.

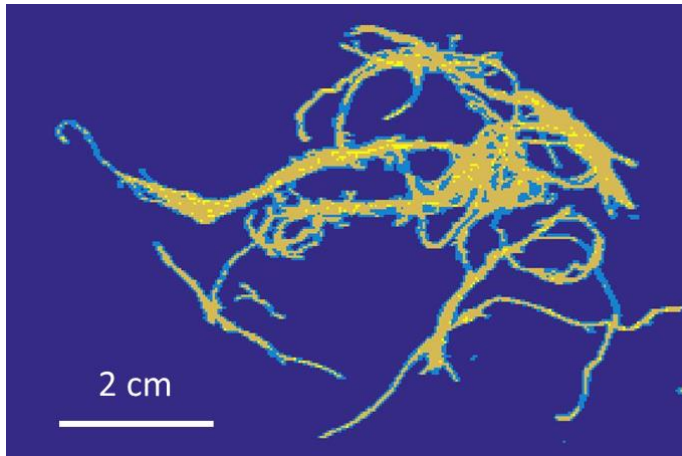


Figure 67: Prédiction d'une image hyperspectrale proche infrarouge acquise sur un échantillon de racines. Un certain nombre de pixels sont mal prédits par les modèles chimiométriques de classification. Les pixels correctement prédits comme étant des racines sont en ocre, ceux qui sont prédits comme étant des résidus de cultures ou de l'ombre sont en jaune et bleu clair respectivement. La bande transporteuse du système d'acquisition est en bleu foncé.

La confusion existe également dans l'autre sens : des pixels non acquis sur des racines (sur des résidus de cultures par exemple) peuvent être prédits comme étant des racines et dès lors entraîner une surestimation de la quantité de racines au sein des échantillons. Les résultats obtenus dans notre étude sur la discrimination des spectres de racines et de résidus de cultures montraient que l'algorithme SVM prédisait en moyenne 17,4 % des pixels de résidus de cultures comme étant des racines.

Une des raisons permettant d'expliquer la classification des spectres de certains pixels dans la mauvaise classe spectrale est la présence des pixels mixtes sur les images hyperspectrales. En effet, les résultats obtenus sur les spectres sélectionnés pour le set de validation des modèles étaient mieux prédits. Ces spectres étaient extraits des images hyperspectrales de telle manière à éviter les pixels mixtes. Dans ce cas, le modèle SVM prédisait correctement 94 % des spectres de résidus de cultures.

Une autre raison qui peut être évoquée pour expliquer la mauvaise classification de certains spectres est la prise en compte d'une diversité trop faible de spectres lors de la construction des modèles. Le cas le plus flagrant est les erreurs de prédictions apparaissant lorsque les modèles sont appliqués sur des images acquises sur des échantillons contenant de nouveaux éléments, tels que des résidus d'autres cultures par exemple. Lorsque ce cas se présente, de nouveaux modèles doivent être construits en y incluant les spectres de ces nouveaux éléments.

Les erreurs de mesures sur des images de racines dues à la confusions des racines et des résidus de cultures avaient déjà été mise en évidence par Richner et al. (2000) face à l'utilisation de logiciels de reconnaissances de formes pour discriminer des résidus de cultures. Dans ce cas, ils expliquaient que bien qu'un rapport de la longueur sur la largeur puisse exclure la majorité des objets étrangers avant de réaliser des mesures morphologiques sur des racines, un nombre inconnu de ces objets était considéré comme étant des racines et un nombre inconnu de racines était exclu à tort.

2.5.Limites de la quantification racinaire basée sur un nombre de pixels

Comme nous l'avons déjà signalé, les racines étudiées dans le cadre de nos recherches ayant fréquemment une épaisseur inférieure à la taille d'un pixel, leurs spectres étaient souvent prédits comme étant de l'arrière-plan étant donné qu'ils contenaient également la signature spectrale de l'ombre située en bordure de ces racines. De telles prédictions menaient à une sous-estimation de la quantité de racines présente au sein des échantillons. Ces pixels acquis sur des racines fines pouvaient également être prédits comme étant des racines alors que celles-ci n'occupaient pas toute la surface des pixels sur lesquels les spectres étaient acquis. Ces prédictions menaient alors à une surestimation de la quantité de racines. De tels effets ont été observés dans des recherches ayant pour but la détection de contaminants dans de la poudre de lait (Lim et al. 2016) et de la farine de blé (Mishra et al. 2015). Des spectres acquis en bordure de racines larges pouvaient mener aux mêmes effets.

Afin de mieux visualiser ces problèmes, deux racines d'épaisseurs différentes ont été représentées à la Figure 68. Dans les deux cas, elles sont centrées sur une surface de neuf pixels. Dans le premier cas, les trois pixels acquis sur la racine contiennent aussi bien l'information de la racine que de l'arrière-plan et peuvent donc potentiellement être classés comme appartenant à l'une ou l'autre classe par les modèles de discrimination. Ces pixels correspondent aux pixels mixtes définis précédemment. Pour rappel, les modèles utilisés sont des modèles de classification binaire. Ils ne considèrent donc les spectres que comme s'ils étaient acquis sur des constituants purs. Si ces spectres sont classés dans l'arrière-plan, cela mène à une sous-estimation de la masse racinaire de l'échantillon. Si par contre ils sont classés comme étant des racines, cela mène à une surestimation de la masse racinaire étant donné la faible surface de racine sur les pixels. Dans le second cas de figure, les trois pixels acquis au centre de la racine ont le plus de chance d'être correctement classés

comme ayant été acquis uniquement sur des racines. Pour les six autres pixels acquis en bordure de la racine, les spectres pourront potentiellement être classés comme étant acquis sur des racines ou l'arrière-plan, comme dans le premier cas de figure.



Figure 68: Représentation schématique de l'acquisition de spectres sur des racines de différentes épaisseurs. Chaque carré représente la surface d'un pixel. Les spectres des pixels qui ne sont pas acquis entièrement sur une racine peuvent être mal prédits par les modèles de discrimination.

Une autre source d'erreur apparaît dans cette estimation de la masse racinaire du fait que les calculs sont basés sur un nombre de pixels, c'est-à-dire une unité de surface (2D), pour estimer la masse de racines qui ont un certain volume (3D).

3. Perspectives dans l'utilisation de l'imagerie hyperspectrale proche infrarouge et de la chimométrie dans l'étude de systèmes racinaires

Au cours des travaux menés durant cette thèse, de nombreuses idées ont germé. Certaines seulement ont pu être testées. Les paragraphes qui suivent présentent un certain nombre d'améliorations qui pourraient être apportées dans l'utilisation de l'imagerie hyperspectrale proche infrarouge et des outils de la chimométrie dans l'étude de systèmes racinaires. Certaines d'entre elles ont été évoquées dans d'autres études et ont parfois déjà été testées.

3.1. Discrimination et quantification de racines vivantes et mortes

Picon-Cochard et al. (2009) ont montré qu'il était possible de déterminer si une racine était vivante ou morte sur base de son spectre visible-proche infrarouge. Leur étude portait sur cinq espèces de graminées prairiales. Cette différenciation était entre autres possible grâce aux teneurs en carbone différentes entre les racines vivantes et mortes. Cette quantification séparée de racines vivantes et mortes n'a pas été réalisée dans le cadre de notre étude; toutes les racines ont été considérées sans faire la distinction entre les racines vivantes et mortes. Or, au cours du développement de la plante, la production de racines mortes a lieu simultanément au développement de nouvelles racines et il y a donc naturellement un mélange de

racines mortes et vivantes dans le sol (Picon-Cochard et al. 2009). Dans les observations que nous avons réalisées, nous avons pu mesurer que la biomasse racinaire était plus faible en fin de cycle des cultures que lors d'autres prélèvements réalisés durant la culture. Plusieurs auteurs ont mis en évidence que le maximum de production racinaire était atteint au moment de la floraison de la culture et qu'il y avait ensuite une disparition d'un certain nombre de racines (Plaza-Bonilla et al. 2014; Muñoz-Romero et al. 2010; Kätterer et al. 1993). Un certain nombre de racines au sein des échantillons étudiés devaient donc être mortes.

Afin d'estimer de manière plus fine la quantité de racines actives lors des prélèvements de carottes de sol, le développement d'un modèle de discrimination supplémentaire permettant de séparer les racines mortes et vivantes au sein des échantillons pourrait constituer une perspective intéressante.

3.2. Mesures de paramètres de l'architecture de systèmes racinaires

La technique mise au point dans cette étude permet d'estimer une masse racinaire au sein d'un volume de sol. Elle ne permet actuellement pas de mesurer des paramètres tels que la longueur racinaire, le diamètre racinaire ou la surface racinaire, des paramètres qui sont souvent étudiés car ils donnent des informations sur la capacité de la plante à prélever l'eau et les nutriments dans le sol (Bengough et al. 2000) et sur la manière dont les racines sont impactées par leur milieu de culture (Cai et al. 2014; Qin et al. 2004).

Ces paramètres peuvent être mesurés par des logiciels d'analyse d'images. Cela n'a pas été testé dans cette étude, mais on pourrait imaginer de coupler la prédiction de la nature de chaque pixel des images hyperspectrales à un logiciel d'analyse d'image qui permettrait de mesurer ces paramètres.

Pour pouvoir réaliser des mesures sur des images, la résolution de l'image doit être élevée. Une résolution de 600 ppp³⁰ est souvent utilisée. Or, avec le système d'acquisition utilisé dans cette étude, la résolution de l'image était de 320 pixels sur une largeur de 10 cm ce qui équivaut à une résolution d'environ 80 ppp. Il faudrait donc augmenter la résolution par un facteur de 7,5 pour pouvoir faire des mesures de longueur et de surface suffisamment précises sur les images. Cependant, une telle résolution d'image n'est pour l'instant pas atteignable en imagerie hyperspectrale.

3.3. Couplage à de la reconnaissance de formes

Afin de réduire la confusion entre les racines et d'autres éléments présents sur les images hyperspectrales, les résidus de cultures par exemple, les images prédites suite à la classification des spectres par les modèles de discrimination pourraient être analysées avec un logiciel de reconnaissance de forme.

Plusieurs études ont déjà montré l'intérêt de la reconnaissance de formes sur des images d'échantillons de racines contenant des matières organiques étrangères

³⁰ ppp: pixels par pouce

(Wasson et al. 2016; Dowdy et al. 1998). Ces études utilisaient la forme des éléments de l'échantillon pour leur discrimination et étaient donc limitées aux éléments ayant des formes bien distinctes.

L'avantage de l'imagerie hyperspectrale proche infrarouge dans la discrimination de racines et d'autres éléments présents au sein d'échantillons est que les analyses sont basées sur les propriétés physico-chimiques des éléments plutôt que sur leur forme. Les confusions sont donc en principe moins importantes. Néanmoins, une certaine confusion a été observée lors de la prédiction de la nature des éléments présents au sein des échantillons analysés dans le cadre de notre étude. Pour rappel, 17 % des pixels acquis au niveau de résidus de cultures étaient prédits comme étant des racines et nous avons pu observer que ces pixels mal prédits se trouvaient principalement en bordure des résidus de cultures (revoir Figure 41 au Chapitre V). On peut imaginer que l'utilisation de logiciels de reconnaissance de formes pourrait permettre de reclasser correctement des pixels présents en bordure ou au centre d'un résidu de culture et prédits comme étant des pixels de racines. Une meilleure classification de ces pixels permettrait de diminuer le nombre de pixels prédits comme étant des racines sans en être, et donc de diminuer la surestimation de la masse de racines qui est due à la présence de résidus de cultures. L'inverse serait vrai aussi. De nombreux pixels acquis au centre de grosses racines étaient prédits comme étant des résidus de cultures et un meilleur classement de ces pixels permettrait de réduire l'écart entre la masse de racine prédite et la masse de racine réellement présente dans l'échantillon de sol.

Une telle combinaison de l'imagerie hyperspectrale avec des logiciels de reconnaissance de forme a déjà été utilisée dans nos recherches pour détecter des nodosités individuelles sur des systèmes racinaires de pois. Elle a également été utilisée par Kamruzzaman et al. (2012) qui utilisa cette combinaison pour la caractérisation et l'authentification de viandes et par (Fernández Pierna et al. 2012) pour détecter la présence de kystes sur des racines de betteraves.

3.4. Identification des longueurs d'onde d'intérêt

Selon Wu et Sun (2013), inclure l'ensemble des longueurs d'onde du spectre n'augmente dans la plupart des cas pas les performances des modèles étant donné que certaines longueurs d'onde contiennent des informations non pertinentes et que d'autres ont un faible rapport signal-sur-bruit. L'élimination de ces longueurs d'onde peut faciliter le processus de calibration et améliorer les résultats en termes de précision et de robustesse. Kamruzzaman et al. (2012) notent également qu'éliminer les régions spectrales n'ayant pas ou peu de contribution lors de la calibration des modèles permet de diminuer le temps nécessaire à l'acquisition et au traitement des images hyperspectrales.

L'utilisation d'un faible nombre de longueurs d'onde a été testée dans nos recherches pour la quantification de la leghémoglobine au sein de nodosités de pois et a montré de très bons résultats. La sélection des longueurs d'ondes y a été réalisée en analysant les coefficients de la régression PLS et en identifiant les longueurs d'ondes qui étaient les plus significatives dans l'explication des variations observées

au sein des teneurs en leghémoglobine. Les potentialités de l'analyse multispectrale n'ont par contre pas été testées sur les racines et les résidus de cultures.

3.5. Amélioration d'autres méthodes de quantification racinaire

La méthode de quantification de racines à l'aide de l'imagerie hyperspectrale proche infrarouge et la chimiométrie a été spécifiquement développée pour réduire le temps nécessaire au tri des racines et de résidus de cultures extraits de carottes de sol. Néanmoins, l'utilisation de l'imagerie hyperspectrale proche infrarouge pourrait être également combinée à d'autres méthodes de quantification racinaire et améliorer l'efficacité de celles-ci.

3.5.1. Utilisation de l'imagerie hyperspectrale proche infrarouge dans des minirhizotrons

Une première amélioration identifiée est celle des minirhizotrons. L'utilisation de caméras hyperspectrales pour faire le suivi racinaire en cours de croissance des plantes pourrait permettre de discriminer les racines et d'autres matières organiques présentes dans les sols, de discriminer des racines de différentes espèces et de discriminer les racines vivantes de racines mortes.

Une proposition similaire avait été faite par Pierret (2008) et Rewald et al. (2012). Selon Rewald et al. (2012), la spectroscopie infrarouge est la méthode la plus prometteuse pour l'identification non destructive des racines et leur quantification en combinaison avec les analyses réalisées en (mini-)rhizotrons. En se basant sur les travaux de Nakaji et al. (2008) qui ont permis la classification d'éléments de la rhizosphère sur base d'images hyperspectrales visible-proche infrarouge, Pierret (2008) écrit que l'inclusion de l'imagerie multispectrale visible-proche infrarouge dans les minirhizotrons permettra d'améliorer cette méthode d'étude des systèmes racinaires.

3.5.2. Utilisation de l'imagerie hyperspectrale proche infrarouge avec la méthode *core break*

Une seconde amélioration qui pourrait être envisagée est celle de la méthode *core break*. Récemment, plusieurs études ont déjà été menées pour réaliser le comptage de racines dans des carottes de sol par spectroscopie, directement au champ et sans lavage (Xu et al. 2016; Wasson et al. 2016; Kusumo et al. 2010). Bien que l'estimation de la quantité de racine produite par ces méthodes soit moins bonne qu'avec un carottage de sol et une extraction des racines par lavage, ces nouvelles méthodes ont l'avantage d'être beaucoup plus rapides étant donné que les spectres sont acquis directement sur les carottes de sol sans devoir les laver. Dans les résultats présentés par ces auteurs, seules les racines d'une espèce étaient étudiées. La mise au point de modèles chimiométriques permettant de discriminer les spectres d'autres éléments présents dans les carottes de sol ou les racines d'autres espèces pourrait être un développement intéressant.

3.6. Amélioration de la qualité des images acquises

Les limitations de l'imagerie hyperspectrale proche infrarouges que nous avons identifiées dans la quantification des racines étaient en grande partie dues au système d'acquisition. Elles étaient dues à la résolution spatiale des images, aux zones d'ombre présentes sur les images et à l'instabilité du système d'acquisition.

Afin d'améliorer la qualité des images hyperspectrales et obtenir ainsi de meilleures prédictions des masses racinaires sur ces images, une augmentation de la résolution spatiale de ces images et une amélioration des conditions d'illumination des échantillons seraient nécessaires. Selon les résultats rapportés dans différentes publications, l'utilisation d'une lumière diffuse permettrait de réduire les zones d'ombre sur les images hyperspectrales (Boldrini et al. 2012) et l'emploi de lampes LED, au lieu des lampes halogènes, permettrait de stabiliser la production lumineuse et d'éliminer les variations au sein des spectres dues aux variations de températures (Wu & Sun 2013).

Conclusions générales

En combinant l'imagerie et la spectroscopie, l'imagerie hyperspectrale proche infrarouge permet d'acquérir un spectre pour chaque pixel d'une image. Combinée à la chimiométrie, elle permet de donner la répartition spatiale des caractéristiques physico-chimique d'un échantillon.

Dans le cadre de notre étude, cette technique d'analyse a démontré sa capacité à discriminer des racines de différentes espèces, des nodosités, des résidus de cultures et des particules de sol extraites par lavage d'échantillons de sol. Nous avons montré que pour obtenir la meilleure discrimination, il était nécessaire d'identifier toutes les sources de variabilités spectrales présentes sur les images hyperspectrales et d'en tenir compte lors de la construction des modèles de discrimination. Cette variabilité était due à la nature des éléments analysés, à la présence de zones d'ombres et de pixels mixtes sur les images hyperspectrales et à la variabilité des mesures induite par le système d'acquisition. Pour diminuer cette dernière source de variabilité, nous avons montré qu'un étalonnage fréquent du système d'acquisition était nécessaire.

Lors de la mise au point des modèles de discrimination, nous avons montré l'intérêt de tester ces modèles sur des images hyperspectrales et non pas uniquement sur des spectres extraits de ces images. Cela nous a permis de valider les modèles de discrimination lorsqu'ils étaient combinés au sein d'un arbre de classification dichotomique, d'identifier les zones de confusion spectrale sur les images et d'estimer la capacité des modèles à discriminer les éléments extraits d'échantillons de sol. Nous avons également montré, lors d'une comparaison des performances des modèles chimiométriques PLS-DA et SVM, que l'algorithme SVM permettait de classer correctement une plus grande part de pixels et donnait de meilleurs résultats dans les zones de confusion spectrale.

Bien que la discrimination des spectres de racines et de résidus de culture n'était pas parfaite, la procédure développée nous a permis de se passer de l'étape de tri manuel qui était jusqu'à présent nécessaire dans les études de systèmes racinaires basées sur le carottage de sol. Cette étape était chronophage, laborieuse et sensible à subjectivité de l'opérateur. Elle constituait dès lors le frein le plus important dans l'étude racinaire par carottage de sol car elle limitait la fréquence d'échantillonnage.

Suite à la détection des pixels de racines et de résidus de cultures sur les images hyperspectrales, nous avons montré qu'il existait une relation linéaire entre le nombre de pixels de ces éléments et leur masse dans les échantillons. Cette relation était d'autant meilleure lorsque les échantillons étaient fortement étalés sur le support d'acquisition des images hyperspectrales.

En se basant sur la détection des racines au sein d'échantillons de sols et le lien avec leur masse, nous avons pu estimer la densité racinaire au sein d'échantillons de sol prélevés par carottage sous des cultures de froment d'hiver et de pois d'hiver menées selon différents itinéraires techniques. Cela nous a permis de suivre le développement de systèmes racinaires entre différents prélèvements réalisés au cours d'une saison culturale, d'observer l'effet du travail du sol sur la répartition du

système racinaire de froment au sein de différents horizons de sol et d'étudier l'effet de la culture en association, couplée à différentes fumures azotées, sur le développement racinaire de cultures de froment et de pois d'hiver.

Enfin, l'imagerie hyperspectrale a été utilisée avec succès pour quantifier la teneur en leghémoglobine au sein de nodosités individuelles de pois protéagineux et a permis d'analyser la répartition des teneurs en leghémoglobine sur des systèmes racinaires de pois.

Ces études réalisées sur des systèmes racinaires de froment et de pois ont donc montré plusieurs opportunités qu'offrait l'imagerie hyperspectrale proche infrarouge combinée à la chimiométrie pour étudier des systèmes racinaires. Ces études ont également montré les limites de cette méthode dans la quantification racinaire. Ces limites étaient entre autres liées aux caractéristiques intrinsèques du système d'acquisition des images hyperspectrales (résolution, stabilité et système d'éclairage). La résolution de ces problèmes techniques grâce à l'évolution des technologies de l'imagerie hyperspectrale permettra d'améliorer les performances des méthodes originales d'étude de systèmes racinaires mises au point dans le cadre de cette thèse.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Abbas, O., Dardenne, P. & Baeten, V., 2012. Near-infrared, mid-infrared, and raman spectroscopy. In Y. Picó, ed. *Chemical Analysis of food: Techniques and Applications*. pp. 61–91.
- Agelet, L.E. & Hurburgh, C.R., 2010. A tutorial on Near Infrared Spectroscopy and its calibration. *Critical Reviews in Analytical Chemistry*, 40(4), pp.246–260.
- Amato, M. & Pardo, A., 1994. Root length and biomass losses during sample preparation with different screen mesh sizes. *Plant Soil*, 161(2), pp.299–303.
- Amigo, J.M., 2010. Practical issues of hyperspectral imaging analysis of solid dosage forms. *Analytical and Bioanalytical Chemistry*, 398, pp.93–109.
- Amigo, J.M., Babamoradi, H. & Elcoroaristizabal, S., 2015. Hyperspectral image analysis. A tutorial. *Analytica Chimica Acta*, 896, pp.34–51.
- Atkinson, D., 2000. Root characteristics: why and what to measure. In A. L. Smit et al., eds. *Root Methods, A Handbook*. Springer - Verlag Berlin Heidelberg, pp. 1–32.
- Baeten, V. et al., 2015. Vibrational spectroscopy for rapid control of agro-food products. In *Handbook of Food Analysis*. pp. 591–621.
- Baeten, V. & Dardenne, P., 2007. Applications of Near-infrared Imaging for monitoring agricultural food and feed products. In R. Bhargava & I. W. Levin, eds. *Spectrochemical Analysis Using Infrared Multichannel Detectors*. pp. 283–301.
- Balasubramaniam & Malathi, A., 1992. Comparative study of hemoglobin estimated by Drabkin's and Sahli's methods. *J Postgrad Med*, 38, p.8.
- Bardgett, R.D., Mommer, L. & De Vries, F.T., 2014. Going underground: Root traits as drivers of ecosystem processes. *Trends in Ecology and Evolution*, 29(12), pp.692–699.
- Barker, M. & Rayens, W., 2003. Partial least squares for discrimination. *Journal of Chemometrics*, 17(3), pp.166–173.
- Bengough, A.G. et al., 2000. Sampling strategies, scaling and statistics. In *Root Methods, A Handbook*. Springer, pp. 147–173.
- Benjamin, J.G. & Nielsen, D.C., 2004. A method to separate plant roots from soil and analyze surface area. *Plant Soil*, 267, pp.225–234.
- Benjamin, J.G. & Nielsen, D.C., 2006. Water deficit effects on root distribution of soybean, field pea and chickpea. *Field Crop Res*, 97, pp.248–253.
- Bertrand, D. & Baeten, V., 2006. Instrumentation. In D. Bertrand & E. Dufour, eds. *La spectroscopie infrarouge, 2ème édition*. Paris, pp. 247–305.
- Bochow, M. et al., 2011. Automatic shadow detection in hyperspectral VIS-NIR images. In *the 7th EARSeL Workshop of the Special Interest Group in Imaging Spectroscopy*. Edinburgh, pp. 9–17.
- Bodner, G. et al., 2017. RGB and spectral root imaging for plant phenotyping and physiological research: Experimental setup and imaging protocols. *Journal of Visualized Experiments*, (126), pp.1–21.

- Böhm, W., 1979. *Methods of studying root systems* Springer-Verlag., Berlin Heidelberg.
- Boldrini, B. et al., 2012. Hyperspectral imaging: A review of best practice, performance and pitfalls for in-line and on-line applications. *J Near Infrared Spec*, 20(5), pp.483–508.
- Bousquet, O., 2001. Introduction aux “Support Vector Machines” (SVM)., <http://www.math-info.univ-paris5.fr/~lomn/Cours/DM/Material/ComplementsCours/SVM.pdf>, (27.04.2018)
- Brown, G.R. & Thilenius, J.F., 1976. A low cost machine for separation of roots from soil material. *Journal of Range Management*, 29(6), pp.506–507.
- Burger, J. & Geladi, P., 2005. Hyperspectral NIR image regression part I: calibration and correction. *Journal of Chemometrics*, 19, pp.355–363.
- Burger, J. & Geladi, P., 2006. Hyperspectral NIR image regression part II: dataset preprocessing diagnostics. *Journal of Chemometrics*, 20, pp.106–119.
- Cai, H. et al., 2014. Effect of subsoil tillage depth on nutrient accumulation, root distribution, and grain yield in spring maize. *Crop J*, 2(5), pp.297–307.
- Caporaso, N., Whitworth, M.B. & Fisk, I.D., 2018. Protein content prediction in single wheat kernels using hyperspectral imaging. *Food Chem*, 240, pp.32–42.
- Cheng, W., Coleman, D.C. & Box, J.E., 1990. Root dynamics, production and distribution in agrosystems on the Georgia Piedmont using minirhizotrons. *J Appl Ecol*, 27(2), pp.592–604.
- Chouhan, S. et al., 2008. Enhancement in Leghemoglobin Content of Root Nodules by Exclusion of Solar UV-A and UV-B Radiation in Soybean. *J Plant Biol*, 51(2), pp.132–138.
- Clarke, F.C. et al., 2002. Determination of the information depth and sample size for the analysis of pharmaceutical materials using reflectance near-infrared microscopy. *Applied Spectroscopy*, 56(11), pp.1475–1483.
- Cornuéjols, A., 2002. Une nouvelle méthode d'apprentissage: Les SVM. Séparateurs à vaste marge. *Bulletin de l'AFIA*, 51.
- Cougnon, M. et al., 2013. NIRS calibration strategies for the botanical composition of grass-clover mixtures. In S. Barth & D. Milbourne, eds. *Breeding strategies for sustainable forage and turf grass improvement*. Springer, Dordrecht, pp. 371–374.
- Dale, L.M. et al., 2012. Chemometric tools for NIRS and NIR Hypespectral Imaging. *Bulletin UASVM Agriculture*, 69(1), pp.70–76.
- Dale, L.M. et al., 2013. Discrimination of grassland species and their classification in botanical families by laboratory scale NIR hyperspectral imaging: preliminary results. *Talanta*, 116, pp.149–154.
- Dale, L.M. et al., 2013. Hyperspectral Imaging Applications in Agriculture and Agro-Food Product Quality and Safety Control: A Review. *Appl Spectrosc Rev*, 48(2), pp.142–159.

- Degrune, F. et al., 2016. No favorable effect of reduced tillage on microbial community diversity in a silty loam soil (Belgium). *Agr Ecosyst Environ*, 224, pp.12–21.
- Delory, B.M. et al., 2016. Root-emitted volatile organic compounds: Can they mediate belowground plant-plant interactions? *Plant and Soil*, 402, pp.1–26.
- de Dorlodot, S. et al., 2007. Root system architecture: opportunities and constraints for genetic improvement of crops. *Trends in Plant Science*, 12(10), pp.474–481.
- Dowdy, R.H. et al., 1998. Automated image analyses for separating plant roots from soil debris elutriated from soil cores. *Plant Soil*, 200, pp.91–94.
- Downie, H.F. et al., 2014. Challenges and opportunities for quantifying roots and rhizosphere interactions through imaging and image analysis. *Plant, Cell and Environment*, 38(7), pp.1213–1232.
- Downie, J.A., 2005. Legume haemoglobins: symbiotic nitrogen fixation needs bloody nodules. *Curr Biol*, 15(6), pp.196–198.
- Drabkin, D.L. & J.H. Austin, 1935. Preparation from washed blood cells; nitric oxide hemoglobin and sulfhemoglobin. *J Biol Chem*, 112, pp.51–65.
- Dramaix-Wilmet, M., 2008. Modèles de prédiction - Intérêt de la validation.
- Drew, M.C. & Saker, L.R., 1980. Assessment of a rapid method, using soil cores, for estimating the amount and distribution of crop roots in the field. *Plant and Soil*, 55(2), pp.297–305.
- Eigenvector, 2016. Using cross-validation., http://wiki.eigenvector.com/index.php?title=Using_Cross-Validation, (01.06.2018)
- Elmasry, G. et al., 2007. Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry. *Journal of Food Engineering*, 81, pp.98–107.
- ElMasry, G., Sun, D.-W. & Allen, P., 2012. Near-infrared hyperspectral imaging for predicting colour, pH and tenderness of fresh beef. *Journal of Food Engineering*, 110, pp.127–140.
- Eshel, A. & Beeckman, T., 2013. *Plant roots. The hidden half*, Taylor and Francis Inc., Boca Raton.
- Eylenbosch, D. et al., 2016. Lifetime stability study of a NIR hyperspectral imaging system. In *IASIM, 6th conference in spectral imaging*. Chamonix-Mont-Blanc, p. 21.
- Eylenbosch, D. et al., 2017. NIR hyperspectral imaging spectroscopy and chemometrics for the discrimination of roots and crop residues extracted from soil samples. *J Chemom*, 32, pp.1–11.
- Eylenbosch, D. et al., 2018. Quantification of leghaemoglobin content in pea nodules based on near infrared hyperspectral imaging spectroscopy and chemometrics. *J. Spectral Imaging*, 7(a9), pp.1–10.

- Faber, K. & Kowalski, B.R., 1997. Propagation of measurement errors for the validation of predictions obtained by principal component regression and partial least squares. *Journal of Chemometrics*, 11, pp.181–238.
- Fan, J. et al., 2016. Root distribution by depth for temperate agricultural crops. *Field Crop Res*, 189, pp.68–74.
- Fehrenbacher, J.B. & Alexander, J.D., 1955. A Method for Studying Corn Root Distribution Using a Soil-Core Sampling Machine and Shaker-Type Washer '. *Agronomy Journal*, 47(10), pp.468–472.
- Fernández Pierna, J.A. et al., 2004. Combination of support vector machines (SVM) and near-infrared (NIR) imaging spectroscopy for the detection of meat and bone meal (MBM) in compound feeds. *J Chemometr*, 18, pp.341–349.
- Fernández Pierna, J.A. et al., 2011. Comparison of various chemometric approaches for large near infrared spectroscopic data of feed and feed products. *Anal Chim Acta*, 705, pp.30–34.
- Fernández Pierna, J.A. et al., 2014. Line Scan Hyperspectral Imaging Spectroscopy for the Early Detection of Melamine and Cyanuric Acid in Feed. *Journal of Near Infrared Spectroscopy*, 22(2), pp.103–112.
- Fernández Pierna, J.A. et al., 2012. NIR hyperspectral imaging spectroscopy and chemometrics for the detection of undesirable substances in food and feed. *Chemometr Intell Lab*, 117, pp.233–239.
- Fernández Pierna, J.A., Baeten, V. & Dardenne, P., 2006. Screening of compound feeds using NIR hyperspectral data. *Chemometr Intell Lab*, 84, pp.114–118.
- Fernández Pierna, J.A., Baeten, V. & Dardenne, P., 2009. Spectroscopic imaging. In S. Brown, R. Tauler, & R. Walczak, eds. *Comprehensive Chemometrics*. Oxford, pp. 173–196.
- Flavel, R.J. et al., 2012. Non-destructive quantification of cereal roots in soil using high-resolution X-ray tomography. *J Exp Bot*, 63(7), pp.2503–2511.
- Frasier, I. et al., 2016. Direct field method for root biomass quantification in agroecosystems. *MethodsX*, 3, pp.513–519.
- Gate, P., 1995. *Ecophysiologie du blé*, Lavoisier, Paris.
- Gowen, A.A. et al., 2007. Hyperspectral imaging - an emerging process analytical tool for food quality and safety control. *Trends in Food Science and Technology*, 18(12), pp.590–598.
- Gregory, P.J. et al., 2013. Contributions of roots and rootstocks to sustainable, intensified crop production. *Journal of Experimental Botany*, 64(5), pp.1209–1222.
- Gregory, P.J., 2006. Roots, rhizosphere and soil: The route to a better understanding of soil science? *European Journal of Soil Science*, 57(1), pp.2–12.
- Guan, D. et al., 2014. Tillage practices affect biomass and grain yield through regulating root growth, root-bleeding sap and nutrients uptake in summer maize. *Field Crop Res*, 157, pp.89–97.

- Hack, H. et al., 1992. Einheitliche Codierung der hänologischen Entwicklungsstadien mono- und dikotyler Pflanzen – Erweiterte BBCH-Skala, Allgemein. *Nachrichtenbl. Deut. Pflanzenschutzd*, 44, pp.265–270.
- Hagen, N. & Kudenov, M.W., 2013. Review of snapshot spectral imaging technologies. *Optical Engineering*, 52(9), p.090901.
- Han, E. et al., 2015. Root growth dynamics inside and outside of soil biopores as affected by crop sequence determined with the profile wall method. *Biology and Fertility of Soils*, 51, pp.847–856.
- Hauggaard-Nielsen, H. & Jensen, E.S., 2005. Facilitative root interactions in intercrops. *Plant and Soil*, 274, pp.237–250.
- He, H.-J., Wu, D. & Sun, D.-W., 2014. Potential of hyperspectral imaging combined with chemometric analysis for assessing and visualising tenderness distribution in raw farmed salmon fillets. *Journal of Food Engineering*, 126, pp.156–164.
- Hearst, M. et al., 1998. Support vector machines. *IEEE Intelligent Systems*, 13, pp.18–28.
- Heeraman, D. a & Juma, N.G., 1993. A Comparison of Minirhizotron, Core and Monolith Methods for Quantifying Barley (*Hordeum-Vulgare* L) and Fababean (*Vicia-Faba* L) Root Distribution. *Plant and Soil*, 148, pp.29–41.
- Hiel, M.-P. et al., 2018. Impact of crop residue management on crop production and soil chemistry after seven years of crop rotation in temperate climate, loamy soils. *PeerJ*, 6, p.e4836.
- Himmelbauer, M.L., Loiskandl, W. & Kastanek, F., 2004. Estimating length, average diameter and surface area of roots using two different Image analyses systems. *Plant and Soil*, 260, pp.111–120.
- Hirte, J. et al., 2017. Overestimation of Crop Root Biomass in Field Experiments Due to Extraneous Organic Matter. *Front Plant Sci*, 8(284), pp.1–12.
- Hochholdinger, F., 2009. The Maize Root System: Morphology, Anatomy, and Genetics. In J. Bennetzen & S. Hake, eds. *Handbook of Maize: Its Biology*. Springer New York, pp. 145–160.
- Hodge, A. et al., 2009. Plant root growth, architecture and function. *Plant and Soil*, 321, pp.153–187.
- Höflich, G. et al., 2000. Influence of agricultural crops and fertilization on microbial activity and microorganisms in the rhizosphere. *J Agron Crop Sci*, 184, pp.49–54.
- Huang, G. -b. et al., 2012. Effects of different tillage systems on soil properties, root growth, grain yield, and water use efficiency of winter wheat (*Triticum aestivum* L.) in arid northwest China. *J Integr Agric*, 11(8), pp.1286–1296.
- Huang, H., Liu, L. & Ngadi, M., 2014. Recent Developments in Hyperspectral Imaging for Assessment of Food Quality and Safety. *Sensors*, 14(4), pp.7248–7276.

- IGN, 2017. Découvrez la photographie aérienne infrarouge. *Géoportail*. Available at: <https://www.geoportail.gouv.fr/actualites/decouvrez-la-photographie-aerienne-infrarouge>, (05.06.2018).
- Izumi, Y., Uchida, K. & Iijima, M., 2004. Crop production in successive wheat-soybean rotation with no-tillage practice in relation to the root system development. *Plant Prod Sci*, 7(3), pp.329–336.
- Jackson, R.B. et al., 1996. A global analysis of root distributions for terrestrial biomes. *Oecologia*, 108, pp.389–411.
- Jose, S. et al., 2001. Comparison of minirhizotron and soil core methods for quantifying root biomass in a temperate alley cropping system. *Agroforestry Syst*, 52(2), pp.161–168.
- Kamruzzaman, M. et al., 2012. Non-destructive prediction and visualization of chemical composition in lamb meat using NIR hyperspectral imaging and multivariate regression. *Innovative Food Science & Emerging Technologies*, 16, pp.218–226.
- Kamruzzaman, M. et al., 2012. Potential of hyperspectral imaging and pattern recognition for categorization and authentication of red meat. *Innov Food Sci Emerg*, 16, pp.316–325.
- Kätterer, T. et al., 2011. Roots contribute more to refractory soil organic matter than above-ground crop residues, as revealed by a long-term field experiment. *Agriculture, Ecosystems & Environment*, 141, pp.184–192.
- Kätterer, T., Hansson, A.-C. & Andrén, O., 1993. Wheat root biomass and nitrogen dynamics- effects of daily irrigation and fertilization. *Plant Soil*, 151, pp.21–30.
- Kemsley, E.K., 1996. Discriminant analysis of high-dimensional data: a comparison of principal components analysis and partial least squares data reduction methods. *Chemometrics and Intelligent Laboratory Systems*, 33, pp.47–61.
- Kosmachevskaya, O. V. & Topunov, A.F., 2007. Method of determination of the content of hemoglobin-like proteins in heterogenic mixtures. *Appl Biochem Micro*, 43(3), pp.313–319.
- Koteen, L.E. & Baldocchi, D.D., 2013. A randomization method for efficiently and accurately processing fine roots, and separating them from debris, in the laboratory. *Plant Soil*, 363, pp.383–398.
- Kuijken, R.C.P. et al., 2015. Root phenotyping: From component trait in the lab to breeding. *Journal of Experimental Botany*, 66(18), pp.5389–5401.
- Kusumo, B.H. et al., 2011. Measuring carbon dynamics in field soils using soil spectral reflectance: prediction of maize root density, soil organic carbon and nitrogen content. *Plant Soil*, 338, pp.233–245.
- Kusumo, B.H. et al., 2010. Predicting pasture root density from soil spectral reflectance: field measurement. *Eur J Soil Sci*, 61, pp.1–13.

- Kusumo, B.H. et al., 2009. The use of Vis-NIR spectral reflectance for determining root density: evaluation of ryegrass roots in a glasshouse trial. *European Journal of Soil Science*, 60(1), pp.22–32.
- Larousse, Dictionnaire de français., <http://www.larousse.fr/dictionnaires/francais>, (18.02.2018).
- Lawrence, K.C. et al., 2004. Comparison between visible/NIR spectroscopy and hyperspectral imaging for detecting surface contaminants on poultry carcasses. In *Proc. SPIE 5271, Monitoring Food Safety, Agriculture, and Plant Health*. p. 35.
- Lebot, V., 2012. Near infrared spectroscopy for quality evaluation of root crops: practical constraints, preliminary studies and future prospects. *J Root Crops*, 38 (1), pp.3–14.
- Lei, P. & Bauhus, J., 2010. Use of near-infrared reflectance spectroscopy to predict species composition in tree fine-root mixtures. *Plant Soil*, 333(1), pp.93–103.
- Li, L. et al., 2006. Root distribution and interactions between intercrop species. *Oecologia*, 147, pp.280–290.
- Lim, J. et al., 2016. Detection of melamine in milk powders using near-infrared hyperspectral imaging combined with regression coefficient of partial least square regression model. *Talanta*, 151, pp.183–191.
- Liu, D., Sun, D.-W. & Zeng, X.-A., 2013. Recent Advances in Wavelength Selection Techniques for Hyperspectral Image Processing in the Food Industry. *Food Bioprocess Tech*, 7(2), pp.307–323.
- Livesley, S.J. et al., 1999. Sieve size effects on root length and biomass measurements of maize (*Zea mays*) and *Grevillea robusta*. *Plant and Soil*, 207, pp.183–193.
- Lobet, G., Pagès, L. & Draye, X., 2011. A novel image-analysis toolbox enabling quantitative analysis of root system architecture. *Plant physiology*, 157, pp.29–39.
- Lynch, J., 1995. Root Architecture and Plant Productivity. *Plant Physiology*, 109, pp.7–13.
- Maeght, J.-L., Rewald, B. & Pierret, A., 2013. How to study deep roots—and why it matters. *Front Plant Sci*, 4(299), pp.1–14.
- Mahesh, S. et al., 2015. Hyperspectral imaging to classify and monitor quality of agricultural materials. *Journal of Stored Products Research*, 61, pp.17–26.
- Majdi, H., 1996. Root sampling methods - application and limitations of the minirhizotron technique. *Plant Soil*, 185, pp.255–258.
- Manley, M., 2014. Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. *Chemical Society Reviews*, 43, pp.8200–8214.
- Martínez, E. et al., 2008. Soil physical properties and wheat root growth as affected by no-tillage and conventional tillage systems in a Mediterranean environment of Chile. *Soil and Tillage Research*, 99(2), pp.232–244.

- McKell, C.M., Wilson, A.M. & Jones, M.B., 1961. A Flotaison method for easy separation of roots from soil samples. *Agron J*, 53, pp.56–57.
- Meinen, C. & Rauber, R., 2015. Root discrimination of closely related crop and weed species using FT MIR-ATR spectroscopy. *Front Plant Sci*, 6(765), pp.1–13.
- Mendiburu, F. & Simon, R., 2015. Agricolae -Ten years of an Open source Statistical tool for experiments in Breeding, agriculture and biology. *PeerJ Prepr*, 3, pp.1–17.
- Menesatti, P. et al., 2009. Supervised multivariate analysis of hyper-spectral NIR images to evaluate the starch index of apples. *Food and Bioprocess Technology*, 2, pp.308–314.
- Metrohm, 2013. *NIR Spectroscopy : A guide to near-infrared spectroscopic analysis of industrial manufacturing processes*, Herisau, Switzerland, Metrohm AG.
- Mishra, P. et al., 2015. Detection and quantification of peanut traces in wheat flour by near infrared hyperspectral imaging spectroscopy using principal-component analysis. *J Near Infrared Spec*, 23(1), pp.15–22.
- Morell, F.J. et al., 2011. Root growth of barley as affected by tillage systems and nitrogen fertilization in a semiarid mediterranean agroecosystem. *Agron J*, 103(4), pp.1270–1275.
- Moule, C., 1971. *Phytotechnie spéciale - II- Céréales La maison rustique.*, Paris.
- Munier-Jolain, N. et al., 2005. *Agrophysiologie du pois protéagineux* Quae., Paris.
- Munkholm, L.J., Hansen, E.M. & Olesen, J.E., 2008. The effect of tillage intensity on soil structure and winter wheat root/shoot growth. *Soil Use and Management*, 24(4), pp.392–400.
- Muñoz-Romero, V. et al., 2009. Effect of tillage system on the root growth of spring wheat. *Plant and Soil*, 326, pp.97–107.
- Muñoz-Romero, V. et al., 2010. Monitoring wheat root development in a rainfed vertisol: Tillage effect. *Eur J Agron*, 33(3), pp.182–187.
- Naganathan, K.G. et al., 2008. Partial least squares analysis of near-infrared hyperspectral images for beef tenderness prediction. *Sensing and Instrumentation for Food Quality and Safety*, 2(3), pp.178–188.
- Nakaji, T., Noguchi, K. & Oguma, H., 2008. Classification of rhizosphere components using visible–near infrared spectral images. *Plant Soil*, 310, pp.245–261.
- Naumann, A., Heine, G. & Rauber, R., 2010. Efficient discrimination of oat and pea roots by cluster analysis of Fourier transform infrared (FTIR) spectra. *Field Crop Res*, 119, pp.78–84.
- Neill, C., 1992. Comparison of soil coring and ingrowth methods for measuring belowground production. *Ecology*, 73(5), pp.1918–1921.
- Newman, E.I., 1966. A method of estimating the total length of root in a sample. *J. Appl. Ecology*, 3, pp.139–145.

- Nkrumah, B. et al., 2011. Hemoglobin estimation by the HemoCue portable hemoglobin photometer in a resource poor setting. *BMC Clinic Pathol*, 11, pp.1–6.
- Van Noordwijk, M. et al., 2000. Trench profile techniques and core break methods. In A. L. Smit et al., eds. *Root methods, A Handbook*. Springer-Verlag Berlin Heidelberg, pp. 211–233.
- Oliveira, M.R.G. et al., 2000. Auger sampling, ingrowth cores and pinboard methods. In A. L. Smit et al., eds. *Root Methods, A Handbook*. Springer-Verlag Berlin Heidelberg, pp. 175–210.
- Orellana, R.G., Fan, F. & Sloger, C., 1978. Tobacco Ringspot Virus and Rhizobium Interactions in Soybean: Impairment of Leghemoglobin Accumulation and Nitrogen Fixation. *Phytopathology*, 68, pp.577–582.
- Picon-Cochard, C. et al., 2009. Use of near-infrared reflectance spectroscopy to predict the percentage of dead versus living grass roots. *Plant Soil*, 317, pp.309–320.
- Pierret, A., 2008. Multi-spectral imaging of rhizobox systems: New perspectives for the observation and discrimination of rhizosphere components. *Plant and Soil*, 310, pp.263–268.
- Pierret, A. et al., 2003. Observing plant roots in their environment: current imaging options and specific contribution of two-dimensional approaches. *Agronomie*, 23, pp.471–479.
- Plaza-Bonilla, D. et al., 2014. Winter cereal root growth and aboveground–belowground biomass ratios as affected by site and tillage system in dryland Mediterranean conditions. *Plant Soil*, 374, pp.925–939.
- Pourmoayyed, P., 2016. *The effect of sulfur deficiency in pea (Pisum sativum) on factors related to nitrogen fixation and on sulfate transporters*. Rheinischen Friedrich Wilhelms Universität Bonn.
- Price, A.H., 2016. Plant roots: New challenges in a changing world. *J Exp Bot*, 67(4), pp.991–993.
- Pu, Y.Y., Feng, Y.Z. & Sun, D.W., 2015. Recent progress of hyperspectral imaging on quality and safety inspection of fruits and vegetables: A review. *Comprehensive Reviews in Food Science and Food Safety*, 14(2), pp.176–188.
- Qin, R., Stamp, P. & Richner, W., 2004. Impact of tillage on root systems of winter wheat. *Agron J*, 96(6), pp.1523–1530.
- R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Ramirez-Garcia, J. et al., 2015. Intercropping effect on root growth and nitrogen uptake at different nitrogen levels. *Journal of Plant Ecology*, 8(4), pp.380–389.
- Rasmussen, I.S., Dresbøll, D.B. & Thorup-Kristensen, K., 2015. Winter wheat cultivars and nitrogen (N) fertilization-Effects on root growth, N uptake efficiency and N use efficiency. *European Journal of Agronomy*, 68, pp.38–49.

- Ravikanth, L. et al., 2015. Classification of contaminants from wheat using near-infrared hyperspectral imaging. *Biosystems Engineering*, 135, pp.73–86.
- Rewald, B. et al., 2012. Root taxa identification in plant mixtures - current techniques and future challenges. *Plant Soil*, 359, pp.165–182.
- Rewald, B. & Meinen, C., 2013. Plant roots and spectroscopic methods - analyzing species, biomass and vitality. *Front Plant Sci*, 4, p.393.
- Richner, W. et al., 2000. Root Image Analysis and Interpretation. In A. L. Smit et al., eds. *Root Methods, A Handbook*. Springer-Verlag, Berlin Heidelberg, pp. 305–341.
- Robinson, D. et al., 1999. Plant root proliferation in nitrogen-rich patches confers competitive advantage. *Proceedings of the Royal Society B: Biological Sciences*, 266(1418), pp.431–435.
- Rodionova, O.Y. et al., 2005. NIR spectrometry for counterfeit drug detection. *Anal Chim Acta*, 549, pp.151–158.
- Roger, J.M. & Ecartot, M., 2016. *Grain 5 : Prétraitements 1*, CheMoocs.
- Roumet, C. et al., 2006. Quantifying species composition in root mixtures using two methods: near-infrared reflectance spectroscopy and plant wax markers. *New Phytol*, 170(3), pp.631–638.
- Rumbaugh, M.D., Clarck, D.H. & Pendery, B.M., 1988. Determination of root mass ratios in alfalfa-grass mixtures using near infrared reflectance spectroscopy. *J Range Manage*, 41(6), pp.488–490.
- Saeyns, W., Mouazen, A.M. & Ramon, H., 2005. Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy. *Biosyst Eng*, 91(4), pp.393–402.
- Sato, T. et al., 1998. Determination of leghemoglobin components and xylem sap composition by capillary electrophoresis in hypernodulation soybean mutants cultivated in the field. *Soil Sci Plant Nutr*, 44(4), pp.635–645.
- Schaepman, M.E., 2009. Imaging Spectrometers. In T. A. Warner, M. Duane Nellis, & G. Foody, eds. *The SAGE Handbook of Remote Sensing*. Sage, London, pp. 166–178.
- Scherer, H.W. et al., 2008. Low levels of ferredoxin, ATP and leghemoglobin contribute to limited N₂ fixation of peas (*Pisum sativum* L.) and alfalfa (*Medicago sativa* L.) under S deficiency conditions. *Biol Fert Soils*, 44(7), pp.909–916.
- Shahin, M.A., Symons, S.J. & Hatcher, D.W., 2014. Quantification of mildew damage in soft red winter wheat based on spectral characteristics of bulk samples: a comparison of visible-near-infrared imaging and near-infrared spectroscopy. *Food Bioprocess Tech*, 7, pp.224–234.
- Sharma, A. & Sharma, S., 2014. Effect of Nitrogen and Sulphur Nutrition on Nitrogen assimilating Enzymes in Soybean Roots and Nodules. *IJAEB*, 7(3), pp.471–480.

- Smit, A.L. et al., 2000. *Root method: a handbook*. Springer-Verlag, Berlin Heidelberg.
- Snee, R.D., 1977. Validation of regression models: methods and examples. *Technometrics*, 19(4), pp.415–428.
- Srivastava, T. et al., 2014. Methods for hemoglobin estimation: A review of “ what works .” *J Hematol Transfus*, 2(3), pp.2005–2006.
- Steyerberg, E.W. et al., 2001. Internal validation of predictive models. *Journal of Clinical Epidemiology*, 54, pp.774–781.
- Strudley, M.W., Green, T.R. & Ascough, J.C., 2008. Tillage effects on soil hydraulic properties in space and time: State of the science. *Soil and Tillage Research*, 99, pp.4–48.
- Taiz, L. et al., 2015. Observing Roots below the ground. *Plant physiology and development*. <http://6e.plantphys.net/topic05.02.html>, (05.06.2018).
- Taiz, L. & Zeiger, E., 2002. *Plant physiology, third edition*, Sunderland, Massachusetts: Sinauer Associates, Inc., Publishers.
- Tardieu, F. & Manichon, H., 1987. Etat structural, enracinement et alimentation hydrique du maïs. II - Croissance et disposition spatiale du système racinaire. *Agronomie*, 7(3), pp.201–211.
- Tennant, D., 1975. A test of a modified line intersect method of estimating root length. *Journal of Ecology*, 63(3), pp.995–1001.
- Tolt, G., Shimoni, M. & Ahlberg, J., 2011. A shadow detection method for remote sensing images using VHR hyperspectral and LIDAR data. In *2011 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, pp. 4423–4426.
- Tracy, S.R. et al., 2012. Quantifying the effect of soil compaction on three varieties of wheat (*Triticum aestivum* L.) using X-ray Micro Computed Tomography (CT). *Plant and Soil*, 353, pp.195–208.
- Tricot-Pellerin, F., Angevin, F. & Crozat, Y., 1994. Elaboration de la biomasse des nodosités: influence de la nutrition carbonée. In B. Ney et al., eds. *Agrophysiologie du pois protéagineux*. UNIP - ITCF, Paris. pp. 75–91.
- Türker-Kaya, S. & Huck, C., 2017. A review of mid-infrared and near-infrared imaging: principles, concepts and applications in plant tissue analysis. *Molecules*, 22, p.168.
- Vacheron, J. et al., 2013. Plant growth-promoting rhizobacteria and root system functioning. *Frontiers in Plant Science*, 4, pp.1–19.
- Vancutsem, F., Pierreux, J. & Bodson, B., 2012. Dynamique de dégradation des résidus de culture dans le sol en fonction de la profondeur d'enfouissement et de la quantité enfouie. Impact sur le développement des cultures en place. <http://hdl.handle.net/2268/116845>, (23.08.2018).
- Vermeulen, P. et al., 2012. Online detection and quantification of ergot bodies in cereals using near infrared hyperspectral imaging. *Food Addit Contam A*, 29(2), pp.232–240.

- Vermeulen, P. et al., 2013. Validation and transferability study of a method based on near-infrared hyperspectral imaging for the detection and quantification of ergot bodies in cereals. *Anal Bioanal Chem*, 405(24), pp.7765–7772.
- Wang, C. et al., 2014. Effects of different irrigation and nitrogen regimes on root growth and its correlation with above-ground plant parts in high-yielding wheat under field conditions. *Field Crop Res*, 165, pp.138–149.
- Wang, W. & Paliwal, J., 2007. Near-infrared spectroscopy and imaging in food quality and safety. *Sens Instrum Food Qual Saf*, 1(4), pp.193–207.
- Wasson, A. et al., 2016. A portable fluorescence spectroscopy imaging system for automated root phenotyping in soil cores in the field. *Journal of Experimental Botany*, 67(4), pp.1033–1043.
- Wasson, A.P. et al., 2014. Soil coring at multiple field environments can directly quantify variation in deep root traits to select wheat genotypes for breeding. *Journal of Experimental Botany*, 65(21), pp.6231–6249.
- Watt, M. et al., 2013. A rapid, controlled-environment seedling root screen for wheat correlates well with rooting depths at vegetative, but not reproductive, stages at two field sites. *Annals of Botany*, 112, pp.447–455.
- Wilhelm, W.W., Mielke, L.N. & Fenster, C.R., 1982. Root Development of Winter-Wheat As Related To Tillage Practice in Western Nebraska. *Agronomy Journal*, 74, pp.85–88.
- Williams, P. et al., 2009. Maize kernel hardness classification by near infrared (NIR) hyperspectral imaging and multivariate data analysis. *Analytica chimica acta*, 653, pp.121–130.
- Wilson, D.O. & Reisenauer, H.M., 1963. Determination of leghemoglobin in legume nodules. *Anal Biochem*, 6, pp.27–30.
- Wise, B.M. et al., 2006. *PLS_Toolbox Version 4.0 for use with MATLAB™*, http://mitr.p.lodz.pl/raman/jsurmacki/pliki/zajecia/LMDiT/cw3/LMDiT_PLS_Manual_4.pdf, (23.08.2018).
- Wold, S., 1995. Chemometrics; what do we mean with it, and what do we want from it? *Chemometrics and Intelligent Laboratory Systems*, 30(1), pp.109–115.
- Wold, S., Sjöström, M. & Eriksson, L., 2001. PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58, pp.109–130.
- Wu, D. et al., 2012. Rapid prediction of moisture content of dehydrated prawns using online hyperspectral imaging system. *Anal Chim Acta*, 726, pp.57–66.
- Wu, D. & Sun, D.W., 2013. Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review - Part I: Fundamentals. *Innovative Food Science and Emerging Technologies*, 19, pp.1–14.
- Xu, S. et al., 2016. Determination of rice root density at the field level using visible and near-infrared reflectance spectroscopy. *Geoderma*, 267, pp.174–184.

- Xu, S. et al., 2017. Determination of rice root density from Vis–NIR spectroscopy by support vector machine regression and spectral variable selection techniques. *Catena*, 157, pp.12–23.
- Zeng, G., Birchfield, S.T. & Wells, C.E., 2006. Detecting and measuring fine roots in minirhizotron images using matched filtering and local entropy thresholding. *Machine Vision and Applications*, 17(4), pp.265–278.
- Zhang, X., 2010. Support vector machines. In C. Sammut & G. Webb, eds. *Encyclopedia of Machines Learning*. Springer, New York. pp. 941–946.
- Zhong, D. et al., 2009. Maize root complexity analysis using a Support Vector Machine method. *Computers and Electronics in Agriculture*, 69, pp.46–50.
- Zhu, J. et al., 2011. From lab to field, new approaches to phenotyping root system architecture. *Current Opinion in Plant Biology*, 14(3), pp.310–317.

