



Les forêts d'arbres décisionnels et la régression linéaire pour étudier les effets du sous-solage et des drains agricoles sur la hauteur des plants de maïs et les nappes d'eau dans un sol à perméabilité réduite

Mémoire

Anicet Djiemon Deuga

Maîtrise en génie agroalimentaire - avec mémoire
Maître ès sciences (M. Sc.)

Québec, Canada

© Anicet Djiemon Deuga, 2019

Les forêts d'arbres décisionnels et la régression linéaire pour étudier les effets du sous-solage et des drains agricoles sur la hauteur des plants de maïs et les nappes d'eau dans un sol à perméabilité réduite

Mémoire

Anicet Djiemon Deuga

Sous la direction de :

Jacques Gallichand, directeur de recherche
Marc-Olivier Gasser, codirecteur de recherche

Résumé

Les travaux de sous-solage qui améliorent le drainage interne et décompactent des horizons rendus pratiquement imperméables par la compaction profonde seraient bénéfiques aux sols de faible perméabilité. Le sous-solage profond exécuté perpendiculairement aux drains avec un béliet (bulldozer) pourrait être plus efficace pour temporairement améliorer le drainage de ces sols qu'une sous-soleuse conventionnelle attelée à un tracteur et opérée en mode parallèle aux drains. Toutefois, les aménagements réalisés pour améliorer le drainage de surface et interne de ces sols rendent complexe l'évaluation de ces pratiques en dispositif expérimental. L'objectif principal de ce projet était de comparer les forêts d'arbres décisionnelles (FAD) à la régression linéaire multiple (RLM) pour détecter les effets du sous-solage et des systèmes de drainage souterrain et de surface sur la hauteur des plants et la profondeur moyenne de la nappe durant la saison de croissance. Un essai de sous solage a été réalisé à l'automne 2014, dans une argile-limoneuse Kamouraska naturellement mal drainée, remodelée en planches arrondies et souffrant de compaction importante. L'essai comparait un témoin sans sous-solage à quatre traitements de sous-solage, soit une sous-soleuse sur béliet ou sur tracteur, opérées parallèlement ou perpendiculairement aux drains. Chaque traitement a été répété trois fois et disposé aléatoirement en autant de blocs. Au printemps 2016, 198 puits ont été creusés à 60 cm de profondeur pour enregistrer la profondeur de la nappe sous chaque traitement entre juin et juillet 2016. La photogrammétrie a été utilisée pour estimer la hauteur des plants de maïs. Les FAD et la RLM permettent de détecter les principaux facteurs affectant la hauteur des plants de maïs et la profondeur moyenne de la nappe, soit les aménagements antérieurs pour améliorer le drainage interne et le drainage de surface des sols. Les coefficients de détermination obtenus avec les FAD ($R^2 \geq 0,94$) étaient toutefois plus élevés que ceux obtenus avec la RLM ($R^2 \geq 0,28$). Aucun traitement de sous-solage n'a amélioré significativement le drainage interne ni la hauteur des plants de maïs par rapport au témoin sans sous-solage. Les FAD permettent en outre de mieux visualiser les relations non linéaires entre les variables prédites et les autres variables, notamment la position sur la planche et la distance aux drains souterrains, et finalement de déterminer les distances aux drains souterrains optimales (< 2 m) et critiques (> 4 m), la distance optimale à la raie de curage (> 8 m) et la profondeur moyenne critique de la nappe (< 0,25 m). Les FAD permettent ainsi de prédire la hauteur des plants de maïs et la profondeur moyenne de la nappe avec une plus grande précision qu'avec la RLM.

Tables des matières

Résumé	iii
Tables des matières	iv
Liste des figures	vii
Liste des tableaux	viii
Remerciements	x
Avant-Propos	xi
Introduction générale.....	1
Chapitre 1 : Revue de la littérature	4
1.1. Origine de la compaction	4
1.1.1. Où se produit la compaction dans le sol?	4
1.1.2. Effet de la compaction sur les rendements et les propriétés des sols.....	7
1.1.3. La compaction des sols argileux moins perméables	9
1.1.4. Remédiation	9
1.1.5. Le sous-solage.....	10
1.1.6. La texture.....	12
1.1.7. Humidité.....	12
1.1.8. L'importance du drainage	13
1.1.9. Profondeur de la compaction.....	13
1.1.10. Topographie du champ.....	14
1.1.11. Des effets mitigés	16
1.1.12. Choix et types de sous-soleuses	17
1.2. La forêt d'arbres décisionnels	18
1.3. Difficulté des méthodes statistiques traditionnelles à faire ressortir des effets non linéaires multiples ou complexes	19
1.4. Photogrammétrie à haute résolution par drone.....	20

1.5.	Hypothèse et objectif	22
1.6.	Références	23
Chapitre 2: Subsoiling and subsurface drainage: Effects on corn plant height and water table depth in a low permeability clay soil		
		27
	Résumé	29
	Abstract	30
2.1.	Introduction.....	32
2.2.	Material and Methods.....	34
2.2.1.	Site and Climate.....	34
2.2.2.	Experimental Setup	34
2.2.3.	Water table monitoring.....	35
2.2.4.	Image Acquisition by Drone and Production of Digital Elevation Models	36
2.2.5.	Map Editing (GIS).....	36
2.3.	Data Analysis	37
2.3.1	Multiple Linear Regression with R.....	37
2.3.2	RF analysis with R	37
2.4.	Results	39
2.4.1	Complete Experimental Setup: Height of Corn Plants	39
2.4.2	Reduced Experimental Setup: Height of Corn Plants	40
2.4.3	Reduced Experimental Setup: Mean Water Table Depth	42
2.5.	Discussion	43
2.6.	Conclusion	45
2.7.	Acknowledgements	47
2.8.	Références.....	57
2.9.	Annexe.....	60

2.9.1 Résultats des analyses de covariance pour tester l'effet des traitements et des covariables.....	60
Conclusion générale et perspectives.....	67

Liste des figures

Fig. 1 Causes de la compaction à différentes profondeurs (a) et influence de la charge par essieu (b) (Duiker, 2004)	4
Fig. 3 Racine pivotante déformée par le compactage du sol à des profondeurs multiples (Raper et al., 2006).....	5
Fig. 4. Effet de la compaction sur les rendements (Duiker, 2004).	8
Fig. 5. Sens du sous-solage perpendiculaire par rapport aux drains souterrains (Weill, 2015)..	14
Fig. 6. Delimitation of the experimental setup with its open drains, subsurface drainage system and two beds.....	48
Fig. 7. Cross-section of the two beds on the experimental setup.....	49
Fig. 8. Subsoiling treatments allocation and observation wells location on experimental setup.	50
Fig. 9 . Plant_height regression models (RF and MLR) on the complete experimental setup (479 obs.). Performance of the Random Forests model for calibration (a, $R^2= 0.97$), validation (b, $R^2= 0.84$) and complete dataset (c, $R^2= 0.98$). Contribution of the predictor variables to the estimate of Plant_height, expressed by purity of nodes (d). Curves of partial dependence on SSDrain_Depth (e), Dir_Y (f), SSDrain_Distance (g) and Dir_X (h). Performance of the MLR on the complete dataset. (i, $R^2= 0.69$).	51
Fig. 10. Plant_height regression models (RF and MLR) on the reduced experimental setup (198 obs.). Performance of the Random Forests model for calibration (a, $R^2= 0.95$), validation (b, $R^2= 0.67$) and complete dataset (c, $R^2= 0.97$). Contribution of the predictor variables to the estimate of Plant_height expressed by the purity of nodes (d). Curves of partial dependence on Mean_WTD (e), OD_Distance (f), Dir_Y (g), SSDrain_Depth (h), Dir_X (i) , SSDrain_Distance (j), SS_direction (k). Performance of the MLR on the complete data set (i, $R^2= 0.69$).	52
Fig. 11. Mean_WTD regression models (RF and MLR) on the reduced experimental setup (198 obs.). Performance of the Random Forests model for calibration (a, $R^2= 0.86$), validation (b, $R^2= 0.45$) and complete dataset (c, $R^2= 0.95$). Contribution of the predictor variables to estimate Mean_WTD expressed by the purity of nodes (d). Curves of partial dependence on Dir_Y (e), OD_Distance (f), Dir_X (g), SSDrain_Distance (h). Performance of the MLR on the complete dataset (i, $R^2= 0.28$). (i, $R^2= 0.28$).	53

Liste des tableaux

Table 1. Poids moyen approximatif à l'essieu de diverses machineries agricoles (Tsague, 2005) .	6
Table 2. Physical properties of soil.....	54
Table 3. Summary of the three case studies for each analytical method	55
Table 4 Spearman correlations between the predictor variables.....	56

Mots clefs

forêts d'arbres décisionnels, régression linéaire multiple, compaction, sous-solage, profondeur de la nappe, hauteur des plants, planches arrondies, maïs

Remerciements

De nombreuses personnes m'ont aidé à produire ce travail, tout d'abord Mrs Marc-Olivier Gasser et Jacques Gallichand, que je remercie pour avoir accepté de me suivre et pour leurs conseils (éclairages) lors des moments d'égarement de la rédaction de ce travail. Merci Marco pendant les moments de rigolades sur les différents sites d'études lors des réalisations des « mini-puits ». Les œufs que je trouvais tout le temps sur mon bureau c'était toujours un grand plaisir merci. Merci Mr Gallichand tout le temps vous étiez toujours de bonne humeur! Et l'heure de nos rencontres 9h vous étiez toujours au rendez-vous et au taquet grand merci.

Je remercie, Michèle Grenier Statisticienne à l'Irda, Abou Niang pour ses conseils et son aide en géomatique, Chirten Moore, pour le temps qu'elle a consacré à la relecture des différentes parties, ainsi que pour ses remarques et conseils judicieux. Je remercie Marie-Ève Tremblay et Juliette Lapeyre pour les manipulations sur le terrain. Merci Mireille pour ton soutien et tes encouragements.

Ces deux années d'étude complémentaire en agronomie ont été pour moi très enrichissantes, que ce soit au niveau de mes connaissances ou sur le plan humain. J'exprime à cet effet toute ma gratitude aux professeurs de maîtrise de l'Université Laval que j'ai côtoyés pendant ces deux ans.

Enfin mes remerciements vont de tout cœur à ma famille et mes amis pour le soutien constant qu'ils m'ont apporté durant ce travail et ces deux années d'étude supplémentaires, avec une attention particulière à King, Alessandro, Biggie sans oublier toutes les personnes qui d'une manière ou d'une autre, ont contribué au bon déroulement de la rédaction de ce mémoire.

Avant-Propos

Ce mémoire de maîtrise est divisé en trois chapitres. Suite à l'introduction générale, le premier chapitre expose la problématique associée au projet ainsi qu'une revue de littérature sur la compaction des sols, l'effet de la compaction sur les rendements et les propriétés des sols, la compaction des sols argileux moins perméables, comment y remédier, le sous-solage, qu'est-ce qu'un arbre de décision, qu'est-ce que le Random Forest, la difficulté des méthodes statistiques traditionnelles à faire ressortir des effets non linéaires multiples ou complexes, et la photogrammétrie à haute résolution par drone.

Le deuxième chapitre est présenté sous forme d'un article scientifique. L'article s'intitule «Subsoiling and subsurface drainage: « Effects on corn plant height and water table depth in a low permeability clay soil» a été soumis à la revue scientifique Soil and Tillage Research. Jacques Gallichand et Marc-Olivier Gasser, respectivement directeur et professeur au département des sols et de génie agroalimentaire de l'Université Laval, et codirecteur, et chercheur à l'Institut de recherche et de développement en agroenvironnement, sont les coauteurs de cet article. Leur contribution fut considérablement importante au niveau de la planification des travaux sur le terrain, de sages suggestions et conseils avertis durant le processus d'écriture de l'article ainsi que lors de sa révision.

Le troisième chapitre présente la conclusion générale et les perspectives reliant les conclusions du chapitre 2 à la problématique du chapitre 1 suite à ce projet de maîtrise.

Introduction générale

La taille de la machinerie n'a cessé d'augmenter durant les 4 dernières décennies et les pratiques culturales se sont intensifiées. Les tracteurs et les équipements agricoles utilisés aujourd'hui sont plus lourds, plus puissants et opèrent à des vitesses plus élevées qu'il y a seulement quelques années. Ainsi, le producteur peut se permettre d'intervenir plus rapidement au champ, et parfois dans des conditions néfastes pour le sol. Les pressions économiques contraignant aussi le producteur à maximiser l'efficacité de sa machinerie continueront d'accroître la pression sur les sols. En l'absence de vérité à ce sujet, la tendance vers un équipement de plus en plus gros risque de se poursuivre (Mehuys et al, 1986).

Au Québec, dans la plupart des régions ce problème est devenu un enjeu capital pour plusieurs producteurs en grandes cultures (Hamzaa et Anderson, 2004; Gasser, 2014). Environ 100 000 ha de sols cultivés souffrent du problème de compaction (Mehuys et al., 1986; Tabi et al., 1990). En 1985 les pertes globales reliées à la compaction des sols étaient estimées à plus de 100 millions de dollars par année au Québec (Mehuys et al., 1986 ; Garon, 2014). En 2018, les coûts liés à la compaction seraient plus élevés, puisque le poids des équipements agricoles circulant dans les champs n'a cessé d'augmenter. Depuis les années quatre-vingt-dix, la charge exercée par des citernes à lisier de 2 à 3 essieux est passée de 7 à 8 tonnes par essieu (Bédard et al, 1996) à 11 tonnes par essieu en 2010. Même si la pression exercée au sol est davantage contrôlée par l'utilisation de pneus radiaux à basse pression et la navigation par système GPS-RTK (real time Kinematic), la pression exercée sur le sol en profondeur a tout de même augmenté (Gaméda et al, 1987).

Par ailleurs, de nombreux champs en grandes cultures présentent encore des problèmes de drainage. Toutes ces conditions favorisent la présence de nappes perchées, l'eau stagne dans des cuvettes ou ne trouve pas d'exutoire. Les problèmes de compaction ne vont que s'accroître si les sols demeurent saturés en eau. D'autres sols ont subi la compaction profonde réduisant leur perméabilité en profondeur (Hakanson et Reeder, 1995).

Le compactage du sol est un processus de densification physique dû aux facteurs externes (pressions externes) (Duiker, 2004) indépendant du tassement naturel. Ce dernier se produit sans pression externe au cours de la saison suivant le travail du sol, par la réorganisation des

particules élémentaires et des agrégats sous leur propre poids soumis simplement aux aléas climatiques (précipitations, cycles de gel-dégel et d'humectation-dessiccation). Dans les deux cas, les petites particules se logent dans les espaces vides laissées entre les grosses. Ces deux phénomènes sont indissociables au champ, mais le tassement naturel n'atteint pas le degré de compacité du tassement artificiel connu comme étant du compactage (Tabi et al, 1999).

La compaction entraîne la réduction du volume du sol, de l'activité biologique ainsi que la productivité des sols pour les cultures agricoles et forestières. La résistance à la pénétration des racines dans le sol augmente, tandis que la distorsion du sol entraîne une perturbation de la structure du sol et diminution de la porosité et de la perméabilité totale. De nombreux autres changements induits dans la matrice du sol et de ses diverses caractéristiques, entraînent aussi des conséquences environnementales hors de la zone immédiate du sol directement touchée (Akker, 2005). Outre une augmentation de la masse volumique apparente du sol, la compaction des sols agricoles modifie ses propriétés hydrauliques dont la vitesse d'infiltration, la conductivité hydraulique, le drainage et la capacité de rétention en eau des sols (Tabi et al, 1990). La résistance à la pénétration des racines (Bowen, 1981), la réduction de la porosité d'air et la réduction de la disponibilité en eau sont parmi les principaux facteurs limitant la croissance des plantes en sol compact, mais leur impact varie selon les conditions de sol, de sa composition et le type de culture (Gaméda et al., 1987 ; Raper, 2013).

Une fois que le sol est devenu compact, plusieurs moyens peuvent être utilisés pour réduire ou éliminer cette condition. Le compactage du sol peu profond (sous la semelle de labour) causé par des processus naturels ou par de l'équipement de terrain peut habituellement être atténué par un chisel à faible profondeur. Toutefois, si le compactage est plus profond, des mesures plus radicales telles que le sous-solage peuvent être nécessaires (Raper et Kirby, 2006; Weill, 2014). Le sous-solage est une technique largement répandue dans le monde et utilisée depuis des décennies pour réduire la compaction des sols (Muller, 1988; Rapper, 2013). Le sous-solage est un processus par lequel une couche compactée du sol est rompue sans retournement de cette dernière (Singh et al., 2013). Ce décompactage du sol doit favoriser la productivité des cultures pour apporter de la biomasse au sol qui favorisera à son tour la régénération de sa structure, augmentant ainsi le rendement des cultures à plus long terme (Pikul et Aase 1999; Pikul et Kristian 2003, Gasser et al.,2015).

Un bon drainage du sol assure l'élimination de l'excès d'eau dans le profil de sol. Le recours au drainage avec la pose de drains agricoles (tuyaux enterrés dans le sol) permet de palier au drainage déficient. Le drainage améliore la productivité des sols mal drainés par l'abaissement de la nappe phréatique, en fournissant plus d'aération au sol (Luthin et al, 1957). Certains sols souffrent de compaction profonde liée à des processus naturels. La présence de couches indurées par exemple peut restreindre la progression des racines, le mouvement de l'eau, et de l'air dans le sol (Raper, 2013), influençant ainsi son état. Par ailleurs dans les planches drainées, les effets du sous-solage sont complexes à évaluer en raison de la présence de drains, de l'effet des planches et des effets du sous-solage sur le drainage. Aussi, les effets du sous solage sur le compactage peuvent rester limités, d'où l'intérêt de travailler sur de nouvelles méthodes d'analyses pour détecter les différents effets, comme les forêts d'arbres décisionnelles (FAD) qui améliorent l'analyse et font ressortir les principaux effets sous formes non linéaires, le cas échéant.

L'amélioration de la qualité du sol pour maintenir une production élevée et réduire les impacts négatifs sur l'environnement est nécessaire pour élaborer des stratégies de production de cultures viables et durables à long terme (Arshad et al, 1999).

L'objectif de ce travail était de comparer à la régression multiple linéaire (RML), la capacité des FAD à détecter les effets des traitements de sous-solage et les effets non-linéaires des drains souterrains et de surface sur la hauteur des plants et la profondeur moyenne de la nappe durant la saison de croissance. Le travail visait aussi d'un côté pratique à déterminer les meilleures techniques à adopter en matière de sous-solage des sols peu perméables en profondeur sur béliet ou conventionnel avec tracteur opérées en modes parallèle ou perpendiculaire aux drains pour améliorer les conditions de drainage et la productivité des sols par rapport à la parcelle témoin.

Chapitre 1 : Revue de la littérature

1.1. Origine de la compaction

1.1.1. Où se produit la compaction dans le sol?

Trois zones de compaction existent (1) en surface où la compaction est liée à la pression de contact des pneus et dont la contrainte verticale est quasi la même que la pression liée au contact du sol (Fig. 1a), (2) une zone intermédiaire où la compaction se stabilise à un niveau relativement élevé en fonction de la pression de contact au niveau du sol et (3) en profondeur, liée à la charge de la roue et la compaction de la masse par essieu (Fig. 1b) (Gameda et al., 1987, Olsen, 1994, Håkansson et al., 1994, Raper et al., 2006). Avec des charges de plus de 10 tm/essieu, elle peut affecter le sol à plus de 50 cm de profondeur (Prosensols, 2011). La compaction de surface est située à 20 cm (Tsague, 2005, Prosensols, 2011). Elle est moins dommageable que celle en profondeur et est causée par les pneus improprement gonflés, le travail excessif du sol et surtout le trafic de la machinerie (Prosensols, 2011). La semelle de labour occasionne une couche compacte plus profonde sous la couche labourée. Les couches compactées plus en profondeur résultent du passage de machineries trop lourdes au niveau de la charge par essieu, à des moments où le sol est trop humide pour supporter les charges élevées.

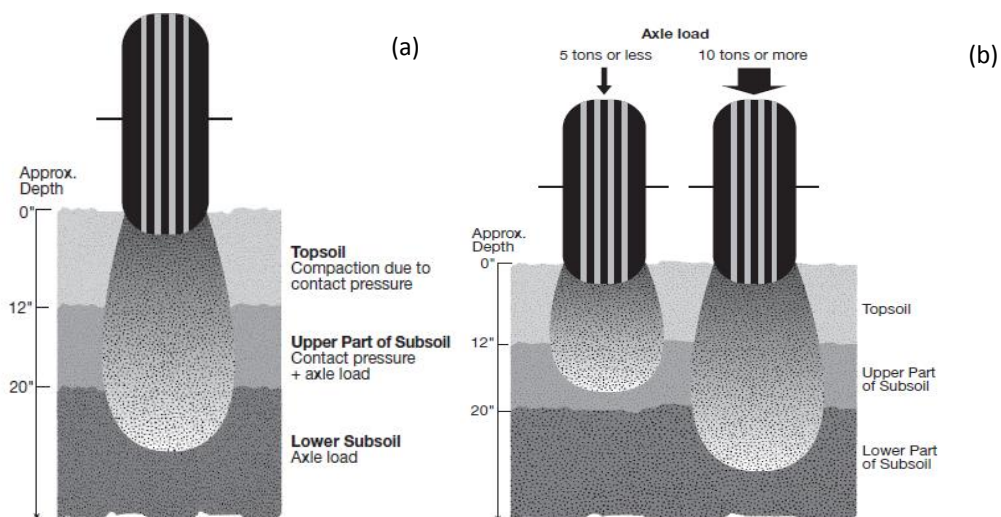


Fig. 1 Causes de la compaction à différentes profondeurs (a) et influence de la charge par essieu (b) (Duiker, 2004)

La compaction entraîne une augmentation importante de la masse volumique apparente du sol et réduit l'espace poral. Ces conditions réduisent la conductivité hydraulique, la vitesse de drainage (mauvais drainage et présence de nappes perchées), les échanges gazeux et la disponibilité de l'air (conditions anaérobies, décomposition des résidus et activité biologique ralentie) et de l'eau, le développement racinaire (Fig. 2) (structure le sol) et les rendements (Fig. 3) (Tsague, 2005, Weill, 2015).

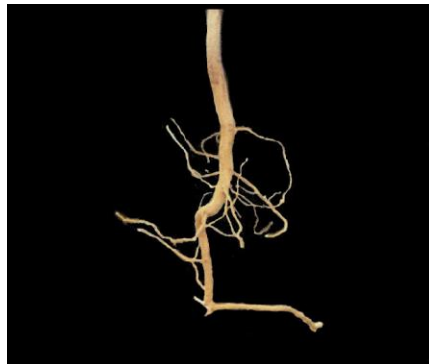


Fig. 2 Racine pivotante déformée par le compactage du sol à des profondeurs multiples (Raper et al., 2006).

L'ensemble du parc de machinerie (tracteurs, moissonneuses, boîtes à grain, etc.) a suivi une tendance à l'augmentation des charges. Certaines opérations culturales sont également exécutées dans des conditions de sol non optimales. De telles charges appliquées au sol entraînent la compaction en profondeur dans bien des sols cultivés. Si la compaction dans la couche de sol travaillée peut être remédiée à plus ou moins court terme par des changements au niveau des pratiques agricoles, en réduisant le travail du sol par exemple, la compaction sous cette couche de surface pourrait persister durant de nombreuses années, causer plus de dommages et donc l'effet sur le rendement pourrait durer une dizaine d'années en l'absence d'un sous-solage voir même ne jamais disparaître (Håkansson et al., 1994, Hamza et al., 2005, Tsague, 2005). La nécessité de semer et de récolter des cultures dans des fenêtres de temps limitées augmente le risque de circuler ou de travailler le sol dans des conditions de teneur en eau inappropriées. Même en réduisant la pression des pneus et en optant pour des pneus radiaux, le poids à l'essieu et la teneur en eau dans le sol restent les principaux facteurs à considérer (Tsague, 2005). La compaction profonde, observée à partir de 30 cm, peut atteindre jusqu'à 60 cm (Tsague, 2005, Prosensols, 2011). Elle augmente en profondeur avec la pression exercée sur le sol en conditions de sol humides (Prosensols, 2011).

Le tableau ci-dessous donne une idée générale du poids moyen à l'essieu de diverses machineries agricoles. À l'heure actuelle, les techniques disponibles en contexte agricole pour limiter la compaction profonde restent inefficaces. Les investigations dans ce domaine continuent leur cours (Batey, 2009).

Table 1. Poids moyen approximatif à l'essieu de diverses machineries agricoles (Tsague, 2005)

MACHINERIES	Poids à l'essieu (tonnes/essieu)
Tracteur, 2 roues motrices	5-7 tonnes
Tracteur, 4 roues motrices	7-13 tonnes
Moissonneuse 6 rangs, vide	10 tonnes
Épandeur de fumier liquide, 5000 gal, 3 essieux	11 tonnes
Moissonneuse 12 rangs, vide	18 tonnes
Wagon à grains rempli, 600 boisseaux, 1 essieu	20 tonnes
Moissonneuse 12 rangs, remplie	24 tonnes
Wagon à grains rempli, 1200 boisseaux, 1 essieu	40 tonnes

Adapté d'une étude de Bill Casady, Uni. Missouri

Quatre-vingts pourcent de la compaction se fait lors d'un premier passage. Toutefois, certains travaux sur l'intensité du trafic ont établi que le nombre de passes joue un rôle important dans le compactage du sol, car les déformations peuvent augmenter avec le nombre de passes (Bakker et al., 1995, Duiker, 2004, Hamza et al., 2005). Des résultats expérimentaux ont aussi montré que tous les paramètres du sol deviennent moins favorables après le passage d'un tracteur (Chygarev et al., 2000) et qu'un certain nombre de passes sur les mêmes lignes de passage d'un tracteur léger peuvent faire autant de dégâts, que le passage d'un tracteur plus lourd avec moins de passes. Selon Jorajuria et al. (2000) le nombre de passes critique est de 10, chiffre au-delà duquel les avantages de l'utilisation d'un tracteur léger sont perdus.

Entre ces deux types de compaction, se trouve généralement une compaction qui se réfère très souvent à la semelle de labour. La semelle de labour est cette couche d'une épaisseur de quelques centimètres de sol, très dense, peu perméable qui se trouve immédiatement sous les passages fréquents de la charrue en dessous de la couche arable (Tsague, 2005, Prosenzols, 2011).

Plusieurs essais menés en Europe et en Amérique du Nord ont confirmé que la compaction des sols peut devenir continue à plus de 40 cm de profondeur dans les sols argileux, même sous l'action du gel/dégel et des fentes de retrait (Håkansson et al., 1994). En fait, l'occurrence et l'intensité de ces facteurs atténuants (gel/dégel, humectation/dessiccation) diminuent rapidement en profondeur dans le sol. Par ailleurs, certains sols issus de tills par exemple, sont naturellement compacts, peu perméables en profondeur et les propriétés physiques du substratum (horizon C) ont peu évolué depuis des millénaires. Ces sols, en général loameux, sont rocheux et massifs sur une grande profondeur. Comme pour les sols légers compacts, les racines et parfois l'eau peuvent être bloquées à la base de la zone de sol travaillée. Dans ces sols, un des buts importants du sous-solage est d'augmenter la profondeur d'enracinement (Weill, 2015). D'autres sols peu perméables en profondeur sont particulièrement sensibles à la compaction, puisqu'ils sont plus souvent saturés en eau au printemps et à l'automne, lors des travaux agricoles. Le seuil critique en deçà duquel la conductivité hydraulique saturée d'un sol en profondeur est limitée et présente des signes de compaction est fixé à 10 cm/jour par l'Agence fédérale allemande de l'environnement (Lebert et al., 2008). Le guide québécois de référence technique en drainage souterrain fixe quant à lui à 3 cm/jour le seuil critique de conductivité hydraulique en deçà duquel l'installation d'un système de drainage souterrain n'est pas justifiée économiquement (CRAAQ, 2005). Selon l'Inventaire des problèmes de dégradation des sols agricoles du Québec, sur les 159 séries de sols analysées, près de 18 % des observations sur celles-ci présentaient entre 40 à 60 cm de profondeur, une conductivité hydraulique saturée (CH) inférieure à 3 cm/jour, tandis que 41 % des observations présentaient une conductivité inférieure à 10 cm/jour (Tabi et al., 1990). D'après ces résultats, plusieurs sols présentaient une perméabilité réduite à plus de 40 cm de profondeur, même ceux sous prairies.

1.1.2. Effet de la compaction sur les rendements et les propriétés des sols

La compaction des sols est un problème important en agriculture qui affecte les propriétés physiques du sol et le drainage de l'eau, et diminue les rendements (Fig. 3). L'étude menée par Duiker (2004) a montré que la compaction de la couche arable de surface (a) et de la partie supérieure du sous-sol (b) a affecté les rendements (Fig. 3) de l'ordre de 15%, mais est restée temporaire et a disparu après 5 à 10 années environ. La compaction profonde quant à elle a affecté les rendements de 3 à 5% et cela s'est maintenu pendant toute la période où les

mesures ont été prises (12 années d'expérience). Le compactage profond du sous-sol (c) est resté quasi permanent et la perte de rendement a duré pendant plusieurs années (Håkansson et al., 1994, Duiker, 2004).

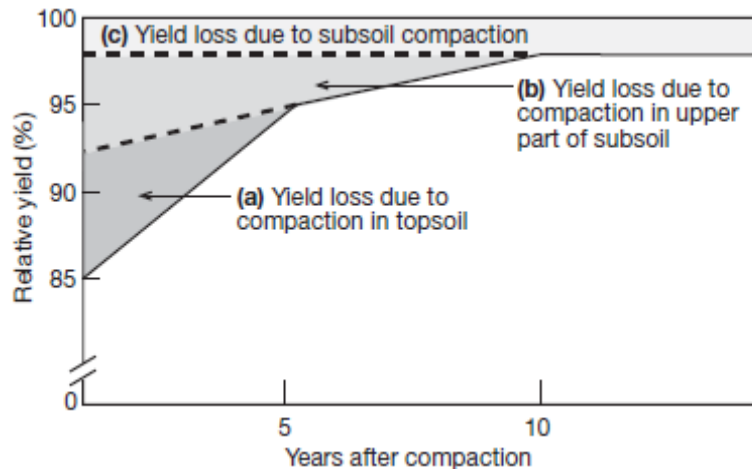


Fig. 3. Effet de la compaction sur les rendements (Duiker, 2004).

Les pertes de rendements associées à la compaction sur l'ensemble des cultures du Québec ont été estimées à 15% (Mehuys, 1986), même si l'effet de la compaction sur les pertes de rendements est difficile à généraliser. Bouthiller-Grenier (2012) a noté une baisse de productivité d'environ 25% du rendement dans les champs entre les drains. Dans un sol compacté, la densité apparente du sol augmente considérablement réduisant ainsi l'espace poral du sol. Cette situation affecte négativement le taux du drainage, la disponibilité de l'air et de l'eau. À titre d'exemple, un sol argileux dont la densité apparente est supérieure à $1,7 \text{ g/cm}^3$ présentera une restriction considérable à l'enracinement de la plupart de plantes. Les études antérieures ont montré une baisse de rendement variant de 10 à 30% due au mauvais développement racinaire. Ce pourcentage s'accroîtra avec le temps si les mêmes pratiques culturales ayant causé la compaction sont maintenues (Tsague, 2005).

L'état du sol en surface (0 - 25 cm) et sa compaction affecte davantage les rendements que la compaction profonde (> 40 cm). Elle persiste moins longtemps, car des facteurs naturels atténuants comme le gel/dégel ou la sécheresse/humectation opèrent plus rapidement à la surface du sol pour changer la donne (Håkansson et al., 1994). La productivité des cultures est généralement mieux corrélée aux propriétés du sol de surface exploité par la majorité des

racines de 0 à 25 cm de profondeur. En revanche, la compaction en profondeur réduit la perméabilité du sol et le maintien dans des conditions humides favorisant la compaction des horizons de surface; un mauvais drainage en profondeur rendant pratiquement nul le mouvement de l'eau dans le sol. Le travail de sol en surface peut temporairement ameublir cette couche sans régler les problèmes en profondeur.

1.1.3. La compaction des sols argileux moins perméables

Les sols argileux de la plaine du Saint-Laurent sont parmi les sols cultivés les plus productifs du Québec lorsque des travaux d'amélioration foncière et de drainage systématique ont été réalisés. L'abaissement de la nappe au printemps se doit d'être suffisamment rapide pour permettre les travaux culturaux et réduire l'asphyxie des plantes au niveau des racines. Toutefois, certains de ces sols sont naturellement moins perméables et seraient particulièrement plus sensibles aux phénomènes de compaction. La présence de nappe perchée entre les drains permet d'identifier des conditions de sol compact (Brunelle et al., 2005), mais cette nappe perchée pourrait résulter d'une perméabilité réduite en profondeur naturelle ou induite par de la compaction profonde.

1.1.4. Remédiation

La compaction est fonction d'un équilibre entre la pression exercée sur le sol et la consistance du sol. Un soc de charrue non affuté ou présentant une façade très arrondie aura tendance à écraser le sol qui se trouve en dessous, la répétition d'un labour à la même profondeur contribuera à consolider cette couche. L'alternance de la profondeur de labour entre 15 cm et 20 cm retardera la formation de la semelle. Un labour plus profond et fréquent n'est toutefois pas conseillé, car en ramenant le sous-sol à la surface, cela contribue à réduire la fertilité du sol. La culture de plantes vivaces à enracinement profond comme la luzerne et le trèfle contribue à fissurer cette semelle de labour, du moins la recoloniser après un sous-solage. Enfin, le passage occasionnel d'un outil combiné de dents rigides (genre Ripper, défonceuse, chisel lourd) atténuera ce problème (Tsague, 2005).

Les travaux de recherche sur la compaction sont effectués aussi en termes de prévention c'est-à-dire en termes de réduction à la source de tout ce qui est charge excessive. L'idéal est de

garder autant que possible le poids à l'essieu à moins de 6 tonnes afin d'alléger la compaction profonde, les sols argileux étant plus sensibles surtout si le champ dont le taux d'humidité est élevé comme au printemps ou après une succession de journées pluvieuses (Vepraskas et al., 1995, Hamza et al., 2005, Tsague, 2005, Raper et al., 2006).

La compaction au niveau des couches supérieures peut être atténuée en réduisant la pression des pneus, en augmentant le nombre d'essieux, et en contrôlant la circulation. En plus de ces solutions mises en place pour atténuer la compaction des sols, d'autres mesures d'évitement comme l'utilisation de chenilles ou de pneus de grand diamètre, pneus basse pression (pneus de flottation, doubles, pneus radiaux), sont préconisées, et finalement l'évitement de la circulation dans des conditions non optimales du sol (Brochu et al., 2000, Duiker, 2004). Certains facteurs naturels tels que les cycles de gel et dégel (Tsague, 2005, Weill, 2015), mouillage/séchage et l'activité biologique des racines et des vers de terre peuvent également jouer un rôle important pour atténuer la compaction en surface.

Toutefois, un bon diagnostic avec des profils de sol et l'analyse de l'égouttement et du drainage des champs demeure essentiel avant d'entreprendre ou de recommander toute intervention d'envergure. Il est généralement recommandé de vérifier et de régler les problèmes d'égouttement de surface et de drainage avant d'entreprendre des travaux de décompaction profonde (Brochu et al., 2000, Brunelle et al., 2005, Weill, 2009).

1.1.5. Le sous-solage

La décompaction mécanique du sol réalisée de 30 à 60 cm de profondeur (sous-solage) a des effets variables sur la compaction profonde et l'augmentation de la productivité (Håkansson et al., 1994). Réalisé dans de bonnes conditions (sols secs, orientation des travaux), le sous-solage sert à initier un long processus de régénération du sol. Le sous-solage à lui seul ne peut régler des problèmes de compaction. Un ensemble de mesure doit être mis en place. La réduction de la circulation de la machinerie et des charges subséquentes est ainsi essentielle, car le sol fragilisé peut se recompacter rapidement et parfois de façon plus importante qu'au départ. Dans les sols à faible perméabilité, l'eau s'accumule dans le sillon sous-solé et les rend ainsi encore plus sensibles à la compaction. Le décompactage ou la fragmentation de la couche compacte doit permettre à l'eau de s'infiltrer plus profondément. Mais si le sol demeure

imperméable en profondeur, l'eau risque de ne pas s'évacuer et alors engendrer d'autres problèmes d'enlèvement et de compaction sous l'effet de la machinerie. L'orientation des travaux de sous-solage en fonction des drains, des pentes du terrain ainsi que de la circulation habituelle permettrait de mieux évacuer cette eau et ainsi réduire les risques de compaction subséquente, mais aussi de favoriser le drainage de nappes perchées (Cochrane et al., 1994).

Le sous-solage est une opération coûteuse qui doit être bien planifiée. Son objectif premier est de briser les horizons denses situés en profondeur et ainsi de faciliter l'aération du sol, la décomposition des débris végétaux, la pénétration des racines (croissances des plants) et l'infiltration de l'eau. Le sous-solage peut être une solution temporaire (Garner et al., 1987, Vepraskas et al., 1995, Raper, 2006, Weill, 2015) ou définitive pour remédier au problème de compaction des sols, mais si les pratiques et les façons culturales ne sont pas modifiées après les travaux, la probabilité que le même problème réapparaisse est élevée.

La recherche indique que deux passes d'un tracteur dans la zone sous-solée provoqueront un retour du sol à son état initial d'avant sous-solage (Blackwell et al., 1989). Par contre, si la circulation est contrôlée, les avantages du sous-solage pourraient être bénéfiques et durables. La gestion globale du système doit être examinée constamment pour déterminer si le compactage du sol est naturel ou induite par la circulation de la machinerie.

Le sous-solage devrait être réalisé au plus une fois tous les deux à trois ans, selon le type et les conditions de sol. Pour avoir une efficacité du système (maintien de la quantité de résidu pour la conservation du sol), cette technique doit être effectuée sans perturber excessivement la surface du sol. Cependant, la plupart des décompacteurs perturbent la surface du sol de façon excessive et ainsi recouvre ses résidus de cultures (Celik et al., 2012).

Toutefois avant de sous-soler, un bon diagnostic avec des profils de sol et l'analyse de l'égouttement et du drainage des champs demeure essentiel (Brochu et al., 2000, Brunelle et al., 2005, Weill, 2015). Afin de connaître les caractéristiques du champ, plusieurs examens de profil de sol doivent être faits, en particulier pour connaître la profondeur de la compaction (Ball et al., 2007, Weill, 2009). Des mesures au pénétromètre peuvent compléter ces examens afin d'évaluer la constance de la présence et de la profondeur de la zone compacte dans l'ensemble du champ; mais ne remplacent pas les examens du profil. Le plan de drainage devrait aussi être examiné ainsi que la topographie du terrain. Les caractéristiques du champ les plus importantes

à considérer sont la texture, le taux du drainage, la profondeur de la compaction et la topographie.

1.1.6. La texture

La texture du sol a un effet important sur sa fragilité à la compaction (Weil, 2015). Sans oublier l'effet de l'humidité sur la texture qui joue un rôle important sur le succès de l'opération de sous solage ou de décompaction. Il convient de préciser que les sols sableux se compactent plus facilement que les sols argileux. Les racines sont davantage bloquées dans les sols sableux, alors que les sols argileux se fissurent et permettent aux racines de pénétrer. En revanche, les sables sont beaucoup plus faciles à décompacter que les argiles. Il est difficile de réussir un bon sous-solage en sol argileux et donc il serait inutile de sous-soler si le drainage n'est pas efficace. Le sous-solage doit être considéré pour diminuer la compaction si le drainage souterrain est adéquat (Weill, 2015).

1.1.7. Humidité

L'humidité n'aide pas dans le processus de décompactations des sols. En revanche, les sols sont plus résistants à la compaction quand ils sont dépourvus d'eau. Le printemps n'est pas le moment idéal pour sous-soler un sol argileux au Québec, car le sol est plus souvent encore saturé d'eau. S'il y a des blocs qui se forment en surface pendant l'opération de sous solage, les cycles de gel et dégel seront nécessaires pour les ameublir et ainsi faciliter les semis au printemps.

Il est important de bien analyser la situation compte tenu de la difficulté de réussir un sous-solage en profondeur dans un sol argileux et humide, et vu le coût important de l'opération de sous-solage. Il ne serait pas nécessaire dans les sols lourds, lorsque ceux-ci sont fissurés (Weill, 2015). Un sol fissuré permet aux racines de mieux se développer en profondeur et favorise une meilleure circulation de l'eau.

L'humidité rend les sols lourds très sensibles à la compaction, mais plus résistants à la compaction lorsqu'ils sont secs. De façon générale, ils résistent mieux à la compaction que les

sols légers, grâce à leur structure en agrégats solides en conditions sèches. Le sous-solage de ces sols doit toujours être fait lorsqu'ils sont secs, soit le plus souvent en été ou au début de l'automne.

1.1.8. L'importance du drainage

Les risques de compaction sont souvent importants quand le drainage n'est pas conforme, et que le travail du sol se fait dans des conditions humides. Cela ne favorise pas les opérations de sous solage et il a comme résultat de compacter de nouveau le sol (Weill, 2015). Étant donné l'importance de l'humidité du sol sur la fragilité à la compaction, en particulier dans les sols argileux, il est essentiel de s'assurer d'un égouttement rapide, c'est à dire de s'assurer que les drainages de surface et souterrain soient fonctionnels et efficaces.

1.1.9. Profondeur de la compaction

1.1.9.1. *Couche compacte peu profonde*

Quand la profondeur de la couche compacte est plus petite que 40 cm, il est facile de décompacter l'ensemble de la couche. Cependant, il est essentiel de décompacter toute l'épaisseur de cette couche imperméable pour éviter d'amplifier le problème. En effet, le sous-solage permet à l'eau de mieux s'infiltrer et de s'écouler à la base de la semelle de labour dans le sol. En absence de sous solage, l'eau s'accumule dans la partie ameublie, car elle est bloquée à la surface de la couche compacte non sous-solée qui est peu voire non perméable (Weill, 2015).

Le sol s'engorge d'eau quand il pleut au printemps, ce qui peut repousser à plus tard les travaux au champ et asphyxier les cultures. De plus, lorsque l'eau stagne, elle augmente l'effet d'anaérobiose et empêche le sol de se restructurer (Weill, 2015).

1.1.9.2. *Couche compacte profonde*

Dans le cas d'une couche compacte profonde, le sous solage est difficile à réaliser, car toute la zone compacte n'est pas éliminée. Toutefois, le sous solage peut être effectué dans le but d'augmenter la profondeur d'enracinement, mais il faut demeurer prudent étant donné que :

1. Le sol est souvent plus mouillé en profondeur, le risque de lissage et de compaction en profondeur dans le sol humide est plus élevé.
2. L'eau peut s'accumuler dans la zone travaillée, lorsque la profondeur du sous-solage est plus faible que la profondeur de la couche compacte, et risque de favoriser la formation de nappes perchées. Le drainage taupe (qui crée des canaux d'écoulement souterrain vers un fossé) dans les sols argileux peut ainsi être plus approprié que le sous-solage pour améliorer la situation. Il est donc important de gérer le sous-solage selon la profondeur de la couche compacte et de la profondeur possible de la nappe perchée.

1.1.10. Topographie du champ

Il est important de choisir le sens du sous-solage selon la topographie du champ et aussi en fonction de la direction des drains souterrains pour que l'évacuation de l'eau soit optimale (Fig. 4). Il faut aussi éviter de diriger l'eau vers des baissières (si elles ne sont pas drainées) afin d'atténuer la compaction induite naturellement. Toutefois, si une partie de la compaction est induite par la machinerie, l'adoption d'une circulation contrôlée ou d'une culture de couverture peut conduire à réduire les interventions de sous-solage (Raper, 2006).

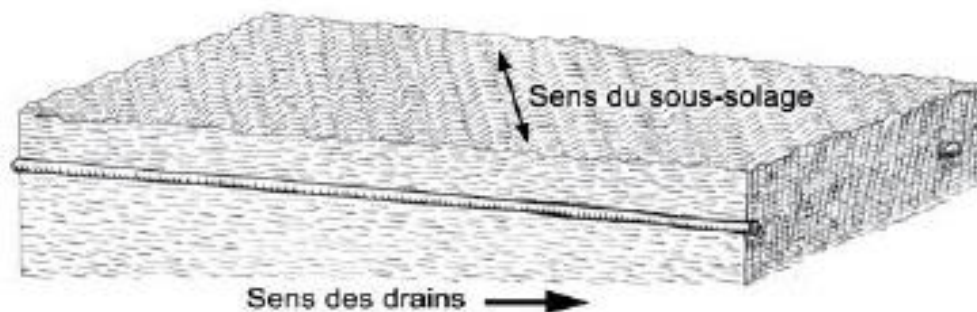


Fig. 4. Sens du sous-solage perpendiculaire par rapport aux drains souterrains (Weill, 2015).

Il est donc important de bien déterminer les causes de la compaction avant de se lancer dans les travaux de toute sorte, pour éviter d'aggraver la situation. Une stratégie inadaptée peut créer plus de dommages que de bienfaits.

Dans certains cas, le but recherché peut être un ameublissement complet du profil de sol alors que dans d'autres, il peut être souhaitable de faire seulement un ameublissement partiel. Il est nécessaire de connaître les caractéristiques du champ afin de bien cibler le but du sous-solage, de décider du moment du sous-solage, de sa profondeur ainsi que de sa direction par rapport à la pente du champ ou aux drains souterrains (Weill, 2015). Le sous-solage doit être adapté afin de ne pas créer une couche gorgée d'eau qui risque de s'assécher très lentement au printemps (Weill, 2015). Suivant Spoor et al. (2003), la présence d'une zone compacte comme une semelle de labour peut protéger le sous-sol de la compaction en offrant une surface ferme qui absorbe et distribue la pression causée par la machinerie. Ces chercheurs recommandent de ne pas trop ameublir la couche compacte et de chercher plutôt à la fissurer ou la casser en blocs assez gros, de façon à laisser passer l'eau et les racines. Ainsi, le sous-solage sur un sol excessivement compacté favorisera un sol ameubli pour la croissance des racines. La profondeur de la croissance des racines est augmentée et les plantes sont mieux à même de résister à des périodes de sécheresse (Raper, 2006).

Le sous-solage contribue souvent à l'augmentation du nombre de macrospores et de l'infiltration (Xu et al., 2001). Même si certains de ces pores disparaîtront après la reconsolidation du sol, beaucoup resteront ouverts et fourniront un stockage accru d'eau et d'oxygène pour les racines des plantes. Cependant, il est important que la circulation des équipements lourds comme les chars à grain et les citernes soit réduite au minimum afin d'éviter des effets néfastes à long terme. Certaines recherches ont indiqué que les avantages du sous-solage étaient perdus après un deuxième passage d'un char lourd à la surface du sol. Cela signifierait que le sous-solage pourrait ne pas être bénéfique pour les cultures, si la circulation venait à s'intensifier après cette opération. Le maintien du profil de sol ameubli et la capacité de stockage accrue de l'eau pourraient être extrêmement précieux pour les racines des plantes pendant des périodes de sécheresse temporaire en été (Raper, 2006).

Dans certains sols où le compactage sévère n'est pas un problème, sous-soler n'est pas toujours synonyme d'augmentation de rendements des cultures. Plusieurs études à ce sujet (Gaultney et al., 1982, Evans et al., 1996) n'ont montré aucune augmentation de rendement, bien que le

compactage du sol a été temporairement réduit. Le sous-solage ne se traduit pas toujours par une augmentation des rendements des cultures lorsque l'irrigation est disponible pour pallier aux conditions de sécheresse (Coates, 1997, Aase et al., 2001, Camp et al., 2002). L'irrigation peut alors compenser le manque d'espace porale dans le sol et l'enracinement profond des cultures.

1.1.11. Des effets mitigés

Plusieurs travaux de recherche sur le sous-solage ont été réalisés à l'échelle mondiale et au Québec. Les effets sur l'augmentation de la productivité des sols sont relativement variables.

L'étude réalisée par Brochu (1986) sur sept sites d'essais de sous-solage en sol argileux dans la région de Saint-Hyacinthe rend compte des baisses de rendements de maïs grain chez certains producteurs, et des gains plus ou moins importants chez d'autres avec un sous-solage réalisé à 45-60 cm de profondeur. Cependant, dans cette étude, les rendements de culture n'ont pas clairement mis en évidence les bienfaits du sous-solage.

Quelques travaux permettent d'anticiper une efficacité supérieure du sous-solage en profondeur dans les sols argileux moins perméables, lorsque les travaux mécaniques sont réalisés perpendiculairement aux drains (Cochrane et al., 1994). En Allemagne, l'amélioration foncière des argiles lourdes d'origine alluviale a été réalisée par la combinaison du drainage systématique et du drainage taupe exécuté perpendiculairement aux drains (Muller, 1988). Le rendement des cultures a augmenté, mettant ainsi en évidence les avantages de cette technique. Muller (1988) recommande de combiner le drainage au sous-solage pour une amélioration des rendements. Cette combinaison devrait être adaptée aux conditions locales du terrain.

Les travaux de recherche de l'Ohio State Extension University sur le sous-solage ont montré que certains types de sol bénéficient d'ameublissement en profondeur. Cependant, tous les sols n'ont généralement pas le même degré de sensibilité au compactage, certains sols même s'ils sont améliorés avec un labour profond, ne produisent pas d'amélioration du rendement (Murdock, 1999). El-Araby et al. (1987) ont révélé que le sous-solage n'est pas recommandé dans certains sols (argiles lourdes), lorsqu'il n'y a pas de couches imperméables dans les 60 premiers centimètres et qu'il n'y a aucune preuve de compactage dans les 70 cm du profil de sol

étudié. Campbell (1974) a cependant montré que lorsque le compactage du sol est un problème, le sous-solage peut être une solution.

Bien que certains travaux de recherche indiquent des chances de succès du sous-solage, peu d'études ont systématiquement évalué cette opération dans des sols argileux de faible perméabilité et l'importance du sens de sous solage par rapport aux drains. Dans ces sols, le sous-solage profond pourrait être bénéfique à la fois pour briser la couche compacte, favoriser l'infiltration de l'eau en profondeur et finalement diriger cette eau en direction des drains. Un abaissement plus rapide des nappes perchées situées sous la couche de sol travaillée aurait un impact majeur sur le potentiel de rendement des cultures. Dans les sols peu perméables, la profondeur du sous-solage aurait aussi son importance pour diminuer la distance entre le drain et le fond du sillon et ainsi favoriser l'infiltration de l'eau.

Les sous-soleuses conventionnelles opèrent avec un tracteur agricole à pneumatiques et travaillent pour la plupart à moins de 50 cm de profondeur en raison de la force exigée et de la traction réduite de ce type de traction. La force de traction nécessaire pour exécuter un travail régulier à des profondeurs de plus de 60 cm serait plus facilement obtenue avec un bélier ou un tracteur à chenilles.

1.1.12. Choix et types de sous-soleuses

L'outil de travail utilisé pour le sous-solage varie considérablement et entraîne des différences dans les exigences de force de traction, et la perturbation des sols (Raper, 2013). De façon générale, le choix de la sous-soleuse doit se faire selon l'objectif poursuivi et les caractéristiques du champ. Les critères suivants doivent être pris en compte :

- Le degré de bouleversement de la surface du sol tolérable après l'opération : dans certains cas, il peut être souhaitable de ne pas bouleverser la surface du sol (semis direct, pâturages, engrais verts ou culture implantée avant le sous-solage);
- Le degré d'ameublissement recherché dans l'ensemble du profil : selon l'objectif poursuivi, l'ensemble du profil doit soit être ameubli, soit seulement créer des passages pour l'infiltration de l'eau;
- La gravité de la compaction et la profondeur de la zone compacte

1.2. La forêt d'arbres décisionnels

La forêt d'arbres décisionnels est une méthode d'analyse de l'intelligence artificielle non paramétrique basée sur l'apprentissage par arbre de décision (CART) utilisant un algorithme aléatoire non linéaire ayant une variable cible prédéfinie. Cette méthode prend en considération les non-linéarités et les interactions à partir des données sans avoir besoin de les modéliser explicitement. Elle fonctionne aussi bien pour les variables continues que catégoriques et est très peu influencée par des valeurs aberrantes et manquantes (Grömping, 2009, Jeong et al., 2016, De'Ath, 2007).

Les algorithmes basés sur les arbres de décisions sont considérés comme de meilleurs outils de prédictions de modèles avec de grandes précisions, très utiles dans l'exploration des données et avec une facilité d'interprétation des résultats. L'algorithme sélectionne un sous-ensemble aléatoire de prédicteurs d'échantillons à partir des données de départ par « bootstrap » (De'Ath, 2007) pour chaque arbre et à chaque nœud (Meinshausen, 2006) où 2/3 des données est utilisé pour la calibration et le reste (1/3) pour la vérification (Jeong et al., 2016, Vincenzi et al., 2011) afin de créer des sous-ensembles pour générer un grand nombre d'arbres au hasard (Francke et al., 2008, Jeong et al., 2016). Grâce à ce processus, seul un petit nombre de variables sélectionnées de façon aléatoire sera disponible pour le partitionnement binaire (Cutler et al., 2007). L'arbre de décision identifie la variable la plus significative et les relations qui existent entre elles afin de fournir le meilleur ensemble homogène de la population. La forêt d'arbres décisionnels se compose d'un grand nombre d'arbres par exemple 1000 (Breiman 2001; Ishwaran 2007) et le fait d'avoir un nombre d'arbres important a été signalé comme étant particulièrement intéressant lors des diagnostics en ce qui concerne l'importance des différentes variables (Breiman 2002).

Les forêts d'arbres décisionnels fonctionnent comme des boîtes noires du fait de la difficulté de compréhension de leurs algorithmes. Mais malgré cela elles s'avèrent très efficaces pour établir des liens entre les données d'entrées et formuler une réponse. En effet il s'agit d'identifier très précisément les relations causales effectives entre les données d'entrées et de sorties. Sur cette base les forêts d'arbres décisionnels peuvent avoir une capacité limitée par rapport aux méthodes statistiques traditionnelles (Huang, Lan et al. 2010), mais elles restent cependant plus performantes que les autres modèles statistiques comme la régression multiple linéaire (MLR).

Par exemple, le carré moyen de l'erreur obtenu avec la forêt d'arbres décisionnels est plus faible comparé aux MLR (Breiman, 2001a, Berk, 2008). Par ailleurs la forêt d'arbres décisionnels peut traiter une taille importante de données avec un grand nombre de variables en identifiant les plus importantes (Jeong et al., 2016). Elle a la capacité de modéliser des interactions complexes entre les variables prédictives (Vincenzi et al., 2011, Fukuda et al., 2013).

La majorité des applications des forêts d'arbres décisionnels servent comme outil de classification. Très peu d'études explorent ses capacités en régression pour prédire la productivité des écosystèmes ou des cultures (Fukuda et al., 2013, Vincenzi et al., 2011). Plusieurs études ont montré un certain nombre d'avantages prometteurs (Jeong et al., 2016, Fukuda et al., 2013, Cutler et al., 2007). Même si la forêt d'arbres décisionnels est un outil assez efficace elle peut par exemple entraîner une perte de précision du modèle lors de la prévision des valeurs extrêmes ou des réponses au-delà des limites des données, lorsqu'utilisée comme outil de régression par rapport aux modèles de régression traditionnelle (Breiman, 2001a, Berk, 2008). À ce jour, les applications de régression par forêt d'arbres décisionnels dans les domaines de l'agronomie et de la science des cultures restent plutôt rares, à quelques exceptions près, (Fukuda et al., 2013, Jeong et al., 2016) et il existe très peu voir pas d'études en agriculture qui n'ont été réalisées pour estimer la hauteur des plants de maïs en lien avec le rendement et la profondeur des nappes d'eau souterraine.

1.3. Difficulté des méthodes statistiques traditionnelles à faire ressortir des effets non linéaires multiples ou complexes

L'application de l'intelligence artificielle et les méthodes d'apprentissage automatique (machine learning) en informatique sont de plus en plus reconnues dans le domaine de l'ingénierie agricole (Huang et al., 2010). Par ailleurs, la nature complexe des données et leur volume ne cessent de croître dans les bases de données, dépassant souvent les capacités de traitement humain et des méthodes statistiques traditionnelles de traitement de données d'où l'intérêt de mettre en place des nouvelles méthodes de traitement comme le Machine Learning, le Data Mining, le Deep Learning, le Réseau de neurones, appliquées aux bases de données (Bendeche et al., 2016).

Un modèle statistique traditionnel utilise un certain nombre d'hypothèses basé sur l'estimation des coefficients. La régression linéaire multiple fait référence à des variables prédictives permettant de prédire une variable de réponse en recherchant une fonction mathématique de type Réponse = constante + paramètre * prédictiveur + ... + paramètre * prédictiveur + r avec r comme résidu (Chatterjee et al., 1991, Grömping, 2009). L'analyse de données traditionnelle se base principalement sur des données multidimensionnelles utilisant les statistiques élémentaires (moyenne, variance), les factorielles, ou l'analyse bivariée (corrélation, régressions) sans tenir compte des effets spatiaux (Bendeckache et al., 2016). En plus, l'analyse des données est très limitée (Yeh, 1999). Elle est connue pour être raisonnablement robuste, si le nombre d'observations est nettement plus grand que le nombre de variables (Grömping, 2009). Un avantage de la régression linéaire multiple est qu'elle est facile à mettre en œuvre, mais il peut arriver que la régression multivariée nécessite des logiciels plus avancés (Francke et al., 2008). Par exemple avec plus de variables que d'observations la régression linéaire multiple devient limitée, mais les méthodes comme la régression de la crête (Hoerl et al., 1970), le Lasso (Tibshirani, 1996) ou le réseau élastique qui est la combinaison des deux (Zou et al., 2005) peuvent remédier à cette situation.

Les méthodes de modélisation statistique basées sur des algorithmes d'apprentissage peuvent servir d'alternatives aux approches de régression traditionnelle et ainsi aider à surmonter certaines de leurs limites. (Jeong et al., 2016). La prédiction des caractéristiques de compactage du sol est nécessaire pour les applications d'ingénierie agricole. Les modèles flous ont l'avantage d'être simples et robustes (Betrie et al., 2013) peut être utilisé comme référence, et peuvent aussi être utilisés comme des outils prometteurs pour prédire la qualité du drainage.

1.4. Photogrammétrie à haute résolution par drone

Le monitoring des cultures est important pour suivre le processus de croissance et maximiser le rendement (Zhao et al., 2016). La détermination avec précision de la hauteur des cultures notamment du maïs est un bon indicateur de la biomasse (Yin et al., 2011, Zhang et al., 2012) et du rendement produit. Cela signifie que pour avoir une bonne corrélation entre le rendement et la hauteur des plants, une grande précision est nécessaire au niveau des informations de la hauteur des plants. Cela vaut non seulement pour le modèle numérique de terrain, mais aussi

pour le modèle d'élévation des plantes (Grenzdörffer, 2014). Les systèmes efficaces, automatiques et bon marché pour la détermination de la hauteur des cultures sont plutôt rares. Cependant, les nuages de points générés par des images de drones hautes résolutions offrent une alternative (Grenzdörffer, 2014). Une large gamme de capteurs à distance est disponible pour une agriculture de précision (par exemple l'imagerie hyper spectrale aéroportée et aérienne, le radar, le Lidar, la photogrammétrie aérienne, etc. (Omasa et al., 2006, Rosell and Sanz, 2011). La photogrammétrie aérienne permet de déterminer les modèles de surface. En utilisant les méthodes dites de « correspondance dense », un triplé de coordonnées peut être calculé pour pratiquement tous les pixels d'une photo aérienne (Haala, 2013). La précision au niveau de la hauteur peut être liée à plusieurs problèmes (la météo, la poussière, le vent, la canopée, etc.) et la limite théorique est généralement autour d'un pixel. La précision réalisable habituellement avec les appareils photo numériques disponibles est comprise entre 0.5 et 2 pixels. L'utilisation d'un angle de l'ordre de 6 à 50 mm est tout à fait réalisable à une hauteur de vol de 50 à 100 m. Toutefois des points de contrôle au sol sont mesurés avec un RTK avec une précision généralement comprise entre 10 et 25 mm (Iqbal et al., 2017, Grenzdörffer, 2014). Ainsi la précision des images réalisées par des photos aériennes est de l'ordre de 15 à 30 mm et pour que cela soit respecté il faudrait que la végétation reste fixe ce qui est malheureusement impossible. Il est possible de séparer le sol de la surface à condition que la végétation ne soit pas trop dense et que l'indice de surface foliaire soit faible.

Il existe deux approches pour la détermination de la hauteur des plants par photogrammétrie :

1. Approche par différence entre un modèle numérique de surface c'est-à-dire la hauteur des plants et un modèle numérique de terrain. Précisons que c'est cette méthode qui a été utilisée pour le cas de cette étude. L'avantage de cette méthode est que la surface de référence reste la même et la mesure en hauteur de la surface de la végétation est précise et fiable tout au long de la saison croissance. Cependant, cela requiert un modèle numérique de terrain de haute précision (Grenzdörffer, 2014, Iqbal et al., 2017)).
2. L'approche par nuage de point en trois dimensions, consiste en la détermination de la hauteur des plants par analyse statistique. La principale difficulté de cette approche est de déterminer la hauteur des plants sans l'existence d'un modèle numérique de terrain. Mais vu que le nuage de point acquis par le drone contient la surface de la végétation,

mais aussi la surface du sol entre les plantes, un modèle numérique de terrain et un modèle numérique de surface est dérivée du même nuage de point.

1.5. Hypothèse et objectif

Hypothèse : Le sous-solage des sols de faible perméabilité améliore leur drainage interne lorsqu'il est exécuté perpendiculairement à l'orientation des drains souterrains et en profondeur. Toutefois, les aménagements réalisés pour améliorer le drainage de surface et interne de ces sols rendent complexe l'évaluation de ces pratiques en dispositif expérimental.

L'objectif principal de ce projet était de comparer l'efficacité des forêts d'arbres décisionnelles (FAD) à la régression linéaire multiple (RLM) pour détecter les effets du sous-solage et des systèmes de drainage souterrain et de surface sur la hauteur des plants de maïs et la profondeur de la nappe durant la saison de croissance.

1.6. Références

- AASE, J., BJORNEBERG, D. & SOJKA, R. 2001. Zone–subsoiling relationships to bulk density and cone index on a furrow–irrigated soil. *Transactions of the ASAE*, 44, 577.
- BAKKER, D., & DAVIS, R. 1995. Soil deformation observations in a Vertisol under field traffic. *Soil Research*, 33, 817-832.
- BALL, B., BATEY, T. & MUNKHOLM, L. J. 2007. Field assessment of soil structural quality—a development of the Peerlkamp test. *Soil use and Management*, 23, 329-337.
- BATEY, T. 2009. Soil compaction and soil management—a review. *Soil use and management*, 25, 335-345.
- BENDECHACHE, M., KECHADI, M.-T. & LE-KHAC, N.-A. Efficient Large Scale Clustering based on data partitioning. Data Science and Advanced Analytics (DSAA), 2016 IEEE International Conference on, 2016. IEEE, 612-621.
- BERK, R. A. 2008. *Statistical learning from a regression perspective*, Springer.
- BETRIE, G. D., TEFAMARIAM, S., MORIN, K. A. & SADIQ, R. 2013. Predicting copper concentrations in acid mine drainage: a comparative analysis of five machine learning techniques. *Environmental monitoring and assessment*, 185, 4171-4182.
- BLACKWELL, P., JAYAWARDANE, N., BLACKWELL, J., WHITE, R. & HORN, R. 1989. Evaluation of soil recompaction by transverse wheeling of tillage slots. *Soil Science Society of America Journal*, 53, 11-15.
- BOUTHILLER-GRENIER 2012. Évaluation de l'efficacité des réseaux de drainage souterrains. Connaître le réseau actuel avant de le doubler. L'élaboration d'une méthode de diagnostique. 6p.
- BREIMAN, L. 2001. Random forests. *Machine learning*, 45, 5-32.
- BROCHU, & MASSICOTTE 2000. Module 7. Diagnostic et correction des problèmes de compaction et de drainage. Feuillet 7A. Dans Guide des pratiques de conservations en grandes cultures. Centre de référence en agriculture et agroalimentaire du Québec. Québec., 20p.
- BROCHU, Y. 1986. Résultats préliminaires de décompaction par sous-solage en sols drainés souterrainement. pp. 69-108. 13e Colloque de génie rural. Université Laval. R. Thériault (éd.). 220 p.
- BRUNELLE, & SAVOIE 2005. Diagnostic et correction de problèmes de drainage. Coordination des clubs-conseils en Agroenvironnement. Longueuil, Qc. 4 p.
- CAMP, C. & SADLER, E. 2002. Irrigation, deep tillage, and nitrogen management for a corn–soybean rotation. *Transactions of the ASAE*, 45, 601.
- CAMPBELL, R. 1974. Physical properties and tillage of Paleudults in the southeastern Coastal Plains. *J. Soil Water Conserv*, 29, 220-224.
- CELIK, A., & RAPER, R. L. 2012. Design and evaluation of ground-driven rotary subsoilers. *Soil and Tillage Research*, 124, 203-210.
- CHATTERJEE, S., & PRICE, B. 1991. *Regression analysis by example*, New York, Wiley.
- CHYGAREV, Y., & LODYATA, S. 2000. Research of tyre rigidity in terms of ecological safety of agricultural landscapes. VII Międzynarodowe Sympozjum Ekologiczne Aspekty Mechanizacji Nawożenia, Ochrony Roślin Uprawy Gleby i Zbioru Roślin Uprawnych, Warszawa, Poland, 18-19 wrzesien 2000, 171-176.
- COATES, W. 1997. Minimum tillage systems for irrigated cotton: Is subsoiling necessary? *Applied Engineering in Agriculture*, 13, 175-179.

- COCHRANE, L., & BRENTON, D., HOLMSTROM, J., VAN ROESTEL, L & CARTER, C., THOMPSON 1994. THE EFFECT OF SUBSOILING AND DRAINAGE TREATMENTS ON SOIL AND CROP CHARACTERISTICS. Final report. CLBRR Contribution No. 94-49. 25p.
- CRAAQ 2005. Guide de référence technique en drainage souterrain et travaux accessoires. 67p.
- CUTLER, D. R., EDWARDS, T. C., BEARD, K. H., CUTLER, A., HESS, K. T., GIBSON, J. & LAWLER, J. J. 2007. Random forests for classification in ecology. *Ecology*, 88, 2783-2792.
- DE'ATH, G. 2007. Boosted trees for ecological modeling and prediction. *Ecology*, 88, 243-251.
- DUIKER, S. W. 2004. Effects of Soil Compaction. Penn State College of Agricultural Sciences research. 12p.
- EL-ARABY, A., EL-HADDAD, Z. & EL-ANSARY, M. 1987. Subsoiling in some heavy clay soils of Egypt. *Soil and Tillage Research*, 9, 207-216.
- EVANS, S., LINDSTROM, M., VOORHEES, W., MONCRIEF, J. & NELSON, G. 1996. Effect of subsoiling and subsequent tillage on soil bulk density, soil moisture, and corn yield. *Soil and Tillage Research*, 38, 35-46.
- FRANCKE, T., LÓPEZ-TARAZÓN, J. A. & SCHRÖDER, B. 2008. Estimation of suspended sediment concentration and yield using linear models, random forests and quantile regression forests. *Hydrological Processes*, 22, 4892-4904.
- FUKUDA, S., SPREER, W., YASUNAGA, E., YUGE, K., SARDSUD, V. & MÜLLER, J. 2013. Random Forests modelling for the estimation of mango (*Mangifera indica* L. cv. Chok Anan) fruit yields under different irrigation regimes. *Agricultural water management*, 116, 142-150.
- GAMEDA, S., RAGHAVAN, G., MCKYES, E. & THERIAULT, R. 1987. Subsoil compaction in a clay soil. I. Cumulative effects. *Soil and tillage research*, 10, 113-122.
- GARNER, T., REYNOLDS, W., MUSEN, H., MILES, G., DAVIS, J., WOLF, D. & PEIPER, U. 1987. Energy requirement for subsoiling coastal plain soils. *Transactions of the ASAE*, 30, 343-0349.
- GAULTNEY, L., KRUTZ, G., STEINHARDT, G. C. & LILJEDAHL, J. 1982. Effects of subsoil compaction on corn yields. *Transactions of the ASAE*, 25, 563-0569.
- GRENZDÖRFFER, G. 2014. Crop height determination with UAS point clouds. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40, 135.
- GRÖMPING, U. 2009. Variable importance assessment in regression: linear regression versus random forest. *The American Statistician*, 63, 308-319.
- HAALA, N. 2013. The landscape of dense image matching algorithms.
- HÅKANSSON, I., & REEDER, R. C. 1994. Subsoil compaction by vehicles with high axle load—extent, persistence and crop response. *Soil and Tillage Research*, 29, 277-304.
- HAMZA, M., & ANDERSON, W. 2005. Soil compaction in cropping systems: a review of the nature, causes and possible solutions. *Soil and tillage research*, 82, 121-145.
- HAMZA, M. A. & ANDERSON, W. K. 2005. Soil compaction in cropping systems: A review of the nature, causes and possible solutions. *Soil and Tillage Research*, 82, 121-145.
- HOERL, A. E., & KENNARD, R. W. 1970. Ridge regression: applications to nonorthogonal problems. *Technometrics*, 12, 69-82.
- HUANG, Y., LAN, Y., THOMSON, S. J., FANG, A., HOFFMANN, W. C. & LACEY, R. E. 2010. Development of soft computing and applications in agricultural and biological engineering. *Computers and Electronics in Agriculture*, 71, 107-127.
- IQBAL, F., LUCIEER, A., BARRY, K. & WELLS, R. 2017. Poppy Crop Height and Capsule Volume Estimation from a Single UAS Flight. *REMOTE SENSING*, 9.
- JEONG, J. H., RESOP, J. P., MUELLER, N. D., FLEISHER, D. H., YUN, K., BUTLER, E. E., TIMLIN, D. J., SHIM, K.-M., GERBER, J. S. & REDDY, V. R. 2016. Random forests for global and regional crop yield predictions. *PLoS one*, 11, e0156571.

- JORAJURIA, D., & DRAGHI, L. 2000. Overcompaction of agricultural soil part I: differential influence of axle load and number of passes. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 4, 445-452.
- LEBERT, M., & BÖKEN, H. 2008. Soil compaction. *The Encyclopedia of Earth*.
- MEHUYS, G. 1986. Mesures de la compaction: aspects physique et économique. *13e Colloque de génie rural*, pp. 19-34.
- MEINSHAUSEN, N. 2006. Quantile regression forests. *Journal of Machine Learning Research*, 7, 983-999.
- MULLER, L. 1988. Efficiency of Subsoiling and Subsurface Drainage in Heavy Alluvial Soils of the G. D. R. *Soil and Tillage Research SOTRD* 5, 12.
- MURDOCK, L. W. 1999. Subsoiling of no-tilled corn.
- OLSEN, H. 1994. Calculation of subsoil stresses. *Soil and tillage research*, 29, 111-123.
- OMASA, K., HOSOI, F. & KONISHI, A. 2006. 3D lidar imaging for detecting and understanding plant responses and canopy structure. *Journal of Experimental Botany*, 58, 881-898.
- PROSENSOLS 2011. Le tassement des sols agricoles Prévenir et remédier. 32p.
- RAPER L 2013. SUBSOILING. Published by Elsevier Ltd. 7p.
- RAPER, R. L. IN-ROW SUBSOILING SOUTHEASTERN SOILS TO REDUCE COMPACTION AND IMPROVE CROP YIELDS. Southern Conservation Systems Conference, Amarillo TX, 2006.
- RAPER, R. L., & MAC KIRBY, J. 2006. *Soil compaction: how to do it, undo it, or avoid doing it*, American Society of Agricultural and Biological Engineers.
- ROSELL, J. R. & SANZ, R. 2011. A review of methods and applications of the geometric characterization of tree crops in agricultural activities. *Computers and Electronics in Agriculture*.
- SPOOR, G., TIJINK, F. & WEISSKOPF, P. 2003. Subsoil compaction: risk, avoidance, identification and alleviation. *Soil and tillage research*, 73, 175-182.
- TABI, M., & TARDIF, L., CARRIER, D., LAFLAMME, G. & ROMPRÉ, M. 1990. Inventaire des problèmes de dégradation des sols agricoles du Québec–Région agricole 6, Richelieu, Saint-Hyacinthe [Quebec inventory of agricultural soil degradation problems–Agricultural Region 6, Richelieu, Saint-Hyacinthe]. Government of Quebec Publication Quebec City, QC.
- TIBSHIRANI, R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267-288.
- TSAGUE, G. E. 2005. La compaction des sols : Les causes et les solutions. 5p.
- VEPRASKAS, M., BUSSCHER, W. & EDWARDS, J. 1995. Residual effects of deep tillage vs. no-till on corn root growth and grain yield. *Journal of production agriculture*, 8, 401-405.
- VINCENZI, S., ZUCCHETTA, M., FRANZOI, P., PELLIZZATO, M., PRANOVI, F., DE LEO, G. A. & TORRICELLI, P. 2011. Application of a Random Forest algorithm to predict spatial distribution of the potential yield of *Ruditapes philippinarum* in the Venice lagoon, Italy. *Ecological Modelling*, 222, 1471-1478.
- WEILL 2015. A guide to successful subsoiling. CETAB+,Victoriaville. Centre de référence en agriculture et agroalimentaire du Québec. 43p.
- WEILL, A. 2009. Guide sur les profils de sol agronomiques: un outil de diagnostic de l'état des sols. CRAAQ.
- XU, D., & MERMOUD, A. 2001. Topsoil properties as affected by tillage practices in North China. *Soil and tillage research*, 60, 11-19.
- YEH, K. Z. L. 1999. Le Data Mining Spatial et les bases de données spatiales. *Revue internationale de géomatique. Volume*, 9.

- YIN, X., MCCLURE, M. A., JAJA, N., TYLER, D. D. & HAYES, R. M. 2011. In-season prediction of corn yield using plant height under major production systems. *Agronomy Journal*, 103, 923-929.
- ZHANG, L., & GRIFT, T. E. 2012. A LIDAR-based crop height measurement system for *Miscanthus giganteus*. *Computers and electronics in Agriculture*, 85, 70-76.
- ZHAO, G., DUAN, Z., MING, L., LI, Y., CHEN, R., HU, J., SVANBERG, S. & HAN, Y. 2016. Reflectance and fluorescence characterization of maize species using field laboratory measurements and lidar remote sensing. *Applied optics*, 55, 5273-5279.
- ZOU, H., & HASTIE, T. 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67, 301-320.

Chapitre 2: Subsoiling and subsurface drainage: Effects on corn plant height and water table depth in a low permeability clay soil

Key words: Random forest, multiple linear regression, subsoiling, water table depth, plant height, subsurface drainage, tile drainage, soil bedding.

Résumé

La compaction des sols représente un enjeu important. Discriminer les paramètres qui affectent la productivité des cultures est essentiel pour développer une agriculture efficace. Nous avons évalué deux méthodes statistiques, les forêts d'arbres décisionnels (FAD) et la régression linéaire multiple (RLM), pour évaluer l'effet de différents paramètres dont le sous-solage sur la hauteur des plants de maïs (Plant_hauteur) et la profondeur moyenne de la nappe (Nappe_moy). Des essais de sous-solage ont été réalisés dans une argile à perméabilité réduite, munie de drains souterrains et remodelée en planches arrondies. Ces essais ont permis de comparer à un témoin sans sous-solage, quatre traitements de sous-solage, soit une sous-soleuse sur béliet ou sur tracteur opérée parallèlement ou perpendiculairement aux drains. Chaque traitement a été répété trois fois et repartit de façon aléatoire. Au printemps 2016, 198 puits de 60 cm de profondeur ont été creusés pour observer la profondeur de la nappe sous chaque traitement. La Plant_hauteur de maïs a été estimée par photogrammétrie. Les résultats ont montré que les FAD et la RLM permettent tous les deux de détecter les principaux facteurs affectant la Plant_hauteur et la Nappe_moy. Les FAD ont toutefois permis de visualiser les relations non linéaires entre la Nappe_moy et les autres variables, notamment la position sur la planche, et de déterminer les distances aux drains critique (<2m) et optimale (< 4m), la profondeur aux drains souterrains optimale (1.4 à 1.7m), la distance optimale à la raie de curage (< 8m) et la profondeur moyenne critique de la nappe (< 0.25m). Les coefficients de détermination obtenus avec les FAD ($R^2 \geq 0.94$) étaient plus élevés que ceux obtenus par RLM ($R^2 \geq 0.28$). Les FAD permettent ainsi de prédire la Plant_hauteur et la Nappe_moy avec une plus grande précision que la RLM.

Abstract

Soil compaction is an agricultural and environmental concern. Understanding the parameters affecting crop productivity is essential for an efficient agriculture. We evaluated two statistical methods, random forest (RF) and multiple linear regression (MLR), to evaluate the effect of various parameters, including subsoiling, on corn plant height (Plant_height) and on seasonal mean water table depth (Mean_WTD). We performed subsoiling treatments in a clay field of low permeability, provided with subsurface drains and remodeled in rounded beds. These experiments allowed comparing a control without subsoiling to four subsoiling treatments: a tractor or bulldozer operated parallel or perpendicular to the subsurface drains. Each treatment was randomized and repeated three times. In the spring of 2016, we drilled 198 60-cm deep water table observation wells. The corn height was estimated by photogrammetry. The results showed that both RF and MLR allowed determining the main factors affecting plant height and seasonal mean water table depth. In addition, RF allowed visualizing nonlinear relationships between mean water table depth, plant height and the independent variables, such as the position on the beds. RF also allowed determining critical (< 2 m) and optimal (<4 m) drain spacing, optimal depth to subsurface drains (1.4 to 1.7 m), optimal horizontal distance to the field open drain (< 8 m) and critical mean water table depth (< 0.25 m). Coefficients of determination were much higher for RF ($R^2 \geq 0.94$) than for MLR ($R^2 \geq 0.28$).

List of abbreviations

SSDrain_Depth: Depth to subsurface drain, m

SSDrain_Distance: Horizontal distance to subsurface drain, m

Dir_Y: Perpendicular direction to the field and to the drainage systems, m

Dir_X: Parallel direction to the field and to the drainage systems, m

OD_Distance: Distance to the field open drain, m

Plant_height: Plant height, m

Mean_WTD: Seasonal mean water table depth, cm

ρ_A : Soil bulk density, g/cm³

φ_t : Soil total porosity, cm³/cm³

SS_direction: Subsoiling direction

RF: Random forest

MLR: Multiple linear regression

DEM: Digital elevation model

SOM: Soil organic matter

2.1.Introduction

The difficulties in using traditional statistical methods to bring out nonlinear effects in a complex experimental field setup makes artificial intelligence and machine learning methods more and more recognized for agricultural research purposes (Huang et al, 2010). New methods of data analysis can identify complex nonlinear relationships between agricultural treatments and factors that regulate soil drainage, soil aeration and, ultimately, crop productivity. Non-parametric models, such as decision trees and random forests (RF), appear to be more promising techniques than classical regressions.

The RF method is simple and robust (Carman, 2008, Taghavifar, 2013, Pentoś, 2017, Hossain et al., 2011) and can be used for classification or regression to predict and generalize complex nonlinear relationships (Fukuda et al, 2013, Vincenzi, 2011, Jeong, 2016). When used for regression, RF loses precision when forecasted values should be beyond the limits of the calibration data (Breiman, 2001, Berk, 2008). The applications of RF in agricultural research are scarce (Fukuda et al, 2013, Jeong, 2016) and no studies appear to have explored relationships between subsoiling practices, drainage systems, plant height and water table depth. Classical analyses of data using multiple linear regression (MLR) are mainly based on exploring multidimensional data using basic statistics (average, variance, regression) (Breiman, 2001 b , Bendeche et al., 2016).

Agricultural soil compaction is a global problem that not only affects the productivity of soils and crops, but also increases greenhouse gases emissions and deteriorates water quality (Håkansson, 1994, Gregorich et al., 2014, Batey, 2009, Batey et al., 2006). Heavy machinery used in wet soil conditions causes a great deal of compaction (Gameda et al, 1987). Soils that have a naturally low permeability at some depth are more susceptible to deep compaction in humid and cold climates (Gameda et al, 1987).

A good soil drainage ensures the elimination of excess water in the soil profile (Spoor, 2006). The use of subsurface drainage allows compensating the deficient drainage. Drainage improves the productivity of the poorly drained soil by lowering the water table, providing more air to the soil and reducing waterlogged conditions that increase compaction risks (Luthin, 1957, Spoor, 2006, Mountain et al., 2009, Ali, 2011). For low permeability soils, surface drainage is also essential to remove depressions that increase compaction risks under waterlogged conditions. Surface drainage and water flow has traditionally been promoted in Québec by bedding systems

with ridges and furrows. This type of surface drainage still persists today even after subsurface drainage implementation in poorly drained soils or soils with low permeability.

Subsoiling is used to break compact soil layers, to improve soil aeration in the root zone and to reduce root resistance to soil penetration. Subsoiling increases crop productivity and improves soil structure by leaving more crop residues and roots in the soil (Hall et al, 2005, Spoor, 2006, Batey, 2009, Weill, 2015). Subsoiling reduces soil compaction if subsurface drainage is adequate (Spoor, 2006; Weill, 2015). Subsoiling allows water to seep into the soil if the layer that impedes water is thin and shallow (i.e. less than 50 to 70 cm from the surface of the soil). This technique is effective because it eliminates the perched water table and promotes infiltration (Spoor, 2006; Weill, 2015). Subsoiling could also help draining the perched water table, if soil tillage is oriented toward a subsurface drain, a ditch, or an open drain. Several studies have been done on the mole drainage to create temporary drainage channels in low permeable clay soils (Rivard, 1993). Despite its limited lasting effect, this technique has allowed to significantly lower the water table depth up to three years after moling in heavy clay soils in Abitibi (Rivard, 1993). Rivard (1993) has demonstrated the relevance of orienting subsoiling toward an outlet. Therefore subsoiling could lower the water table in low soils permeability. On soils featuring bedding and subsurface drainage, water depth may be influenced by several parameters other than subsoiling only.

The crop yield may also be affected by compaction and the state of drainage of the soil on a complex soil rounded beds and a drainage system. For example, maize yields vary according to the horizontal distance to the subsurface drains, but these variations depend also on the prevalence of wet weather (Kladivko et al, 2005). This variation of performance according to the horizontal distance to the subsurface drains would be indicative of compaction or poor drainage problems between the subsurface drains (Kladivko et al, 2005).

The objectives of this study were to compare, for a low permeable soil, the effectiveness of random forests and multiple linear regressions to determine the effects of subsoiling direction and surface/subsurface drainage systems parameters on 1) plant height and 2) water table depth. The effects on plant height were investigated within a complete field of two beds (complete design), whereas a reduced design, in the lowest areas of the field, was used to determine the effect of parameters on both plant height and water table depth. The hypothesis

was that subsoiling of low permeability soils improves their internal drainage when oriented perpendicularly to the direction of the surface and subsurface drains.

2.2. Material and Methods

2.2.1. Site and Climate

The experimental site was located at Saint-Henri-de-Lévis near Québec City, Canada (latitude 46° 41' 15.72 "and longitude -71° 0' 57.60") and subsoiling was performed in the fall of 2014. The elevation of the site varies from 85.6 to 87.5 m above the mean sea level. The climate is wet and cold temperate (annual average temperature of 4.4°C). The normal annual rainfall is 1253 mm; (local weather station located 13 km from the experimental site). The soil is from the Kamouraska soil series (clay soil with natural low permeability < 20 cm/day)) and shows significant compaction at a depth from 30 to 60 cm below soil surface. Soil texture varies from clay loam (38.9% clay, 39.4% silt) at the surface (0-30 cm depth), to a silty loam (45.6% clay, 47.2% silt) at the 30 to 90 cm depth (Table 2). After the installation of subsurface drains (10 cm in diameter and 13 m spacing) in 1980, the field had been remodeled to improve the flow of surface water, forming two rounded beds of 40 m wide and 1 m high. Fig. 5 shows the topography of the rounded beds and the position of subsurface drains. The vertical distance to the subsurface drains located near the open drains varies from 46 to 70 cm from the soil surface, whereas the subsurface drains located near the top of the beds ranges from 140 to 150 cm from the surface of the soil.

2.2.2. Experimental Setup

Conventional subsoilers pulled by farm tractors on tires, generally operate at variable depths due to the required strength and reduced traction. By contrast, a subsoiler mounted on a bulldozer operates at a more constant depth due to its more constant traction. Working soil at more than 70 cm depth should be more constant with bulldozer.

The experimental field is shown in Fig. 5, which gives an overview of the field boundaries, open and subsurface drains, and a general idea of surface topography. Fig. 6 shows a cross section of

the field, which represents the direction perpendicular to both surface (open drains) and subsurface drainage systems, the water table observation wells and the two beds. The complete setup is the entire field (Fig. 5) and was used to estimate the effect of various parameters on plant height. It also allowed comparing the five subsoiling treatments, repeated three times in randomized complete block design: (1) no subsoiling, (2) parallel subsoiling with a bulldozer, (3) parallel subsoiling with a tractor, (4) perpendicular subsoiling with a bulldozer, and (5) perpendicular subsoiling with a tractor. The terms “parallel” and “perpendicular” refer to the direction with respect to the direction of the drainage systems, and abbreviated “Dir_X” and “Dir_Y”, respectively. The reduced setup refers to a specific part of the field where we measured plant height within a radius of 0.75 m around each observation well. Observation wells were located within 3 m; 6 m and 9 m respectively each side of the open drains in the lowest part of the beds (Fig. 6). To optimize cracking, subsoiling was performed on 15 September 2014, before the rainy season, when the soil is dry and less plastic. Deep subsoiling was executed at a constant depth of 80 cm with a John Deere 850J LGP 200 HP RAM bulldozer using three straight teeth without fins, 120 cm apart. Conventional subsoiling, to a depth ranging from 60 to 80 cm, was achieved with a farm tractor JCB 8280 Fastrack of 280 HP and a LaSole subsoiler with 5 dishes straight teeth without fins, 75 cm apart. The locations of all plots are shown in.

2.2.3. Water table monitoring

A set of 198 observation wells were dug in May 2016 with an Edelman Dutch Auger of 5 cm in diameter, at a depth of 60 cm, in the lower part of the field between the two beds (Fig. 7); this was done in order to monitor the water table level as affected by the various subsoiling treatments. The bottom of the beds is the location in the experimental setup that is most affected by the effects of the high water tables. For plots closest to an open drain, 9 observation wells were installed at the quarter- and half-subsurface drain spacing, as shown in Fig. 7.

We covered each well with a polystyrene cap to protect it from direct precipitation. Water table measurements were taken manually with a laser rangefinder (GLM 100C, Bosch tools, it) on eight dates from 3 May to 2 August 2016, after rainfall events; the average water table depth from these eight dates was used in this study.

2.2.4. Image Acquisition by Drone and Production of Digital Elevation Models

High-resolution photogrammetry offers the advantage of accurate determination of plants height. It is a good indicator of biomass productivity and yield (Yin et al, 2011, Zhang, 2012, Geipel, 2014). However, this requires the digital elevation models to be accurate both for the soil surface and the canopy elevations (Grenzdorffer, 2014). A drone Iris (3d Robotics Iris +, Berkeley, California) has been used to acquire images and create digital elevation (DEM) models of the soil surface and top of the maize plants. A flight carried out on 24 August 2016 was used to create a DEM of the top of the maize plants in the R1 (silks are visible outside the husk) of grain filling stage. Another flight was carried out on 19 November 2016, after the corn was harvested to create a DEM of the soil surface. These DEMs had horizontal and vertical resolutions less than 5.5 and 8 cm, respectively. The height of corn plants was calculated by difference between the soil surface and canopy surface DEMs.

Several studies (Lopes et al., 2012, My et al., 2001, Raun et al., 2001, Freeman et al., 2007, Yin et al., 2011, Geipel, 2014, Sulik, 2016), as well as measures with an experimental thresher on some rows of our experimental setup, show a good correlation ($R^2= 0.7$) between yield and the height of corn plants. To estimate the height of the plants on the complete setup and to compare them to yield measured in some locations, the average plant height was calculated on areas that were equivalent to those of yield measurements, i.e. two rows 0.762 m wide by 20 m long centered inside each experimental plot. In order to avoid bias in the calculation of plants height, the edges of the plots and some areas showing a lack of plant emergence or inadequate seeding rate, were subtracted from the analyzed surface. For the reduced setup, evaluation the plants height around the water table observation wells was done for circular surfaces with a radius of 0.762 m around the well. All analyses of images were made with version 2.18.11 of QGIS (QGIS Development Team, 2002).

2.2.5. Map Editing (GIS)

The field data was georeferenced using a DGPS (Differential Global Positioning System) receiver with a horizontal accuracy of 0.60 m and a GPS - RTK (real time Kinematic) receiver with a

horizontal precision less than 0.05 m. The cartographic data processing was carried out with version 10.4 of ArcMap (ESRI, USA, 1969) and version 2.14.11 of QGIS (QGIS Development Team, 2002).

2.3.Data Analysis

The data was processed and analyzed with the R program (version 3.4.1) on RStudio (R Core Team, 1993).

We choosed to deal with that strategy because the results of the covariance analysis (block effect and interaction) on different treatments were very low , so non-significant (see the results of the analysis in annex).

2.3.1 Multiple Linear Regression with R

The MLR models included response variables of plant height (Plant_height) and mean water table depth (Mean_WTD), and the predictive variables were depth to subsurface drains (SSDrain_Depth), horizontal distance to subsurface drains (SSDrain_Distance), horizontal distance to the field open drains (OD_Distance), position on the field represented by the perpendicular (Dir_Y) and parallel (Dir_X) directions to the subsurface drainage systems, and direction of subsoiling (SS_direction). The partial R^2 were calculated with the package realimpo (Groemping, 2013).

2.3.2 RF analysis with R

The randomForest package in R was used for RF modeling using the same variables as for MLR with parameters mtry = 3, ntree = 2000, and nodesize = 1 (Liaw et al, 2002). The RF method requires relatively few parameters to be set; the value of mtry = 3 was determined with the function tuneRF of the same package. Precision of the models did not improve for ntree values above around 1000, so a safe ntree value of 2000 was used in the analyses. The RF allows exploring the relationships between the response and predictive variables; these relationships

were determined by the function “importance” of the RF package, in particular for predictive variables: direction of subsoiling (SS_direction), SSDrain_Distance, Dir_X, Dir_Y and SSDrain_Depth which may be important in explaining the corn plant performance. All of the data was randomly split into a calibration dataset (75% of the data) and a validation dataset (25% of the data). The calibration dataset was used to define the RF model. The validation dataset was not used in calibration. As RF is not very sensitive to overfitting, a total model that was based on the complete dataset was also computed.

2.4. Results

2.4.1 Complete Experimental Setup: Height of Corn Plants

The height of corn plants was estimated on 479 polygons distributed across the complete experimental setup and averaged 1.32 m, with minimum and maximum values of 0.56 and 2.1 m, respectively.

The MLR model included six variables, four of which were significant in explaining Plant_height (Table 3a). SSDrain_Depth was the variable most significantly correlated to Plant_height ($p \leq 0.001$); resulting in a partial R^2 contribution of 0.36 to the model, with a relative importance of 52.3%; its positive relationship indicated an increase in Plant_height with SSDrain_Depth. SSDrain_Distance was the second significant variable ($p \leq 0.001$) in the model, with a partial R^2 contribution of 0.23, with a relative importance of 33.8%. Its negative relationship with Plant_height indicates that corn plants were smaller away from subsurface drains. Dir_X brings a R^2 partial contribution of about 0.07 ($p \leq 0.001$), with a relative importance of 10.4%, while Dir_Y ($p \leq 0.001$) explains 0.02 of the R^2 and a relative importance of 2.5%. There was a gradient of productivity along the two beds on the X axis (i.e. in the direction parallel to the field and to the drainage systems), which seems linked to the general elevation of the land (Fig. 5). SS_direction (non-significant) affects very slightly Plant_height with a R^2 partial contribution of 0.006 and relative importance of 0.9% (Table 3a). OD_Distance does not significantly contribute to the model. The effects of subsoiling are insignificant compared to the effects of the position on the beds, that is mainly reflected by SSDrain_Depth, and distance to the drains SSDrain_Distance.

The RF model provided a good prediction of Plant_height, comparing results using the calibration data ($R^2= 0.97$, Fig. 8a) and the verification data ($R^2= 0.84$, Fig. 8b). Using all the data, RF could predict Plant_height with a $R^2= 0.98$ (Fig. 8c), compared to MLR ($R^2= 0.69$, Fig. 8i). Fig. 8d shows the index of increase in nodes purity (IncNodePurity) which ranks the contribution of the different predictor variables included in the model. SSDrain_Depth appears as the most important variable (Fig. 8e), with a relative importance of 33.7% (Table 3a). It is followed by Dir_Y with a 25.2% contribution to the model, and SSDrain_Distance with a 23.2% contribution. Dir_X contributed 11.2%, and OD_Distance 5.2%. SS_direction contributed with only 1.5% to the model, indicating a negligible effect of subsoiling on plant height. The graphs of the selected

variables show how Plant_height can be accurately estimated using RF, even if some variables are correlated with each other and not necessarily linearly related to the predicted variable. For example, SSDrain_Depth matches the elevation on the beds which is also expressed by Dir_Y. Plant_height roughly follows a linear relationship with SSDrain_Depth (Fig. 8e), whereas its relationship with Dir_Y is non-linear; it increases and decreases according to the position on the two beds (Fig. 8f). Fig. 8e, however, suggests an increase in Plant_height, as soon as SSDrain_Depth is greater than approximately 1.1 m, which is on the upper part of the beds. Plant_height also sharply declines as soon as SSDrain_Distance exceeds 2 to 3 m from the drains (Fig. 8g). This observation supports the idea that a spacing of about 6 to 7 m between subsurface drains would maximize Plant_height; the actual distance being about 13 m. Plant_height variation according to Dir_X suggests weaker yields in the lower part of the field (Fig. 8h). This could be due either to surface soil stripping when the beds were constructed, to less well-drained conditions, or to the greater machinery traffic in this part of the field. OD_Distance and SS_direction only contributed marginally to the model and their effects on Plant_height were negligible compared to the effects of SSDrain_Depth, Dir_Y, and SSDrain_Distance.

2.4.2 Reduced Experimental Setup: Height of Corn Plants

On this part of the setup, the height of corn plants averaged 0.91 m with minimum and maximum values of 0.38 and 2.2 m, respectively. Mean_WTD, which may affect Plant_height, was included in the independent variables.

The MLR model included five significant variables. Table 3b shows the contribution of each independent variable to Plant_height. Mean_WTD was best correlated with Plant_height (partial $R^2 = 0.26$, $p \leq 0.001$, relative importance = 37.6%). The positive relationship between Mean_WTD and Plant_height indicated that corn plants were smaller when the water table was near the surface of the soil. OD_Distance (partial $R^2 = 0.20$, $p \leq 0.001$) and SSDrain_Depth (partial $R^2 = 0.12$, $p \leq 0.05$) had relative importance of 29.5% and 17.6%, respectively. The positive relationship between OD_Distance and Plant_height indicates that corn plants were taller as the OD_Distance increased. Dir_X (partial $R^2 = 0.06$, $p \leq 0.001$) and SSDrain_Distance (partial $R^2 = 0.03$, $p \leq 0.001$) made lower contributions with respective relative importance of

8.7% and 5.1%. Finally, effects of Dir_Y (partial $R^2 = 0.006$, non-significant) and subsoiling (partial $R^2 = 0.003$, non-significant) explained only a very small part of the variation of Plant_height. Overall, on the reduced experimental setup, Plant_height increased with Mean_WTD, OD_Distance, and SSDrain_Depth, and decreased with SSDrain_Distance.

With calibration (Fig. 9a) and verification data (Fig. 9b), the RF model predicted Plant_height with R^2 values of 0.95 and 0.69, respectively. The RF model based on the complete dataset (198 observations) resulted in a higher R^2 ($R^2 = 0.97$, Fig. 9c) than the MLR ($R^2 = 0.69$, Fig. 9k). The order of importance of each independent variable is presented in Fig. 9d. Mean_WTD and OD_Distance, were the two most important variables affecting Plant_height with relative importance of 24.1 and 23.6%. Dir_Y contributed to 15.5%, SSDrain_Depth to 12.3%, Dir_X to 12.1%, SSDrain_Distance to 10.7%, but the SS_direction had very little effect. Mean_WTD positively affected Plant_height, but non-linearly. When the mean water table depth was less than 25 cm from the surface (Fig. 9e), Plant_height decreased exponentially; denoting at what depth corn plant roots become very sensitive to saturated water conditions. At the opposite, when the water table was on average more than 40 cm deep, Plant_height was maximum. A reduction of OD_Distance reduced Plant_height (Fig. 9f). Within 6 m of the open drain, Plant_height was always less than 0.9 m. Beyond that distance, Plant_height gradually increased and stabilized at a little more than 1.1 m. A large area near the open drain, a 12-m width, produced low Plant_height values, therefore low corn yields. Plant_height varied with Dir_Y, increasing when moving away from the open drain, centered at $y = 40$ m (Fig. 9g). Plant_height was positively linked to SSDrain_Depth (Fig. 9h). SSDrain_Depth is in fact related to the elevation on the beds and to OD_Distance. Subsurface drains were located at a 0.60-m depth at the bottom of the beds near the open drains, and at a 1.40-m depth higher on the beds. Fig. 9h shows an increase in the height of the plants as soon as SSDrain_Depth exceeds 0.80 m. Plant_height was relatively constant with Dir_X (Fig. 9i) up to about 250 m along the setup; values of Plant_height averaged 0.90 m, but beyond 250 m, Plant_height increased progressively. Plant_height decreased with the increase of SSDrain_Distance in this area near the open drains and significantly after 3.5 m (Fig. 9j) from the subsurface drains. A subsurface drain spacing of 7 m would maximize Plant_height and therefore corn yield, in this area near the open drains. SS_direction effect on Plant_height was relatively unimportant (Fig. 9k). Mean_WTD, and the variables related to the position on the rounded beds (OD_Distance, SSDrain_Depth, Dir_Y and Dir_X) mostly influenced Plant_height.

2.4.3 Reduced Experimental Setup: Mean Water Table Depth

The MLR model included three significant variables for predicting Mean_WTD in the reduced setup (Table 3c). Dir_Y was the variable with the most positive and significant effect (partial $R^2 = 0.09$, $p \leq 0.001$, with a relative importance of 3.8%). Its positive relationship with Mean_WTD indicated that Mean_WTD was deeper on the top of the beds. It was followed by OD_Distance (partial $R^2 = 0.06$, $P > 0.10$, relative importance =22.5%). The positive relationship of OD_Distance with Mean_WTD indicates that Mean_WTD was deeper away from the open drains. SSDrain_Distance and SSDrain_Depth both had a negative impact on Mean_WTD with partial R^2 values of 0.04 ($p \leq 0.001$, relative importance =14.8%) and 0.03 (non-significant, relative importance =10.8%), respectively. The negative relationship between Mean_WTD and SSDrain_Distance indicates that Mean_WTD was shallower further away from the subsurface drains, reflecting the effect of a conventional drainage system. Mean_WTD would also increase with Dir_Y and OD_Distance, but these relationships were not significant.

With the RF model, calibration data (Fig. 10a) predicted Plant_height with a R^2 of 0.86. Validation data produced a $R^2 = 0.45$ (Fig. 10b). Using the complete dataset, RF could predict Mean_WTD with a $R^2 = 0.95$ (Fig. 10c), much higher than the MLR ($R^2 = 0.28$; (Fig. 10i)). Fig. 10d shows the order of importance of the various independent variables included in the model. Variable Dir_Y was the most important with a relative importance of 25% (Table 3c). It was followed by OD_Distance (20.4%), Dir_X (19.4%), SSDrain_Distance (17.3%), SSDrain_Depth (14.5%), and SS_direction (3.5%). The variable Dir_Y showed that at the bottom of the beds, close the open drains (open drains at $y = 40$ m), Mean_WTD was closest to the surface (Fig. 10e), while farther away from the open drains, Mean_WTD increased. Mean_WTD was also closer to the soil surface when OD_Distance was smallest. This effect was noticeable to a distance of approximately 10 m from the open drains, after which water table depth remained on average, deeper than 21 cm from the soil surface (Fig. 10f). Dir_X had a negative effect on Mean_WTD within the first 100 m along the reduced setup (Fig. 10g). The beds as well as the open drains, predominantly influenced Mean_WTD. SSDrain_Distance had a negative effect on Mean_WTD, but in nonlinear way (Fig. 10h). Mean_WTD decreased abruptly at 3.5 m from the subsurface drains, showing a perched water table at a further distance from the drains.

2.5. Discussion

The use of MLR could, in some cases, give the same order of importance than RF. This was the case for SSDrain_Distance and SS_direction in the Plant height analysis on the complete experimental setup (Table 3a), as well as Mean_WTD and OD_Distance in the Plant_height analysis on the reduced experimental setup (Table 3b). This was also the case for Dir_Y, OD_Distance, Dir_X, SSDrain_Distance, SSDrain_Depth and SS_direction in the analysis of Mean_WTD on the reduced setup (Table 2c). However, the number of significant variables was relatively less for MLR than for RF. In fact, multicollinearity between some predictor variables can reduce their contribution to the MLR model (Table 4).

RF is more robust than MLR for prediction, because it better integrates the complexity of relationships between data (Breiman 2001). The RF and the MLR have been used and compared to predict productivity of ecological systems including crop yield (Fukuda et al., 2013, Vincenzi, 2011, Jeong, 2016). RF highlighted the effects of the position on the beds, reflected by SSDrain_Depth and Dir Y in the complete setup, and the critical value of SSDrain_Distance (2 to 3 m). Plant_height and Mean_WTD also were mostly influenced by the position on the beds in the reduced setup, reflected by the optimal distance to the open drains (> 8 m) and Dir_Y. RF also highlighted the distance of subsurface drains (> 4 m) in the reduced setup, and the critical mean water table depth (< 0.25 m). This depth corresponds to the threshold below which the corn plants would suffer from water saturated conditions in soil. The R^2 obtained with RF are much larger than those obtained with MLR in predicting Plant_height and Mean_WTD, using either calibration or validation datasets. Unlike MLR, RF seems to capture the nonlinear relationships between the independent variables and the dependent variable. RF also provided useful information about the importance of independent variables and their relationship with the dependent variable. Partial dependency graphs are useful in assessing the relationship between each predictor and the response variable (Friedman, 2001, John read, 2010). The purity of nodes was used to evaluate the most important variables for Plant_height and Mean_WTD (Table 3). For the complete experimental setup, graphs of variables importance and of partial dependencies showed that Plant_height was more affected by the position relative to the beds and the subsurface drainage systems indicated by SSDrain_Depth, Dir_Y and SSDrain_Distance (Fig. 9). On the reduced setup, near the open drains, Plant_height was more affected by the position on the beds (OD_Distance) and Mean_WTD (Fig. 10). Mean_WTD was more affected by

the position on the beds (OD_Distance and Dir_Y) (Fig. 11). These results highlight the usefulness of RF to detect and rank various factors related to farm practices such as surface and subsurface drainage implementation and subsoiling on crops. Multiple benefits exist for the use of RF compared to other methods of analysis, including providing better information on the mechanisms involved from various sources of data (Breiman, 2001b). The results of this study strongly support this for predicting Plant_height and Mean_WTD. The inclusion of other predictor variables such as minimum or maximum water depth could improve the performance of the MLR (data not shown), but the performance did not reach the levels of RF; Mean_WTD was therefore used in order to compare the two statistical methods on the same basis (Butler, 2013). When variables are highly correlated with each other, RF can be very effective to bring out certain differences, where MLR does not (Berk, 2008, Gromping, 2009).

Subsoiling tests effects on Plant_height and Mean_WTD were minor in this study. Regarding the different subsoiling treatments, including SS_direction, results were mostly non-significant. RF and MLR results showed that subsoiling did not have any significant effect on Plant_height or Mean_WTD compared to other factors. However, subsoiling perpendicular to drains seemed to improve very slightly Plant_height and Mean_WTD. The Bulldozer kept the shanks at a more constant depth than the conventional subsoiler operated with a tractor, but no significant effect was detected (data not reported).

Several hypotheses could explain the mediocre performance of subsoiling. The first is related to the ephemeral effect of mechanical work on soil. For instance, without additional improvement of soil by colonization of dense root systems and biomass return, soil quickly returns to its former state of compaction, following renewed traffic on the soil and natural settling of the worked soil. Traffic control of tractors, spreaders, combines and grain carts etc. would be essential to extend the benefit of subsoiling operations (Raper et al., 2007). Some clay soils are also naturally compact or show particle sizes with shapes of clay and silt that are not as efficient as other clays of the St. Lawrence Plain to develop their structure at deeper layers.

2.6. Conclusion

This study aimed 1) to assess how subsoiling and subsurface and surface drainage systems affect the height of corn plants and the water table depths in low permeable clay soils and 2) and to compare the capability of RF and MLR to detect these effects.

Subsoiling tests carried out in this study, more specifically subsoiling direction did not have a very strong effect on plant productivity or water table depth. RF and MLR both were able to detect the main factors affecting plant productivity or water table depth, except that RF included additional factors and was useful to highlight the nonlinear effects of the different predictor variables, some of them related to the position on the beds. Coefficients of determination obtained with RF ($R^2 \geq 0.98, 0.97, 0.95$) were higher than those obtained with MLR ($R^2 \geq 0.69; 0.69; 0.28$). RF also helped to visualize critical values for horizontal distance to drains or water table depth.

On the complete experimental setup in MLR, Plant_height was mainly affected by its position on the beds reflected by SSDrain_Depth, while in RF, it was also affected by Dir_Y. An improvement of productivity in terms of Plant_height is seen as SSDrain_Depth exceeds 1.1 m, near the top of the beds. SSDrain_Depth is mainly determined by the position on the beds which is also expressed by Dir_Y.

Plant_height was negatively influenced by SSDrain_Distance and significantly from 2 to 3 m from subsurface drains. A spacing of about 6-7 m between subsurface drains would maximize Plant_height.

On the reduced setup in MLR, Plant_height was mainly affected by Mean_WTD, OD_Distance, and SSDrain_Depth, while in RF, it is also influenced by Dir_Y. When the water table was shallower than 0.25 m approximately, Plant_height became sensitive to waterlogged conditions and when water table level was more than 0.40 m depth, Plant_height was at its maximum. As on the complete experimental setup, Plant_height was mainly affected by its position on the beds, reflected by SSDrain_Depth, OD_Distance and Dir_Y. SSDrain_Dept was 0.60 m at the bottom of the beds, near the open drains, and 1.40 m higher on the beds. When SSDrain_Depth exceeded 0.80 m, an increase in Plant_height was observed. Water logging mostly affected corn productivity on the lower part of the beds. Within 6 m from the open drains (OD_Distance),

Plant_height was less than 0.9 m and beyond this distance, Plant_height increased gradually and stabilized at slightly more than 1.1 m. Plant_height was more or less constant (0.90 m) with Dir_X up to about 250 m along the setup, but over 250 m, Plant_height increased progressively. As on the complete experimental setup, Plant_height decreased with an increase of SSDrain_Distance in the area near the open drains and significantly dropped at 3.5 m from the subsurface drains. A spacing of 7 m between subsurface drains would maximize Plant_height, in this area near the open drains.

On the reduced setup, Mean_WTD was mainly influenced by position on the beds (Dir_Y, OD_Distance and SSDrain_Depth), X direction (Dir_X) and SSDrain_Distance, evidenced either in MLR or RF. However, some of these relationships were not statistically significant in MLR. Mean_WTD was higher at the bottom of the beds near the open drains. With RF, a Mean_WTD less than 22 cm from the soil surface was evidenced within 10 m from the open drains (OD_Distance).

These results highlight the usefulness of RF for detecting different effects or factors related to land improvement and agricultural practices, such as drainage and subsoiling, on crop productivity and environment. For this reason, better cropping practices are needed to ensure long-term effects of subsoiling and the return on investment. Integrating cover crops and autumn cereals that will help dry out the soil in spring and minimize work when the fields are waterlogged, and controlling machinery traffic and weight should be strongly encouraged. As a result, surface and subsurface conditions were a lot more important than subsoiling effects the second year later. RF and MLR showed similar performances in ranking relative importance of these variables, with the only difference that RF was more powerful in terms of prediction and in highlighting the main correlations and the non-linear effects of variables on Plant_height or Mean_WTD.

2.7. Acknowledgements

This research project was achieved with financial assistance from the program Innov' agri-food Action, a program from the agreement of growing the future; agreement concluded between the Ministry of Agriculture, fisheries and food and Agriculture and Agri-Food Canada.

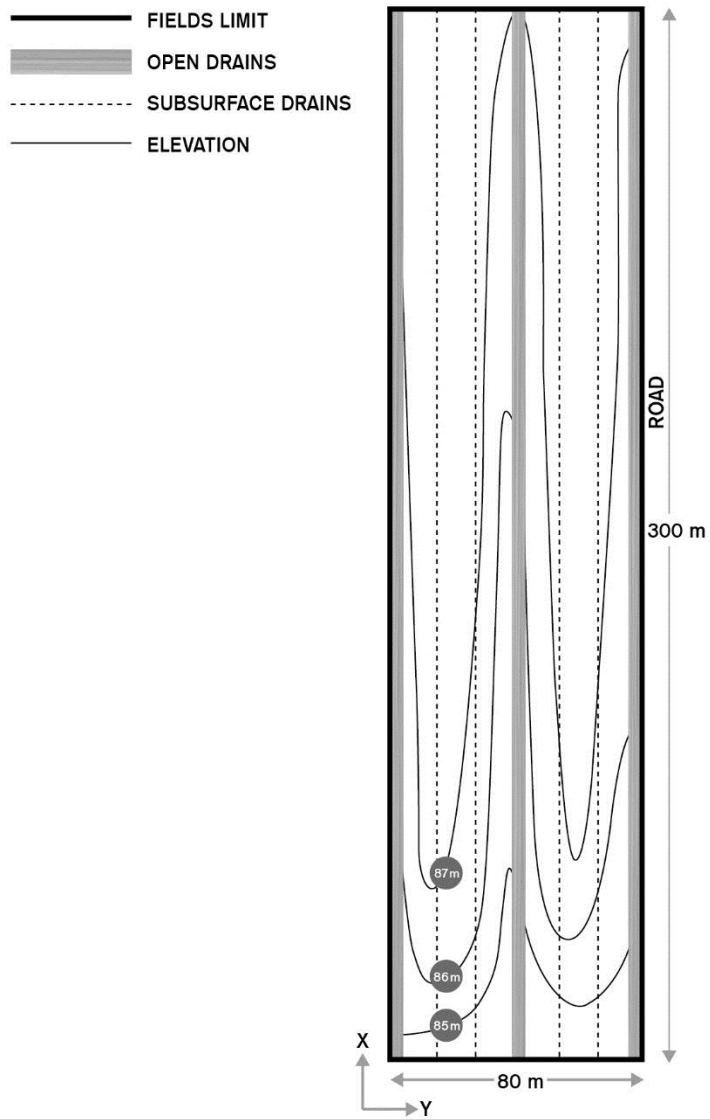


Fig. 5. Delimitation of the experimental setup with its open drains, subsurface drainage system and two beds.

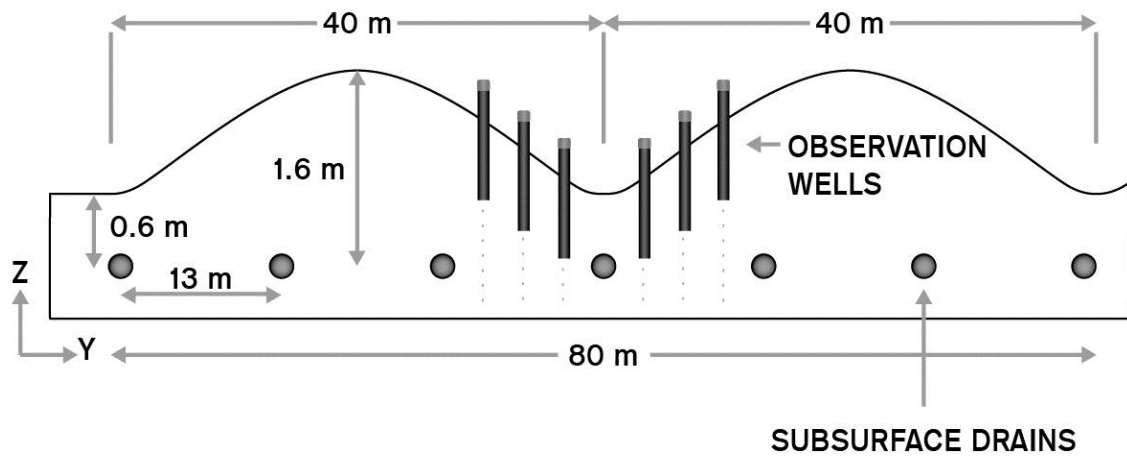


Fig. 6. Cross-section of the two beds on the experimental setup.

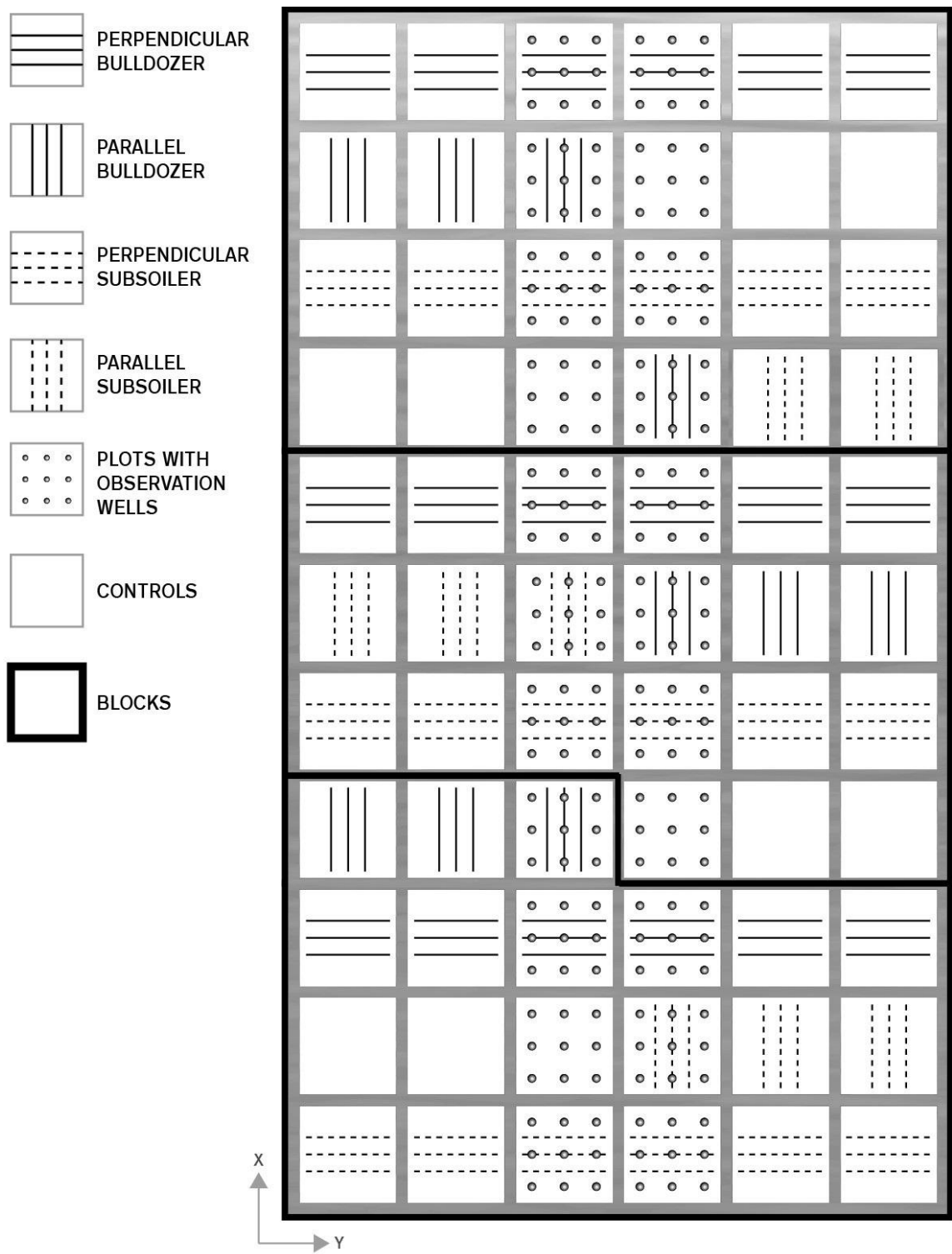


Fig. 7. Subsoiling treatments allocation and observation wells location on experimental setup.

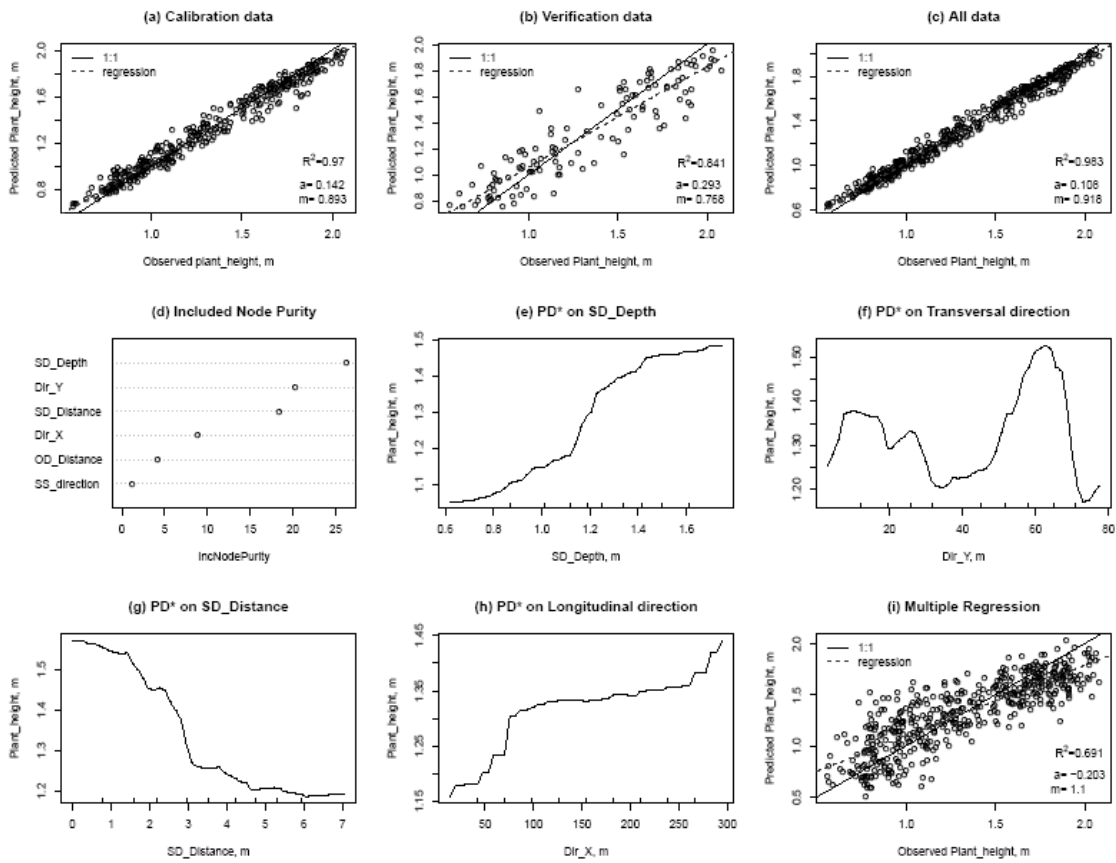


Fig. 8 . Plant_height regression models (RF and MLR) on the complete experimental setup (479 obs.). Performance of the Random Forests model for calibration (a, $R^2= 0.97$), validation (b, $R^2= 0.84$) and complete dataset (c, $R^2= 0.98$). Contribution of the predictor variables to the estimate of Plant_height, expressed by purity of nodes (d). Curves of partial dependence on SSDrain_Depth (e), Dir_Y (f), SSDrain_Distance (g) and Dir_X (h). Performance of the MLR on the complete dataset. (i, $R^2= 0.69$).

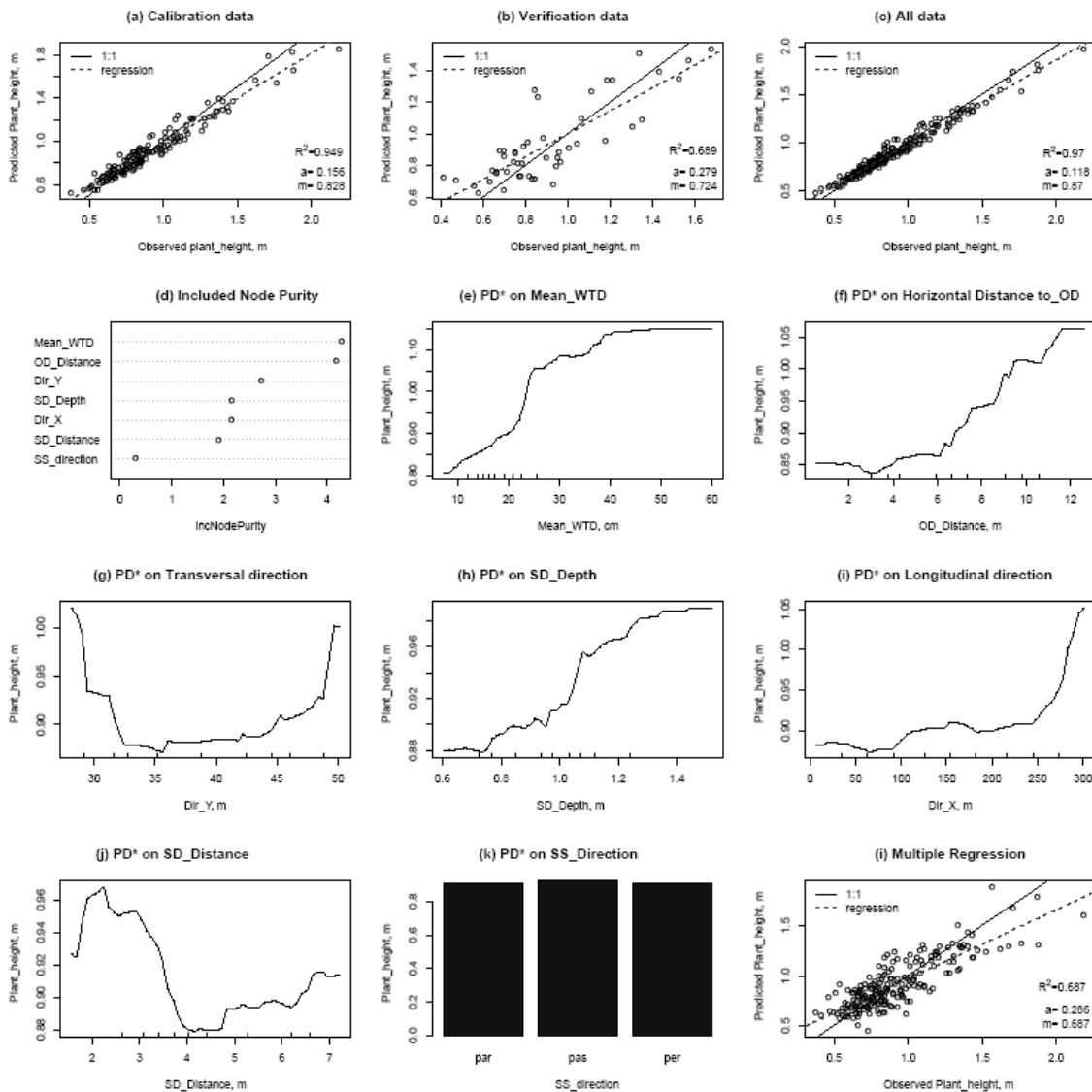


Fig. 9. Plant_height regression models (RF and MLR) on the reduced experimental setup (198 obs.). Performance of the Random Forests model for calibration (a, $R^2=0.95$), validation (b, $R^2=0.67$) and complete dataset (c, $R^2=0.97$). Contribution of the predictor variables to the estimate of Plant_height expressed by the purity of nodes (d). Curves of partial dependence on Mean_WTD (e), OD_Distance (f), Dir_Y (g), SSDrain_Depth (h), Dir_X (i), SSDrain_Distance (j), SS_direction (k). Performance of the MLR on the complete data set (l, $R^2=0.69$).

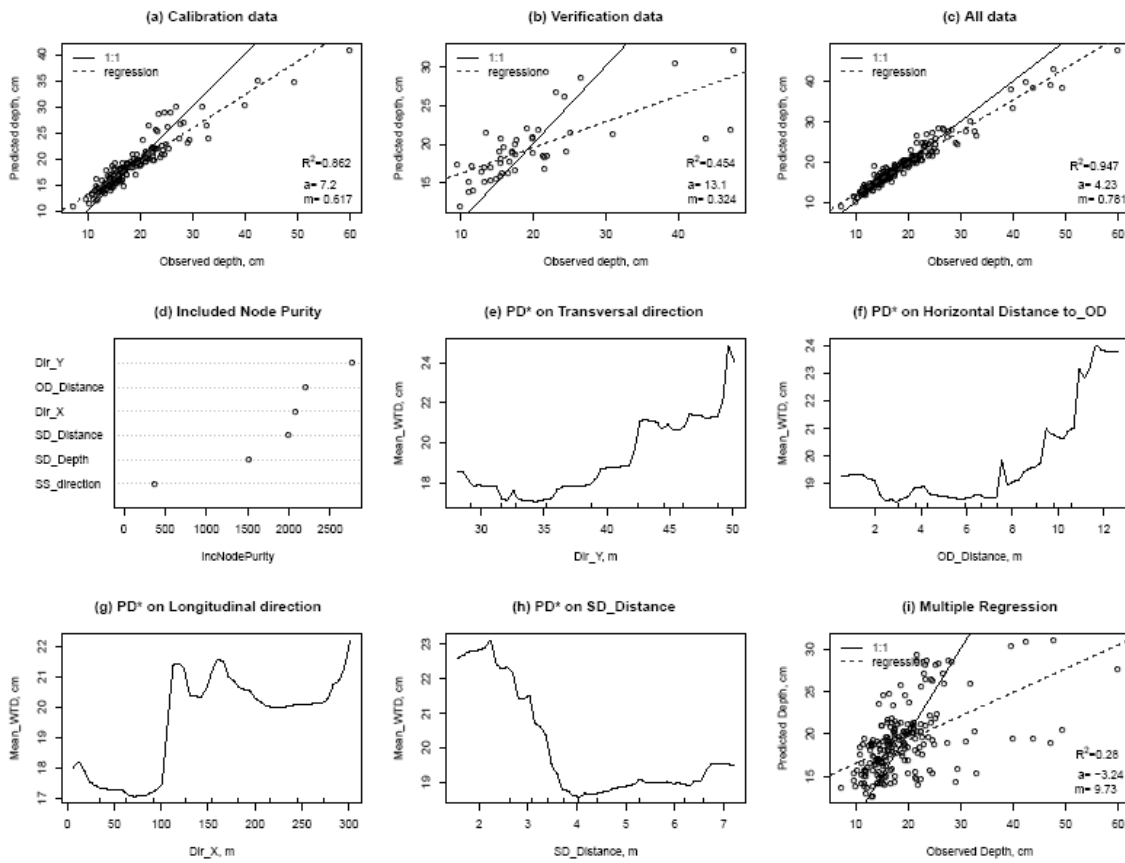


Fig. 10. Mean_WTD regression models (RF and MLR) on the reduced experimental setup (198 obs.). Performance of the Random Forests model for calibration (a, $R^2= 0.86$), validation (b, $R^2= 0.45$) and complete dataset (c, $R^2= 0.95$). Contribution of the predictor variables to estimate Mean_WTD expressed by the purity of nodes (d). Curves of partial dependence on Dir_Y (e), OD_Distance (f), Dir_X (g), SSDrain_Distance (h). Performance of the MLR on the complete dataset (i, $R^2= 0.28$). (i, $R^2= 0.28$).

Table 2. Physical properties of soil

Depth cm	OM - (%) -	Sand ----- % -----	Silt ----- % -----	Clay ----- % -----	Textural class	ρ_A g/cm ³	ϕ_t cm ³ /cm ³
0-15	6.7	21.0	40.0	39.0	Clay loam	1.27	0.52
15-30	6.2	22.5	38.8	38.8	Clay loam	1.31	0.50
30-60	1.6	11.6	43.0	45.4	Silty clay	1.55	0.42
60-90	1.0	1.8	52.5	45.8	Silty clay	1.49	0.44

OM: Soil organic matter; ρ_A : Bulk density; ϕ_t : Total porosity

Table 3. Summary of the three case studies for each analytical method

	RF		MLR		
		Relative importance (%)	partial R ²	Meaning of the correlation	Relative importance (%)
(a) Full setup Plant_height					
SSDrain_Depth	26.6	33.7	0.3613	+	52.3***
SSDrain_Distance	18.4	23.2	0.2338	-	33.8***
OD_Distance	4.1	5.2	0.0005	+	0.1 ^{NS}
SS_direction	1.2	1.5	0.0060	+	0.9 ^{NS}
Dir_X	8.8	11.2	0.0718	+	10.4***
Dir_Y	19.9	25.2	0.01758	+	2.5***
Total	79.0	100	0.691	+	100
(b) Reduce setup					
Plant_height					
SSDrain_Depth	2.2	12.3	0.1209	+	17.6*
SSDrain_Distance	1.9	10.7	0.0353	-	5.1***
OD_Distance	4.2	23.6	0.2022	+	29.5**
Mean_WTD	4.3	24.1	0.2584	+	37.6***
Dir_X	2.1	12.1	0.0598	+	8.7***
Dir_Y	2.7	15.5	0.0066	+	1.0 ^{NS}
SS_direction	0.3	1.7	0.0032	+	0.5 ^{NS}
Total	17.7	100	0.687	+	100
(c) Reduce setup					
Mean_WTD					
SSDrain_Depth	1589.2	14.5	0.0302	+	10.8 ^{NS}
SSDrain_Distance	1894.8	17.3	0.0413	-	14.8***
OD_Distance	2230.1	20.4	0.0628	+	22.5 ^{NS}
Dir_X	2122.9	19.4	0.0418	+	15.0**
Dir_Y	2733.7	25.0	0.0943	+	33.8***
SS_direction	382.0	3.5	0.0090	+	3.2 ^{NS}
Total	10952.7	100	0.28	+	100

Probability: NS = non-significant, * = ≤ 0.05 ; ** = ≤ 0.01 ; *** = ≤ 0.001 .

Table 4 Spearman correlations between the predictor variables

(a) complete setup	SSDrain_Depth	SSDrain_Distance	OD_Distance	Dir_X	Dir_Y
SSDrain_Depth	1				
SSDrain_Distance	-0.05	1			
OD_Distance	-0.03	0.05	1		
Dir_X	-0.37	0.00	0.05	1	
Dir_Y	-0.19	-0.08	-0.05	-0.06	1

(b) reduce setup	SSDrain_Depth	SSDrain_Distance	OD_Distance	Dir_X	Dir_Y
SSDrain_Depth	1				
SSDrain_Distance	0.07	1			
OD_Distance	0.81	0.00	1		
Dir_X	-0.34	0.08	0.10	1	
Dir_Y	-0.07	-0.03	0.03	0.03	1

2.8. Références

- Ali, H. 2011. Practices of Irrigation & On-farm Water Management: Volume 2. Practices of Irrigation & On-farm Water Management. Springer.
- Batey, T. 2009. Soil compaction and soil management-a review. *Soil use and management*, 25, 335-345.
- Batey, T., & McKenzie, D. 2006. Soil compaction: identification directly in the field. *Soil Use and Management*, 22, 123-131.
- Bendecheche, M., Kechadi, Stommel. & Le-Khac, N.-A. Efficient Large Scale Clustering based on data partitioning. Data Science and Advanced Analytics (DSAA), IEEE International Conference 2016, 2016. IEEE, 612-621.
- Berk, R. A. 2008. *Statistical learning from a regression perspective*, Springer.
- Breiman, L. 2001. Random forests. *Machine learning*, 45, 5-32.
- Breiman, L. 2001 b. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science*, 16, 199-231.
- Butler, E. E., & Huybers, P. 2013. Adaptation of US maize to temperature variations. *Nature Climate Change*, 3, 68-72.
- Carman, K. 2008. Prediction of soil compaction under pneumatic tires a using fuzzy logic approach. *Journal of Terramechanics*, 45, 103-108.
- Freeman, K. W., Girma, k., Arnall, D. B., Mullen, R. W., Martin, K. L., Teal'C, R. K. & Raun, W. R. 2007. By-plant prediction of corn drilling biomass and nitrogen uptake at various growth stages using remote sensing and plant height. *Agronomy Journal*, 99, 530-536.
- Friedman, J. H. 2001. Greedy function approximation: a gradient boosting machine. *Annals of statistics* , 1189-1232.
- Fukuda, S., Spreer, W., Yasunaga, E., Yuge, K., Sardud, V. & Muller, J. 2013. Random Forests modeling for the estimation of mango (*Magnifera indica* l. cv. Chok Anan) fruit yields under different irrigation regimes. *Agricultural water management*, 116, 142-150.
- Gameda, S., Raghavan, G., McKyes, E., & Thériault, R. 1987. Subsoil compaction in a clay soil. I. Cumulative effects. *Soil and tillage research*, 10, 113-122.
- Geipel, J., Link, J., & Claupein, W. 2014. Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models with year acquired unmanned aircraft system. *Remote Sensing*, 6, 10335-10355.

Gregorich, E.G., McLaughlin, N.B., Lapen, D.R., Ma, B.L. & Rochette, P. Soil compaction, both year environmental and agronomic culprit: nitrous oxide emissions and reduced plant nitrogen uptake Increased. *Soil Science Society of America Journal*, 78, 1913-1923.

Grenzdorffer, G. 2014. Crop height determination with UAS point clouds. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40, 135.

Gromping, U. 2009. Importance assessment variable in régression: linear regression versus random forest. *The American Statistician*, 63, 308-319.

Håkansson, I., & Reeder, R. C. 1994. Subsoil compaction by vehicles with high axle load-extent, persistence and crop response. *Soil and Tillage Research*, 29, 277-304.

Hall, M., & Anderson, W. 2005. Soil compaction in cropping systems: a review of the nature, causes and possible solutions. *Soil and tillage research*, 82, 121-145.

Hossain, A., Rahman, A. & Mohiuddin, A. K. M. 2011. Fuzzy expert system for controlling swamp field smart air-cushion tracked vehicle. *International Journal of Automotive Technology*, 12, 745.

Huang, Y., Lan, Y., Thomson, S. J., Fang, A., Hoffmann, W. C. & Lacey, R. E. 2010. Development of soft computing and applications in agricultural and biological engineering. *Computers and Electronics in Agriculture*, 71, 107-127.

Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., Timlin, D. J., Shim, K.-M., Gerber, J. S. & Reddy, V. R. 2016. Random forests for global and regional crop yield predictions. *PloS one*, 11, e0156571.

John Lu, Z. 2010. The elements of statistical learning: data mining, inference, and prediction. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173, 693-694.

Kladivko, E., Willoughby, G. & Santini, J. 2005. Corn growth and yield response to subsurface drain spacing we Clermont silt loam soil. *Agronomy journal*, 97, 1419-1428.

Liaw, A., & Wiener, M. 2002. Classification and regression by randomForest. *R news*, 2, 18-22.

Lopes, M. S., & Reynolds, M. P. 2012. Stay-green in spring wheat can be determined by spectral reflectance measurements (normalized difference vegetation index) independently from phenology. *Journal of Experimental Botany*, 63, 3789-3798.

Luthin 1957. Measurement of hydraulic conductivity In in situ: Drainage of agricultural Lands. *American Society of Agronomy*, Madison, WI, ASA Agronomy Monography 7, PP. 420-430.

Ma, B., DWYER, I. M., COSTA, C., COBER, E. R., & MORRISON, M. J. 2001. Early prediction of soybean yield from canopy reflectance measurements. *Agronomy Journal*, 93, 1227-1234.

Mountain, D., Cornu, S., Le Foresterier, I. & Cousin, I. 2009. Soil Drainage as year Active Agent of Recent Soil Evolution: A Review * 1 * 1Project supported by the Centre Region, France. *Pedosphere*, 19, 1-13

Pentoś, K., & Pieczarka, K. 2017. Year applying artificial neural network approach to the analysis of tractive properties in changing soil conditions. *Soil and Tillage Research*, 165, 113-120.

Raper, R., & Bergtold, J. 2007. In - row subsoiling: a review and suggestions for reducing cost of this conservation tillage operation. *Applied engineering in agriculture*, 23, 463-471.

Raun, W. R., Solie, J. B., Johnson, G. V., Stone, M. L., Lukina, E. V., Thomason, W. E. & Schepers, J. S. 2001. In-season prediction of potential grain yield in winter wheat using canopy reflectance. *Agronomy Journal*, 93, 131-138.

Rivard, P. J. 1993. Assessment of the technique of mole-drainage in heavy clay of the Abitibi-Témiscamingue region.

Spoor, G. 2006. Alleviation of soil compaction: requirements, and technical equipment. *Soil Use and Management*, 22, 113-122.

Sulik, J. J., & Long, D. S. 2016. Spectral considerations for modeling yield of canola. *Remote Sensing of Environment*, 184, 161-174.

Taghavifar, H., Mardani, A. & Taghavifar, I. 2013. A hybridized artificial neural network and imperialist competitive algorithm optimization approach for prediction of soil compaction in soil bin facility. *Measurement*, 46, 2288-2299.

Vincenzi, S., Zucchetta, M., Franzoi, P., Pellizzato, M., Pranovi, F., Leo, G. A. & Torricelli, P. 2011. Application of a Random Forest algorithm to predict spatial distribution of the potential yield of *Ruditapes philippinarum* in the Venice lagoon, Italy. *Ecological Modelling*, 222, 1471-1478

Weill. 2015. Guide to successful subsoiling. CETAB +, Victoriaville. Reference center in agriculture and agri-food sector in Quebec. 43 p.

Yin, X., McClure, M. A., Jaja, N., Tyler, D. D. & Hayes, R. M. 2011. In-season prediction of corn yield using plant height under major production systems. *Agronomy Journal*, 103, 923-929.

Zhang, I., & Grift, T. E. 2012. A LIDAR-based crop height measurement system for *Miscanthus giganteus*. *Computers and electronics in Agriculture*, 85, 70-76.

2.9. Annexe

2.9.1 Résultats des analyses de covariance pour tester l'effet des traitements et des covariables

Des modèles d'analyse de covariance ont été appliqués afin d'étudier les effets des traitements (deux types de sous-soleuses x deux directions de sous-solage et témoin sans sous-solage) sur la hauteur des plants de maïs sur les dispositifs complets et réduits, de même que la profondeur moyenne de la nappe sur le dispositif réduit. La distance à la raie et la distance au drain ont été utilisées comme covariables dans les modèles. Les effets aléatoires des blocs et des parcelles ont également été intégrés afin de tenir compte du dispositif expérimental. Dans tous les modèles appliqués, quelle que soit la covariable utilisée, la statistique F du test de l'effet des traitements est très faible et ne permet pas de conclure à un effet significatif. Chaque covariable utilisée a toutefois un effet très significatif sur la hauteur des plants et la profondeur moyenne de la nappe.

Hauteur des plants de maïs en fonction des traitements et de la distance au drain (dispositif complet)

Sur le dispositif complet, les Tests de F des effets de type 3 ne révèlent pas d'effet significatif des traitements (Prob. F = 0,3912) sur la hauteur de plants, mais un effet très significatif de la distance au drain (Prob. F <0.0001).

Tests des effets fixes de type 3				
Effet	DDL num.	DDL den.	Valeur F	Pr > F
trait	4	9.05	1.16	0.3912
DrainST_dist	1	464	172.74	<.0001

L'effet de la covariable distance au drain est linéaire, la hauteur des plants de maïs diminue de 11 cm à chaque m de distance au drain. Les effets de chaque traitement de sous-solage ne sont pas significativement différents du témoin.

Solution pour effets fixes						
Effet	trait	Estimation	Erreur type	DDL	Valeur du test t	Pr > t
Intercept		1.6930	0.06193	19.4	27.34	<.0001
trait	bélier_par	-0.08903	0.07835	12.6	-1.14	0.2771
trait	bélier_per	-0.00987	0.07145	8.69	-0.14	0.8933
trait	ss_par	0.02592	0.07650	11.5	0.34	0.7408
trait	ss_per	0.07215	0.07095	8.49	1.02	0.3373

Solution pour effets fixes						
Effet	trait	Estimation	Erreur type	DDL	Valeur du test t	Pr > t
trait	tem_pas	0
DrainST_dist		-0.1103	0.008396	464	-13.14	<.0001

Les comparaisons par contrastes des types de sous-soleuses (Bélier vs Sous-soleuse) ou du sens du sous-solage (parallèle ou perpendiculaire aux drains) de même que l'interaction et la comparaison au témoin ne révèlent pas non plus d'effet significatif des traitements.

Moyennes et contrastes calculés pour une valeur moyenne de la distance au drain (3.44 m)						
Libellé	Estimation	Erreur type	DDL	Valeur du test t	Pr > t	
Moyenne perpendiculaire	1.3445	0.04004	5.16	33.58	<.0001	
Moyenne parallèle	1.2818	0.04542	8.51	28.22	<.0001	
perpendiculaire VS parallèle	0.06270	0.05391	8.94	1.16	0.2750	
perpendiculaire VS témoin	0.03114	0.06201	9.58	0.50	0.6269	
parallèle VS témoin	-0.03156	0.06566	12.1	-0.48	0.6394	
Moyenne bélier	1.2639	0.04325	6.99	29.22	<.0001	
Moyenne sous-soleuse conventionnelle	1.3624	0.04238	6.48	32.15	<.0001	
Bélier VS sous-soleuse conventionnelle	-0.09848	0.05393	8.96	-1.83	0.1013	
Bélier VS témoin	-0.04945	0.06428	11	-0.77	0.4579	
Sous-soleuse conventionnelle VS témoin	0.04903	0.06345	10.6	0.77	0.4566	
Bélier parallèle VS bélier perpendiculaire	-0.07916	0.07720	9.36	-1.03	0.3310	
Sous-sol parallèle VS Sous-sol perpendiculaire	-0.04623	0.07528	8.54	-0.61	0.5551	
Interaction (bélier par - bélier per) VS (ss par - ss per)	-0.03293	0.1078	8.94	-0.31	0.7671	

Les moyennes des moindres carrés sont présentées ici à titre indicatif, mais ne sont pas statistiquement différentes. Le traitement témoin produit des valeurs intermédiaires de hauteurs de plants par rapport aux traitements de sous-solage

Moyennes des moindres carrés (pour dist_drain moyenne=3.44m)						
Effet	Trait	Estimation	Erreur type	DDL	Valeur du test t	Pr > t
trait	bélier_par	1.2240	0.06220	15	19.68	<.0001
trait	bélier_per	1.3031	0.05342	8.22	24.40	<.0001
trait	ss_par	1.3389	0.06020	13.3	22.24	<.0001

Moyennes des moindres carrés (pour dist_drain moyenne=3.44m)						
Effet	Trait	Estimation	Erreur type	DDL	Valeur du test t	Pr > t
trait	ss_per	1.3852	0.05293	7.96	26.17	<.0001
trait	tem_pas	1.3130	0.05464	11.6	24.03	<.0001

Hauteur des plants de maïs en fonction des traitements et de la distance à la raie (dispositif réduit)

Sur le dispositif réduit, les Tests de F des effets de type 3 ne révèlent pas d'effet significatif des traitements (Prob. F = 0.9923) sur la hauteur de plants de maïs, mais un effet quadratique très significatif de la distance à la raie (Prob. F <0.0001), de même qu'un effet linéaire significatif de cette distance à la raie (Prob. F <0.0199).

Tests des effets fixes de type 3				
Effet	DDL num.	DDL den.	Valeur F	Pr > F
Trait	4	9.5	0.06	0.9923
Raie_dist	1	182	5.51	0.0199
Raie_dist2	1	182	26.66	<.0001

Les effets linéaire et quadratique de la covariable distance à la raie donne une augmentation positive et exponentielle de la hauteur des plants en fonction de la distance à la raie. Les effets de chaque traitement de sous-solage ne sont pas significativement différents du témoin.

Solution pour effets fixes						
Effet	Trait	Estimation	Erreur type	DDL	Valeur du test t	Pr > t
Intercept		0.8260	0.1078	34.8	7.67	<.0001
trait	bélier_par	-0.03336	0.1121	10.4	-0.30	0.7718
trait	bélier_per	-0.01848	0.1092	9.35	-0.17	0.8693
trait	ss_par	-0.03272	0.1121	10.4	-0.29	0.7760
trait	ss_per	0.01072	0.1092	9.34	0.10	0.9239
trait	tem_pas	0
Raie_dist		-0.05676	0.02417	182	-2.35	0.0199
Raie_dist2		0.009005	0.001744	182	5.16	<.0001

Les comparaisons simples des différents traitements de sous-solage incluant le témoin ne révèlent pas d'effet significatif des traitements.

Différences des moyennes des moindres carrés							
Effet	trait	_trait	Estimation	Erreur type	DDL	Valeur du test t	Pr > t
trait	bull_par	bull_per	-0.01488	0.1157	9.32	-0.13	0.9004
trait	bull_par	ss_par	-0.00063	0.1183	10.2	-0.01	0.9958
trait	bull_par	ss_per	-0.04407	0.1156	9.31	-0.38	0.7117
trait	bull_par	tem_pas	-0.03336	0.1121	10.4	-0.30	0.7718
trait	bull_per	ss_par	0.01425	0.1157	9.31	0.12	0.9046
trait	bull_per	ss_per	-0.02919	0.1129	8.46	-0.26	0.8022
trait	bull_per	tem_pas	-0.01848	0.1092	9.35	-0.17	0.8693
trait	ss_par	ss_per	-0.04344	0.1157	9.31	-0.38	0.7156
trait	ss_par	tem_pas	-0.03272	0.1121	10.4	-0.29	0.7760
trait	ss_per	tem_pas	0.01072	0.1092	9.34	0.10	0.9239

Hauteur des plants de maïs en fonction des traitements et de la distance au drain (dispositif réduit)

Sur le dispositif réduit, les Tests de F des effets de type 3 ne révèlent pas d'effet significatif des traitements (Prob. F = 0.9973) sur la hauteur de plants de maïs, mais un effet linéaire très significatif de la distance au drain (Prob. F = 0.0004).

Tests des effets fixes de type 3				
Effet	DDL num.	DDL den.	Valeur F	Pr > F
trait	4	9.88	0.03	0.9973
DrainST_dist	1	184	13.11	0.0004

L'effet de la covariable distance au drain est linéaire, la hauteur des plants de maïs diminue de 5 cm à chaque m de distance au drain. Les effets de chaque traitement de sous-solage ne sont pas significativement différents du témoin.

Solution pour effets fixes						
Effet	Trait	Estimation	Erreur type	DDL	Valeur du test t	Pr > t
Intercept		1.1251	0.1065	19	10.56	<.0001
trait	bélier_par	-0.01936	0.1209	11.6	-0.16	0.8755
trait	bélier_per	-0.01332	0.1153	9.54	-0.12	0.9104
trait	ss_par	-0.02295	0.1209	11.6	-0.19	0.8527
trait	ss_per	0.01565	0.1153	9.54	0.14	0.8949
trait	tem_pas	0
DrainST_dist		-0.04678	0.01292	184	-3.62	0.0004

Les comparaisons simples des différents traitements de sous-solage incluant le témoin ne révèlent pas d'effet significatif des traitements.

Différences des moyennes des moindres carrés							
Effet	trait	_trait	Estimation	Erreur type	DDL	Valeur du test t	Pr > t
trait	bull_par	bull_per	-0.00604	0.1226	9.66	-0.05	0.9617
trait	bull_par	ss_par	0.003594	0.1279	11.4	0.03	0.9781
trait	bull_par	ss_per	-0.03501	0.1226	9.66	-0.29	0.7813
trait	bull_par	tem_pas	-0.01936	0.1209	11.6	-0.16	0.8755
trait	bull_per	ss_par	0.009636	0.1226	9.66	0.08	0.9390
trait	bull_per	ss_per	-0.02896	0.1171	8.03	-0.25	0.8108
trait	bull_per	tem_pas	-0.01332	0.1153	9.54	-0.12	0.9104
trait	ss_par	ss_per	-0.03860	0.1226	9.66	-0.31	0.7596
trait	ss_par	tem_pas	-0.02295	0.1209	11.6	-0.19	0.8527
trait	ss_per	tem_pas	0.01565	0.1153	9.54	0.14	0.8949

Profondeur de la nappe en fonction des traitements et de la distance au drain (dispositif réduit)

Sur le dispositif réduit, les Tests de F des effets de type 3 ne révèlent pas d'effet significatif des traitements (Prob. F = 0.8467) sur la profondeur de la nappe, mais un effet quadratique très significatif de la distance au drain (Prob. F = 0.0004).

Tests des effets fixes de type 3				
Effet	DDL num.	DDL den.	Valeur F	Pr > F
trait	4	7.83	0.34	0.8467
DrainST_dist	1	164	9.95	0.0019

L'effet de la covariable distance au drain est linéaire, la profondeur de la nappe diminue de 0.65 cm à chaque m de distance au drain, donc débute à 21 cm de la surface du sol sur le drain et se rapproche de la surface en s'éloignant du drain. Les effets de chaque traitement de sous-solage ne sont pas significativement différents du témoin.

Solution pour effets fixes						
Effet	Trait	Estimation	Erreur type	DDL	Valeur du test t	Pr > t
Intercept		20.9561	1.9140	6.73	10.95	<.0001
trait	bull_par	-0.9069	1.6310	9.9	-0.56	0.5905
trait	bull_per	0.9803	1.5097	7.28	0.65	0.5361
trait	ss_par	0.06236	1.6015	9.24	0.04	0.9698
trait	ss_per	0.03023	1.5164	7.4	0.02	0.9846
trait	tem_pas	0
DrainST_dist		-0.6466	0.2050	164	-3.15	0.0019

Les comparaisons simples des différents traitements de sous-solage incluant le témoin ne révèlent pas d'effet significatif des traitements.

Différences des moyennes des moindres carrés								
Effet	trait	_trait	DrainST_dist	Estimation	Erreur type	DDL	Valeur du test t	Pr > t
trait	bull_par	bull_per	4.02	-1.8871	1.6405	7.98	-1.15	0.2833
trait	bull_par	ss_par	4.02	-0.9692	1.7253	9.76	-0.56	0.5869
trait	bull_par	ss_per	4.02	-0.9371	1.6458	8.08	-0.57	0.5846
trait	bull_par	tem_pas	4.02	-0.9069	1.6310	9.9	-0.56	0.5905
trait	bull_per	ss_par	4.02	0.9179	1.6140	7.5	0.57	0.5862
trait	bull_per	ss_per	4.02	0.9500	1.5279	6	0.62	0.5569
trait	bull_per	tem_pas	4.02	0.9803	1.5097	7.28	0.65	0.5361
trait	ss_par	ss_per	4.02	0.03213	1.6193	7.59	0.02	0.9847

Différences des moyennes des moindres carrés								
Effet	trait	_trait	DrainST_dist	Estimation	Erreur type	DDL	Valeur du test t	Pr > t
trait	ss_par	tem_pas	4.02	0.06236	1.6015	9.24	0.04	0.9698
trait	ss_per	tem_pas	4.02	0.03023	1.5164	7.4	0.02	0.9846

Conclusion générale et perspectives

La présente étude avait pour objectif de valider les meilleures techniques à adopter en matière de sous-solage des sols peu perméables en profondeur. Pour ce faire, l'efficacité d'une sous-soleuse montée sur béliet a été comparée à une sous-soleuse conventionnelle montée sur tracteur et les deux ont été opérées en modes parallèles ou perpendiculaires aux drains pour améliorer les conditions de drainage et la productivité des sols par rapport à la parcelle témoin. Les forêts d'arbres décisionnelles et la régression linéaire multiple ont été utilisées et comparées pour détecter les effets du sous-solage et des systèmes de drainage souterrain et de surface sur la hauteur des plants et la profondeur moyenne de la nappe durant la saison de croissance.

Les essais de sous-solage réalisés dans cette étude notamment le sens du sous-solage n'a pas eu un effet très marqué sur la hauteur des plants et la profondeur moyenne de la nappe. La nature peu perméable du sous-sol rend les conditions optimales pour le sous-solage peu fréquent. Les sols demeurent souvent trop humides en profondeur pour se fissurer adéquatement sous l'effet du sous-solage. Ces essais de sous-solage ont été réalisés le 17 et 18 septembre 2014 à une période où le sol était encore à un état assez sec.

Le type de sous-soleuse ne semble pas non plus avoir eu d'effet important sur les performances de l'opération. Le béliet permettait juste de maintenir les étançons à des profondeurs plus constantes que les sous-soleuses conventionnelles opérées avec des tracteurs, mais dans les deux cas les profondeurs atteintes pouvaient être supérieures à 60 cm.

Les FAD et la RLM permettent tous les deux de détecter les principaux facteurs affectant la Plant_hauteur et la Nappe_moy. Sauf que les FAD ont permis de faire ressortir les relations non-linéaires entre la Nappe_moy, Plant_hauteur et les variables indépendantes à l'étude, notamment la position sur la planche.

Sur le dispositif complet en régression linéaire multiple (RLM), la hauteur des plants dépend principalement de la position sur la planche symbolisée par DrainST_prof et de la distance aux drains (DrainST_dist). Tandis qu'en FAD, la direction perpendiculaire à la planche (Dir_Y) s'inscrit en deuxième position entre DrainST_prof et DrainST_dist.

Sur le dispositif réduit, Plant_hauteur dépend principalement de Nappe_moy, Raie_dist, et DrainST_prof en RLM, tandis qu'en FAD, Raie_dist, Nappe_moy et la forme du terrain en planches arrondies affectent le plus Nappe_moy. Le bas de la planche est l'endroit sur le dispositif où la nappe est la plus haute.

Sur le dispositif réduit, Nappe_moy dépend principalement de Dir_Y et Raie_dist en RLM, tandis qu'en FAD c'est la forme du terrain en planches arrondies ainsi que la raie de curage qui influencent Nappe_moy. Nappe_moy est affectée de façon négative proche de la raie, alors que loin de la raie, Nappe_moy augmente.

Ces résultats mettent en évidence l'utilité des FAD pour détecter les effets de différents facteurs liés à l'aménagement du terrain et à une pratique agricole, et les effets du sous-solage, sur les cultures. C'est pourquoi la modification des pratiques agricoles par un contrôle du trafic et de la circulation des tracteurs, épandeurs, moissonneuses et chars à grain serait donc essentielle pour prolonger le bénéfice des opérations de sous-solage (Raper et Bergtold, 2007). Lorsqu'aucun contrôle de la circulation de la machinerie n'est exercé dans le champ après des opérations de sous-solage, le sol reviendrait rapidement à son état compact initial. Par ailleurs, pour être efficace, le sous-solage devrait être suivi de cultures pouvant bénéficier de l'aération du sous-sol pour l'investir de leurs racines et produire un maximum de biomasse pour alimenter la méso et macrofaune et favoriser la structuration du sol. L'allègement de la machinerie est aussi fortement souhaitée. Les FAD et la RLM ont des performances plus ou moins similaires selon l'ordre d'importance relative des variables à la seule différence que les FAD sont plus puissants en matière d'interpolation des données et à faire ressortir les principales corrélations et les effets non linéaires entre les variables

Les présents résultats portent à la recommandation suivante : les distances aux drains critiques (> 2 m) et optimaux (> 4 m), la profondeur aux drains souterrains optimale (1.4 à 1.7m), la distance optimale à la raie de curage (> 8 m) et la profondeur moyenne critique de la nappe (< 0.25 m). Une amélioration de la productivité en termes de hauteur des plants se fait ressentir dès que la distance verticale aux drains souterrains dépasse les 1.1 m, soit sur le dessus de la planche. La distance horizontale aux drains souterrains est principalement déterminée par le dénivelé sur la planche qui est aussi exprimée par la direction perpendiculaire au champ et aux systèmes de drainage. La distance horizontale aux drains souterrains a influencé à la baisse la

hauteur des plants et de façon marquée à partir d'une distance de 2 à 3 m des drains et qu'un écartement des drains d'environ 6 à 7 m favoriserait une maximisation de la hauteur des plants.

Comme perspective, les futurs travaux devraient explorer dans quelle mesure la planche arrondie pourrait être remodelée, étudier aussi une éventuelle possibilité d'intensification du système de drainage et enfin prospecter la possibilité d'explorer le lien entre le rendement du maïs et les bandes multi spectrales provenant des images utilisées pour établir la hauteur des plants de maïs.