



Caractérisation des sols organiques cultivés du sud-ouest de Montréal : Formation de groupes de gestion des sols à des fins de conservation

Mémoire

Raphaël Deragon

Maîtrise en sols et environnement - avec mémoire
Maître ès sciences (M. Sc.)

Québec, Canada

Caractérisation des sols organiques cultivés du sud-ouest de Montréal

Formation de groupes de gestion des sols à des fins de conservation

Mémoire

Raphaël Deragon

Sous la direction de :

Jean Caron, directeur de recherche

Résumé

Les sols organiques cultivés des basses-terres du Saint-Laurent sont parmi les sols maraîchers les plus productifs de la province et fournissent les aliments de base d'une grande partie de la population québécoise et du nord-est des États-Unis. Par leur mise en culture, ces sols sont exposés à différents processus de dégradation et de perte de hauteur de sol. Par conséquent, un important plan de conservation vise les sols dégradés et ceux à risque. Ce mémoire a pour objectifs de faire un bilan des propriétés physiques, chimiques et pédologiques de cinq tourbières de la région en plus de définir et de former des groupes de gestion des sols à des fins de conservation. L'hypothèse de départ supposait l'existence de groupes de sols liés à la qualité agronomique ou au grand groupe pédologique de 120 sites échantillonnés en 2019. Au terme de l'étude, une MANOVA a permis de conclure que la profondeur cultivable était une propriété clé pouvant classer les sites en fonction de leur dégradation. Les sols ayant une profondeur cultivable de moins de 60 cm montraient des comportements significativement différents, résultats supportés par 17 propriétés discriminantes entre les groupes. Un second seuil pratique de 100 cm, celui-ci étant relié au drainage, a aussi été retenu. Ensuite, un second volet axé sur la cartographie numérique a permis de générer des cartes prédictives à l'échelle régionale de la profondeur au sol minéral, de l'épaisseur de la couche de coprogène et de la profondeur de sol cultivable en combinant les deux premières. Les données cumulées permettront de suivre l'évolution des sols dans le temps, alors que les cartes serviront d'outils complémentaires d'aide à la décision afin de cibler des zones d'intervention prioritaires pour conserver les ressources en sol de la région.

Abstract

Organic soils converted to agricultural fields in the southwest plain of Montreal are amongst the most productive soils for horticulture in Quebec, not only supplying food for the province but also to the north-eastern part of the United-States. After their initial drainage to allow agriculture, Organic soils are susceptible to many forms of degradation and soil loss. Therefore, an important soil conservation plan is targeting degraded soils and soil at risk. The objectives of this master's thesis are to characterize soil physical, chemical and pedological properties of five peatlands in addition to forming soil conservation management groups. The initial hypothesis supposed the existence of latent groups related to soil quality or pedological Great Groups of 120 sampling sites visited in 2019. As a result of that work, a MANOVA lead to the conclusion that the peat maximum thickness of cultivation is a key property that is related to soil degradation. Soils with a peat maximum thickness of cultivation lower than 60 cm showed significantly more degradation signs. These results are supported by 17 discriminant properties between the groups. A second critical threshold of 100 cm was also established. The second part of this thesis used digital soil mapping as a tool to generate predictive maps at a regional scale of the depth to the mineral layer, coprogenous layer thickness and peat maximum thickness of cultivation by subtracting the first two maps. The rich dataset formed in this study will be used to monitor soil evolution while the maps will serve as complementary decision tools to target fields for a prioritized soil conservation intervention.

Table des matières

Résumé	ii
Abstract.....	iii
Table des matières	iv
Liste des figures.....	vii
Liste des tableaux.....	xi
Remerciements.....	xiii
Avant-propos	xiv
Introduction	1
Chapitre 1 – Revue de littérature	2
1.1 Les sols organiques.....	2
1.1.1 La formation des sols organiques	2
1.1.2 L'évolution des sols organiques cultivés	3
1.2 Caractérisation des sols	5
1.2.1 Pédologie	5
1.2.2 Physique et hydrologie des sols organiques	8
1.2.3 Chimie des sols.....	15
1.3 En terrain connu.....	18
1.4 Propriétés discriminantes de la qualité du sol et classes de gestion.....	22
1.5 La cartographie des sols	25
1.6 Hypothèses et objectifs	28
Bibliographie	31
Chapitre 2 – Groupes de gestion basés sur la profondeur de sol cultivable	35
2.1 Introduction.....	35
2.2 Matériel et méthode.....	35
2.2.1 Sélection des sites	35
2.2.2 Caractéristiques pédologiques	37
2.2.3 Analyses physiques	38
2.2.4 Analyses chimiques	40
2.2.5 Statistiques et analyses cartographiques.....	41
2.3 Résultats	42
2.3.1 Analyse réursive de l'association entre la qualité et les propriétés des sols	42

2.3.2	Analyse réursive de l'association entre le grand groupe et les propriétés des sols.....	45
2.3.3	Approche par modèle de mélange	49
2.3.4	Approche par profondeur de sol cultivable.....	49
2.4	Discussion	61
2.5	Conclusion.....	63
	Bibliographie.....	64
	Chapitre 3 - Mapping cultivated Organic soils in the southwest plain of Montreal: Determining peat maximum thickness of cultivation at a regional scale.....	66
3.1	Résumé	67
3.2	Abstract	67
3.3	Introduction.....	68
3.4	Methodology	70
3.4.1	Study area.....	70
3.4.2	Soil data	70
3.4.3	Environmental covariates.....	72
3.4.4	Covariate reduction	75
3.4.5	Modeling approaches.....	76
3.4.6	Model tuning and validation	78
3.4.7	Prediction uncertainty.....	80
3.5	Results	81
3.5.1	Covariate reduction	81
3.5.2	LOBOCV block size	81
3.5.3	Accuracy assessment and model selection	83
3.5.4	Covariates importance in the final models	87
3.5.5	Regional peat maximum thickness of cultivation map and its uncertainty.....	88
3.6	Discussion	91
3.6.1	Relevant covariates to predict coprogenous layer thickness and depth to the mineral layer	91
3.6.2	Peat maximum thickness of cultivation as a tool to guide soil conservation practices and its uncertainty.....	92
3.7	Conclusions	92
	Acknowledgements.....	93
	References	94
	Conclusion générale	99
	Bibliographie générale	100

Annexe 1 - Description des variables de l'étude.....	103
Annexe 2 - Statistiques descriptives des variables continues retenues dans les analyses	104
Annexe 3 – Autres variables significatives à $\alpha = 0,05$ selon la série d'ANOVAs.....	105

Liste des figures

- Figure 1 - Profondeur de sol organique estimée à l'aide d'une a) induction électromagnétique suivie d'une inversion 1D, et d'un b) géoradar combiné à un sondage manuel (Airo et al., 2014). Le titre de la figure a été traduit. Figure sous licence : Creative Commons CC-BY 4.0. 7
- Figure 2 - Conductivité hydraulique saturée (10^{-3} cm sec⁻¹) en fonction du degré d'humification de trois types de composition botanique. L'échelle de von Post est utilisée afin d'évaluer la décomposition du sol (1 = peu décomposé à 10 = très décomposé). Reproduit à partir de « Wetland Soils : Genesis, Hydrology, Landscapes, and Classification » par Vepraskas, M.J. & Craft, C.B. © 2000, p.351, CRC Press. Reproduit avec la permission de Taylor and Francis Group, LLC, une division d'Informa plc. 10
- Figure 3 - Conductivité hydraulique saturée (cm jour⁻¹) en fonction a) de la masse volumique apparente (g cm⁻³) et b) du volume de solide (%) de sols organiques selon plusieurs études. Reproduit à partir de « Organic soils and peat materials for sustainable agriculture » par Parent L.E. & Ilnicki, P. © 2002, p.55, CRC Press. Reproduit avec la permission de Taylor and Francis Group, LLC, une division d'Informa plc. 11
- Figure 4 - Courbes de rétention a) d'une tourbe de sphaigne, b) de joncs et de roseaux (Päivänen, 1973), et c) des trois structures de la couche de Moorsh, soit tourbeuse (peaty), humique (humic) et granulaire (grainy) (adapté de Brandyk et al., 2003). Le titre des figures a) et b) a été traduit. Figures sous licence : Creative Commons CC BY-SA 4.0. Figure c) reproduite à partir de « Organic soils and peat materials for sustainable agriculture » par Parent L.E. & Ilnicki, P. © 2002, p.50, CRC Press. Reproduit avec la permission de Taylor and Francis Group, LLC, une division d'Informa plc. 13
- Figure 5 - Porosité totale a) et macroporosité (diamètre équivalent de 50 µm) b) de différentes compositions botaniques en fonction de la masse volumique apparente (Liu & Lennartz, 2019). On y retrouve la sphaigne (sphagnum), le bois (woody), la laïche (sedge) et d'autres origines tourbeuses (others). Copyright © 1969 Reproduit avec la permission de John Wiley and Sons, Inc. 14
- Figure 6 - Nuage de points de la teneur en azote total (TN) en fonction de la teneur en carbone organique total (TOC) pour différents régimes hydriques (Anshari et al., 2010). Le titre de la figure a été traduit. Figure sous licence : Creative Commons CC-BY 4.0. 17
- Figure 7 - Profils de teneur en cendres avec datation au carbone pour un site a) peu affecté par l'homme, b) utilisé extensivement en prairie et c) utilisé intensivement en prairie (Kruger et al., 2015). À chaque site, trois sondages ont eu lieu afin d'obtenir des répliques (ex. : NW1, NW2 et NW3). Le titre de la figure a été traduit. Figure sous licence : Creative Commons CC-BY 4.0. 17
- Figure 8 - Résistance à l'enfoncement mesurée dans trois sites au degré de décomposition différent : a) Profils de résistance à l'enfoncement et b) valeur maximale enregistrée par tranche de 10 cm (Dessureault-Rompre et al., 2018). Différentes lettres signifient une différence significative entre les types de sol selon le test de Tukey DSH. Le titre de la figure a été traduit. Figure sous licence : Creative Common CC-BY-NC-ND 4.0. ... 19

Figure 9 - Courbe de rétention moyenne basé sur la valeur médiane des paramètres calibrés selon le modèle de van Genuchten pour le groupe 1 a) et le groupe 2 b) (Hallema et al., 2015b). Copyright © 1969 Reproduit avec la permission de John Wiley and Sons, Inc.....	20
Figure 10 – Schéma de l’approche conceptuelle utilisée reprenant les hypothèses de départ représentées par trois étapes successives. Les étapes 1 et 2 se retrouvent dans le chapitre 2, alors que l’étape 3 sera abordée au chapitre 3.	30
Figure 11 - Résumé des opérations et des types de prélèvements effectués chez 14 fermes partenaires à l’été 2019.....	35
Figure 12 - Pédons réalisés en Montérégie-Ouest à l’été 2019. Les points en orange représentent les 120 sites. Crédit d’image de fond : Google Satellite.	37
Figure 13 - Arbre récursif issu de la classification de 114 sites en fonction de leur qualité, soit 71 sites de bonne qualité et 43 de mauvaise qualité. Chaque branche mène à une feuille contenant un certain nombre de sites (n). Les bons sols sont représentés par la couleur grise et les mauvais sols par la couleur noire. À droite des feuilles se trouve une échelle permettant d’évaluer la proportion des sites appartenant à chacune des qualités. Une feuille d’une seule couleur indique une classification parfaite alors qu’un mélange de couleurs indique qu’une proportion des sites a été mal classée. On peut y voir cinq variables, soit le pourcentage massique de matière organique entre 40-60 cm (MO_3), le pourcentage massique de sol de surface de diamètre inférieur à 0,84 mm (P_inf_0.84mm_1), la résistance à l’enfoncement moyenne (RE_Moy), le pourcentage volumique pondéré de fibres non frottées (FIB) et la conductivité hydraulique saturée à 0-5 cm (ksat_1).....	43
Figure 14 - Arbre récursif issu de la classification de 80 sites en fonction de leur qualité. On peut y voir trois variables, soit la profondeur de sol cultivable (Cult), le pourcentage massique de sol de surface de diamètre inférieur à 0,84 mm (P_inf_0.84mm_1) et la conductivité électrique de la couche 0-20 cm (CE_1).....	44
Figure 15 - Arbre récursif issu de la classification de 113 sites en fonction de leur grand groupe, soit 75 humisols et 38 mésisols. On peut y voir neuf variables, soit le pourcentage de fibres frottées moyen du profil pondéré par l’épaisseur de la couche (FIB_frot), l’indice au pyrophosphate de la couche 40-60 cm (pyro_3), la conductivité hydraulique saturée à 0-5 cm et à 50-55 cm (ksat_1 et ksat_3), le pH de la couche 0-20 cm et 40-60 cm (pH_1 et pH_3), la teneur en eau volumique à 3 kPa à 0-5 cm (Tee.3kPa_1), l’épaisseur de coprogène présent dans le premier 1,6 m de sol (Copro_160) et le contenu en fragments grossiers de la couche qui en contient le plus dans le premier mètre de sol (FRGT_max).....	48
Figure 16 – Relations exploratoires de type exponentielle inverse entre quatre propriétés et la profondeur cultivable.....	50
Figure 17 - Relations exploratoires de type fonction inverse entre deux propriétés et la profondeur cultivable.....	50

Figure 18 – Diagrammes en violon du contenu en matière organique en fonction du groupe de profondeur cultivable pour trois profondeurs d'échantillonnage : a) 0 à 20 cm de profondeur, b) 20 à 40 cm et c) 40 à 60 cm.....	56
Figure 19 - Diagrammes en violon de la masse volumique apparente en fonction du groupe de profondeur cultivable pour deux profondeurs d'échantillonnage : a) 0 à 5 cm de profondeur et b) 50 à 55 cm.....	56
Figure 20 - Diagrammes en violon de la conductivité électrique en fonction du groupe de profondeur cultivable pour deux profondeurs d'échantillonnage : a) 20 à 40 cm de profondeur et b) 40 à 60 cm.....	57
Figure 21 - Diagrammes en violon a) de la teneur en air à 5 kPa et b) de la réserve en eau utilisable à 10 kPa en fonction du groupe de profondeur cultivable selon un échantillonnage à 50-55 cm.	57
Figure 22 - Diagrammes en violon a) du pourcentage de fibres pondéré et b) du pourcentage de fibres frottées pondéré en fonction du groupe de profondeur cultivable.....	58
Figure 23 - Diagrammes en violon du ratio carbone sur azote en fonction du groupe de profondeur cultivable pour trois profondeurs d'échantillonnage : a) 0 à 20 cm de profondeur, b) 20 à 40 cm et c) 40 à 60 cm.....	58
Figure 24 - Diagrammes en violon de la teneur en carbone total en fonction du groupe de profondeur cultivable pour trois profondeurs d'échantillonnage : a) 0 à 20 cm de profondeur, b) 20 à 40 cm et c) 40 à 60 cm.....	59
Figure 25 – Carte régionale des groupes de profondeur de sol cultivable à des fins de conservation des sols.	60
Figure 26 – Study area and sampling sites categorized by the nature of the information collected. Overlaid is a square grid of 1.5 km size used to spatially partition the datasets to allow leave-one-block-out spatial cross-validation.....	71
Figure 27 - Experimental variograms (dots) of the two predicted features and their best fit model (line). (A) Variogram of the coprogenous layer thickness using 255 square-root-transformed observations. (B) Variogram of the depth to the mineral layer using 4,488 observations.....	81
Figure 28 - Experimental residuals variogram (dots) of the two predicted features and their best fit model (line). (A) Variogram of the coprogenous layer thickness residuals using square-root-transformed observations. (B) Variogram of the depth to the mineral layer residuals. Both prediction residuals are from Cubist models.	84
Figure 29 – Map of the predicted coprogenous layer thickness (cm) across the study area using the Cubist model.....	85
Figure 30 – Depth to the mineral layer (cm) predicted by four models: a) Cubist, b) Random Forest and c) k-Nearest Neighbor.....	86

Figure 31- Relative importance of each covariate (%) on the final model A) of the coprogenous layer thickness and B) of the depth to the mineral layer. Refer to Table 14 for a description of covariates' abbreviations. 87

Figure 32 - Peat maximum thickness of cultivation map obtained by subtracting the coprogenous layer thickness to the depth to the mineral layer..... 89

Figure 33 - Peat maximum thickness of cultivation uncertainty maps. A) Upper and B) Lower bounds of the 90 % prediction interval..... 90

Liste des tableaux

Tableau 1 - Propriétés physiques de tourbes fenniques en fonction de leur degré de décomposition (Okruszko & Ilnicki, 2003). R1 = fibrique (Of), R2 = mésique (Om) et R3 = humique (Oh). Source originale : Okruszko, H. & Szymanowski, M. 1992. Proc. 9th Int. Peat Congr., Uppsala, Sweden, 3 :106-115. Reproduit avec la permission de la International Peatland Society.	12
Tableau 2 - Porosité moyenne d'une tourbe fibrique (R1) de mousse et de laîche, d'une tourbe d'aulne humique (R3) et des structures de la couche de Moorsh de sols de la Pologne (Ilnicki & Zeitz, 2003).....	13
Tableau 3 - Analyses chimiques pour deux couches de sol en fonction de la distance des sites avec la rivière (Pawluczuk et al., 2019). Le titre du tableau a été traduit. Tableau sous licence : Creative Common CC-BY-NC-SA 4.0.....	16
Tableau 4 - Valeur médiane des paramètres en fonction de leur groupe issu d'un partitionnement hiérarchique (adapté de Hallema et al., 2015b). Copyright © 1969 Reproduit avec la permission de John Wiley and Sons, Inc.	20
Tableau 5 - Statistiques descriptives de différentes propriétés : MVA (pb), MVR(pp), porosité (n), teneur en MO (fo), profondeur de sol organique (Do), profondeur de la couche compacte (zcl), la Ksat (Ks), la réserve en eau utilisable (SWAC) pour la laitue et l'année de mise en culture (Reclam. yr) (Hallema et al., 2015a). Copyright © 1969 Reproduit avec la permission de John Wiley and Sons, Inc.	22
Tableau 6 - Détails du nombre de sites d'échantillonnage par ferme ainsi que de leur qualité	36
Tableau 7 - Propriétés pédologiques étudiées lors de la réalisation du profil de sol	38
Tableau 8 - Série de MANOVAs avec un seuil incrémenté par tranches de 10 cm. Le tableau rapporte aussi le nombre d'observations dans chaque groupe ($N1/N2$), la trace de Pillai ainsi que les bornes de son intervalle de confiance de 95% et le nombre d'ANOVAs significatives sur 53 entre les groupes au seuil testé.	52
Tableau 9 - Série de MANOVAs avec deux seuils, le premier fixé à 60 cm et le second incrémenté par tranches de 10 cm. Le tableau rapporte aussi le nombre d'observations dans chaque groupe ($N1/N2$), la trace de Pillai ainsi que les bornes de son intervalle de confiance de 95% et le nombre d'ANOVAs significatives sur 53 entre les groupes au seuil testé.	54
Tableau 10 – Moyenne (Moy) et erreur-type (SE) des 17 variables significativement différentes entre le groupe 1 (< 60 cm) et le groupe 2 (> 60 cm).	55
Tableau 11 – Superficie prédites des groupes de profondeurs à l'échelle régionale.....	60

Table 12 - Summary statistics of the coprogenous layer thickness (CLT) and depth to the mineral layer (DML) 72

Table 13 – Bands and indices from Landsat 8 and Sentinel 2 satellites initially used as covariates. 75

Table 14 - Final selection of covariates for the two models after stepwise variance inflation factor and recursive feature selection 82

Table 15 – Tuning and leave-one-block-out cross validation results for the machine learners. Results for the final model of the coprogenous layer thickness (CLT) and for the three machine learners for the depth to the mineral layer (DML) are shown. Standard deviation of Lin’s concordance correlation coefficient (CCC) and of root mean square of error (RMSE) are in parentheses. RMSE, observed and predicted range are in centimeter. 83

Remerciements

Je tiens tout d'abord à remercier le Dr. Jean Caron que j'ai eu la chance de côtoyer dès ma première année de baccalauréat. Durant mon parcours, j'ai pu bénéficier de ses précieux conseils et de son encadrement à plusieurs niveaux. Il a su croire en mon potentiel et m'a poussé à le développer en m'intégrant au monde de la recherche en sciences du sol. Je lui en serai toujours reconnaissant.

Je ne peux passer sous silence la contribution de Michaël Leblanc qui m'a transmis sa passion pour la pédologie et sans qui le troisième chapitre de ce mémoire aurait été impossible. C'est aussi grâce à lui que j'ai pu rencontrer Lucie Grenon, Daniel D. Saurette et Brandon Heung, que je remercie d'ailleurs pour leur patience et le partage de leurs connaissances. Je remercie aussi Anne-Sophie Julien pour ses conseils en statistiques.

Évidemment, je tiens à souligner le travail de Diane Bulot, Geneviève Montmigny, Alexandre McCutcheon, Jacynthe Dessureault-Rompré et Christophe Libbrecht. Ces derniers ont grandement contribué à la gestion du travail au terrain, au traitement des échantillons en laboratoire, en plus d'avoir participé aux nombreuses discussions concernant l'analyse des données. Le projet était ambitieux et c'est grâce à leur support qu'il est maintenant achevé. Dans la même veine, je tiens à remercier tous les auxiliaires de recherche et les étudiants gradués de l'équipe qui m'ont aidé de près ou de loin avec mon projet.

Un grand merci à Antoine Karam qui a su me guider dans mes démarches tout au long de mon parcours afin de me faciliter les choses. J'ai grandement apprécié nos échanges et sa dévotion pour les étudiants du département.

Je tiens aussi à remercier ma famille et mes amis qui m'ont encouragé à me dépasser et qui ont su me motiver tout au long de mon parcours. Je me compte chanceux de vous avoir. Merci pour votre support.

En terminant, je tiens à remercier le CRSNG, le FRQNT ainsi que les 14 fermes partenaires du projet, soit Les Fermes Hotte & Van Winden Inc., Les Fermes R.R. et fils Inc., Maraîchers J.P.L. Guérin & Fils Inc., Production Horticole Van Winden Inc., Productions Maraîchères Breizh Inc., R. Pinsonneault et fils Ltée, Vert Nature Inc., Delfland Inc, Patates Isabelle Inc., La Production Barry Inc., Le Potager Montréalais Ltée, Le Potager Riendeau Inc., Les Jardins A. Guérin et fils Inc. et Les Fermes du Soleil Inc. pour leur support financier et leur participation active au projet de recherche. Ce fut un plaisir de collaborer avec eux et d'apprendre à connaître ces agriculteurs de ma région natale.

Avant-propos

Ce mémoire est constitué de trois chapitres. Le premier est une revue de littérature portant sur la dégradation des sols organiques cultivés et l'évolution de leurs propriétés. Le chapitre 2 traite de la formation de groupes de gestion des sols organiques cultivés afin de mieux orienter leur conservation. Ce chapitre sera par la suite traduit en anglais et soumis pour publication dans une édition spéciale du *Canadian Journal of Soil Science* portant sur l'évaluation de la santé des sols. Le chapitre 3 constitue un article scientifique rédigé en anglais qui sera soumis pour publication en décembre 2021 dans une édition spéciale du *Canadian Journal of Soil Science* portant sur la pédologie et la classification des sols.

Cet article s'intitule « *Mapping cultivated Organic soils in the southwest plain of Montreal: Determining peat maximum thickness of cultivation at a regional scale* ». Il suit la séquence logique du mémoire en se fondant sur l'importance de la profondeur cultivable des sols afin de cartographier cette propriété à l'échelle régionale. L'objectif est de cibler des zones moins profondes afin d'orienter les interventions de conservation et offrir aux producteurs des cartes à jour de la profondeur des sols.

J'en serai le premier auteur, suivi de Daniel D. Saurette, MSc. (Ontario Ministry of Agriculture Food and Rural Affairs, OMAFRA), de Brandon Heung, PhD. (Faculty of Agriculture, Department of Plant, Food, and Environmental Sciences, Dalhousie University) et de Jean Caron, PhD. (Département des sols et de génie agroalimentaire, Université Laval).

J'ai réalisé la prise de données sur le terrain, puis réalisé le traitement des données. Ce traitement a été dirigé par M. Saurette et M. Heung, qui ont aussi fourni certains scripts R et autres ressources. Ils ont aussi participé à la révision de l'article. M. Caron a élaboré la problématique de départ, a fourni l'équipement pour la prise des données, le financement et a aussi participé à la révision de l'article.

La version finale de l'article peut différer de celle présentée dans ce mémoire.

Introduction

Les sols organiques cultivés résultent d'un drainage anthropique de tourbières ou de marécages. Ces sols sont prisés pour leur fertilité et leur rétention en eau élevées, permettant des rendements importants au plan maraîcher. Ne couvrant que 4 % du territoire du Québec méridional, ces sols ont une grande importance en ce qui a trait à l'indépendance alimentaire du Québec et à l'économie agricole (Parent & Gagné, 2010). Or, leur mise en culture entraîne des pertes de sol par différents processus, entre autres sous forme de CO₂ contribuant au réchauffement climatique (Ilnicki & Zeitz, 2003). Citons par exemple le tassement, l'érosion et la minéralisation (Kroetsch et al., 2011). Ces pertes entraînent la détérioration des propriétés physiques, chimiques et hydrodynamiques des sols en plus de constituer un risque pour les cours d'eau environnants. L'évolution des propriétés rend les cartes et certains résultats d'analyse rapidement désuets (Kempen et al., 2009). De plus, à long terme, la mise en culture de certains champs ne sera plus rentable puisque la perte annuelle moyenne de 2,5 cm épuisera l'épaisseur de sol organique cultivable (Ilnicki, 2003; Esselami, Boudache & Grenon, 2014). Ces sols sont une ressource non renouvelable à protéger.

Par conséquent, une importante chaire de recherche industrielle en conservation et en restauration des sols a été mise sur pied. Entre autres, elle vise l'utilisation de biomasse afin de contrebalancer les pertes de sol et d'autres mesures préventives pour lutter contre l'érosion. Pour étudier et orienter ces mesures, les sols doivent être caractérisés au préalable. Grâce à la participation de 14 entreprises maraîchères de la Montérégie-Ouest, une importante banque de données a pu être amassée afin d'établir un bilan initial des sols à l'étude dans le cadre de ce projet. Ce besoin est justifié par le volume d'informations que nécessite le plan en conservation des sols et par le manque d'informations dans la littérature portant sur la caractérisation complète des sols sur le long terme en intégrant une grande quantité de paramètres chimiques et physiques. Encore moins d'études portent sur la détermination de seuils de qualité pour les propriétés retenues. De nouvelles connaissances doivent être acquises dans ce domaine.

Au début de ce projet, des données préliminaires internes (non publiées) indiquaient que quatre groupes de sols semblaient se distinguer par leurs propriétés physiques et hydrodynamiques. L'étude présentée dans ce document, réalisée à plus grande échelle et intégrant davantage de paramètres (chimiques et pédologiques), avait pour but d'étudier cette observation et de former des groupes de gestion des sols adaptés à la variabilité présente dans les cinq tourbières étudiées.

Chapitre 1 – Revue de littérature

1.1 Les sols organiques

Au Canada, les sols organiques représentent l'un des 10 ordres de sols retrouvés sur le territoire (GTCS, 2002). Au plan taxonomique, 4 grands groupes organiques divisent l'ordre : les humisols, les mésisols et les fibrisols qui sont formés par des plantes de milieu humide, alors que les folisols, le quatrième grand groupe, est formé en zone sèche en milieu forestier des hautes terres. Ce dernier est retrouvé majoritairement dans l'Ouest canadien. Les sols forestiers de la forêt boréale ne sont pas cartographiés, donc on ne connaît pas l'importance des folisols dans ce vaste territoire.

Quant à leur classification, les règles sont précises : il faut un minimum de 40 cm de matériau organique pour les sols à surface humique ou mésique et un minimum de 60 cm pour les sols fibriques, en plus d'avoir une teneur minimale de 17 % de carbone organique (p/p) (GTCS, 2002). Le grand groupe est déterminé selon la dominance d'horizons humique, mésique ou fibrique entre 40 et 120 cm de profondeur. La grande distinction entre les grands groupes se situe au niveau de leur dégradation et du pourcentage de fibres frottées dans l'horizon dominant. Les grands groupes sont associés à du matériel organique de plus en plus dégradé dans l'ordre suivant : fibrique < mésique < humique.

À l'échelle du globe, l'agriculture est la principale utilisation des sols organiques, suivie par les domaines de la foresterie, de l'énergie (combustion) et de l'horticulture (sacs de terre). Le Canada abrite environ le quart de la superficie mondiale de tourbières (Vepraskas & Craft, 2015). En termes de superficie, environ 8 % du pays sont couverts par des sols organiques issus de milieux humides (Waddell, 2010, cité dans Kroetsch et al., 2011). Comme seule une fraction de cette superficie est utilisée à des fins agricoles et que ces sols ont une formation et une évolution bien particulière, il est important de bien comprendre les processus en jeu afin de progresser dans les prochaines sections.

1.1.1 La formation des sols organiques

Les sols organiques se forment par l'un des deux processus pédogénétiques suivants. La paludification, particulièrement présente en région boréale, entraîne la formation d'un milieu humide souvent marqué par la colonisation de sphaigne en zone terrestre a priori sèche à la suite d'un changement de régime hydrique. Le comblement, deuxième processus, survient lorsqu'une dépression se remplit graduellement d'eau et de matériel organique qui, en condition d'anaérobiose, ne pourra pas se décomposer. Cependant, ce dernier processus n'est pas le responsable principal de la formation des sols organiques au pays, la paludification étant plus importante (Kroetsch et al., 2011).

Ces sols se développent sur des dépôts minéraux dont le drainage peut être qualifié de mauvais à très mauvais. Ils sont issus principalement de matière organique végétale provenant d'espèces ayant poussé sur le site d'accumulation (Kroetsch et al., 2011). Plus en détails, la formation de ces sols résulte de quatre processus de formation plus généraux : l'accumulation (feuilles, racines, tiges, branches, matériaux organiques transportés par l'eau), la perte (érosion hydrique, éolienne, décomposition physique et microbienne), le transfert (dissolution, lessivage, relocalisation par la faune) et la transformation (décomposition) de matières organiques (Kroetsch et al., 2011).

Deux types de tourbière sont distinguables : les bogs (ombrotrophiques), qui sont maintenus humides par la pluie et qui ont une eau moins oxygénée et renouvelée, et les fens (minérotrophes) qui sont généralement alimentés par une eau souterraine riche en éléments minéraux et moins acide. Il est possible de distinguer ces milieux humides par leur pH, leur forme, leur richesse en nutriments et ultimement, leur potentiel agricole (Warner & Rubec, 1997). En Montérégie-Ouest, les sols organiques sont issus principalement du comblement de dépressions laissées par les glaciers et sont davantage des bogs et des tourbières boisées (des marécages principalement constitués de tourbe forestière bien décomposée, recouvrant quelquefois de la tourbe fennique) (Lucie Grenon, communication personnelle, 2019).

En terminant, il est possible de retrouver chez les tourbières formées par comblement une couche de sédiments lacustres entre le matériel minéral et le sol organique. Le coprogène, parfois appelé gyttja, est une couche de matériel limnique se déposant en eaux tranquilles tel le fond d'un lac. Ce matériel sédimentaire de plancton, de déjections de poissons et d'autres résidus de plantes ou d'animaux peut aussi contenir des restes de diatomées et du carbonate de calcium (Okruszko & Ilnicki, 2003). Une telle couche est non cultivable et s'avère imperméable. Elle nuit au drainage et à la mise en culture de sols peu profonds (Kroetsch et al., 2011).

1.1.2 L'évolution des sols organiques cultivés

Les conditions de formation et d'accumulation anaérobique de la tourbe permettent, sur des centaines d'années, d'obtenir ces sols riches et convoités pour la culture maraîchère. Pour les mettre en culture, leur drainage est nécessaire, mais les prive de leurs conditions de conservation. D'importantes modifications irréversibles surviennent à ce moment.

À la suite du drainage, il y a consolidation du sol (retrait initial), perte de matière organique (MO) par décomposition et affaissement du sol dans les années qui suivent (Vepraskas & Craft, 2015). Le retrait a lieu dans les premières années seulement. Cette perte de hauteur est associée à la perte de portance du sol à la suite du retrait de l'eau en surface et aux forces capillaires qui ne permettent pas la reprise de volume lors de la réhumectation (Driessen et al., 2000; Brandyk et al., 2003). Le sol ne parvient plus à supporter son propre

poids et il s'affaisse (Brandyk et al., 2003; Ilnicki, 2003). Le retrait peut atteindre 53 à 70 % du volume initial de la tourbe. Le degré de décomposition et la composition botanique vont affecter son importance (Zeitz & Velty, 2002; Ilnicki & Zeitz, 2003).

Après le drainage initial, un nouveau processus pédogénétique transforme le sol à sa surface, soit le processus de formation de la couche de Moorsh (PFCM) (Okruszko & Ilnicki, 2003). Le PFCM est caractérisé par une humification et une minéralisation intenses. Plus ce processus est avancé, plus le matériel tourbeux s'affine et la présence d'agrégats se fait plus importante dans la couche de Moorsh. Ces modifications morphologiques sont principalement à la surface du sol puisqu'elles sont liées à la présence d'oxygène, de microorganismes et aux cycles d'humectation-dessiccation qui amplifient le PFCM (Zeitz & Velty, 2002; Ilnicki, 2003; Okruszko & Ilnicki, 2003). Cette couche, surtout retrouvée dans les premiers 30 cm de sol, se situe fréquemment au-dessus d'une couche compacte. Elle est rarement divisible en sous-couches au Canada puisque les sols sont travaillés, homogénéisant les effets du phénomène de dégradation (Okruszko & Ilnicki, 2003).

L'oxydation du sol entraîne à la fois sa minéralisation et son humification. Tel que le précisent Kroetsch et al. (2011), le fait de cultiver un sol organique sur le long terme entraîne un changement de son état de décomposition et donc de sa classification. Les sols organiques sont en constante évolution, modulée par leur teneur initiale en carbone et leur utilisation (Mann, 1986). Cette évolution est unidirectionnelle et orientée vers la dégradation des sols organiques. La minéralisation du sol, principalement en surface, est fonction de plusieurs facteurs tels le degré de décomposition de la tourbe, la température, sa teneur en eau et en air, et le ratio des nutriments disponibles aux microorganismes (Ilnicki & Zeitz, 2003).

Sur le plus long terme, un affaissement entre 1 et 5 cm par an est constaté par Ilnicki (2003), ce qui rejoint le 1 à 4 cm mesuré par Esselami, Boudache et Grenon (2014). L'érosion hydrique et éolienne, ainsi que la décomposition du sol, en sont les principales causes (Kroetsch et al., 2011). La température et la profondeur de la nappe semblent exercer une influence des plus importantes sur l'affaissement. En ce sens, les cycles saisonniers influencent l'évolution des sols organiques. Les cycles de gel-dégel mèneraient à un affaissement irréversible du sol, alors que les cycles d'humectation-dessiccation causeraient la formation de fentes de retrait qui favorisent l'aération du sol en profondeur (Ilnicki, 2003). Quant à elle, la profondeur de la nappe influence le régime hydrique du sol et son oxygénation. Une nappe à 90 cm entraînerait une émission maximale de CO₂, mesure fortement liée à la décomposition du sol (Ilnicki & Zeitz, 2003; Okruszko & Ilnicki, 2003).

Les fentes de retrait peuvent se former lorsque le sol s'assèche et tombe sous les 65% (v/v) de teneur en eau. Cela augmente l'aération en profondeur et contribue au PFCM (Okruszko & Ilnicki, 2003). Une teneur en air volumique de 20 à 30 % semble optimale pour augmenter l'intensité du processus de dégradation (Ilnicki & Zeitz, 2003).

En conditions plus sèches, les sols organiques deviennent hydrophobes, ce qui contribue à limiter la remontée capillaire, favorisant la décomposition du sol. Ce phénomène d'assèchement est amplifié par une restructuration porale. En effet, les mésopores sont affectés par une diminution marquée de leur nombre au bénéfice principal des micropores qui retiennent plus fortement l'eau, limitant la remontée capillaire. (Ilnicki & Zeitz, 2003; Okruszko & Ilnicki, 2003).

1.2 Caractérisation des sols

Plusieurs études portent sur le suivi des propriétés des sols organiques à la suite du drainage des tourbières. Cette section traite des mesures les plus fréquemment utilisées afin d'évaluer et de comparer l'évolution de ces sols en fonction des différents processus en cause, des différents matériaux rencontrés et de leur utilisation. Des relations entre l'évolution de certaines propriétés sont aussi couvertes.

1.2.1 Pédologie

Le sol est sujet à une transformation génétique lorsque la présence d'oxygène permet sa décomposition. Les concepts de la couche de Moorsh et du PFCM proviennent des classifications polonaise et allemande (Okruszko & Ilnicki, 2003). Au Canada, le système de classification y réfère respectivement en tant qu'horizon Ohp et d'une humification des couches de surface (GTCS, 2002). Par humification, il est entendu que le sol évolue d'une dominance fibrique à mésique puis à humique de façon irréversible en fonction de son accumulation initiale.

Avec le temps, la structure des résidus tourbeux change et s'affine de façon à ce qu'on ne puisse plus reconnaître les fragments des végétaux qui la composaient. C'est entre autres avec le contenu en fibres frottées (en volume) que l'on peut évaluer l'état de la décomposition des horizons de tourbe (particulièrement pour les horizons de surface en milieu drainé et cultivé). En milieu naturel non drainé, le degré de décomposition peut être évalué par l'échelle de von Post après avoir observé la couleur et la turbidité de l'eau s'écoulant de la main lorsqu'une poignée de terre est comprimée (GTCS, 2002). Cette méthode comporte 10 classes allant de non décomposée (H1) à complètement décomposée (H10) (Brandyk et al., 2003). Pour un horizon de sol fibrique (Of - peu décomposé) dont le contenu en fibres frottées est de 40 % et plus, l'échelle donne une valeur de 1 à 4 ; un contenu en fibres frottées de 10 à 40 % donne habituellement une valeur de 5 ou 6 pour un horizon mésique (Om – modérément décomposé); un contenu en fibres frottées plus petit que 10% donne une valeur de 7 à 10 pour un horizon humique (Oh - très décomposé) (GTCS, 2002). L'indice au pyrophosphate est une autre méthode permettant d'évaluer le degré de décomposition des horizons de tourbe (Carter & Gregorich, 2007). Comme il s'agit d'une caractéristique importante des couches de sol, il importe de l'évaluer au champ et/ou en laboratoire (Vepraskas & Craft, 2015).

Le PFCM a comme impact majeur le changement de la structure du sol, propriété fréquemment évaluée lors de la réalisation d'un profil de sol, ce qui affecte grandement les propriétés physiques et hydrauliques. Tel que mentionné plus haut, le sol subit un affaissement continu et une restructuration porale à la suite du retrait de l'eau. L'affaissement initial est rapide, mais décélère à mesure que la masse volumique apparente augmente. La faible densité initiale et le grand volume de pores avant drainage (78-93%) sont des facteurs propices au retrait du sol (Driessen et al., 2000; Ilnicki & Zeitz, 2003; van Asselen, Stouthamer, & Smith, 2010; Liu, Janssen, & Lennartz, 2016). De même, à long terme, la décomposition du sol ralentit alors que les substrats facilement décomposables laissent place à une proportion d'humus plus importante et à une teneur en cendres plus élevée.

Trois structures différencient la couche de Moorsh du matériel parental. Ces structures, d'un sol plus jeune à un sol plus décomposé, sont dites tourbeuses (peaty), humiques (humic) et granulaires (grainy) (Okruszko & Ilnicki, 2003). La modification de la structure influence la porosité totale et la capacité de rétention en eau.

En se dégradant, les fibres laissent place à des grains de 4 mm de diamètre pouvant se briser en de plus petites particules (Okruszko & Ilnicki, 2003). Les grains peuvent former des agrégats de 5-10 mm de diamètre, poreux et généralement hydrophobes. En se fragmentant en particules, la structure évolue de granulaire moyenne à inexistante, c'est-à-dire particulaire, tel du marc de café (GTCS, 2002). Les grains en surface sont plus petits que ceux en profondeur, uniformisant la structure du sol en surface. Ce dernier devient alors sensible à l'érosion éolienne puisque très fin et peu cohésif en condition sèche. De plus, l'hydrophobicité augmente le ruissellement en limitant l'infiltration et diminue la teneur en eau du sol. Cela contribue à fournir des conditions plus sèches au sol qui sera plus sensible à la transformation.

La structure d'un sol plus dégradé (granulaire fine à particulaire) est associée à une porosité totale moindre et à une plus grande rétention en eau au point de flétrissement permanent ($pF = 4.2$), témoignant d'une plus grande microporosité (Brandyk et al., 2003). Cette eau n'est pas nécessairement disponible aux plantes puisqu'elle est fortement retenue par les micropores (Vepraskas & Craft, 2015). Plus en détails, une perte de porosité totale de 3% survient dans les sols moins dégradés (structure tourbeuse), alors que cette perte triplerait pour les couches humiques dégradées à structure granulaire (Ilnicki & Zeitz, 2003).

Les facteurs de formation de ces sols (ex. : le climat, les organismes, etc.) varient selon la localisation du site d'accumulation, surtout à l'échelle mondiale. Cela peut mener à d'importantes différences quant à la quantité de matériel accumulé. Plusieurs méthodes existent pour évaluer l'épaisseur de sol organique (Parry et al., 2014). En plus de l'utilisation d'une tige de métal pour le sondage manuel, il existe une tarière pouvant retirer des échantillons verticaux de sol. Cette tarière est une tarière russe (« Russian peat borer »), aussi appelée sonde Macauley.

D'autres méthodes par proxi-détection (« proximal sensing ») et télédétection (« remote sensing ») existent afin d'évaluer l'épaisseur de sol organique. Ces méthodes sont relativement abordables (Ji et al., 2019). La proxi-détection pourrait offrir des résultats plus précis, compte tenu de la grande variabilité locale. Par exemple, divers instruments peuvent être utilisés afin de mesurer les variations de conductivité électrique et de constante diélectrique afin de cartographier la profondeur du sol organique (McBratney, Mendonça Santos, & Minasny, 2003; Minasny et al., 2019). En voici un exemple :

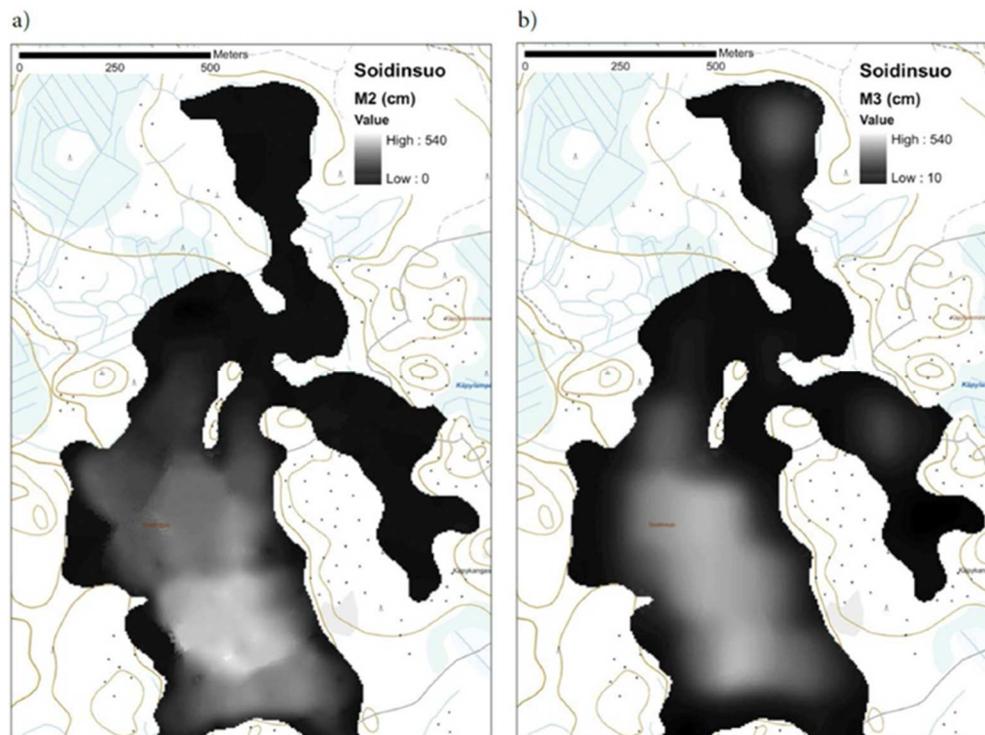


Figure 1 - Profondeur de sol organique estimée à l'aide d'une a) induction électromagnétique suivie d'une inversion 1D, et d'un b) géoradar combiné à un sondage manuel (Airo et al., 2014). Le titre de la figure a été traduit. Figure sous licence : Creative Commons CC-BY 4.0.

Un patron est observable quant à la profondeur de sol sur la figure 1. Peu importe la méthode utilisée, soit la proxi-détection en a) et la télédétection en b), la profondeur semble plus importante vers le centre de la tourbière qu'à ses limites (Minasny et al., 2019). Un constat similaire a été partagé par Hallema et al. (2015a). Cependant, la distance entre une zone de grande profondeur et le contour de la tourbière n'est pas homogène. Cela suggère qu'une étude de la profondeur de sol organique soit requise afin de cibler avec précision les champs moins profonds qui ont, par le fait même, une durée de vie utile estimée plus faible.

La profondeur des bassins de sol organique varie à l'échelle mondiale. Minasny et al. (2019) présentent un tableau dans lequel la profondeur moyenne de tourbières est listée pour des études mondiales et pour des tourbières tropicales spécifiquement en rassemblant les données de 10 études différentes. Une moyenne de 1,85 m et de 3,1 m de profondeur respectivement sont rapportées selon les études. De gros écarts sont observés, étant possiblement dus à la profondeur minimale requise pour être considéré un sol organique qui peut varier entre les systèmes de classification. De plus, certaines tourbières sont plus anciennes que d'autres, telle que celle sondée par Buffam et al. (2010) au Wisconsin qui comptait plus de 10 m d'accumulation à certains endroits. En fonction de la profondeur originale, un sol aura une durée de vie estimée plus ou moins longue selon une perte annuelle de quelques centimètres.

L'origine botanique des dépôts organiques est un facteur pédogénétique qui a aussi son rôle à jouer concernant la profondeur et l'évolution des tourbières. Anciennement, les tourbières naturelles du comté de Napierville étaient recouvertes de conifères comme des épinettes, mélèzes, pins et cèdres. La flore s'est adaptée au drainage des tourbières et des feuillus ont pu coloniser les anciens milieux humides comme les bouleaux blancs et gris, les peupliers, les érables rouges, les ormes et les saules (Lamontagne, Martin, & Nolin, 2014). Selon les mêmes auteurs, on y retrouve surtout des tourbières de bassin (bogs) ou des marécages de bordure. Des tourbes forestières et herbacées dominent la composition de la tourbe (LaSalle 1963; Grenon 1988). L'origine botanique du matériel parental tourbeux a une influence à bien des égards sur les propriétés physicochimiques et hydrauliques du sol. Bien que cet impact soit lié à la pédologie et qu'il soit abondamment évoqué dans la littérature, il sera plutôt traité dans les prochaines sous-sections afin d'éviter les répétitions.

1.2.2 Physique et hydrologie des sols organiques

Selon Vepraskas & Craft (2015), les propriétés physiques telles que la conductivité hydraulique saturée (K_{sat}), la masse volumique apparente (MVA) et la courbe de rétention en eau (CR) sont en grande partie fonction de la teneur en MO et du degré de décomposition de la tourbe. Ces paramètres sont de bons indicateurs de l'évolution du sol et de sa qualité, et influencent la productivité agricole et le comportement du sol. Si des stratégies de réennoiment sont utilisées afin de conserver les sols organiques, il est important de faire des simulations en zone saturée et insaturée à l'aide des paramètres hydrauliques nécessaires. Il est donc pertinent d'en traiter et d'en avoir une bonne compréhension puisqu'ils diffèrent de ceux en sol minéral (Dettmann et al., 2014).

1.2.2.1 La conductivité hydraulique saturée (K_{sat})

La K_{sat} est une mesure de la perméabilité du sol, soit la facilité avec laquelle un certain volume d'eau s'infiltré au travers d'une surface de sol en un temps donné. Ses unités sont habituellement en centimètre par seconde.

La Ksat dépend de l'origine botanique et de son degré de décomposition (Gnatowski et al., 2010). En guise de valeur de référence, une tourbe peu décomposée affiche des valeurs de 10^{-1} à 10^{-3} cm s⁻¹, alors qu'un matériel végétal très décomposé peut avoir une valeur inférieure à 10^{-6} cm s⁻¹ (Rycroft et al., 1975a, 1975b). Les espèces végétales qui se sont accumulées peuvent réduire l'infiltration de l'eau lorsque les résidus sont empilés sous forme de couche laminaire comme les laïches (Brandyk et al., 2003), lorsque des sphaignes composent majoritairement les résidus (Dessureault-Rompere et al., 2018) ou lorsque les racines des cultures se développent à l'horizontal près de la couche compacte (Hallema et al., 2015a). Cela peut mener à l'anisotropie du sol où la Ksat verticale ne sera pas égale à la Ksat horizontale. En effet, le ratio v/h est en général de 2 pour un fen peu décomposé, alors que ce même ratio est de 1:1 lorsque le sol est dans un état de dégradation avancé (plus homogène). Cependant, il faut préciser que ces valeurs diffèrent selon la méthode de mesure et l'importance des sphaignes dans la composition botanique. Cette espèce crée une stratification importante du sol (Brandyk et al., 2003).

Selon la figure 2, il est possible de constater que peu importe le degré de décomposition, la tourbe de joncs, de roseaux et de laïches (« Reed and sedge peat ») démontre les valeurs les plus élevées de Ksat alors que la tourbe de sphaignes (« sphagnum peat ») fait état des valeurs les plus faibles. Pour une valeur d'humification donnée, la plage de valeurs possibles est plus importante pour la tourbe de joncs, de roseaux et de laïches, indiquant une plus grande variabilité reliée à ces espèces végétales. Des tendances similaires ont été observées par Liu & Lennartz (2019). La porosité active, celle qui participe au drainage et autres flux d'eau dans le sol par une interconnexion entre les pores, est plus élevée dans le sol à base de mousses que le sol boisé (Gnatowski et al., 2010). Les sols à prédominance de plantes herbacées auraient une porosité active intermédiaire. Comme cette porosité permet l'infiltration de l'eau, elle est liée à la Ksat.

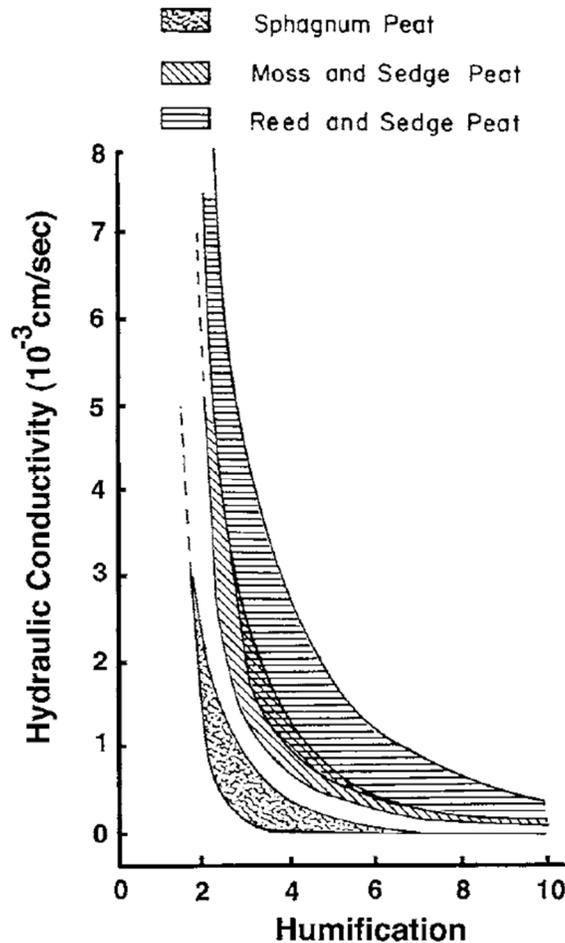


Figure 2 - Conductivité hydraulique saturée ($10^{-3} \text{ cm sec}^{-1}$) en fonction du degré d'humification de trois types de composition botanique. L'échelle de von Post est utilisée afin d'évaluer la décomposition du sol (1 = peu décomposé à 10 = très décomposé). Reproduit à partir de « Wetland Soils : Genesis, Hydrology, Landscapes, and Classification » par Vepraskas, M.J. & Craft, C.B. © 2000, p.351, CRC Press. Reproduit avec la permission de Taylor and Francis Group, LLC, une division d'Informa plc.

De plus, la K_{sat} diminue avec la profondeur (Brandyk et al., 2003). En surface, le travail de sol et la formation de canaux préférentiels incluant les biopores permettent de créer davantage de porosité active et continue. Le PFCM a aussi un impact négatif sur la K_{sat} . Cette dernière montre des valeurs plus faibles dans un sol dégradé (Zeitz, 1991; Sauerbrey and Zeitz, 1999, cités dans Ilnicki & Zeitz, 2003). Tel que discuté à la prochaine section, la MVA augmente avec la dégradation du sol. Il est donc possible de visualiser la baisse de la K_{sat} en fonction de la MVA à la figure 3 a). Selon la figure, il semble que la K_{sat} soit généralement supérieure dans les fens.

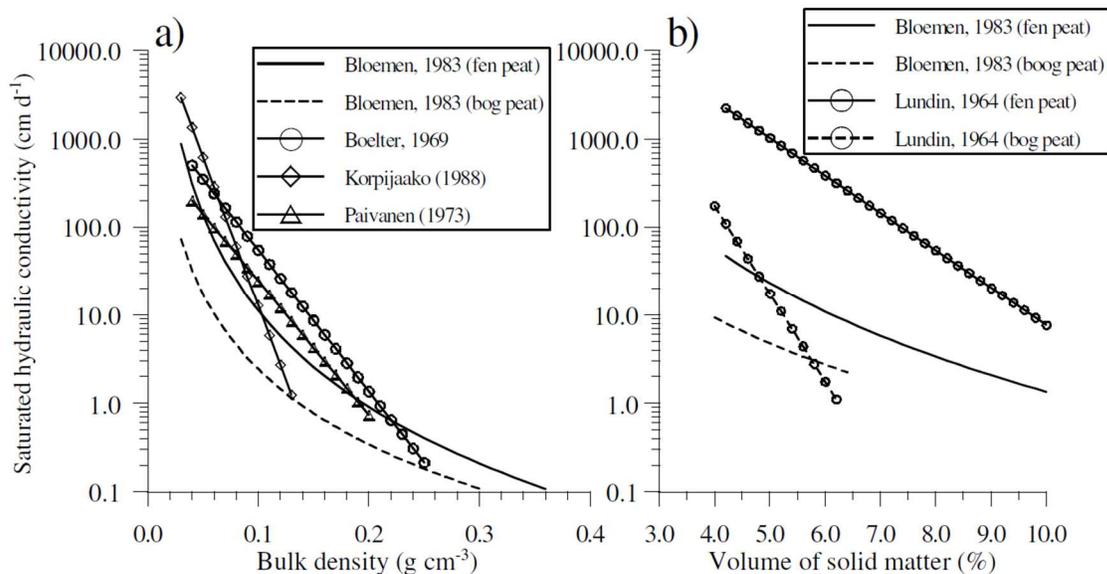


Figure 3 - Conductivité hydraulique saturée (cm jour⁻¹) en fonction a) de la masse volumique apparente (g cm⁻³) et b) du volume de solide (%) de sols organiques selon plusieurs études. Reproduit à partir de « Organic soils and peat materials for sustainable agriculture » par Parent L.E. & Ilnicki, P. © 2002, p.55, CRC Press. Reproduit avec la permission de Taylor and Francis Group, LLC, une division d'Informa plc.

Cependant, il faut être prudent lors de l'analyse de ce paramètre du sol. Dans la littérature, il est aisé de constater la grande variabilité des K_{sat} rapportées, mais aussi des méthodes de mesure employées (Gnatowski et al., 2010). La K_{sat} peut être mesurée en laboratoire ou au champ. Plusieurs chercheurs ont démontré que les valeurs observées en laboratoire sont généralement plus élevées (Brandyk et al., 2003). Ce biais peut être relié au transport des échantillons de sol, en théorie, non perturbés. Entre les sites d'échantillonnage, l'utilisation du sol et les pratiques culturales peuvent varier, ce qui contribue à la variance élevée observée pour les données de K_{sat} (Liu & Lennartz, 2019). Une autre problématique rencontrée lors de l'analyse et la comparaison de la K_{sat} est sa faible corrélation spatiale et temporelle (Brandyk et al., 2003). Une solution afin de retrancher les valeurs aberrantes liées à la grande variabilité des K_{sat} est de conserver les observations entre les 2,5^e et 97,5^e percentiles, tel qu'exécuté par Hallema et al. (2015a).

1.2.2.2 Masse volumique apparente

La MVA augmente avec le drainage, la décomposition du matériel parental, la profondeur et la compaction (Brandyk et al., 2003; Ilnicki & Zeitz, 2003; Anshari et al., 2010).

En comparant les données de couches de sol sujettes au PFCM au matériel tourbeux parental, Okruszko & Ilnicki (2003) rapportent que le sol est plus dense de 0,2 à 0,3 g cm⁻³ et que la porosité y est plus faible. Il est à noter qu'une grande variabilité spatiale existe d'une tourbière à l'autre et que des différences significatives de

MVA peuvent exister entre les couches d'un même profil de sol (Vepraskas & Craft, 2015). En effectuant des variogrammes sur la MVA, Brandyk et al. (2003) ont obtenu une portée de 157 m pour la MVA, indiquant la présence d'une corrélation spatiale importante. Cependant, la pépite était très élevée (entre 63 % et 83 % de la portée). Il est possible de conclure à une forte variation des propriétés physiques des couches de sol à l'échelle locale d'échantillonnage (soit 10 m dans le cas de cette étude).

Tableau 1 - Propriétés physiques de tourbes fenniques en fonction de leur degré de décomposition (Okruszko & Ilnicki, 2003). R1 = fibrique (Of), R2 = mésique (Om) et R3 = humique (Oh). Source originale : Okruszko, H. & Szymanowski, M. 1992. Proc. 9th Int. Peat Congr., Uppsala, Sweden, 3 :106-115. Reproduit avec la permission de la International Peatland Society.

Degree	Observations No.	Ash (kg kg ⁻¹)	Bulk Density (g cm ³)	Porosity	Moisture Content (m ³ m ⁻³)		
					pF ^a 2.0	pF 2.7	pF 4.2
R ₁	138	0.084	0.114	0.927	0.665	0.396	0.132
R ₂	171	0.114	0.141	0.912	0.715	0.496	0.195
R ₃	177	0.151	0.175	0.893	0.686	0.527	0.252

^a pF = log₁₀(ψ_M), where ψ_M is the matric potential in centimeters.

Selon le tableau 1, un degré d'humification plus important vient de pair avec une augmentation du compactage par un changement de MVA, tout comme une altération des propriétés hydrauliques du sol au niveau de la remontée capillaire, de la capacité de rétention en eau, de la porosité et de la Ksat. C'est pourquoi la MVA est un bon paramètre physique pouvant être utilisé pour prédire la Ksat et la rétention en eau (Thompson & Waddington, 2014).

1.2.2.3 La courbe de rétention et la porosité

La courbe de rétention (CR) en eau illustre le comportement du sol lors du drainage. La courbe met en relation le potentiel matriciel du sol et sa teneur en eau volumétrique. Plus en détails, la courbe permet de connaître la valeur de la porosité totale du sol (pF = 0), sa teneur en eau à capacité au champ (pF = 0,7 à 2,7 selon le sol) et au point de flétrissement permanent (pF = 4,2), puis la réserve en eau utilisable par soustraction (capacité au champ – point de flétrissement permanent). De plus, en comparant plusieurs courbes d'un même type de sol, il est possible de comparer leur niveau de compaction (Brandyk et al., 2003). En effet, une diminution de la capacité de rétention en eau est observable lorsque le sol est dégradé (Zeitz & Velty, 2002). De plus, la taille des particules et la structure affectent la CR (Shein et al., 2018).

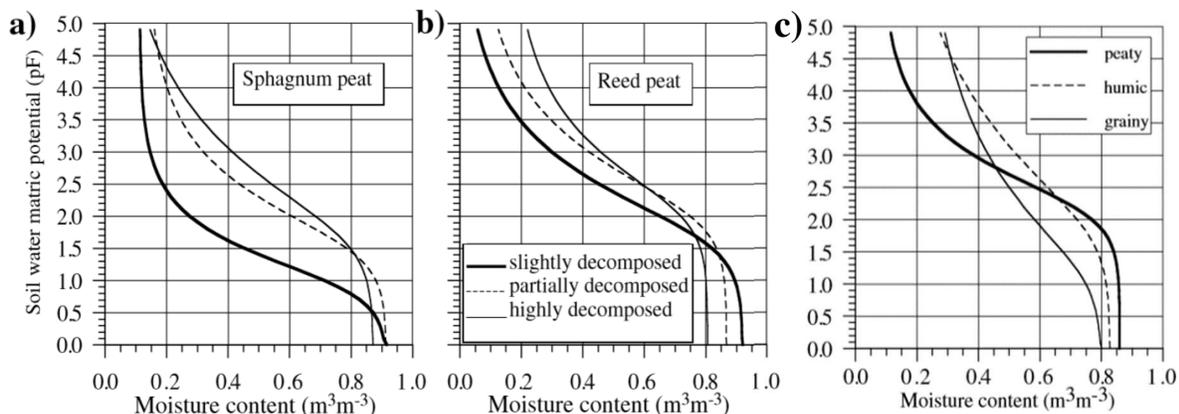


Figure 4 - Courbes de rétention a) d'une tourbe de sphaigne, b) de joncs et de roseaux (Päivänen, 1973), et c) des trois structures de la couche de Moorsh, soit tourbeuse (peaty), humique (humic) et granulaire (grainy) (adapté de Brandyk et al., 2003). Le titre des figures a) et b) a été traduit. Figures sous licence : Creative Commons CC BY-SA 4.0. Figure c) reproduite à partir de « Organic soils and peat materials for sustainable agriculture » par Parent L.E. & Ilnicki, P. © 2002, p.50, CRC Press. Reproduit avec la permission de Taylor and Francis Group, LLC, une division d'Informa plc.

À la figure 4, une baisse de la porosité totale à saturation est observable, ainsi qu'une plus grande rétention en eau à des tensions élevées plus la décomposition est avancée en a) et en b). Ce raisonnement est aussi valide en c), selon l'évolution de la structure primaire. Cependant, cette eau n'est pas nécessairement disponible aux plantes puisqu'elle est retenue plus fortement par la nouvelle microporosité plus importante à la suite du changement de structure vers une structure particulaire (Brandyk et al., 2003).

Tableau 2 - Porosité moyenne d'une tourbe fibrique (R1) de mousse et de laïche, d'une tourbe d'aulne humique (R3) et des structures de la couche de Moorsh de sols de la Pologne (Ilnicki & Zeitz, 2003).

Pour des raisons de droits d'auteur, la figure a été retirée de cet ouvrage. Le lecteur est invité à consulter le tableau 2.5 du livre « *Organic soils and peat materials for sustainable agriculture* » par Parent et Ilnicki (2003).

Le tableau 2 permet de constater l'évolution de la porosité totale et des différentes classes de porosité. Tel qu'observé à la figure 4, la porosité totale diminue avec la décomposition (R1 vers R3) et avec la dégradation de la structure. La mésoporosité diminue d'environ 50%, alors que la microporosité augmente de près du tiers. D'autres facteurs influencent la courbe de désorption en eau du sol. L'hydrophobicité affecte la rétention en eau des couches de surface en modifiant l'angle de contact des gouttes avec le sol. Elle est surtout fonction de la teneur en eau, mais elle dépend aussi du degré de décomposition et de la profondeur (Brandyk et al., 2003).

Plus le sol s'assèche, plus il devient hydrophobe. En profondeur, l'hydrophobicité maximale atteinte est plus importante et survient à des teneurs en eau plus faibles.

Quant aux propriétés hydrodynamiques, l'hystérèse devient plus faible dans les sols plus dégradés. Il s'agit d'une dissymétrie entre la courbe de rétention en eau au moment du séchage et de la réhumectation qui est fortement liée à la macroporosité du sol ($>50 \mu\text{m}$) (Ilnicki & Zeitz, 2003). Il semble qu'en évoluant vers une structure plus granulaire ou particulaire, le sol s'uniformise en diminuant l'hystérèse. Ce phénomène est causé par les forces hydrophobes et l'écoulement préférentiel (Hallema et al., 2015a).

La composition botanique a aussi son importance sur la capacité de rétention en eau du sol (Boelter, 1968, cité dans Vepraskas & Craft, 2015; Weiss et al., 1998; Gnatowski et al., 2010). En effet, une tourbe herbacée plus décomposée, ayant donc une MVA plus élevée, est caractérisée par un espace poral plus petit et une capacité de rétention en eau plus faible. Shein et al. (2018) ont déterminé que la teneur en eau est fonction de la MVA, de la teneur en cendres et du degré de décomposition du sol dans cet ordre d'importance. Comme ces relations ne sont pas linéaires, les auteurs mentionnent qu'il s'agit de l'influence d'autres facteurs non étudiés comme la composition botanique.

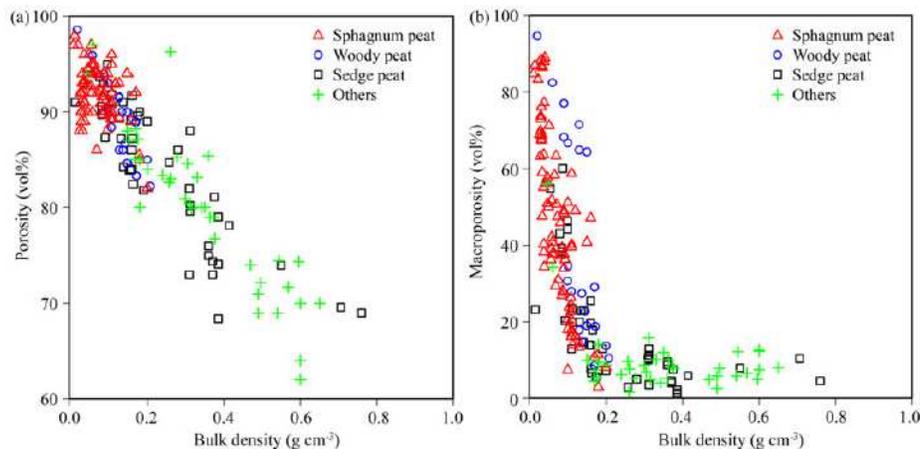


Figure 5 - Porosité totale a) et macroporosité (diamètre équivalent de $50 \mu\text{m}$) b) de différentes compositions botaniques en fonction de la masse volumique apparente (Liu & Lennartz, 2019). On y retrouve la sphaigne (sphagnum), le bois (woody), la laïche (sedge) et d'autres origines tourbeuses (others). Copyright © 1969 Reproduit avec la permission de John Wiley and Sons, Inc.

En ce sens, les matériaux organiques semblent partager des fourchettes de MVA différentes selon la figure 5. Alors que la sphaigne et la tourbe boisée varient moins en termes de porosité totale et de MVA, il semble que leur macroporosité diminue autant que celle d'une tourbe à base de laïche ou d'autres matériaux organiques.

Selon une autre étude, un sol à prédominance de mousse aurait la plus faible rétention en eau (Gnatowski et al., 2010). Un sol d'origine botanique boisée montrait les plus fortes valeurs d'humidité résiduelle avec des valeurs aussi hautes que $0,4$ à $0,5 \text{ m}^3 \text{ m}^{-3}$, comparativement à $0,05 \text{ m}^3 \text{ m}^{-3}$ pour le sol d'origine mousseuse. Cependant, la porosité active du sol boisé était la plus faible. Le sol herbacé semble montrer des propriétés intermédiaires à ce niveau et avait les valeurs les plus élevées de K_{sat} (intermédiaires pour le bois et plus faibles pour la mousse).

1.2.3 Chimie des sols

Cette section couvre les principaux indicateurs chimiques de la dégradation du sol tels qu'énumérés par Ilnicki & Zeitz (2003). Parmi ces indicateurs figurent la composition du sol en éléments organiques et inorganiques. Selon une étude sur une ancienne forêt en tourbière tropicale, l'augmentation du pH et de l'azote totale (Nt) seraient des indicateurs chimiques importants de dégradation, si l'apport d'engrais azoté est pris en considération (Anshari et al., 2010). Une augmentation en oligoéléments adsorbés au sol de surface est observable. La couche de Moorsh contiendrait aussi plus d'éléments inorganiques (Si, Fe, P et Al) (Ilnicki & Zeitz, 2003). Cependant, il faut se questionner sur le choix des meilleurs indicateurs chimiques de dégradation en fonction de la rapidité de la méthode d'analyse et de son coût. L'étude des oligoéléments est très peu répandue dans la littérature.

La teneur en carbone organique (C_{org}) est liée à la teneur en MO par un facteur de conversion de 0,58 généralement accepté dans la littérature. La méthode par perte au feu devrait être privilégiée afin d'évaluer la teneur en MO d'un sol selon Minasny et al. (2019). Quant au facteur de conversion, il ne serait pas fixe, mais il dépendrait plutôt de la composition botanique du sol comme le rapportent plusieurs auteurs (Farmer et al., 2014; Klingenuß et al., 2014) et pourrait même être obtenu en utilisant la MVA pour une certaine plage de valeurs par une fonction de pédotransfert (Warren et al., 2012; Farmer et al., 2014; Rudiyanto et al., 2016a).

Une équipe a suivi durant 3 ans une tourbière drainée à l'aide de tranchées drainantes dans le but d'étudier les variations de concentration de solutés dans le temps et l'espace. Cette équipe a notamment étudié le C_{org} et le Nt. L'étude a été réalisée sur des transects de 6 m où l'élévation était mesurée en plusieurs points. Une analyse de surfaces par tendance des propriétés du sol a permis d'expliquer entre 73 % et 92 % de la variance des deux paramètres. En d'autres mots, les différences d'élévation de la surface du sol à l'échelle du transect sont principalement responsables des différences observées au niveau des teneurs mesurées. Dans un troisième transect relativement plat, cela correspondait à 34 % de la variance pour le Nt et 42 % pour le C_{org} (Tiemeyer et al., 2007).

Plus en détails, les valeurs observées de C_{org} variaient entre 0,8% et 41%, et celles de Nt entre 0,08% et 2,87% (Tiemeyer et al., 2007). Il faut cependant se rappeler qu'au Canada, il faut un minimum de 17% de C_{org} pour être considéré comme un sol organique. Il appert que la distance avec la tranchée drainante et l'élévation relative de la surface sont deux paramètres très importants qui affectent les valeurs observées selon les auteurs (Tiemeyer et al., 2007). Les analyses ont aussi révélé la présence d'une corrélation spatiale entre des mesures de C_{org} ou de Nt espacées de 5 à 26 m. Au-delà de cette distance, il est impossible de prédire la valeur qui sera observée, rendant l'extrapolation des données à l'échelle d'un champ plus risquée.

Une autre équipe a suivi trois prairies établies sur des sols organiques à 50 m (pré coupé quelques fois), 150 m (prairie récoltée fréquemment) et 300 m (culture plus intensive de foin et pâturage) des berges d'une rivière en Pologne. Le but de cette recherche était de mieux comprendre les variations des propriétés physicochimiques en fonction du lieu et de la régie de culture (Pawluczuk, Stepien, & Alberski, 2019). Le tableau des résultats des analyses chimiques est présenté plus bas.

Tableau 3 - Analyses chimiques pour deux couches de sol en fonction de la distance des sites avec la rivière (Pawluczuk et al., 2019). Le titre du tableau a été traduit. Tableau sous licence : Creative Common CC-BY-NC-SA 4.0.

Specification	Layer (cm)	Distance from the river (m)		
		50 m (MtIc _{pl180})	150 m (MtIc _{pl160})	300 m (MtIc _{pl140})
N-total (g kg ⁻¹)	5-10	1.39a	1.27ab	0.61b
	35-40	1.73a	1.15b	1.23b
P- total (g kg ⁻¹)	5-10	2.66a	1.20b	0.71c
	35-40	1.74a	1.08a	0.87a
K- total (g kg ⁻¹)	5-10	1.00a	1.00a	1.00a
	35-40	0.83a	0.83a	0.83a
Mg- total (g kg ⁻¹)	5-10	1.33b	1.57b	2.05a
	35-40	1.27a	1.50a	1.87a

a, b, c, ... values with the same letter are not significantly different according to the Tukey's test ($P \leq 0.05$)

Les résultats démontrent une différence significative pour le Nt, le Pt et le Mgt selon la distance avec la rivière à deux profondeurs. La teneur en Nt est associée à la hauteur de la nappe et à la teneur en air (Pawluczuk et al., 2019), c'est pourquoi la minéralisation du N se retrouve davantage dans la couche de Moorsh en surface (Ilnicki & Zeitz, 2003). Cependant, l'utilisation des sites n'a pas eu d'impact significatif sur la dégradation des sols (Pawluczuk et al., 2019). Sachant que le sol était couvert par la prairie dans tous les cas, les résultats ne sont pas nécessairement comparables à des champs maraîchers sur butte sans culture de couverture.

L'aération de surface a un effet sur la teneur en Nt (figure 6). Les sols drainés forment un groupe qui se distingue clairement des deux autres quant à leur teneur plus élevée en Nt.

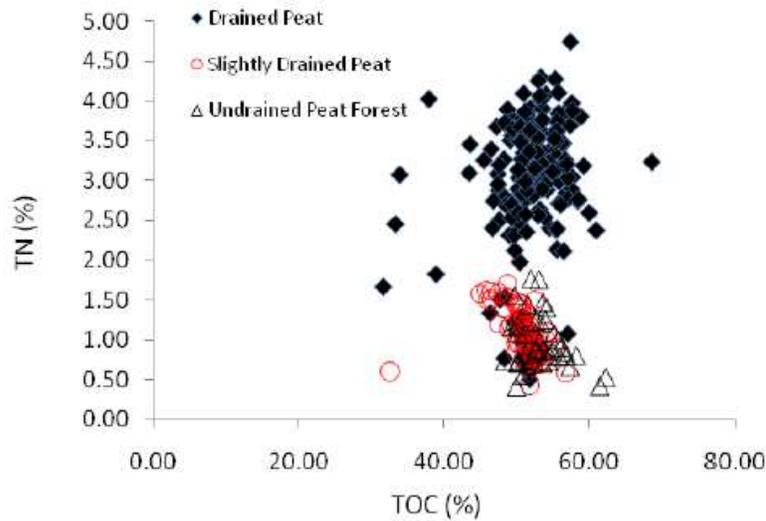


Figure 6 - Nuage de points de la teneur en azote total (TN) en fonction de la teneur en carbone organique total (TOC) pour différents régimes hydriques (Anshari et al., 2010). Le titre de la figure a été traduit. Figure sous licence : Creative Commons CC-BY 4.0.

Des chercheurs ont étudié les indicateurs de dégradation pour un bog en Allemagne dans trois sites en fonction de leur utilisation plus ou moins intensive. Les paramètres étudiés sont le C¹³, N¹⁵, la teneur en cendres, le ratio C/N et la MVA (Kruger et al., 2015).

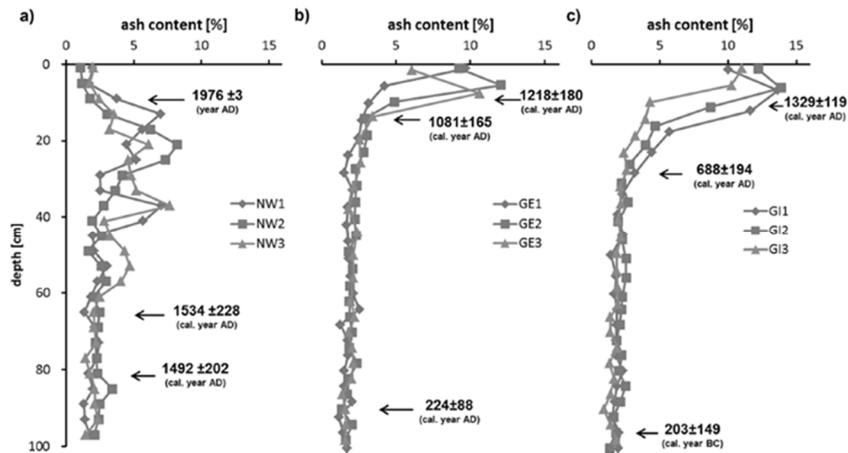


Figure 7 - Profils de teneur en cendres avec datation au carbone pour un site a) peu affecté par l'homme, b) utilisé extensivement en prairie et c) utilisé intensivement en prairie (Kruger et al., 2015). À chaque site, trois sondages ont eu lieu afin d'obtenir des répliques (ex. : NW1, NW2 et NW3). Le titre de la figure a été traduit. Figure sous licence : Creative Commons CC-BY 4.0.

Cette figure démontre que la minéralisation se produit en surface lorsque le sol est cultivé, mais aussi que les sols peuvent montrer des signes de minéralisation ayant eu lieu lors de l'accumulation de la tourbe. De plus, plus l'activité humaine est importante, plus la teneur en cendres est élevée sur une profondeur plus importante, témoignant d'une plus grande minéralisation à la suite du drainage et de la mise en culture du sol. Cela concorde avec une perte de C_{org} plus prononcée dans les sols plus dégradés (baisse de la MO) (Kruger et al., 2015).

Pour finir, la différence entre le pH tampon SMP d'un échantillon de sol humide et séché à l'air ou à l'étuve est un indicateur de l'avancement du PFCM (Parent & Tremblay, 2003). En effet, une tourbière non perturbée voit son pH tampon SMP augmenter de 0,22 à 0,25 unité à la suite du séchage. Des mécanismes de condensation et d'hydrophobicité seraient en cause. Alors que l'acidité de réserve diminue pour une tourbière intacte, l'inverse est observé pour un horizon sujet au PFCM. Parent & Tremblay (2003) rapportent une baisse du pH tampon SMP de 0,12 unité associée à ce type d'horizon. Afin d'évaluer l'intensité de la dégradation subie par le sol, il serait possible d'observer la variation du pH tampon. Plus la valeur diminue après le séchage, plus le sol est dans un processus avancé de décomposition.

1.3 En terrain connu

Cette section résume les travaux issus de trois études ayant eu lieu dans les mêmes bassins de sol organique que ceux visités dans le cadre du présent projet afin d'avoir un aperçu des propriétés de la région.

La première étude date de 2015 et portait sur la relation entre la résistance à l'enfoncement (RE) et la K_{sat} pour trois sites : un fribrisol limnique (1), un humisol mésique terrique (2) et un humisol terrique (3) (Dessureault-Romppe et al., 2018). La résistance à l'enfoncement, tel que son nom l'indique, est une mesure de la résistance du sol en MPa à l'enfoncement d'une tige métallique munie d'un cône à son extrémité. Généralement, un pénétromètre manuel est utilisé pour sa mesure. La valeur de résistance est mesurée pour chaque centimètre jusqu'à 80 cm. Cette mesure permet de localiser et d'évaluer l'intensité de la couche compacte.

Alors qu'aucune corrélation n'existait entre les deux paramètres, les auteurs ont toutefois pu observer le développement d'une couche compacte de plus en plus importante en fonction de la dégradation des sites. C'est ce que permet d'observer la figure 8. Cette couche était plus compacte entre 30 et 40 cm de profondeur.

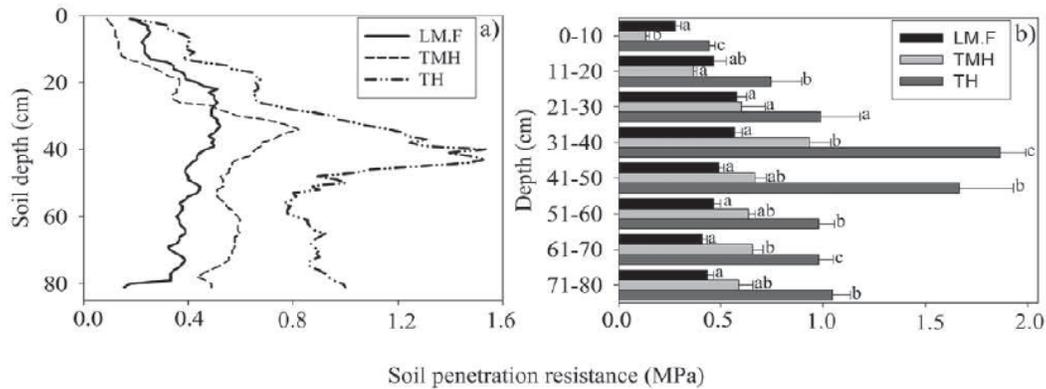


Figure 8 - Résistance à l'enfoncement mesurée dans trois sites au degré de décomposition différent : a) Profils de résistance à l'enfoncement et b) valeur maximale enregistrée par tranche de 10 cm (Dessureault-Romppe et al., 2018). Différentes lettres signifient une différence significative entre les types de sol selon le test de Tukey DSH. Le titre de la figure a été traduit. Figure sous licence : Creative Common CC-BY-NC-ND 4.0.

Plus le sol est humifié (LM.F < TMH < TH), plus la couche compacte est importante selon la figure 8a. De même, les sols moins profonds (sous-groupe terrique, ici représenté par TMH et TH) montrent des valeurs en général significativement plus élevées que le sol plus profond en b) qui est profond d'au moins 1,6 m (Dessureault-Romppe et al., 2018).

Dans d'autres études complémentaires, 207 profils de sol répartis sur 5 fermes ont été étudiés dans le but d'évaluer des modèles d'interpolation (Lafond et al., 2015). Les auteurs ont mesuré la profondeur au sol minéral, la MVA à 4 profondeurs (5-10, 20-25, 35-40, 45-50 cm), le % de MO, la porosité totale, la Ksat (20 cm et 30 cm) et ont généré la courbe de rétention.

La profondeur de sol variait entre 20 et 260 cm (médiane 127,5 cm), alors que la MVA variait entre 0,132 et 0,562 g cm⁻³ (médiane 0,262). La teneur en MO affichait des valeurs entre 42,6 et 96,7 % (médiane 85,9 %). La porosité totale se trouvait entre 0,706 et 0,912 cm³ cm⁻³ et la valeur de la Ksat entre 1,14 x 10⁻⁵ et 6,32 x 10⁻² cm s⁻¹ (médiane 1,00 x 10⁻³ cm s⁻¹) (Hallema et al., 2015a; Lafond et al., 2015). Quant à la CR, le modèle de rétention en eau offrant la meilleure calibration est le modèle van Genuchten (Hallema et al., 2015b). Le même modèle a aussi été utilisé par (Liu & Lennartz, 2019). Une couche compacte se trouvait en moyenne entre 29 et 37 cm dans 50 % des cas (Hallema et al., 2015a, 2015b).

Afin de pousser plus loin les analyses, Hallema et al. (2015b) ont réalisé un partitionnement hiérarchique en utilisant les médianes des paramètres physiques et hydriques évalués. Il en découle deux groupes de qualité différente, le nombre de groupes ayant été choisi arbitrairement lors de l'analyse des résultats.

Tableau 4 - Valeur médiane des paramètres en fonction de leur groupe issu d'un partitionnement hiérarchique (adapté de Hallema et al., 2015b). Copyright © 1969 Reproduit avec la permission de John Wiley and Sons, Inc.

Propriété	Groupe 1	Groupe 2
<i>N</i>	30	55
Masse volumique apparente (g cm ⁻³)	0,316	0,257
Masse volumique réelle (g cm ⁻³)	1,645	1,628
Porosité (%)	80,9	84,4
Matière organique (%)	86,0	88,5
Profondeur de sol organique (cm)	145	200
Profondeur de la couche compacte (cm)	32,5	34,0
Ksat (cm/s)	1,462 x 10 ⁻⁴	2,290 x 10 ⁻³
Teneur en eau à saturation (m ³ m ⁻³) selon le modèle van Genuchten	0,799	0,826

Le groupe 2, présenté au tableau 4, semble d'une meilleure qualité comparativement au groupe 1. Ce dernier montre des signes de compaction qui ont pu favoriser le groupement des sites. Le groupe 2 est caractérisé par une MVA plus faible, une plus grande porosité, des sols plus profonds (tout comme une couche compacte plus profonde) et une Ksat plus élevée. De même, les CR comparatives obtenues selon le modèle de van Genuchten présentées à la figure 9 permettent de constater la compaction.

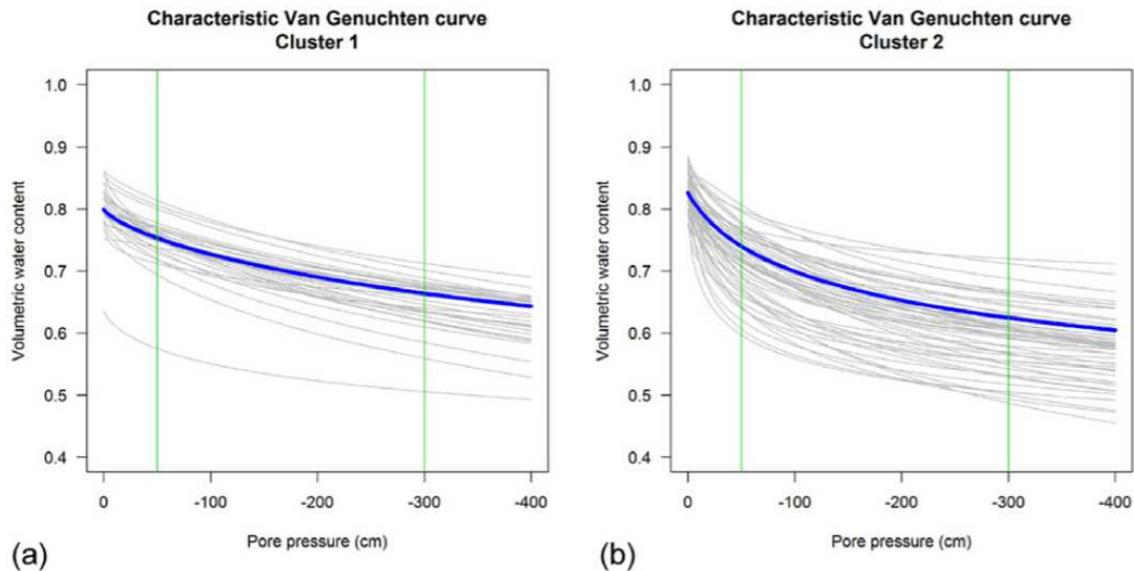


Figure 9 - Courbe de rétention moyenne basé sur la valeur médiane des paramètres calibrés selon le modèle de van Genuchten pour le groupe 1 a) et le groupe 2 b) (Hallema et al., 2015b). Copyright © 1969 Reproduit avec la permission de John Wiley and Sons, Inc.

La porosité totale est plus faible en a) et le sol a une teneur en eau plus élevée à de faibles tensions, signifiant une augmentation de la microporosité. Il est aussi possible d'observer une plus grande variabilité en b), comparativement au groupe présenté en a) qui semble plus uniforme à la suite d'une compaction dans la plupart des sites échantillonnés.

Dans une dernière étude, Hallema et al. (2015a) ont étudié l'effet de 40 ans de culture sur les propriétés hydrauliques et physiques des mêmes 5 fermes. Ils ont utilisé l'année de mise en culture comme facteur discriminant en créant deux groupes, soit les champs datant d'avant 2000 et d'après 2000. Ce qui sépare les deux groupes de champs était la MVA plus faible et la teneur en MO plus élevée pour les champs jeunes. La Ksat semble diminuer avec le temps. De plus, selon une matrice de corrélation, il existe une forte corrélation négative entre la MVA et la porosité (-0,983), et entre la MVA et le contenu en MO (-0,748), alors qu'il existe une corrélation positive entre la porosité et la teneur en MO (0,627). La corrélation la plus forte avec l'année de mise en culture est avec la teneur en MO (0,366). Liu & Lennartz (2019) se sont prêtés au même exercice et ont obtenu une forte corrélation négative entre la MVA et la porosité (-0,91), la MVA et le contenu en MO (-0,89), et une forte corrélation positive entre la porosité et la teneur en MO (0,77). Ces valeurs sont semblables entre les études. De plus, ils ont obtenu une corrélation négative modérée entre le logarithme de la Ksat et la MVA (-0,47), alors que Hallema et al. (2015a) ont obtenu une faible corrélation négative de -0,118.

D'autres observations sont faites par Hallema et al. (2015a). Ils rapportent l'absence de relation entre la profondeur de la couche compacte et l'épaisseur de sol organique, bien que la MVA soit plus élevée là où le sol est moins profond. De plus, la Ksat était plus élevée à 35 cm qu'à 20 cm.

Voici un tableau de statistiques descriptives réalisé avec les données de l'étude. Hormis l'année de mise en culture, les statistiques sont basées sur des jeux de données de 207 à 886 observations, ce qui constitue un résumé statistique important de la situation entre les étés 2008 et 2012.

Tableau 5 - Statistiques descriptives de différentes propriétés : MVA (ρ_b), MVR(ρ_p), porosité (n), teneur en MO (f_o), profondeur de sol organique (D_o), profondeur de la couche compacte (z_{cl}), la Ksat (K_s), la réserve en eau utilisable (SWAC) pour la laitue et l'année de mise en culture (Reclam. yr) (Hallema et al., 2015a). Copyright © 1969 Reproduit avec la permission de John Wiley and Sons, Inc.

Statistic	ρ_b	ρ_p	n	f_o	D_o	z_{cl}	K_s	SWAC	Reclam. yr
	g cm ⁻³		cm cm ⁻³		cm		cm s ⁻¹	cm	
Observations, no.	886	661	877	878	220	222	233	207	102 fields
Mean	0.263	1.669	0.845	0.830	134.2	33.6	4.13×10^{-3}	13.74	1971
Min.	0.132	1.572	0.706	0.426	20	10	1.14×10^{-5}	2.00	1955
Max.	0.562	2.035	0.912	0.967	260	85	6.32×10^{-2}	34.29	2006
Median	0.262	1.643	0.846	0.859	127.5	32.0	1.00×10^{-3}	13.50	1965
First quartile	0.190	1.612	0.813	0.801	85.0	29.3	2.79×10^{-4}	9.37	1960
Third quartile	0.316	1.683	0.883	0.903	190.8	36.8	3.74×10^{-3}	18.07	1970
SD	0.089	0.090	0.045	0.109	60.7	9.2	8.61×10^{-3}	6.51	13.6
CV	0.338	0.054	0.053	0.131	0.452	0.273	2.08	0.47	6.88×10^{-3}
Skewness	0.825	1.900	-0.487	-1.557	0.154	1.799	4.07	0.46	1.605

L'étude du coefficient de variation (CV) permet d'apprécier la variation importante quant à la MVA, la profondeur de sol organique, la profondeur de la couche compacte et la Ksat. Ce tableau est constitué de valeurs de référence qui pourront être utilisées dans le présent projet. Il est aussi intéressant de constater une profondeur médiane de 127,5 cm, ce qui signifie qu'au moins la moitié des 220 points sondés sont qualifiés de terriques au niveau de leur classification pédologique (GTCS, 2002).

En résumé, les études dans le domaine semblent comprendre l'analyse de plusieurs propriétés, souvent de nature physique ou hydrodynamique. Beaucoup de données quant à la distribution de la valeur des propriétés sont à disposition afin de faciliter les comparaisons, bien qu'une grande variabilité existe pour certaines propriétés en sols organiques. Cela est notamment dû à la composition botanique et aux processus de décomposition. Notons toutefois l'absence de seuils de qualité liés à ces propriétés.

1.4 Propriétés discriminantes de la qualité du sol et classes de gestion

Cette section se veut une récapitulation des sections précédentes où les liens les plus évidents entre les propriétés du sol ont été mis de l'avant. Quelques méthodes fréquemment utilisées afin de discriminer les sols selon certaines propriétés seront ensuite données en exemple.

1.4.1 Quelques propriétés importantes et seuils associés

À la lumière des dernières sections, l'importance de la composition botanique et du degré de décomposition de celle-ci est importante. Son évaluation au champ (méthode des fibres frottées) et en laboratoire (indice au pyrophosphate) semble des plus importantes afin de qualifier l'état des champs à l'étude (Kroetsch et al., 2011). À cela s'ajoutent d'autres analyses récurrentes qui témoignent de leur potentiel, telle que l'étude de la morphologie, de la MVA et de la porosité qui sont de bons indicateurs de l'évolution des propriétés du sol (Ilnicki & Zeitz, 2003; Okruszko & Ilnicki, 2003). Les mesures liées à la compaction et la profondeur de sol sont aussi à considérer.

Afin de discriminer les sols en fonction de leurs propriétés, l'utilisation de seuils critiques serait une première étape d'analyse en mode univarié. Alors que plusieurs indices de qualité et de dégradation pour les sols organiques ont été traités plus haut, l'utilisation de seuils critiques de qualité est pratiquement inexistante dans la littérature. Néanmoins, il est possible de trouver quelques seuils de références. Une méta-analyse basée sur la publication de 24 articles en hydrologie a démontré que la composition botanique et la décomposition du sol affectaient significativement les paramètres hydrauliques (Liu & Lennartz, 2019). Une valeur critique de MVA a été avancée au terme de cette étude, soit $0,2 \text{ g cm}^{-3}$. De part et d'autre de ce seuil, la macroporosité et les paramètres hydrauliques se comportent différemment en relation avec la MVA lorsque des fonctions de pédo-transfert sont utilisées. Lafond et al. (2015) ont utilisé une valeur limitante de K_{sat} de $10^{-4} \text{ cm s}^{-1}$ dans leur étude puisque des valeurs supérieures sont requises pour un bon drainage en milieu agricole et pour éviter l'asphyxie racinaire. Il est aussi possible de s'inspirer des milieux artificiels à base de tourbe où une teneur en air à 1 kPa inférieur à $0,10\text{--}0,15 \text{ cm}^3 \text{ cm}^{-3}$ est problématique pour la culture (Caron et Michel, 2021).

Selon Bengough & Mullins (1990), des racines en croissance peuvent exercer une force maximale de 0,9 à 1,3 MPa sur le sol pour se frayer un chemin dans ce dernier. La littérature offre peu de valeurs précises sur la valeur de RE limitante, encore moins en sol organique. Il serait donc prudent d'assumer une valeur de 1,5 MPa comme limitante à la croissance des racines, stressant suffisamment le plant pour limiter l'exploration du sol en profondeur ou même les rendements. À noter que cette valeur peut varier en fonction du type de sol et de la culture.

1.4.2 Quelques analyses fréquentes

Deux types d'analyses statistiques reviennent fréquemment dans la littérature : l'analyse en composantes principales (ACP) et le partitionnement hiérarchique. Les modèles de mélange ainsi que le partitionnement récursif seront aussi discutés étant donné leur potentiel.

L'ACP a été utilisée dans une étude afin de déterminer l'importance relative de certaines propriétés du sol et évaluer la redondance de ces dernières (Brown et al., 1990). Cette méthode, ou une matrice de corrélation, peut être utilisée comme première étape pour déterminer quelles propriétés mesurées devraient être conservées comme propriétés discriminantes. Une sélection de variables ou les composantes issues de l'analyse peuvent être introduites dans un modèle statistique au terme de l'ACP.

L'utilisation du partitionnement hiérarchique a déjà été abordée plus haut afin de séparer en groupes des observations en mode multivarié, c'est-à-dire tenir compte de l'influence simultanée de plusieurs propriétés sur la variable réponse (Brown et al., 1990; Ye & Wright, 2010; Hallema et al., 2015b). Cette méthode est utile lorsqu'aucune variable de groupement n'existe pour les observations a priori. Dans une autre étude, des valeurs de référence pour des éléments majeurs, secondaires et mineurs en fonction de la dégradation du sol ont été déterminées à l'aide du partitionnement hiérarchique (Lima et al., 2016). Le dendrogramme produit a permis de diviser les observations en trois groupes de dégradation. Le partitionnement hiérarchique réunit à chaque embranchement deux groupes (et ultimement deux observations) en mode multivarié selon une méthode d'agglomération. Une de ces méthodes qui a été utilisée dans le domaine de la recherche en sols est la méthode de Ward (Lima et al., 2016; Zawieja & Glina, 2017). Cette dernière vise à minimiser la somme des carrés par regroupement. La distance euclidienne est fréquemment utilisée afin de mesurer la distance entre les groupes, soit leur similarité (Gnatowski et al., 2010; Lima et al., 2016; Zawieja & Glina, 2017). Afin d'évaluer la transformation des sols organiques en montagne, cette méthode s'est montrée efficace pour regrouper les sites par type de tourbière. Les auteurs suggèrent son utilisation dans le domaine de l'évolution des sols organiques, surtout pour les sols d'origine botanique différente (Zawieja & Glina, 2017). Le choix du nombre de groupes à considérer au terme du partitionnement revient à l'analyste, bien que des indices soient disponibles afin de déterminer le nombre de groupes significatifs. Par exemple, la librairie R NbClust (Charrad et al., 2014) utilise 30 indices reconnus dans le domaine du partitionnement (Milligan et Cooper, 1985) afin de suggérer le nombre optimal de groupements significatifs.

Un modèle de mélange, une analyse de classes latente, est aussi envisageable. Le logiciel Mplus (Muthén & Muthén, 2010) permet ce genre d'analyse en offrant plus de flexibilité que la librairie R mclust (Scrucca et al., 2016) par exemple, quoique la licence soit payante. Lorsque les groupes ne sont pas connus a priori, ce type de modèle permet de diviser un jeu de données en classes probables selon l'espace des données multivariées. La sortie Mplus fournit plusieurs tests statistiques afin d'inférer sur la qualité du modèle, tel que l'entropie (évalue la séparation des classes) ou encore le test de Lo-Mendell-Rubin (qui permet de déterminer le nombre de classes significatives), ce qui donne certains avantages à ce modèle comparativement à d'autres analyses de partitionnement, en plus d'être basé sur les données elles-mêmes (Oberski, 2016).

Le partitionnement récursif permet de construire un modèle en deux étapes afin de former des groupes par régression (variables continues) et classification (variables catégorielles) sous forme d'arbres avec des embranchements binaires (Hothorn & Everitt, 2010; Therneau & Atkinson, 2019). Ces arbres peuvent être élagués afin de s'assurer que le modèle produit ne soit pas sur-ajusté ou sous-ajusté au jeu de données. Cette approche est robuste aux relations non linéaires entre les prédicteurs et la variable dépendante, et permet de représenter clairement les relations à l'aide d'un arbre de classification.

Plus en détails, le jeu de données est divisé prédicteur par prédicteur à l'aide d'un seuil numérique pour les variables continues, divisant en deux les observations, ou à l'aide de l'appartenance à une catégorie pour les prédicteurs catégoriels (Hothorn & Everitt, 2010). Un critère de fin est spécifié par l'utilisateur afin que le modèle puisse savoir à quel moment il faut arrêter de subdiviser le jeu de données. L'algorithme est construit de sorte qu'il recherche le prédicteur parmi tous ceux présents permettant de diviser le jeu de données en deux groupes les plus semblables aux groupements réels. Ensuite, il procède avec un autre prédicteur afin de faire grandir l'arbre. Pour finir, l'arbre est élagué (Therneau & Atkinson, 2019).

Quant au pouvoir prédictif de l'arbre, chaque sous-groupe est caractérisé par des statistiques descriptives comme la moyenne ou la médiane pour les données continues, ou la proportion des classes pour les variables catégorielles. Un problème peut cependant survenir si les seuils ne sont pas marqués dans les données originales puisque l'arbre force un seuil précis (Hothorn & Everitt, 2010). Les méthodes d'ensemble permettent alors de générer un grand nombre d'arbres à l'aide de l'approche par *bootstrap* avec les mêmes variables. Une nouvelle observation est alors classifiée par chacun des arbres et la moyenne de la prédiction est calculée.

En terminant, d'autres méthodes de classifications existent, telles que le *random forest*, les réseaux neuronaux, la méthode *K-means*, etc. L'autoapprentissage (« *machine learning* »), que ce soit pour la régression ou la classification, est un outil puissant et flexible utilisé dans le domaine des sciences du sol. Il est conseillé de comparer les résultats de plusieurs modèles qui peuvent représenter et comprendre différemment les relations entre les variables.

1.5 La cartographie des sols

Dans le domaine de la pédologie, la cartographie de base des types de sol (classes générales) est complétée pour l'ensemble du globe (FAO, 2020). Par conséquent, les nouveaux efforts dans le domaine visent la cartographie des propriétés des sols et de leurs fonctions (rétention du carbone, filtration de l'eau, etc.) (Lemercier, 2019) à différentes échelles spatiales. Concernant les sols organiques, des efforts en cartographie

des tourbières ont déjà été déployés pour une partie des superficies au Canada (18 %), en Russie (12 %), en Indonésie (11 %) et en Angleterre (11 %) (Minasny et al., 2019).

Afin guider la cartographie des sols, il faut maîtriser certains concepts liés à leur pédogénèse. Cela permettra de mieux guider l'échantillonnage et le choix des covariables utilisées dans le modèle prédictif. Le sol est la résultante de cinq facteurs sur sa pédogénèse, soit le climat, la topographie, la présence d'organismes (incluant les humains), le matériau parental et le temps, tels que définis par Jenny (1941). Ce modèle conceptuel a été rapidement accepté par les pédologues et a servi de base aux études pédologiques (Vepraskas & Craft, 2015). Il faut se rappeler que le travail d'un pédologue consiste à créer un modèle conceptuel de la formation des sols dans sa tête en fonction de différents facteurs à la suite d'un échantillonnage et de l'utilisation de photographies aériennes, ce qui rend son modèle peu explicite (McBratney et al., 2003; Kerry et al., 2012; Lemerrier, 2019).

La cartographie traditionnelle a maintenant laissé place à la cartographie numérique (CN) des sols par modélisation statistique. Cette approche est définie par Lemerrier (2019) comme l'intégration de covariables environnementales et de données de calibration dans un modèle statistique informatique permettant la prédiction d'un phénomène d'intérêt dans une région donnée. Plusieurs modèles sont basés sur des relations impliquant les facteurs de formation du sol, mais principalement sur la topographie (McBratney et al., 2003). Aujourd'hui, la CN englobe entre autres les SIG, les géostatistiques et différentes méthodes d'acquisition de données dont la télédétection (Arrouays et al., 2020).

La CN utilise un modèle de description empirique et quantitatif basé sur une fonction de prédiction spatiale découlant de celle de (Jenny, 1941), appelée *scorpan* selon l'équation suivante : « $S_{c\text{ ou }a} = f(s,c,o,r,p,a,n) + e$ » (Lemerrier, 2003; McBratney et al., 2003; Lemerrier et al., 2012; Minasny et al., 2019). Ce modèle tente de prédire des classes de sol (S_c) ou des propriétés sélectionnées (S_a) à une position donnée en fonction d'observations et de covariables liées au sol (s), au climat (c), aux organismes (o), au relief (r), au matériel parental (p), au temps (a) et à la position spatiale (n), tout en considérant l'erreur spatialement corrélée (e) (Minasny et al., 2019). Tel que spécifié dans l'équation, il est possible d'utiliser des propriétés du sol afin d'en prédire d'autres ou encore pour prédire un type de sol. Dans la formule, le « f » représente la fonction utilisée dans le modèle prédictif qui peut être une régression linéaire, une régression logistique ou des méthodes par autoapprentissage tels que les arbres de classification ou le *random forest*.

Comme plusieurs propriétés du sol évoluent dans le temps, l'utilisation seule de cartes conventionnelles peut être limitative. De plus, l'information reliée à leur incertitude est rarement disponible. En matière de cartographie numérique, la cartographie de l'incertitude des estimés est presque aussi importante que la cartographie des

estimés eux-mêmes. L'incertitude est très élevée lorsque les erreurs de mesure, de localisation et celles liées à la modélisation sont considérées (Lemercier, 2019; Minasny et al., 2019). La CN permet d'utiliser les cartes historiques afin d'intégrer de nouveaux paramètres que les pédologues ne pouvaient prendre en considération à l'époque grâce à l'accessibilité des covariables et à la puissance de calcul des ordinateurs (Arrouays et al., 2020). D'autres avantages sont attribués à la CN afin de prédire des classes de sol ou des propriétés du sol. Par exemple, la reproductibilité, le faible coût et la possibilité de connaître les facteurs pédogénétiques principalement responsables des observations (Arrouays et al., 2020).

Le travail du cartographe numérique débute par la définition de la zone à l'étude, la collecte des covariables en fonction des données disponibles et la caractérisation de ces données (Lemercier, 2019). L'utilisation de données historiques par leur désagrégation permet de limiter le rééchantillonnage et de diminuer les coûts en CN (Kerry et al., 2012). Lorsque cette méthode est utilisée, il est essentiel de calibrer les données à certains endroits clés de la zone à l'étude afin de s'assurer de la validité des mesures, surtout en changeant de zone (Ji et al., 2019). L'utilisation de cartes historiques afin de prédire les propriétés des sols organiques doit dépendre de leur l'âge (Kempen et al., 2009). Cela est dû à l'oxydation rapide et à l'évolution constante de ces sols. Un compromis doit être fait entre l'utilisation de vieilles données et la perte de sites de calibration pour le modèle.

Tel que précisé précédemment, des covariables sont nécessaires afin de prédire les types de sol. Il est fortement conseillé de bien évaluer la pertinence de chaque covariable utilisée et de s'assurer d'avoir un nombre minimal d'observations au champ pour la surface ciblée lors de la CN d'une région. De plus, ne pas rendre disponible les estimés d'incertitude ou négliger les techniques de validation des données sont des manquements importants qu'il faut éviter afin de produire une carte de qualité et faire avancer la méthodologie de cette approche cartographique moderne (Arrouays et al., 2020). McBratney et al. (2003) mentionnent l'utilisation de la proxi-détection et de la télédétection pour obtenir des covariables avec une grande densité d'information sur le territoire. Il n'est parfois pas nécessaire de payer davantage pour des fichiers de données d'une plus grande résolution à utiliser comme covariable, sauf si la covariable est importante au niveau de la prédiction de la variable en sortie (Samuel-Rosa et al., 2015). Un rééchantillonnage peut être une solution.

Selon leur revue de littérature, McBratney et al. (2003) mentionnent que la CN est utilisée dans 70% des cas afin d'estimer des attributs du sol plutôt que des classes de sol (30%). Dans les études, le nombre d'observations au terrain varie entre 30 et 2448, pour une médiane de 180. Plusieurs études utilisent des cartes historiques comme source principale de données. La topographie et le sol sont les deux facteurs du modèle *scorpan* les plus utilisés dans les études. De plus, un seul facteur est utilisé dans 40% des études et deux facteurs dans un autre 40% des études. Les modèles numériques de terrain (pente, élévation, etc.), les données issues de

télédéttection et des cartes de sol historiques sont les sources de données les plus utilisées dans la littérature (McBratney et al., 2003). La CN est relativement récente et est toujours en développement (Minasny & McBratney, 2016). Il faut toutefois demeurer conservateur sur l'interprétation de ces statistiques qui datent de 2003.

Il existe cependant des lacunes dans cette approche. Par exemple, l'importance du nombre et de la localisation des points échantillonnés, tout comme le fait que les covariables ne soient qu'une approximation des conditions réelles ayant formé le sol (Samuel-Rosa et al., 2015). D'autres évoquent aussi la multiplicité des jeux de données et le transfert vers les utilisateurs (Lemerrier, 2019) ou encore la baisse de la qualité de la prédiction lorsque les unités de sol ou les conditions de formation sont trop semblables (Ellili et al., 2019). Il faut s'assurer de la qualité des données ainsi que de leur quantité, porter attention au surajustement des modèles et aux références circulaires de la donnée en entrée et en sortie (McBratney et al., 2003).

1.6 Hypothèses et objectifs

Au terme de cette revue de littérature, 3 hypothèses et 3 objectifs ont été formulés afin d'orienter ce mémoire :

1.6.1 Hypothèses

1- Il existe une association significative à $p < 0,05$ entre la qualité de sol organique et ses propriétés physiques, ses propriétés chimiques et ses propriétés pédologiques.

2- Les sols organiques peuvent se répartir en au moins deux classes de sensibilité à la dégradation en fonction de leur caractérisation physico-chimique afin d'en faciliter leur conservation.

3- Il existe un ensemble de covariables environnementales significativement corrélées à $p < 0,05$ aux sols organiques afin de cartographier les groupes de gestion à l'aide d'algorithmes d'autoapprentissage.

1.6.2 Objectifs

1-Établir un portrait initial (valeurs de référence) de zones d'intervention sur les 14 fermes partenaires de la Chaire de recherche industrielle en conservation et restauration des sols organiques cultivés à l'aide d'une caractérisation des propriétés physico-chimiques et de profils de sols pour permettre un suivi temporel de l'état des sols.

2-Définir des classes de gestion de sols en fonction de propriétés discriminantes afin d'appuyer un outil d'aide à la décision aux interventions de conservation des sols.

3-Cartographier les ressources en sols et les classes de gestion en s'appuyant sur un échantillonnage à faible intensité des champs et des données satellitaires.

1.6.2 Résumé schématique du processus d'analyse

Afin de mieux visualiser l'approche conceptuelle mise de l'avant dans les chapitres suivants, un schéma faisant la corrélation entre les approches utilisées ainsi que les hypothèses et objectifs est présentée à la page suivante.

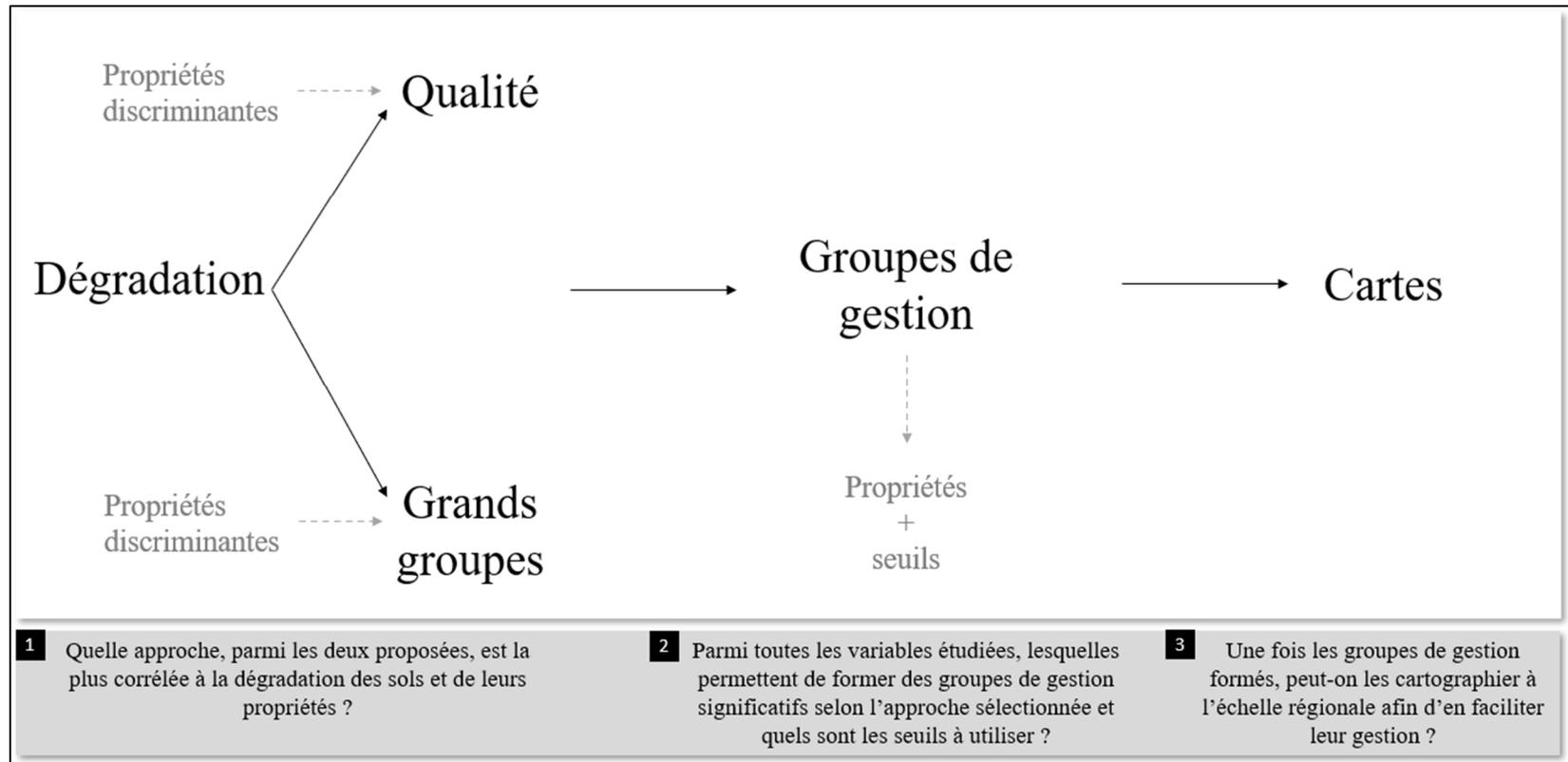


Figure 10 – Schéma de l'approche conceptuelle utilisée reprenant les hypothèses de départ représentées par trois étapes successives. Les étapes 1 et 2 se retrouvent dans le chapitre 2, alors que l'étape 3 sera abordée au chapitre 3.

Bibliographie

- Airo, M.-L., Hyvönen, E., Lerssi, J., Leväniemi, H. & Ruotsalainen, A. 2014. Tips and tools for the application of GTK's airborne geophysical data. Geological Survey of Finland, Report of Investigation 215, 33 pages https://www.researchgate.net/publication/274008539_Tips_and_tools_for_the_application_of GTK's_airborne_geophysical_data
- Arrouays, D., McBratney, A., Bouma, J., Libohova, Z., Richer-de-Forges, A. C., Morgan, C. L. S., . . . Mulder, V. L. (2020). Impressions of digital soil maps: The good, the not so good, and making them ever better. *Geoderma Regional*, 20. doi:10.1016/j.geodrs.2020.e00255
- Bengough, A. G., & Mullins, C. E. (1990). Mechanical impedance to root growth: a review of experimental techniques and root growth responses. *Journal of Soil Science*, 41(3), 341-358. doi:10.1111/j.1365-2389.1990.tb00070.x
- Brown, D., Mathur, S., Brown, A., & Kushner, D. (1990). Relationships between some properties of organic soils from the southern canadian shield. *Canadian journal of soil science*, 70(3), 363-377. doi:10.4141/cjss90-037
- Buffam, I., Carpenter, S. R., Yeck, W., Hanson, P. C., & Turner, M. G. (2010). Filling holes in regional carbon budgets: predicting peat depth in a north temperate lake district. *Journal of Geophysical Research: Biogeosciences*, 115(G1). <https://doi.org/10.1029/2009JG001034>
- Charrad, M., Ghazzali, N., Boiteau, V. & Niknafs, A. (2014). NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, 61(6), 1-36. <http://www.jstatsoft.org/v61/i06/>
- Dessureault-Rompré, J., Theriault, L., Guedessou, C., & Caron, J. (2018). Strength and Permeability of Cultivated Histosols Characterized by Differing Degrees of Decomposition. *Vadose Zone Journal*, 17(1). doi:10.2136/vzj2017.08.0156
- Dettmann, U., Bechtold, M., Frahm, E., & Tiemeyer, B. (2014). On the applicability of unimodal and bimodal van Genuchten–Mualem based models to peat and other organic soils under evaporation conditions. *Journal of Hydrology*, 515, 103-115. doi:10.1016/j.jhydrol.2014.04.047
- Driessen, P., Deckers, J., Spaargaren, O., & Nachtergaele, F. (2000). *Lecture notes on the major soils of the world*: Food and Agriculture Organization (FAO).
- Ellili, Y., Malone, B. P., Michot, D., Minasny, B., Vincent, S., Walter, C., & Lemerrier, B. (2019). Comparing three approaches of spatial disaggregation of legacy soil maps based on DSMART algorithm. *SOIL Discussions*, 1-39. doi:10.5194/soil-2019-36
- Farmer, J., Matthews, R., Smith, P., Langan, C., Hergoualc'h, K., Verchot, L., & Smith, J. U. (2014). Comparison of methods for quantifying soil carbon in tropical peats. *Geoderma*, 214-215, 177–183. <https://doi.org/10.1016/j.geoderma.2013.09.013>

- Hothorn, T., & Everitt, B.S. (2010). Chapter 9 - Recursive Partitioning: Predicting Body Fat and Glaucoma Diagnosis A Handbook of Statistical Analyses Using R (2nd ed.). Chapman and Hall/CRC. <https://doi.org.acces.bibl.ulaval.ca/10.1201/9781420079340>
- Jenny, H. (1941). Factors of Soil Formation. *Soil Science*, 52(5), 415. doi:10.1097/00010694-194111000-00009
- Kerry, R., Goovaerts, P., Rawlins, B. G., & Marchant, B. P. (2012). Disaggregation of legacy soil data using area to point kriging for mapping soil organic carbon at the regional scale. *Geoderma*, 170, 347-358. doi:10.1016/j.geoderma.2011.10.007
- Klingenuß, C., Roßkopf, N., Walter, J., Heller, C., & Zeitz, J. (2014). Soil organic matter to soil organic carbon ratios of peatland soil substrates. *Geoderma*, 235-236, 410–417. <https://doi.org/10.1016/j.geoderma.2014.07.010>
- Lafond, J. A., Gumiere, S. J., Hallema, D. W., Périard, Y., Jutras, S., & Caron, J. (2015). Spatial Distribution Patterns of Soil Water Availability as a Tool for Precision Irrigation Management in Histosols: Characterization and Spatial Interpolation. *Vadose Zone Journal*, 14(6), 1-13. doi:10.2136/vzj2014.10.0140
- Lemercier, B. (2019). La cartographie numérique des sols par modélisation statistique (CSMS) : Bilan et perspective en France. [Vidéo en ligne]. Repéré à l'adresse suivante <https://vimeo.com/367222374>
- Lemercier, B., Lacoste, M., Loum, M., & Walter, C. (2012). Extrapolation at regional scale of local soil knowledge using boosted classification trees: A two-step approach. *Geoderma*, 171-172, 75-84. doi:10.1016/j.geoderma.2011.03.010
- Lima, E., Amaral Sobrinho, N., Paiva, F., Coutinho, I., Pereira, M., & Zonta, E. (2016). Quality reference values of trace elements in Brazilian organosols. *Environmental Monitoring and Assessment*, 188(7), 1-10. doi:10.1007/s10661-016-5436-7
- Liu, H., Janssen, M., & Lennartz, B. (2016). Changes in flow and transport patterns in fen peat following soil degradation. *European Journal of Soil Science*, 67(6), 763-772. doi:10.1111/ejss.12380
- Liu, H., & Lennartz, B. (2019). Hydraulic properties of peat soils along a bulk density gradient-A meta study. *Hydrological Processes*, 33(1), 101-114. doi:10.1002/hyp.13314
- Mann, K. L. (1986). Changes in soil carbon storage after cultivation. *Soil Science*, 142(5), 279-288. doi:10.1097/00010694-198611000-00006
- Milligan G.W., Cooper MC. (1985). An Examination of Procedures for Determining the Number of Clusters in a Data Set. *Psychometrika*, 50(2), 159-179.
- Minasny, B., & McBratney, A. B. (2016). Digital soil mapping: A brief history and some lessons. *Geoderma*, 264, 301-311. doi:10.1016/j.geoderma.2015.07.017
- Oberski D. (2016) Mixture Models: Latent Profile and Latent Class Analysis. In: Robertson J., Kaptein M. (eds) Modern Statistical Methods for HCI. Human–Computer Interaction Series. Springer, Cham. https://doi.org/10.1007/978-3-319-26633-6_12

- Päivänen J. (1973). Hydraulic conductivity and water retention in peat soils. *Acta Forestalia Fennica*, 129:1-70. <https://doi.org/10.14214/aff.7563>
- Parent, L-É. & C. Tremblay. (2003). Soil acidity Determination Methods for organic soils and peat materials. *Organic soils and peat materials for sustainable agriculture*. CRC Press, Boca Raton, Florida. pp. 93-104.
- Rudiyanto, R., Minasny, B., & Setiawan, B. I. (2016). Further results on comparison of methods for quantifying soil carbon in tropical peats. *Geoderma*, 269, 108–111. <https://doi.org/10.1016/j.geoderma.2016.01.038>
- Rycroft, D. W., Williams, D. J. A., & Ingram, H. A. P. (1975a). The transmission of water through peat: I. Review. *Journal of Ecology*, 63(2), 535–556.
- Rycroft, D. W., Williams, D. J. A., & Ingram, H. A. P. (1975b). The transmission of water through peat: II. Field experiments. *Journal of Ecology*, 63(2), 557–568.
- Samuel-Rosa, A., Heuvelink, G. B. M., Vasques, G. M., & Anjos, L. H. C. (2015). Do more detailed environmental covariates deliver more accurate soil maps? *Geoderma*, 243-244(C), 214-227. doi:10.1016/j.geoderma.2014.12.017
- Scrucca L, Fop M, Murphy TB, Raftery AE (2016). mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal*, 8(1), 289–317.
- Shein, E., Pozdnyakova, A., Shvarov, A., Il'in, L., & Sorokina, N. (2018). Hydrophysical Properties of the High-Ash Lowmoor Peat Soils. *Eurasian Soil Science*, 51(10), 1214-1219. doi:10.1134/S1064229318100113
- Therneau, T.M. & Atkinson E.J. (2019) An Introduction to Recursive Partitioning Using the RPART Routines. Mayo Foundation. <https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf>
- Thompson, D. K., & Waddington, J. M. (2014). A Markov chain method for simulating bulk density profiles in boreal peatlands. *Geoderma*, 232-234, 123-129. doi:10.1016/j.geoderma.2014.04.032
- Tiemeyer, B., Frings, J., Kahle, P., Köhne, S., & Lennartz, B. (2007). A comprehensive study of nutrient losses, soil properties and groundwater concentrations in a degraded peatland used as an intensive meadow – Implications for re-wetting. *Journal of Hydrology*, 345(1), 80-101. doi:10.1016/j.jhydrol.2007.08.002
- van Asselen, S., Stouthamer, E., & Smith, N. (2010). Factors controlling peat compaction in alluvial floodplains: A case study in the cold-temperate cumberland marshes, Canada. *Journal of Sedimentary Research*, 80(1-2), 155-166. doi:10.2110/jsr.2010.015
- Warner, B., & Rubec, C. (1997). Système de classification des terres humides au Canada. *Centre de recherche sur les terres humides, Université de Waterloo, Waterloo, Ontario, Canada*.
- Warren, M. W., Kauffman, J. B., Murdiyarso, D., Anshari, G., Hergoualc'h, K., Kurnianto, S., Purbopuspito, J., Gusmayanti, E., Afifudin, M., Rahajoe, J., Alhamd, L., Limin, S., & Iswandi, A. (2012). A cost-efficient method to assess carbon stocks in tropical peat soil. *Biogeosciences*, 9(11), 4477–4485. <https://doi.org/10.5194/bg-9-4477-2012>
- Weiss, R., Alm, J., Laiho, R., & Laine, J. (1998). Modeling moisture retention in peat soils. *Soil Science Society of America Journal*, 62(2), 305–313. <https://doi.org/10.2136/sssaj1998.03615995006200020002x>

- Ye, R., & Wright, A. L. (2010). Multivariate analysis of chemical and microbial properties in histosols as influenced by land-use types. *Soil & Tillage Research*, 110(1), 94-100. doi:10.1016/j.still.2010.06.013
- Zawieja, B., & Glina, B. (2017). Application of multivariate statistical methods in the assessment of mountain organic soil transformation in the central Sudetes. *Biometrical Letters*, 54(1), 43-59. doi:10.1515/bile-2017-0003
- Zeitz, J., & Veltz, S. (2002). Soil properties of drained and rewetted fen soils. *Journal of Plant Nutrition and Soil Science*, 165(5), 618-626. doi:10.1002/1522-2624(200210)165:5<618::AID-JPLN618>3.0.CO;2-W

Chapitre 2 – Groupes de gestion basés sur la profondeur de sol cultivable

2.1 Introduction

À la lumière des précédentes sections, il appert que plusieurs analyses physiques et chimiques soient nécessaires à la réalisation des différents objectifs de cette étude. De même, des profils de sol permettront d'obtenir de l'information sur la composition botanique et la morphologie du sol. Les différents échantillons pourront aussi être corrélés à la description pédologique.

2.2 Matériel et méthode

Cette section a pour but de décrire les analyses retenues ainsi que les traitements réalisés au laboratoire. La figure 11 résume les opérations effectuées au champ à l'été 2019 qui sont détaillées dans les sous-sections suivantes.

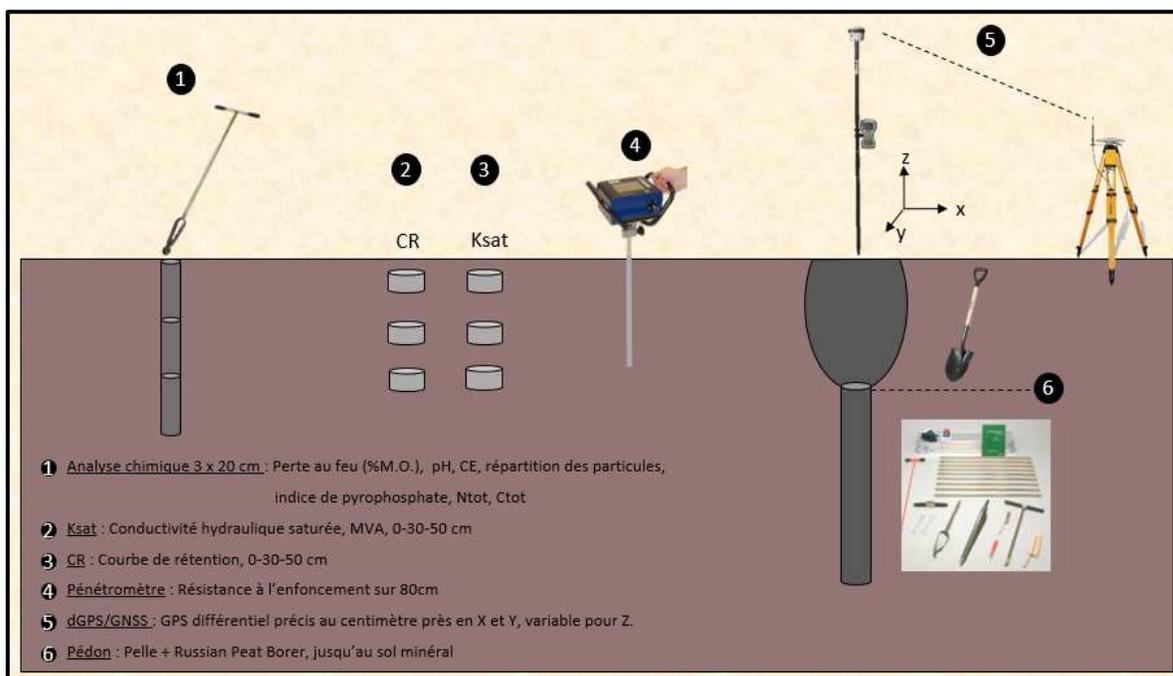


Figure 11 - Résumé des opérations et des types de prélèvements effectués chez 14 fermes partenaires à l'été 2019.

2.2.1 Sélection des sites

Les producteurs des 14 fermes partenaires de la chaire de recherche industrielle ont été consultés afin de classer empiriquement certains de leurs champs pour mieux répartir les pédons dans des zones à contraste. Pour ce faire, ils ont été appelés à diviser leurs champs en deux groupes selon leur qualité agronomique :

a) Bons champs (B) : Les champs de type B sont de bonne qualité. Ce sont eux qui offrent les meilleurs rendements et qui ne démontrent aucun problème particulier au niveau de leurs propriétés physiques et chimiques.

b) Mauvais champs (M) : Les champs de type M font état d'une qualité agronomique moindre en raison de problèmes de drainage, de fertilité ou de compaction par exemple.

Parmi les 120 sites échantillonnés, 73 étaient de type B et 47 ont été identifiés comme appartenant au type M. À certains moments, les sites M ne représentaient qu'un problème localisé dans une portion d'un champ, alors que pour d'autres sites échantillonnés, il représentait l'ensemble du champ. Ainsi, certains champs ont été échantillonnés à deux sites, l'un étant B et l'autre M.

Le nombre de site d'échantillonnage par ferme a été déterminé au prorata du financement de la ferme. Par conséquent, le nombre de site varie grandement d'une ferme à l'autre et il n'y a pas un ratio égal de site de type B et de type M pour chaque ferme. Le tableau 6 présente le nombre de sites visités par ferme.

Tableau 6 - Détails du nombre de sites d'échantillonnage par ferme ainsi que de leur qualité

Identifiant de la ferme	Nombre de sites	Type B	Type M
1	6	3	3
2	4	2	2
3	6	3	3
4	4	2	2
5	11	7	4
6	20	15	5
7	14	9	5
8	8	5	3
9	9	6	3
10	9	5	4
11	18	10	8
12	2	1	1
13	6	3	3
14	3	2	1
Total:	120	73	47

2.2.2 Caractéristiques pédologiques

Pour chacun des 120 sites, un profil de sol a été effectué. L'emplacement de chaque site a été enregistré à l'aide d'un récepteur GPS différentiel (GENEQ GNSS F90) et d'un collecteur de données GPS (SXPpad 1000P). La figure 12 permet de localiser tous les sites sur le territoire couvert.

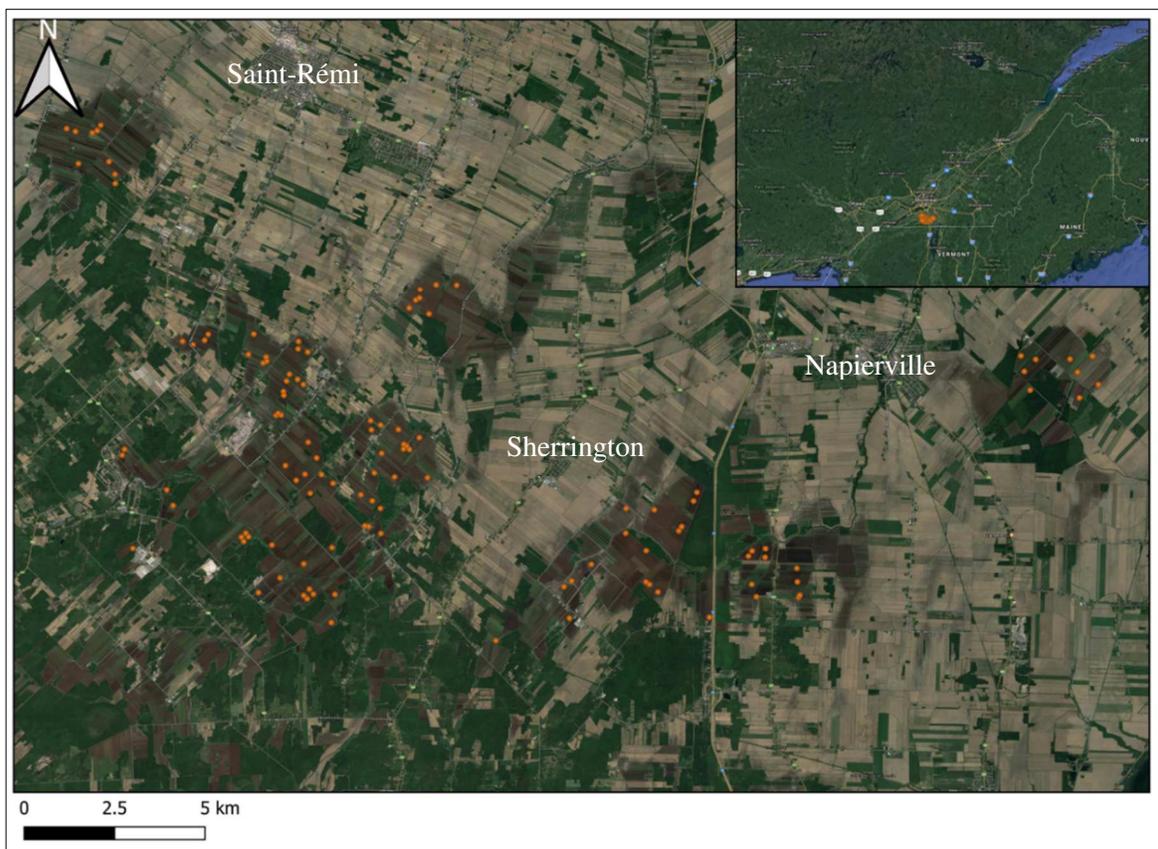


Figure 12 - Pédons réalisés en Montérégie-Ouest à l'été 2019. Les points en orange représentent les 120 sites.
Crédit d'image de fond : Google Satellite.

Un des paramètres mesurés était la profondeur jusqu'au sol minéral. Pour se faire, le pédon ne s'est pas limité à la profondeur de la coupe témoin en sol organique qui est de 160 cm (GTCS, 2002). Le profil de sol était débuté à la pelle jusqu'à environ 70 cm, puis une tarière Macauley (Eijkelkamp peat sampler) a été utilisée afin de sortir des carottes de sol de 50 cm jusqu'à atteindre le sol minéral.

Plusieurs données pédologiques ont été décrites à chaque site. Le tableau 7 résume, sous forme de liste succincte, la description pédologique effectuée pour chaque profil de sol :

Tableau 7 - Propriétés pédologiques étudiées lors de la réalisation du profil de sol

Propriété de la couche	Description
Profondeur	Limite supérieure et inférieure de la couche (en cm)
Désignation de l'horizon	Code abrégé de la couche
Texture	Texture de la couche (si la couche était minérale, la texture était évaluée à la main)
Fragments grossiers	Le pourcentage par couche et le type de fragment étaient précisés
Degré de décomposition	Évalué par la méthode des fibres frottées au champ
Composition botanique	Composition, en pourcentage, de la teneur en matériau amorphe, de mousse, de laïche et de bois
Famille de tourbe	Code qui correspond à la dominance de la composition botanique
Couleur de la matrice	La teinte, la saturation et la luminosité étaient déterminées
Marbrures	Détermination de la couleur. Ensuite, l'abondance, la dimension et le contraste des marbrures étaient évalués
Structure	Le type, la classe et le grade étaient spécifiés
Consistance	Évaluation manuelle de la résistance à la déformation/rupture
Effervescence	Évaluée seulement pour le coprogène et le sol minéral à l'aide d'une bouteille de HCl 10 %
Commentaire	Tout commentaire pertinent était noté

Le code Munsell (Munsell Color, 2010) a été utilisé afin de qualifier la couleur des couches et des marbrures. La description des profils de sols se base sur le Manuel de description des sols sur le terrain (Comité canadien d'experts sur la prospection, 1983) et le Guide canadien de classification des sols (GTCS, 2002).

À l'aide du guide canadien de classification des sols, le grand groupe et le sous-groupe des profils ont été déterminés et ajoutés à la base de données pédologique. Avec une fiche synthèse des séries de sols organiques de la région (André Martin, Lucie Grenon et Luc Lamontagne, communication personnelle), la série de sol a été identifiée puis ajoutée dans la base de données.

2.2.3 Analyses physiques

Pour 30 des 120 sites, seule une description pédologique a été réalisée puisque l'ensemble de l'échantillonnage a eu lieu en 2018. Ces sites ont été revisités en 2019 afin de diminuer l'échantillonnage.

2.2.3.1 Distribution des particules

La distribution des particules est un indice de l'érodibilité du sol en surface. Il est calculé à partir du sol prélevé entre 0-20 cm (voir section analyse chimique). La distribution des particules est obtenue en suivant une méthode adaptée de celle proposée par le CPVQ (1997b). Ensuite, la fraction érodable est caractérisée selon López et al. (2007). La méthode a été adaptée aux sols organiques en utilisant plutôt un tamis de 0,85 mm.

2.2.3.2 Conductivité hydraulique saturée

Au champ, trois cylindres en aluminium de 82 mm de diamètre et de 55 mm de hauteur sont enfoncés dans le sol en limitant la perturbation du sol. Ensuite, ils sont retirés à l'aide d'une pelle et le sol est creusé jusqu'à la profondeur suivante. Les cylindres sont pris à 0-5 cm, à 30-35 cm et à 50-55 cm. Un couteau est utilisé afin de couper l'excédent et les deux côtés du cylindre sont recouverts d'un tissu retenu par un élastique afin de contenir le sol et en conserver l'humidité. Le cylindre est ensuite transféré dans un sac fermé hermétiquement avant d'être transporté au laboratoire. Il sera conservé à 4°C en attendant la prise de mesures.

La méthode employée afin de mesurer la Ksat est celle définie par Carter & Gregorich (2007). Le logiciel Loggernet (version 3.4.1, Campbell Scientific, 2007) a permis l'acquisition des données et leur sauvegarde. Le calcul de la valeur de Ksat s'est fait dans Microsoft Excel.

2.2.3.3 Masse volumique apparente

La MVA est calculée avec les mêmes cylindres de sol utilisés pour la Ksat, une fois cette dernière mesurée. En laboratoire, la méthode de séchage et de calcul utilisée provient de Carter et Gregorich (2007). Cette méthode a été adaptée et l'équipe fait plutôt sécher les échantillons à 65°C. En effet, dans la littérature, il est possible de trouver des références suggérant 60°C (Vepraskas & Craft, 2015) ou une valeur entre 50°C et 80°C (Carter & Gregorich, 2007). Pour les sols organiques, le séchage à l'air est trop long, alors que le séchage traditionnel à 105°C est dommageable pour certaines propriétés et peut faire brûler la matière organique. Cela peut fausser les résultats.

2.2.3.4 Courbe de rétention en eau

La CR peut être mesurée de deux façons : en laboratoire sur des échantillons de sol non perturbé posés sur une table de billes de verre puis déposés dans une marmite, ou in situ à l'aide d'un tensiomètre et d'une sonde TDR (Brandyk et al., 2003). Dans les deux cas, deux paramètres sont évalués, soit la teneur en eau en fonction de la tension (potentiel matriciel du sol).

La méthode retenue est celle des cylindres de sols non perturbés pris côte-à-côte aux mêmes profondeurs que les cylindres destinés à la mesure de la Ksat/MVA. Leur manipulation jusqu'au laboratoire est identique.

Le montage est adapté de celui du CPVQ (1997a) pour les premiers points de la courbe sur une table de tension, mais des balances sont utilisées au lieu des marmites pour les derniers points de la courbe selon la méthode par évaporation (Šimůnek, Genuchten, & Wendroth, 1998). Le logiciel Loggernet est utilisé afin de faire l'acquisition des données et leur sauvegarde. Certains échantillons ont été placés à la marmite afin d'avoir une courbe complète pour les échantillons de 2018.

2.2.3.5 Résistance à l'enfoncement

Un pénétromètre numérique à cône statique de Eijkelkamp (Eijkelkamp Soil and Water, Giesbeek, Pays-Bas) a été utilisé. Il permet la détection de couches compactes en sondant les 80 premiers centimètres du sol. L'appareil produit une courbe de la RE en MPa en fonction de la profondeur en temps réel. De plus, l'appareil géoréférence chaque courbe à l'aide d'un récepteur GPS intégré. Il est possible de sauvegarder ces données sous forme de tableau et de les exporter par la suite à l'aide du logiciel PenetroViewer (version 6.08, Eijkelkamp, 2011).

À chaque site, cinq courbes ont été effectuées. Par la suite, la moyenne a été faite par centimètre afin d'obtenir une seule courbe par site. Pour finir, les données ont été filtrées à l'aide d'un script R (Guedessou, 2020) afin de corriger les biais de l'expérimentateur.

2.2.4 Analyses chimiques

Quant aux analyses chimiques, trois sacs de sols étaient remplis à chaque site à l'aide d'une tarière hollandaise. Le premier sac était constitué de sol de 0-20 cm de profondeur, le deuxième sac de sol de 20-40 cm de profondeur et le troisième sac de sol de 40-60 cm de profondeur. Ces sacs étaient ensuite fermés hermétiquement et stockés dans une pièce fraîche en attendant d'être apportés au laboratoire à Québec. En attendant d'être traités, les échantillons sont laissés au réfrigérateur à 4°C.

Toutes les analyses chimiques requièrent une préparation. Le sol doit être broyé et tamisé à 2 mm au préalable. La méthode du CPVQ (1988) est utilisée.

2.2.4.1 Perte au feu

Cette méthode permet d'obtenir le pourcentage de matière organique d'un sol. La perte au feu a été calculée selon une méthode adaptée du Centre d'expertise en analyse environnementale du Québec (MDDELCCQ, 2017). Les échantillons de 10 g sont mis dans de petits béchers puis identifiés. Ils sont ensuite déposés dans un four à moufles (modèle FA1730, Barnstead International) muni d'un contrôleur de température Furnatrol II (type 53600, Barnstead International).

2.2.4.2 Pyrophosphate de sodium

Cette analyse chimique permet d'évaluer le degré de dégradation du sol. La méthode utilisée est celle de Carter et Gregorich (2007). Le spectrophotomètre est un UV-VIS de type Cary 100 Bio, opérant avec le logiciel CarywinUV (Agilent Technologies Inc., version et année non disponibles).

2.2.4.3 Azote total et carbone total

L'azote totale et le carbone total sont dosés par combustion LECO sur un appareil CN 828 (LECO, Michigan, États-Unis) selon le protocole du fabricant (Bernard Montminy, communication personnelle). Cette analyse est réalisée par le laboratoire de l'IRDA.

2.2.4.4 pH eau

La méthode du pH à l'eau est utilisée selon le CEAEQ (2014). Cependant, le ratio sol-solution utilisé est 1:5 (10 cc de sol équivaut à environ 5 g, donc 5 g de sol pour 25 ml) selon Carter et Gregorich (2007). En effet, les sols contenant beaucoup de matière organique ont tendance à former une pâte au lieu d'une solution liquide ce qui nuit la mesure du pH. L'appareil utilisé est un pH-mètre symphony SB70P (VWR, Pennsylvanie, États-Unis).

2.2.4.5 CE

La conductivité électrique du sol est une mesure de sa salinité. La méthode employée est identique à celle du pH (CEAEQ, 2014) et le ratio 1:5 a aussi utilisé (Carter & Gregorich, 2007). L'appareil utilisé est un symphony SB70C (VWR, Pennsylvanie, États-Unis).

2.2.5 Statistiques et analyses cartographiques

Les analyses et manipulations spatiales sont réalisées avec le logiciel QGIS (QGIS Development Team, 2020). Les analyses statistiques sont réalisées dans R (R Core Team, 2020), tout comme la cartographie numérique abordée au chapitre 3.

2.3 Résultats

2.3.1 Analyse réursive de l'association entre la qualité et les propriétés des sols

Afin d'évaluer la corrélation entre les propriétés mesurées et la variable de qualité d'un site, une analyse réursive a été réalisée sur 114 sites, une fois les données manquantes retirées. Les sites échantillonnés sont ainsi classés en fonction de leur qualité, variable prédite grâce à la meilleure sélection possible des propriétés étudiées. L'analyse permet de retirer les propriétés redondantes et ainsi éviter la multicolinéarité. La librairie *rpart* de l'anglais « Recursive PARTitioning » (Therneau & Atkinson, 2019) a été utilisée dans le logiciel R. Voir la section 1.4 pour plus de détails concernant le fonctionnement de cette analyse.

La figure 13 présente l'arbre produit par l'analyse. Cet arbre est élagué et contient un nombre de variables optimal selon le modèle. Cinq variables sur 61 ont été conservées afin de prédire efficacement la qualité du sol. L'exactitude globale de la prédiction est de 79 % (93% pour les bons champs et 56% pour les mauvais champs). À l'aide de 5 000 bootstraps, les trois meilleures racines de l'arbre ont été déterminées, soit MO_1, MO_3 et FIB, bien qu'il y ait une seule racine par arbre. Ces trois propriétés permettraient la meilleure classification initiale des sites en fonction de leur qualité lorsqu'une d'entre elles est utilisée dans l'arbre réursif. Ces trois variables sont liées à la nature botanique du site et à la durée d'exploitation et seraient donc étroitement liées à la qualité des sites.

Dans cet arbre se trouvent des variables liées à la chimie, à la physique/l'hydrologie et à la pédologie du sol. Il semble donc y avoir une corrélation significative entre les trois types de données et la qualité du sol. Le faible nombre de variables laisse présager que certaines sont meilleures que d'autres afin de classer un site en fonction de sa qualité. Cela était attendu, avec une matrice de données aussi volumineuse, la multicolinéarité était probable.

En s'attardant aux seuils critiques suggérés par l'arbre, il est possible de constater quelques divergences avec les corrélations attendues entre certains paramètres et leur qualité. Plus en détails, une *ksat_1* plus faible et un pourcentage supérieur de sol de surface de diamètre < 0,84 mm mènent en majorité à des sols de bonne qualité. En pratique, une corrélation inverse serait attendue. Ces résultats suggèrent un biais possible quant à la variable de qualité. De plus, la grande variabilité de la *ksat* à l'échelle du champ peut compliquer la généralisation des résultats (Gnatowski et al., 2010).

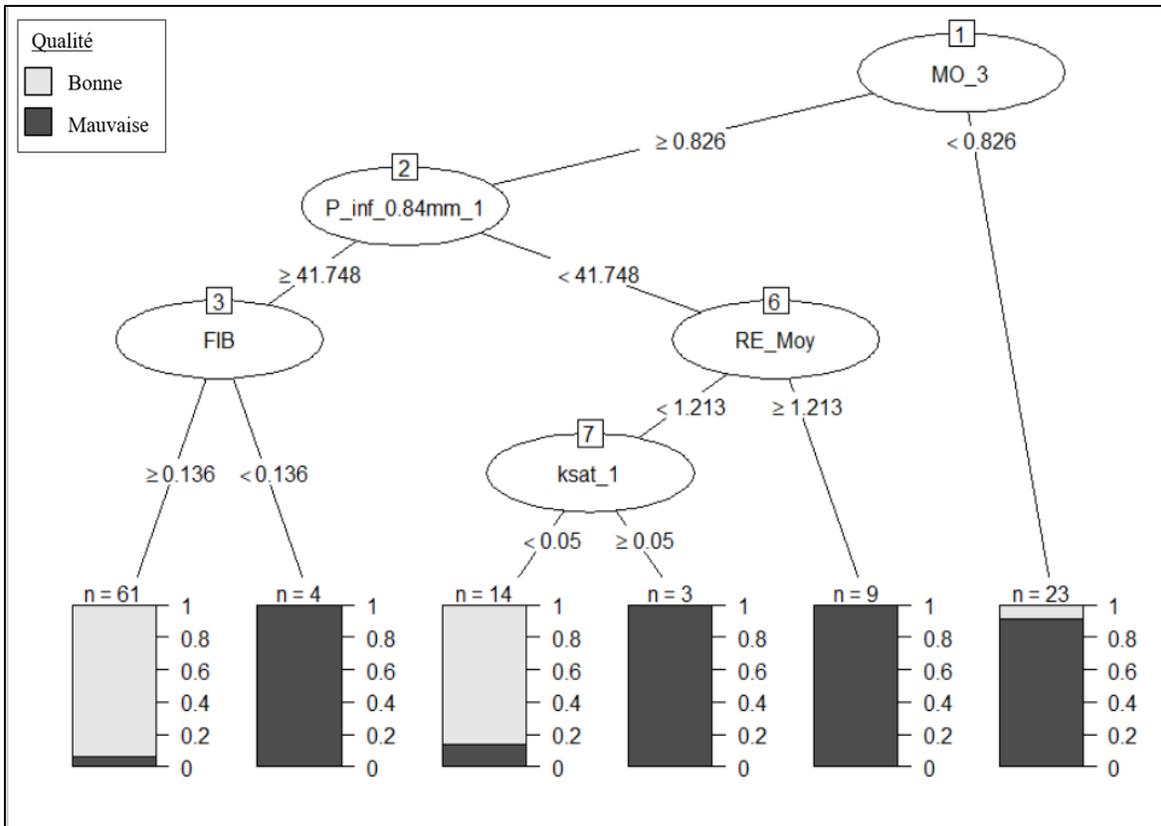


Figure 13 - Arbre récurif issu de la classification de 114 sites en fonction de leur qualité, soit 71 sites de bonne qualité et 43 de mauvaise qualité. Chaque branche mène à une feuille contenant un certain nombre de sites (n). Les bons sols sont représentés par la couleur grise et les mauvais sols par la couleur noire. À droite des feuilles se trouve une échelle permettant d'évaluer la proportion des sites appartenant à chacune des qualités. Une feuille d'une seule couleur indique une classification parfaite alors qu'un mélange de couleurs indique qu'une proportion des sites a été mal classée. On peut y voir cinq variables, soit le pourcentage massique de matière organique entre 40-60 cm (MO_3), le pourcentage massique de sol de surface de diamètre inférieur à 0,84 mm (P_inf_0.84mm_1), la résistance à l'enfoncement moyenne (RE_Moy), le pourcentage volumique pondéré de fibres non frottées (FIB) et la conductivité hydraulique saturée à 0-5 cm (ksat_1).

En effet, un nombre inégal de sites ont été visités par ferme. Il est possible que les fermes ayant une plus grosse superficie et étant échantillonnées davantage aient plus de variabilité dans la qualité de leurs sols. Il serait alors plus aisé de séparer les bons champs des mauvais champs. À l'inverse, dans les plus petites fermes, les champs se ressemblent davantage. De plus, l'interprétation de la qualité d'un sol peut varier d'une ferme à l'autre. Afin de valider cette hypothèse, une deuxième analyse récurif a été réalisée en utilisant seulement 80 sites provenant des six fermes ayant au moins 9 sites d'échantillonnage (figure 14).

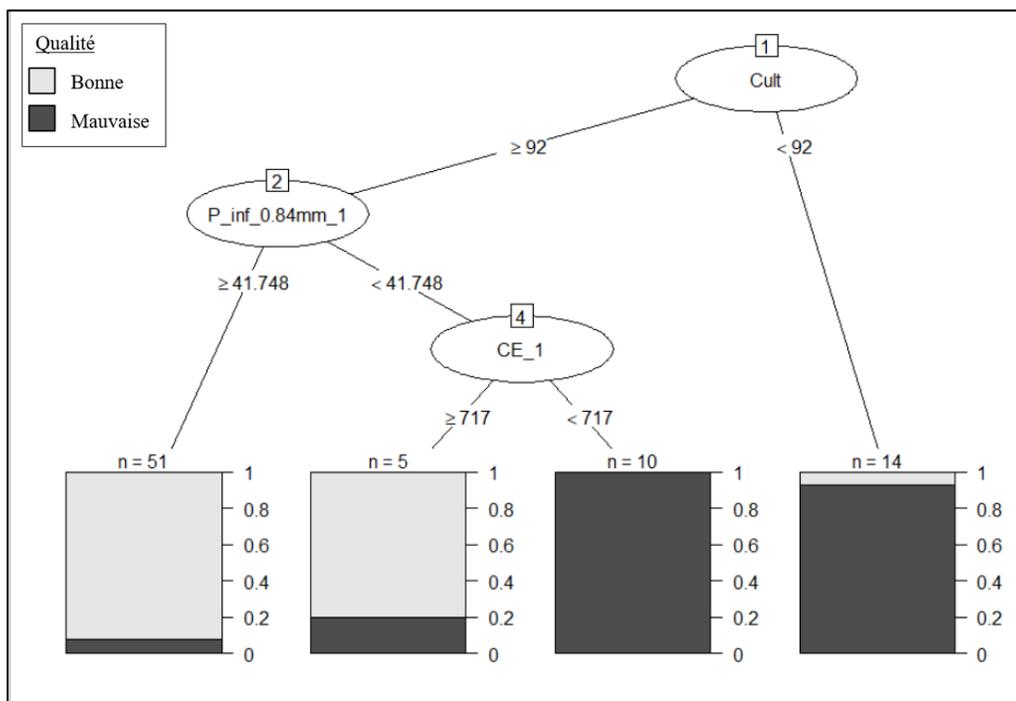


Figure 14 - Arbre récuratif issu de la classification de 80 sites en fonction de leur qualité. On peut y voir trois variables, soit la profondeur de sol cultivable (Cult), le pourcentage massique de sol de surface de diamètre inférieur à 0,84 mm (P_inf_0.84mm_1) et la conductivité électrique de la couche 0-20 cm (CE_1).

Dans ce deuxième arbre, une légère augmentation de l'exactitude est observée afin d'atteindre 80 % (68% pour les mauvais sites et 87% pour les bons sites, figure 14). En comparaison au premier arbre, on remarque la répétition de P_inf_0.84mm_1, mais les deux autres variables semblent différer a priori. Cependant, l'analyse en composantes principales réalisée à la prochaine section révèle que Cult est corrélé à MO_3 ($r = 0,54$), ce qui signifie qu'une partie des arbres est corrélée.

Il est intéressant de constater que le seuil associé à P_inf_0.84mm_1 est contre-intuitif, c'est-à-dire qu'une quantité plus grande de sol érodable en surface semble liée aux sols de meilleure condition. En ce sens, il est important de noter que l'horizon de surface est presque toujours humique (Ohp, relativement dégradé), et ce, peu importe la classification pédologique du reste du profil de sol.

Bien que les arbres soient en partie corrélés et que le nombre de variables ait été réduit, il y a perte de robustesse quant à la prédiction de la qualité d'un site. Cela semble supporter l'hypothèse qu'un biais existe chez les plus petites fermes ayant moins de sites. Il faut alors limiter la portée des conclusions. En ce sens, les variables discriminantes et les seuils critiques proposés par l'analyse ne seront pas interprétés ni utilisés afin de former des groupes de gestion.

Comme mentionné précédemment, cela peut être dû à la définition même de la qualité et de la productivité agricole qui peut varier d'une entreprise à l'autre. De même, le nombre et la distribution des sites n'étaient pas optimaux. Le dispositif d'échantillonnage (« sampling design ») n'était pas assez rigoureux. Le critère de la qualité d'un site n'a pas été mesuré directement à l'aide de mesures de rendement, mais indirectement en consultant les producteurs. L'approche utilisée ne permet que de faire des inférences en lien avec la qualité relative des champs d'une ferme à l'autre. Cela introduit nécessairement de la variabilité dans les indicateurs de qualité qui ont été explorés. Aucun groupe de gestion des sols ne peut être formé avec cette approche. Avant d'invalider l'hypothèse 1, l'approche des grands groupes de sol est tentée.

2.3.2 Analyse récursive de l'association entre le grand groupe et les propriétés des sols

En guise de validation interne, l'association entre les propriétés évaluées et le grand groupe de chaque site a été étudiée. Cela permettra de valider la classification des profils de sol en plus de voir si certaines tendances se dégagent des relations entre les propriétés. En d'autres mots, il serait possible que les grands groupes puissent servir à former des groupes de gestion des sols. Avant de traiter des résultats de l'analyse récursive sur les grands groupes, voici une brève comparaison entre les résultats de la description pédologique de cette étude comparée à l'étude du comté de Napierville.

Les études pédologiques de la région ont été consultées afin de déterminer si l'échantillonnage était représentatif des inventaires réalisés par le passé. Il faut savoir que les 5 bassins s'étendent dans quatre études pédologiques : les études des comtés de Châteauguay (1950), de Laprairie (2000), de St-Jean (2001) et de Napierville (2014). La majorité de la superficie des tourbières se trouve dans les études de Napierville et de Châteauguay. Comme les propriétés des sols organiques évoluent rapidement (Kempen et al., 2009) et qu'il existe une grande disparité entre les études quant à leur année de réalisation, seules les données du comté de Napierville sont utilisées à des fins exploratoires.

Selon le rapport pédologique du comté de Napierville (Lamontagne et al., 2014), on compte 6 234 ha de sol organique sur le territoire (16,8 % du comté). Le territoire compte 3,2 % de fibrisols, 18,2 % de mésisols et 78,6 % d'humisols selon le rapport pédologique. L'échantillonnage de 2019 rapporte plutôt 0,9 % de fibrisols, 33,3 % de mésisols et 65,8 % d'humisols si l'on ne conserve que les 117 profils de sols organiques.

Quant aux sous-groupes, l'étude de Napierville rapporte que 60,5 % des superficies en sol organique sont terriques (2,8 % attribuables aux mésisols et 57,7 % aux humisols). En rappel, ce sous-groupe correspond aux sols organiques minces ayant la limite supérieure du sol minéral à l'intérieur de la coupe témoin de 1,6 m. Quant aux sols limniques (présence de terre coprogène d'une épaisseur minimale de 5 cm dans le premier 1,6 m), l'étude rapporte 11,4 % de la superficie et uniquement des mésisols limniques. Ces données diffèrent considérablement de celles de cette étude. Il faut toutefois comprendre qu'au niveau de la classification canadienne, dès lors qu'un sol est terrique, même s'il est aussi limnique, le sous-groupe terrique domine (Kroetsch et al., 2011). Le sous-groupe limnique devient alors une variante. Cela peut expliquer une divergence quant aux données de l'étude et celles de ce projet. En effet, plusieurs sites terriques et limniques ont été observés, mais l'information ne peut être extraite aussi facilement du rapport pédologique. Dans cette étude, 18 sites sur 120 étaient terriques et limniques.

Les écarts entre les études peuvent être attribuables à des erreurs d'évaluation des fibres frottées, caractère relativement subjectif, ou encore au patron d'échantillonnage qui diffère substantiellement de celui du rapport pédologique. Néanmoins, la suite des analyses n'est pas compromise, puisque la description des profils en 2019 n'est pas en elle-même aberrante.

Tel que discuté à la section 1.1, la différence entre les trois grands groupes de sols organiques retrouvés au Québec se situe au niveau des fibres frottées et du degré d'humification de l'étage intermédiaire (entre 40 et 120 cm). Une analyse réursive sur les grands groupes se fonde sur le fait que ces derniers soient intrinsèquement liés à l'état de dégradation des sols organiques (GTCS, 2002).

Dans ce nouvel arbre réursif, le fibrisol (seul de sa catégorie) a été retiré afin de conserver uniquement les humisols (75 sites) et les mésisols (38 sites). Cela simplifie l'analyse en passant de trois à deux groupes à prédire. L'arbre produit est présenté à la figure 15. L'exactitude du classement global est de 68% (83% pour les humisols et 39% pour les mésisols). Un grand nombre de variables semble nécessaire, malgré une performance modérée du modèle. Les fibres frottées et le pyrophosphate permettent à eux seuls de classer 42% des sites. Cela confirme l'utilisation de ces variables dans la littérature (GTCS, 2002; Carter & Gregorich, 2007).

Il est possible de remarquer une corrélation entre les variables, particulièrement au niveau du pH, de la ksat et du coprogène. Les seuils respectent les comportements attendus des grands groupes. Mentionnons que la variable liée au coprogène revient à deux reprises, ce qui peut signifier un effet quadratique ou un sur-ajustement du modèle puisque les classes produites contiennent peu d'observations. La teneur en eau à 3 kPa en surface est plus importante pour les humisols. Cela rejoint les observations de Brandyk et al. (2003) qui mentionnent

une redistribution porale avec l'affinement de la tourbe en particules. La macroporosité laisse place à une microporosité plus importante et à une rétention en eau plus importante.

Le nombre important de variables conservées dans le modèle et l'exactitude modérée peuvent être liés à la façon dont les groupes sont formés. Le critère de base de distinction entre les grands groupes est le pourcentage de fibres frottées des couches dominantes entre 40 et 120 cm. Par conséquent, un mésisol avec un étage à 11% de fibres frottées peut être similaire à un humisol à 8% de fibres frottées, bien que leur grand groupe soit différent. En ce sens, cela aura très peu d'impact sur bien des paramètres. Cela ajoute donc une variabilité que le modèle récursif basé sur un seuil de classification binaire n'arrive pas à décrire.

En résumé, le modèle sous-entend la présence d'une corrélation entre des variables physiques, chimiques et pédologiques, et le grand groupe des sites sondés. On retrouve certaines variables aussi observées à la section précédente. Cependant, bien que la corrélation existe, l'utilisation des grands groupes comme groupes de gestion des sols est à proscrire. Comme mentionné plus haut, l'évolution d'un mésisol vers un humisol se fait progressivement bien que les mésisols puissent partager des propriétés identiques en surface et parfois en profondeur. Le grand groupe n'arrive pas à lui seul à considérer les différences dans la décomposition des fibres et leur origine (Gnatowski et al., 2010; Vepraskas & Craft 2015), la porosité et la compaction (Brandyk et al., 2003; Okruszko & Ilnicki, 2003) qui sont au cœur de l'évolution et de l'humification de ces sols.

Pour toutes ces raisons, l'utilisation du grand groupe ne permet pas de former des groupes de gestion liés à la dégradation des sols organiques cultivés selon le contexte de l'étude. Concernant la variable de qualité, son caractère qualitatif se prête mal à l'étude et génère un biais difficile à évaluer avec cette approche. Cependant, les résultats ont démontré certaines relations intéressantes (par exemple FIB et MO) ayant un bon pouvoir prédictif et discriminatoire qui pourraient être explorées davantage. **L'hypothèse 1 est donc rejetée.** La formation de groupes de gestion de conservation des sols devra être basée sur une autre approche plus complexe et être capable de caractériser la variabilité des sols en sélectionnant un nombre plus restreint de propriétés.

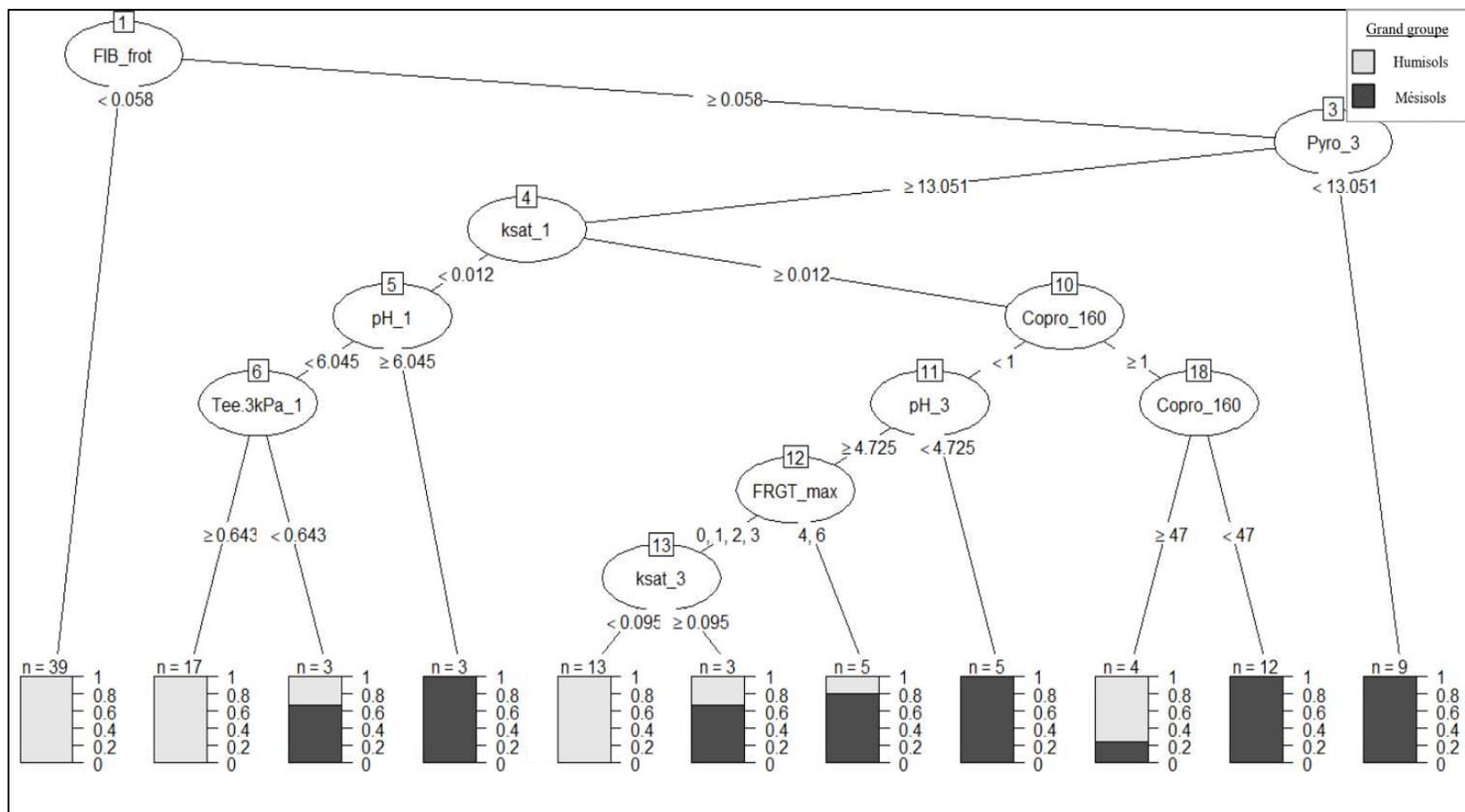


Figure 15 - Arbre récursif issu de la classification de 113 sites en fonction de leur grand groupe, soit 75 humisols et 38 mésisols. On peut y voir neuf variables, soit le pourcentage de fibres frottées moyen du profil pondéré par l'épaisseur de la couche (FIB_frot), l'indice au pyrophosphate de la couche 40-60 cm (pyro_3), la conductivité hydraulique saturée à 0-5 cm et à 50-55 cm (ksat_1 et ksat_3), le pH de la couche 0-20 cm et 40-60 cm (pH_1 et pH_3), la teneur en eau volumique à 3 kPa à 0-5 cm (Tee.3kPa_1), l'épaisseur de coprogène présent dans le premier 1,6 m de sol (Copro_160) et le contenu en fragments grossiers de la couche qui en contient le plus dans le premier mètre de sol (FRGT_max).

2.3.3 Approche par modèle de mélange

Deux approches intuitives ont été évaluées afin de trouver des propriétés discriminantes et former des groupes de gestion sans succès tel que vu à la section précédente. Ensuite, une approche basée sur la réduction de dimensions a été tentée à l'aide d'une analyse en composantes principales. Les variables conservées ont été introduites dans un modèle de mélange (Mplus version 6.1, Muthén & Muthén, 2010) qui n'a pas convergé. Cela signifie que des maximums locaux dans l'espace des données peuvent fausser la génération des classes latentes. Les résultats ne sont pas fiables. Le test de Lo-Mendell-Rubin rapporte un p-value = 0,1045, signifiant que deux classes ne sont pas significativement différentes d'une seule. En d'autres mots, selon cette analyse, il ne semble pas exister de groupes cachés dans le jeu de données.

2.3.4 Approche par profondeur de sol cultivable

Une dernière approche plus objective a été évaluée, cette fois-ci basée sur la profondeur de sol cultivable. Étant inspirée de celle d'Hallema et al. (2015a), cette dernière approche vise à segmenter le jeu de données en groupes de profondeur différentes. Plusieurs seuils ont été successivement utilisés comme variable indépendante dans une MANOVA afin de voir s'il existe une profondeur optimale à partir de laquelle les propriétés du sol (les variables dépendantes) varient significativement.

La profondeur cultivable se définit, dans le cadre de ce projet, comme étant l'épaisseur de sol tourbeux au-dessus du sol minéral ou au-dessus d'une couche de coprogène lorsque présente. À titre d'exemple, un sol ayant 128 cm de sol tourbeux, 25 cm de coprogène et le sol minéral à 153 cm de profondeur aurait une profondeur cultivable de 128 cm. Il est ensuite possible de mettre en relation chacune des propriétés mesurées au champ avec la profondeur cultivable des sites.

L'hypothèse derrière cette approche est que les sols moins profonds vont montrer des signes de dégradation plus importants, puisqu'ils sont sujets à des pertes de sols importantes et à une décomposition marquée sur le long terme. Il faut cependant noter que ces sites peuvent être simplement associés à des dépôts plus minces. En guise d'analyse exploratoire, les figures 16 et 17 mettent en relation une sélection de variables étudiées et la profondeur cultivable. Des relations s'approchant de fonctions exponentielles inverse et inverse permettent d'observer un changement de comportement des propriétés vers 50 cm de profondeur.

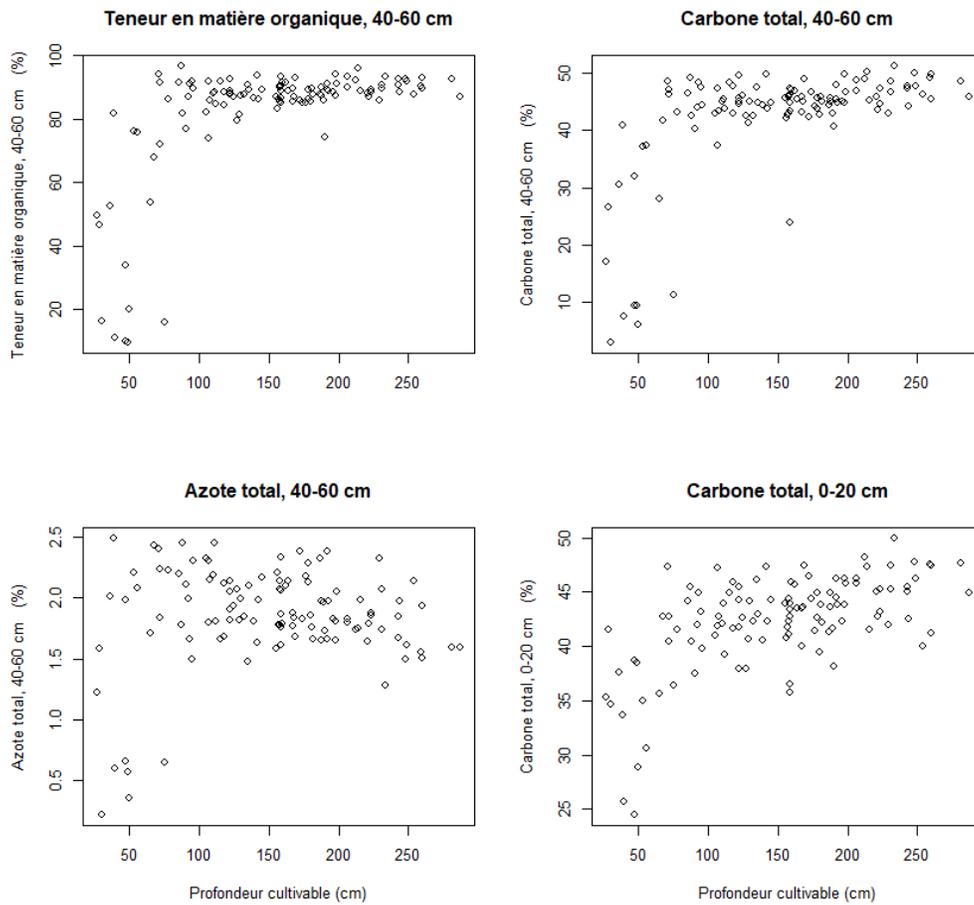


Figure 16 – Relations exploratoires de type exponentielle inverse entre quatre propriétés et la profondeur cultivable.

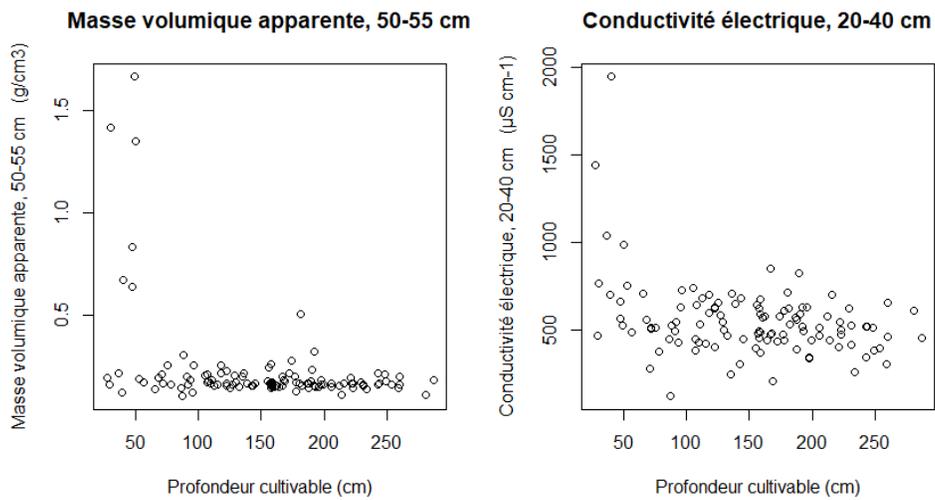


Figure 17 - Relations exploratoires de type fonction inverse entre deux propriétés et la profondeur cultivable.

Le lecteur aura remarqué que certaines observations ont une profondeur cultivable plus faible que la profondeur d'échantillonnage. En effet, les trois profondeurs d'échantillonnage ont été fixées, peu importe la profondeur cultivable d'un site. Cette décision se fonde sur le fait que les racines des cultures maraîchères s'implantent surtout dans ces couches de sol. Il est donc pertinent d'évaluer la MVA à 50-55 cm même si cette couche est composée de sol minéral ou de coprogène de manière à pouvoir quantifier le changement des propriétés par rapport à un sol ayant 3 m de tourbe.

Il est aussi possible de remarquer que certaines propriétés mesurées plus près de la surface (0-5 cm ou 0-20 cm) montrent aussi des signes de dégradation du sol même s'il s'agit de sol tourbeux. Cela confirme les tendances observées graphiquement, bien qu'il soit possible que les effets observés soient associés aux profondeurs d'échantillonnage pour les échantillons plus profonds. Les analyses physiques sont faites à un maximum de 55 cm alors que les analyses chimiques sont faites à un maximum de 60 cm. Un échantillonnage à 70 ou 80 cm aurait peut-être permis d'observer d'autres signes de dégradation, mais le chapitre 1 a permis de conclure que le phénomène de dégradation se produit surtout en surface, dans les premiers 30 cm. Il aurait été plus laborieux et difficiles à justifier d'échantillonner à des profondeurs plus importantes.

Il est important de remarquer que cette exploration est faite en mode univarié et que ces relations seront peut-être différentes en mode multivarié. De plus, seules quelques variables ont été présentées ici puisque la majorité ne montraient aucune relation appréciable visuellement. Néanmoins, ces figures supportent l'approche de la profondeur cultivable qui sera détaillée dans les prochaines sous-sections.

2.3.4.1 Préparation des données

La matrice contenait initialement 66 variables, mais certaines jugées moins pertinentes ont été retirées, dont les catégorielles et binaires, afin d'en conserver seulement 53. Ensuite, après un examen des données, un grand nombre d'entre elles avait des valeurs de dissymétrie (« skewness ») et de voussure (« kurtosis ») dépassant 2 (Tabachnick & Fidell, 2013). Afin de résoudre ce problème, la librairie R *bestNormalize* (Peterson, 2019) a été utilisée afin de trouver la meilleure transformation pour chacune des variables. Un tableau définissant les acronymes des variables se trouve à l'annexe 1, alors qu'un tableau contenant les statistiques descriptives des 53 variables est présenté à l'annexe 2.

Afin de s'approcher du ratio VD:N supérieur à 1:10 et de diminuer la multicolinéarité entre les variables (Tabachnick & Fidell, 2013), une analyse en composantes principales (ACP) a été réalisée sur la matrice de corrélation des variables avec la librairie *psych* (Revelle, 2020). Une rotation varimax a été utilisée. Concernant

les postulats, la normalité et la factoriabilité de la matrice de corrélation sont respectées. Les postulats en lien avec le nombre d'observations et la multicollinéarité ne sont pas respectés, mais les résultats de cette étape préparatoire ont justement pour but d'assurer le respect de ces deux postulats pour la MANOVA. L'ACP est donc utilisée comme processus de sélection de variables et non pas comme une fin en soi. Le non-respect de ces postulats peut être toléré dans ce cas.

Au terme de l'ACP, 10 composantes ont été retenues, chacune ayant une valeur propre > 1. Cela signifie que chaque composante permet d'expliquer au moins autant de variance qu'une variable brute. Pour chacune de ces composantes, la variable qui y était la plus corrélée, sans l'être à aucune autre composante, était conservée. Les dix variables sont les suivantes : la teneur en carbone total et en azote total (0-20 cm), le pH (20-40 cm), la teneur en eau à 3 kPa (0-5 cm), la réserve en eau utilisable à 10 kPa (0-5 cm), la teneur en air à 5 kPa (30-35 cm et 50-55 cm), la porosité totale (50-55 cm), la profondeur de la couche compacte et la résistance à l'enfoncement maximale (0-80 cm). Elles représentaient 73,3 % de la variance du jeu de données initial. La matrice finale utilisée pour la MANOVA comporte donc 10 variables faiblement corrélées correspondant à 73,3 % de la variance du jeu de données original. Le test de Mahalanobis (valeur critique du χ^2 à $\alpha = 0,001$ pour 10 d.l.l.) a révélé la présence d'aucune donnée extrême multivariée.

2.3.4.2 Détermination d'un seuil de profondeur cultivable optimal à l'aide de MANOVAs

Afin de déterminer le meilleur seuil de profondeur, des analyses multivariées de variance (MANOVAs) ont été réalisées sur les dix variables issues de l'ACP (les variables dépendantes). La variable indépendante était l'appartenance d'un site à un groupe de profondeur (groupe 1 ou groupe 2). Chaque groupe était formé à l'aide d'un seuil qui était successivement augmenté (50 cm, 60 cm, ..., 120 cm et 130 cm) afin de couper en deux le jeu de données. Le tout a été réalisé dans R avec la fonction `manova` (R Core Team, 2020). Un total de 114 observations étaient disponibles pour l'analyse. Le seuil d'alpha = 0,05 a été utilisé.

Tableau 8 - Série de MANOVAs avec un seuil incrémenté par tranches de 10 cm. Le tableau rapporte aussi le nombre d'observations dans chaque groupe (N1/N2), la trace de Pillai ainsi que les bornes de son intervalle de confiance de 95% et le nombre d'ANOVAs significatives sur 53 entre les groupes au seuil testé.

Seuil testé (cm)	40	50	60	70	80	90	100	110	120	130
N1/N2	5/109	9/105	12/102	14/100	19/95	22/92	27/87	31/83	37/77	44/70
Trace de Pillai	0,260	0,351	0,495	0,498	0,431	0,382	0,387	0,363	0,337	0,358
Pillai limite inf. (2,5 %)	0,142	0,236	0,362	0,382	0,308	0,275	0,294	0,289	0,264	0,293
Pillai limite sup. (97,5 %)	0,498	0,585	0,675	0,684	0,636	0,593	0,593	0,556	0,534	0,537
ANOVAs significatives /53	6	18	18	17	15	13	14	14	13	16

Note : Toutes les MANOVAs sont significatives, $p < 0,001$.

Au terme des analyses, le seuil retenu a été 60 cm, divisant le jeu de données en deux groupes de 12 et de 102 observations (tableau 8). Puisque toutes les MANOVAs du tableau 8 se sont révélées significatives, le choix du seuil optimal s'est fondé sur la trace de Pillai (valeur élevée désirée) et du nombre d'ANOVAs significatives pour un seuil donné (valeur élevée désirée). En effet, à la suite de la MANOVA, des ANOVAs se sont intéressées au nombre de propriétés qui diffèrent significativement entre les groupes (détails plus bas). Les VDs combinées sont significativement affectées par le groupe de profondeur, $F(10, 103) = 10,099$, $p < 0,001$. Pour le seuil de 60 cm, le postulat d'homogénéité des matrices de variance-covariance est rejeté (Box's M-test : $p < 0,05$). La normalité multivariée est respectée pour le kurtosis et est rejetée pour la dissymétrie. L'absence de multicollinéarité entre les variables est assumée.

Afin de considérer ce manque de précision, l'intervalle de confiance de la trace de Pillai est fournie. L'utilisation de la trace de Pillai est conseillée lorsque des postulats de la MANOVA sont rejetés (Olson, 1974). Cette statistique peut prendre une valeur entre 0 et 1, où une valeur plus élevée reflète un effet plus grand de la variable explicative sur les variables réponses. Les résultats reflètent une association modérée entre le groupe de profondeur et la combinaison des VDs, trace de Pillai = 0,495 avec un intervalle de confiance de 95% obtenu par 5 000 bootstraps compris entre 0,36 et 0,68. En bref, un bootstrap utilise les observations du jeu de données initial avec remise pour générer aléatoirement de nouvelles populations desquelles des MANOVAs seront produites. La trace de Pillai sera extraite pour chaque MANOVA et l'intervalle de confiance peut ainsi être généré en agrégeant l'information de ces 5 000 itérations.

Le seuil de 60 cm est aussi supporté par d'autres résultats. En effet, Caron et al. (2021, communication personnelle) ont montré par simulation qu'un sol organique cultivé montre une baisse importante de l'efficacité d'un système de drainage dès qu'il atteint une profondeur d'environ 50 cm. De plus, un sol cultivé doit avoir une profondeur minimale de 40 cm afin de faire partie de l'ordre organique si on assume qu'il s'agisse d'un humisol ou d'un mésisol (GTCS, 2002). Lorsque l'on combine ces deux informations aux graphiques exploratoires présentés plus haut, il est possible de voir une certaine convergence autour du seuil de 60 cm proposé.

Une fois le seuil de 60 cm retenu, une deuxième série de MANOVAs a été réalisée dans le but de tester s'il serait préférable de former trois groupes de gestion en fonction de la profondeur cultivable. Le tableau 9 rapporte les résultats de cette deuxième série de MANOVAs. Tous les seuils étaient significatifs, mais la trace de Pillai (0,295 à 0,334 : tableau 9) était dans tous les cas inférieure à celle du seuil retenu pour la MANOVA à deux groupes. Cela dit, il est nécessaire de considérer un deuxième seuil selon une approche pratique, à défaut d'être supportée par une approche statistique. En effet, si le seuil de 60 cm répond à la question : « à partir de quelle

profondeur le sol est-il suffisamment dégradé afin de voir ses propriétés modifiées ? », il ne répond pas à la question : « à partir de quelle profondeur doit-on commencer à mettre en place des mesures de conservation des sols ? ». Comme les mesures de conservation ont un coût associé, il est impératif d'appliquer ces mesures lorsque le sol est encore de bonne qualité et avant qu'il n'atteigne une profondeur de 60 cm.

Tableau 9 - Série de MANOVAs avec deux seuils, le premier fixé à 60 cm et le second incrémenté par tranches de 10 cm. Le tableau rapporte aussi le nombre d'observations dans chaque groupe ($N1/N2$), la trace de Pillai ainsi que les bornes de son intervalle de confiance de 95% et le nombre d'ANOVAs significatives sur 53 entre les groupes au seuil testé.

Seuil testé (cm)	60/80	60/90	60/100	60/110	60/120	60/130	60/140	60/150	60/160	60/170
$N1/N2/N3$	12/7/95	12/10/92	12/15/87	12/19/83	12/25/77	12/32/70	12/37/65	12/40/62	12/51/51	12/58/44
Trace de Pillai	0,295	0,297	0,302	0,299	0,306	0,324	0,326	0,330	0,331	0,334
Pillai limite inf. (2,5 %)	0,253	0,264	0,276	0,277	0,278	0,297	0,294	0,291	0,297	0,295
Pillai limite sup. (97,5 %)	0,459	0,461	0,458	0,450	0,459	0,474	0,478	0,481	0,478	0,482
ANOVAs significatives /53	17	17	18	18	18	19	19	19	19	18

Toutes les MANOVAs sont significatives, $p < 0,001$.

Considérant que la pose des drains se fait fréquemment à 90 cm, il serait logique de considérer une profondeur plus importante, telle que 100 cm comme seuil critique d'intervention. Plus en détails, afin d'éviter une dégradation des propriétés, il serait conseillé de conserver les sols dès qu'ils deviennent moins profond qu'un mètre, alors qu'il serait plus coûteux et moins rentable de conserver les sols de moins de 60 cm, déjà affectés par une baisse marquée de leur productivité. Comme il s'agit d'un seuil pratique et non statistique, il serait approprié, après plusieurs années de tests in-situ, d'utiliser les connaissances acquises sur l'efficacité des mesures de conservations et d'évaluer économiquement si un autre seuil de conservation serait plus approprié que le 100 cm proposé plus haut. Le seuil de 100 cm sera intégré à la carte de gestion des sols présentée plus loin, mais ne sera pas utilisé dans les statistiques qui suivent.

2.3.4.3 Étude des ANOVAs significatives entre les groupes de profondeur cultivable

La MANOVA révèle une différence significative entre les groupes au seuil de 60 cm de profondeur cultivable, mais ne révèle pas la nature des variables qui diffèrent entre les groupes. Pour répondre à cette question, une série d'ANOVAs a été réalisée sur chacune des 53 variables transformées. Un ajustement de Bonferroni a permis de contrôler l'erreur de type I (Tabachnick & Fidell, 2013). Plus en détails, le seuil $\alpha = 0,05$ a été divisé par le nombre total d'ANOVAs, soit 53, afin d'obtenir une nouvelle valeur critique de significativité. Ce nouveau seuil α est de 0,000943. Ainsi, au seuil de profondeur cultivable de 60 cm, 18 variables sont significatives entre les groupes. Ces variables sont principalement reliées à la teneur en carbone, au ratio C/N, à la MVA, à la salinité et aux courbes de rétention.

Les postulats de normalité, d'indépendance des observations et d'homogénéité de la variance ont été vérifiés pour chacune des ANOVAs. Certaines variables montraient des signes d'hétéroscédasticité, ce qui a justifié la réalisation d'ANOVAs à variance hétérogène qui est plus robuste au non-respect du postulat d'homogénéité de la variance entre groupes. L'une des 18 variables, MVA_2, n'était plus significative et a été retirée. Concernant les autres variables ne respectant pas ce postulat, il semble que le non-respect soit lié au faible nombre d'observations : les résultats demeurent donc valides. Le tableau 10 présente les 17 variables significatives entre les groupes. À titre indicatif seulement, les autres ANOVAs significatives au seuil $\alpha = 0,05$ se retrouvent à l'annexe 3. Le lecteur doit comprendre que le risque d'erreur de type I est important et que ces valeurs sont présentées uniquement puisqu'elles peuvent avoir un impact sur la mise en place d'un plan de drainage ou de conservation des sols. Ces variables doivent être vues comme des indicateurs possibles de dégradation, mais non significatifs pour le moment.

Tableau 10 – Moyenne (Moy) et erreur standard (SE) des 17 variables significativement différentes entre le groupe 1 (< 60 cm) et le groupe 2 (> 60 cm).

Variable	Valeur F $F(1;112)$	p value	Groupe 1 $N = 12$		Groupe 2 $N = 102$	
			Moy	SE	Moy	SE
Ctot_1	50,59	< 0.0001	33,74	0,93	43,31	0,32
Ctot_2	47,29	< 0.0001	34,05	1,00	43,98	0,34
Ctot_3	51,5	< 0.0001	21,39	1,87	44,92	0,64
MO_1	44,59	< 0.0001	66,27	1,68	84,60	0,58
MO_2	45,78	< 0.0001	66,19	1,88	85,63	0,64
MO_3	49,53	< 0.0001	40,30	3,54	87,09	1,21
MVA_1	25,76	< 0.0001	0,38	0,02	0,29	0,01
MVA_3	15,83	0,000123	0,64	0,05	0,18	0,02
CN_1	39,08	< 0.0001	17,27	0,97	22,87	0,33
CN_2	41,44	< 0.0001	17,05	1,10	23,44	0,38
CN_3	62,47	< 0.0001	15,53	1,21	23,96	0,42
CE_2	19,71	< 0.0001	862,25	53,58	515,98	18,38
CE_3	31,41	< 0.0001	1272,50	81,09	513,47	27,82
Tea.5kPa_3	15,33	0,000156	0,11	0,02	0,19	0,01
REU.10kPa_3	22,88	< 0.0001	0,10	0,02	0,18	0,01
FIB	56,46	< 0.0001	4,16	2,18	23,13	0,75
FIB_frot	43,02	< 0.0001	1,43	1,27	8,52	0,44

Les variables montrent toutes des signes de dégradation de la qualité du sol dans le groupe de faible profondeur cultivable. Selon ces variables, les sols peu profonds sont plus humifiés (Ctot, MO, C/N, FIB et FIB_frot plus faibles), plus compactes (MVA plus élevée, Tea et REU plus faibles) et plus salins (CE plus élevée). Sans surprise, la présence de sol minéral et de coprogène près de la surface a affecté les propriétés du sol.

Afin de compléter ce tableau de résultats, les figures 18 à 24 permettent d'observer la distribution des observations par groupe de profondeur pour les 17 variables significatives grâce à des diagrammes en violon. Cette représentation graphique des observations permet d'apprécier la distribution de celles-ci à l'aide d'une boîte à moustaches (« boxplot ») combinée à un double histogramme de densité.

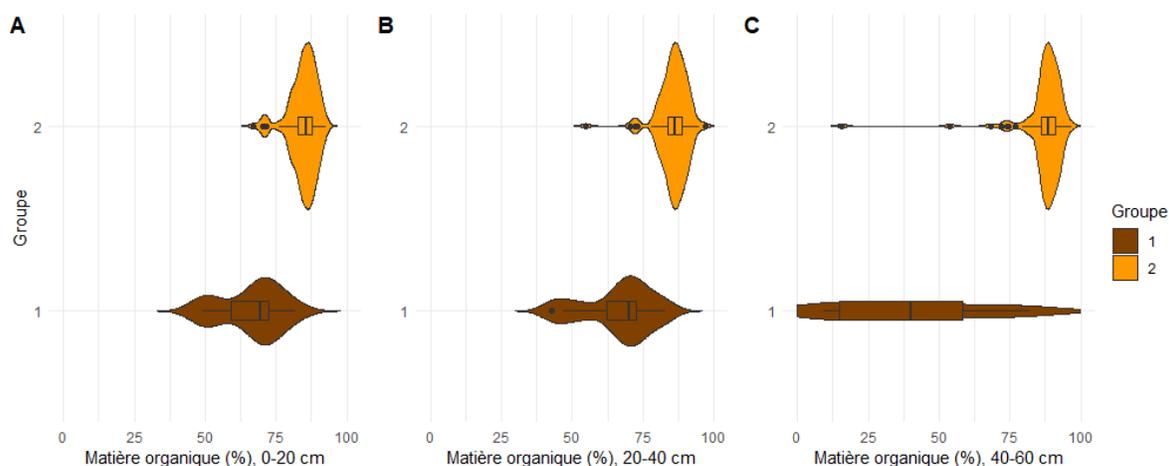


Figure 18 – Diagrammes en violon du contenu en matière organique en fonction du groupe de profondeur cultivable pour trois profondeurs d'échantillonnage : a) 0 à 20 cm de profondeur, b) 20 à 40 cm et c) 40 à 60 cm.

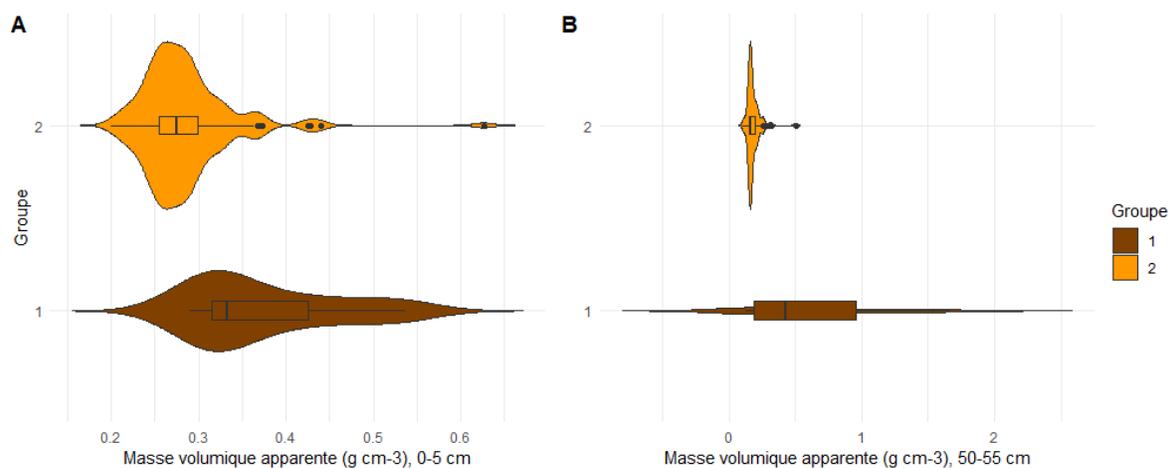


Figure 19 - Diagrammes en violon de la masse volumique apparente en fonction du groupe de profondeur cultivable pour deux profondeurs d'échantillonnage : a) 0 à 5 cm de profondeur et b) 50 à 55 cm.

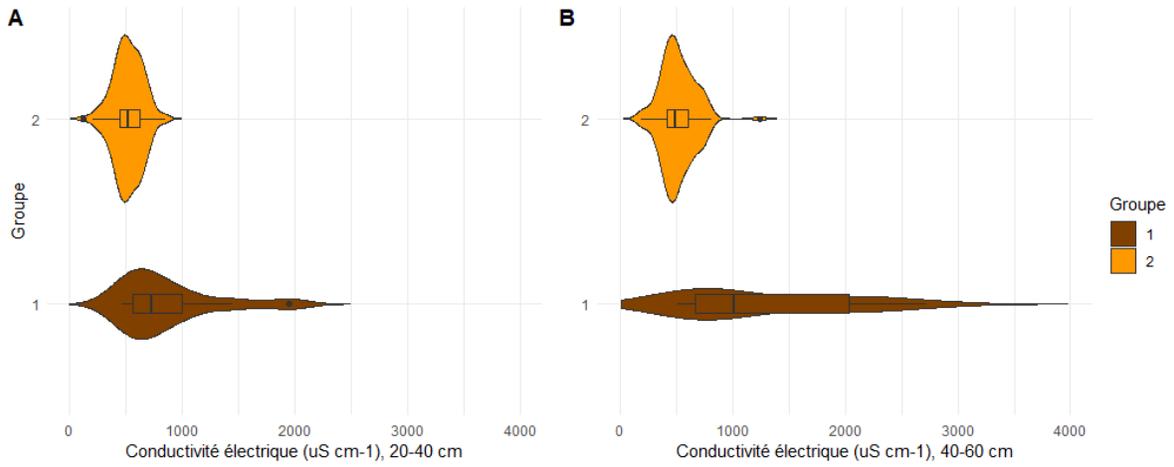


Figure 20 - Diagrammes en violon de la conductivité électrique en fonction du groupe de profondeur cultivable pour deux profondeurs d'échantillonnage : a) 20 à 40 cm de profondeur et b) 40 à 60 cm.

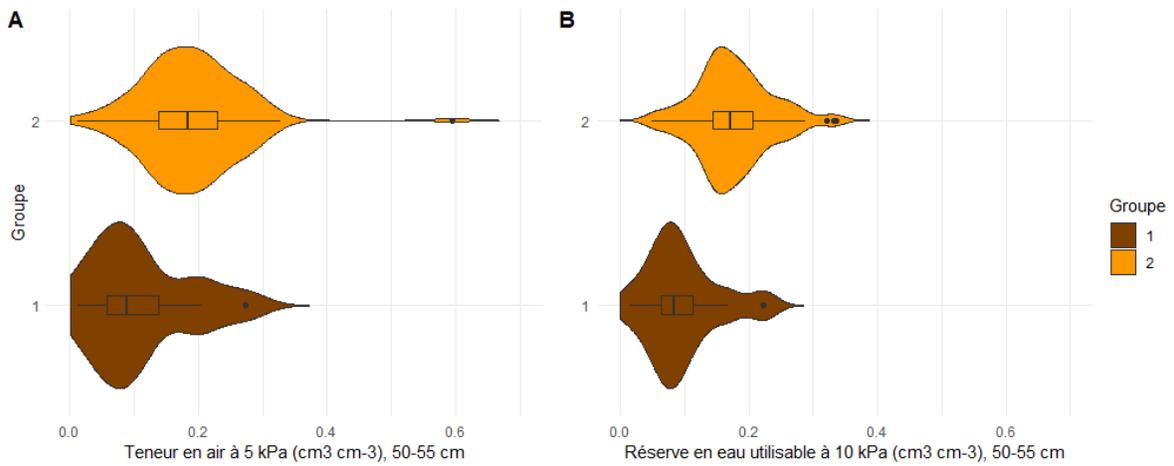


Figure 21 - Diagrammes en violon a) de la teneur en air à 5 kPa et b) de la réserve en eau utilisable à 10 kPa en fonction du groupe de profondeur cultivable selon un échantillonnage à 50-55 cm.

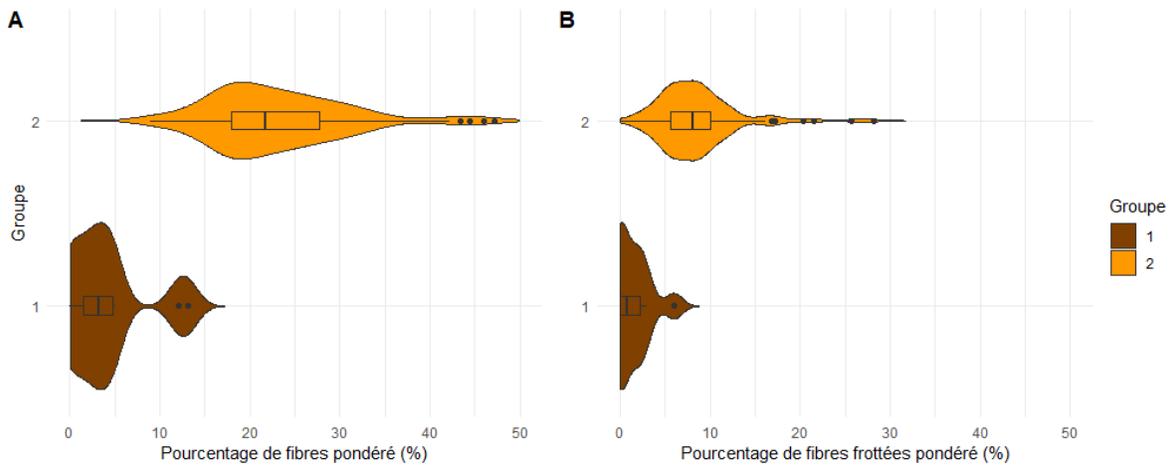


Figure 22 - Diagrammes en violon a) du pourcentage de fibres pondéré et b) du pourcentage de fibres frottées pondéré en fonction du groupe de profondeur cultivable.

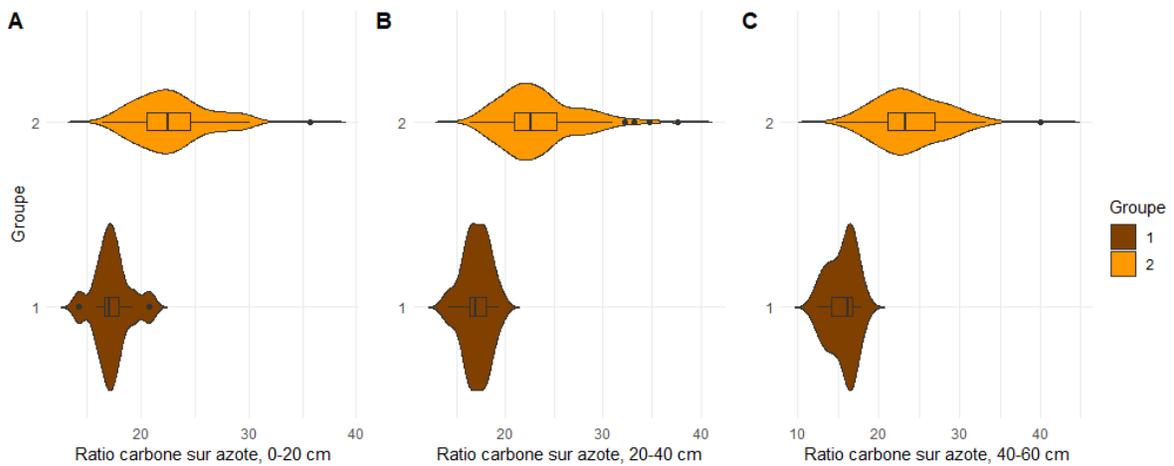


Figure 23 - Diagrammes en violon du ratio carbone sur azote en fonction du groupe de profondeur cultivable pour trois profondeurs d'échantillonnage : a) 0 à 20 cm de profondeur, b) 20 à 40 cm et c) 40 à 60 cm.

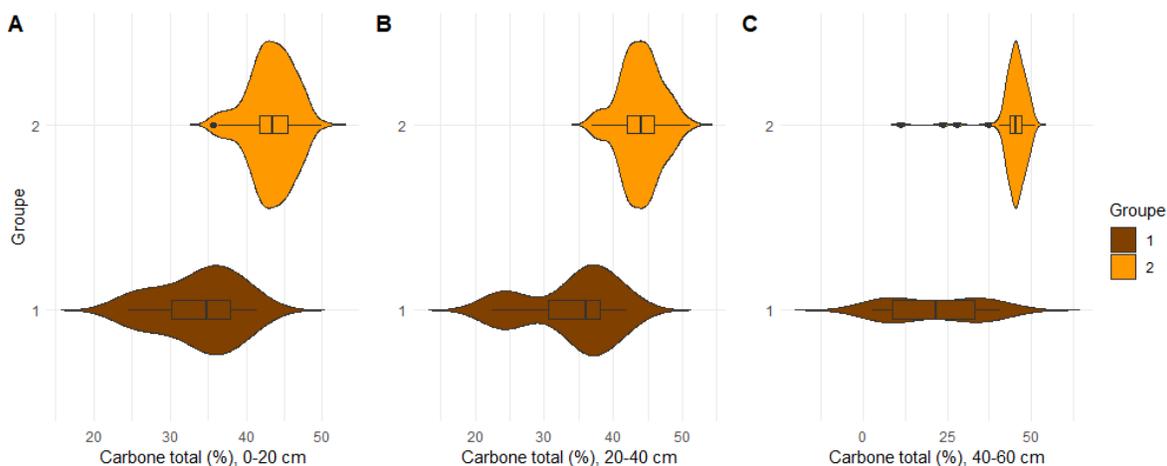


Figure 24 - Diagrammes en violon de la teneur en carbone total en fonction du groupe de profondeur cultivable pour trois profondeurs d'échantillonnage : a) 0 à 20 cm de profondeur, b) 20 à 40 cm et c) 40 à 60 cm.

2.3.4.4 Distribution spatiale des sols en fonction de leur groupe de profondeur cultivable à l'échelle régionale

Une carte de la profondeur au sol minéral et une autre de l'épaisseur de la couche de coprogène ont été réalisées en utilisant les données de cette étude et d'autres données acquises par l'équipe. Les détails de la réalisation de ces cartes sont présentés au chapitre 3. En soustrayant la carte de l'épaisseur de la couche coprogène à la carte de profondeur au sol minéral, la profondeur cultivable est obtenue. À noter que les erreurs de prédiction liées aux deux cartes s'additionnent, augmentant l'erreur en chaque point de la carte finale. Les seuils de 60 cm et de 100 cm obtenus à la sous-section précédente peuvent être appliqués comme classes afin de former la légende de la carte de la profondeur cultivable des sols afin de représenter visuellement la distribution de ces groupes de profondeur sur le territoire (figure 25).

Tout d'abord, il est possible de remarquer une entrée supplémentaire dans la légende, soit celle de profondeurs négatives. La carte de la profondeur cultivable comporte de grandes erreurs de prédiction, 19 à 387 cm d'écart-type par cellule de 10 x 10 m. Le modèle final prévoit parfois une épaisseur plus importante de coprogène qu'il n'y a de sol organique. Par conséquent, une valeur négative résulte de la soustraction du modèle minéral et du modèle coprogène. Cette superficie correspond toutefois à 0,1% de la zone à l'étude.

Ensuite, les sols plus dégradés (< 60 cm) se retrouvent majoritairement en bordure des tourbières, ce qui était attendu, sans se limiter à ce simple patron spatial. Une zone dégradée à l'intérieur des terres est visible dans la tourbière au sud-ouest de la carte. Cette zone est associée à une épaisseur prédite de coprogène importante.

Quant à la tourbière à sa droite, elle contient une concentration de sol dégradé plus importante que les autres tourbières. Cela peut être associé au dépôt qui originellement était moins important ou à l'année de drainage et de mise en culture des champs.

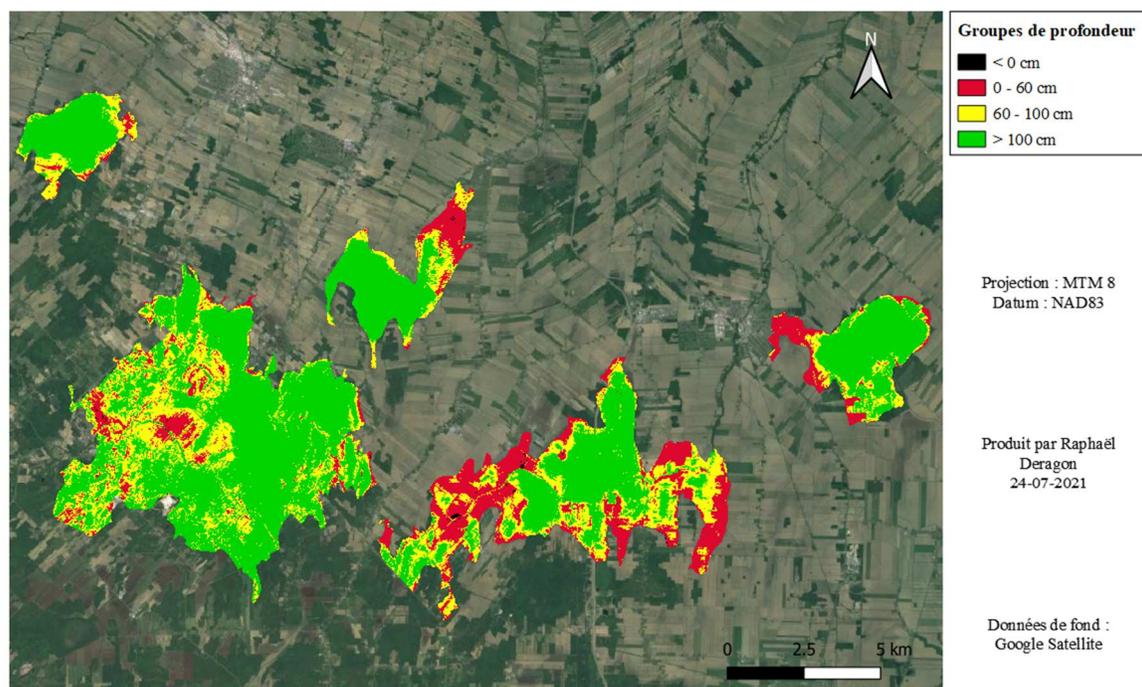


Figure 25 – Carte régionale des groupes de profondeur de sol cultivable à des fins de conservation des sols.

Quant aux implications de ces résultats face à la conservation de ces sols, près de 60 % de la superficie est d'une profondeur cultivable de plus d'un mètre (tableau 11). À cela s'ajoute environ 25 % de sols à risque (60-100 cm) alors que 15 % des superficies sont déjà dégradées (< 60 cm).

Tableau 11 – Superficie prédites des groupes de profondeurs à l'échelle régionale.

Profondeur cultivable (cm)	Superficie (ha)	Pourcentage
< 0	14	0,1%
0-60	1946	16,3%
60-100	2906	24,3%
> 100	7083	59,3%
Total :	11949	100%

2.4 Discussion

La méthode proposée, basée sur la profondeur cultivable du sol, a permis de définir un seuil de changement des propriétés associé à la profondeur critique de 60 cm et un seuil critique de 100 cm associé aux caractéristiques de drainage. De part et d'autre du premier seuil, les propriétés varient significativement. Les variables du tableau 10 indiquent une dégradation du sol pour le groupe de faible profondeur cultivable. Ces résultats semblent cohérents avec les quelques seuils discutés au chapitre 1 et l'évolution attendue des propriétés avec la dégradation du sol. La baisse de porosité observée concorde avec l'évolution attendue des sols organiques (Brandyk et al., 2003; Ilnicki & Zeitz, 2003; Okruszko & Ilnicki, 2003; Hallema et al., 2015b). L'humification du sol en surface entraîne un déséquilibre du ratio C/N, ce qui a été observé, puisque de l'azote est libéré et du carbone est relâché par la respiration des microorganismes (Anshari et al., 2010; Ilnicki & Zeitz, 2003; Kruger et al., 2015). Quant à la salinité qui augmente, cela peut être explicable par une remontée des sels de l'horizon minéral sous-jacent ou encore du coprogène. Quant au contenu en fibres, il est attendu qu'un sol mince ait du matériel organique plus dégradé à la fois chimiquement et physiquement, respectivement causé par l'action des microorganismes en présence d'oxygène et par le travail du sol (Ilnicki & Zeitz, 2003; Kroetsch et al. 2011; Pawluczuk et al., 2019).

Plus en détails, la MVA en surface dépasse le seuil de $0,2 \text{ g cm}^{-3}$ pour les deux groupes de sols. Cela s'explique par la dégradation du sol qui se fait majoritairement à la surface. Avec le travail de sol, les particules s'affinent et s'imbriquent plus facilement. Cela se produit peu importe la profondeur du sol. La situation change si la MVA entre 50-55 cm est étudiée. La MVA des sols profonds est en deçà du seuil, alors que la présence de sol minéral entraîne une MVA en moyenne trois fois plus dense pour les sols moins profonds. Quant à la teneur en air du sol, propriété physique liée à la compaction et essentielle à la diffusivité des gaz, elle serait inférieure au seuil évoqué par Caron et Michel (2021) pour les sols peu profonds, en plus d'avoir une réserve en eau utile inférieure. Concernant la salinité, les cultures maraîchères comme la carotte ($1000 \mu\text{S cm}^{-1}$), l'oignon ($1200 \mu\text{S cm}^{-1}$), le céleri ($1800 \mu\text{S cm}^{-1}$), la laitue ($2000 \mu\text{S cm}^{-1}$) et l'épinard ($2000 \mu\text{S cm}^{-1}$) sont sensibles et modérément sensibles à la salinité du sol (Machado & Serralheiro, 2017). Des baisses de rendement sont attendues lorsque la salinité du sol dépasse ces seuils, ce qui était le cas dans la majorité des sols peu profonds. La remontée capillaire et l'apport d'eau d'irrigation saline favorisent cette accumulation de sels à la surface qui peut nuire aux cultures sur le long terme.

En pratique, les résultats suggèrent que les sols moins profonds, dans lesquels les cultures ont inévitablement une profondeur d'enracinement plus faible, ont aussi une faible réserve en eau en plus de limiter la diffusion des

gaz. Il s'agit d'une combinaison de propriétés qui ne sont pas désirées au niveau agronomique. Cela entraîne des conséquences majeures sur l'irrigation des cultures, autant sur la quantité d'eau apportée que sur le moment optimal de l'irrigation. Ces variables significatives entre les groupes de profondeur pourront être utilisées afin d'orienter la prise de décisions et de mesures sur le terrain dans un contexte de conservation des sols. Évidemment, d'autres variables peuvent s'ajouter telle la Ksat afin de générer un plan de drainage par exemple.

L'utilisation de la profondeur cultivable comme outil d'aide à la décision a toutefois une incertitude associée. Peu de données sont disponibles quant à l'affaissement et la perte de sol annuelle à l'échelle du champ, sans compter les variations de la hauteur du sol qui est due aux cycles saisonniers et aux cycles d'humectation-dessiccation (Ilnicki & Zeitz, 2003). L'incertitude entourant la mesure de la profondeur cultivable d'un site d'une année à l'autre demeure donc inconnue. L'intensité et la durée des cycles diffèrent de site en site, nécessitant des modèles de prédiction calibrés avec des échantillons sur le terrain (Camporese et al., 2006; Morton & Heinemeyer, 2019). Les variations de la hauteur du sol peuvent influencer les mesures de conservation si la profondeur mesurée se trouve près de l'un des deux seuils proposés. Plusieurs méthodes sont disponibles afin de mesurer la profondeur du sol, mais elles ne permettent pas toutes la discrimination entre le coprogène, le sol minéral et le sol tourbeux de surface. Elles ont des coûts associés et des incertitudes qui varient (Parry et al., 2014) que ce soit avec des méthodes manuelles ou la détection de proximité (Ji et al., 2019).

De plus, trois facteurs viennent nuancer les résultats de la carte de gestion proposée. Premièrement, ces superficies prédites sont soumises à la précision des prédictions, tel que discuté précédemment. Par conséquent, l'approche n'est pas suffisante en elle-même afin de supporter l'agriculture de précision (plans de drainage et plan de conservation précis). L'utilisation de technologies de détection de proximité serait une alternative à l'approche par télédétection. Dans les prochaines années, de nouveaux points seront ajoutés au jeu de données ce qui pourra contribuer à la couverture de la zone à l'étude, à la performance des modèles prédictifs et à augmenter la robustesse des conclusions. En effet, le patron d'échantillonnage initial, axé autour de la variable de qualité, ne couvre peut-être pas l'ensemble des valeurs de profondeur et des gradients rencontrés dans ces tourbières.

Deuxièmement, le tableau 11 suppose que les 14 producteurs agricoles concernés aient des sols équitablement répartis parmi ces classes. Or, tel que vu plus haut, la profondeur cultivable varie entre les tourbières et certaines sont plus dégradées que d'autres. Pour ces entreprises, l'implication sur les mesures de conservation est importante.

Troisièmement, il est important de noter que la carte présentée et les statistiques qui en découlent comprennent des forêts, des habitations et voies publiques en plus de deux carrières à ciel ouvert. Les superficies présentées au tableau 11 ne sont donc pas absolues. Tel que discuté à la prochaine section, malgré les erreurs de prédiction, cette carte permet toutefois de comprendre les tendances et de repérer les gradients. La carte de la profondeur au sol minéral est basée sur un jeu de données plus important et elle comporte un terme d'erreur de prédiction légèrement plus faible et une concordance plus élevée que la carte de l'épaisseur de coprogène. Il pourrait être conseillé de se référer à cette carte afin de planifier les interventions de conservation des sols dans les zones que l'on sait exemptes de coprogène afin de réduire l'erreur de la prise de décision. Une analyse économique des gains en fonction des coûts de conservation et de la ressource en sol restante d'un champ donné pourra idéalement compléter l'outil.

Ainsi, il est maintenant possible d'affirmer que **l'hypothèse 2 est acceptée**. En effet, trois groupes de gestion ont été formés à l'aide de 17 propriétés discriminantes. Ces seuils sont critiques (statistiquement et en pratique) pour la gestion des sols organiques cultivés dans un contexte de conservation des sols.

2.5 Conclusion

L'approche par profondeur de sol cultivable a été retenue afin de lier la dégradation des sols organiques cultivés à certaines propriétés discriminantes couramment associées à l'évolution de ces sols. Deux seuils de dégradation (60 cm) et un seuil d'intervention (100 cm) sont proposés selon les analyses réalisées, permettant de former trois groupes de gestion des sols. Une carte de la répartition de ces groupes a été produite afin de faciliter la mise en place de mesures de conservation en ciblant des zones d'intervention prioritaires. Cependant, afin d'améliorer la performance de l'outil logiciel, il est impératif d'intégrer des données connexes qui permettront une mise en place efficace d'un plan de conservation des sols. Notamment, les facteurs de dégradation des sols (quantification des pertes par érosion éolienne et hydrique, et quantification des pertes par décomposition microbiennes) et le potentiel des méthodes de conservation (efficacité estimée et diminution de la perte de sol annuelle) devront être intégrés au processus de décision. En effet, certaines zones plus profondes pourraient être soumises à des pertes éoliennes plus importantes qu'un autre site moins profond où la perte de sol est plus lente. En d'autres mots, la profondeur cultivable ne doit pas être l'unique critère utilisé afin de cibler des zones d'intervention.

Bibliographie

- Camporese, M., Ferraris, S., Putti, M., Salandin, P., & Teatini, P. (2006). Hydrological modeling in swelling/shrinking peat soils. *Water Resources Research*, 42(6).
<https://doi.org/10.1029/2005WR004495>
- Caron, J. & Michel, J.-C. (2021). Understanding and optimizing the physical properties of growing media for soilless cultivation. 10.19103/AS.2020.0076.04.
- Centre d'expertise en analyse environnementale du Québec. (2014). Détermination du pH : méthode électrométrique, MA. 100 – pH 1.1, Rév. 3, Ministère du Développement durable, de l'Environnement, de la Faune et des Parcs du Québec, 11 p.
<http://www.ceaeq.gouv.qc.ca/methodes/pdf/MA100pH11.pdf>
- Comité canadien d'experts sur la prospection, p. (1983). Système d'informatique des sols au Canada (SISCan) : manuel de description des sols sur le terrain : rapport technique. Ottawa: Direction générale de la recherche, Agriculture Canada.
- Conseil des Productions Végétales du Québec inc. (1997a). Méthodes d'analyse des propriétés physiques des milieux artificiels, des tourbes et des sols organiques. Courbe de rétention. p.10
- Conseil des Productions Végétales du Québec inc. (1997b). Méthodes d'analyses des propriétés physiques des milieux artificiels, des tourbes et des sols organiques. Estimation de la distribution des particules (granulométrie). p.4
- Conseil des productions végétales du Québec inc. (Juin 1988). Préparation des échantillons de sol, Agdex 533, Méthode PR-1, 2 p.
- Guedessou, C. V. (2020). Modélisation de systèmes de drainage des sols organiques cultivés sujets à la compaction et au tassement (Thèse Ph.D.). Université Laval, Québec.
- López, M. V., de Dios Herrero, J. M., Hevia, G. G., Gracia, R., & Buschiazso, D. E. (2007). Determination of the wind-erodible fraction of soils using different methodologies. *Geoderma*, 139(3), 407-411.
doi:10.1016/j.geoderma.2007.03.006
- Machado, R., & Serralheiro, R. (2017). Soil salinity: effect on vegetable crop growth. management practices to prevent and mitigate soil salinization. *Horticulturae*, 3(2), 30–30.
<https://doi.org/10.3390/horticulturae3020030>
- Ministère du Développement Durable, de l'Environnement et de la Lutte contre les Changements Climatiques du Québec. (2017). Détermination des solides totaux et des solides totaux volatils : méthode gravimétrique, MA. 100 – S.T. 1.1, Rév. 5, Centre d'expertise en analyse environnementale du Québec, 13 p.

- Morton, P. A., & Heinemeyer, A. (2019). Bog breathing: the extent of peat shrinkage and expansion on blanket bogs in relation to water table, heather management and dominant vegetation and its implications for carbon stock assessments. *Wetlands Ecology and Management*, 27(4), 467–482.
- Munsell Color (Firm). (2010). *Munsell soil color charts : with genuine Munsell color chips*. Grand Rapids, MI :Munsell Color.
- Olson, C. L. (1974). Comparative Robustness of Six Tests in Multivariate Analysis of Variance. *Journal of the American Statistical Association*, 69:348, 894-908.
- Peterson, R.A. (2019). Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. *Journal of Applied Statistics*, 1-16.
- Revelle, W. (2020) *psych: Procedures for Personality and Psychological Research*, Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=psych>, R package version 2.0.12
- Šimůnek, J., Genuchten, M. T., & Wendroth, O. (1998). Parameter Estimation Analysis of the Evaporation Method for Determining Soil Hydraulic Properties. *Soil Science Society of America Journal*, 62(4), 894-905. doi:10.2136/sssaj1998.03615995006200040007x
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson Education.
- Therneau, T. & Atkinson, B. (2019). *rpart: Recursive Partitioning and Regression Trees*. R package version 4.1-15. <https://CRAN.R-project.org/package=rpart>

Chapitre 3 - Mapping cultivated Organic soils in the southwest plain of Montreal: Determining peat maximum thickness of cultivation at a regional scale

Raphaël Deragon^a, Daniel D. Saurette^{bc}, Brandon Heung^d, Jean Caron^a

^a Département des sols et de génie agroalimentaire, Université Laval, 2325 Rue de l'Université, Québec, QC, G1V 0A6

^b Ontario Ministry of Agriculture Food and Rural Affairs, 1 Stone Road West, 3rd Floor SE Guelph, Ontario N1G 4Y2.

^c School of Environmental Sciences, University of Guelph, Guelph, Ontario, N1G 2W1.

^d Faculty of Agriculture, Department of Plant, Food, and Environmental Sciences, Dalhousie University, 50 Pictou Rd., Truro, NS, B2N 5E3

Keywords

Digital soil mapping, Machine-learning, Organic soils, Peat cultivation thickness, Coprogenous soil

3.1 Résumé

Les sols organiques cultivés sont sujets à une dégradation constante et à une perte de sol annuelle menaçant leur utilisation sur le long terme. La profondeur de sol cultivable (PSC) est une variable clé afin de cibler des zones d'intervention pour la conservation de ces sols. Grâce à la cartographie numérique des sols, une carte à l'échelle régionale a été produite à l'aide de la profondeur au sol minéral (PSM) et de l'épaisseur de la couche coprogène (ECC). Trois modèles d'auto-apprentissage et une méthode géostatistique ont été comparées. Le modèle Cubist a le mieux performé, rapportant des coefficients de concordance de 0.69 et de 0.07 (PSM et ECC respectivement). Cependant, l'incertitude associée à la PSC était très importante. De nouvelles covariables doivent être étudiées et des approches par détection de proximité sont conseillées afin de mieux comprendre la distribution et l'épaisseur du matériel coprogène.

3.2 Abstract

Reclaimed Organic soils are subject to constant degradation and annual soil loss threatening their long-term cultivation. Peat maximum thickness of cultivation (PMTc) is a key feature that can help target priority intervention zones for their conservation. In order to produce a regional PMTC map, it is necessary to know the depth to the mineral layer (DML) and the coprogenous layer thickness (CLT). Using digital soil mapping, a map of DML and CLT were produced and combined to obtain PMTC using open-access covariates. Three machine-learning models and a geostatistical method were compared. Cubist model performed the best, yielding coefficients of concordance of 0.69 and 0.07 (DML and CLT). However, the uncertainty associated with PMTC predictions was very large. New covariates need to be investigated and proximity sensing approaches are advised to better understand the distribution and thickness of the coprogenous layer.

3.3 Introduction

Canada is home to about a quarter of the world's peatland extent (Vepraskas & Craft, 2015). Cultivated Organic soils cover only 4 % of the Province of Quebec's southern territory; yet, it greatly contributes to food production and to the economy of the province with export to northeastern United States (Groupe AGÉCO, 2007; Parent & Gagné, 2010). These soils are prized for vegetable production but they are affected by intense soil loss processes unique to the evolution of peat materials. Soil loss occurs primarily by subsidence, oxidation, and erosion after land conversion for agriculture, and is enhanced by the drainage of peatlands (Kroetsch et al., 2011; Vepraskas & Craft, 2015). Over the past decades, with an annual estimated soil loss of 2.5 cm (Ilnicki, 2003; Esselami, Boudache & Grenon, 2014), degraded, shallower fields will transition to less productive mineral soils, as if the mineral boundary had been moving towards the surface.

Furthermore, coprogenous material (also known as gyttja) can also be found between the peaty layer and the mineral layer (Lamontagne et al., 2014). When the material is found within the first 160 cm and has a minimum thickness of 5 cm, it is referred to as a 'limnic' layer (SCWG, 1998). This gelatinous, impervious material is unsuitable for agriculture and specific to lacustrine organic deposits (Kroetsch et al, 2011). When it dries, this material shrinks and does not rewet (SCWG, 1998). In Poland, calcareous limnic deposits can limit root growth (Ilnicki, 2003). Limnic materials can be of different types, namely coprogenous earth (sedimentary peat), diatomaceous earth, or marl (SCWG, 1998). In the southwest plain of Montreal, coprogenous earth is most common and sometimes with a small layer of marl. The latter is effervescent due to the presence of shells and precipitated CaCO_3 , while the former can be a mineral (< 17 % organic carbon), or organic deposit (\geq 17 % organic carbon) enriched by algae or aquatic life plants transformation products (SCWG, 1998). Since pedological surveys focused on the spatial extent of peat and soil series description, little is known about the depth to the mineral layer (DML) and the coprogenous layer thickness (CLT). To mitigate the impact of long term soil loss, soil conservation approaches, such as the addition of biomass crop amendments, are needed (Dessureault-Rompré et al., 2020); however, this approach has a related cost because the biomass is not an infinite resource. Therefore, priority management zones need to be defined in order to guide the application of a regional soil conservation plan. The effective peat thickness that can be used for agriculture, hereinafter referred to as the 'peat maximum thickness of cultivation' (PMTC), could be used to define these priority management zones. In other words, the PMTC can be defined as the thickness of the peaty layer of an Organic soil, therefore excluding coprogenous and mineral materials. This definition would better reflect the real cultivation thickness of a field and help manage shallower soils if the spatial distribution of the PMTC is known.

Modern digital soil mapping (DSM) techniques can provide such information by leveraging technological advances for predicting a soil attribute or class using georeferenced field measurements and a suite of environmental covariates obtained via remote or proximal sensing (McBratney et al., 2003). While global map products are available, new mapping efforts are targeting soil properties and soil functions (FAO, 2020). In DSM studies related to, the thickness of the peat and organic carbon content are the most frequently predicted features (Minasny et al., 2019). To our knowledge, predictive mapping of coprogenous materials has never been explored and the relevant covariates are unknown. Yet, the information that could be acquired through a predictive map on a regional scale is crucial to the estimation of PMTC, and therefore the delineation of priority management zones. Manual probing over large areas is labor intensive and relatively slow (Parry et al., 2014). A regional approach, relying on a field calibration dataset could provide useful maps when combined with a relevant set of covariates. Many combinations of covariates are commonly used in regression and classification mapping of peatlands and could be investigated as potential candidates for predicting the thickness of the coprogenous layer. Peat thickness and its extent are often evaluated by combining a digital elevation model (DEM) and its derivatives, airborne gamma radiometric data, electromagnetic data, and satellite data (Gatis et al., 2019; Minasny et al., 2019; Rudiyanto et al., 2018; Siemon et al., 2020). Most of the covariates are the product of remote sensing techniques, while ground penetrating radar, gamma radiometric data and soil electrical resistivity or conductivity can be obtained with proximal sensing techniques (Beucher et al., 2020; Comas et al., 2015; Parry et al., 2014; Rosa et al 2009). Hence, given the lack of data concerning DML and CLT for the study area and the abundance of covariates, the main objective of this study was to integrate open-access, remote sensing covariates and field data to predict the spatial distribution of the PMTC, including mineral and coprogenous materials' depth. Here, the specific objectives were to (1) determine contributing covariates to the prediction of DML and CLT; and (2) to derive a map of PMTC from the mineral and coprogenous material predictions as a tool to guide soil conservation practices.

3.4 Methodology

3.4.1 Study area

The study area is comprised of five drained and cultivated peatlands in the southwest plain of Montreal, Quebec, covering approximately 90 km² (45.1°N to 45.3°N latitude and -73.3°W to -73.7°W longitude). Figure 26 shows the study area and the spatial distribution of sampling sites. Pristine forests are still present but are likely affected by agricultural drainage. Most of the soils are used for horticulture, while a small proportion are used for producing gardening soil bags. The five peatlands are part of three separate watersheds (Lamontagne et al., 2014). The elevation ranges between 45 and 70 m above mean sea level; however, the agricultural fields are level and relatively flat across the peatlands.

Fieldwork was carried out on 14 partnering research farms, which were converted to agricultural use between 1950 and 2010. Most fields were classified as Humisols and Mesisols, while a small area of recently converted field was classified as Fibrisols (SCWG, 1998). The organic materials found in the southwest plain of Montreal were deposited in channels or in depressions. Basin bogs or shore swamps, the latter having shallower deposits, were the result of hundreds of years of accumulation (Lamontagne et al., 2014). Peatlands in this region are mainly composed of forest peat or herbaceous materials (LaSalle 1963; Grenon 1988). Limnic materials are found throughout the study area, but not in every peatland: the northwestern-most peatland appeared to be devoid of limnic material whereas the other peatlands had regions of limnic materials.

3.4.2 Soil data

Two data sets were used to build the predictive models. The first one consisted of 255 sites that were visited between 2019 and 2021 (shown as white dots in Figure 26). The coprogenous layer thickness (CLT) and the depth to the mineral layer (DML) were obtained by manually extracting soil cores with a Macauley corer (Eijkelkamp peat sampler) until the mineral horizon was reached.



Figure 26 – Study area and sampling sites categorized by the nature of the information collected. Overlaid is a square grid of 1.5 km size used to spatially partition the datasets to allow leave-one-block-out spatial cross-validation.

The second dataset consisted of 4,488 observations. Here, 4,286 locations were manually-probed sites (shown as grey dots in Figure 26) between 2010 and 2014 (Jean Caron, personal communication) and were combined with the first dataset. Manual probing consists of inserting a thin metal rod in the soil and when a change of resistivity is felt, the depth is recorded. Only the DML was recorded at these locations. It should be noted that coprogenous materials do not offer notable resistance when a rod is inserted into the soil in comparison to the important change in density when the probe reaches the underlying mineral soil. Manual probing is sensitive to measurement errors due to wood pieces (Parry et al., 2014). Nonetheless, manual probing can be done rapidly and requires less energy to do manually than the Macauley sampling technique. Furthermore, the sites are clustered in three of the five peatlands. It was believed that the addition of this dataset to this project could contribute to the predictive power of the models when generating DML predictions even though manual probing might be prone to more measurement bias or error. Since some sampling locations were close to each other, observations that fell into the same raster cell were averaged. The spatial resolution of this project and all resulting maps is 10 x 10 m. Summary statistics of the two datasets are presented in Table 12.

Table 12 - Summary statistics of the coprogenous layer thickness (CLT) and the depth to the mineral layer (DML)

	CLT	DML
<i>N</i>	255	4,488
Minimum (cm)	0	5
Maximum (cm)	206	392
Mean (cm)	34.16	136.23
Median (cm)	25.00	130.00
Standard deviation (cm)	37.22	70.97
Coefficient of variation (%)	108.83	52.11
Variance (cm ²)	1,385.62	5,036.64
Skewness	1.26	0.46
Kurtosis	1.74	-0.42

The high coefficient of variation (CV) of the CLT (109 %) indicates high heterogeneity in the dataset, compared to the moderate to high heterogeneity of the DML (52 %). This can be explained by the fact that CLT could be considered as a zero-inflated variable with a wide range of values. A large proportion of the dataset was composed of sites without a coprogenous layer.

3.4.3 Environmental covariates

When generating a training dataset to be fitted with a predictive model, the georeferenced soil observations are spatially intersected with a suite of geospatial environmental layers (i.e., covariates or predictors). Here, the predictive models were used to establish the relationships between the covariate and soil properties to predict the spatial phenomenon of interest. Here we present the five categories of covariates used, which were selected based on the SCORPAN model of McBratney et al. (2003) and a review of the digital mapping of peatlands by Minasny et al. (2019). The description and pre-processing details will be given for the final selection of covariates.

Digital elevation model (DEM) data

A DEM is a common dataset that can be used to derive a suite of geomorphological and hydrological covariates (McBratney et al., 2003; Minasny et al., 2019). Since organic soils are formed in lowlands, depressions, and water saturated conditions, the topographical covariates are expected to be effective predictors of peat accumulation and the delineation of peatlands. The DEMs were produced and distributed by the *Ministère*

des Forêts, de la Faune et des Parcs where it was downloaded from the *Données Québec* repository. Here, the multiple DEM tiles were mosaicked. The DEM data was derived from a LiDAR survey and made available at a 1 m spatial resolution.

The DEM was pre-processed in five steps using the *whitebox* package in R (Wu, 2020). Missing data were first filled-in and following this, a feature preserving smoothing tool was used to remove short-scale noise due to the use of the ultra fine resolution DEM data. Afterwards, a mean filter with a 5 m x 5 m window was used to further smooth the DEM and remove artifacts. To reduce the computational burden and memory requirements, the DEM was aggregated to a 10 m resolution. This covariate became the reference to which all other covariates were resampled to match its extent and resolution. Single cell pits were filled to remove local artifacts and to produce a more continuous output. The breach depressions least cost tool was used to prepare a DEM to be derived in hydrological covariates (Lindsay, 2016). It breaches and then fills depressions in the DEM to ensure continuous flowpaths through depressions with a least cost pathway method based on a breaching algorithm by Lindsay (2016).

After pre-processing, both the *RSAGA* (Brenning et al., 2018) and *whitebox* packages were used to generate 70 covariates. Some of these covariates included the difference and deviation from mean elevation, topographic wetness index, multi-resolution valley bottom flatness (MRVBF), multi-resolution ridge top flatness (MRRTF), catchment areas, hillshade, slope, aspect, and curvature (Beven & Kirkby, 1979; Gallant and Dowling, 2003; Lindsay et al., 2015; McBratney et al., 2003). In addition to the aforementioned covariates, Euclidean distance fields were generated for the study area to provide spatial context (Behrens et al., 2018).

Aerial gamma-ray spectroscopy data

Six raster layers were downloaded from Natural Resources Canada's GEOSCAN database (Natural Resources Canada, 2019). The layers represented surface concentrations of potassium (K, %), equivalent uranium (URA, ppm), equivalent thorium (THO, ppm), ratios eU/eTh (RUT), eU/K (RUK), and eTh/K (RTK). Radiometric data has proven to be useful for differentiating mineral soils from organic soils given their difference in parent material, porosity, and water content (Beamish, 2013; Keaney et al., 2013). Water filled peat soil attenuates bedrock geology radioactivity and can provide an indication of the magnitude of the deposit. The six rasters were resampled from a 250 m to 10 m spatial resolution using a bilinear method to match the extent and resolution of the other covariates.

Sentinel 2 L2-A data

Sentinel 2 is a constellation of two satellites with a combined repeat cycle of 5 days over the same area. Common indices and raw bands were tested as potential covariates. Gholizadeh et al. (2018) obtained significant correlations (-0.74 to -0.36) between bands 4, 5, 11, 12 and soil organic carbon; hence, raw bands could serve as discriminant covariates to delineate a peatland's extent. L2-A level of pre-processing generates a bottom-of-atmosphere corrected reflectance raster for all 13 bands. This level of pre-processing includes radiometric, geometrics (including ortho-rectification) and atmospheric corrections. The rasters were pre-processed by the European Space Agency and made available by the Copernicus Open Access Hub. All final rasters were resampled to a 10 m spatial resolution using the bilinear method since the raw bands have a native resolution of 10, 20, or 60 m; furthermore, the rasters were realigned with the other covariates.

A multitemporal approach was adopted to capture moisture gradients over the study area. Therefore, median and standard deviation layers were produced for all raw bands and indices that were gathered for the following dates: March 31st, April 20th and 25th, and May 20th of 2020. These dates were selected because they had neither snow nor crops covering the fields to better differentiate between the organic and mineral soils, as well as forests. The dates were also selected based on data availability and cloud cover over the study area.

Landsat 8 OLI_TIRS C2 Level-2 data

Landsat 8 is a satellite with a 16-day repeat cycle over the same area. C2 Level-2 data is referred to as analysis ready data, being pre-processed by the USGS Earth Resources Observation and Science (EROS) Center and then made freely available. After Level-2 processing, surface reflectance (bottom of atmosphere) values were obtained. More information on Landsat missions and pre-processing can be found in Young et al. (2017).

This satellite provides remote sensing band layer of Earth's temperature and land surface images. The data acquired from 2 thermal infrared sensors (100 m resolution) and 9 spectral bands (30 m resolution with one at 15 m resolution) are sensitive to different wavelength ranges. As was done with Sentinel 2 data, Landsat 8 raw bands and indices were resampled to a 10 m spatial resolution. A multitemporal approach was also used to capture seasonal and moisture gradients. Three series of satellite images were chosen based on cloud percentage, absence of snow and crops. Images from May 6th of 2015, May 14th and 30th of 2018 were used. The median and standard deviation over the three dates were used as covariates. The same indices were generated for both satellites, using their respective bands, and some were computed for one satellite only (Table 13). All indices were computed in QGIS (QGIS.org, version 3.10).

Table 13 – Bands and indices from Landsat 8 and Sentinel 2 satellites explored as potential covariates.

Bands and indices	Landsat	Sentinel	Source
Brightness index	x	x	(Escadafal, 1994)
Normalized difference vegetation index (NDVI)	x	x	(Rouse et al., 1994)
Soil color index	x	x	(Poggio & Gimona, 2017)
Soil moisture index	x	x	(Poggio & Gimona, 2017)
Bare soil index	x		(Rikimaru et al., 2002)
Combined spectral response index	x		(Taghizadeh-Mehrjardi et al., 2021)
Normalized difference moisture index	x		(Gao, 1996)
Salinity index	x		(Khan et al., 2005)
Raw bands 2 to 7	x		
Normalized Difference Salinity Index (NDSI)		x	(Khan et al., 2005)
Soil Adjusted Vegetation Index (SAVI)		x	(Huete, 1988)
Transformed vegetation index (TVI)		x	(Nellis & Briggs, 1992)
Raw bands 2 to 8, 8A, 9, 11, 12		x	

Distance and directionality covariates

The center of a peatland is often the point where peat accumulation is at its maximum, whereas near the border of the peatland, one can expect thinner deposits. Therefore, the last two covariates were project-specific distance layers made in QGIS with the Multi-distance buffer plugin (Håvard, 2018). The first layer was produced by finding a peatland's centroid, then generating 10 m buffers around it until the peatland's border was reached. The product was then rasterized with a resolution of 10 m. The second layer was produced by generating 10 m buffers from the peatland's border inwards, which was then rasterized. One might think the results would be the same, but the produced covariates exhibit significantly different patterns due to the shape of each peatland. Eastness and northness rasters are shape and directionality-related covariates that were also generated in this study.

3.4.4 Covariate reduction

The 142 covariates produced above were reduced with a stepwise variance inflation factor (VIF) procedure implemented with the *vifstep* function (Naimi et al., 2014) to reduce collinearity as a general first approach (O'brien, 2007). Since it required the use of standardized covariates, the *scale* function from the *raster* package was used. This approach takes each of the standardized covariate and uses the remaining ones as independent predictors in a multiple linear regression. Then, the coefficient of determination of the covariate acting as dependent variable (R_i^2) is calculated. This process is repeated for every covariate. With equation 1, VIF is computed for every i_{th} predictor where a high proportion of variance explained by a given combination of

predictors will result in a higher VIF score. The model then evaluates if one or more covariates has a VIF score that exceeds a predetermined threshold value; in our case, a score of 5 was selected although 10 is also a common threshold (O'brien, 2007; James et al., 2017; Bian et al., 2020). If so, the covariate with the highest score is removed, and the process is repeated until all remaining covariates are below the threshold.

$$VIF_i = \frac{1}{1 - R_i^2} \quad (1)$$

Recursive feature elimination (RFE) was performed with the *rfe* function (Kuhn, 2020) to further reduce the number of covariates for training the final models. Briefly, RFE incorporates resampling to measure model performance with a reduced selection of covariates. It is a backwards feature selection method, meaning that the model initially includes all predictors, and some are removed iteratively to simplify the model, and in some cases, improve its performance. To fit the RFE models, a random forest function was used with 10-fold cross-validation repeated five times. This allowed the computation of RMSE (see equation 2) that was used as the metric to determine the best model. The most notable difference between VIF and RFE is that VIF is independent of the response variable, whereas the RFE reduces covariates as a function of the model performance in predicting the response variable.

3.4.5 Modeling approaches

Three machine learning algorithms (MLAs) commonly used in digital soil mapping were compared for both CLT and DML variables (Minasny et al., 2019; Rudiyanto et al., 2017). The selection includes Cubist, Random Forest (RF) and k-Nearest Neighbor (kNN). Although this section will briefly summarize each model, more details concerning the MLAs are provided in Heung et al. (2016) where the models were reviewed and compared. All the modeling was done using the R statistical software (R Core Team, 2020) and the caret package, which included all the tested models (Kuhn, 2020). Final maps were made in QGIS.

Cubist

Cubist is a rule-based model producing regression trees that relies on the M5 model (Quinlan, 1992), and was subsequently adapted by Kuhn & Johnson (2013) in R. At each node of the tree, the dataset is split based on a set of rules based on the value of one or many covariates, and form groups that are defined with low internal error. The terminal nodes (i.e., leaves) of the resulting tree are replaced with multivariate linear models, which are applied to the covariates to make predictions. The hybridization of piecewise linear models and the hierarchical tree models allow Cubist to capture both linear and non-linear relationships between predictors and the response variables (Malone et al., 2017). This model has two hyperparameters to be tuned by training the model (Kuhn & Johnson, 2013). Over and under-predictions can be accounted for with a boosting technique

using committees, the first hyperparameter. It specifies the number of similar trees to be produced sequentially and aggregated to optimize the set of rules. This process is explained in greater details in Kuhn & Johnson (2013). The second hyperparameter is the number of neighbors to be considered in a nearest-neighbor search through the original dataset to find the closest observation to the predicted value (Quinlan, 1993). This final step is useful to adjust the value at the predicted location based on the training dataset.

Random Forests

Based on ensemble theory (Zhang & Ma, 2012), RF is a nonparametric decision-tree model where uncorrelated predictive trees are combined (Breiman, 2001). Multiple trees are generated with the use of bagging; random sampling of the original dataset with replacement generates a new dataset for each tree. This newly formed dataset will be split at each node of the tree based on a set of rules applicable to a selection of covariates to maximize the within-node homogeneity and the between node heterogeneity. In the *caret* package implementation, *mtry* is the only tuning parameter and is used to control the number of randomly selected covariate at each node. Higher the number of covariates (*mtry*), higher the chances of correlation between the trees, but lower will be the variance of the prediction (Hastie et al., 2009). The trees from the forest are then aggregated to obtain an average prediction for each new observation (Genuer & Poggi, 2020). Predictions can be regarded as more effective than those from single tree-learners due to RF's ability to mitigate bias (Kuhn & Johnson, 2013).

k-Nearest Neighbors

kNN is a distance-based learner with one hyperparameter to fit: *k*, the number of neighbors of the unobserved location to be used for the prediction (Hastie et al., 2009). The rationale behind this MLA is that observations that have similar properties (e.i., covariates values) will tend to have a similar value for the response variable of interest. Therefore, kNN predictions are made using *k* observations from training data that have similar values to those of the predicted site in the covariates multivariate feature space. In order to find neighbors of the predicted value, a distance metric must be used, such as the Euclidean distance, to assess the closeness between observations. In a regression context, the *k* closest training data will be averaged, but if *k* = 1, the value of the training data will be attributed to the predicted location. Since covariates can have different range relatively to their metric, they were standardized for this MLA (Hastie et al., 2009).

3.4.6 Model tuning and validation

Spatial cross-validation

As stated in a review of peat mapping studies (Minasny et al., 2019), 64 % of the 90 reviewed studies did not perform validation, which may result in overfitting and overoptimistic models' performance metrics. Therefore, each of the MLAs were trained to find the optimal hyperparameters for DML and CLT models. Considering recent papers concerning overoptimistic performance estimation of MLAs in DSM applications, spatial cross-validation was used to evaluate these models (Meyer et al., 2018; Ploton et al., 2020; Pohjankukka et al., 2017; Schratz et al., 2019). This approach has proven more effective at measuring internal error and controlling for spatial autocorrelation in the dataset, although this statement was recently nuanced in Wadoux et al. (2021). Clustered data makes it easier for a model to predict nearby observations during the training process and, therefore, estimates of model performance can be biased.

Both datasets have clustered observations over the study area. To allow a fair assessment of model performance, leave-one-block-out cross validation (LOBOCV) was used (Roberts et al., 2017). This form of spatial cross-validation is similar to leave-one-out cross-validation where one observation from a dataset is partitioned out during model training and is used as a test observation. During LOBOCV, one spatial block is put aside as a validation set. The model is trained using all observations except for those from the testing block. The model is then applied to this unseen data and error measures are derived. The process is repeated by changing each validation block for a new one at each fold. The number of folds is equal to the number of blocks, respectively 45 (CLT) and 44 (DML) for this study.

To form relevant spatial block units with the discrete observations, the range from theoretical variograms of both predicted features were calculated. Oliver and Webster's (2014) approach to fitting variograms was used. The experimental and theoretical variograms were computed using the gstat package (Gräler, Pebesma & Heuvelink, 2016; Pebesma, 2004). Weighted least residual sum of squares and visual fit were used to select the best model (Oliver & Webster, 2014). If fitted models all had an adequate visual fit, the one with the lowest weighted sum of squared errors was selected. Four outliers were identified in the DML dataset, but they were associated to deep peat deposits and were kept in the model. Four outliers were identified in the CLT and the data was skewed (Table 12). Since outliers are not mistakes nor do they belong to another population, they were kept in the model and a square root transformation was applied (Oliver & Webster, 2014). This transformation was more effective at reducing skewness than a logarithmic transformation. No outliers remained after transformation. The grid was then applied to the area and can be seen in Figure 26. The same

grid was applied to both predicted features under the assumption that CLT and DML would have shared similar autocorrelation if CLT dataset had more observations.

Model performance

Two quality indices were used to select the best hyperparameters based on accuracy and precision: root mean square of error (RMSE) and Lin's concordance correlation coefficient (CCC).

$$RMSE = \sqrt{\sum_{i=1}^N \frac{(O_i - P_i)^2}{N}} \quad (2)$$

where O is the observed value and P the corresponding predicted value and N is the total number of observations.

$$CCC = \frac{2s_{op}}{(\bar{o} - \bar{p})^2 + s_o^2 + s_{op}^2} \quad (3)$$

where s_{op} is the covariance between predicted and observed values, s^2 are their corresponding variance, \bar{o} is the mean of observed values and \bar{p} is the mean of predicted values.

CCC can vary between -1 and 1 while a value near 0 indicates a lack of concordance between two variables. This index is similar to R^2 , because it measures agreement between two variables, but with the notable difference that it corrects for systematic bias if the relationship departs from the 1:1 line (Lawrence & Lin, 1989). CCC can never be higher than the absolute value of the Pearson correlation coefficient in the presence of a bias. The CCC and RMSE statistics were obtained with `oss.getCCC` function (Saurette, 2021) for the MLAs.

Data splitting

The CLT dataset was used without modification in LOBOCV because of its low number of observations ($N = 255$). On the other hand, the DML dataset was split in two using a standard 70/30 % partition with conditioned Latin hypercube sampling on the X and Y coordinates (Minasny & McBratney, 2006). Before merging the two data sets (255 + 4,286), each of them were split based on their spatial coordinates to ensure a similar coverage of the study area in the training and test partitions. Then, splits from both datasets were combined. The greater number of observations allowed to perform LOBOCV and compute internal performance assessment with a test dataset yet unseen by the model.

3.4.7 Prediction uncertainty

Many peat mapping papers do not include uncertainty estimates of the final product (Minasny et al., 2019). To produce a prediction uncertainty map to support PMTC predictions, a bootstrapping approach was used. The code was adapted from Malone et al. (2017). Since the PMTC map is the result of the difference between the DML and CLT maps, propagation of error had to be considered to adequately capture uncertainty.

As such, a bootstrapping approach was implemented where 100 realizations were produced for both CLT and DML final models. This allowed prediction of the mean, variance, and covariance of bootstrap predictions, assuming a normal distribution of the statistics for each cell. Each raster was generated by randomly sampling with replacement the original dataset with an equal number of observations.

Equation 4 describes the standard deviation value was calculated for each cell of the final map, using the bootstrap-produced maps. It is assumed that both variables were dependent (i.e., correlated), therefore a term corresponding to the covariance between the features was added to the generic equation. The uncertainty of the PMTC prediction becomes the square root of the sum of the variance of both features per cell, minus twice the covariance between CLT and DML final models for a given cell (Ku, 1966).

$$\sigma_{PMTC} = \sqrt{\sigma_M^2 + \sigma_C^2 - 2\sigma_{MC}} \quad (4)$$

where σ_{PMTC} is the standard deviation of the peat maximum thickness of cultivation for a single cell, σ^2 are the variances of the mineral and coprogenous predictions for 100 bootstraps and σ_{MC} is the covariance, as defined by equation 5.

$$\sigma_{MC} = \sum_{i=1}^N \frac{(M_i - \bar{M})(C_i - \bar{C})}{N-1} \quad (5)$$

where M_i and C_i are the value of the i th cell of one bootstrap raster, and \bar{M} and \bar{C} are the mean value of the 100 rasters for that cell, respectively for mineral and coprogenous maps.

Finally, to obtain the upper bound (95th percentile) and lower bound (5th percentile) of the 90 % prediction interval, the corresponding z value of 1.645 was multiplied to the standard deviation of the PMTC. It was then subtracted and added to the predicted PMTC value of a given cell (Equation 6). One can observe what the range of predicted values is expected to be with the produced uncertainty maps.

$$Prediction\ limits = \overline{X_{PMTC}} \pm 1.645 * \sigma_{PMTC} \quad (6)$$

where $\overline{X_{PMTC}}$ is the mean PMTC of a given cell obtained with bootstraps.

3.5 Results

3.5.1 Covariate reduction

The VIF procedure reduced the number of covariates from 142 to 59. Since this number was still unsatisfactory to proceed with the MLAs, the RFE procedure further removed covariates for both DML and CLT models. The final covariate selection thus differs for the DML and CLT models: 14 were kept for the former and 7 for the latter. Table 14 lists covariates for both models. Only two types of satellite covariate were kept in the models: soil moisture from Sentinel 2 and band 2 layer from Landsat 8. DEM derivatives are still present in large numbers. Distance layers and gamma radiometrics are also present in the final selection.

3.5.2 LOBOCV block size

A spherical model was used to fit the CLT data while an exponential model was used to fit the DML data to evaluate spatial autocorrelation. The nugget-to-sill ratio provide insight about sampling density and measurement errors when the ratio is large, but no other inference should be made from this ratio (Oliver & Webster, 2014). The nugget-to-sill ratio of variograms presented in Figure 27 are 15.8 % and 2.0 %, respectively. Since the lag interval was 100 m, the nugget-to-sill ratio was informative of the percentage of variance not accounted for in the model for a distance between points smaller than 100 m. While no clear reference value seems to exist concerning interpretation of the ratio, it is safe to assume that the CLT suffered from a low number of observations spread over an extent far wider than the lag distance. Moreover, the accuracy of the CLT empirical variogram was lower compared to the second, more voluminous dataset. This variability in Figure 27a was expected given the distance separating sites and their fewer number.

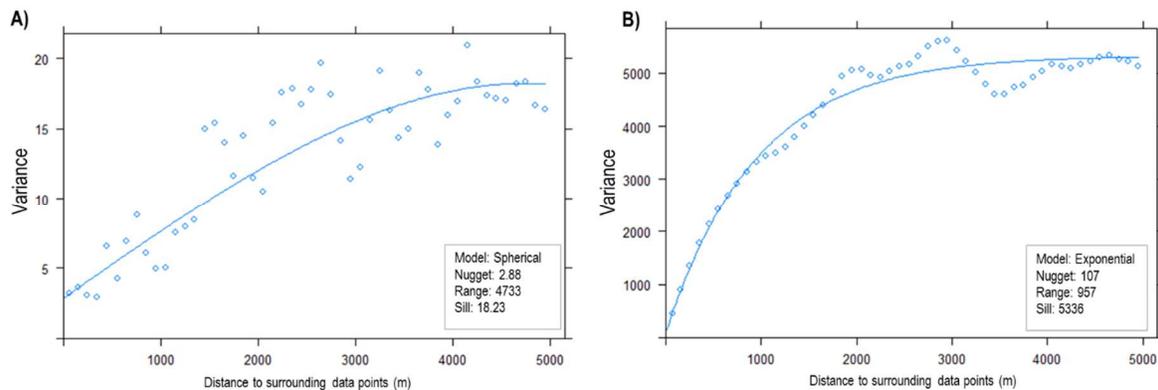


Figure 27 - Experimental variograms (dots) of the two predicted features and their best fit model (line). (A) Variogram of the coprogenous layer thickness using 255 square-root-transformed observations. (B) Variogram of the depth to the mineral layer using 4,488 observations.

Table 14 - Final selection of covariates for the two models after stepwise variance inflation factor and recursive feature selection

Covariate name	Description	Provider	Scorpan factor	Coprogenous model (7)	Mineral model (14)
RTK	Ratio of equivalent thorium to potassium from airborne gamma ray spectrometry	Natural Resources Canada	P		x
RUT	Ratio of equivalent uranium to equivalent thorium from airborne gamma ray spectrometry	Natural Resources Canada	P		x
THO	Surface concentration of equivalent thorium (ppm)	Natural Resources Canada	P	x	x
Buffer	Distance layer of 10 m buffers generated from each peatland's boundaries	-	N	x	x
Center	Distance layer of 10 m buffers generated from each peatland's center	-	N		x
ElevPercent219	Elevation percentile from a 2190 m filter kernel	Derived from DEM	R		x
LBand2_med	Multitemporal median of Landsat's band 2 over 3 dates	USGS Earth Explorer	O	x	x
LBand7_sd	Multitemporal standard deviation of Landsat's band 7 over 3 dates	USGS Earth Explorer	S		x
MaxDiffMeanScale2187	Maximum Difference from Mean Elevation for a maximum search neighbourhood radius of 21870 m	Derived from DEM	R		x
MaxElevDevScale2187	Maximum elevation deviation for a maximum search neighbourhood radius of 21870 m	Derived from DEM	R	x	
MaxElevDevScale656	Maximum elevation deviation for a maximum search neighbourhood radius of 6560 m	Derived from DEM	R		x
MRRTF	Multiresolution index of the ridge top flatness	Derived from DEM	R	x	x
MRVBF	Multiresolution index of valley bottom flatness	Derived from DEM	R	x	x
RSP	Relative slope position	Derived from DEM	R		x
SSoilMoist_med	Multitemporal median of the soil moisture over 4 dates using Sentinel-2 images	Copernicus Open Access Hub	S	x	x

The range for the CLT was 4,733 m and 957 m for the DML. This range provides the maximum distance at which two measurements of the same property are related. The results indicate the presence of autocorrelation amongst both features that need to be captured during model tuning. One should note that for the exponential model, it is actually the effective range (lag corresponding to 95 % of the sill) that was defined.

To account for autocorrelation during the cross-validation step, the ranges obtained above were used to generate spatial clusters of sites. The study area was divided in squares using a grid of 1,000, 1,500 and 2,000 m for comparison. For each model, LOBOCV was performed. This approach was compared to a non-spatial 10-fold cross-validation. The latter returned overoptimistic values of concordance and RMSE compared to LOBOCV (data not shown), reinforcing the need of a spatial cross-validation approach. It was determined that a distance of 1,500 m gave better results and was used to form blocks.

3.5.3 Accuracy assessment and model selection

Internal validation results of the three MLAs are summarized in Table 15. Models were compared based on their CCC and RMSE obtained after LOBOCV. It was clear that CLT models underperformed compared to DML models. While RMSE is similar, CCC is lower. This result could indicate potential bias or low correlation with the covariates.

The Cubist model had the best performance for CLT predictions, with a CCC = 0.07 and RMSE = 30 cm (other CLT tuning statistics not shown). For the DML predictions, model performance of the Cubist model (CCC = 0.69 and RMSE 28 cm) was similar to the RF model (CCC = 0.68 and RMSE = 26 cm). Both models had visible artifacts that could be seen in Figure 30. The predicted range and presence of artifacts lead to the selection of the Cubist model as the final DML model. The range of predicted values by Cubist was closer than RF's to the actual range of observed values.

Table 15 – Tuning and leave-one-block-out cross validation results for the models. Results for the final model of the coprogenous layer thickness (CLT) and for the three MLAs for the depth to the mineral layer (DML) are shown. Standard deviation of Concordance and of root mean square of error (RMSE) are in parentheses. RMSE, observed and predicted range are in centimeter.

<i>N</i>	Feature	Model	Hyperparameter	Concordance	RMSE	Observed range	Predicted range
255	CLT	Cubist	committees = 50 neighbors = 5	0.07 (0.23)	30 (29)	[0, 206]	[0, 210]
		Cubist	committees = 100 neighbors = 5	0.69 (0.29)	28 (18)		[1, 390]
4488	DML	KNN	k = 4	0.62 (0.30)	34 (22)	[5, 392]	[7, 353]
		RF	mtry = 6	0.68 (0.24)	26 (12)		[13, 331]

To assess the impact of LOBOCV at controlling autocorrelation in the datasets, residuals variograms were produced for both final models (Figure 28). CLT residuals were transformed by subtracting the minimum value to all residuals to ensure they were positive, then their square root was computed. Skewness was reduced from 1.51 to -0.23 by this procedure. DML residuals were not transformed (skewness = 0.03).

Concerning DML, Figure 28b almost shows a pure nugget effect, indicating low autocorrelation in the residuals. On the other hand, a more significant range of 594 m remains for CLT residuals, meaning that the grid size might not have been of the correct dimension. Nonetheless, autocorrelation was significantly reduced compared to the usual use of k-fold cross-validation without spatial consideration when forming training and validation folds. Indeed, overoptimistic concordance up to 0.93 were obtained with non-spatial cross-validation techniques (data not shown).

Since both final models are selected, prediction maps have been produced and are shown in Figures 29 and 30. Other DML models' map are presented to allow visual comparison between the outputs.

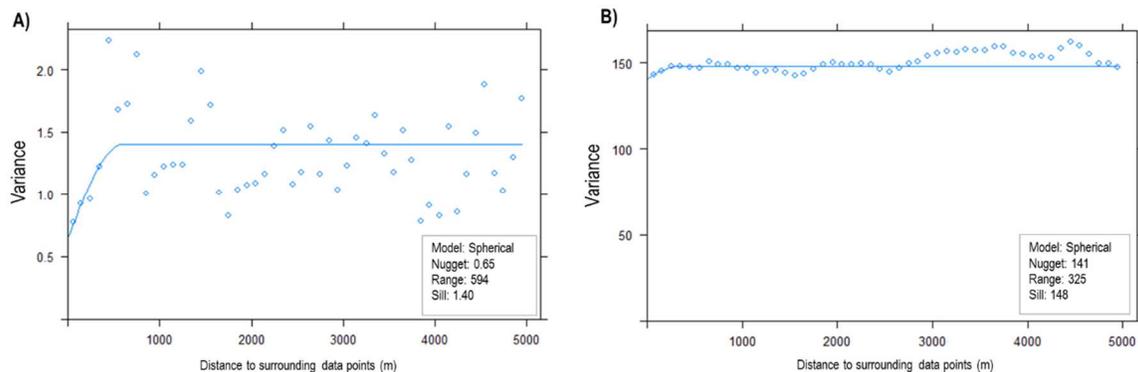


Figure 28 - Experimental residuals variogram (dots) of the two predicted features and their best fit model (line). (A) Variogram of the coprogenous layer thickness residuals using square-root-transformed observations. (B) Variogram of the depth to the mineral layer residuals. Both prediction residuals are from Cubist models.

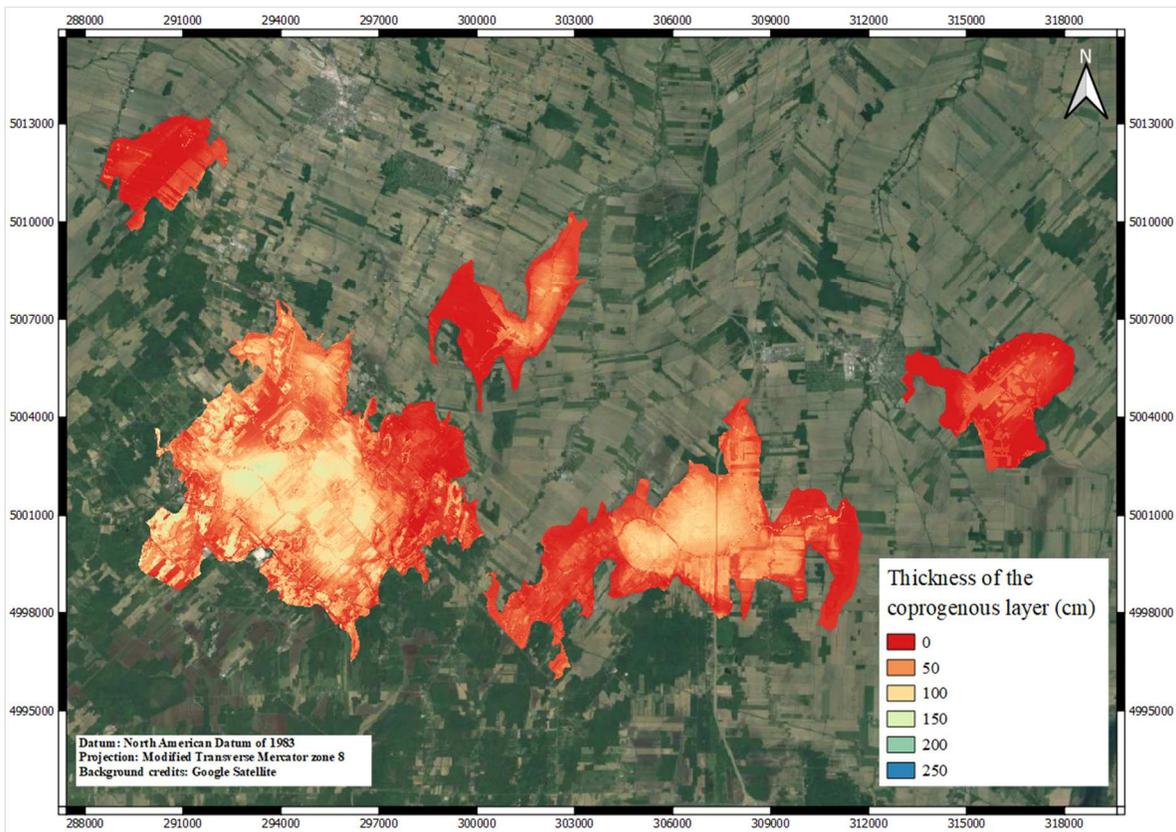


Figure 29 – Map of the predicted coprogenous layer thickness (cm) across the study area using the Cubist model.

According to Figure 29, thicker layers of coprogenous soil can be found in the two lower peatlands and tend to be thicker towards the center of those peatlands. This trend is consistent with field observations. Nevertheless, no uniform gradient can be found, indicating localized accumulation sites across the peatlands. The reader can still see important artifacts related to the MaxElevDevScale2187 covariate present in the final model. It is also interesting to note a prediction range of 0-60 cm of coprogenous material in the top-left peatland that is reportedly exempt of such layer according to our field sampling survey.

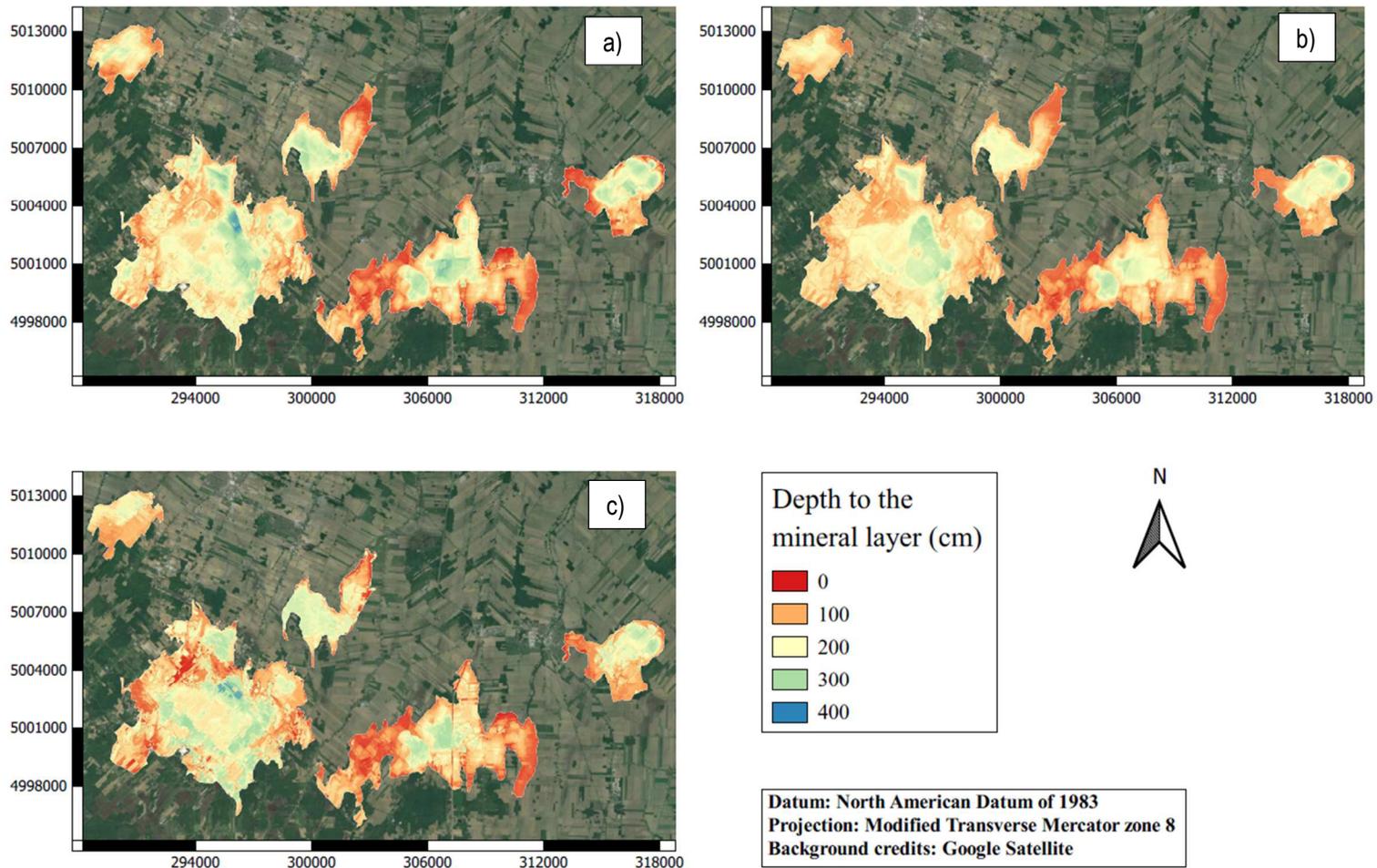


Figure 30 – Depth to the mineral layer (cm) predicted by four models: a) Cubist, b) Random Forest and c) k-Nearest Neighbor.

Concerning predicted DML, a smoother set of maps can be the result of more observations in the models (Figure 30). All four maps seem to share similar gradients (e.i., similar zones are deep and similar zones are shallow), except KNN's map depicting a different gradient in the center of the southwest peatland. It is interesting to note that the deepest zone can be found closer to the boundary of the peatland than to its center. Similarities between different models are mainly found in areas with a greater density of observations.

3.5.4 Covariates importance in the final models

Concerning the use of covariates during model training, relative variable importance plots are shown in Figure 31. These plots show the percentage of times where each covariate was used in a Cubist tree. For the CLT, MRVBF (100 %) contributed the most to the final model, followed by MaxElevDevScale2187 (86 %) and LBand2_med (78 %). Then, a sharp decline in relative importance is observed for MRRTF (44 %), SSoilMoist_med (8 %), Buffer (8 %) and THO (0 %).

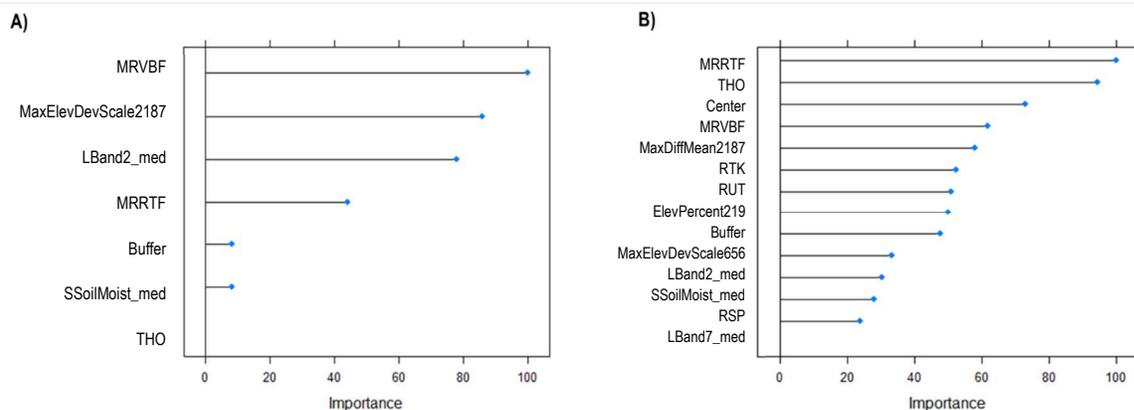


Figure 31- Relative importance of each covariate (%) on the final model A) of the coprogenous layer thickness and B) of the depth to the mineral layer. Refer to Table 14 for a description of covariates' abbreviations.

MRRTF (100 %) was the most used covariate in the final DML model, followed by THO (94 %), Center (73 %), MRVBF (62 %), MaxDiffMean2187 (57 %), RTK (52 %), RUT (51%), ElevPercent219 (50 %), Buffer (48 %), MaxElevDevScale656 (33 %), LBand2_med (30 %), SSoilMoist_med (28 %), and RSP (24 %).

The most important CLT covariates are related to relative height of the soil for a given spatial window. Despite being levelled, the fields on a medium to large scale express variation of altitude correlated with the accumulation of coprogenous material. The multitemporal median of Landsat's band 2 is also a significant predictor. Band 2 is associated with blue visible light wavelengths. It is unclear how this covariate can effectively predict the thickness of a buried coprogenous layer. One hypothesis could be the ease to separate organic soil from mineral soil and forests. To some extent, one could argue that the interpretation of the relevance of

covariates in the final CLT model is useless since its predictive power was almost none ($CCC = 0.07$). This could suggest, amongst other possible causes, that tested remote sensing tools, except DEM-related covariates, are not useful to predict this feature that is more related to terrain and spatial trends. Once again, the number of observations and their coverage of the study area might have been another limiting factor to model performance.

In the DML top covariates, gamma radiometrics and DEM-related covariates are predominant. MRRTF was the most important covariate. As stated by Gallant & Dowling (2003), MRRTF and MRVBF can be important predictors for hydrologic and geomorphic processes that are related to valley, depressions, and slopes, like sedimentary soil deposits (i.e., coprogenous material and later on peat deposits). Distance to the center of each peatland was also an important predictor, meaning higher deposits are generally found near the center and thinner deposits near the peatland border.

3.5.5 Regional peat maximum thickness of cultivation map and its uncertainty

Finally, PMTC map at a regional scale was derived by subtraction CLT's map to DML'S map to reflect the real thickness of the peaty layer (Figure 32). PMTC ranges from -5 to 348 cm. Approximately 0.1% of the PMTC map are cells with a negative PMTC. This is due to a predicted coprogenous layer thicker than the predicted depth to the mineral layer. As previously stated, higher imprecision in the CML model linked to a poor coverage of the study area and a weak relationship between covariates and the predicted feature could be responsible for model uncertainty. Peatlands differ quite significantly on their PMTC. This has major implications from a soil conservation perspective, because degraded and shallow soils are not uniformly distributed nor confined to a peatlands' border. Therefore, farmers are not affected equally and will have to use conservation measures accordingly to their situation.

Concerning the uncertainty maps (Figure 33), the upper prediction limit ranges between 29 and 909 cm, while the lower prediction limit ranges between -406 cm and -22 cm. With increasing PMTC, uncertainty tends to be equal or higher. The high imprecision can be attributed to the propagation of error between DML and CLT models. Data clusters and poor spatial coverage of the study area's extent could have limited the model performance in less sampled areas.

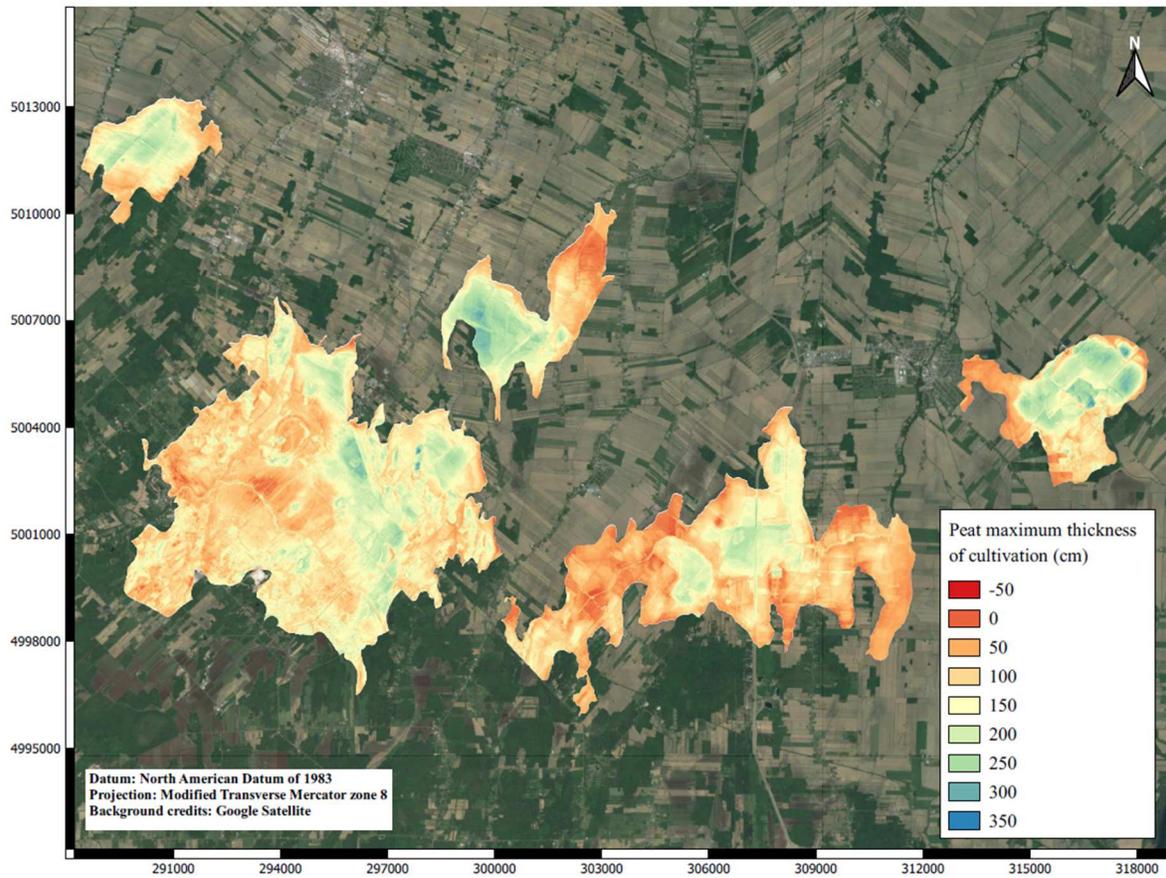


Figure 32 - Peat maximum thickness of cultivation map obtained by subtracting the coprogenous layer thickness to the depth to the mineral layer

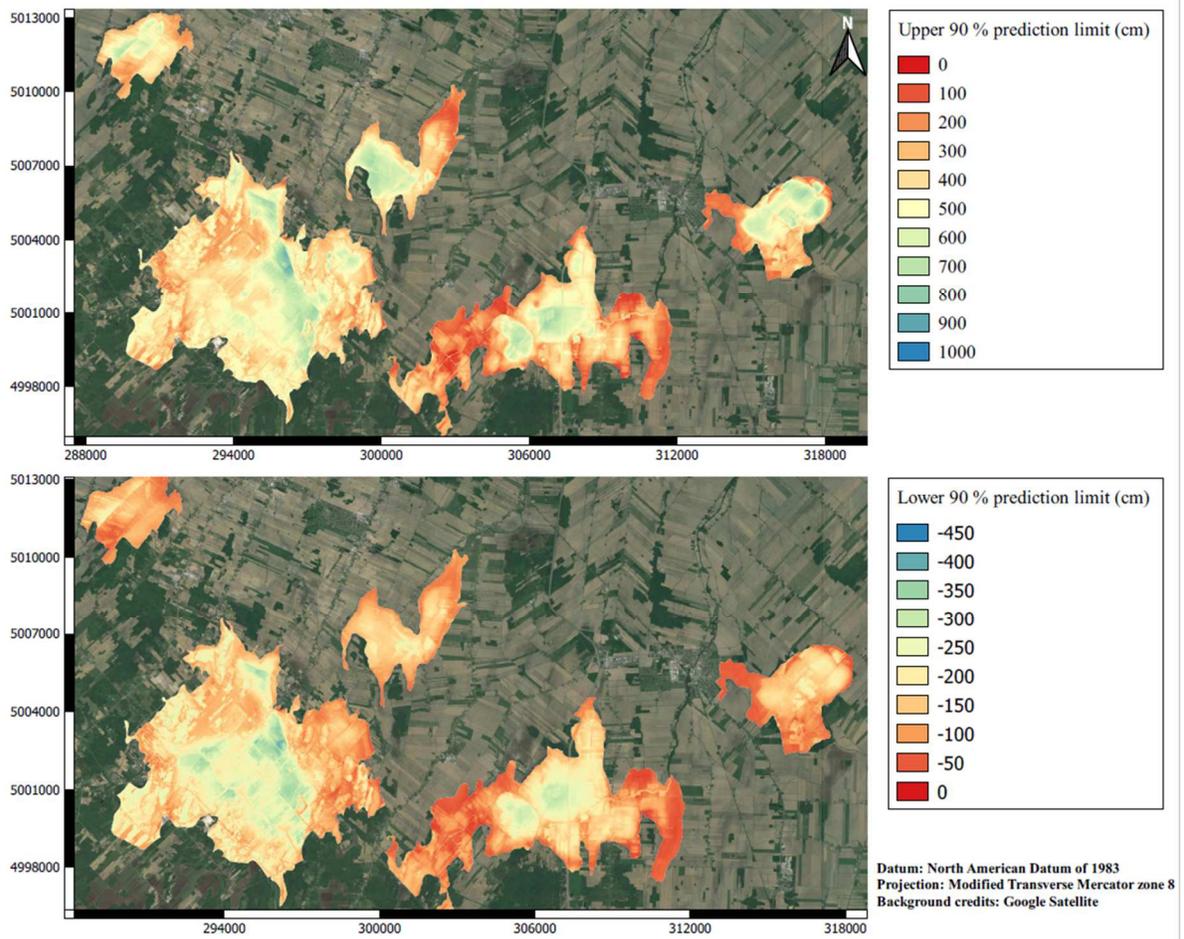


Figure 33 - Peat maximum thickness of cultivation uncertainty maps. A) Upper and B) Lower bounds of the 90 % prediction interval.

3.6 Discussion

3.6.1 Relevant covariates to predict coprogenous layer thickness and depth to the mineral layer

The use of open access, regional covariates has proven useful for predicting DML, but an underperformance of CLT models was observed. As recently covered in Wadoux et al. (2021), spatial cross-validation of autocorrelated data is not always the solution to obtaining a representative assessment of the bias of a model. Standard cross-validation might yield better results without being overoptimistic. This could partially explain why the CCC obtained with LOBOCV for the CLT model was as low as 0.07 in comparison with standard cross-validation. Rudiyanto et al. (2018) tested 14 machine-learning models to map peat thickness in Indonesian peatlands. They were able to predict DML with an average error of 30 cm with 159 observations at a regional scale (50,000 ha), similarly to the 31 cm RMSE of Gatis et al. (2019) with 1779 observations. This error is of the same magnitude than the ones observed in this study during model validation. Furthermore, the same authors had better results with Cubist and RF models, and MRVBF was ranked 2nd most important covariate. This supports the importance of this DEM-derived covariate on the prediction of deposited materials. Some soil-forming factors affect limnic and organic deposits differently depending on the scale at which phenomenon are studied (Behrens et al., 2018). Multiresolution covariates can provide this information to the model to allow a better understanding of the magnitude of the deposits.

Furthermore, gamma radiometrics contribution was expected for delineating peat extent and predicting peat thickness (Minasny et al., 2019). The same authors also suggested multi-temporal satellite covariate might provide useful information on peat thickness. In this study, most of Landsat 8 and Sentinel 2 predictors were highly correlated, possibly due to a short time window. The ones that made it to the RFE step were set aside for their low additive prediction effectiveness. Nonetheless, it seems that the multi-temporal median of Landsat 8 band 2 was able to provide useful information to both models. DML had more observations but had twice the number of covariates CLT model had. A weak relationship between this feature and the suit of tested covariate may explain why so few were kept in the final model. This could potentially suggest that other sampling surveys should be made at a shorter scale. The benefits of the use of proximal sensing tools could be assessed in the future as more relevant covariates.

The study area encompasses four pedological surveys for which digital maps are freely available online. Legacy soil data were considered at first due to their popular use in digital soil mapping but were not included as covariates due to temporal incoherence between the surveys (1950, 2000, 2001 and 2014). Evolution of Organic soils over time can lead to substantial biases and render maps useless, even though they can be updated with modern techniques (Kempen et al., 2009). Otherwise, this covariate could have been useful to predict areas with coprogenous soil and to delineate peat extent.

3.6.2 Peat maximum thickness of cultivation as a tool to guide soil conservation practices and its uncertainty

The results suggested a potential limited use of PMTC maps produced at a regional scale for soil management purposes given the high uncertainty of the combined CLT and DML predictions. Artifacts in DML and CLT maps create sharp transitions between neighboring cells that are not seen in nature. That can affect interpretation of the final PMTC map for fields near them or by propagating error that could be covered by combining maps. This is heavily limiting its use for precision agriculture and for conservation projects. Moreover, the accuracy is not sufficient to predict at which depth drain tiles could be put or if soil conservation PMTC thresholds for intervention are reached. Even the average RMSE of 30 cm is not accurate enough.

Many sources of uncertainty are present in this study (Heuvelink, 2017). Two methods of sampling with a different precision were combined. In addition, the dataset from 2010 was mainly clustered in two peatlands. Since model tuning will inevitably fit the model on a majority of points coming from those peatlands, the final model might not be as good to predict the features found in the other three peatlands if their formation and evolution differs significantly (i.e., botanical origin, soil management, groundwater level variations based on differences related to the watershed and elevation, average reclamation year, etc.). Pre-processing of the original data to produce covariates introduces error in the model as well (e.i., resampling, projecting, and smoothing raster layers). Indeed, as stated in Samuel-Rosa (2015), the number and the localization of calibration points matter, like the fact that covariates are only approximations of real soil-forming conditions. These errors are complicated if not impossible to quantify and are propagated at each new step of the workflow. A better sampling design would have been needed from the start to cover the full feature space of DML and CLT of the study area, and a better spatial distribution of observation points.

Therefore, PMTC should be used with caution, as one decision tool amongst others to manage Organic soils conservation. Yet, not considering coprogenous materials would greatly overestimate the remaining soil resource that can be used to produce crops.

3.7 Conclusions

1. The proposed selection of freely available covariates and workflow did not provide a sufficient correlation with the coprogenous layer thickness to be able to produce maps that can be used for precision agriculture (soil conservation plans and drainage plans).
2. This study was able to provide the first map of the depth to the mineral layer for this region with a fairly good accuracy, comparable to other DSM studies.

3. The map of peat maximum thickness of cultivation at a regional scale can be used as a guide to determine priority intervention zones, along with other decision tools, although it should be used with caution.
4. Proximal sensing tools must be investigated as a more relevant source of covariates to produce coprogenic layer thickness maps at a local scale instead of relying on remote sensing tools at a regional scale.

Acknowledgements

The authors are thankful for the financial support from a Canadian Graduate Scholarship program by the Natural Sciences and Engineering Research Council of Canada (NSERC) and a Master's Scholarship program (B1X) by the Fonds de recherche du Québec – Nature et technologies granted to R. Deragon.

We also acknowledge the financial support of the NSERC through an Industrial Research Chair Grant in Conservation and Restoration of Cultivated Organic Soils (IRCPJ 411630-17) in partnership with Delfland Inc, Productions maraîchères Breizh inc., La Production Barry inc., Les Fermes R.R. et fils inc., Le Potager Montréalais Itée, R. Pinsonneault et fils Itée, Patate Isabelle inc., Les Fermes du Soleil Inc., Les Jardins A. Guérin et Fils inc., Le Potagers Riendeau inc., Vert Nature Inc, Fermes Hotte et Van Winden Inc, Production Horticole Van Winden and Maraichers J.P.L. Guerin & fils.

The authors would also like to thank Michaël Leblanc and Lucie Grenon for their early contribution to this project and John Lindsay for his advice on DEM pre-processing.

References

- Beamish, D. (2013). Gamma ray attenuation in the soils of Northern Ireland, with special reference to peat. *J. Environ. Radioact.* 115, 13–27.
- Behrens, T., Schmidt, K., Viscarra Rossel, R. A., Gries, P., Scholten, T., & MacMillan, R. A. (2018). Spatial modelling with euclidean distance fields and machine learning. *European Journal of Soil Science*, 69(5), 757–770. <https://doi.org/10.1111/ejss.12687>
- Gräler, B., Pebesma, E., & Heuvelink, G. (2016). Spatio-Temporal Interpolation using gstat. *The R Journal* 8(1), 204-218
- Beucher, A., Koganti, T., Iversen, B. V., & Greve, M. H. (2020). Mapping of peat thickness using a multi-receiver electromagnetic induction instrument. *Remote Sensing*, 12(15), 2458–2458. <https://doi.org/10.3390/rs12152458>
- Beven, K.J., & Kirkby, M.J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.* 24, 43–69.
- Bian, Z., Guo, X., Wang, S., Zhuang, Q., Jin, X., Wang, Q., & Jia, S. (2020). Applying statistical methods to map soil organic carbon of agricultural lands in northeastern coastal areas of china. *Archives of Agronomy and Soil Science*, 66(4), 532–544. <https://doi.org/10.1080/03650340.2019.1626983>
- Breiman, L. (2001). “Random Forests.” *Machine Learning*, 45, 5–32.
- Brenning, A., Bangs, D., & Becker, M. (2018). RSAGA: SAGA Geoprocessing and Terrain Analysis. R package version 1.3.0. <https://CRAN.R-project.org/package=RSAGA>
- Comas, X., Terry, N., Slater, L., Warren, M., Kolka, R., Kristijono, A., Sudiana, N., Nurjaman, D., & Darusman, T. (2015). Imaging tropical peatlands in Indonesia using ground penetrating radar (GPR) and electrical resistivity imaging (ERI) : implications for carbon stock estimates and peat soil characterization. *Biogeosciences*, 11(10), 2995–3007.
- Dessureault-Rompré, J., Libbrecht, C., & Caron, J. (2020). Biomass crops as a soil amendment in cultivated histosols: can we reach carbon equilibrium? *Soil Science Society of America Journal*, 84(2), 597–608. <https://doi.org/10.1002/saj2.20051>
- Escadafal, R. (1994). Soil spectral properties and their relationships with environmental parameters-examples from arid regions. In *Imaging Spectrometry—A Tool for Environmental Observations*; Springer: Dordrecht, The Netherlands. pp. 71–87.
- Gallant, J.C., & Dowling, T.I. (2003). A multiresolution index of valley bottom flatness for mapping depositional areas. *Water Resources Research*, 39/12:1347-1359

- Gao, B.-c. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3), 257–266. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3)
- Gatis, N., Luscombe, D. J., Carless, D., Parry, L. E., Fyfe, R. M., Harrod, T. R., Brazier, R. E., & Anderson, K. (2019). Mapping upland peat depth using airborne radiometric and lidar survey data. *Geoderma*, 335, 78–87. <https://doi.org/10.1016/j.geoderma.2018.07.041>
- Genuer, R., & Poggi, J.-M. (2020). *Random forests with r (Ser. Use r!)*. Springer. <https://doi.org/10.1007/978-3-030-56485-8>
- Gholizadeh, A., Ziz̄ala Daniel, Saberioon, M., & Boruřka Luboš. (2018). Soil organic carbon and texture retrieving and mapping using proximal, airborne and sentinel-2 spectral imaging. *Remote Sensing of Environment*, 218, 89–103. <https://doi.org/10.1016/j.rse.2018.09.015>
- Groupe AGÉCO. (2007). *Portrait et priorités du secteur marařcher quřbecois; Rapport final*. [En ligne] https://www.mapaq.gouv.qc.ca/fr/Publications/Portrait_secteurmaraicher.pdf
- Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning : data mining, inference, and prediction (Second edition, corrected 7th printing, Ser. Springer series in statistics)*. Springer.
- Heung, B., Ho, H. C., Zhang, J., Knudby, A., Bulmer, C. E., & Schmidt, M. G. (2016). An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. *Geoderma*, 265, 62–77. <https://doi.org/10.1016/j.geoderma.2015.11.014>
- Heuvelink, G. (2017). *Uncertainty. Soil Organic Carbon Mapping Cookbook*. 1st ed. Rome, Food and Agriculture Organisation of the United Nations. pp. 109–121.
- Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3), 295–309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X)
- James, G., Witten, D., Hastie, T., & Tibshirani, R., (2014). *An introduction to statistical learning : with applications in r (Corrected at 4th printing 2014, Ser. Springer texts in statistics)*. Spriger.
- Keaney, A., McKinley, J., Graham, C., Robinson, M., & Ruffell, A. (2013). Spatial statistics to estimate peat thickness using airborne radiometric data. *Spatial Statistics* 5, 3–24.
- Khan, N.M., Rastoskuev, V.V., Sato, Y., & Shiozawa, S. (2005). Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators. *Agric. Water Manag.* 77, 96–109.
- Ku, H. H. (1966). Notes on the use of propagation of error formulas. *Journal of Research of the National Bureau of Standards*, 70(4), 263-273.
- Kuhn, M. (2020). *caret: Classification and Regression Training*. R package version 6.0-86. <https://CRAN.R-project.org/package=caret>

- Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer. <https://doi.org/10.1007/978-1-4614-6849-3>
- Lawrence, I., & Lin, K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45, 255–268.
- Lindsay, J. B. (2016). Efficient hybrid breaching-filling sink removal methods for flow path enforcement in digital elevation models. *Hydrological Processes*, 30(6), 846–857. <https://doi.org/10.1002/hyp.10648>
- Lindsay, J. B., Cockburn, J. M. H., & Russell, H. A. J. (2015). An integral image approach to performing multi-scale topographic position analysis. *Geomorphology*, 245, 51–61. <https://doi.org/10.1016/j.geomorph.2015.05.025>
- Malone, B. P., Minasny, B., & McBratney, A. B. (2017). *Using r for digital soil mapping* (Ser. Progress in soil science). Springer Nature. <https://doi.org/10.1007/978-3-319-44327-0>
- Minasny, B., & McBratney, A. B. (2006). A conditioned latin hypercube method for sampling in the presence of ancillary information. *Computers and Geosciences*, 32(9), 1378–1388. <https://doi.org/10.1016/j.cageo.2005.12.009>
- Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., & Nauss, T. (2018). Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environmental Modelling & Software*, 101, 1–9.
- Naimi, B., Hamm, N.A.S., Groen, T.A., Skidmore, A.K., & Toxopeus, A.G. (2014). Where is positional uncertainty a problem for species distribution modelling?, *Ecography* 37 (2): 191-203.
- Natural Resources Canada. (2019). Geoscience data repository for geophysical data. Magnetic-Radiometric-EM datasets [WWW Document]. URL. <http://gdr.agg.nrcan.gc.ca/gdrdap/dap/search-eng.php>.
- Nellis, M. D., & Briggs, J. M. (1992). Transformed vegetation index for measuring spatial variation in drought impacted biomass on konza prairie, kansas. *Transactions of the Kansas Academy of Science (1903-)*, 95(1-2), 93–99.
- O'brien, R. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity*, 41(5), 673–690.
- Oliver, M. A., & Webster, R. (2014). A tutorial guide to geostatistics: computing and modelling variograms and kriging. *Catena*, 113, 56–69. <https://doi.org/10.1016/j.catena.2013.09.006>
- Pebesma, E.J. (2004). Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*, 30: 683-691.

- Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois Gaëlle, Bayol, N., Lyapustin, A., Gourlet-Fleury, S., & Pélissier, R. (2020). Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nature Communications*, 11(1). <https://doi.org/10.1038/s41467-020-18321-y>
- Pohjankukka, J., Pahikkala, T., Nevalainen, P., & Heikkonen, J. (2017). Estimating the prediction performance of spatial models via spatial k-fold cross validation. *International Journal of Geographical Information Systems*, 31(10), 2001–2019.
- Poggio, L., & Gimona, A. (2017). Assimilation of optical and radar remote sensing data in 3d mapping of soil properties over large areas. *Science of the Total Environment*, 579, 1094–1110. <https://doi.org/10.1016/j.scitotenv.2016.11.078>
- Quinlan, J.R. (1992) Learning with Continuous Classes. Proceedings of Australian Joint Conference on Artificial Intelligence, Hobart 16-18 November 1992, 343-348.
- Quinlan, J.R. (1993) Combining instance-based and model-based learning. Proceedings of the Tenth International Conference on Machine Learning pp. 236-243
- Rikimaru, A.; Roy, P.S.; Miyatake, S. Tropical forest cover density mapping. *Trop. Ecol.* **2002**, 43, 39–47. [\[Google Scholar\]](#)
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillerá-Aroita, G., Hauenstein, S., Lahoz-Monfort José J, Schröder Boris, Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929. <https://doi.org/10.1111/ecog.02881>
- Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. NASA special publication, 351(1974), 309.
- Rosa, E., Larocque, M., Pellerin, S., Gagné S, & Fournier, B. (2009). Determining the number of manual measurements required to improve peat thickness estimations by ground penetrating radar. *Earth Surface Processes and Landforms*, 34(3), 377–383.
- Rudiyanto, Minasny, B., Setiawan, B. I., Saptomo, S. K., & McBratney, A. B. (2018). Open digital mapping as a cost-effective method for mapping peat thickness and assessing the carbon stock of tropical peatlands. *Geoderma*, 313, 25–40. <https://doi.org/10.1016/j.geoderma.2017.10.018>
- Saurette, D. (2021). onsoilsurvey: Making PDSM in Ontario Better. R package version 0.0.0.9000.

- Schratz, P., Muenchow, J., Iturrutxa, E., Richter, J., & Brenning, A. (2019). Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecological Modelling*, 406, 109–120. <https://doi.org/10.1016/j.ecolmodel.2019.06.002>
- Simon, B., Ibs-von Seht, M., & Frank, S. (2020). Airborne electromagnetic and radiometric peat thickness mapping of a bog in northwest Germany (Ahlen-Falkenberger Moor). *Remote Sensing*, 12(2), 203–203. <https://doi.org/10.3390/rs12020203>
- Soil Classification Working Group (SCWG). (1998). *The Canadian System of Soil Classification*, 3rd edition. Agriculture and Agri-Food Canada, Publication 1646, 187 pp.
- Taghizadeh-Mehrjardi, R., Emadi, M., Cherati, A., Heung, B., Mosavi, A., & Scholten, T. (2021). Bio-Inspired Hybridization of Artificial Neural Networks: An Application for Mapping the Spatial Distribution of Soil Texture Fractions. *Remote Sens.* 13, 1025. <https://doi.org/10.3390/rs13051025>
- Taghizadeh-Mehrjardi, R., Toomanian, N., Khavaninzadeh, A., Jafari, A., & Triantafyllis, J. (2016). Predicting and mapping of soil particle-size fractions with adaptive neuro-fuzzy inference and ant colony optimization in central Iran. *Eur. J. Soil Sci.* 67, 707–725.
- Tveite, H. 2018. The QGIS Multi-distance buffer Plugin, Version 3.2.4. <http://plugins.qgis.org/plugins/MultiDistanceBuffer/>
- Wadoux, A. M. J. C., Heuvelink, G. B. M., Bruin, D. S., & Brus, D. J. (2021). Spatial cross-validation is not the right way to evaluate map accuracy. *Ecological Modelling*, 457.
- Webster, R., & Oliver, M. A. (2007). *Geostatistics for environmental scientists* (2nd ed., Ser. Statistics in practice). John Wiley & Sons.
- Wu, Q. (2020). whitebox: 'WhiteboxTools' R Frontend. R package version 1.4.0/r27. <https://R-Forge.R-project.org/projects/whitebox/>
- Young, N. E., Anderson, R. S., Chignell, S. M., Vorster, A. G., Lawrence, R., & Evangelista, P. H. (2017). A survival guide to Landsat preprocessing. *Ecology*, 98(4), 920–932. <https://doi.org/10.1002/ecy.1730>
- Zhang, C., & Ma, Y. (2012). *Ensemble machine learning : methods and applications*. Springer. <https://doi.org/10.1007/978-1-4419-9326-7>

Conclusion générale

De précédents travaux de recherche de l'équipe proposaient l'existence de 4 groupes de sols organiques, aux comportements physiques et hydrodynamiques variés. Ce projet de maîtrise s'est intéressé à l'existence de groupes de sols qui diffèrent quant à leurs propriétés afin de former des groupes de gestion dans un contexte de conservation des sols. Au terme d'une campagne d'échantillonnage, un portrait initial des propriétés des sols de 14 fermes partenaires a été réalisé. Les propriétés évaluées ont été intégrées dans une base de données permettant le suivi de leur évolution et leur intégration dans divers projets cartographiques en lien avec la conservation des ressources en sol.

Les résultats obtenus suggèrent plutôt la mise en place de trois groupes sur la base d'un seuil statistique et d'un seuil pratique, tous deux axés sur la profondeur de sol organique cultivable. Ces seuils sont facilement interprétables sur le terrain, et supportés par des analyses en laboratoire de 17 propriétés qui agissent à titre de propriétés discriminantes. L'étude simultanée de ces propriétés permettra de classifier l'état de dégradation de nouveaux sites et permet de limiter l'échantillonnage nécessaire.

Plusieurs approches ont été tentées afin de former des groupes de gestion fondés sur l'impact de la dégradation de sols sur leurs propriétés physiques, chimiques et pédologiques. La première hypothèse sous-entendait l'utilisation d'approches simples basées sur une variable empirique de la qualité des sols ou encore le grand groupe pédologique d'un site. Cette tentative s'est montrée biaisée et limitée face au processus continu de dégradation des sols.

Ensuite, des approches statistiques plus complexes ont été explorées. Une analyse de classe latente a été incapable de former des groupes avec le jeu de données récoltées. C'est finalement une MANOVA qui a permis de déterminer un seuil de profondeur, soit 60 cm, à partir duquel des propriétés fortement liées à la dégradation des sols montraient des signes importants de détérioration. Par la suite, un seuil pratique de 100 cm a été formulé, facilitant la sélection des zones d'intervention. Les trois groupes de gestion permettent le respect de la deuxième hypothèse de recherche.

Pour finir, la distribution spatiale de ces groupes a été possible grâce à l'utilisation de méthodes de cartographie numérique modernes et de covariables environnementales disponibles gratuitement en ligne. L'hypothèse 3 est validée pour le sol minéral, mais est invalidée pour le coprogène. Cela limite l'utilisation des cartes de l'épaisseur de sol coprogène et de l'épaisseur de sol cultivable. Elles permettent cependant d'observer les grandes tendances spatiales et seront de nouveaux outils d'aide à la décision à intégrer dans le plan de conservation des sols. Ces travaux supportent l'utilisation de la profondeur de sol cultivable comme variable clé liée à la dégradation des sols organiques cultivés, variable qui doit toutefois être explorée davantage.

Bibliographie générale

- Anshari, G., Affudin, M., Nuriman, M., Gusmayanti, E., Arianie, L., Susana, R., . . . Rafiastanto, A. (2010). Drainage and land use impacts on changes in selected peat properties and peat degradation in West Kalimantan Province, Indonesia. *Biogeosciences*, 7(11), 3403-3419. doi:10.5194/bg-7-3403-2010
- Brandyk, T., Szatyłowicz, J., Oleszczuk, R., & T. Gnatowski. (2003). Water-related physical attributes of organic soils. *Organic soils and peat materials for sustainable agriculture*. CRC Press, Boca Raton, Florida. pp. 33-66.
- Carter, M. R., & Gregorich, E. G. (2007). *Soil sampling and methods of analysis* (2nd ed.). Boca Raton, FL: CRC Press.
- Esselami, D., M. Boudache et L. Grenon. (2014). L'évolution des terres noires et le problème de la compaction. Prisme Consortium. Présentation donnée aux journées horticoles. https://www.mapaq.gouv.qc.ca/SiteCollectionDocuments/Regions/Monteregion-Ouest/Journees_horticoles_2014/4_decembre/Terres_noires/9h05_b_JH2014_profil_compaction_DEsselami.pdf (PDF consulté en ligne le 15 avril 2020).
- FAO (2020). Soil Maps and Databases : Other Global Soil Maps and Databases. [En ligne] Repéré à l'adresse suivante : <http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/other-global-soil-maps-and-databases/en/>
- Gnatowski, T., Szatyłowicz, J., Brandyk, T., & Kechavarzi, C. (2010). Hydraulic properties of fen peat soils in Poland. *Geoderma*, 154(3), 188-195. doi:10.1016/j.geoderma.2009.02.021
- Grenon, L. 1988. Répartition des terres humides dans la plaine du Saint-Laurent. Équipe pédologique du Québec, Centre de recherches sur les terres. Direction générale de la recherche, Agriculture Canada, Sainte-Foy. 3 cartes à l'échelle 1 : 250 000.
- Groupe de travail sur la classification des sols (GTCS). (2002). Le système canadien de classification des sols, 3e édition. Direction générale de la recherche, Ministère de l'Agriculture et Agro-Alimentaire Canada. Publication 1646, Ottawa (ON). 196 p.
- Hallema, D., Lafond, J., Périard, Y., Gumiere, S. J., Sun, G., & Caron, J. (2015a). Long-Term Effects of Peatland Cultivation on Soil Physical and Hydraulic Properties: Case Study in Canada. *Vadose Zone Journal*, 14(6). doi:10.2136/vzj2014.10.0147
- Hallema, D. W., Périard, Y., Lafond, J. A., Gumiere, S. J., & Caron, J. (2015b). Characterization of Water Retention Curves for a Series of Cultivated Histosols. *Vadose Zone Journal*, 14(6), 1-8. doi:10.2136/vzj2014.10.0148
- Ilnicki, P. (2003). Agricultural Production Systems for Organic Soil Conservation. *Organic soils and peat materials for sustainable agriculture*. CRC Press, Boca Raton, Florida. pp. 209-221.
- Ilnicki, P., & J. Zeitz. (2003). Irreversible loss of organic soil functions after reclamation. *Organic soils and peat materials for sustainable agriculture*. CRC Press, Boca Raton, Florida. pp. 15-32.

- Ji, W., Adamchuk, V. I., Chen, S., Mat Su, A. S., Ismail, A., Gan, Q., . . . Biswas, A. (2019). Simultaneous measurement of multiple soil properties through proximal sensor data fusion: A case study. *Geoderma*, 341, 111-128. doi:10.1016/j.geoderma.2019.01.006
- Kempen, B., Brus, D. J., Heuvelink, G. B. M., & Stoorvogel, J. J. (2009). Updating the 1:50,000 Dutch soil map using legacy soil data: A multinomial logistic regression approach. *Geoderma*, 151(3), 311-326. doi:10.1016/j.geoderma.2009.04.023
- Kroetsch, D. J., Geng, X., Chang, S. X., & Saurette, D. D. (2011). Organic soils of Canada: Part 1. Wetland organic soils. *Canadian journal of soil science*, 91(5), 807-822.
- Kruger, J. P., Leifeld, J., Glatzel, S., Szidat, S., & Alewell, C. (2015). Biogeochemical indicators of peatland degradation - a case study of a temperate bog in northern Germany. *Biogeosciences*, 12(10), 2861-2871. doi:10.5194/bg-12-2861-2015
- Lamontagne, L., Martin, A. et Nolin, M.C. (2014). Étude pédologique du comté de Napierville (Québec). Laboratoires de pédologie et d'agriculture de précision, Centre de recherche et de développement sur les sols et les grandes cultures, Direction générale des sciences et de la technologie, Agriculture et Agroalimentaire Canada, Québec (Québec).
- LaSalle, P. (1963). Géologie de la région de Verchères. Dépôts meubles. Rapport préliminaire. Ministère des Richesses Naturelles du Québec. 9 pages + 2 cartes au 1 : 50 000.
- McBratney, A. B., Mendonça Santos, M. L., & Minasny, B. (2003). On digital soil mapping. *Geoderma*, 117(1), 3-52. doi:10.1016/S0016-7061(03)00223-4
- Minasny, B., Berglund, O., Connolly, J., Hedley, C., Vries, de F., Gimona, A., Kempen, B., Kidd, D., Lilja, H., Malone, B., McBratney, A., Roudier, P., O'Rourke, S., Rudiyanto, Padarian, J., Poggio, L., Caten, ten, A., Thompson, D., Tuve, C., & Widyatmanti, W. (2019). Digital mapping of peatlands – a critical review. *Earth-Science Reviews*, 196.
- Muthén, L.K. and Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén
- Okruszko, H., & P. Ilnicki. (2003). The moorsh horizons as quality indicators of reclaimed organic soils. *Organic soils and peat materials for sustainable agriculture*. CRC Press, Boca Raton, Florida. pp. 1-14.
- Parent, L.-É., & Gagné, G. (2010). Guide de référence en fertilisation (2e éd.). Québec, Canada : Centre de référence en agriculture et agroalimentaire du Québec (CRAAQ).
- Parry, L. E., West, L. J., Holden, J., & Chapman, P. J. (2014). Evaluating approaches for estimating peat depth. *Journal of Geophysical Research: Biogeosciences*, 119(4), 567-576. doi:10.1002/2013JG002411
- Pawluczuk, J., Stepień, A., & Alberski, J. (2019). Physical and chemical properties of organic soils in connection with habitat conditions and the land use in the Dolina Rzeki Pasłęki Natura 2000 Site. *Journal of Elementology*, 24(2).

QGIS Development Team (2020). QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.org>

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Vepraskas, M. J., & Craft, C. B. (2015). *Wetland Soils: Genesis, Hydrology, Landscapes, and Classification* (2nd ed.): CRC Press.

Annexe 1 - Description des variables de l'étude

Variables chimiques	
Nom	Description
CE	Conductivité électrique (au laboratoire)
Ntot	Azote total (analysé par l'IRDA)
Ctot	Carbone total (analysé par l'IRDA)
Pyro	Indice au pyrophosphate de sodium
MO	Pourcentage de matière organique par perte au feu
pH	pH eau (au laboratoire)
Variables physiques	
Nom	Description
Porotot	Porosité totale
Tea-5kPa	Teneur en air à -5 kPa
REU-10kPa	Réserve en eau utilisable à -10 kPa
Tee-1kPa	Teneur en eau à -1 kPa
Tee-3kPa	Teneur en eau à -3 kPa
ksat	Conductivité hydraulique saturée
MVA	Masse volumique apparente issue des cylindres de Ksat
P_inf_0,84mm	Pourcentage de sol de diamètre inférieur à 0,84 mm après passage dans une série de tamis
RE_moy	Résistance à l'enfoncement moyenne pour les premiers 80 cm
RE_max	RE maximale rencontrée dans les premiers 80 cm
RE_max_d	Profondeur à laquelle la RE_max est observée
RE_max_Ohp	RE maximale rencontrée près de la surface (entre 10 et 40 cm)
RE_max_Ohp_d	Profondeur à laquelle la RE_max_Ohp est observée
Variables pédologiques	
Nom	Description
Copro	Épaisseur de la couche de coprogène
Copro_160	Épaisseur de coprogène dans le premier 1,6 m (coupe témoin du profil de sol)
Cult	Épaisseur cultivable (épaisseur au minéral - épaisseur de coprogène)
FIB	Pourcentage de fibres non frottées pour les couches cultivables, pondéré par l'épaisseur de la couche pour l'ensemble de la profondeur cultivable
FIB_frot	Pourcentage de fibres frottées pour les couches cultivables, pondéré par l'épaisseur de la couche pour l'ensemble de la profondeur cultivable
STT_limit	Nombre de couches limitantes pour le drainage dans les premiers 100 cm (STT = 0, 1, 2, 7, 10)

Note 1: Les _1, _2 et _3 à la suite d'une variable signifient une profondeur différente (respectivement 0-20, 20-40 et 40-60 cm) pour les analyses chimiques.

Note 2: Les _1, _2 et _3 à la suite d'une variable signifient une profondeur différente (respectivement 0-5, 30-35 et 50-55 cm) pour certaines des analyses physiques

Note 3: Les RE ont été filtrées par tranche de 5 cm à l'aide d'un algorithme de correction (Guedessou, 2020). Deux modèles polynomiaux (D3 et D5) ont été utilisés.

Annexe 2 - Statistiques descriptives des variables continues retenues dans les analyses

Variable	n	Moyenne	Écart-type	Médiane	Minimum	Maximum	Skewness	Kurtosis	Unité
CE_1	114	665,55	243,24	638,00	233,00	1630,00	1,25	2,58	μS cm ⁻¹
CE_2	114	552,43	213,38	520,00	118,40	1948,00	3,08	16,63	
CE_3	114	593,37	364,63	507,00	183,70	2721,00	3,55	14,80	
CN_1	114	22,28	3,76	22,05	14,24	35,69	0,58	0,44	-
CN_2	114	22,77	4,28	22,26	14,24	37,66	0,77	0,69	
CN_3	114	23,07	4,92	22,81	12,50	40,06	0,32	0,20	
Copro	114	43,68	44,92	32,00	0,00	206,00	0,91	0,39	cm
Copro_160	114	14,54	23,14	0,00	0,00	89,00	1,44	0,94	
Ctot_1	114	42,30	4,37	42,80	24,50	49,96	-1,54	3,40	% (p/p)
Ctot_2	114	42,93	4,60	43,50	22,50	51,22	-1,81	5,44	
Ctot_3	114	42,44	9,71	45,10	2,90	51,28	-2,68	6,65	
Cult	114	150,32	63,70	158,00	27,00	287,00	-0,05	-0,80	cm
FIB	114	21,14	9,52	20,59	0,00	47,11	0,15	0,53	% (v/v)
FIB_frot	114	7,77	4,89	7,38	0,00	28,16	1,32	3,27	
ksat_1	114	0,03	0,03	0,02	0,00	0,14	1,75	3,69	cm s ⁻¹
ksat_2	114	0,02	0,02	0,02	0,00	0,10	1,57	2,43	
ksat_3	114	0,03	0,05	0,01	0,00	0,28	3,06	10,81	
MO_1	114	82,67	8,08	84,66	49,17	92,63	-2,06	5,03	% (p/p)
MO_2	114	83,58	8,82	85,55	42,72	96,91	-2,31	6,50	
MO_3	114	82,17	18,89	88,14	9,62	96,73	-2,78	6,92	
MVA_1	114	0,29	0,07	0,28	0,20	0,63	2,21	6,77	g cm ⁻³
MVA_2	114	0,25	0,07	0,25	0,13	0,45	0,47	0,09	
MVA_3	114	0,23	0,23	0,17	0,10	1,66	4,64	22,26	
Ntot_1	114	1,92	0,22	1,92	1,40	2,42	0,08	-0,60	% (p/p)
Ntot_2	114	1,92	0,24	1,92	1,36	2,44	-0,03	-0,51	
Ntot_3	114	1,85	0,41	1,86	0,22	2,49	-1,63	3,86	
P_inf_0.84mm_1	114	46,88	14,84	50,03	2,48	74,83	-0,73	0,26	% (p/p)
pH_1	114	5,56	0,35	5,56	4,52	6,52	-0,28	0,37	-
pH_2	114	5,54	0,37	5,60	4,56	6,36	-0,68	0,11	
pH_3	114	5,36	0,43	5,44	3,44	6,11	-1,13	2,33	
Porotot_1	114	0,84	0,03	0,84	0,76	0,92	0,19	0,45	cm ³ cm ⁻³
Porotot_2	114	0,85	0,04	0,86	0,77	0,94	-0,18	-0,65	
Porotot_3	114	0,88	0,08	0,90	0,38	0,93	-4,27	19,54	
Pyro_1	114	54,29	23,52	53,81	13,76	111,66	0,26	-0,89	-
Pyro_2	114	49,76	24,18	46,34	9,32	112,87	0,51	-0,61	
Pyro_3	114	33,86	20,83	27,74	4,17	96,78	0,95	0,03	
RE_max	114	0,93	0,27	0,87	0,44	2,14	1,38	3,59	Mpa
RE_max_d	114	51,58	23,04	52,50	10,00	75,00	-0,24	-1,58	cm
RE_max_Ohp	114	0,81	0,24	0,77	0,39	2,04	1,61	5,34	MPa
RE_max_Ohp_d	114	31,49	8,67	35,00	10,00	40,00	-0,84	-0,26	cm
RE_Moy	114	0,59	0,18	0,58	0,27	1,65	2,16	9,32	MPa
REU.10kPa_1	114	0,25	0,06	0,26	0,02	0,39	-0,64	0,92	cm ³ cm ⁻³
REU.10kPa_2	114	0,19	0,06	0,19	0,06	0,33	-0,02	-0,53	
REU.10kPa_3	114	0,17	0,06	0,16	0,01	0,34	0,23	0,13	
Tea.5kPa_1	114	0,22	0,06	0,22	0,07	0,41	0,35	0,57	cm ³ cm ⁻³
Tea.5kPa_2	114	0,19	0,07	0,19	0,04	0,36	0,01	-0,06	
Tea.5kPa_3	114	0,18	0,08	0,18	0,01	0,59	0,99	4,52	
Tee.1kPa_1	114	0,77	0,07	0,78	0,49	0,92	-1,09	2,87	cm ³ cm ⁻³
Tee.1kPa_2	114	0,77	0,04	0,76	0,66	0,92	0,44	0,94	
Tee.1kPa_3	114	0,78	0,09	0,79	0,38	0,92	-2,14	6,58	
Tee.3kPa_1	114	0,68	0,07	0,68	0,43	0,83	-0,92	1,75	cm ³ cm ⁻³
Tee.3kPa_2	114	0,70	0,05	0,70	0,58	0,90	1,06	1,91	
Tee.3kPa_3	114	0,73	0,09	0,73	0,36	0,89	-1,50	4,20	

Annexe 3 – Autres variables significatives à alpha = 0,05 selon la série d'ANOVAs

Variable	Valeur F F(1;112)	pvalue	Groupe 1 n = 12		Groupe 2 n =102	
			Moy	Sd	Moy	Sd
Pyro_1	4,963	0,02790	67,38	14,69	52,75	23,93
CE_1	5,597	0,01970	821,42	292,3	647,21	231,65
pH_1	7,223	0,00830	5,81	0,23	5,53	0,35
Porotot_1	10,35	0,00169	0,82	0,03	0,84	0,03
Tea.5kPa_1	4,876	0,02930	0,19	0,06	0,23	0,06
Pyro_2	3,998	0,04800	61,24	19,20	48,41	24,42
Porotot_2	5,776	0,01790	0,83	0,04	0,86	0,04
Tea.5kPa_2	5,247	0,02390	0,14	0,06	0,19	0,07
ksat_2	5,41	0,02180	0,0093	0,0087	0,0224	0,0209
MVA_2	12,02	0,00075	0,31	0,09	0,24	0,06
Pyro_3	4,402	0,03820	26,30	22,31	34,75	20,58
Ntot_3	6,606	0,01150	1,33	0,82	0,91	0,29
Porotot_3	7,306	0,00794	0,74	0,20	0,89	0,02
Tee.1kPa_3	9,919	0,00210	0,67	0,17	0,79	0,06
Tee.3kPa_3	7,544	0,00702	0,65	0,16	0,74	0,07
RE_Moy	4,733	0,03170	0,50	0,17	0,60	0,18

Note : Contrairement aux résultats présentés au chapitre 2, le respect des postulats de ces ANOVAs n'a pas été évalué comme il s'agit de données internes à titre indicatif seulement. De plus, les écarts-type sont présentés au lieu des erreurs standards.