Quantifying the Spatial Distribution of Soil Organic Carbon and Nitrogen Using Reflectance Spectroscopy

by

Preston Thomas Sorenson

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Soil Science

Department of Renewable Resources
University of Alberta

Abstract

Soil is a critical component of global biogeochemical cycles, and there is an increasing need for cost effective tools to measure soil carbon stocks and determine soil nitrogen contents. Reflectance spectroscopy can deliver large volumes of soil carbon data, with potential applications for understanding soil carbon distribution and assessing reclaimed soils. Reflectance spectra in the SWIR range were collected on a range of soil samples, including intact cores, using a SisuROCK automated hyperspectral imaging system in a laboratory setting. Samples were also analyzed for soil organic carbon and total nitrogen concentrations by dry combustion to prepare a training data set to use as inputs for predictive models. Predictive models were built using continuous wavelet processing along with Cubist and Bayesian Regularized Neural Net models. Overall, soil organic carbon was more aggregated in Chernozemic soils and in B and C horizons compared to A horizons. Nitrogen in turn showed more aggregation for all soil types and horizons compared to soil organic carbon. Additionally, crop rotations were revealed to influence both the concentration and distribution of carbon and nitrogen. Continuous forage rotations were found to have the highest soil organic carbon (SOC) and total nitrogen (TN) contents compared to an agro-ecological rotation for only the top 3 and 4 cm, respectively. These two rotations had comparable concentrations for both parameters for the rest of the topsoil, which was greater than the concentration of SOC and TN in a continuous grain rotation to depths of approximately 12 cm. Increases in both SOC and TN were associated with increased spatial aggregation at fine spatial scales. Reflectance spectroscopy data was also

found to be valuable for reclaimed soil assessments using a Cubist statistical model. The root mean square error (RMSE), R^2, and ratio of performance to deviation (RPD) values for SOC were 0.60%, 0.80, and 2.2, respectively. The TN model results were 0.05%, 0.81 and 2.5, and pH model results were 0.44, 0.69 and 1.8. In addition to a reflectance spectroscopy system, a simple two-band reflectance sensor was also evaluated for use assessing reclaimed soils. This two-band sensor could only be used for general qualitative comparisons amongst soil zones. Specifically, to identify areas with statistically significant differences in organic matter, cation exchange capacity and water content. This system could be used to map out zones with significant soil variation as part of reclamation monitoring, and then used to guide laboratory analytical sampling. Overall, these results indicate that reflectance-based sensing tools can be used to successfully measure soil properties and support the assessment of reclaimed soils.

PREFACE

Chapters 2 and 3 were conceived by Preston Sorenson. All field and laboratory work, data analysis, and manuscript preparation were completed by Preston Sorenson.

Chapters 4 and 5 of this research were conducted in collaboration with Alberta Innovates and Mr. Arnold Janz from Alberta Environment. Mr. Arnold Janz conceived the research and Alberta Innovates completed the field work for the project. Preston Sorenson developed the data analysis approach, conducted all data analysis, and prepared the manuscripts.

Chapters 2, 4 and 5 have been published and Chapter 3 is under review in Geoderma. The citations for each manuscript are as follows:

Chapter 2 was published as: Sorenson, P. T., Quideau, S.A., Rivard, B., 2018. High resolution measurement of soil organic carbon and total nitrogen with laboratory imaging spectroscopy. Geoderma 315, 170–177. https://doi.org/10.1016/j.geoderma.2017.11.032

Chapter 3 has been submitted as: Sorenson, P., Quideau, S.A., Rivard, B., and Dyck, M. Distribution mapping of soil profile carbon and nitrogen with laboratory imaging spectroscopy

Chapter 4 was published as: Sorenson, P.T., Small, C., Tappert, M.C., Quideau, S.A., Drozdowski, B., Underwood, A., Janz, A., 2017. Monitoring organic carbon, total nitrogen, and pH for reclaimed soils using field reflectance spectroscopy. Can. J. Soil Sci. 97, 241–248. https://doi.org/10.1139/cjss-2016-0116

Chapter 5 was published as: Sorenson, P T, Small, C., Quideau, S.A., Underwood, A., Janz, A., 2018. Assessment of reclaimed soils by unsupervised clustering of proximal sensor data. Can. J. Soil Sci. 98, 688–695.

ACKNOWLEDGEMENTS

Funding for this work was provided by a MITACS Accelerate Grant along with funding to Dr. Sylvie Quideau as part of a NSERC Discovery Grant.

Alberta Innovates provided datasets for the reclamation assessment portion of this research.

Thank you to Allan Harms and Conrad David for assistance with this research.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Soil Carbon

Soil carbon is a critical component of the global carbon cycle as soil contains 4.5 times the amount of carbon present in biomass (Jobbágy and Jackson, 2000). Soil carbon is a critical part of global biogeochemical cycles and climate change, because soils can act as a source or sink for carbon. Quantifying soil carbon is essential for understanding how much carbon will be released from areas where soils will become a source for carbon and to understand how much carbon can be stored in soils acting as carbon sinks. There is a need for more cost-effective methods to quantitatively measure and characterize soil carbon. The increasing need for soil data has been referred to as a soil data crisis (McBratney et al., 2006). Reflectance spectroscopy has the potential to solve this crisis as it is a rapid and non-destructive analysis that can be utilized both in the laboratory and in the field. There is extensive literature on the use of reflectance spectroscopy to quantify soil carbon in other regions of the world (Bartholomeus et al., 2008; Ben-Dor and Banin, 1990, 1995; Chang et al., 2001; Ge et al., 2014; Gomez et al., 2008b; McBratney et al., 2006; McCarty et al., 2002; Rossel et al., 2010) and inorganic carbon has been measured using reflectance spectroscopy since 1990 (Ben-Dor and Banin, 1990). While some work has been done analyzing soil carbon with reflectance spectroscopy in Canada (Martin et al., 2002; Xie et al., 2011), the literature is not as extensive and improved results can be obtained using machine learning as compared to the partial least squares regression models used in previous Canadian studies (Doetterl et al., 2013; Rossel et al., 2010).

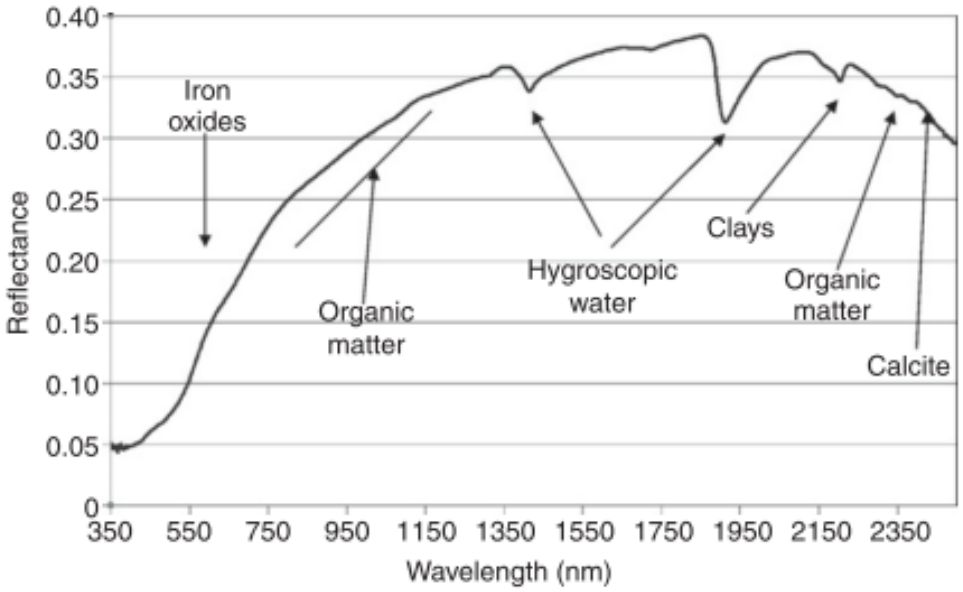
There is a large global demand for more rapid and cost-effective soil data analysis solutions due to industrial oil and gas activities. Alberta alone has approximately 470,000 well sites that will require reclamation (Alberta Energy Regulator, 2016a), more than 39,000 oil and gas facilities that require regular environmental monitoring (Alberta Energy Regulator, 2016b; Government of Alberta, 2009), and has experienced 17,605 pipeline leaks between 1990 and 2012 (Alberta Energy Regulator, 2013). Construction of these facilities is associated with soil disturbance and degradation in soil quality, particularly a loss of soil organic carbon (Hammermeister et al., 2003). This project is focused on developing reflectance spectroscopy methods to quantify and characterize natural soil organic carbon in soil. Measuring soil carbon stocks is critical for understanding the effects of climate change on soil and to quantify the impacts of industrial disturbance on soil. Soil data are needed to effectively manage soils, and reflectance spectroscopy will improve upon current methodologies by providing faster and cheaper data.

1.2. Reflectance Spectroscopy

Soil is a three-phase system that consists of solid, liquid and gaseous components. The solid portions of soil consist of both mineral and organic components in varying concentrations. The main components of interest to soil scientists that are spectrally active between 350 nm and 2500 nm (which includes visible light to shortwave infrared light) are: iron oxides, organic matter, water, clays, and carbonates (Figure 1-1). Fundamentally, reflectance spectroscopy is due to

covalent bonds preferentially absorbing different wavelengths of light. Specifically, many parameters of interest to soil scientists (i.e. Organic Carbon, Nitrogen, Carbonates, Clay Minerals) will have fundamental absorptions in the infrared portion of the electromagnetic spectrum. Overtones of these fundamental absorptions then occur within the near-infrared (750-1000 nm) and the short-wave infrared (1000-2500 nm) portion of the electromagnetic spectrum.

Figure 1-1. Example soil spectrum and illustration of locations where spectral properties associated with different soil properties and structural water can be observed (Ben-Dor et al., 2008).



Increases in soil organic matter are associated with decreases in the reflectance throughout the visible light region due to wide conduction bands associated with organic matter (St. Luce et al., 2014). This has long been quantitatively observed in Soil Science, with the colour of topsoil used to infer soil organic matter content and to classify soil (Soil Classification

Working Group, 1998). Increases in organic matter content is associated with an overall decrease in near infrared (NIR) and shortwave infrared reflectance (SWIR) regions (St. Luce et al., 2014), along with specific absorption features in these regions. These specific absorption features are due to the absorbance of specific wavelengths of light by different covalent bonds.

Organic matter absorption features in the SWIR region are overtones of the fundamental vibrations that occur within the infrared regions (Klavarioti et al., 2014). An overtone is the absorption feature that results when a molecule transitions from the ground state to a second or higher excited state. The fundamental absorption feature, by comparison, results from the transition from the ground state to the first excited state.

There are many organic matter spectral features compared to mineral components, such as clay, because soil organic matter is compositionally diverse. Soil organic matter has a variety of different types of chemical bonds, and as decomposition proceeds the structure of the organic molecules becomes increasingly amorphous (Schaumann, 2006), which shifts the position of different absorption features.

Soil organic matter has a variety of covalent bonds present, and these bonds have different specific absorption features (Table 1-1**Error! Reference source not found.**).

- Aromatic carbon has a distinct absorption feature at 1650 nm, with weaker signatures at 1100 and 825 nm (Rossel et al., 2010).
- Phenolic compounds have a separate absorption feature at 1961 nm due to the C-OH bonds in phenols that are absent in other types of aromatic carbon.

- Alkyl carbon has an asymmetric-symmetric doublet that is centered near 1706 and 1754 nm, associated with alkyl C-H bonds, and four additional weaker features at lower wavelengths are observable.
- Aliphatic compounds have a distinct absorption feature at 2275 nm and a weaker feature at 1706 nm due to the C-H bonds.
- Methyl groups have a wide range of absorption features ranging from 2307-2469 nm and 1730-1852 nm. This variance in methyl group feature locations is likely due to the wide range of methyl group configurations in organic matter. Methyl groups can be found in several different organic molecules and can be attached to different types of elements and carbon bonds. This leads to the wide variance in where methyl spectral features occur.
- The C=O bond leads to absorption features at 1930 nm and 1449 nm in carboxylic acids and at 2033 and 1524 nm in amides. The shift in position of the absorption feature can be attributed to the nitrogen molecule bonded to the carbon in amides that is not present in carboxylic acids.
- Amines have bonds at 2060 nm and 1500 nm due to the N-H bonds present in these molecules, and weaker overtones at 1000 nm and 751 nm.
- Polysaccharides and carbohydrates have distinct absorption features at 2137 nm and 2381 nm, respectively. Both compounds are present in fresh organic matter and have low mean residence times in soil and are largely absent in degraded organic matter (Hoyos-Santillan et al., 2015).

Table 1-1. Visible, Near-Infrared and Short-Wave Infrared (VIS-NIR) wavelength responses for key soil properties along with their associated fundamental infrared absorption locations (Rossel et al., 2010).

Inorganic Constituents			
Soil constituent	**Vis-NIR Wavelength Response (nm)**	**Fundamental (cm^{-1})**	**vis-NIR Mode**
Goethite	434, 480, 650, 920		Electronic Transitions
Haematite	404, 444, 529, 650, 884		Electronic Transitions
Water	1915 1455 1380, 1135 940	ν1 O-H: 3278 ν2 H-O-H: 1645 ν3 O-H: 3848	$\nu_2 + \nu_3$ $2\nu_2 + \nu_3$ $\nu_1 + \nu_3$ $\nu_1 + \nu_2 + \nu_3$ $2\nu_1 + \nu_3$
Hydroxyl	1400 930 700	ν1 O-H: 3575	$2\nu_1$ $3\nu_1$ $4\nu_1$
Kaolin doublet	1395 1415 2160 2208	 ν1a O–H 3695 cm^{-1} ν1b O–H 3620 cm^{-1} δ Al–OH 915 cm^{-1}	$2\nu_{1a}$ $2\nu_{1b}$ ν1a+δ ν1b+δ
Smectite	2206 2230	ν1 O–H 3620 cm^{-1} δ_a Al–OH 915 cm^{-1} δ_b AlFe–OH 885 cm^{-1}	$\nu_1+\delta_a$ $\nu_1+\delta_b$
Illite	2206, 2340 2450	ν1 O–H 3620 cm^{-1}	$\nu_1+\delta$ Poorly defined
Carbonate	2336	ν3 CO–3 2 1415 cm^{-1}	$3\nu_3$

Organic Constituents			
Soil constituent	**Vis-NIR Wavelength Response (nm)**	**Fundamental (cm^{-1})**	**vis-NIR Mode**
Aromatics	1650 1100 825	 ν_1 C–H 3030 cm^{-1}	$2\nu_1$ $3\nu_1$ $4\nu_1$
Amine	2060 1500 1000 751	δN–H 1610 cm^{-1} ν_1 N–H 3330 cm^{-1}	$\nu_1+\delta$ $2\nu_1$ $3\nu_1$ $4\nu_1$
Alkyl asymmetric-symmetric doublet	1706 1754 1138 1170 853 877	ν_3 C–H 2930 cm^{-1} ν_1 C–H 2850 cm^{-1}	$2\nu_3$ $2\nu_1$ $3\nu_3$ $3\nu_1$ $4\nu_3$ $4\nu_1$
Carboxylic acids	1930 1449	ν_1 C=O 1725 cm^{-1}	$3\nu_1$ $4\nu_1$
Amides	2033 1524	ν_1 C=O 1640 cm^{-1}	$3\nu_1$ $4\nu_1$
Aliphatics	2275 1706	ν_1 C–H 1465 cm^{-1}	$3\nu_1$ $4\nu_1$
Methyls	2307-2469 1730-1852	ν_1 C–H 1445–1350 cm^{-1}	$3\nu_1$ $4\nu_1$
Phenolics	1961	ν_1 C–OH 1275 cm^{-1}	$4\nu_1$
Polysaccharides	2137	ν_1 C–O 1170 cm^{-1}	$4\nu_1$
Carbohydrates	2381	ν_1 C–O 1050 cm^{-1}	$4\nu_1$

Reflectance spectroscopy has primarily focused on point spectroscopy systems. Most research regarding the use of reflectance spectroscopy to analyze soil has taken in place in Israel, Europe, Australia and the United States (i.e. Ben-Dor and Banin, 1995; Chang and Laird, 2002; Viscarra Rossel and Lark, 2009; Doetterl et al., 2013). Work regarding soil analysis with

reflectance spectroscopy in Canada has been fairly limited, with studies in Manitoba, Ontario and Quebec (Martin et al., 2002; St. Luce et al., 2014; Xie et al., 2011). The Canadian studies have used partial least squared regression, rather then newer machine learning methods such as CUBIST models (Kuhn et al., 2014). Machine learning models have outperformed partial least squared regression in other regions, and evaluating their performance on Canadian soils, particularly the Canadian Prairies, has not been done. Additionally, Canada has relatively high carbon soils compared to the regions where most of this research has taken place. Evaluating the performance of these models on high carbon Canadian soils, and testing if the signal saturates under high carbon soil conditions requires investigation.

Spectral pre-processing of near infrared reflectance spectroscopy (NIRS) spectra for soil has typically focused on methods such as standard normal variates, multiplicative scatter corrections, and Savitksy-Golay smoothing. Wavelet transform signal processing has been used in other fields, such as geology, and has been shown to improve reflectance spectroscopy results (Rivard et al., 2008). The use of continuous wavelet signal processing has been very limited in soil spectroscopy research, with a couple examples of discrete wavelets being used (Rossel et al., 2010; Viscarra Rossel and Lark, 2009). Wavelet transforms can be continuous or discrete. The advantage of continuous wavelet transforms is that they are directly comparable to the original spectrum, which facilitates manual feature selection prior to machine learning. The use of wavelets transforms on soil spectra generally has been very limited, particularly continuous wavelets, and part of the goal of this research is to investigate their use on spectra from Canadian soils with machine learning models.

Imaging spectroscopy applications for soil have focused on airborne and spaceborne platforms (Chabrillat et al., 2002; Gomez et al., 2008b) rather than at the soil sample or soil profile scale. The extensive research regarding soil chemometrics has focused on point spectroscopy. The application of imaging spectroscopy at the soil profile scale has not been extensively investigated. A study was conducted in Germany (Steffens and Buddenbaum, 2013), with a spectral range of 400 to 990 nm and no wavelet processing of the data. The short wave infrared range was not investigated, where most organic matter features are located (Rossel et al., 2010). There are no examples of imaging spectroscopy applications at the soil sample or profile scale in Canada.

The objectives of this research were therefore to investigate the following:

1. To determine if shortwave infrared imaging spectroscopy could be used to measure soil organic carbon and total nitrogen on intact and unground samples in the laboratory for a variety of Canadian soil samples. Further, to use the imaging spectroscopy results to characterize the spatial distribution of SOC and TN at the soil aggregate scale, and determine if the distribution of SOC and TN varies based on soil type and horizon at fine spatial scales.
2. To identify at what depth changes could be detected in SOC and TN using imaging spectroscopy. Specifically, following the use of different crop rotations, and if these changes in SOC and TN were associated with a change in the spatial distribution of these parameters at the same fine spatial scales.

3. To determine if reflectance spectroscopy data collected in the field with a drill-rig mounted spectroscopy system could be used to measure SOC, TN and pH as part of reclamation assessments.
4. To investigate if a simple two-band reflectance sensor could be used to generate quantitative SOC results or support reclamation assessments by successfully identifying different zones of soil organic matter content in reclaimed soil.

1.3. References

Alberta Energy Regulator, 2016a. ST37: List of Wells in Alberta Monthly Updates [WWW Document]. URL https://www.aer.ca/data-and-publications/statistical-reports/st37 (accessed 4.26.16).

Alberta Energy Regulator, 2016b. Statistical Reports [WWW Document]. URL https://www.aer.ca/data-and-publications/statistical-reports (accessed 5.18.16).

Alberta Energy Regulator, 2013. Report 2013-B: Pipeline Performance in Alberta, 1990-2012. Calgary, Alberta.

Bartholomeus, H.M., Schaepman, M.E., Kooistra, L., Stevens, A., Hoogmoed, W.B., Spaargaren, O.S.P., 2008. Spectral reflectance based indices for soil organic carbon quantification. Geoderma. https://doi.org/10.1016/j.geoderma.2008.01.010

Ben-Dor, E., Banin, A., 1995. Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj1995.03615995005900020014x

Ben-Dor, E., Banin, A., 1990. Near-infrared reflectance analysis of carbonate concentration in soils. Appl. Spectrosc. https://doi.org/10.1366/0003702904086821

Ben-Dor, E., Taylor, R.G., Hill, J., Demattê, J.A.M., Whiting, M.L., Chabrillat, S., Sommer, S., 2008. Imaging Spectrometry for Soil Applications. Adv. Agron. https://doi.org/10.1016/S0065-2113(07)00008-9

Chabrillat, S., Goetz, A.F.H., Krosley, L., Olsen, H.W., 2002. Use of hyperspectral images in the identification and mapping of expansive clay soils and the role of spatial resolution. Remote Sens. Environ. https://doi.org/10.1016/S0034-4257(02)00060-3

Chakraborty, S., Weindorf, D.C., Morgan, C.L.S., Ge, Y., Galbraith, J.M., Li, B., Kahlon, C.S., 2010. Rapid Identification of Oil-Contaminated Soils Using Visible Near-Infrared Diffuse Reflectance Spectroscopy. J. Environ. Qual. https://doi.org/10.2134/jeq2010.0183

Chang, C.-W., Laird, D.A., 2002. Near-Infrared Reflectance Spectroscopic Analysis of Soil C and N. Soil Sci. 167, 110–116. https://doi.org/10.1097/00010694-200202000-00003

Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2001.652480x

Doetterl, S., Stevens, A., Van Oost, K., van Wesemael, B., 2013. Soil Organic Carbon Assessment at High Vertical Resolution using Closed-Tube Sampling and Vis-NIR Spectroscopy. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2012.0410n

Forrester, S.T., Janik, L.J., McLaughlin, M.J., Soriano-Disla, J.M., Stewart, R., Dearman, B., 2013. Total Petroleum Hydrocarbon Concentration Prediction in Soils Using Diffuse Reflectance Infrared Spectroscopy. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2012.0201

Ge, Y., Morgan, C.L.S., Ackerson, J.P., 2014. VisNIR spectra of dried ground soils predict properties of soils scanned moist and intact. Geoderma. https://doi.org/10.1016/j.geoderma.2014.01.011

Gomez, C., Viscarra Rossel, R.A., McBratney, A.B., 2008. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. Geoderma. https://doi.org/10.1016/j.geoderma.2008.06.011

Government of Alberta, 2009. Soil Monitoring Directive. Edmonton, Alberta.

Hoyos-Santillan, J., Lomax, B.H., Large, D., Turner, B.L., Boom, A., Lopez, O.R., Sjögersten, S., 2015. Getting to the root of the problem: litter decomposition and peat formation in lowland Neotropical peatlands. Biogeochemistry. https://doi.org/10.1007/s10533-015-0147-7

Jobbágy, E.G., Jackson, R.B., 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecol. Appl. https://doi.org/10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2

Klavarioti, M., Kostarelos, K., Pourjabbar, A., Ghandehari, M., 2014. In situ sensing of subsurface contamination-part I: Near-infrared spectral characterization of alkanes, aromatics, and chlorinated hydrocarbons. Environ. Sci. Pollut. Res. https://doi.org/10.1007/s11356-013-2478-z

Kuhn, M., Weston, S., Keefer, C., Coulter, N., 2014. Cubist: Rule- and Instance-Based Regression Modeling.

Laakso, K., Rivard, B., Rogge, D., 2016. Enhanced detection of gossans using hyperspectral data: Example from the Cape Smith Belt of northern Quebec, Canada. ISPRS J. Photogramm. Remote Sens. https://doi.org/10.1016/j.isprsjprs.2016.02.004

Martin, P.D., Malley, D.F., Manning, G., Fuller, L., 2002. Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy. Can. J. Soil Sci. https://doi.org/10.4141/S01-054

McBratney, A.B., Minasny, B., Viscarra Rossel, R., 2006. Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. Geoderma. https://doi.org/10.1016/j.geoderma.2006.03.051

McCarty, G.W., Reeves, J.B., Reeves, V.B., Follett, R.F., Kimble, J.M., 2002. Mid-Infrared and Near-Infrared Diffuse Reflectance Spectroscopy for Soil Carbon Measurement. Soil Sci. Soc. Am. J. 66, 640. https://doi.org/10.2136/sssaj2002.0640

Okparanma, R.N., Coulon, F., Mouazen, A.M., 2014. Analysis of petroleum-contaminated soils by diffuse reflectance spectroscopy and sequential ultrasonic solvent extraction-gas chromatography. Env. Pollut 184, 298–305. https://doi.org/10.1016/j.envpol.2013.08.039

Rivard, B., Feng, J., Gallie, A., Sanchez-Azofeifa, A., 2008. Continuous wavelets for the improved use of spectral libraries and hyperspectral data. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2008.01.016

Rossel, R.A.A.V., Behrens, T., Viscarra Rossel, R.A., Behrens, T., 2010. Using data mining to model and interpret soil diffuse reflectance spectra. Geoderma 158, 46–54. https://doi.org/10.1016/j.geoderma.2009.12.025

Scafutto, R.D.P.M., de Souza Filho, C.R., Rivard, B., 2016. Characterization of mineral substrates impregnated with crude oils using proximal infrared hyperspectral imaging. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2016.03.033

Schaumann, G.E., 2006. Soil organic matter beyond molecular structure Part II: Amorphous nature and physical aging. J. Plant Nutr. Soil Sci. https://doi.org/10.1002/jpln.200521791

Soil Classification Working Group, 1998. The Canadian System of Soil Classification. Can. Syst. Soil Classif. 3rd ed. Agric. Agri-Food Canada Publ. 1646 187.

St. Luce, M., Ziadi, N., Zebarth, B.J., Grant, C.A., Tremblay, G.F., Gregorich, E.G., 2014. Rapid determination of soil organic matter quality indicators using visible near infrared reflectance spectroscopy. Geoderma. https://doi.org/10.1016/j.geoderma.2014.05.023

Steffens, M., Buddenbaum, H., 2013. Laboratory imaging spectroscopy of a stagnic Luvisol profile - High resolution soil characterisation, classification and mapping of elemental concentrations. Geoderma. https://doi.org/10.1016/j.geoderma.2012.11.011

Viscarra Rossel, R.A., Lark, R.M., 2009. Improved analysis and modelling of soil diffuse reflectance spectra using wavelets. Eur. J. Soil Sci. https://doi.org/10.1111/j.1365-2389.2009.01121.x

Xie, H.T., Yang, X.M., Drury, C.F., Yang, J.Y., Zhang, X.D., 2011. Predicting soil organic carbon and total nitrogen using mid- and near-infrared spectra for Brookston clay loam soil in Southwestern Ontario, Canada. Can. J. Soil Sci. https://doi.org/10.4141/cjss10029

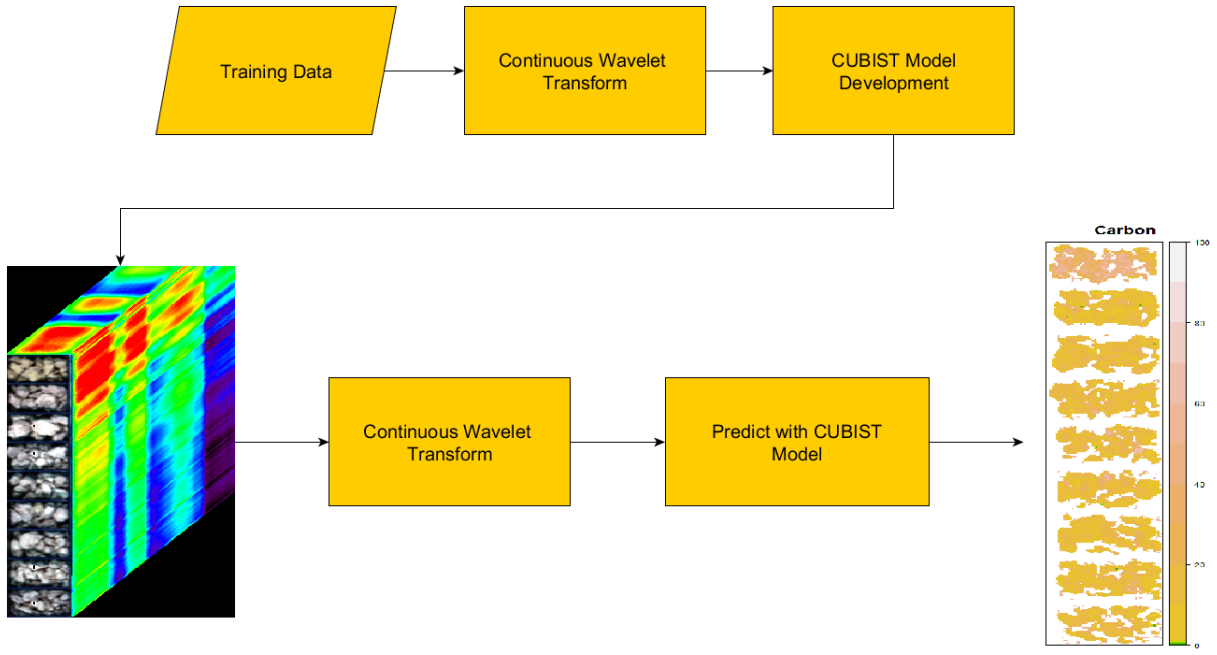
2. High Resolution Measurement of Soil Organic Carbon and Total Nitrogen with Laboratory Imaging Spectroscopy

2.1. Abstract

Soil is a critical component of global biogeochemical cycles, and there is an increasing need for cost effective tools to measure soil carbon stocks and determine soil nitrogen contents. Reflectance spectroscopy can deliver large volumes of soil carbon data. However, as soil carbon concentrations can be spatially heterogeneous, imaging spectroscopy presents the best potential to provide high resolution measurements and accurately characterize soil carbon heterogeneity. For this study, discrete, intact and unground soil samples were collected and analyzed using a SisuROCK automated hyperspectral imaging system in a laboratory setting, focused on the shortwave infrared portion of the electromagnetic spectrum. Samples were also analyzed for soil organic carbon and total nitrogen concentrations by dry combustion to prepare a training data set. Predictive models were built using continuous wavelet processing along with partial least squares regression and CUBIST models. Spatial variation of carbon and nitrogen was determined using Moran's i and comparisons of spatial variations among soil types and horizons were made using a spatial generalized least squares model. Overall, soil organic carbon was more aggregated in Chernozemic soils and in B and C horizons compared to A horizons. Nitrogen in turn showed more aggregation for all soil types and horizons compared to soil organic carbon. Results indicated that imaging spectroscopy can be successfully used to

measure and characterize the spatial variability of soil carbon and nitrogen at the soil aggregate scale.

Graphical Abstract



Highlights

- Continuous wavelets with CUBIST models produce accurate soil predictive models.
- Imaging spectroscopy can provide high spatial resolution soil C and N measurements
- Imaging spectroscopy can quantify fine scale spatial clustering in soil C and N
- B and C horizon carbon showed more spatial aggregation than A horizon carbon
- Soil N was more spatially aggregated than soil C

Keywords

Imaging spectroscopy; carbon; nitrogen; spatial statistics

2.2. Introduction

Soil organic carbon (SOC) is a critical component of the global carbon cycle as soil contains 4.5 times the amount of carbon present in aboveground biomass (Jobbágy and Jackson, 2000), and soils can act as a source or sink for carbon. Quantifying SOC is essential for understanding how much carbon will be released from areas where soils will become a source for carbon and to understand how much carbon can be stored in soils acting as carbon sinks. In addition to SOC, nitrogen plays an essential role in biogeochemical cycles with nitrogen limitation widespread in terrestrial ecosystems (Vitousek and Howarth, 1991). There is a need for more rapid and cost-effective methods to quantitatively measure and characterize SOC and total nitrogen (TN). The increasing need for soil data has been referred to as a soil data crisis (McBratney et al., 2006).

Reflectance spectroscopy has the potential to solve this crisis, as it is a rapid and non-destructive analysis that can be utilized both in the laboratory and in the field. There is an extensive literature on the use of point measurement reflectance spectroscopy in many regions of the world to quantify SOC (Bartholomeus et al., 2008; Ben-Dor and Banin, 1990, 1995; Chang et al., 2001; Ge et al., 2014; Gomez et al., 2008b; McBratney et al., 2006; McCarty et al., 2002; Rossel et al., 2010) and total nitrogen (TN) (Chang et al., 2001; Chang and Laird, 2002; Morellos et al., 2016; St. Luce et al., 2014). While some work has been done analyzing SOC and TN with point reflectance spectroscopy in Canada (Martin et al., 2002; Xie et al., 2011), the literature is not as extensive.

Previous research has established that machine learning models can outperform partial least squares regression when used to build predictive models to measure soil properties from reflectance spectra (Doetterl et al., 2013; Nawar et al., 2016; Rossel and Behrens, 2010; Sorenson et al., 2017). Additionally, an alternative signal processing method, continuous wavelet transforms, has improved results derived from the analysis of reflectance spectra compared to conventional signal processing techniques such as Savitsky-Golay smoothing, derivatives and multiplicative scatter correction (Rossel et al., 2010; Viscarra Rossel and Lark, 2009). Combining wavelet analysis with machine learning models has been successful in producing SOC and TN predictive models for point spectroscopy systems with Canadian soils (Sorenson et al., 2017).

Imaging spectroscopy has the advantage over point measurements of providing high resolution continuous spatial measurements. Imaging spectroscopy studies in soil science have tended to focus on airborne or space borne applications (Gomez et al., 2008b, 2012; Melendez-Pastor et al., 2010; Ouerghemmi et al., 2011) as compared to measurements at the soil sample or profile scale where higher spatial resolutions can be achieved (e.g.; less than 1 mm per pixel). These studies have mapped different soil properties using imaging spectroscopy including: clay, sand, silt, SOC, and inorganic carbon content. Imaging spectroscopy can also be utilized at the laboratory scale to measure variation within samples, as well as to obtain high vertical resolution data. Lastly, laboratory imaging spectroscopy has been used to characterize the spatial variability of SOC and TN for a soil profile in Germany (Steffens and Buddenbaum,

2013). There are no studies using imaging spectroscopy systems to measure SOC and TN in Canada, specifically using wavelet analysis to analyze intact and unground samples.

Soil organic matter varies spatially both horizontally and vertically, and conventional analytical methods have limited ability to measure SOC and TN in fine spatial resolution. Reflectance spectroscopy has been shown to be a valuable tool to measure SOC in high vertical spatial resolution (Doetterl et al., 2013). High vertical resolution measurements can illustrate how SOC changes with depth depending on, for instance, the underlying parent material, and how these changes may be affected by distinct parent materials. Approximately 50% of SOC is contained below 20 cm in boreal forest soils and 59% is below 20 cm in croplands (Jobbágy and Jackson, 2000). Subsoil organic matter corresponds to a substantial proportion of the global soil carbon pool, and its characterization is essential for accurate soil carbon budgets. Other work has found that approximately 63% of SOC and 64% of TN are contained below 30 cm (Wang et al., 2017). For these reasons, tools that can accurately and cost effectively measure SOC at depth are very valuable for developing soil carbon budgets.

As imaging spectroscopy contains spatial as well as spectral information, the spatial variability of several distinct soil attributes can be analyzed at a variety of spatial scales. Imaging spectroscopy has not been used to characterize the soil aggregate scale spatial structure of SOC and TN in Canadian soils, and to determine how spatial structure in SOC and TN may vary between soil types and horizons. Canadian soils tend to be distinct from other regions that have been a focus of reflectance spectroscopy research in that they have relatively high organic

carbon contents, can be frequently water saturated and are relatively young due to recent glaciation. Based on the success of point spectra to measure SOC and TN, and the success of imaging spectroscopy to measure soil and geological properties, the first objective of this study was to determine if shortwave infrared imaging spectroscopy could be used to measure SOC and TN on intact and unground samples in the laboratory for a variety of Canadian soil samples. The second objective was to use the imaging spectroscopy results to characterize the spatial distribution of SOC and TN at the soil aggregate scale, and determine if the distribution of SOC and TN varies based on soil type and horizon at fine spatial scales.

2.3. Materials and Methods

Sample collection and preparation

A total of 370 soil samples was collected in Alberta and Saskatchewan, Canada in August and September 2015 and May 2016 (Table 2-1). To capture a range of soil samples common across the Canadian Prairies, samples were collected from the main soil orders that occur in this region: Chernozems, Gleysols and Luvisols (Table 2-1). In total, 121 Chernozemic samples (58 A horizon samples, 48 B horizon samples and 17 C horizon samples), 137 Gleysolic samples (47 A horizons samples, 36 B horizon samples and 54 C horizon sample) and 96 Luvisolic samples (36 A horizon samples, 33 B horizon samples, and 27 C horizon samples) were collected. Additionally, 16 Brunisolic samples (5 A horizon samples, 5 B horizon samples, and 6 C horizon samples) were collected as they were encountered opportunistically during soil sampling. Less

Brunisolic samples were collected because the focus of the study was to collect samples from Luvisolic, Chernozemic and Gleysolic soils. The Brunisols were encountered during sampling for the Luvisolic samples where glaciofluvial inclusions were present within a predominantly glacial till landscape.

Samples were collected from catenas with forested vegetation or from arable cropland. Catenas with forested vegetation consisted of Luvisols or Brunisols at upper slope positions depending on soil texture, and Gleysols in lower slope positions. Catenas from arable croplands consisted of either Luvisols or Chernozems, depending on latitude, with Gleysols in lower slope positions. Samples were collected to 1.0 m at each sampling location in 2015. A soil pit was dug to 30 cm and one sample was collected from each horizon to 30 cm. A Dutch auger was then used to collect samples from 30 cm to 1.0 m in 10 cm increments. An additional 120 A and 80 B horizon samples were collected in May 2016 to obtain more data with SOC concentrations above one percent. These samples were collected from soil pits dug to 30 cm. All samples were air dried prior to oven drying. Samples were oven dried at 105°C until reaching a stable weight to ensure that moisture had been removed. Samples were then placed as intact aggregates and unground into polypropylene chip trays for spectral analysis.

Spectral Measurements

Spectral data were collected using a SisuROCK automated hyperspectral imaging system developed by Spectral Imaging (Specim) Ltd., Finland. The SisuROCK collects data with two

high-resolution spectral cameras. The first spectral camera collects reflectance data in the visible near-infrared light range (VNIR; 400-1000 nm). The second camera collects reflectance data in the shortwave infrared range (SWIR; 1000-2500 nm). The SWIR data were used for this study as SOC and TN are primarily spectrally active in this range (Rossel et al., 2010), and automated feature selection using all bands only selected bands from the SWIR region. Data in the SWIR range were collected in 256 spectral bands with a spectral resolution of 10 nm and a spatial resolution of 0.2 mm. Quartz halogen lamps provided the SWIR illumination to each sample and the measured light spectrum for each pixel was converted to reflectance via normalization to the average light spectrum from a Spectralon® panel.

Laboratory Analyses

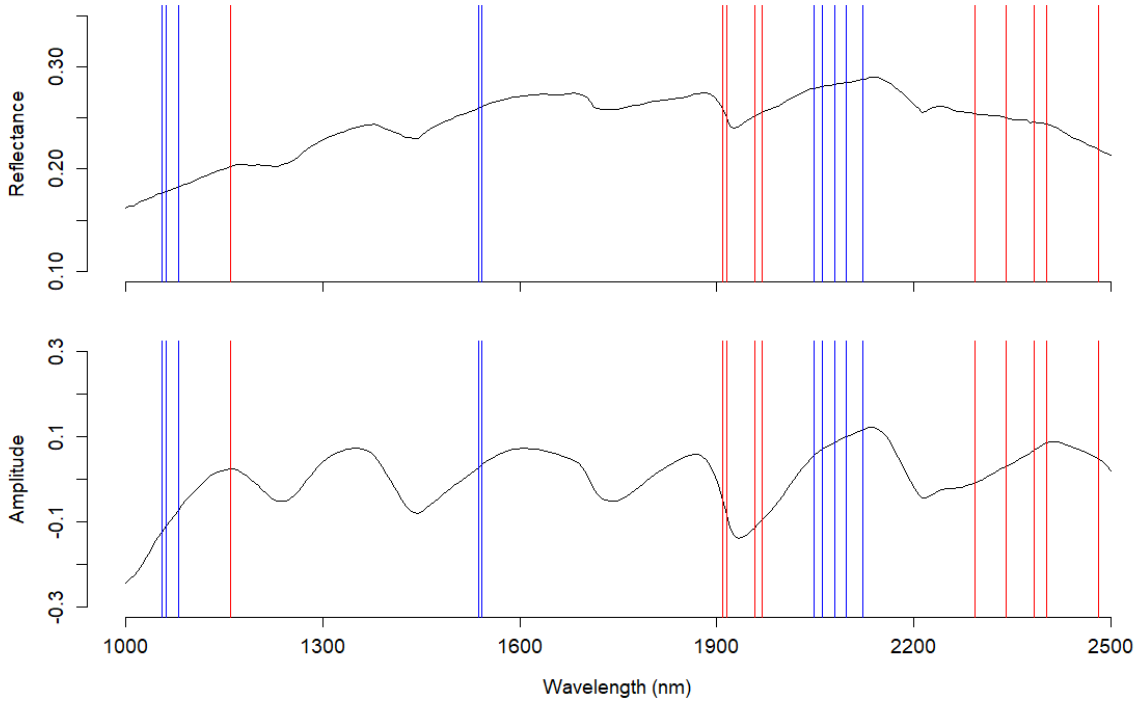
Prior to analysis for SOC and TN, samples were ground with a Retsch MM200 ball mill grinder. Samples were then analyzed for SOC and TN by dry combustion on a Costech ECS 4010 Elemental Analyzer (EA) equipped with a thermocouple detector (Costech Analytical Technologies Inc., Valencia, USA). As none of the A and B horizon sampless would contain carbonates, samples were not pretreated for carbonate removal. Samples were tested with dilute hydrochloric acid to confirm carbonates were not present in the samples. In total 201 samples were analyzed for SOC and TN. Only 197 samples were used for the TN model development as four samples had TN values below detection limits.

Development of carbon and nitrogen predictive models

Processing of spectral data

To generate predictive models, soil laboratory analyses of SOC and TN were compared to the spectral signature of 201 samples using all available soil particle pixels in the ENVI Software Platform (Harris Geospatial Solutions, Melbourne, USA). The shortwave infrared spectrum of each sample was calculated by taking the average of each soil particle pixel for a given sample. The average spectra were then processed using continuous wavelet transforms (CWT) (e.g.; (Rivard et al., 2008; Scafutto et al., 2016; Tappert et al., 2015) using the wmtsa package in R (Percival et al., 2016), to reduce the influence of non-compositional effects such as varying particle size on the analysis. The CWT outputs were calculated using an eight scale second order Gaussian transform. Scales 2, 3 and 4 were then summed used for the model development (Figure 2-1).

Figure 2-1. Example shortwave infrared spectral signature and the sum of wavelet scales 2, 3, and 4 from a second order Gaussian Wavelet Transform. The sample is a Chernozemic Ah horizon containing 35.8 g kg-1 of soil organic carbon and 3.1 g kg-1 of total nitrogen. The upper plot is the unprocessed spectrum and the lower plot is the plot of the wavelet coefficients. The red lines indicate the spectral bands used for prediction of soil organic carbon. The blue lines indicate the spectral bands used for prediction of total nitrogen.



Model Development

Predictive models were developed by training the spectral data with SOC and TN data sets. Models were developed using the *caret* package (Kuhn et al., 2016) in R (R Core Team, 2018). A Cubist model (Kuhn et al., 2014) was used to develop SOC and TN predictive models as previous work has identified that the Cubist model can successfully be used with spectral data (Doetterl et al., 2013; Minasny and McBratney, 2008; Rossel et al., 2010). Partial least squares

regression (PLSR) models (Mevik et al., 2015) were also run for comparison, as PLSR models have been the standard modelling tool for developing chemometric models with soil spectral data (e.g. Chang et al., 2001; Gomez et al., 2008). Details on both Cubist and PLSR models are available in Kuhn and Johnson (2013). Spectra were spectrally subset prior to development of the predictive models. Initially, the spectra were subset to spectral regions known to have SOC related spectral features for the SOC model and regions known to have organic nitrogen related features for the TN model (Rossel et al., 2010). Features were then further sub selected using the feature selection routine in the multivariate adaptive regression splines (MARS) model as per the methodology in Rossel and Behrens (2010). Only features selected by the MARS model were included in the CUBIST and PLSR model development using the *caret* package. The spectral bands used for model development are provided in Table 2-2**Error! Reference source not found.**. All models were run with the default values defined by the *caret package*, which were then automatically optimized using leave-one-out cross-validations. The minimal root mean square error (RMSE) was used as the optimization criterion. The Cubist model was optimized in terms of the number of committees and neighbors in the model, and the PLSR model was optimized in terms of the number of latent variables used.

To evaluate the performance of the models on an independent dataset, a second analysis was performed where 24 samples were withheld from the training dataset. Samples from one field containing Chernozemic soils and another field containing Luvisolic soils were removed from the training dataset and used as an independent test data set. This approach was to ensure

that the training data came from a spatially distinct area compared to the validation data. Predictive models for carbon and nitrogen were built using the Cubist model in R (Kuhn et al., 2014). The spectral bands included in the model were those identified during the cross-validation step and the model was built with 10 committees and 5 neighbors, as this was the optimal number identified during cross-validation model tuning.

Model Evaluation

Model accuracy was evaluated with the RMSE, R^2 and the ratio of performance to deviation (RPD), which is the ratio of the standard deviation to the RMSE. Generally, models with RPD values greater than 2 indicate it can be used accurately for prediction. Models with RPD values between 1.4 and 2 are satisfactory but could use improvement and models with values less than 1.4 have no predictive capability (Chang et al., 2001).

Carbon and nitrogen image maps

Image Processing

Three processing steps took place prior to predicting SOC and TN for each pixel of the sample images. The hyperspectral images were first processed with a 3 x 3 median filter for noise reduction. Following smoothing, a sequential maximum angle convex cone (SMACC) endmember analysis was conducted on each image to select image endmembers in ENVI (Harris Geospatial, Melbourne, USA). Spectral angle mapper analysis was then conducted to classify each image to the endmembers selected by the SMACC process. The resulting classification image was used to create a mask to remove the tray and heavily shaded areas

from the image prior to further processing. The masked image was then imported into R (R Core Team, 2018), where each pixel was processed using an eight scale second order Gaussian transform, and scales 2, 3 and 4 were then summed. Lastly the resulting wavelets coefficients were analyzed using the CUBIST model generated during the statistical model development to create the SOC and TN concentration images.

Spatial Analysis

To investigate the within sample spatial distribution of SOC and TN, Moran's i values were calculated for each sample using the raster package in R (Hijmans, 2016). Following the generation of the predicted SOC and TN images, Moran's i values were calculated for each soil sample for both SOC and TN. Moran's i provides a measure of spatial autocorrelation, with a value of 1 indicating perfect aggregation, values of -1 indicating a regular spatial distribution and a value of 0 indicating a random spatial distribution.

Following calculation of within sample Moran's i values for SOC and TN, a spatial generalized least squares (GLS) model was run to determine if there was a difference between horizons or soil orders regarding the distribution of SOC and TN using the nlme package in R (Pinheiro et al., 2016) . Data were analyzed at the order level because of insufficient replication at the soil subgroup level. Additionally, soil horizon and order interactions were not assessed as the GLS equations were not solvable since the GLS fit was singular due to insufficient degrees of freedom for the interaction tests. The underlying concept behind a spatial GLS model is that

residual spatial patterns can be used as a surrogate for unmeasured variables (McIntire and Fajardo, 2009). Results from the spatial GLS model further indicate if a relationship between factors is present after removing unwanted variables by including location as a correlation structure in the GLS model (Sorenson et al., 2017). A significant intercept indicates that unaccounted for variables are affecting dependent variables. The parameter estimate indicates the direction and magnitude of the relationship.

2.4. Results and Discussion

Predictive models

SOC concentrations in the samples used in this study ranged from 0.9 to 75.9 g kg^{-1} and total TN concentrations varied from 0.1 to 7.59 g kg^{-1} (Table 2-1**).** The Cubist model produced a more accurate predictive model compared to PLSR for SOC (Table 2-2) with the Cubist model producing a predictive model with an R^2 of 0.93, an RMSE of 4.92 g kg^{-1}, and an RPD of 3.79 (Table 2-2, Figure 2-2). By comparison, the PLSR model had an R^2 of 0.74, an RMSE of 9.51, and an RPD of 1.96 (Table 2-2). The Cubist model also produced a more accurate predictive model for TN with an R^2 of 0.90, and RMSE of 0.53 g kg^{-1} and an RPD of 3.20 (Table 2-2, Figure 2-2). The PLSR model had an R^2 of 0.78, an RMSE of 0.80 g kg^{-1} and an RPD of 2.13 (Table 2-2). Regarding its performance on the independent validation data, the Cubist model produced a more accurate predictive model for TN compared to SOC. The SOC predictive model had an R^2

of 0.85, an RMSE of 6.34 g kg^{-1} and an RPD value of 2.64 (Figure 2-3). The TN predictive model had an R^2 of 0.89, an RMSE of 0.51 g kg^{-1} and an RPD value of 3.11.

Table 2-1. Soil taxonomic classification and properties for the samples used to build predictive models.

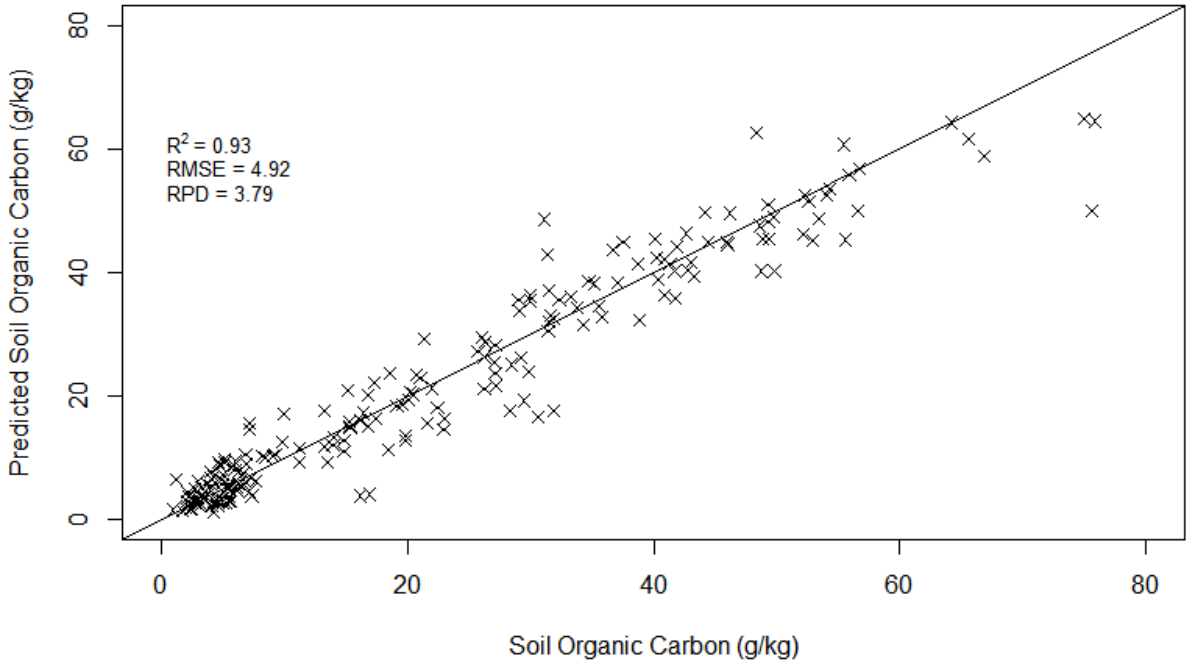
Soil Classification (Canadian) (Soil Classification Working Group, 1998)	**Soil Classification (WRB)** (IUSS Working Group WRB, 2014)	**Horizon**	**Number of Samples**	**Soil Organic Carbon (g kg^{-1})**	**Total Nitrogen (g kg^{-1})**
Orthic Gray Luvisol	Albic Luvisol	A	5	4.8 – 16.4	0.6 – 1.3
		B	10	2.4 – 7.6	0.4 – 0.9
		C	10	2.3 – 6.2	0.2 – 0.6
Dark Gray Luvisol	Albic Luvisol	A	28	2.7 – 26.3	0.4 – 2.3
Gleyed Dark Gray Luvisol	Gleyed Luvisol	A	3	4.4 – 20.1	0.5 – 1.7
		B	3	3.2 – 9.0	0.4 – 0.6
Humic Luvic Gleysol	Planosol	A	4	4.5 – 43.3	1.7 – 3.4
		B	5	4.4 – 29.8	0.7 – 2.2
		C	1	22.9	
Orthic Luvic Gleysol	Planosol	A	2	2.9 – 5.0	0.5 – 0.6
		B	5	1.9 – 6.5	0.4 – 0.9
		C	5	2.1 – 6.2	0.3 – 0.9
Eutric Brunisol	Eutric Cambisol	A	5	5.8 – 37.1	0.5 – 2.0
		B	2	2.9 – 3.7	0.3 – 0.4
		C	4	0.9 – 2.1	0.1 – 0.2
Gleyed Black Chernozem	Gleyic Chernozem	A	1	35.2	3.1
		B	3	8.2 – 19.8	0.5 – 0.6
Orthic Black Chernozem	Chernozem	A	51	11.2 – 64.2	1.0 – 6.3
		B	4	11.2 – 31.9	0.8 – 1.2
Rego Black Chernozem	Chernozem	A	3	29 – 52.6	2.7 – 5.0
Orthic Gleysol	Eutric Gleysol	B	1	7.1	0.5
		C	5	3.5 – 8.4	0.3 – 0.5
Orthic Humic Gleysol	Mollic Gleysol	A	29	9.2 – 75.9	0.8 – 7.6
		B	1	0.61	0.6
		C	2	2.5 – 3.2	0.4
Rego Gleysol	Gleysol	C	3	6.8 – 20.2	0.6 – 2.0
Rego Humic Gleysol	Mollic Gleysol	A	6	55.6 – 75.6	5.2 – 7.4

Table 2-2. Cross-validation results for all soil samples soil organic carbon (SOC) and total nitrogen (TN) using reflectance spectroscopy data. Model results are based on the root mean square error (RMSE), R, and Ratio of Performance to Deviation (RPD) obtained during a leave-one-out cross-validation. The spectral bands used for model development are the band center wavelengths in nm.

Model	**Cross-Validation Results**		
	$RMSE_{cv}$	**R^2_{cv}**	**RPD**
Carbon - Partial Least Squares Regression	9.51 g kg^{-1}	0.74	1.96
Carbon - Cubist	4.92 g kg^{-1}	0.93	3.79
Nitrogen - Partial Least Squares Regression	0.80 g kg^{-1}	0.78	2.13
Nitrogen - Cubist	0.53 g kg^{-1}	0.90	3.20
SOC Bands (nm)	1093, 2312, 2374, 1874, 2355, 1917, 1930, 2493, 2262, 1867		
TN Bands (nm)	2043, 1010, 985, 991, 1484, 2024, 2011, 2062, 2087, 1490		

Figure 2-2. Cross-validated predicted versus observed (a) soil organic carbon (SOC) content (g kg^{-1}), (b) total nitrogen as produced by the Cubist model. The model was developed using 201 samples which were analyzed for SOC, and 196 samples analyzed for TN. The solid line indicates the 1:1 line to illustrate deviations between predicted and measured data.

(a)



(b)

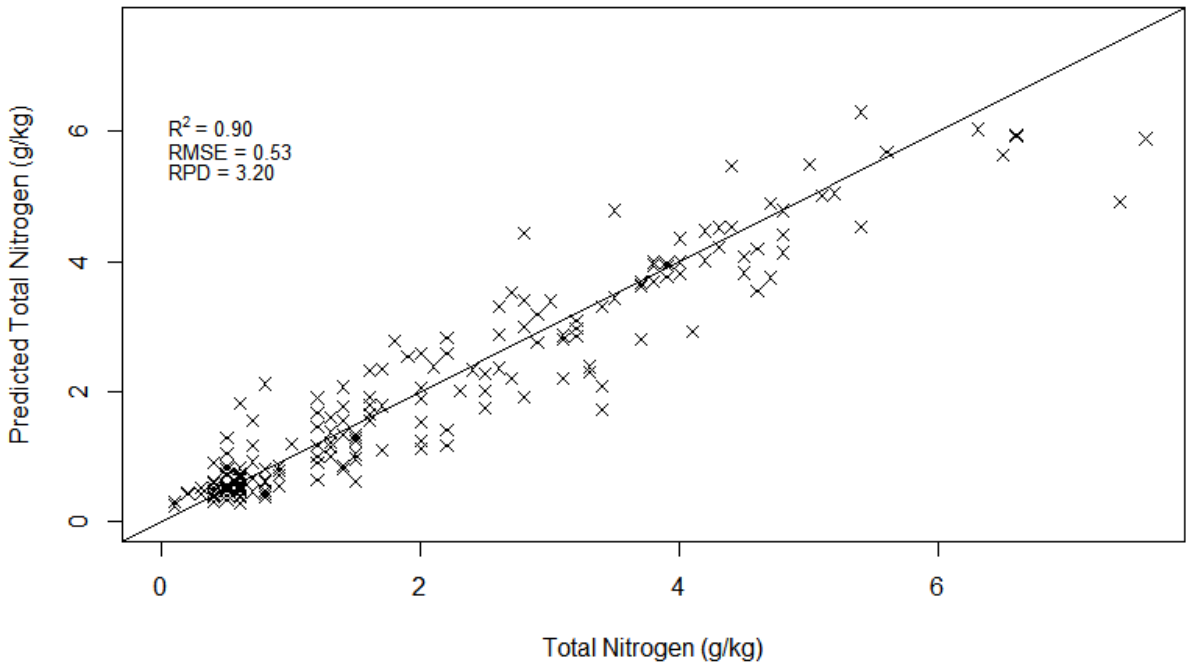
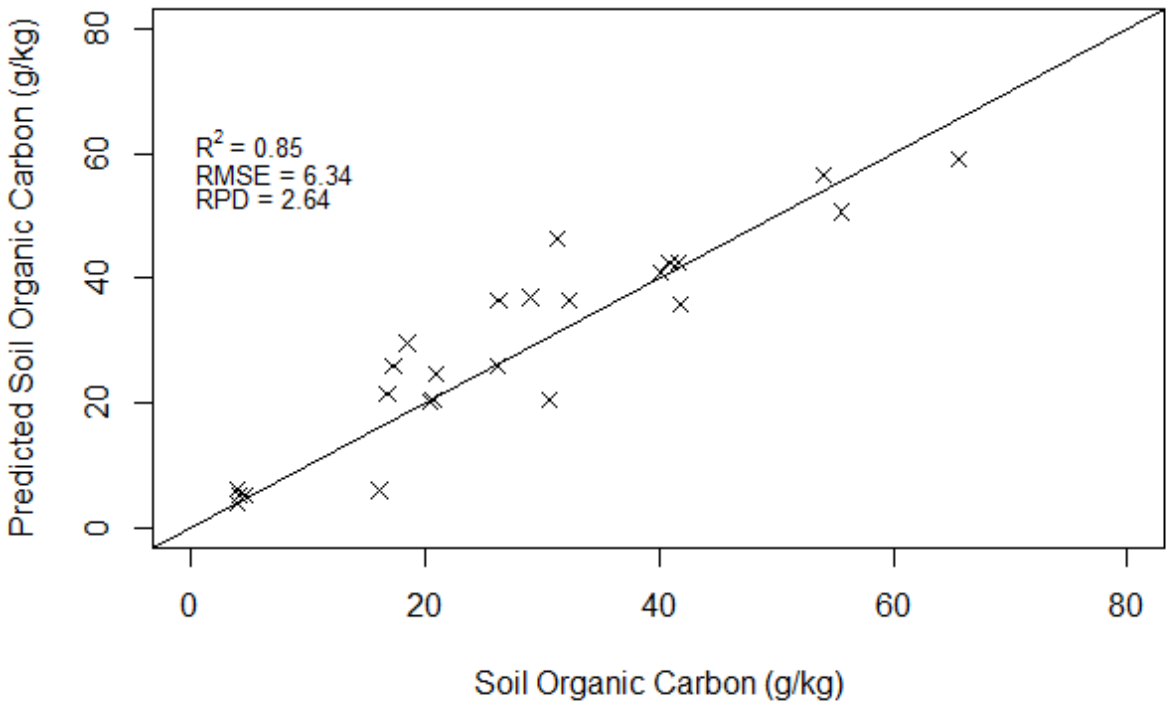
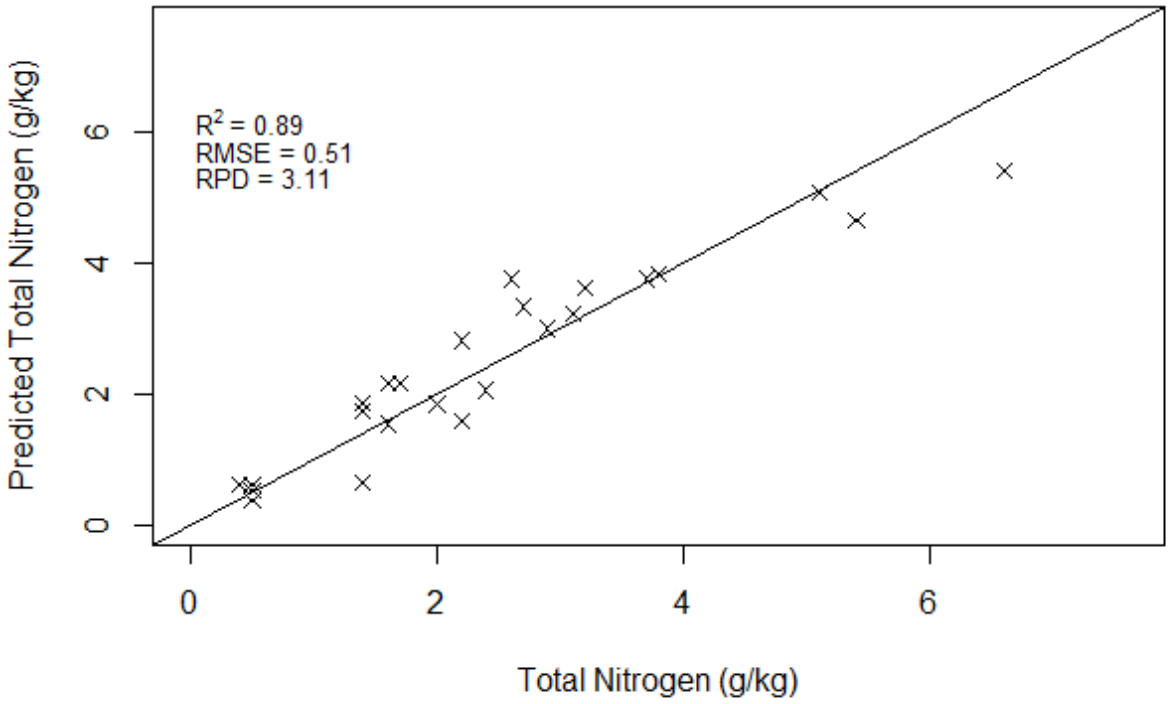

R^2 = 0.90
RMSE = 0.53
RPD = 3.20
Predicted Total Nitrogen (g/kg)
Total Nitrogen (g/kg)
0
2
4
6

Figure 2-3. Independent validation dataset predicted versus observed (a) soil organic carbon (SOC) content (g kg^{-1}), (b) total nitrogen as produced by the Cubist model. The model was developed using 183 samples, and validated using 24 samples, which were analyzed for SOC and for TN. The solid 1:1 line illustrates deviations between predicted and measured data.

(a)



(b)


R^2 = 0.89
RMSE = 0.51
RPD = 3.11
Predicted Total Nitrogen (g/kg)
Total Nitrogen (g/kg)
0
2
4
6

Previous work in Canada has obtained accurate PLSR prediction models, with RMSE values of 1.75 g kg^{-1} and an RPD of 4.03 for SOC and an RMSE value of 0.12 g kg^{-1} and RPD value of 4.12 for TN (Xie et al., 2011). It is important to note that these samples were all from a single soil type, and that samples were ground and homogenized, which is likely the reason for the more accurate predictive models compared to the models developed in the current study that focus on unground soil samples from multiple soil types. Results from a study using PLSR to predict SOC in Manitoba, Canada was less successful with an RMSE value of 3.45 g kg^{-1} and an RPD value of 1.97. A TN predictive model could not be successfully built due to errors associated with the calibration data (Martin et al., 2002). Spectra from that study were obtained from homogenized soil in the laboratory. A more recent laboratory spectroscopy study in Canada using processed and homogenized samples was able to obtain RPD values of 3.70 for SOC and 3.83 for TN using PLSR (St. Luce et al., 2014).

Our results are comparable to results from research in Belgium where soil cores were scanned directly with a point spectrometer, and the Cubist model produced a more accurate predictive model compared to PLSR with an RMSE of 1.3 g kg^{-1} and an RPD of 3.7 (Doetterl et al., 2013). The Belgian study differed from our study in that a point spectrometer was used as compared to an imaging system, and wavelet analysis was not used to pre-process the data. Morellos et al. (2016) also found that Cubist models produced more accurate predictive models compared to PLSR models when analyzing minimally processed soil samples with a point spectrometer for both SOC and TN. Subsetting the spectra using a feature selection process has

been reported to improve prediction results, and reduce the likelihood of overfitting. Feature selection using MARS, prior to analysis improved results of both MARS and artificial neural network (ANN) SOC prediction models in previous studies (Rossel et al., 2010). This process was confirmed to improve results in the current study using imaging spectroscopy data, and with the additional step of an initial manual feature selection step to limit the spectra available for feature selection to regions that are known to contain absorption features related to organic matter carbon bonds.

The predictive models produced in the current study has higher predictive capability than another study in Alberta using field spectroscopy. Using CWT analysis and a cubist model, SOC predictive models were produced with RMSE values of 6.0 g kg^{-1} and an RPD of 2.2 for SOC and 0.5 g kg^{-1} and an RPD value of 2.5 for TN (Sorenson et al., 2017); spectra from that study were collected under field conditions, and the spectrometer used only collected to 2200 nm. Important spectral features for SOC were identified as being above 2200 nm in the current study (Table 2-2**Error! Reference source not found.**), which could explain the poorer model p erformance obtained in (Sorenson et al., 2017).

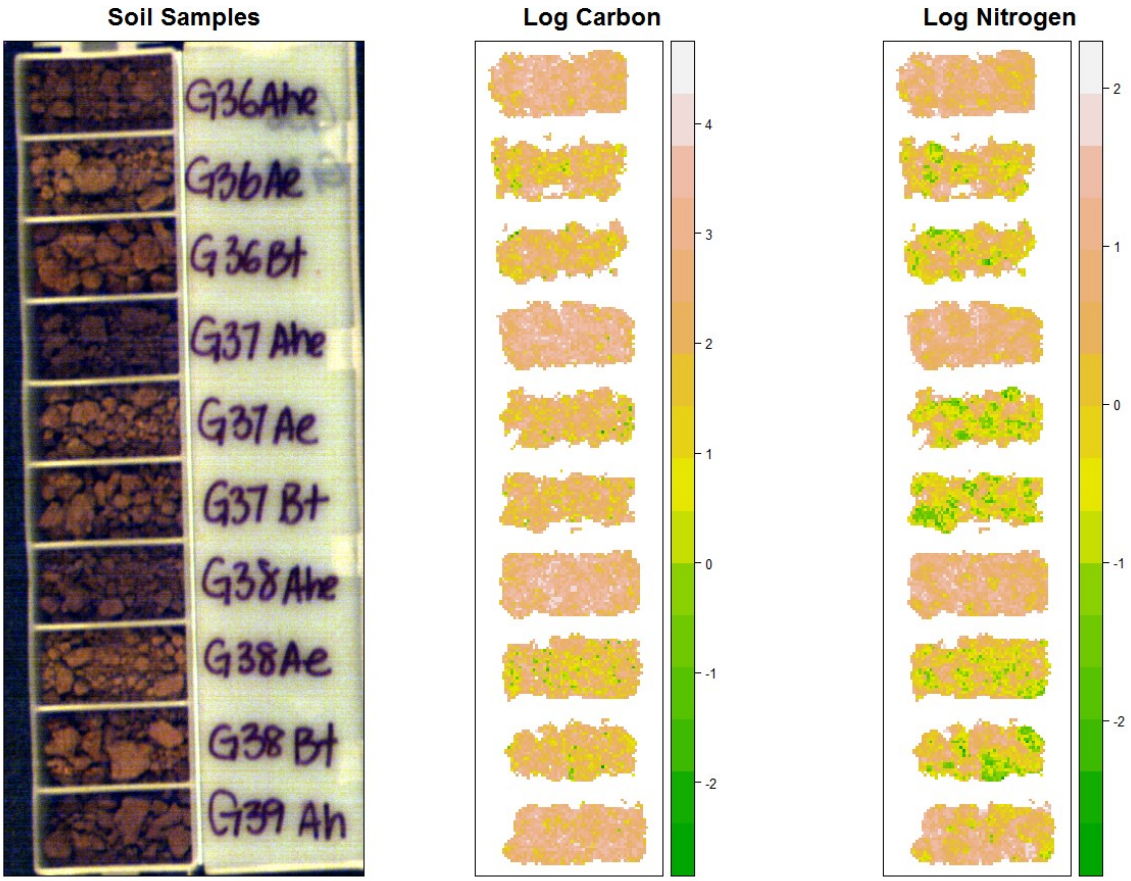
Spatial Analysis

An advantage of using imaging spectroscopy as opposed to point measurements is the ability to achieve finer spatial resolution. In this study, data were collected at a resolution of 0.2 mm. Following development with training data sets, the predictive models were applied to

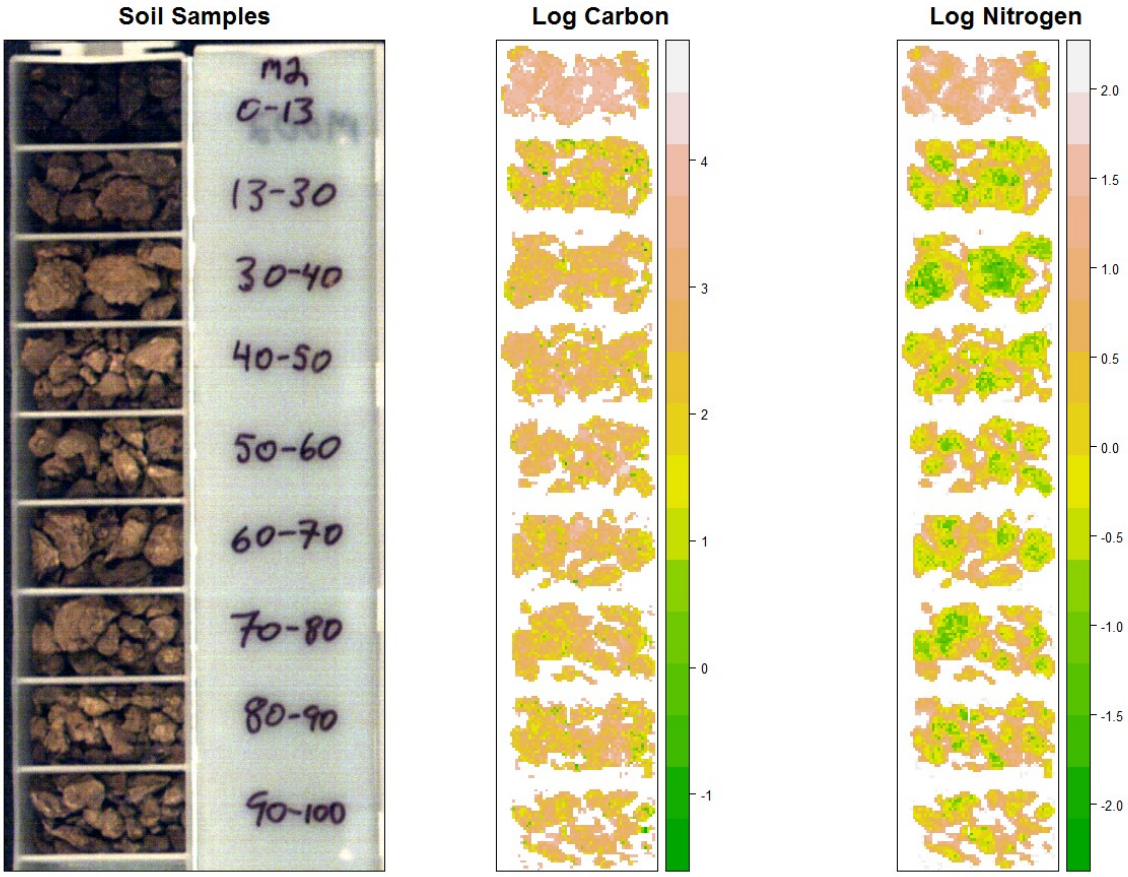
images to produce SOC and TN measurements for each pixel (Figure 2-4). Imaging spectroscopy allows the characterization of SOC and TN vertical changes within soil profiles and can accurately display their spatial variability. Samples from deeper positions in the profile have clear distinct inclusions of SOC and TN and their concentrations are not uniform. This within sample variation is also apparent in the topsoil samples (Figure 2-4). The spatial distribution of SOC within aggregates varied among horizons and soil types, with TN being universally more aggregated than SOC (Table 2-3). Results of the spatial analysis are illustrated on Figure 2-4. While the subsoil samples show patches of SOC and TN, the TN patches are larger than the SOC patches for each sample.

Figure 2-4. Total nitrogen (TN) and soil organic carbon (SOC) concentrations as predicted by a CUBIST model following wavelet transformation of the spectrum for each pixel. The image on the right is the true colour image generated from the hyperspectral imaging system. The label before the horizon indicates the sample location ID. The middle image is the log transformed SOC measurements ranging from 0 to 70 g kg^{-1} for each sample. The image on the left is the log transformed TN concentrations ranging from 0 to 7 g kg^{-1}. Samples in each image include: (a) Ahe, Ae and Bt horizons were collected from three different Dark Gray Luvisols used for crop production, (b) Orthic Black Chernozem developed on glaciolacustrine sediments used for crop production, (c) Orthic Gray Luvisol development on glacial till sediments under boreal forest vegetation.

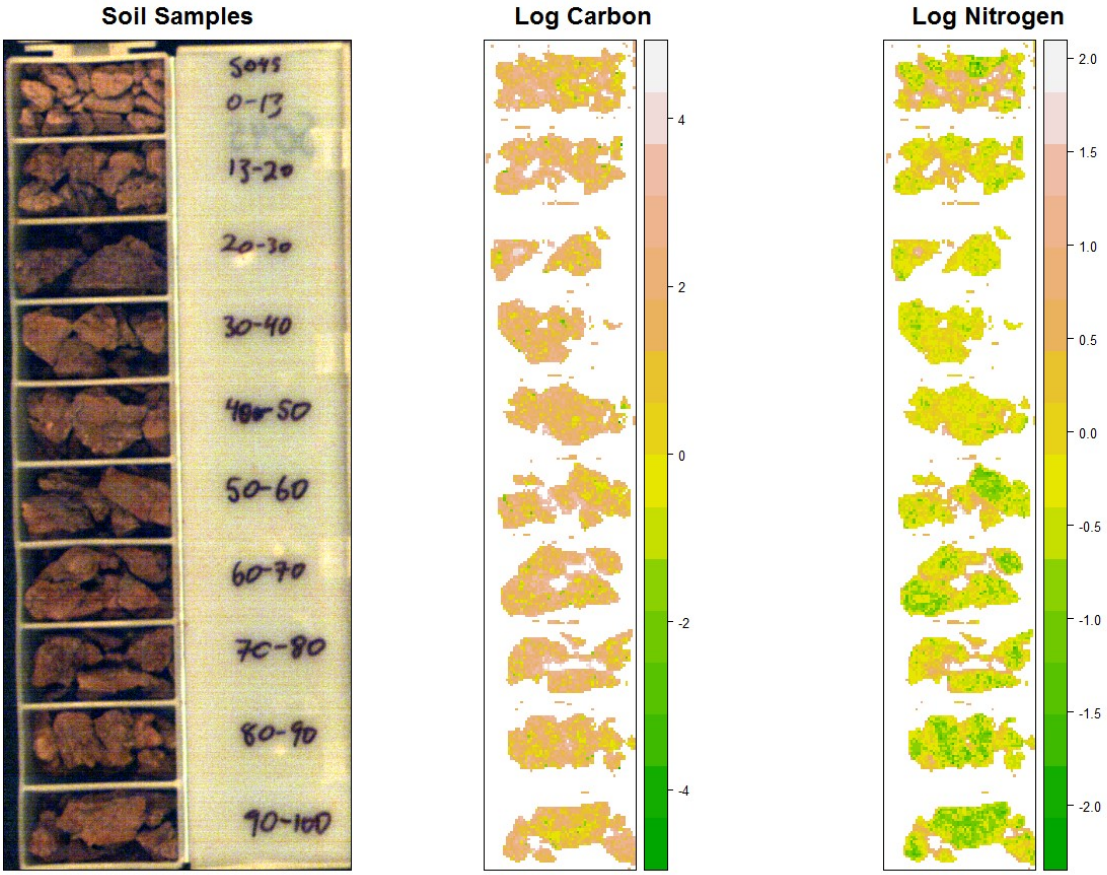
(a)



(b)


Soil Samples
M2
0-13
13-30
30-40
40-50
50-60
60-70
70-80
80-90
90-100
Log Carbon
4
3
2
1
0
-1
Log Nitrogen
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0

(c)


Soil Samples
Log Carbon
Log Nitrogen
Soils
0-13
13-20
20-30
30-40
40-50
50-60
60-70
70-80
80-90
90-100
4
2
0
-2
-4
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0

Brunisolic soils had a similar level of SOC aggregation compared to Chernozemic soil samples (Table 2-4). SOC in Gleysolic and Luvisolic soil samples was less aggregated, with the Luvisols having the least aggregated SOC; in these soils, SOC in both B and C horizons was more aggregated than in topsoil samples (Table 2-4). Visually, these patterns were illustrated in Figure 2-4, where little heterogeneity in the topsoil SOC is visible. Significantly more patchiness was visible in the distribution of SOC in the subsoil samples. Additionally, the Bm horizon in particular (13-30 cm depth in the Chernozemic sample in Figure 2-4b) had a higher number of low SOC patches (Figure 2-4).

Table 2-3. Average Moran's i values for soil organic carbon (SOC) and total nitrogen (TN) for each of the four soil orders included in this study, for all 370 samples analyzed using the SisuROCK automated hyperspectral imaging system. Moran's i was calculated on each soil sample to evaluate the relative aggregation of SOC and TN. A value of 1 indicated complete aggregation, a value of -1 indicates a completely regular spatial distribution and a value of 0 indicated random spatial distribution.

Soil Order	Horizon	SOC	TN
Brunisol	A	0.37	0.54
	B	0.46	0.66
	C	0.36	0.55
Chernozem	A	0.40	0.49
	B	0.33	0.47
	C	0.25	0.40
Gleysol	A	0.33	0.45
	B	0.36	0.46
	C	0.33	0.43
Luvisol	A	0.29	0.41
	B	0.32	0.44
	C	0.36	0.51

Table 2-4. Spatial generalized least squares model parameter estimates and p-values. Comparisons are relative to the intercept term, which is the Chernozem samples for the soil order comparisons and the A horizons for the soil horizon comparisons. The magnitude of the parameter estimate indicated the direction of the effect. A positive value is associated with higher Moran's i values, and a negative value is associated with lower Moran's i values relative to the default cases in the intercept term. Results are from all 370 samples analyzed using the SisuROCK automated hyperspectral imaging system.

Parameter	**SOC**		**TN**	
	Parameter Estimate	**P-Value**	**Parameter Estimate**	**P-value**
Intercept	0.27	0.06	0.13	0.40
Order – Brunisol	0.14	0.79	0.32	0.53
Order – Gleysol	-0.39	0.06	-0.20	0.34
Order – Luvisol	-0.65	0.01	-0.44	0.07
Soil – B Horizons	0.38	0.02	0.35	0.09
Soil – C Horizons	0.44	0.02	0.20	0.41

Given the low concentration of SOC in the C horizons, and considering that C horizons have limited pedogenesis, a possible explanation for the aggregated C horizon SOC could be due to heterogeneous and limited roots at depth, leaving limited amounts of aggregated SOC residues. Most common boreal tree species are known to produce vertical roots as an adaptation to different boreal conditions (Strong and La Roi, 1983). For example, aspen are known to grow taproots to access deep water resources and white spruce grow oblique lateral roots for mechanical support (Strong and La Roi, 1983).

Differences in SOC distribution among soil orders can be attributed to different pedogenic processes. Chernozemic soils developed under grassland conditions, whereas Gleysolic and Luvisolic soils developed under woody vegetation. Grasslands have a greater proportion of SOC at lower depths due to greater carbon inputs from root turnover (Jobbágy and Jackson, 2000), and deep-rooting grasses could explain the more regular SOC distribution at depth in Chernozemic soil samples. On the other hand, the higher degree of SOC aggregation in Chernozemic A horizons can be attributed to the higher biological activity typically observed in these soils compared to Luvisolic A horizons. This SOC aggregation is likely linked to the important role organic matter plays in soil aggregate formation. Organic matter has been observed to relatively rapidly become mineral associated and a part of soil aggregates in Chernozemic soils (Jastrow, 1996).

Interestingly, spatial relationships for TN were somewhat different from those observed for SOC (Tables 2-3 and 2-4). While TN was consistently more aggregated than SOC (Figure 2-4),

differences among soil orders and horizons were less apparent (Table 2-4). Luvisolic soil samples showed a weakly established ($p=0.07$) decrease in TN aggregation compared to other soil orders. Additionally, B horizons had a weakly established increase in TN aggregation compared to A horizons. The C horizons did not show an increase in aggregation compared to A horizons (Table 2-4). The TN aggregation was visible in each of the subsoil samples in Figure 2-4, with large patches of low TN concentrations in each subsoil sample.

While most of soil nitrogen occurs in association with carbon in soil organic matter, results from this study indicated that their distribution was not uniform, and that TN was more aggregated than SOC at the aggregate scale. Possible explanations for the increased TN aggregation could be linked to plant activity and uptake or due to differences in the extent of soil organic matter decomposition. Nitrogen limitation is widespread across many ecosystems, including boreal forests (LeBauer and Treseder, 2008). Therefore, biological activity may be part of a positive feedback loop where concentrated activity around areas of higher nitrogen leads to increased aggregation as nitrogen is alternately mineralized and immobilized. Alternatively, carbon to nitrogen ratios decrease as decomposition proceeds (Berg, 2000), and the ratio of microbially-derived nitrogen compounds compared to plant-derived compounds increases (Bingham and Cotrufo, 2016). Fine scale variation in factors such as water availability and litter inputs influences soil organic matter decomposition, and therefore the relative concentration of nitrogen, which may explain the increased aggregation of TN observed in our study. Further

work where these factors are controlled is needed to investigate their potential influence on TN aggregation in the profile.

2.5. Conclusions

Results from this study demonstrate the importance of imaging spectroscopy as a valuable tool for investigating SOC and TN patterns at the aggregate scale. The high resolution of measurements (<1 mm) enabled spatial patterns in the distribution SOC and TN to become visible, with TN being observed to be significantly more aggregated than SOC. These fine spatial scale relationships were found to be linked to larger spatial scale variation in soil type, which was likely due to distinct pedogenic processes. To build on this study, future work is needed to test this methodology on intact soil cores as compared to discrete samples to obtain high resolution SOC inventories of whole soil profiles.

Acknowledgements

We would like to acknowledge the National Science and Engineering Council (NSERC) for providing financial support for this project with a Discovery Grant to Sylvie Quideau (RGPIN-2014-04693).

2.6. References

Alberta Energy Regulator, 2016. Statistical Reports [WWW Document]. URL https://www.aer.ca/data-and-publications/statistical-reports (accessed 5.18.16).

Alberta Energy Regulator, 2011. Directive for Monitoring the Impact of Sulphur Dust on Soils. Edmonton, Alberta.

Anderson, D., Cerkowniak, D., 2010. Soil Formation in the Canadian Prairie Region. Prairie Soils Crop. 3, 57–64.

Bartholomeus, H.M., Schaepman, M.E., Kooistra, L., Stevens, A., Hoogmoed, W.B., Spaargaren, O.S.P., 2008. Spectral reflectance based indices for soil organic carbon quantification. Geoderma. https://doi.org/10.1016/j.geoderma.2008.01.010

Ben-Dor, E., Banin, A., 1995. Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj1995.03615995005900020014x

Ben-Dor, E., Banin, A., 1990. Near-infrared reflectance analysis of carbonate concentration in soils. Appl. Spectrosc. https://doi.org/10.1366/0003702904086821

Berg, B., 2000. Litter decomposition and organic matter turnover in northern forest soils. For. Ecol. Manage. 133, 13–22. https://doi.org/http://dx.doi.org/10.1016/S0378-1127(99)00294-7

Bingham, A.H., Cotrufo, M.F., 2016. Organic nitrogen storage in mineral soil: Implications for policy and management. Sci. Total Environ. 551–552, 116–126. https://doi.org/10.1016/j.scitotenv.2016.02.020

Blake, L., Goulding, K.W.T., 2002. Effects of atmospheric deposition, soil pH and acidification on heavy metal contents in soils and vegetation of semi-natural ecosystems at Rothamsted Experimental Station, UK. Plant Soil. https://doi.org/10.1023/A:1015731530498

Bremner, J.M., 1996. Nitrogen - Total, in: Sparks, D.L., Page, A.., Helmke, P.A., Loeppert, R.H., Soltanpour, P.N., Tabatabai, M.A., Johnston, C.T., Sumner, M.E. (Eds.), Methods of Soil Analysis Part 3—Chemical Methods. Madison, WI, pp. 1085–1123.

Chang, C.-W., Laird, D.A., 2002. NEAR-INFRARED REFLECTANCE SPECTROSCOPIC ANALYSIS OF SOIL C AND N. Soil Sci. https://doi.org/10.1097/00010694-200202000-00003

Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2001.652480x

Dessureault-Rompré, J., Zebarth, B.J., Burton, D.L., Georgallas, A., 2015. Predicting soil nitrogen supply from soil properties. Can. J. Soil Sci. https://doi.org/10.4141/cjss-2014-057

Doetterl, S., Stevens, A., Van Oost, K., van Wesemael, B., 2013. Soil Organic Carbon Assessment at High Vertical Resolution using Closed-Tube Sampling and Vis-NIR Spectroscopy. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2012.0410n

Ge, Y., Morgan, C.L.S., Ackerson, J.P., 2014. VisNIR spectra of dried ground soils predict properties of soils scanned moist and intact. Geoderma. https://doi.org/10.1016/j.geoderma.2014.01.011

Gomez, C., Lagacherie, P., Coulouma, G., 2012. Regional predictions of eight common soil properties and their spatial structures from hyperspectral Vis-NIR data. Geoderma. https://doi.org/10.1016/j.geoderma.2012.05.023

Gomez, C., Lagacherie, P., Coulouma, G., 2008a. Continuum removal versus PLSR method for clay and calcium carbonate content estimation from laboratory and airborne hyperspectral measurements. Geoderma. https://doi.org/10.1016/j.geoderma.2008.09.016

Gomez, C., Viscarra Rossel, R.A., McBratney, A.B., 2008b. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. Geoderma. https://doi.org/10.1016/j.geoderma.2008.06.011

He, Y., Huang, M., García, A., Hernández, A., Song, H., 2007. Prediction of soil macronutrients content using near-infrared spectroscopy. Comput. Electron. Agric. https://doi.org/10.1016/j.compag.2007.03.011

Hijmans, R.J., 2016. raster: Geographic Data Analysis and Modeling.

Indorante, S.J., Jansen, I.J., Boast, C.W., 1981. Surface mining and reclamation : Initial changes in soil character. J. Soil Water Conserv. 36, 347–351.

IUSS Working Group WRB, 2014. World reference base for soil resources 2014. International soil classification system for naming soils and creating legends for soil maps, World Soil Resources Reports No. 106. https://doi.org/10.1017/S0014479706394902

Jastrow, J.D., 1996. PIk soo38-0717(!spo159-x SOIL AGGREGATE FORMATION AND THE ACCRUAL OF PARTICULATE AND MINERAL-ASSOCIATED ORGANIC MATTER. Soil Biol. Biochem 28, 665–676.

Jobbágy, E.G., Jackson, R.B., 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecol. Appl. https://doi.org/10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2

Karatzoglou, A., Smola, A., Hornik, K., Seileis, A., 2004. kernlab - An S4 Package for Kernal Methods in R. J. Stat. Softw. 11, 1–20.

Kinoshita, R., Moebius-Clune, B.N., van Es, H.M., Hively, W.D., Bilgilis, A.V., 2012. Strategies for Soil Quality Assessment Using Visible and Near-Infrared Reflectance Spectroscopy

in a Western Kenya Chronosequence. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2011.0307

Kuhn, M., Johnson, K., n.d. Applied Predictive Modeling.

Kuhn, M., Weston, S., Keefer, C., Coulter, N., 2014. Cubist: Rule- and Instance-Based Regression Modeling.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., R Core Team, Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., 2016. caret: Classification and Regression Training.

Lagacherie, P., Baret, F., Feret, J.B., Madeira Netto, J., Robbez-Masson, J.M., 2008. Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2007.06.014

LeBauer, D.S., Treseder, K.K., 2008. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. Ecology 89, 371–379. https://doi.org/10.1890/06-2057.1

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 2, 18–22.

Martin, P.D., Malley, D.F., Manning, G., Fuller, L., 2002. Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy. Can. J. Soil Sci. https://doi.org/10.4141/S01-054

McBratney, A.B., Minasny, B., Viscarra Rossel, R., 2006. Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. Geoderma. https://doi.org/10.1016/j.geoderma.2006.03.051

McCarty, G.W., Reeves, J.B., Reeves, V.B., Follett, R.F., Kimble, J.M., 2002. Mid-Infrared and Near-Infrared Diffuse Reflectance Spectroscopy for Soil Carbon Measurement. Soil Sci. Soc. Am. J. 66, 640. https://doi.org/10.2136/sssaj2002.0640

McIntire, E.J.B., Fajardo, A., 2009. Beyond description: the active and effective way to infer processes from spatial patterns 90, 46–56.

McLean, E.O., 1982. Soil pH and Lime Requirement, in: Page, A.., Miller, R.H., Keeney, D.R. (Eds.), Methods of Soil Analysis Part 2, Chemical and Microbiological Properties. Madison, WI, pp. 199–225.

Melendez-Pastor, I., Navarro-Pedreño, J., Koch, M., Gómez, I., 2010. Applying imaging spectroscopy techniques to map saline soils with ASTER images. Geoderma. https://doi.org/10.1016/j.geoderma.2010.02.015

Mevik, B.H., Wehrens, R., Liland, K.H., 2015. pls: Partial Least Squares and Principal Component Regression.

Milborrow, S., 2016. earth: Multivariate Adaptive Regression Splines.

Minasny, B., McBratney, a B., Pichon, L., Sun, W., Short, M.G., 2009. Evaluating near infrared spectroscopy for field prediction of soil properties. Aust. J. Soil Res. 47, 664–673. https://doi.org/Doi 10.1071/Sr09005

Minasny, B., McBratney, A.B., 2008. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. Chemom. Intell. Lab. Syst. https://doi.org/10.1016/j.chemolab.2008.06.003

Morellos, A., Pantazi, X.-E., Moshou, D., Alexandridis, T., Whetton, R., Tziotzios, G., Wiebensohn, J., Bill, R., Mouazen, A.M., 2016. Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. Biosyst. Eng. https://doi.org/10.1016/j.biosystemseng.2016.04.018

Nawar, S., Buddenbaum, H., Hill, J., Kozak, J., Mouazen, A.M., 2016. Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy. Soil Tillage Res. https://doi.org/10.1016/j.still.2015.07.021

Nelson, D.W., Sommers, L.E., 1996. Total Carbon, Organic Carbon, and Organic Matter, in: Sparks, D.L., Page, A.., Helmke, P.A., Loeppert, R.H., Soltanpour, P.N., Tabatabai, M.A., Johnston, C.T., Sumner, M.E. (Eds.), Methods of Soil Analysis Part 3—Chemical Methods. Soil Science Society of America, Madison, WI, pp. 961–1011.

Ouerghemmi, W., Gomez, C., Naceur, S., Lagacherie, P., 2011. Applying blind source separation on hyperspectral data for clay content estimation over partially vegetated surfaces. Geoderma. https://doi.org/10.1016/j.geoderma.2011.04.019

Percival, D.B., Walden, A.T., William, A., Percival, D., Constantine, M.W., 2016. Wavelet Methods for Time Series Analysis.

Pinheiro, J., Bates, D., Debroy, S., Sarkar, D., R Core Team, 2016. _nlme: Linear and Nonliner Mixed Effects Models_.

R Core Team, 2018. R: A language and environment for statistical computing.

Reeves, D.W., 1997. The role of soil organic matter in maintaining soil quality in continuous cropping systems. soil& Tillage Res. Soil Tillage Res. Reeues/Soil Tillage Res. 43, 131–167.

Rivard, B., Feng, J., Gallie, A., Sanchez-Azofeifa, A., 2008. Continuous wavelets for the improved use of spectral libraries and hyperspectral data. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2008.01.016

Roger, J.M., Chauchard, F., Bellon-Maurel, V., 2003. EPO-PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. Chemom. Intell. Lab. Syst. 66, 191–204. https://doi.org/10.1016/S0169-7439(03)00051-0

Rossel, R.A.A.V., Behrens, T., Viscarra Rossel, R.A., Behrens, T., 2010. Using data mining to model and interpret soil diffuse reflectance spectra. Geoderma 158, 46–54. https://doi.org/10.1016/j.geoderma.2009.12.025

Scafutto, R.D.P.M., de Souza Filho, C.R., Rivard, B., 2016. Characterization of mineral substrates impregnated with crude oils using proximal infrared hyperspectral imaging. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2016.03.033

Shrestha, R.K., Lal, R., 2007. Soil Carbon and Nitrogen in 28-Year-Old Land Uses in Reclaimed Coal Mine Soils of Ohio. J. Environ. Qual. https://doi.org/10.2134/jeq2007.0071

Soil Classification Working Group, 1998. The Canadian System of Soil Classification. Can. Syst. Soil Classif. 3rd ed. Agric. Agri-Food Canada Publ. 1646 187.

Sorenson, P.T., MacKenzie, M.D., Quideau, S.A., Landhausser, S.M., 2017. Can spatial patterns be used to investigate aboveground-belowground links in reclaimed forests? Ecol. Eng. 104, 57–66. https://doi.org/10.1016/j.ecoleng.2017.04.002

Sorenson, P.T., Quideau, S.A., MacKenzie, M.D., Landhäusser, S.M., Oh, S.W., 2011. Forest floor development and biochemical properties in reconstructed boreal forest soils. Appl. Soil Ecol. https://doi.org/10.1016/j.apsoil.2011.06.006

Sorenson, P.T., Small, C., Tappert, M.C., Quideau, S.A., Drozdowski, B., Underwood, A., Janz, A., 2017. Monitoring organic carbon, total nitrogen, and pH for reclaimed soils using field reflectance spectroscopy. Can. J. Soil Sci. https://doi.org/10.1139/cjss-2016-0116

St. Luce, M., Ziadi, N., Zebarth, B.J., Grant, C.A., Tremblay, G.F., Gregorich, E.G., 2014. Rapid determination of soil organic matter quality indicators using visible near infrared reflectance spectroscopy. Geoderma. https://doi.org/10.1016/j.geoderma.2014.05.023

Steffens, M., Buddenbaum, H., 2013. Laboratory imaging spectroscopy of a stagnic Luvisol profile - High resolution soil characterisation, classification and mapping of elemental concentrations. Geoderma. https://doi.org/10.1016/j.geoderma.2012.11.011

Strong, W.L., La Roi, G.H., 1983. Root-system morphology of common boreal forest trees in Alberta, Canada. Can. J. For. Res. 13, 1164–1173.

Tappert, M.C., Rivard, B., Fulop, A., Rogge, D., Feng, J., Tappert, R., Stalder, R., 2015. Characterizing kimberlite dilution by crustal rocks at the Snap Lake diamond mine (Northwest Territories, Canada) using SWIR (1.90-2.36 μm) and LWIR (8.1-11.1 μm) hyperspectral imagery collected from drill core. Econ. Geol. https://doi.org/10.2113/econgeo.110.6.1375

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics, Fourth. ed. Springer, New York.

Viscarra Rossel, R.A., Lark, R.M., 2009. Improved analysis and modelling of soil diffuse reflectance spectra using wavelets. Eur. J. Soil Sci. https://doi.org/10.1111/j.1365-2389.2009.01121.x

Vitousek, P.M., Howarth, R.W., 1991. Nitrogen Limitation on Land and in the Sea : How Can It Occur ? Nitrogen limitation on land and in the sea : How can it occur ? 13, 87–115. https://doi.org/10.1007/BF00002772

Wang, S., Zhuang, Q., Wang, Q., Jin, X., Han, C., 2017. Mapping stocks of soil organic carbon and soil total nitrogen in Liaoning Province of China. Geoderma. https://doi.org/10.1016/j.geoderma.2017.05.048

Xie, H.T., Yang, X.M., Drury, C.F., Yang, J.Y., Zhang, X.D., 2011. Predicting soil organic carbon and total nitrogen using mid- and near-infrared spectra for Brookston clay loam soil in Southwestern Ontario, Canada. Can. J. Soil Sci. https://doi.org/10.4141/cjss10029

3. Distribution Mapping of Soil Profile Carbon and Nitrogen With Laboratory Imaging Spectroscopy

3.1. Abstract

Conversion of arable cropland to forage crops has been proposed as a potential method to increase soil organic carbon (SOC) stocks to sequester carbon and improve soil quality. In this study, intact soil cores were collected from long-term boreal forest soil research plots established in 1980 consisting of: a mixed arable crop and forage agroecological rotation (AE), continuous forage (CF), and continuous grain (CG) rotations. These cores were analyzed using a SisuROCK automated hyperspectral imaging system in a laboratory setting collecting shortwave infrared reflectance data. Samples were then analyzed for SOC and total nitrogen (TN) contents by dry combustion to prepare a training data set. Predictive models were successfully built for SOC and TN using a combination of wavelet analysis and Bayesian Regularized Neural Nets. The CF rotation was found to have the highest SOC and TN contents compared to AE rotation for only the top 3 and 4 cm, respectively. These two rotations had comparable contents for both parameters for the rest of the topsoil, which was greater than the SOC and TN contents in the CG rotation to depths of approximately 12 cm. Increases in both SOC and TN were associated with increased spatial aggregation at fine spatial scales. These results indicate that adding forages to rotations in boreal forest soils increases SOC and TN, however these changes were concentrated in the surface depths.

3.2. Introduction

Soil Organic Carbon (SOC) and nitrogen both play essential roles in terrestrial ecosystems. The amount of carbon in the form of SOC is 4.5 times that of biomass (Jobbágy and Jackson, 2000). Additionally, nitrogen plays an important role as nitrogen limitation is widespread in terrestrial ecosystems (Vitousek and Howarth, 1991), and it can be an important limitation on ecosystem productivity. The management of SOC has global implications and has become an important public policy priority. Increasing soil carbon sequestration with soil management has been proposed as a strategy to offset potentially 5 to 15 percent of global fossil fuel emissions (Lal, 2004). As a result, the influence of land use and management on carbon stocks has been an increasing research focus.

A range of land use changes have been documented to increase SOC. One strategy that has been proposed is the conversion of land from arable cropping to forage crop production (Guo and Gifford, 2002). Rates of SOC accumulation following conversion to forages vary widely depending on vegetation, soil conditions and management history (Post and Kwon, 2000). The maximum rates of SOC accumulation occur during the early perennial vegetation aggrading stage, with average SOC accumulation rates of 0.332 g C m^{-2} yr^{-1} observed (Post and Kwon, 2000). The incorporation of a range of management practices associated with forage production such as improved grazing, fertilization, increased proportion of legumes, and improved grass species and conversion from cultivation has led to C accumulation rates of 0.105 to more than 1 Mg C ha^{-1} yr^{-1} (Post and Kwon, 2000).

The conversion of land from arable crops to pasture has been documented to influence total nitrogen (TN) in addition to SOC. The cultivation of soil with perennial grasses has been shown to lead to a clear decline in TN (Malhi et al., 2003). The decline is particularly present in the low-density organic matter fraction that is not complexed with the soil mineral matrix which is of particular significance as it is the main source of readily mineralizable organic nitrogen in the soil (Malhi et al., 2003). Previous work at the same study plots as those in this study have found that rotation with a mix of forages and arable crops lead to a similar increase in TN as continuous forages when compared to continuous arable cropping (Ross et al., 2008).

Reflectance spectroscopy is increasingly maturing as a tool for SOC and TN measurement (i.e. Chang et al., 2001; McBratney et al., 2006; Nocita et al., 2015; Viscarra Rossel et al., 2006), and imaging spectroscopy presents opportunities to measure soil properties at spatial scales previously thought impossible (Hobley et al., 2018; Sorenson et al., 2018; Steffens and Buddenbaum, 2013). Laboratory imaging spectroscopy has been used to measure the quantity and spatial distribution of SOC and TN at the soil aggregate scale with soil samples collected in the Canadian Prairies (Sorenson et al., 2018). Additionally, emergence of new data processing methods has improved the accuracy of soil information retrieved from reflectance spectra. Specifically, the combination of wavelet analysis and machine learning models have produced more accurate results compared to conventional reflectance spectroscopy analysis methods such as Savitsky-Golay smoothing, derivatives and multiplicative scatter correction (Rossel et al., 2010; Sorenson et al., 2017; Sorenson et al., 2018; Viscarra Rossel and Lark, 2009).

Laboratory imaging spectroscopy has been used to estimate SOC and TN in an intact soil profile (Steffens and Buddenbaum, 2013), and to identify regions of SOC enrichment in the subsoil of intact soil cores (Hobley et al., 2018). However, it has not yet been used to characterize effects of land use change on SOC and TN at fine spatial scales in the soil profile and to identify specifically where changes in the profile are present.

The objectives of this study were to investigate the use of laboratory imaging spectroscopy for measuring SOC and TN at fine spatial scales throughout intact soil profiles that had experienced different land management practices. While previous work has investigated changes in soil properties due to conversion to forage crops, the tools used in these studies did not allow for investigation of where changes occurred in the soil profile at fine spatial scales. This study was focused on identifying at what depths changes could be detected in SOC and TN, and if these changes in SOC and TN were associated with a change in the spatial distribution of these parameters at the same fine spatial scales.

3.3. Materials and Methods

Sample collection and preparation

A total of 200 soil samples were collected from the University of Alberta Hendrigan Plots near Breton, Alberta, Canada (53.08N, 114.44W) in May 2018, and from selected unfertilized plots from the 5-year rotation of the Breton Classical Plots. The soil at these plots developed on glacial till parent material under boreal forest vegetation. The long-term research plots were established in 1929, and the Hendrigan Plots were established in 1980. The Classical Plots were

established due to concerns at the time with cultivating Gray Luvisols, particularly their low A horizon carbon contents and high B horizon clay contents relative to the Chernozemic soils cultivated in the Canadian Prairies at the time (Dyck et al., 2012).

The Hendrigan plots consist of three cropping system treatments: (i) continuous perennial grass-legume forage (CF), (ii) a continuous grain system planted with barley or triticale (CG), and (iii) an agroecological 8-year rotation (AER), with a rotation of barley-barley-faba beans-barley-hay-hay-hay-hay. The Classical Plots host a 2-year wheat-fallow rotation and a 5-year cereal-forage rotation with various fertility treatments (Dyck et al. 2012) are a wheat-oats-barley-hay-hay rotation – sampling was for this work was limited to the 5-year rotation. A detailed description of the three Hendrigan treatments and their management is available in Ross et al. (2008). Annual precipitation averages 547 mm with a mean annual temperature of 2.1°C. The soil consists of an Orthic Gray Luvisol (Albic Luvisol in WRB, and Typic Cryoboralf in USDA Soil Taxonomy System). One intact soil core was collected per treatment for the three Hendigran Plots (15 cores) and 6 cores from the classical plots, for a total of 21 cores. All cores were collected with Giddings Soil Coring Equipment (Giddings Machine Company Inc.) to a depth of one meter.

Spectral Measurements

Spectral data were collected using a SisuROCK automated hyperspectral imaging system developed by Spectral Imaging (Specim) Ltd., Finland. The SisuROCK collects data with two high-resolution spectral cameras. The first spectral camera collects reflectance data in the visible

near-infrared light range (VNIR; 400-1000 nm). The second camera collects reflectance data in the shortwave infrared range (SWIR; 1000-2500 nm). The SWIR data were used for this study as SOC, TN and Clay are primarily spectrally active in this range (Rossel et al., 2010), and automated feature selection using all bands only selected bands from the SWIR region. Data in the SWIR range were collected in 256 spectral bands with a spectral resolution of 10 nm and a spatial resolution of 0.2 mm. Quartz halogen lamps provided the SWIR illumination to each sample and the measured light spectrum for each pixel was converted to reflectance via normalization to the average light spectrum from a Spectralon® panel. The acquisition time for each core took place in less than a minute.

Laboratory Analysis

Prior to analysis for SOC and TN, 1 cm thick samples collected from the soil cores were ground with a Retsch MM200 ball mill grinder. Samples were then analyzed for SOC and TN by dry combustion on a Costech ECS 4010 Elemental Analyzer (EA) equipped with a thermocouple detector (Costech Analytical Technologies Inc., Valencia, USA). As none of the samples were from horizons that would contain carbonates, samples were not pretreated for carbonate removal. Samples were tested with dilute hydrochloric acid to confirm carbonates were not present in the samples. In total 200 samples were analyzed for SOC and TN.

All clay content analyses were performed using a Beckman Coulter LS 13 320 Laser Diffraction Particle Size Analyzer. Samples were first lightly ground with a mortar and pestle to break up large aggregates. Approximately 0.5 g of soil sample was then placed in a pyrex test

tube along with 2 ml of ten percent sodium hexametaphosphate and 10 ml of deionized water. The sample was sonicated for 60 seconds in the tube and then added to the aqueous liquid module well. The sample was sonicated for an additional 60 seconds in the well, soil particle size measurements were taken for 60 seconds and averaged for 93 fraction sizes. The clay content was then determined by summing the percentages for fraction sizes below 2 μm.

Processing of Spectral Data

To generate the predictive models, laboratory analysis of SOC, TN and clay for the 200 samples we compared to the spectra of each sample obtained from the average of all pixels within the image of the 1 cm soil interval in the ENVI Software Platform (Harris Geospatial Solutions, Melbourne, USA). The average spectra were then processed using continuous wavelet transforms (CWT) (e.g.; (Rivard et al., 2008; Scafutto et al., 2016; Tappert et al., 2015) using the wmtsa package in R (Percival et al., 2016), to reduce the influence of non-compositional effects such as varying particle size on the analysis. The CWT outputs were calculated using an eight scale, second order Gaussian transform. Scales 2, 3 and 4 were then summed and used for the model development (***Error! Reference source not found.***).

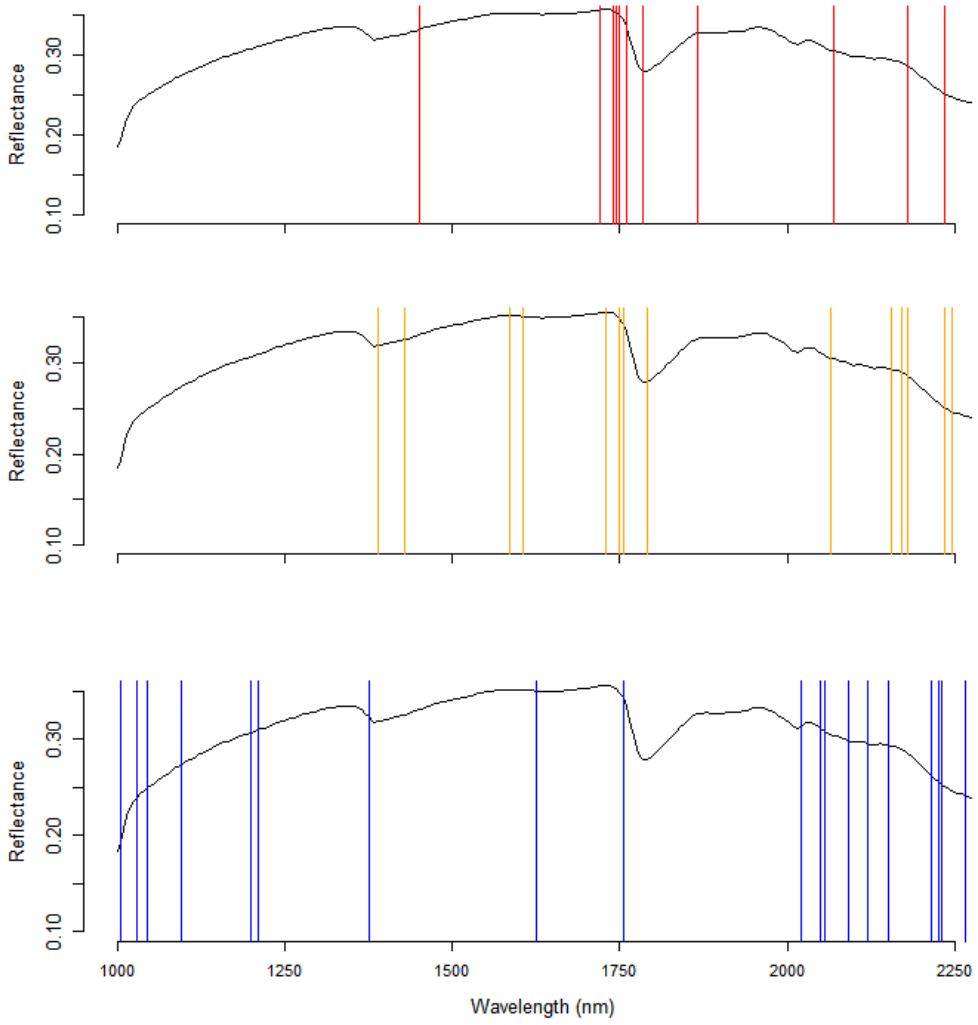
Model Development

Three predictive models were developed using the spectral data and their corresponding SOC, TN or clay data set. Models were developed using R (R Core Team, 2018). Seventy five percent of the dataset or 150 samples were used in the training data set. The remaining fifty

samples were selected to be in the test data set. The test data set was subset using the Kennard-Stone Algorithm in the prospectr package in R (Stevens and Ramirez-Lopez, 2013). Data was selected for training and testing datasets based on Mahalanobis distances.

A Bayesian regularized neural net (BRNN) model (Rodriguez and Gianola, 2016) was used to develop the SOC, TN and clay content predictive models. Spectra were spectrally subset prior to development of the predictive models. All spectral features were selected using the feature selection routine in the multivariate adaptive regression splines (MARS) model (Milborrow, 2018). Eleven features were selected for the SOC model, 14 features for the TN model and the twenty most important features selected by the MARS model for clay were used. These features were then used to build the BRNN model. The spectral features used for model development are illustrated in Figure 3-1. The BRNN models were developed with 5 neurons as this minimized the validation data set root-mean-square-error (RMSE).

Figure 3-1. Example shortwave infrared spectral signature. The sample is an Ahe horizon containing 5.9% soil organic carbon, 0.46% total nitrogen and 11% clay. The red lines on the upper plot indicate the spectral bands used for prediction of soil organic carbon. The orange lines on the middle plot indicate the spectral bands used for prediction of total nitrogen. The blue lines on the lower plot indicated the spectral bands used for prediction of clay content



Model Evaluation

Model accuracy was evaluated with the RMSE, R^2 and the ratio of performance to deviation (RPD), which is the ratio of the standard deviation to the RMSE. Generally, models with RPD values greater than 2 can be used accurately for prediction. Models with RPD values between 1.4 and 2 are satisfactory but could use improvement and models with values less than 1.4 have no predictive capability (Chang et al., 2001).

Carbon and Nitrogen Image Maps

Image Processing

Three processing steps took place prior to predicting SOC and TN for each pixel of the sample images. The hyperspectral images were first processed with a 3 x 3 median filter for spatial noise reduction. Following smoothing, a mask was created to remove the tray from the image prior to further processing. As the tray material has reflectance values peaking within the first 10 bands, contrary to the soil spectra, pixels with peak reflectance values in the first 10 bands were masked. The mask was created in R (R Core Team, 2018) using the raster package (Hijmans, 2017). Each pixel was then processed using an eight scale second order Gaussian transform, and wavelet scales 2, 3 and 4 were then summed. Lastly, the resulting wavelets coefficients were analyzed using the BRNN model generated during the statistical model development to estimate the SOC, TN and Clay content of each pixel in the images.

Spatial Analysis

The spatial distribution of SOC and TN was investigated by calculating Moran's i values each in a series of depth intervals for each soil core using the raster package in R (Hijmans,

2017). The spatial distribution of clay content was not investigated as it was expected to be influenced by pedogenesis and not recent additions of carbon. The depth intervals corresponded to approximately 3 cm, as the width of the images on along the x-axis was approximately 3 cm. This depth interval was selected such that the Moran's i calculations would be conducted on a roughly square image. Moran's i provides a measure of spatial autocorrelation, with a value of 1 indicating perfect aggregation, value of -1 indicating a regular spatial distribution and a value of 0 indicating a random spatial distribution. Additionally, the scale of spatial structure was determined by fitting a spherical semivariogram model and calculating the range for each depth interval using the geoR package in R (Ribeiro and Diggle, 2016).

Following the calculation of Moran's i values for SOC and TN, a spatial generalized least squares (GLS) model was run for each depth interval to determine if there was a difference between crop rotations in terms of SOC, TN, Moran's i for SOC and TN, the carbon to nitrogen ratio. All spatial GLS models were performed using the nlme package in R (Pinheiro et al., 2016). Analyses were performed separately for each depth interval so that depth where significant treatment effects occurred to could be identified. The underlying concept behind a spatial GLS model is that residual spatial patterns can be used as a surrogate for unmeasured variables (McIntire and Fajardo, 2009). Results from the spatial GLS model further indicate if a relationship between factors is present after removing unwanted variables by including location as a correlation structure in the GLS model (Sorenson et al., 2017). A significant intercept

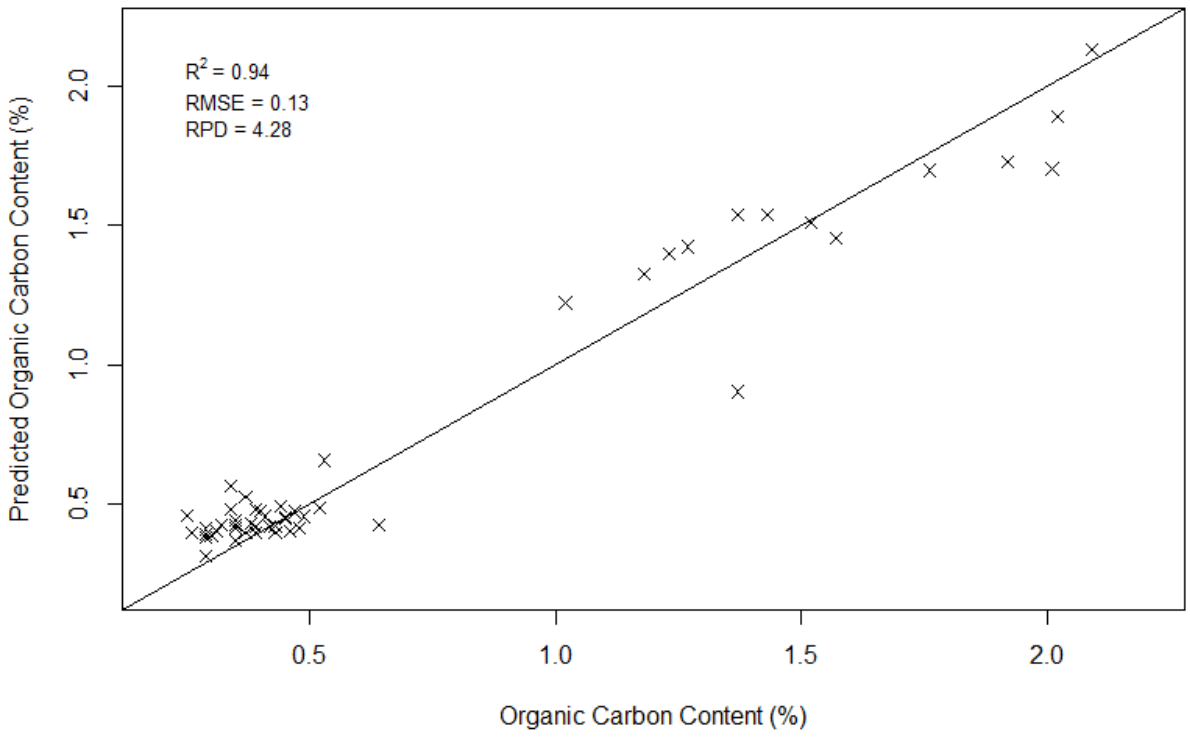
indicates that unaccounted for variables are affecting dependent variables. The parameter estimate indicates the direction and magnitude of the relationship.

3.4. Results

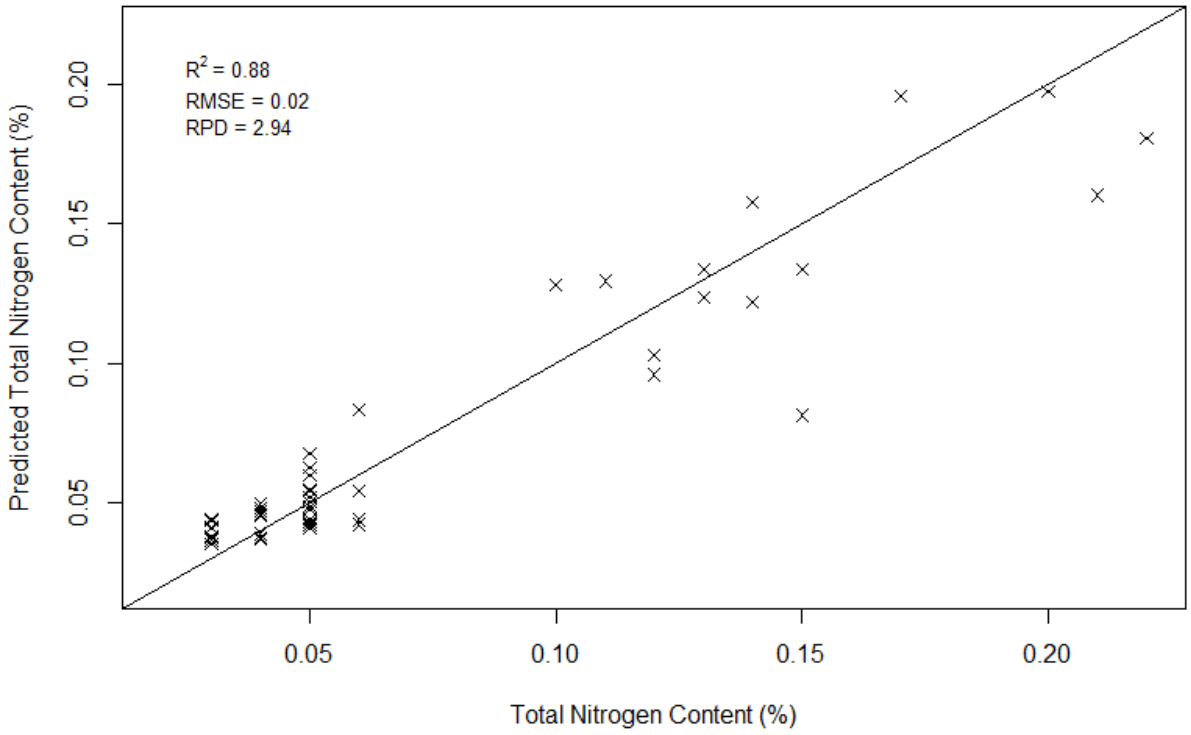
Models were successfully built for all three soil parameters, and model accuracy was determined based on the results of the independent validation data set. The BRNN model for SOC produced an R^2 value of 0.94, a RMSE of 0.14% and an RPD value of 4.28 (Figure 3-2a). Results for the TN model had an R^2 value of 0.88, an RMSE of 0.02% and an RPD value of 2.94 (Figure 3-2b). The clay model had an R^2 of 0.80, an RMSE of 2.70% and an RPD value of 2.23 (Figure 3-2c). Based on RPD threshold of 2, models for all three parameters can be considered accurate for prediction.

Figure 3-2. Independent validation dataset predicted versus observed (a) soil organic carbon (SOC) content (g kg^{-1}), (b) total nitrogen, and (c) clay content as produced by the Bayesian Regularized Neural Net model. The model was developed using 150 samples, and validated using 50 samples, which were analyzed for SOC, TN, and Clay Content. The solid 1:1 line illustrates deviations between predicted and measured data.

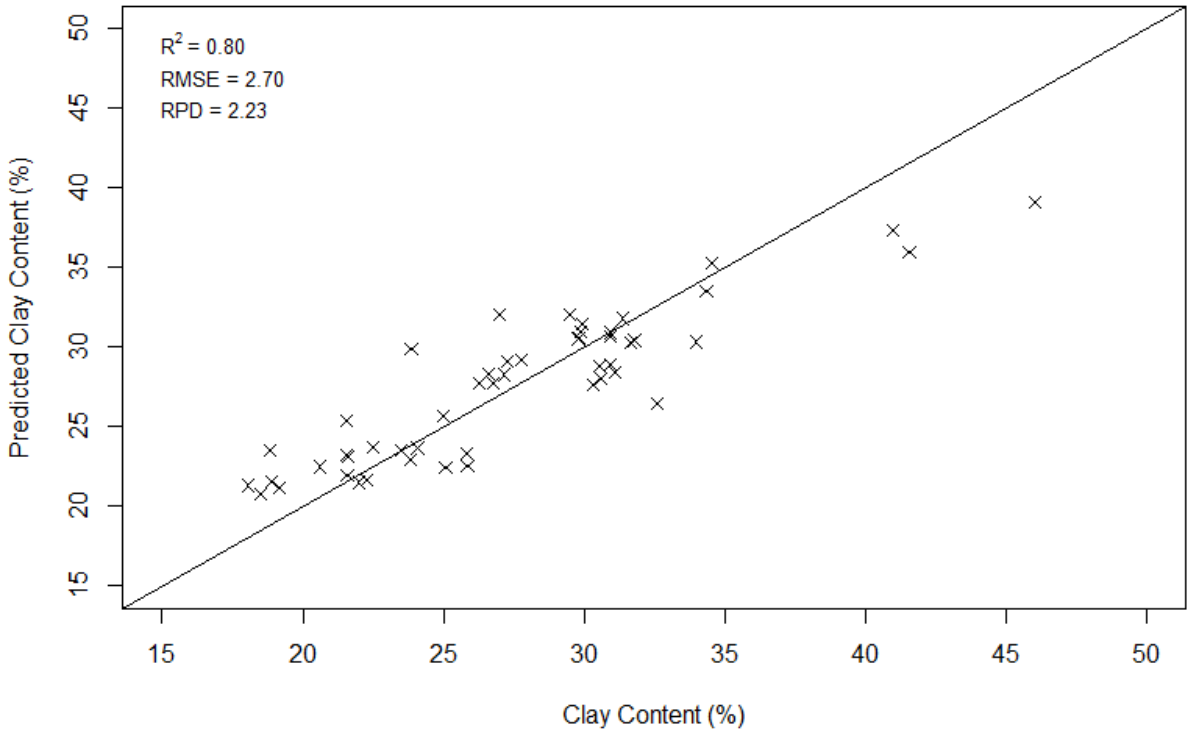
(a)



(b)


R^2 = 0.88
RMSE = 0.02
RPD = 2.94
Predicted Total Nitrogen Content (%)
0.05
0.10
0.15
0.20
Total Nitrogen Content (%)

(c)



The concentrations of SOC varied significantly amongst rotations in the topsoil and subsoil. The highest SOC concentrations were in the CF rotation, with an average SOC content of 2.3% (Table 3-1), followed by the AE rotation with an average SOC content of 2.2%. The CG and Classical rotations had SOC contents of 1.6% and 1.4%, respectively. The AE rotation and CF rotation had significantly more SOC than the CG and classical rotations (p-values <0.01). No differences in the subsoil SOC were detected with SOC concentrations of 0.4% in all rotations (Table 3-1). Based on these concentrations the total mass of SOC in the soil profile to 1 m on a megagrams per hectare (Mg ha^{-1}) basis ranged from a maximum of 120 Mg ha^{-1} in the CF

rotation and 119 Mg ha^{-1} in the agro-ecological rotation, to 98 Mg ha^{-1} in the CG rotation and 91 Mg ha^{-1} in the classical rotation (Table 3-1). The overall gain in SOC stocks was 29 Mg ha^{-1} for the CF and AE rotations compared to the CG rotation. This increase equates to a 22% increase in SOC stocks.

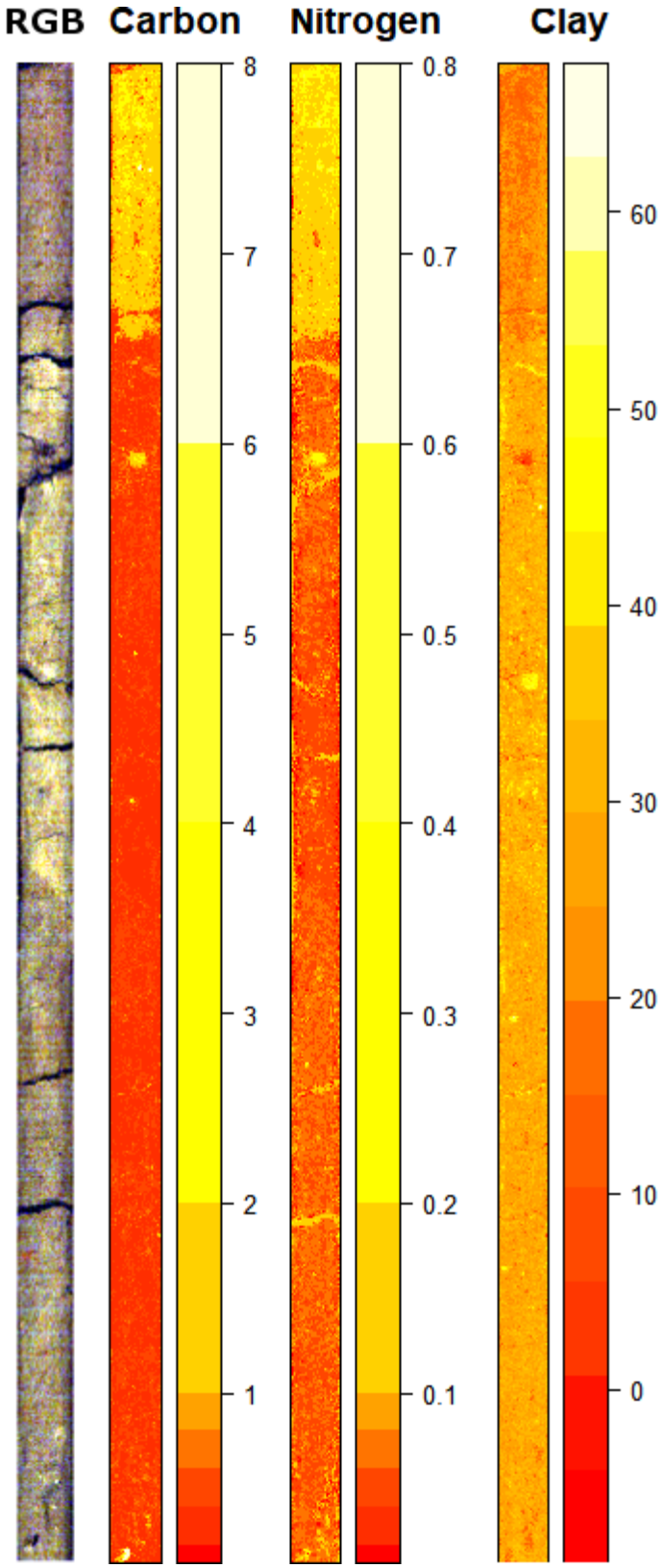
Table 3-1. Summary of mean and standard deviation laboratory soil parameter values for each rotation.

Rotation	Horizon	Soil Organic Carbon (%)	Total Nitrogen (%)	Clay Content (%)	Bulk Density (g cm^{-3})	Horizon Depth (cm)	Carbon (Mg ha^{-1})	Nitrogen (Mg ha^{-1})
Agro-ecological	A Horizon	2.2 ± 1.2	0.22 ± 0.10	21 ± 5	1.4 ± 0.1	21 ± 4	65 ± 0.5	6.5 ± 2.9
	B and C Horizon	0.4 ± 0.2	0.05 ± 0.03	30 ± 6	1.7 ± 0.2	79 ± 4	54 ± 0.2	6.7 ± 4.0
Continuous Forage	A Horizon	2.3 ± 1.8	0.22 ± 0.13	21 ± 5	1.2 ± 0.4	25 ± 3	69 ± 2.2	6.6 ± 3.9
	B and C Horizon	0.4 ± 0.2	0.06 ± 0.04	32 ± 7	1.7 ± 0.2	75 ± 3	51 ± 0.1	7.7 ± 5.1
Continuous Grain	A Horizon	1.6 ± 0.7	0.18 ± 0.07	22 ± 4	1.4 ± 0.2	20 ± 2	44 ± 0.3	5.0 ± 2.0
	A Horizon	0.4 ± 0.3	0.06 ± 0.03	30 ± 6	1.7 ± 0.3	80 ± 2	54 ± 0.2	8.2 ± 4.1
Classical	B and C Horizon	1.4 ± 0.7	0.15 ± 0.6	23 ± 4	1.4 ± 0.1	18 ± 4	35 ± 0.3	4.5 ± 0.8
	A Horizon	0.4 ± 0.3	0.06 ± 0.03	29 ± 6	1.7 ± 0.2	82 ± 4	56 ± 0.2	8.4 ± 4.1

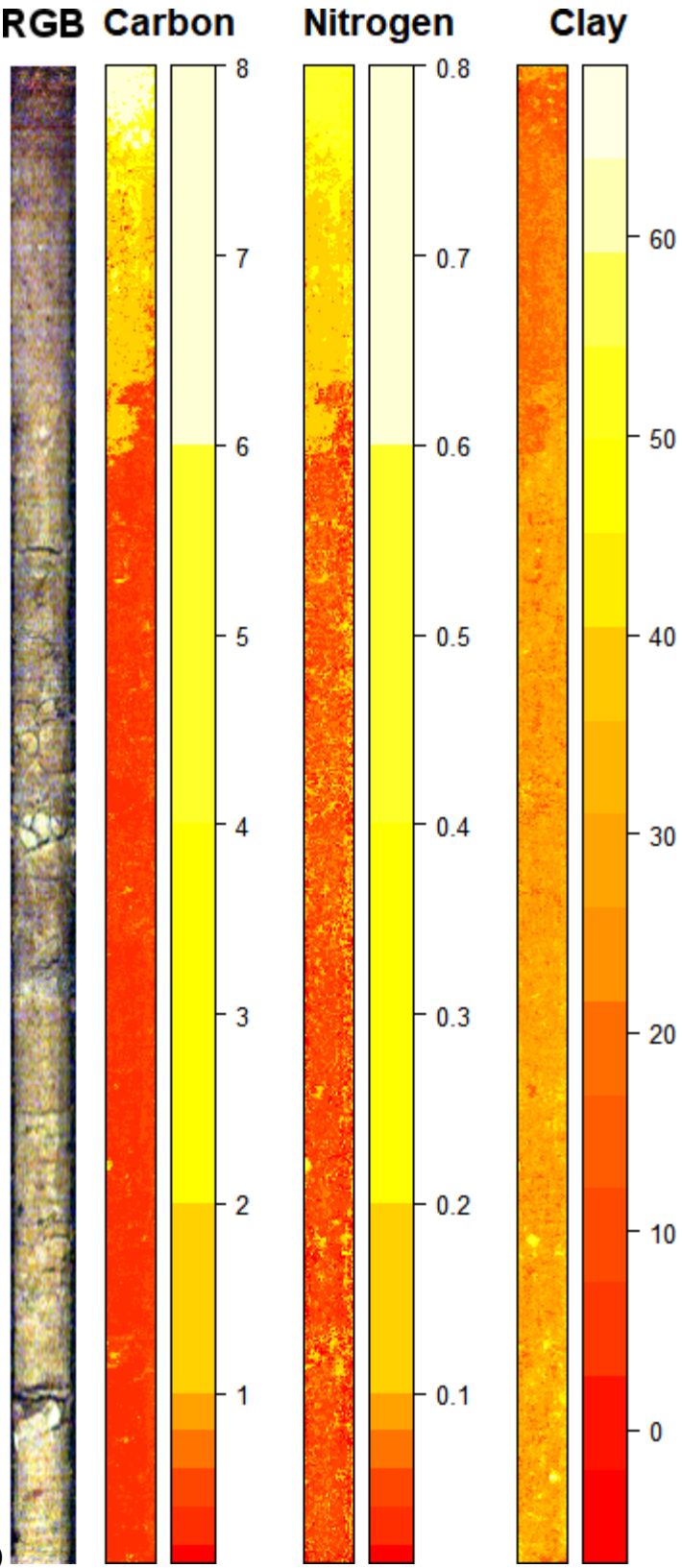
The imaging spectroscopy results illustrate that fine scale variance in SOC, TN and clay content is apparent in these soils. These soils have patches of higher SOC and TN concentration, particularly in the lower parts of the topsoil (Figure 3-3). Small isolated patches of elevated SOC and TN are clearly apparent in the subsoil as well. Additionally, the CF treatment clearly has a

more discontinuous A/B horizon boundary compared to the other treatments. These results also highlight both the lower clay content of the A horizon in these soils and that the subsoil has clear pockets of higher clay content, and heterogeneity of clay content at fine spatial scales.

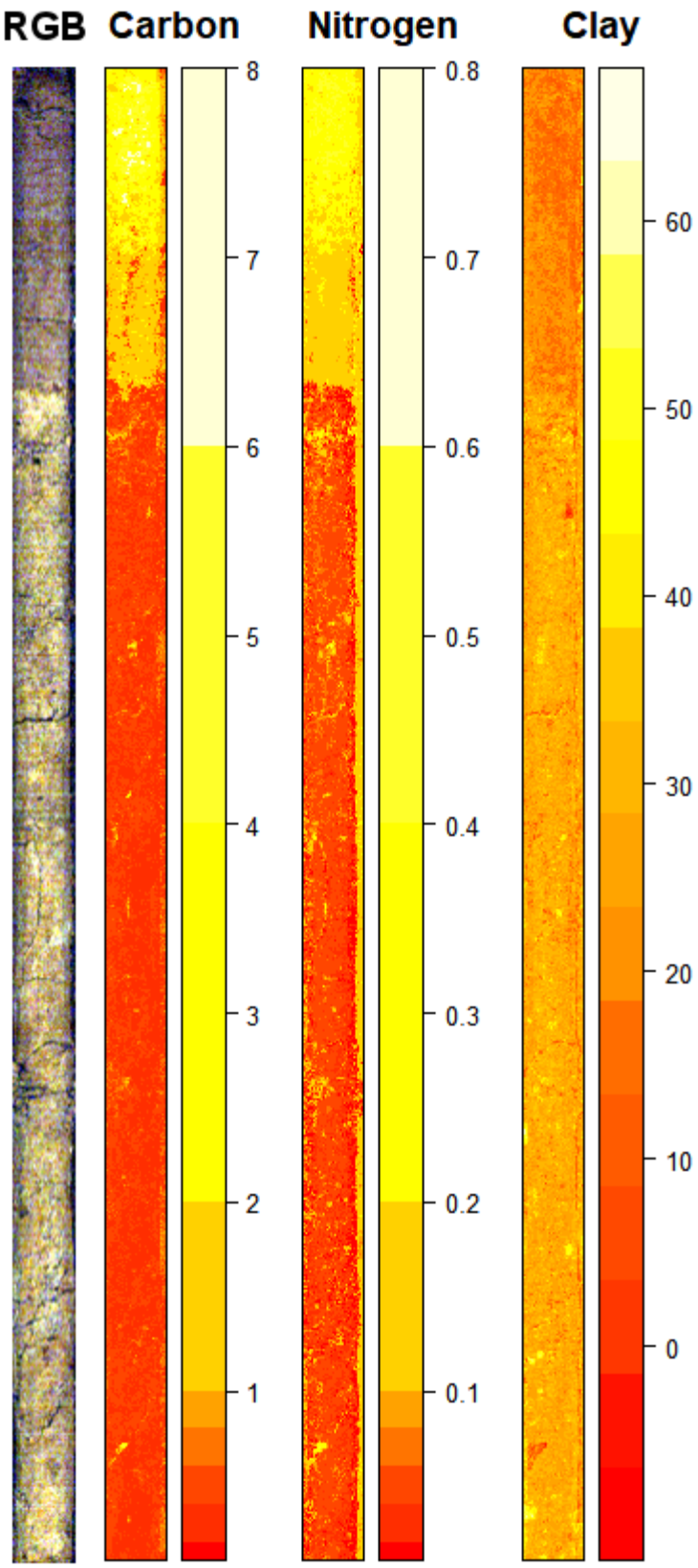
Figure 3-3. True Colour (RGB), soil organic carbon (SOC), nitrogen (TN), and clay contents as predicted by a Bayesian Regularized Neural Net model following wavelet transformation of the spectrum for each pixel for (a) a continuous grain profile, (b) a continuous forage profile and (c) an agro-ecological profile. The image on the left is the true colour image generated from the hyperspectral imaging system. The second image from the left is the predicted per pixel SOC, the third image from the left is the predicted per pixel TN concentration and the image on the right is the predicted per pixel clay content. Adjacent to each soil map is a color table for the corresponding soil parameter.


RGB
Carbon
Nitrogen
Clay
8
7
6
5
4
3
2
1
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
60
50
40
30
20
10
0

(a)


RGB
Carbon
Nitrogen
Clay
8
7
6
5
4
3
2
1
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
60
50
40
30
20
10
0

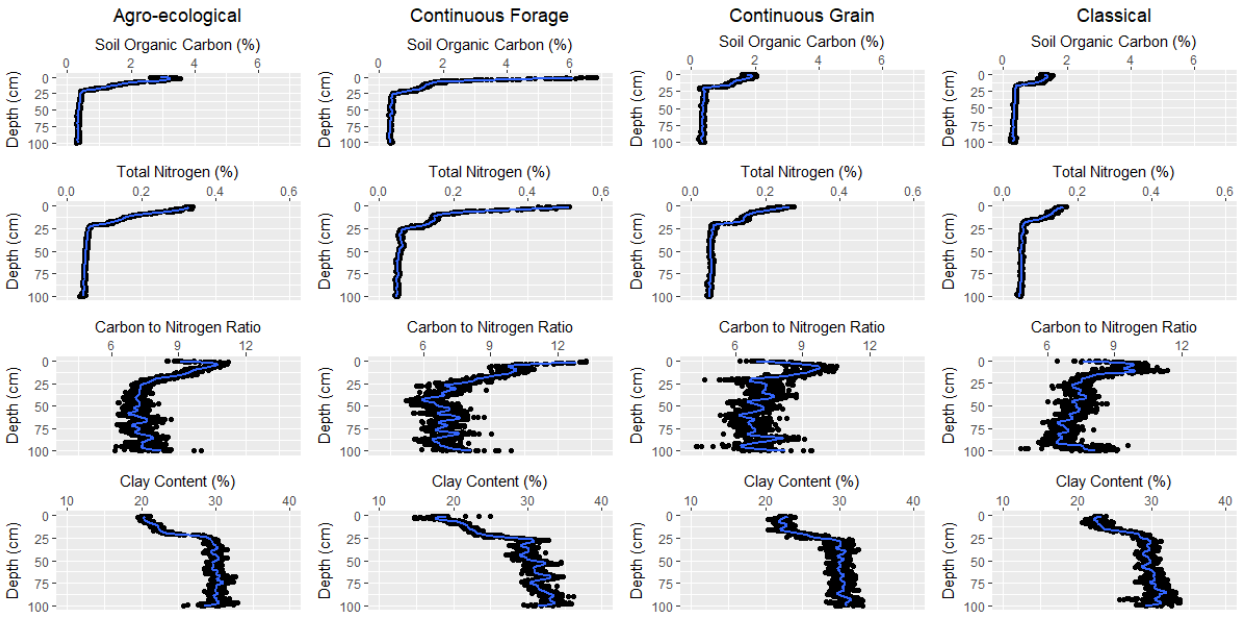
(b)


RGB
Carbon
Nitrogen
Clay
8
7
6
5
4
3
2
1
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
60
50
40
30
20
10
0

(c)

Each of the four treatments had a clear decreasing SOC content with depth that stabilized at concentrations near 0.4% in the subsoil (Figure 3-4). A similar trend was observable between TN and depth. The carbon to nitrogen ratio also showed a similar trend, however higher variance is observable in the subsoil compared to SOC and TN (Figure 3-4). The CF treatment did showed a continuous decrease in C:N content with depths, compared to the initial increase followed by a decrease observed in the other treatments (Figure 3-4). The clay content showed a gently sloping increase in clay content with depth in the A horizon. The clay content then increased sharply near the horizon boundary between the A and B horizons. This increase in clay content with the B horizon is consistent with Luvisolic soils, which are characterized by clay translocation from the topsoil to the subsoil.

Figure 3-4. Average depth profiles for soil organic carbon, nitrogen, carbon to nitrogen ratio, and clay content for each rotation: agro-ecological, continuous forage, continuous grain, and classical rotations. Measurements were obtained for approximately 1300 distinct depth intervals. Locally weighted scatterplot smoothing was used for plotting the line on each plot with a span of 0.1, equal to a 10 percent smoothing span. Depth profiles were created by averaging the values from each core by treatment.



SOC was found to be more aggregated in the topsoil for the AE, CF and CG rotations (Figure 3-5). However, no clear SOC spatial aggregation trend was present in the classical rotation. The SOC aggregation showed a decreasing trend with depth to a depth of approximately 50 cm. A slight increase in SOC spatial aggregation was then observable from 50 cm to 100 cm. This trend was most pronounced in the AE rotation. The AE, CF and CG rotations also showed a decrease in spatial aggregation of TN with depth (Figure 3-5). The increasing trend with depth after 50 cm present with SOC was not visible with TN. TN spatial aggregation was slightly lower in all rotations compared to SOC except for the CF rotation, which had similar Moran's i values for SOC and TN. The scale of spatial aggregation of SOC ranged from 1.3 cm to 2.4 cm, nitrogen ranged from 2.2 to 3.0 cm, and clay from 1.6 to 3.2 cm (Table 3-2). There were no significant differences amongst treatments. The average standard deviation for SOC Moran's i values was 0.07 for topsoil and 0.06 for subsoil. For nitrogen, the average Moran's i value standard deviation were 0.11 for topsoil and 0.07 for subsoil.

Figure 3-5. Soil organic carbon (SOC) and Total Nitrogen (TN) Moran's I values indicating the spatial aggregation of SOC and TN for each of the four rotations. A value of 1 indicates complete spatial aggregation, and a value of -1 indicates completely regularly distributed. Locally weighted scatterplot smoothing was used for plotting the line on each plot with a span of 0.1, equal to a 10 percent smoothing span.

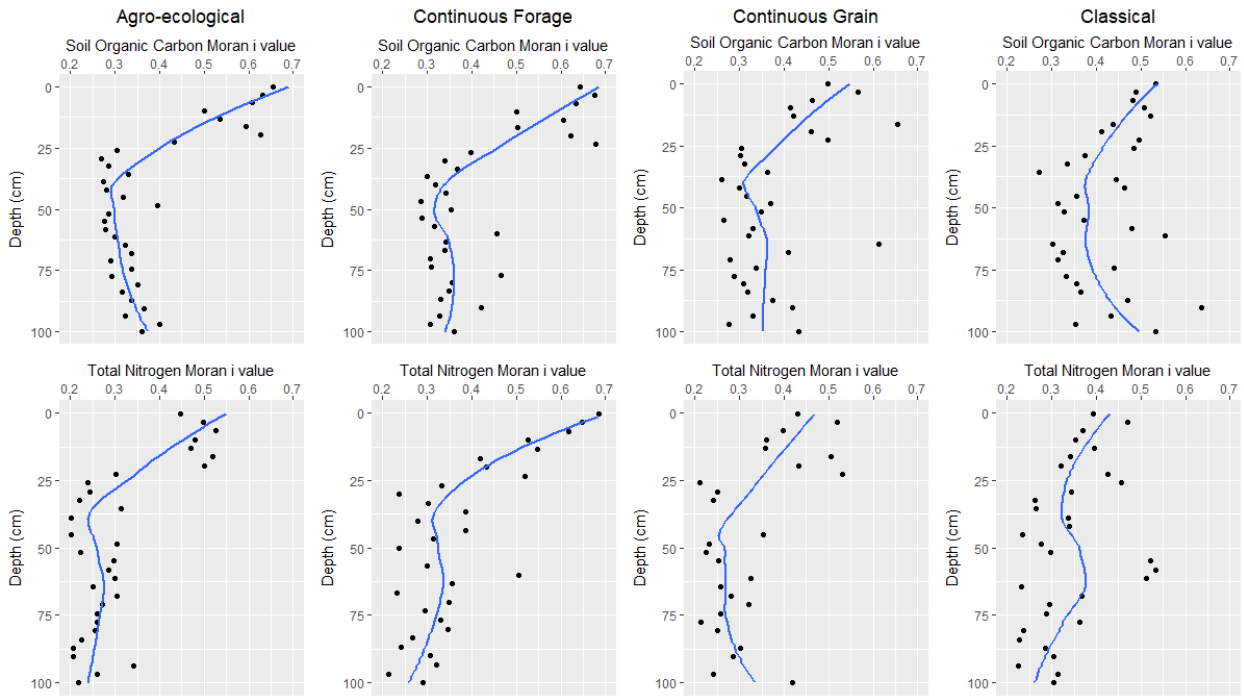


Table 3-2. Semivariogram range mean and standard deviation values for soil organic carbon, total nitrogen and clay content for each rotation. The semivariogram range indicates the distance where points are spatially independent from each other, and is related to the scale of spatial aggregation of the soil properties of interest. There were no significant differences across treatments or horizon.

Rotation	Horizon	Soil Organic Carbon (cm)	Total Nitrogen (cm)	Clay Content (cm)
Agro-ecological	A Horizon	2.1 ± 0.7	2.2 ± 0.9	3.2 ± 0.9
	B and C Horizon	1.8 ± 0.5	2.6 ± 0.3	1.8 ± 0.5
Continuous Forage	A Horizon	1.9 ± 0.8	2.7 ± 1.3	2.9 ± 0.5
	B and C Horizon	2.3 ± 0.4	2.5 ± 0.3	1.6 ± 0.4
Continuous Grain	A Horizon	1.3 ± 0.9	3.0 ± 0.2	2.6 ± 0.6
	B and C Horizon	1.7 ± 0.5	2.9 ± 0.2	1.8 ± 0.6
Classical	A Horizon	2.4 ± 1.2	2.4 ± 0.7	2.0 ± 0.7
	B and C Horizon	2.1 ± 0.4	2.6 ± 0.3	2.0 ± 0.2

While on average the AE and CF rotations had similar topsoil SOC concentrations, the CF rotation had significantly more SOC in the top 3.5 cm compared to the AE rotation (Figure 3-6; Table 3-3). The AE and CF rotations had significantly more SOC than the CG and Classical rotations to depths of 11.9 cm and 15.5 cm, respectively. The depth of the rotation effects was slightly deeper for TN compared to SOC. The CF rotation had significantly more TN than the AE rotation to a depth of 4.4 cm (Figure 3-6; Table 3-3). The CG and Classical rotations had significantly less TN to depths of 12.1 and 16.1 cm, respectively. The depth of any rotation effects on the carbon to nitrogen ratio were less compared to SOC and TN (Figure 3-6; Table 3-3). The CF rotation had a significantly higher carbon to nitrogen ratio to a depth of 2.5 cm, and

the CG and Classical rotations had significantly lower carbon to nitrogen rations to depths of 5.5 and 4.1 cm, respectively.

Figure 3-6. Soil Organic Carbon (SOC), Total Nitrogen (TN), Carbon to Nitrogen Ratio, SOC Moran's i, and TN Moran's i plots. Each plot shows data for the agro-ecological (AE), continuous forage (CF), continuous grain (CG), and classical treatments. Plots are provided for a depth of 0 to 25 cm as consistent change in soil parameters occurred within this depth range.

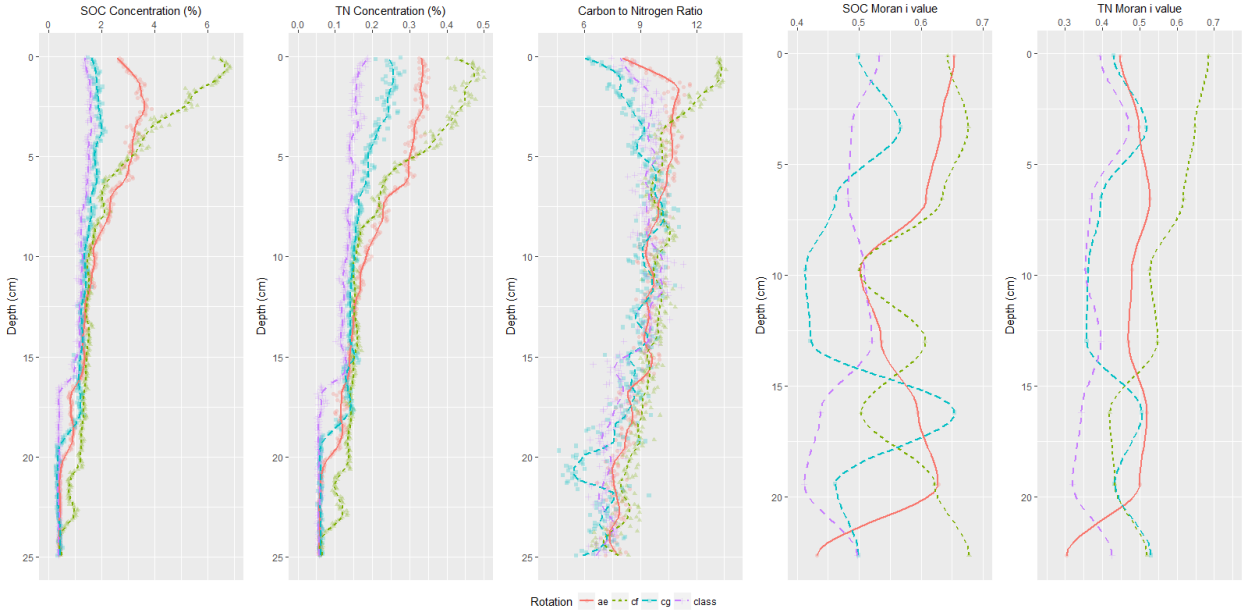


Table 3-3. Spatial generalized least squares model parameter estimates and p-values. Comparisons are relative to the intercept term in the model, which is the agro-ecological rotation. The magnitude of the parameter estimates indicates the direction of the effect. A positive value is associated with higher values relative to the intercept, and a negative value is associated with lower values relative to the default cases in the intercept term. The Parameter Estimate and P-values are the average for all the of the GLM models that were run separately for each depth from the surface to the depth where the rotation effect was no longer significant (alpha = 0.05). Where the depth of effect is NA, the Parameter Estimate and p-values are the average of all depth intervals.

Parameter	SOC			TN		
	Depth of Effect (cm)	Parameter Estimate	P-Value	Depth of Effect (cm)	Parameter Estimate	P-value
Continuous Forage	3.5	1.5	<0.01	4.4	1.1	<0.01
Continuous Grain	11.9	-1.1	0.02	12.1	-1.1	0.02
Classical	15.5	-1.4	0.01	16.1	-1.5	0.01

	Carbon Moran's i			Nitrogen Moran's i			C:N		
	Depth of Effect (cm)	Parameter Estimate	P-Value	Depth of Effect (cm)	Parameter Estimate	P-Value	Depth of Effect (cm)	Parameter Estimate	P-Value
Continuous Forage	NA	0.1	0.5	3.4	1.9	<0.01	2.5	1.4	0.01
Continuous Grain	3.4	-1.9	<0.01	NA	-0.1	0.5	5.5	-1.5	0.01
Classical	9.8	-1.1	0.04	NA	0.1	0.4	4.1	-1.0	0.03

The increased SOC content in the first 3.5 cm for the CF rotation compared to the AE rotation was not associated with a difference in the spatial aggregation of SOC (Figure 3-6; Table 3-3). However, both rotations had significantly more spatially aggregated SOC concentrations than the CG and Classical rotations. SOC was more aggregated in the first 3.4 cm compared to the CG rotation and the first 9.8 cm compared to the Classical rotation. A different trend was observed when comparing the spatial aggregation of TN by rotation (Figure 3-6; Table 3-3). The CF rotation had more spatially aggregated TN than the other rotations for the first 3.4 cm. No difference in the spatial aggregation of TN was detected among the AE, CG and Classical rotations.

3.5. Discussion

Previous work on the three rotations sampled in this work observed increases in SOC and TN with time (Wanit et al., 1994). More recent work on the Hendrigan Plots focused on TN found that the AE rotation had higher TN compared to the CF rotation (Ross et al., 2008). The Ross et al. (2008) study involved collecting either two discrete samples per plot or one composite sample per plot taken from 0-15 cm and 15-30 cm. This study did not find a significant difference between the two rotations based on average TN contents in the topsoil. However, the CF rotation did have significantly more TN in the top 4.1 cm compared to the AE rotation based on our data. The difference in results can be attributed to the finer spatial resolution data that was collected within our study, which enabled differences to be detected that were obscured with coarser spatial resolutions. Given the limited number of samples taken

from each plot in this study and the study done by Ross et al. (2008) it is possible the differences in results could be due to lateral spatial variation within the treatment plots.

The increased SOC and TN in the top 3-4 cm in the CF rotation was not associated with an overall increase in SOC stocks. Both rotations had a similar increase in total SOC stocks of 22 percent compared to the CG rotation. This is lower, but comparable, to other studies that have investigated the influence of conversion of cultivated lands to forages. On average an increase of SOC stocks of 39.2% have been observed (Conant et al., 2017). The lower change in this study may be attributable to the sharp increase in clay content and bulk density in the subsoil. Fibrous grass roots have been found to be inhibited in high clay content soils with bulk densities similar to those in this study (Chen and Weil, 2010).

The addition of SOC was largely in the topsoil, and the addition of SOC to the soil profile was associated with increased spatial aggregation of SOC. Of note in this study is that the CF treatment was not tilled, and the AE and CG rotations are tilled to a depth of 10 cm every five years (AE) or annually (CG) to incorporate manure and fertilizers. This management difference could explain the different trends in C:N ratios with depth between the CF rotation and other treatments (Figure 4). Other studies have found that the majority of SOC increases due to management changes occur at or near the surface. A study looking at the interaction between fertilization and no-till versus mouldboard plowing found only detectable differences in top 20 cm, and not when looking at profile to 60 cm (Poirier et al., 2009). Luo et al. (2010), in a meta-analysis of 69 studies, found that, on average, adoption of no till increased SOC

accumulation in the surface 30 cm. When depths greater than 40 cm were considered, the total SOC content was almost stable. Other work has found that the magnitude of SOC changes with management is depth-dependent (Conant et al., 2017), with an average increase of 23% in the surface 20 cm, and 12% lower in soil profiles. Overall the results are comparable in this study, with high spatial resolution measurements identifying increases in SOC to depths of 11 cm or 15 cm for the AE and CF treatments, depending on the conventional management practice.

The addition of TN to the soil was also largely concentrated in the topsoil, with the same pattern as with SOC, and an increase in TN was associated with increased spatial aggregation of TN. Previous work looking at the spatial aggregation of TN found on average TN more aggregated than SOC (Sorenson et al., 2018). However, that was not observed in this study, which may be because a narrower range of soil types were investigated. The increase in TN being higher in the topsoil compared to the subsoil is consistent with other studies. Research in Saskatchewan, Canada found that soil cultivation led to decreases in SOC and TN, particularly in the light fraction. These effects were strongest in the top 5 cm compared to deeper depth intervals (Malhi et al., 2003). Contrary to the findings in this study, a study in Northwest China did not find increased TN after conversion from annual crops to alfalfa, but did find an increase in particulate organic matter nitrogen (Su, 2007). However, that study was conducted four years after conversion to alfalfa, which may have been insufficient time for increases in TN to be detectable.

The results from this study indicate that the AE and CF rotations lead to similar increases in SOC compared to continuous grain or cereal-forage rotations.. The CF rotation was not associated with additional gains in SOC (Table 3-3) over the entire topsoil on average, with the CF rotation only having higher SOC concentrations in the top 3 cm. These increases in SOC were concentrated in the topsoil, and particularly in the shallower depths of the topsoil (Figure 3-4) and changes in SOC were not detectable deeper in the soil profile. It is possible that additions of SOC deeper in the profile may be occurring, however those additions could be leading to microbial stimulation and a loss of older stabilized carbon (Fontaine et al., 2007). The mechanisms associated with priming effects, namely either microbial nitrogen mining or stochiometric decomposition effects, has been observed in controlled laboratory studies (Chen et al., 2014). Other research has found mixed results regarding priming following the addition of fresh organic matter, with the need for priming effects to explain decomposition rates being highly dependant on the type of decomposition model being used (Cardinael et al., 2015). However, the high bulk density of the subsoil in these soils is likely reducing root penetration (Chen and Weil, 2010), and therefore the addition of SOC to the subsoil.

Overall, the results of this study indicate that SOC and TN have increased throughout the topsoil in the CF and AE plots compared to the CG and the unfertilized classical treatments. These increases were limited to the topsoil, likely due to the high bulk density and clay content of the subsoil characteristic of the Luvisolic soils of the study area. This does highlight the importance of soil type in terms of where SOC changes can be expected to occur following a

management change. Finally, while the CF rotation did have more SOC and TN in the top 3 to 4 cm compared to the AE rotation, these changes were not enough to lead to a change in the concentration of these parameters throughout the topsoil. Based on these results, the SOC and TN benefits from forages can likely be gained by introducing them into a rotation and there appears to be limited additional benefit from converting land permanently to forage crops in these soils.

3.6. Conclusions

Imaging spectroscopy is a valuable tool for investigating the fine spatial scale distribution of SOC and TN in the soil profile. Based on the imaging spectroscopy results from this study, SOC and TN are increased due to increasing frequency of forage crops in rotations. These gains are concentrated in the topsoil, and within the shallower depths of the topsoil. These results have important management implications, because if the majority of the additional SOC is concentrated in the topsoil, it may be more susceptible to loss if management practices are reversed. Disturbance of the topsoil will affect where most of the SOC gains have occurred, and reintroduction of management practices such as frequent tilling could quickly lead to a loss of the accumulated SOC in these soils.

Acknowledgements

We would like to acknowledge the National Science and Engineering Council (NSERC) for providing financial support for this project with a Discovery Grant to Sylvie Quideau (RGPIN-

2014-04693). We would also like to acknowledge the MITACS Accelerate program for funding to Preston Sorenson, and Dick Puurveen for project support.

3.7. References

Cardinael, R., Eglin, T., Guenet, B., Neill, C., Houot, S., Chenu, C., 2015. Is priming effect a significant process for long-term SOC dynamics? Analysis of a 52-years old experiment. Biogeochemistry 123, 203–219. https://doi.org/10.1007/s10533-014-0063-2

Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2001.652480x

Chen, G., Weil, R.R., 2010. Penetration of cover crop roots through compacted soils. Plant Soil 331, 31–43. https://doi.org/10.1007/s11104-009-0223-7

Chen, R., Senbayram, M., Blagodatsky, S., Myachina, O., Dittert, K., Lin, X., Blagodatskaya, E., Kuzyakov, Y., 2014. Soil C and N availability determine the priming effect: Microbial N mining and stoichiometric decomposition theories. Glob. Chang. Biol. 20, 2356–2367. https://doi.org/10.1111/gcb.12475

Conant, R.T., Cerri, C.E.P., Osborne, B.B., Paustian, K., 2017. Grassland management impacts on soil carbon stocks: a new synthesis. Ecol. Appl. 27, 662–668. https://doi.org/10.1002/eap.1473

Dyck, M.F., Roberston, J.A., Puurveen, D., 2012. The University of Alberta Breton Plots. Prairie Soils Crop. J. 5, 96–115.

Fontaine, S., Barot, S., Barré, P., Bdioui, N., Mary, B., Rumpel, C., 2007. Stability of organic carbon in deep soil layers controlled by fresh carbon supply. Nature. https://doi.org/10.1038/nature06275

Guo, L.B., Gifford, R.M., 2002. Soil carbon stocks and land use change: A meta analysis. Glob. Chang. Biol. https://doi.org/10.1046/j.1354-1013.2002.00486.x

Hijmans, R.J., 2017. raster: Geographic Data Analysis and Modeling.

Hobley, E., Steffens, M., Bauke, S.L., Kögel-Knabner, I., 2018. Hotspots of soil organic carbon storage revealed by laboratory hyperspectral imaging. Sci. Rep. 8, 13900. https://doi.org/10.1038/s41598-018-31776-w

Jobbágy, E.G., Jackson, R.B., 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecol. Appl. https://doi.org/10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2

Lal, R., 2004. Soil Carbon Sequestration Impacts on Global Climate Change and Food Security. Am. Assoc. Adv. Sci. 304, 1623–7. https://doi.org/10.1126/science.1097396

Luo, Z., Wang, E., Sun, O.J., 2010. Can no-tillage stimulate carbon sequestration in agricultural soils? A meta-analysis of paired experiments. Agric. Ecosyst. Environ. 139, 224–231. https://doi.org/10.1016/j.agee.2010.08.006

Malhi, S.S., Brandt, S., Gill, K.S., 2003. Cultivation and grassland type effects on light fraction and total organic C and N in a Dark Brown Chernozemic soil. Can. J. Soil Sci. 83, 145–153. https://doi.org/10.4141/S02-028

McBratney, A.B., Minasny, B., Viscarra Rossel, R., 2006. Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. Geoderma. https://doi.org/10.1016/j.geoderma.2006.03.051

McIntire, E.J.B., Fajardo, A., 2009. Beyond description: the active and effective way to infer processes from spatial patterns 90, 46–56.

Milborrow, S., 2018. earth: Multivariate Adaptive Regression Splines.

Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., Dor, E. Ben, Brown, D.J., Clairotte, M., Csorba, A., Dardenne, P., Demattê, J.A.M., Genot, V., Guerrero, C., Knadel, M., Montanarella, L., Noon, C., Ramirez-Lopez, L., Robertson, J., Sakai, H., Soriano-Disla, J.M., Shepherd, K.D., Stenberg, B., Towett, E.K., Vargas, R., Wetterlind, J., 2015. Soil Spectroscopy: An Alternative to Wet Chemistry for Soil Monitoring, in: Advances in Agronomy. https://doi.org/10.1016/bs.agron.2015.02.002

Percival, D.B., Walden, A.T., William, A., Percival, D., Constantine, M.W., 2016. Wavelet Methods for Time Series Analysis.

Pinheiro, J., Bates, D., Debroy, S., Sarkar, D., R Core Team, 2016. _nlme: Linear and Nonliner Mixed Effects Models_.

Poirier, V., Angers, D.A., Rochette, P., Chantigny, M.H., Ziadi, N., Tremblay, G., Fortin, J., 2009. Interactive Effects of Tillage and Mineral Fertilization on Soil Carbon Profiles. Soil Sci. Soc. Am. J. 73, 255. https://doi.org/10.2136/sssaj2008.0006

Post, M., Kwon, K.C., 2000. Soil Carbon Sequestration and Land-Use Change: Processes and Potential. Glob. Chang. Biol. 6, 317–328. https://doi.org/10.1046/j.1365-2486.2000.00308.x

R Core Team, 2018. R: A language and environment for statistical computing.

Ribeiro, P.J., Diggle, P.J., 2016. geoR: Analysis of Geostatistical Data.

Rivard, B., Feng, J., Gallie, A., Sanchez-Azofeifa, A., 2008. Continuous wavelets for the improved use of spectral libraries and hyperspectral data. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2008.01.016

Rodriguez, P.P., Gianola, D., 2016. brnn: Bayesian Regularization for Feed-Forward Neural Networks.

Ross, S.M., Izaurralde, R.C., Janzen, H.H., Robertson, J.A., McGill, W.B., 2008. The nitrogen balance of three long-term agroecosystems on a boreal soil in western Canada. Agric. Ecosyst. Environ. https://doi.org/10.1016/j.agee.2008.04.007

Rossel, R.A.A.V., Behrens, T., Viscarra Rossel, R.A., Behrens, T., 2010. Using data mining to model and interpret soil diffuse reflectance spectra. Geoderma 158, 46–54. https://doi.org/10.1016/j.geoderma.2009.12.025

Scafutto, R.D.P.M., de Souza Filho, C.R., Rivard, B., 2016. Characterization of mineral substrates impregnated with crude oils using proximal infrared hyperspectral imaging. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2016.03.033

Sorenson, P.T., MacKenzie, M.D., Quideau, S.A., Landhausser, S.M., 2017. Can spatial patterns be used to investigate aboveground-belowground links in reclaimed forests? Ecol. Eng. 104, 57–66. https://doi.org/10.1016/j.ecoleng.2017.04.002

Sorenson, P.T., Quideau, S.A., Rivard, B., 2018. High resolution measurement of soil organic carbon and total nitrogen with laboratory imaging spectroscopy. Geoderma. https://doi.org/10.1016/j.geoderma.2017.11.032

Sorenson, P.T., Small, C., Tappert, M.C., Quideau, S.A., Drozdowski, B., Underwood, A., Janz, A., 2017. Monitoring organic carbon, total nitrogen, and pH for reclaimed soils using field reflectance spectroscopy. Can. J. Soil Sci. https://doi.org/10.1139/cjss-2016-0116

Steffens, M., Buddenbaum, H., 2013. Laboratory imaging spectroscopy of a stagnic Luvisol profile - High resolution soil characterisation, classification and mapping of elemental concentrations. Geoderma. https://doi.org/10.1016/j.geoderma.2012.11.011

Stevens, A., Ramirez-Lopez, L., 2013. An introduction to the prospectr package.

Su, Y.Z., 2007. Soil carbon and nitrogen sequestration following the conversion of cropland to alfalfa forage land in northwest China. Soil Tillage Res. 92, 181–189. https://doi.org/10.1016/j.still.2006.03.001

Tappert, M.C., Rivard, B., Fulop, A., Rogge, D., Feng, J., Tappert, R., Stalder, R., 2015. Characterizing kimberlite dilution by crustal rocks at the Snap Lake diamond mine (Northwest Territories, Canada) using SWIR (1.90-2.36 μm) and LWIR (8.1-11.1 μm) hyperspectral imagery collected from drill core. Econ. Geol. https://doi.org/10.2113/econgeo.110.6.1375

Viscarra Rossel, R.A., Lark, R.M., 2009. Improved analysis and modelling of soil diffuse reflectance spectra using wavelets. Eur. J. Soil Sci. https://doi.org/10.1111/j.1365-2389.2009.01121.x

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma. https://doi.org/10.1016/j.geoderma.2005.03.007

Vitousek, P.M., Howarth, R.W., 1991. Nitrogen Limitation on Land and in the Sea : How Can It Occur? Biogeochemistry 13, 87–115. https://doi.org/10.1007/BF00002772

Wanit, S.P., Mcgill, W.B., Robertson, J.A., Thurston, J.J., 1994. Improved soil quality and barley yields and fababeans, manure, forages and crop rotation on a Gray Luvisol. Can. J. Soil Sci. 74, 75–84.

4. Monitoring organic carbon, total nitrogen and pH for field reclaimed soils using reflectance spectroscopy

4.1. Abstract

Assessing the success of soil reclamation programs can be costly and time consuming due to the cost of traditional soil analytical techniques. One cost-effective tool that has been successfully used to efficiently analyze a range of soil parameters is reflectance spectroscopy. We used reflectance data to analyze natural and reclaimed soils in the field, examining three key soil parameters: soil organic carbon (SOC), total nitrogen (TN), and soil pH. Continuous wavelet transforms combined with machine learning models were used to predict these parameters. Based on the root mean square error (RMSE), R^2 value and the Ratio of Performance to Deviation (RPD), the Cubist model produced the most accurate models for SOC, TN, and pH. The RMSE, R^2, and RPD values for SOC were 0.60%, 0.80, and 2.2, respectively. The TN model results were 0.05%, 0.81 and 2.5, and pH model results were 0.44, 0.69 and 1.8. Overall, the optimal model can be used to predict SOC and TN accurately, and improvements in the pH model are needed as pH values less than 6.5 were consistently overpredicted.

Key Words: Reflectance Spectroscopy; Reclamation; Carbon; Nitrogen

4.2. Introduction

Conventional methodologies for assessing soils before and after disturbance are time consuming and costly requiring large volumes of samples to be sent to a commercial laboratory for analytical measurement. There is a need for technologies that can provide accurate data in a more timely and cost-effective manner. In Alberta, the energy industry alone is responsible for soil disturbance associated with 470,000 wellsites, 39,000 other oil and gas facilities and 24 open pit coat mines (Alberta Energy Regulator, 2016b). In addition to the facility footprint, there is additional soil disturbance due to borrow pits, transmission lines, pipelines and access roads that are associated with these facilities.

Evaluating the success of soil reconstruction and reclamation requires measuring key soil parameters. Three soil parameters of particular relevance are soil organic carbon (SOC), total nitrogen (TN) and soil pH, that affect a number of key functions (i.e., nutrient cycling, plant productivity, soil structure development) important for the development of sites undergoing reclamation (Blake and Goulding, 2002; Dessureault-Rompré et al., 2015; Reeves, 1997). Research conducted on reclaimed mine topsoil in the Athabasca Oil Sands Region has shown that SOC and TN can change in response to disturbance and reclamation. SOC and TN have been shown to increase with time, following reclamation due to development of the forest floor (Sorenson et al., 2011). Similar findings were observed in agricultural mine soils, reclaimed forests, pasture and hay land uses which tend to have lower SOC content relative to undisturbed soils (Indorante et al., 1981; Shrestha and Lal, 2007). Soil pH can also be affected by

soil disturbance. For example, in the Canadian prairies, soils underlain by carbonate rich parent material, which are admixed with overlying topsoil during a disturbance, can lead to an elevated soil pH (Anderson and Cerkowniak, 2010). Furthermore, in locations such as Alberta, where sulfur dust is a common by-product of the desulfurization of oil and natural gas, low pH can also result from the addition of elemental sulfur to soil or as a result of industrial emissions of sulfur dioxide (Alberta Energy Regulator, 2011).

Unfortunately, obtaining laboratory data for SOC, TN, and pH can be costly and time consuming, which can serve as a barrier to reclamation and monitoring programs. More affordable options include using field systems to collect reflectance spectroscopy data in a rapid and non-destructive manner. Reflectance spectroscopy has been used to predict SOC, TN and pH in a number of studies (Bartholomeus et al., 2008; Ben-Dor and Banin, 1995; Chang et al., 2001; Ge et al., 2014; Gomez et al., 2008b; McBratney et al., 2006; Rossel et al., 2010). SOC has been measured with error rates as low as 1.5 g/kg on dried intact soil cores and on samples ranging from 0.1% to 10.4% TOC (Doetterl et al., 2013; McCarty et al., 2002). While work conducted in the Canadian provinces of Manitoba and Ontario has successfully shown that reflectance spectroscopy can be used to predict SOC and TN (Martin et al., 2002; Xie et al., 2011), there is a lack of information on the use of reflectance spectroscopy on Canadian soils and more specifically, Alberta soils.

The processing of soil spectral data has focused on using partial least squares regressions (e.g. Gomez et al. 2008a, 2008b; Xie et al. 2011; Kinoshita et al. 2012) and using signal processing

techniques such as Savitsky-Golay smoothing, derivatives, and multiplicative scatter correction (e.g. He et al. 2007; Minasny et al. 2009; Doetterl et al. 2013; Nawar et al. 2016). Recently, two data processing techniques have been applied to spectral data and have improved the accuracy of model predictions. Machine learning tools have been successfully applied to spectral data and generated more accurate predictions than partial least squares regression models. Specifically, multivariate adaptive regression splines, support vector machines, and CUBIST models have tended to produce the most accurate predictions (Doetterl et al., 2013; Nawar et al., 2016; Rossel et al., 2010). Additionally, the application of an alternative signal processing method, the wavelet transform, has improved the use of spectral soil data to accurately predict different soil properties (Rossel et al., 2010; Viscarra Rossel and Lark, 2009). Wavelet analyses can be discrete or continuous, while the continuous wavelet transform can be highly redundant, it is directly comparable to the original spectra and therefore manual inspection of the wavelet results is easier (Rivard et al., 2008). Detailed information about the application of the continuous wavelet transform to spectral data is available in Rivard et al. (2008).

Overall, the main objective of this research was to determine if reflectance spectroscopy data collected in the field with a drill-rig mounted spectroscopy system cosuld be used to measure SOC, TN and pH. Specifically, this study focused on using machine learning models, which have been applied to soils in other regions (Doetterl et al., 2013; Rossel et al., 2010), along with wavelet transforms, and determining if they can be successfully applied to reflectance spectroscopy data from Canadian soils.

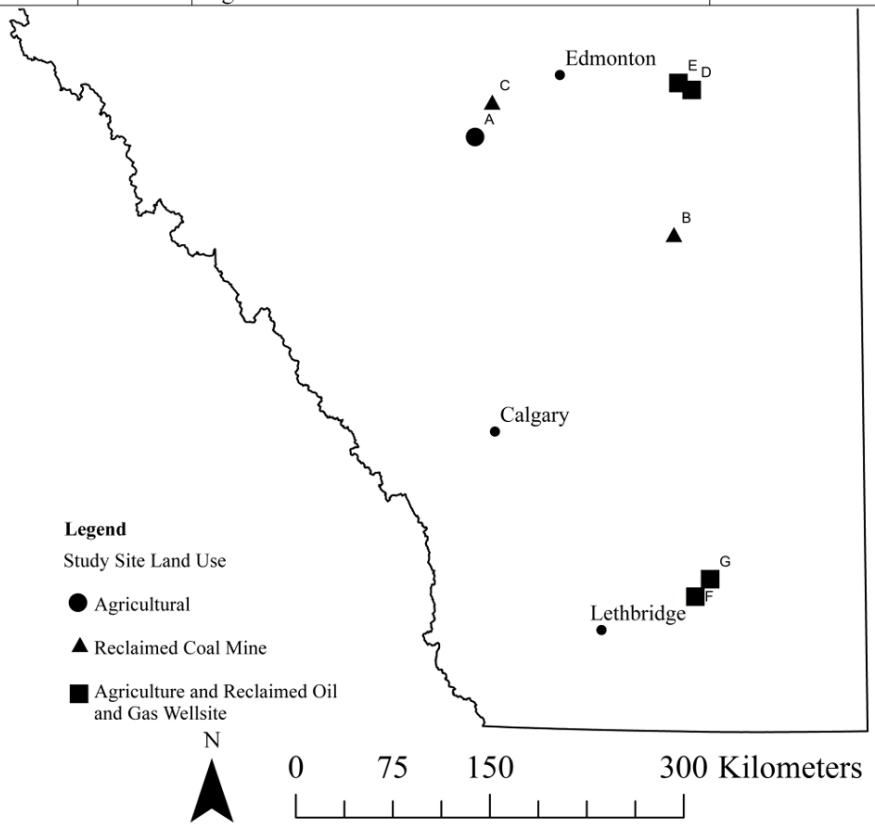
4.3. Materials and Methods

Studied Soils

Sampling locations were selected to include a range of soil types, with variations in SOC, TN and pH. Samples were collected from natural reference locations and from two common land disturbances in Alberta. Soil samples were collected from agricultural soils, reclaimed oil and gas wellsites, and reclaimed coal mines at seven sites across central and southern Alberta, Canada; 144 were collected from agricultural soil and 104 were collected from reclaimed soils (Figure 4-1). Agricultural soil samples were collected at a set of permanent agricultural plots established in 1929 that investigate farming practices near Breton, Alberta; at two separate locations near Vegreville, Alberta; and, at two locations near Taber, Alberta. Reclaimed soils were collected from two reclaimed coal mines, one near Paintearth, Alberta and the other near Edmonton, Alberta. The remaining soil samples were collected from reclaimed oil and gas wellsites located near Vegreville and Taber, Alberta.

Figure 4-1. Location of study sites sampled using the Veris ® spectrophotometer P4000 probe in Alberta, Canada. Source: North American State Province Boundaries from Esri, TomTom.

Site	Location	Samples	Land Use	Soil
A	Breton	16	Agriculture	Gray and Dark Gray Luvisols
B	Paintearth	12	Reclaimed Coal Mine	Dark Brown Solodized Solonetz
C	Edmonton	16	Reclaimed Coal Mine	Gray and Dark Gray Luvisols
D	Vegreville	72	Agriculture	Black Chernozems
E	Vegreville	72	Agriculture and Reclaimed Oil and Gas Wellsite	Black Chernozems
F	Taber	28	Agriculture and Reclaimed Oil and Gas Wellsite	Brown Chernozems
G	Taber	32	Agriculture and Reclaimed Oil and Gas Wellsite	Brown Chernozems



Spectral Measurements

Spectral measurements were collected *in-situ* using a P4000 drill rig mounted spectrometer (Veris® Technologies, Salina, Kansas). The P4000 contact probe has a sapphire window and an integrated light source. Detectors for the P4000 include a 3648 element Toshiba TCD1304AP

Linear CCD Array and a 256 element InGaAs Linear Image Sensor G9206-02. The P4000 has a response range of 350-2200 nm, 384 wavelength bands and a spectral resolution of 8 nm. Reflectance values were calculated by adjusting the soil radiance based on the radiance measurement of a black and a white reference panel provided for regular calibration of the P4000. Reference readings were taken when initializing the equipment and then every ten minutes using the integrated light source. Spectral measurements were collected every 0.15 m continuously throughout each borehole. Each borehole was completed to 1 meter below ground surface. Samples for the training datasets were collected from discrete 15 or 30 cm increments from the borehole at 3 or 4 sample locations. Samples were collected from A, B and C horizons in the profile. If samples were collected from a 30 cm interval, the two spectra were averaged for that training data point. If spectra collected at a borehole did not have a corresponding lab data point, then it was not included in the analysis.

Laboratory Analyses

Laboratory analyses were run on samples collected from the boreholes to generate a training data set. Soil samples were oven dried and then ground with a ball mill prior to laboratory analysis. Total carbon and TN were determined by dry combustion according to the methods described in Nelson and Sommers (1996) and Bremner (1996). Samples were analyzed for total carbon and nitrogen on a Elementar Vario Max CNS Analyzer (Elementar Americas Inc., Mt. Laurel, NJ). Soil samples were analyzed for pH according to the saturated paste method

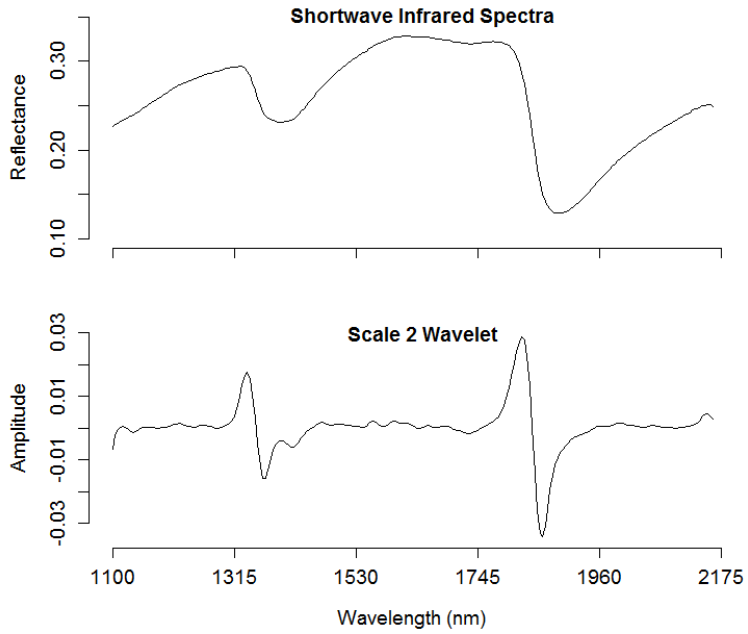
(McLean, 1982). In total, 248 samples were analyzed for total carbon (TC), 164 samples were run for TN, and 232 samples were analyzed for pH.

Statistical Analysis

Data Processing

The analysis focused on two spectral ranges: (1) 700 nm and 950 nm, and (2) 1097 nm and 2183 nm. These spectral ranges were selected as the near infrared (NIR) and shortwave infrared (SWIR) regions have spectral features associated with organic carbon. Data were trimmed at the end of the NIR and the beginning of the SWIR region due to a low signal to noise ratio associated with detector switch between the NIR and SWIR regions. The end of the SWIR region was also removed due to noise associated with being at the edge of the detector range. Each band in both of these wavelength ranges was used in the model development. Each spectrum was processed using continuous wavelet transforms (CWT) (e.g., Rivard et al., 2008; Tappert et al., 2015; Scafutto et al., 2016) using the wmtsa package in R (Percival et al., 2016). The CWT outputs were calculated for each spectral range using an eight scale second order Gaussian transform. Between 700 and 950 nm, scale four was deemed to best capture the spectral absorption features relating to composition and used to develop the statistical model, and between 1097 and 2183 nm, scale two was selected (Figure 4-2).

Figure 4-2. Example shortwave infrared spectral signature and Scale 2 Wavelet for an Ah horizon sample collected near Vegreville, Alberta. Scale 2 Wavelet is a second order Gaussian Wavelet Transform. The sample contains 4.14% soil organic carbon, 0.36% total nitrogen and has a pH of 6.6.



Model Development

The spectral data was calibrated against SOC, TN and soil pH calibration data sets. All model development was conducted using the *caret* package (Kuhn et al., 2016) in R (R Core Team, 2018), with models previously used to predict soil parameters using spectroscopic data (Doetterl et al., 2013; Minasny and McBratney, 2008; Rossel et al., 2010). Models included in the analysis are: Multivariate Adaptive Regression Splines (Milborrow, 2016), Artificial Neural Networks (Venables and Ripley, 2002), Radial Basis Support Vector Machines (Karatzoglou et al., 2004), Partial Least Squares Regression (Mevik et al., 2015), Random Forests (Liaw and Wiener, 2002), and Cubist models (Kuhn et al., 2014). All models were initially run with the default values defined by the *caret* package, with the final model parameters automatically selected by the *caret* package during model training (Kuhn et al., 2016). Model parameters are automatically selected by the *caret* package during a leave-one-out cross-validation, with the model parameters producing the smallest root mean square error (RMSE) of cross-validation being selected.

Model Evaluation

Model accuracy was evaluated based on the RMSE, R^2, and ratio of performance to deviation (RPD). The RPD value is the ratio of the standard deviation to the RMSE. Generally, RPD values greater than 2 indicate that a model can be accurately used for prediction, values between 1.4 and 2 are satisfactory, but improvement would be valuable, and values less than 1.4 indicate the model has no prediction capability (Chang et al., 2001). The best model calibration

was determined based on the model that minimized the RMSE value. Models were created using the combined reclaimed and natural soil data. The cross-validation results for the natural and reclaimed sites were separated to investigate how well the model performed in both contexts.

4.4. Results

Soil Organic Carbon (SOC)

SOC concentrations ranged from 0.25% to 6.14% with an average concentration of 1.93% in the natural background soils (Table 4-1). SOC concentrations in the reclaimed soils ranged from 0.43% to 4.40% and had an average concentration of 1.63% (Table 4-1). The Cubist model produced the lowest RMSE value (0.60%) (Figure 4-3a), followed by the random forest model (0.62%), multivariate adaptive regression splines (0.66%), support vector machines (0.67%), partial least squares regression (0.90%) and finally the artificial neural nets (1.56%) model (Table 4-2). The RPD values for the Cubist (2.2), random forest (2.1), multivariate adaptive regression splines (2.0) and support vector machines (2.0) models indicate that these models can accurately be used for prediction. The partial least squares regression (1.5) produced satisfactory RPD values, whereas the artificial neural nets (0.9) model produced results that cannot be used for accurate prediction of SOC.

Table 4-1. Soil organic carbon (SOC), total nitrogen (TN) and soil pH statistics for natural and reclaimed soils analyzed in this study. A total of 248 samples were analyzed for SOC, 164 samples for TN and 232 samples for soil pH.

Property	Natural			Reclaimed		
	Mean	Range	Std. Dev	Mean	Range	Std. Dev
Soil Organic Carbon (%)	1.93	0.25-6.14	1.49	1.63	0.43-4.40	1.06
Total Nitrogen (%)	0.21	0.04-0.54	0.13	0.15	0.04-0.40	0.10
pH	7.1	5.1-8.6	0.80	7.7	5.6-8.5	0.68

Table 4-2. Cross-validation results for Soil Organic Carbon prediction using reflectance spectroscopy data. Model results are evaluated based on the root mean square error (RMSE), R2, and the Ratio of Performance to Deviation (RPD).

Model	Cross-Validation Results		
	RMSE	R^2	RPD
Multivariate Adaptive Regression Splines	0.66	0.76	2.0
Artificial Neural Nets	1.56	0.01	0.9
Support Vector Machines	0.67	0.75	2.0
Partial Least Squares Regression	0.90	0.54	1.5
Random Forest	0.62	0.78	2.1
Cubist	0.60	0.80	2.2

Soil Nitrogen

Soil nitrogen concentrations ranged from 0.04% to 0.54% in the natural background soils and from 0.43% to 0.40% in the reclaimed soils (Table 4-1). Natural background soils on average contained 0.21% TN and the reclaimed soils had an average TN content of 0.15% (Table 4-3). The Cubist model produced the lowest RMSE value (0.05%) (Figure 4-3b), followed by multivariate adaptive regression splines (0.06%), random forest (0.06%), support vector machines (0.06%), partial least squares regression (0.07%), and artificial neural nets model (0.12%). The RPD values for the Cubist (2.5), multivariate adaptive regression splines (2.1), random forest (2.1) and support vector machines (2.1) models indicate that they can accurately be used for prediction. The partial least squares regression (1.8) model was found to satisfactory based on the RPD value. However, the artificial neural nets (1.0) model did not accurately predict TN.

Table 4-3. Cross-validation results for Total Nitrogen prediction using reflectance spectroscopy data. Model results are evaluated based on the root mean square error (RMSE), R2, and the Ratio of Performance to Deviation (RPD).

Model	**Cross-Validation Results**		
	RMSE	**R^2**	**RPD**
Multivariate Adaptive Regression Splines	0.06	0.75	2.1
Artificial Neural Nets	0.12	0.20	1.0
Support Vector Machines	0.06	0.78	2.1
Partial Least Squares Regression	0.07	0.67	1.8
Random Forest	0.06	0.78	2.1
Cubist	0.05	0.81	2.5

Soil pH

Soil pH in the natural background soils ranged from 5.1 to 8.6, with an average pH of 7.1 (Table 4-1). The pH in the reclaimed soils ranged from 5.6 to 8.5 with an average pH of 7.7. The Cubist model produced the lowest RMSE value (0.44) (Figure 4-3c), followed by random forest (0.47) and support vector machines model (0.51). Multivariate adaptive regression splines were next (0.54), followed by partial least squares regression (0.68) and artificial neural networks (6.39). No model had an RPD value above 2.0. The cubist (1.8), random forest (1.7), support vector machines (1.6) and multivariate adaptive regression splines models (1.5) had RPD values greater than 1.4 indicating that they could satisfactorily be used for prediction. The partial least squares regression (1.2) and artificial neural nets (0.1) models could not be successfully used for prediction of soil pH.

Table 4-4. Cross-validation results for soil pH prediction using reflectance spectroscopy data. Model results are evaluated based on the root mean square error (RMSE), R2, and the Ratio of Performance to Deviation (RPD).

Model	**Cross-Validation Results**		
	RMSE	**R^2**	**RPD**
Multivariate Adaptive Regression Splines	0.54	0.58	1.5
Artificial Neural Nets	6.39	0.01	0.1
Support Vector Machines	0.51	0.60	1.6
Partial Least Squares Regression	0.68	0.31	1.2
Random Forest	0.47	0.67	1.7
Cubist	0.44	0.69	1.8

Prediction accuracy for natural and reclaimed soils

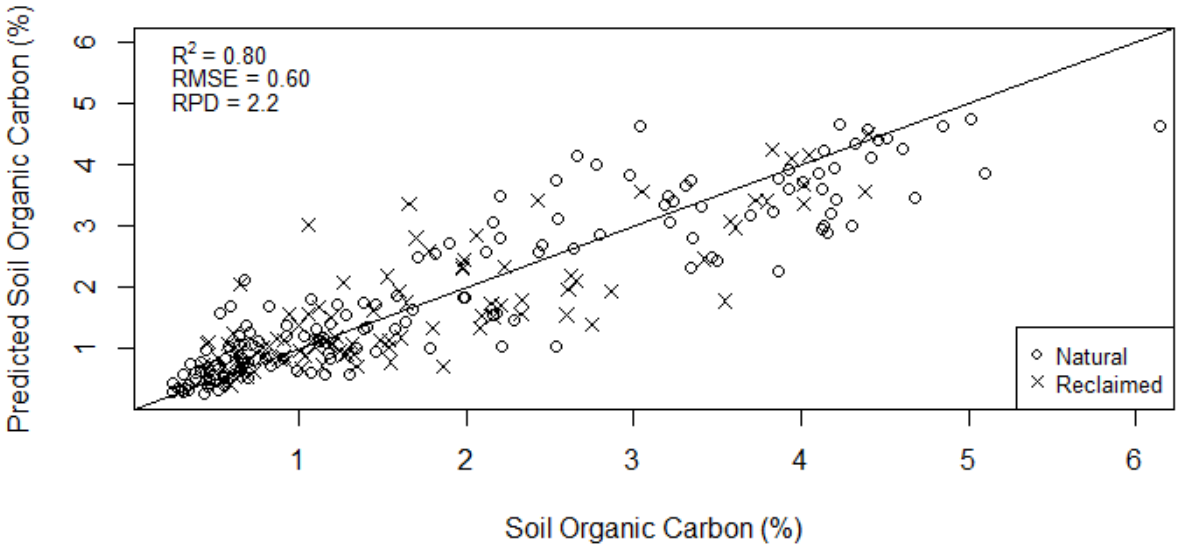
The accuracy of the model predictions varied slightly between the natural background and reclaimed soils for SOC, TN and pH values (Table 4-5). While the RMSE value for natural soil carbon (0.60) was slightly higher than the RMSE value for reclaimed soil carbon (0.59), the predicative capability for carbon on natural soils was slightly higher when compared to reclaimed soils. The same pattern was also observed for nitrogen. Soil pH model results for reclaimed soils had a lower RMSE value and a higher RPD compared to the natural soils.

Table 4-5. Cross-validation results from the CUBIST model for samples collected from natural and reclaimed soils. Model results are evaluated based on the root mean square error (RMSE), R2, and the Ratio of Performance to Deviation (RPD).

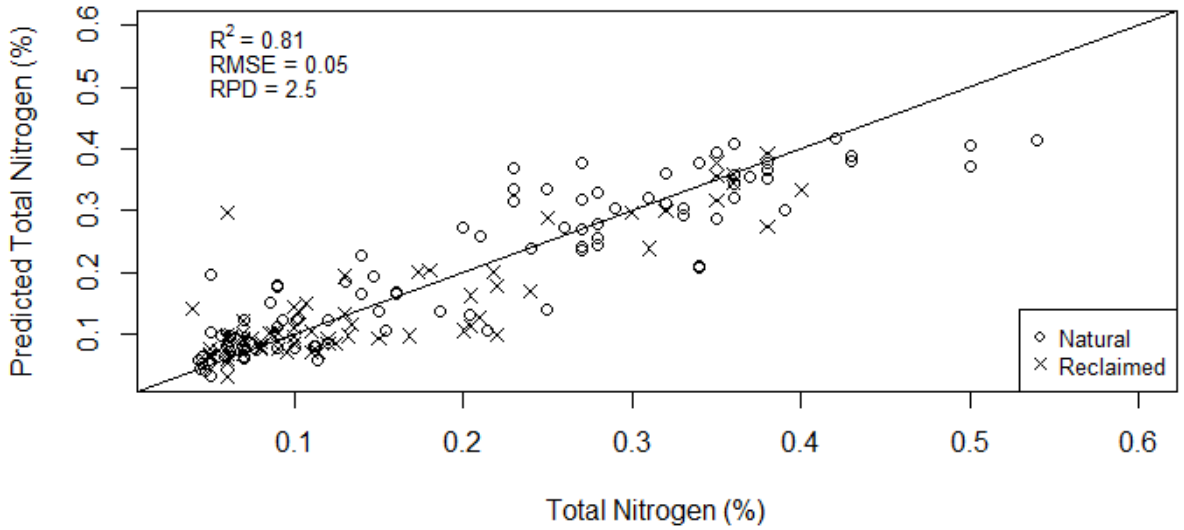
Parameter	**CUBIST Model Cross-Validation Results**		
	RMSE	**R^2**	**RPD**
Natural Soil Carbon	0.60	0.84	2.5
Reclaimed Soil Carbon	0.59	0.70	1.8
Natural Soil Nitrogen	0.06	0.82	2.3
Reclaimed Soil Nitrogen	0.05	0.76	2.0
Natural Soil pH	0.48	0.62	1.6
Reclaimed Soil pH	0.39	0.68	1.8

Figure 4-3. Cross-validated predicted versus observed soil values for (a) soil organic carbon, (b) total nitrogen, and (c) soil pH for both natural and reclaimed soils as produced by the Cubist model. The solid line indicates the 1:1 line to illustrate deviations between predicted and measured data.

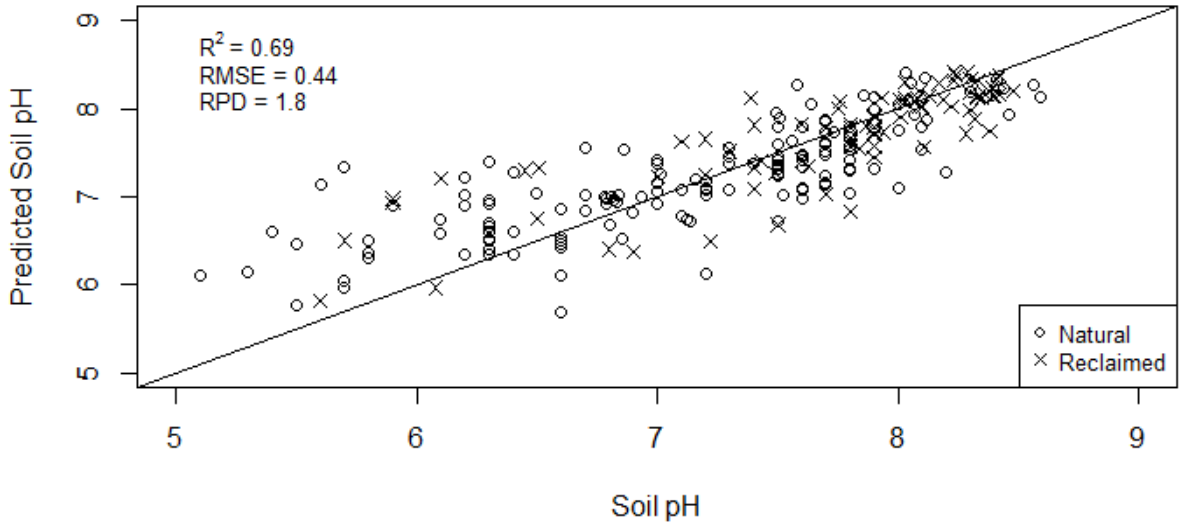
(a)



(b)


R^2 = 0.81
RMSE = 0.05
RPD = 2.5
Natural
Reclaimed
Predicted Total Nitrogen (%)
Total Nitrogen (%)
0.1
0.2
0.3
0.4
0.5
0.6

(c)


R^2 = 0.69
RMSE = 0.44
RPD = 1.8
Natural
Reclaimed
Predicted Soil pH
Soil pH
5
6
7
8
9

4.5. Discussion

Overall, our results indicate that there is potential for using field spectroscopy to monitor reclaimed soil properties. Specifically, our models could be used to predict SOC and TN concentrations from field collected spectra based on the criteria outlined in Chang et al. (2001). For soil pH, the model is adequate for prediction at pH values above 6.5, but improvements are required to accurately predict lower pH values as there was consistent over prediction of pH at values less than 6.5 (Figure 4-3c). The pH model may be improved by using a spectrometer that collects spectral data to 2500 nm, as key carbonate features are found near 2336 nm (Ben-Dor and Banin, 1990).

For our study, model results were comparable for natural and reclaimed soils for TN and soil pH (Table 4-5). The reclaimed SOC samples had a lower RPD value than the natural samples, and the success of the model on reclaimed soil samples could use improvement. Modelling results can likely be improved by obtaining more data from reclaimed sites at SOC concentrations above 4 percent, as the reclaimed samples had less data at these concentrations.

The results of the Cubist models for predicting SOC and TN under field conditions were satisfactory and compared favourably to results from studies using dried and homogenized soil samples in the laboratory (Chang et al., 2001; Martin et al., 2002; Xie et al., 2011). Our results compare favourably to a study in Australia that used a portable spectrometer to measure SOC contents in the field, as we were able to obtain a RPD value of 2.2 compared to 1.92 in the study in Australia (Gomez et al., 2008b). An important point about predictive modelling of reflectance

spectroscopy data is that a model developed with more homogenous training data sets tend to be more accurate with lower RMSE values, but also less transferable to new sites. Therefore, it is important to consider the RPD value when assessing results and not just the RMSE. Overall, this study using a drilling rig mounted spectrophotometer probe had comparable results to other laboratory and field spectroscopy studies.

Additional work is needed to investigate methods that improve modelling results using field collected spectra. Variation in soil moisture in the field can have a substantial effect on overall reflectance. Other researchers have successful used external parameter orthogolonization (EPO) to correct spectra based on soil moisture variation (Minasny et al., 2009). Initially developed to remove the effects of temperature when analyzing fruit for sugar contents with reflectance spectroscopy (Roger et al., 2003), EPO corrects the spectra based on spectral areas affected by soil moisture and then projects the spectra orthogonal to the soil moisture variation (Minasny et al., 2009). Further work is needed to compare results using low scale wavelets and EPO corrections to remove the effects of soil moisture from reflectance spectra. Modelling results may also be improved by using a detector that allows data collection up to 2500 nm, as SOC has important spectral features between 2200 and 2500 nm (Rossel et al., 2010).

Improving the system to simultaneously measure carbonates, salinity and bulk density would increase the benefit of these tools through the dual measurement of soil chemical and physical parameters in the field. The measurement of carbonates using reflectance spectroscopy

is well established (Ben-Dor and Banin, 1990; Gomez et al., 2012; Lagacherie et al., 2008), and does not represent a significant technical challenge. This parameter could help inform reclamation practitioners of admixing, since the presence of carbonate-rich material from parent materials is common in reconstructed soils in the Canadian Prairies (Anderson and Cerkowniak, 2010). The Veris ® P4000 Spectrophotometer Probe has dipole EC contacts and a load cell force sensor integrated into the probe, suitable for measuring soil bulk density. Determining how accurately dipole EC contacts and load cell force sensors can also be used to measure soil salinity and bulk density *in situ* is necessary before deploying this system for reclamation monitoring.

Acknowledgements

We would like to thank Alberta Environment and Parks (formally the Alberta Environmental Monitoring and Evaluation Agency (AEMERA)) and Alberta Innovates-Technology Futures for providing financial assistance, soil samples and soil proximal soil data for the project. We would also like to acknowledge the National Science and Engineering Council (NSERC) for providing financial support for this project with a Discovery Grant to S.A. Quideau (no 462352-2014).

4.6. References

Alberta Energy Regulator, 2016. Statistical Reports [WWW Document]. URL https://www.aer.ca/data-and-publications/statistical-reports (accessed 5.18.16).

Alberta Energy Regulator, 2011. Directive for Monitoring the Impact of Sulphur Dust on Soils. Edmonton, Alberta.

Anderson, D., Cerkowniak, D., 2010. Soil Formation in the Canadian Prairie Region. Prairie Soils Crop. 3, 57–64.

Bartholomeus, H.M., Schaepman, M.E., Kooistra, L., Stevens, A., Hoogmoed, W.B., Spaargaren, O.S.P., 2008. Spectral reflectance based indices for soil organic carbon quantification. Geoderma. https://doi.org/10.1016/j.geoderma.2008.01.010

Ben-Dor, E., Banin, A., 1995. Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj1995.03615995005900020014x

Ben-Dor, E., Banin, A., 1990. Near-infrared reflectance analysis of carbonate concentration in soils. Appl. Spectrosc. https://doi.org/10.1366/0003702904086821

Berg, B., 2000. Litter decomposition and organic matter turnover in northern forest soils. For. Ecol. Manage. 133, 13–22. https://doi.org/http://dx.doi.org/10.1016/S0378-1127(99)00294-7

Bingham, A.H., Cotrufo, M.F., 2016. Organic nitrogen storage in mineral soil: Implications for policy and management. Sci. Total Environ. 551–552, 116–126. https://doi.org/10.1016/j.scitotenv.2016.02.020

Blake, L., Goulding, K.W.T., 2002. Effects of atmospheric deposition, soil pH and acidification on heavy metal contents in soils and vegetation of semi-natural ecosystems at Rothamsted Experimental Station, UK. Plant Soil. https://doi.org/10.1023/A:1015731530498

Bremner, J.M., 1996. Nitrogen - Total, in: Sparks, D.L., Page, A.., Helmke, P.A., Loeppert, R.H., Soltanpour, P.N., Tabatabai, M.A., Johnston, C.T., Sumner, M.E. (Eds.), Methods of Soil Analysis Part 3—Chemical Methods. Madison, WI, pp. 1085–1123.

Chang, C.-W., Laird, D.A., 2002. NEAR-INFRARED REFLECTANCE SPECTROSCOPIC ANALYSIS OF SOIL C AND N. Soil Sci. https://doi.org/10.1097/00010694-200202000-00003

Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2001.652480x

Dessureault-Rompré, J., Zebarth, B.J., Burton, D.L., Georgallas, A., 2015. Predicting soil nitrogen supply from soil properties. Can. J. Soil Sci. https://doi.org/10.4141/cjss-2014-057

Doetterl, S., Stevens, A., Van Oost, K., van Wesemael, B., 2013. Soil Organic Carbon Assessment at High Vertical Resolution using Closed-Tube Sampling and Vis-NIR Spectroscopy. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2012.0410n

Ge, Y., Morgan, C.L.S., Ackerson, J.P., 2014. VisNIR spectra of dried ground soils predict properties of soils scanned moist and intact. Geoderma. https://doi.org/10.1016/j.geoderma.2014.01.011

Gomez, C., Lagacherie, P., Coulouma, G., 2012. Regional predictions of eight common soil properties and their spatial structures from hyperspectral Vis-NIR data. Geoderma. https://doi.org/10.1016/j.geoderma.2012.05.023

Gomez, C., Lagacherie, P., Coulouma, G., 2008a. Continuum removal versus PLSR method for clay and calcium carbonate content estimation from laboratory and airborne hyperspectral measurements. Geoderma. https://doi.org/10.1016/j.geoderma.2008.09.016

Gomez, C., Viscarra Rossel, R.A., McBratney, A.B., 2008b. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. Geoderma. https://doi.org/10.1016/j.geoderma.2008.06.011

He, Y., Huang, M., García, A., Hernández, A., Song, H., 2007. Prediction of soil macronutrients content using near-infrared spectroscopy. Comput. Electron. Agric. https://doi.org/10.1016/j.compag.2007.03.011

Hijmans, R.J., 2016. raster: Geographic Data Analysis and Modeling.

Indorante, S.J., Jansen, I.J., Boast, C.W., 1981. Surface mining and reclamation : Initial changes in soil character. J. Soil Water Conserv. 36, 347–351.

IUSS Working Group WRB, 2014. World reference base for soil resources 2014. International soil classification system for naming soils and creating legends for soil maps, World Soil Resources Reports No. 106. https://doi.org/10.1017/S0014479706394902

Jastrow, J.D., 1996. PIk soo38-0717(!spo159-x SOIL AGGREGATE FORMATION AND THE ACCRUAL OF PARTICULATE AND MINERAL-ASSOCIATED ORGANIC MATTER. Soil Biol. Biochem 28, 665–676.

Jobbágy, E.G., Jackson, R.B., 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecol. Appl. https://doi.org/10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2

Karatzoglou, A., Smola, A., Hornik, K., Seileis, A., 2004. kernlab - An S4 Package for Kernal Methods in R. J. Stat. Softw. 11, 1–20.

Kinoshita, R., Moebius-Clune, B.N., van Es, H.M., Hively, W.D., Bilgilis, A.V., 2012. Strategies for Soil Quality Assessment Using Visible and Near-Infrared Reflectance Spectroscopy

in a Western Kenya Chronosequence. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2011.0307

Kuhn, M., Johnson, K., n.d. Applied Predictive Modeling.

Kuhn, M., Weston, S., Keefer, C., Coulter, N., 2014. Cubist: Rule- and Instance-Based Regression Modeling.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., R Core Team, Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., 2016. caret: Classification and Regression Training.

Lagacherie, P., Baret, F., Feret, J.B., Madeira Netto, J., Robbez-Masson, J.M., 2008. Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2007.06.014

LeBauer, D.S., Treseder, K.K., 2008. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. Ecology 89, 371–379. https://doi.org/10.1890/06-2057.1

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 2, 18–22.

Martin, P.D., Malley, D.F., Manning, G., Fuller, L., 2002. Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy. Can. J. Soil Sci. https://doi.org/10.4141/S01-054

McBratney, A.B., Minasny, B., Viscarra Rossel, R., 2006. Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. Geoderma. https://doi.org/10.1016/j.geoderma.2006.03.051

McCarty, G.W., Reeves, J.B., Reeves, V.B., Follett, R.F., Kimble, J.M., 2002. Mid-Infrared and Near-Infrared Diffuse Reflectance Spectroscopy for Soil Carbon Measurement. Soil Sci. Soc. Am. J. 66, 640. https://doi.org/10.2136/sssaj2002.0640

McIntire, E.J.B., Fajardo, A., 2009. Beyond description: the active and effective way to infer processes from spatial patterns 90, 46–56.

McLean, E.O., 1982. Soil pH and Lime Requirement, in: Page, A.., Miller, R.H., Keeney, D.R. (Eds.), Methods of Soil Analysis Part 2, Chemical and Microbiological Properties. Madison, WI, pp. 199–225.

Melendez-Pastor, I., Navarro-Pedreño, J., Koch, M., Gómez, I., 2010. Applying imaging spectroscopy techniques to map saline soils with ASTER images. Geoderma. https://doi.org/10.1016/j.geoderma.2010.02.015

Mevik, B.H., Wehrens, R., Liland, K.H., 2015. pls: Partial Least Squares and Principal Component Regression.

Milborrow, S., 2016. earth: Multivariate Adaptive Regression Splines.

Minasny, B., McBratney, a B., Pichon, L., Sun, W., Short, M.G., 2009. Evaluating near infrared spectroscopy for field prediction of soil properties. Aust. J. Soil Res. 47, 664–673. https://doi.org/Doi 10.1071/Sr09005

Minasny, B., McBratney, A.B., 2008. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. Chemom. Intell. Lab. Syst. https://doi.org/10.1016/j.chemolab.2008.06.003

Morellos, A., Pantazi, X.-E., Moshou, D., Alexandridis, T., Whetton, R., Tziotzios, G., Wiebensohn, J., Bill, R., Mouazen, A.M., 2016. Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. Biosyst. Eng. https://doi.org/10.1016/j.biosystemseng.2016.04.018

Nawar, S., Buddenbaum, H., Hill, J., Kozak, J., Mouazen, A.M., 2016. Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy. Soil Tillage Res. https://doi.org/10.1016/j.still.2015.07.021

Nelson, D.W., Sommers, L.E., 1996. Total Carbon, Organic Carbon, and Organic Matter, in: Sparks, D.L., Page, A.., Helmke, P.A., Loeppert, R.H., Soltanpour, P.N., Tabatabai, M.A., Johnston, C.T., Sumner, M.E. (Eds.), Methods of Soil Analysis Part 3—Chemical Methods. Soil Science Society of America, Madison, WI, pp. 961–1011.

Ouerghemmi, W., Gomez, C., Naceur, S., Lagacherie, P., 2011. Applying blind source separation on hyperspectral data for clay content estimation over partially vegetated surfaces. Geoderma. https://doi.org/10.1016/j.geoderma.2011.04.019

Percival, D.B., Walden, A.T., William, A., Percival, D., Constantine, M.W., 2016. Wavelet Methods for Time Series Analysis.

Pinheiro, J., Bates, D., Debroy, S., Sarkar, D., R Core Team, 2016. _nlme: Linear and Nonliner Mixed Effects Models_.

R Core Team, 2018. R: A language and environment for statistical computing.

Reeves, D.W., 1997. The role of soil organic matter in maintaining soil quality in continuous cropping systems. soil& Tillage Res. Soil Tillage Res. Reeues/Soil Tillage Res. 43, 131–167.

Rivard, B., Feng, J., Gallie, A., Sanchez-Azofeifa, A., 2008. Continuous wavelets for the improved use of spectral libraries and hyperspectral data. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2008.01.016

Roger, J.M., Chauchard, F., Bellon-Maurel, V., 2003. EPO-PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. Chemom. Intell. Lab. Syst. 66, 191–204. https://doi.org/10.1016/S0169-7439(03)00051-0

Rossel, R.A.A.V., Behrens, T., Viscarra Rossel, R.A., Behrens, T., 2010. Using data mining to model and interpret soil diffuse reflectance spectra. Geoderma 158, 46–54. https://doi.org/10.1016/j.geoderma.2009.12.025

Scafutto, R.D.P.M., de Souza Filho, C.R., Rivard, B., 2016. Characterization of mineral substrates impregnated with crude oils using proximal infrared hyperspectral imaging. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2016.03.033

Shrestha, R.K., Lal, R., 2007. Soil Carbon and Nitrogen in 28-Year-Old Land Uses in Reclaimed Coal Mine Soils of Ohio. J. Environ. Qual. https://doi.org/10.2134/jeq2007.0071

Soil Classification Working Group, 1998. The Canadian System of Soil Classification. Can. Syst. Soil Classif. 3rd ed. Agric. Agri-Food Canada Publ. 1646 187.

Sorenson, P.T., MacKenzie, M.D., Quideau, S.A., Landhausser, S.M., 2017. Can spatial patterns be used to investigate aboveground-belowground links in reclaimed forests? Ecol. Eng. 104, 57–66. https://doi.org/10.1016/j.ecoleng.2017.04.002

Sorenson, P.T., Quideau, S.A., MacKenzie, M.D., Landhäusser, S.M., Oh, S.W., 2011. Forest floor development and biochemical properties in reconstructed boreal forest soils. Appl. Soil Ecol. https://doi.org/10.1016/j.apsoil.2011.06.006

Sorenson, P.T., Small, C., Tappert, M.C., Quideau, S.A., Drozdowski, B., Underwood, A., Janz, A., 2017. Monitoring organic carbon, total nitrogen, and pH for reclaimed soils using field reflectance spectroscopy. Can. J. Soil Sci. https://doi.org/10.1139/cjss-2016-0116

St. Luce, M., Ziadi, N., Zebarth, B.J., Grant, C.A., Tremblay, G.F., Gregorich, E.G., 2014. Rapid determination of soil organic matter quality indicators using visible near infrared reflectance spectroscopy. Geoderma. https://doi.org/10.1016/j.geoderma.2014.05.023

Steffens, M., Buddenbaum, H., 2013. Laboratory imaging spectroscopy of a stagnic Luvisol profile - High resolution soil characterisation, classification and mapping of elemental concentrations. Geoderma. https://doi.org/10.1016/j.geoderma.2012.11.011

Strong, W.L., La Roi, G.H., 1983. Root-system morphology of common boreal forest trees in Alberta, Canada. Can. J. For. Res. 13, 1164–1173.

Tappert, M.C., Rivard, B., Fulop, A., Rogge, D., Feng, J., Tappert, R., Stalder, R., 2015. Characterizing kimberlite dilution by crustal rocks at the Snap Lake diamond mine (Northwest Territories, Canada) using SWIR (1.90-2.36 μm) and LWIR (8.1-11.1 μm) hyperspectral imagery collected from drill core. Econ. Geol. https://doi.org/10.2113/econgeo.110.6.1375

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics, Fourth. ed. Springer, New York.

Viscarra Rossel, R.A., Lark, R.M., 2009. Improved analysis and modelling of soil diffuse reflectance spectra using wavelets. Eur. J. Soil Sci. https://doi.org/10.1111/j.1365-2389.2009.01121.x

Vitousek, P.M., Howarth, R.W., 1991. Nitrogen Limitation on Land and in the Sea : How Can It Occur ? Nitrogen limitation on land and in the sea : How can it occur ? 13, 87–115. https://doi.org/10.1007/BF00002772

Wang, S., Zhuang, Q., Wang, Q., Jin, X., Han, C., 2017. Mapping stocks of soil organic carbon and soil total nitrogen in Liaoning Province of China. Geoderma. https://doi.org/10.1016/j.geoderma.2017.05.048

Xie, H.T., Yang, X.M., Drury, C.F., Yang, J.Y., Zhang, X.D., 2011. Predicting soil organic carbon and total nitrogen using mid- and near-infrared spectra for Brookston clay loam soil in Southwestern Ontario, Canada. Can. J. Soil Sci. https://doi.org/10.4141/cjss10029

5. Assessment of reclaimed soils by unsupervised clustering of proximal sensor data

5.1. Abstract

The application of soil proximal sensors on reclaimed sites presents a novel method for assessing the quality of reclaimed landscapes; improving assessment reliability, information management, and environmental assurance. One proximal sensing system that could be used to provide high spatial resolution measurements of soil parameters is an on-the-go optical sensor that collects data at two wavelengths: 660 nm and 940 nm. Proximal soil sensing data were collected at 27 sites, where organic matter, cation exchange capacity and soil water content were collected from 221 soil samples from 0 to 15 cm. The proximal soil sensor data were then automatically clustered using a combination of self-organizing maps and random uniform forests. Overall, the proximal sensor data combined with this data analysis approach created maps with either three or four soil zones. On average, soil zones had statistically significant differences in organic matter, cation exchange capacity and water content. This system could be used to map out zones with significant soil variation as part of reclamation monitoring, and then used to guide laboratory analytical sampling. Future work should focus on development of on-the-go reflectance spectroscopy systems to provide quantitative soil data with high spatial resolution.

Key Words: Proximal soil sensing; soil organic matter; unsupervised classification

5.2. Introduction

Evaluation and assessment of disturbed soils is required for ensuring soil quality for both environmental management and reclamation. The energy industry in Alberta is responsible for soil disturbance at over 470 000 wellsites, 39 000 other oil and gas facilities and 24 open pit coal mines (Alberta Energy Regulator, 2016b). These primary facilities have additional soil disturbances because of associated development activities such as borrow pits, transmission lines and access roads. As per the regulatory requirements in Alberta, all sites and associated facilities must be eventually reclaimed and restored to a state of equivalent land capability (Government of Alberta, 2010). Heightened expectations from stakeholder groups and the public, regarding the need for more timely and comprehensive information on the status of provincial reclamation activities, is pressuring the government and industry to pursue new technologies and applications for acquiring information on the status of reclaimed sites.

A potential solution to the need for more timely and spatially comprehensive information is the use of soil proximal sensors. The potential of reflectance spectroscopy to measure key soil parameters such as organic carbon and total nitrogen has been extensively investigated within the broader context of soil analysis (i.e. Ben-Dor and Banin 1995; Chang et al. 2001; Chang and Laird 2002; Rossel and Behrens 2010; Sorenson et al. 2018) and more specifically for reclamation monitoring (P.T. Sorenson et al., 2017). Additionally, simpler lower cost two-band reflectance sensors have been explored for potential use in precision agriculture

for mapping changes in soil organic matter contents at high spatial resolutions (Kweon et al., 2013; Piikki et al., 2016).

Determining soil organic carbon is part of wellsite reclamation assessment procedures in Alberta (Alberta Environment, 2010). Previous work in Alberta has identified that wellsite construction can be associated with a decrease in soil organic carbon and overall changes in soil properties (Hammermeister et al., 2003). Soil bulk density has been found to be higher on a number of reclaimed wellsites, which in turn affects infiltration and soil water content (Hammermeister et al., 2003). Soil reclamation strategies on oil and gas wellsites often focus on increasing soil porosity and water holding capacity (McConkey et al., 2012). Cation exchange capacity (CEC) is closely related to organic matter content in soil, and both parameters are closely tied to soil fertility (Parfitt et al., 1995). Therefore, all three of these parameters can be affected by construction activities, and the ability to detect changes in these parameters is important for monitoring if a site is moving toward reclamation success.

Soil proximal sensing generates large volumes of data, which necessitates the development of techniques to process this information into a format that facilitates decision making. While building site-specific calibration models with two-band optical reflectance sensors has met some degree of success, universal calibration equations have not been successfully developed so far (Kweon et al., 2013). Rather than quantitative analysis, two-band reflectance data could potentially be used for qualitative measurements using unsupervised machine learning tools. One such technique for unsupervised data analysis and dimensionality

reduction is self-organizing maps (Wehrens and Buydens, 2007). Self-organizing maps are similar to multi-dimensional scaling, but rather than trying to preserve relative distance between points, self-organizing maps focus on preserving the topology of the data structure, and concentrate more on mapping similarity rather than dissimilarity (Wehrens and Buydens, 2007). Self-organizing maps have been used successfully to solve other challenges associated with complex soil data analysis, such as the assessment of soil biological quality (Mele and Crowley, 2008).

The main objective of this research was to determine if proximal sensing data collected with a two-band reflectance sensor with bands in the red and near-infrared could be used to map relative differences in key soil attributes. While the spectral features for organic matter and water are strongest in the short-wave infrared region, some features are present in near infrared region and increases in both properties decrease overall reflectance (Rossel et al., 2010). Specifically, this study focused on combining two unsupervised machine learning methods, self-organizing maps and random uniform forests, to classify soil proximal sensing data. The overall goal of this research is to validate automated spatial zoning of soil data to support reclamation assessments.

5.3. Materials and Methods

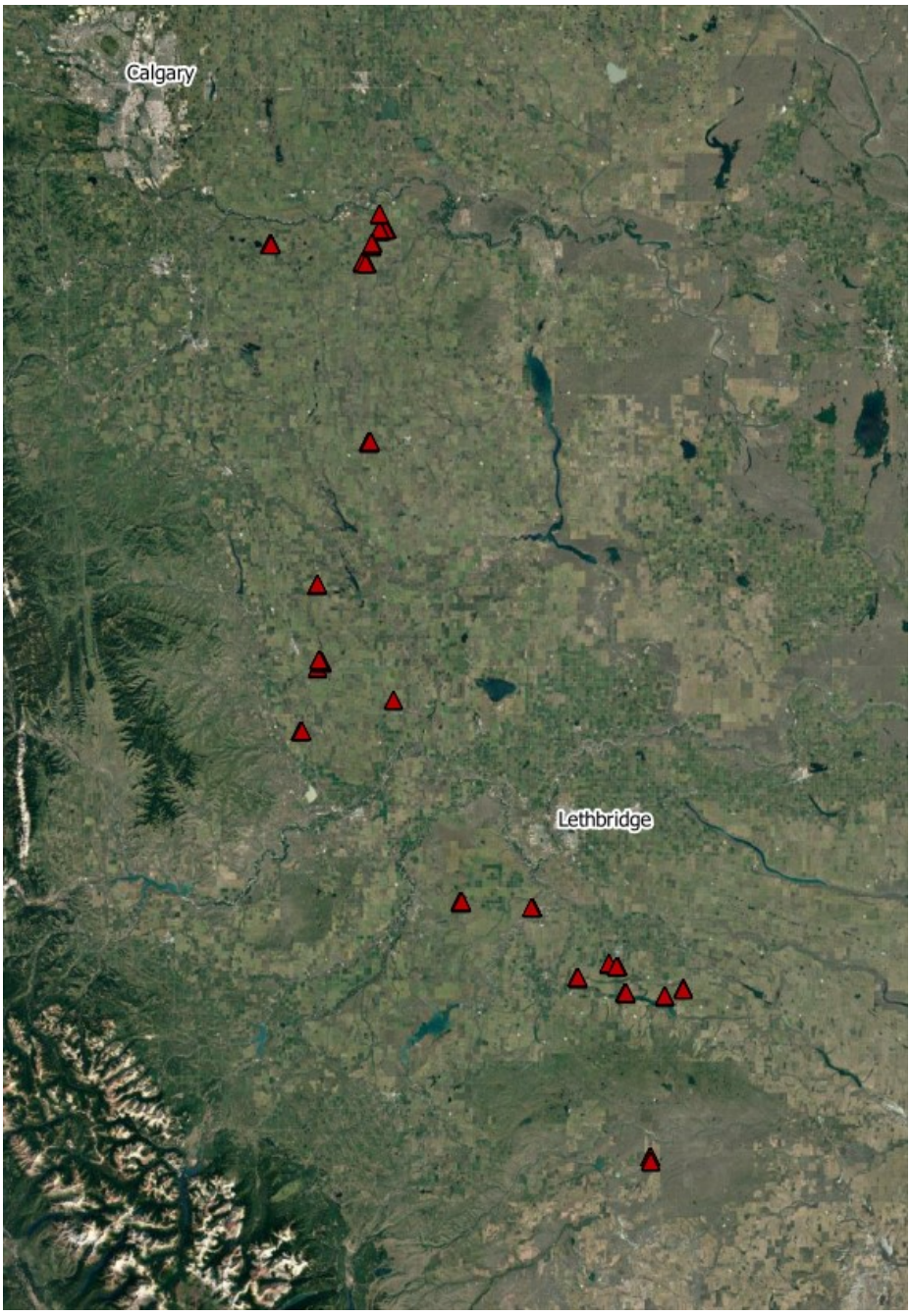
Studied Soils

Sampling locations included 27 sites (Table 5-1) at locations ranging from southeast of Calgary, Alberta to south of Lethbridge, Alberta (Figure 5-1). All sites contained disturbed and undisturbed soils as they consisted of reclaimed oil and gas wellsites, with the center of each assessed area consisting of the wellsite. An area surrounding the wellsite, which included undisturbed agriculture soils, was also included in the assessment area. At each site, 5 to 10 samples were collected for laboratory analyses. In total, 221 samples were collected from the 27 sites. Each site has samples collected from disturbed and undisturbed soils.

Table 5-1. Mean values of soil and site characteristics, with the range of values observed across all sites indicated in parentheses.

Sites	**Organic matter content (%)**	**Cation exchange capacity (meq/100g)**	**Gravimetric water content (%)**	**Area (ha)**	**Slope (%)**	**Topographic Position Index**
27	4.54 (2.28 – 8.27)	27.9 (1.24 – 71.76)	18 (8 – 29)	14.4 (7.8 – 27)	2 (0 – 8)	0 (0.67 – 0.68)

igure 5-1. Location of study sites sampled using the Veris ® Optic Mapper in Alberta, Canada. Research site locations are indicated with the red triangles. Source: © 2018 Google Imagery.



Proximal sensing data

Proximal soil sensing data were acquired using the Veris Technologies OpticMapper®, which consists of a two-band reflectance sensor along with a GPS unit to collect location and elevation data. Reflectance data were collected by the sensor in two bands, within the red portion of the electromagnetic spectrum from 650 nm to 670 nm and within the near infrared region from 930 nm to 950 nm. A fluted coulter on the optic mapper cuts through crop residues and opens a slot in the soil where measurements take place. All measurements take place on the surface of the exposed soil in the slot. Data were collected in northeast to southwest passes spaced 10 m across the site. Calibration checks for the instrument were performed as per manufacturer specifications. Light and dark reference panels were used to check sensor performance prior to data collection and sensor response was in accordance to manufacturer specifications for each site prior to data collection.

Laboratory Analyses

Each of the 221 soil samples collected was analyzed for organic matter content and for CEC; 179 of the samples were additionally analyzed for gravimetric water content. Gravimetric water content was analyzed rather than volumetric water content, as previous reflectance spectroscopy research has demonstrated that gravimetric water content can be more accurately measured than volumetric water content with reflectance-based measurement methods (Ji et al., 2016). Samples were collected the same day the proximal sensor measurements were taken. All soil samples were collected from a depth of 0 to 15 cm to correspond to the range of the optical

measurements that were collected, based on the depth of the furrow created by the fluted coulter. Organic matter content was determined by the loss on ignition method described in Nelson and Sommers (1996). CEC was quantified using the methods described in Hendershot et al. (2008), and gravimetric water content was determined using the method described in Topp et al. (2008).

Statistical Analysis

Data Processing

All data processing was performed using R (R Core Team, 2018). The raw reflectance data and elevation point data were converted to rasters using the inverse distance weighting function in the gstat package of R (Pebesma, 2004). The reflectance and elevation data were then smoothed using a 3x3 focal median window to reduce noise using the raster package in R (Hijmans, 2016). Following smoothing, the ratio of near infrared to red reflectance was calculated. Slope and topographic position index values were calculated from the elevation data using the terrain function in the raster package of R (Hijmans, 2016).

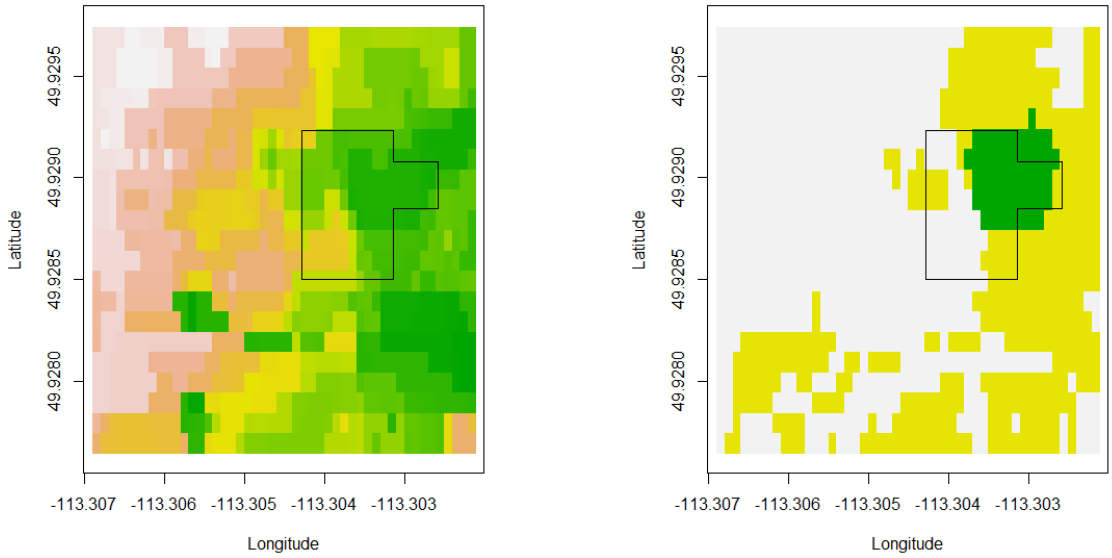
Model Development

For each site, a raster stack was created with a raster for each of the following parameters: red reflectance, near infrared reflectance, ratio of red to near infrared reflectance, elevation, slope, and topographic position index. The raster stack was first processed using a self-organizing map with the kohonen package in R (Wehrens and Buydens, 2007). A 10 by 10 hexagonal grid topology was specified for the self organizing map. Following initial clustering

of the data into the self organizing map grid, final clustering into a smaller number of clusters was performed using an unsupervised random uniform forest in R (Ciss, 2015). The advantage of this analysis is that it can be used for unsupervised clustering of data. Additionally, random uniform forests will identify the optimal number of clusters for the data without user specification of cluster numbers.

Random uniform forests uses a multiple step approach to perform the unsupervised clustering. The random uniform forest first grows a forest of decision trees using random subsampling and random cut points according to a continuous uniform distribution, followed by multidimensional scaling and clustering with either k-means or hierarchical clustering (Ciss, 2015). In this study the hierarchical clustering step was used, and the optimal number of clusters was automatically selected based on where the maximum lagged difference in cluster heights occurred. The unsupervised clustering using random forests was performed on the self-organizing map grid cell codes, leading to a final cluster number for each grid cell and each associated raster cell in each grid cell. An example of the clustering results from a site are illustrated in Figure 5-2.

Figure 5-2. Results from the self organizing map and random uniform forest. The map on the left illustrates the results from the self organizing map. The image on the right illustrates the results of the random uniform forest clustering of the self organizing map results. The black polygon within each image illustrates the approximated area disturbed by construction activities.



Model Evaluation

Model performance was evaluated on the unsupervised classification results using a linear mixed effects model with the nlme package in R (Pinheiro et al., 2016). For each site, the cluster that each laboratory analytical sample was collected from was determined. The clusters were then relabelled from lowest to highest to correspond to the highest to lowest organic matter concentrations, CEC or gravimetric water content as the specific number sequence is randomized during the classification process. This step was necessary to allow for the comparison of relative effect differences among clusters across all sites. Organic matter, CEC

and gravimetric water content data were then scaled by subtracting the mean and dividing by standard deviation. This transformation permitted comparison between sites regardless of the differences in magnitude across sites. Mean centered data were then compiled and analyzed using a linear mixed effects model with site as random factor. The data were normally distributed and showed homogeneity of variance. Linear mixed effects models were run for organic matter, CEC, and gravimetric water content to test if significant differences were present among clusters.

5.4. Results and Discussion

The clustering exercise yielded three clusters for most sites (Table 5-2), with 26 of 27 sites including three clusters and only one site clustering into four clusters. The average cluster size was 2.40 ha compared to an average site size of 14.40 ha (Table 5-1). Soil organic matter values ranged from 2.28 to 8.27 percent with an average value of 4.54 percent. Overall, there was a significant difference among clusters in terms of their organic matter content (Overall F-value = 22.33, p value of <0.01). Additionally, each cluster had significant differences in organic matter content relative to all other clusters (Table 5-3). The greatest differences in organic matter content was between clusters one and two, with cluster three only having slightly more organic matter compared to cluster two (Figure 5-3a). On average, organic matter contents in cluster one were 0.52 standard deviations below the mean, while cluster two included organic matter contents 0.15 standard deviations above the mean and cluster three included organic matter

contents 0.46 standard deviations also above the mean. For the one site with four clusters (cluster four) organic matter contents were 1.13 standard deviations above the mean.

Table 5-2. Cluster numbers at each site and average cluster sizes.

Number of clusters	Number of sites	Average cluster size (ha)	Minimum cluster size (ha)	Maximum cluster size (ha)
3	26	2.4	0.02	9.4
4	1	1.5	0.44	2.7

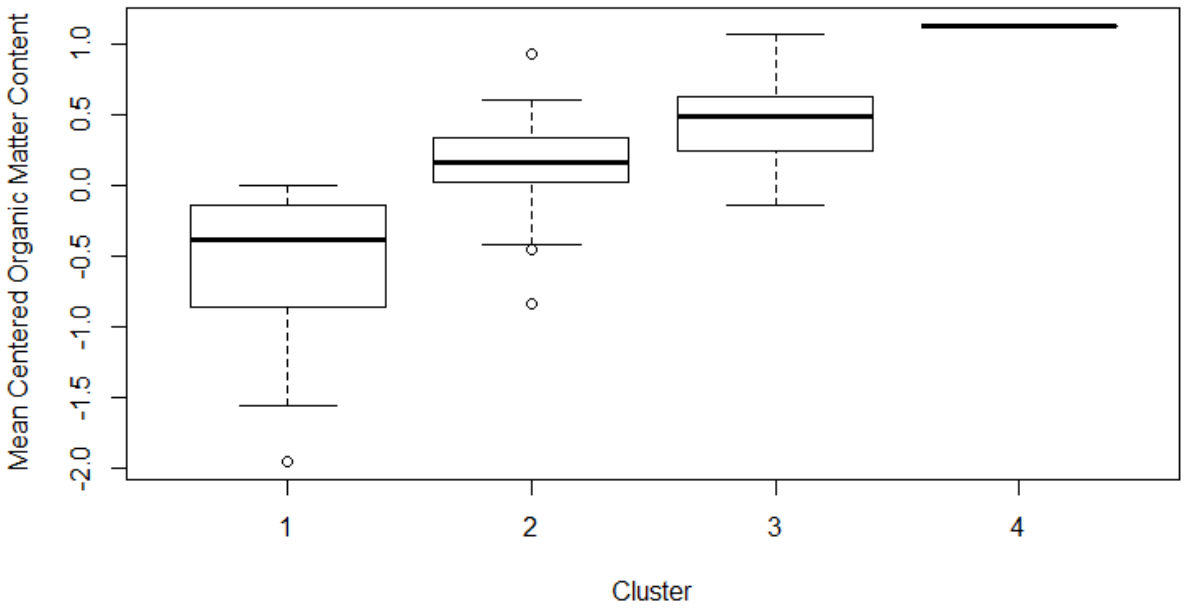
Table 5-3. Results from the linear mixed effects model examining if significant differences exist between organic matter content, cation exchange capacity and gravimetric water content across clusters.

Cluster	Parameter estimate	t-Value	p-value
Organic matter			
Intercept	-0.52	-5.76	<0.01
2	0.68	5.67	<0.01
3	1.07	6.53	<0.01
4	1.95	4.46	<0.01
Cation exchange capacity			
Intercept	-0.31	-2.87	<0.01
2	0.47	2.99	<0.01
3	0.88	4.22	<0.01
4	1.01	1.86	0.08
Gravimetric water content			
Intercept	-0.44	-5.76	<0.01
2	0.62	6.28	<0.01
3	1.08	8.00	<0.01
4	1.49	4.12	<0.01

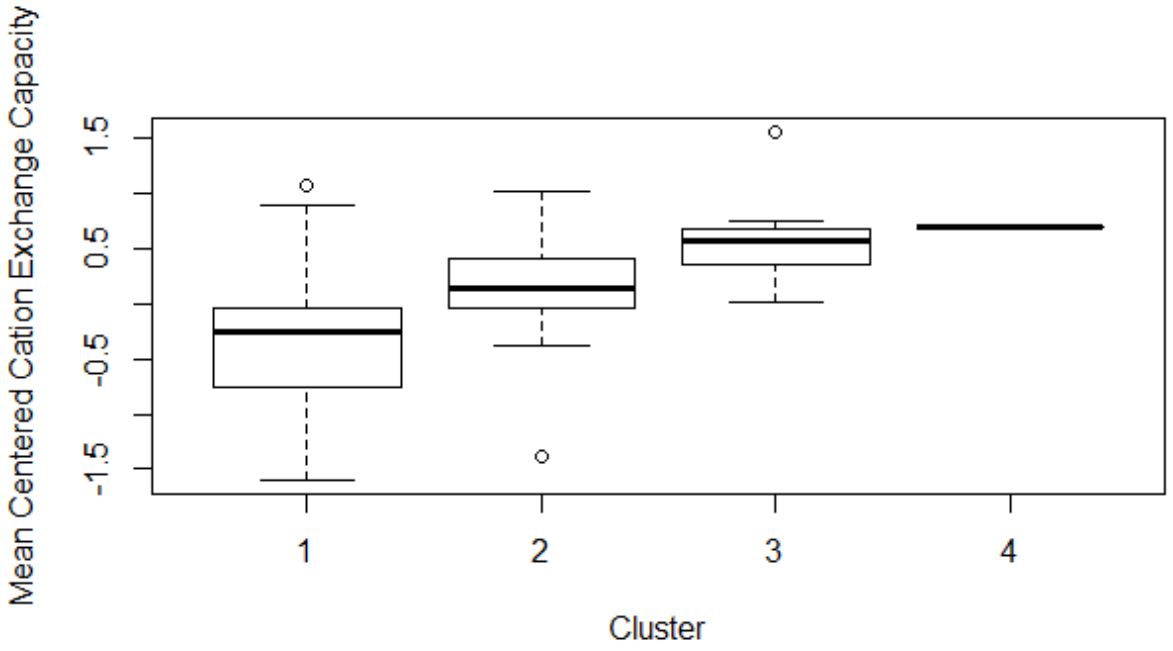
The CEC ranged from 1.24 to 71.76 meq/100g with an average value of 27.90 meq/100g across all sites (Table 5-1). The average gravimetric water content was 18 percent with values ranging from 8 to 29 percent. There was a significant difference among clusters in terms of CEC (F-value = 7.29 p-value<0.01) and gravimetric water content (F-value = 28.60, p-value<0.01). There were significant differences amongst all clusters for both parameters, with the exception of cluster four for CEC, which was not significantly different from cluster three (Table 5-3). On average cluster one had CEC values 0.31 standard deviations below the mean, CEC values for cluster two were 0.16 standard deviations above the mean and for cluster three 0.57 standard deviations above the mean (Figure 5-3b). For gravimetric water content, cluster one had average values 0.04 standard deviations below the mean, cluster two had average values at the mean, and cluster three had average values 0.25 standard deviations above the mean (Figure 5-3c).

Figure 5-3. Results for each unsupervised cluster across all study sites for (a) organic matter content, (b) cation exchange capacity and (c) soil moisture content. Values have been scaled and centered around the mean for each site to allow comparisons across sites with different magnitudes of soil values. No distribution can be provided for class 4 as only one site had four classes after the unsupervised cluster analysis. The circles indicate values that are 1.5 times the interquartile range above the third quartile or below the first quartile.

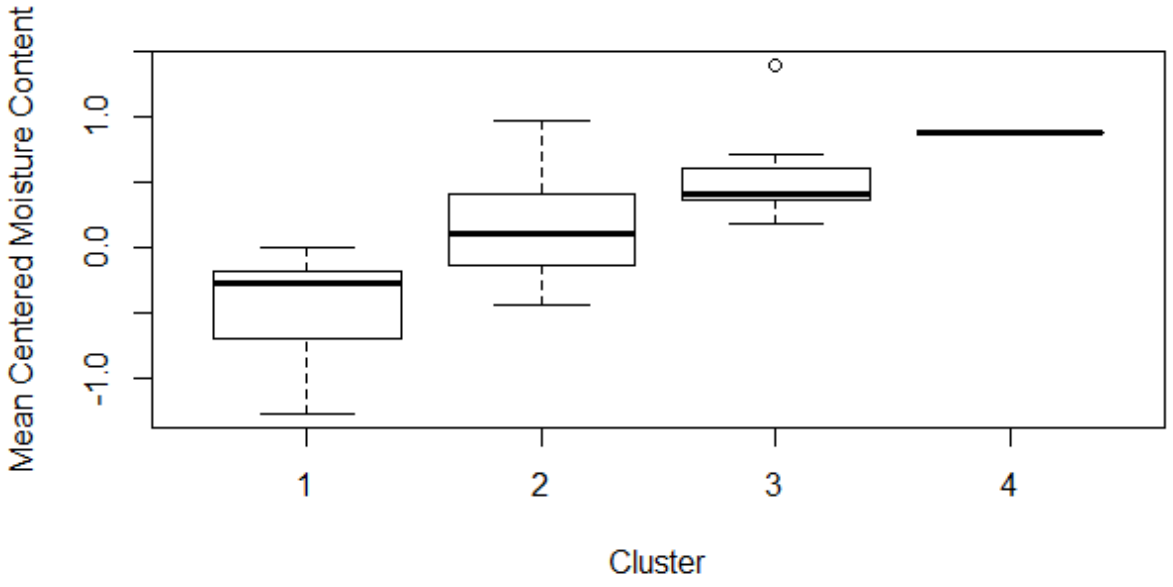
(a)



(b)


Mean Centered Cation Exchange Capacity
1.5
0.5
-0.5
-1.5
1
2
3
4
Cluster

(c)


Mean Centered Moisture Content
1.0
0.0
-1.0
1
2
3
4
Cluster

Previous work investigating the same sensors as the ones used in this study had been successful in building site-specific calibration models (Kweon et al., 2013; Piikki et al., 2016). However, calibration models that generalized to new areas could not be built in either of these two studies. In this present study, site-specific calibration models could not be successfully built for any of the analyzed soil parameters. Specifically, Piikki et al. (2016) used the optic mapper sensor with portable x-ray fluorescence and electromagnetic induction to build site-specific calibrations for total carbon. However, predictions for new sites were poor for total carbon, and soil texture could not be modelled with any accuracy. Kweon et al. (2013) concluded that field specific calibrations are possible and sufficiently accurate based on leave-one-out cross validation results. However, they could not develop a universal calibration model. While a quantitative model could not be developed in our study either, our results indicated that an unsupervised classification approach can be used to successfully classify a reclaimed area into zones with different organic matter, CEC and moisture contents based on soil reflectance data from two distinct bands along with elevation data. A possible explanation for why a quantitative model could not be built for this study is that 30-40 samples were collected from each site in Piikki et al. (2016), and with more data collected from each site a site-specific calibration model may have been possible with this study. Additionally, compared to the Kweon et al. (2013) study, the average standard deviation in organic matter was higher in this

study, likely due to the presence of disturbed soils. More calibration samples are needed to build an accurate calibration model for these sites.

It is important to note that while results from the current study demonstrate that proximal sensor data can be used to automatically zone soil with different parameters, the methodological approach used does not provide any information on the direction or magnitude of difference among classes. Alternative sensor arrangements have been used by other researchers to quantitively measure soil organic carbon and other soil parameters under field conditions. Viscarra Rossel et al. (2017) successfully used reflectance spectroscopy and gamma ray attenuation to measure soil organic stocks from soil cores collected in the field. Reflectance spectroscopy has been combined with x-ray fluorescence to successfully take multi-parameter measurements of soil (Duda et al., 2017). Additionally, field reflectance spectroscopy has been used under similar site conditions to successfully measure soil organic carbon and nitrogen, but not soil pH (Sorenson et al., 2017).

Construction activities have been documented to lower soil organic matter content and increase bulk density (Hammermeister et al., 2003). These changes in turn influence both the porosity of the soil along with the CEC, which is closely linked to organic matter content (Parfitt et al., 1995). Overall, these results show that two band reflectance measurements, along with elevation data, have a role to play for monitoring soil reclamation when combined with the appropriate data analysis techniques. While quantitative models could not be successfully built using the sensors in this study, automated clustering of the data allowed the identification of

zones with different organic matter, CEC and water contents (Table 5-3). In practice, follow-up investigations would be required for quantitative determination of soil organic matter, CEC and water content for each soil zone. This analysis could consist of either collecting samples for conventional laboratory analysis, or by collecting further point data using a higher spectral resolution reflectance spectroscopy system (P.T. Sorenson et al., 2017).

Ultimately, while there is value in qualitatively delineating reclaimed soils into zones to identify variance in key soil characteristics, quantitative measurements would be an improvement. Rather than using two reflectance bands, high resolution visible light near infrared reflectance spectroscopy could be utilized to provide quantitative measurements. Soil organic carbon, CEC and water content have all been shown to be successfully measurable using reflectance spectroscopy (Soriano-Disla et al., 2014). Point spectroscopy has been used on some of these sites to measure soil carbon quantitatively in-field as well (P.T. Sorenson et al., 2017).

5.5. Conclusion

The results of this study indicate that broad-band reflectance data combined with unsupervised classification approaches can be used to qualitatively and automatically map a field into soil zones with differences in organic matter, CEC and water contents. However, more site-specific data than were collected in this study or alterative types sensors are needed to quantitatively map soil organic matter content. Future work should focus on integrating reflectance spectroscopy sensors, rather than on two-band reflectance sensors, into a hardware

configuration similar to the Optic Mapper® to provide high spatial resolution quantitative measurements of soil parameters for reclamation monitoring.

Acknowledgements

We would like to thank Alberta Environment and Parks (formally the Alberta Environmental Monitoring and Evaluation Agency (AEMERA)) and InnoTech Alberta for providing financial assistance, soil samples and soil proximal soil data for the project. We would also like to acknowledge MITACS for providing financial support for this project with a MITACS Accelerate to S.A. Quideau (no IT08009).

5.6. References

Alberta Energy Regulator, 2016. Statistical Reports [WWW Document]. URL https://www.aer.ca/data-and-publications/statistical-reports (accessed 5.18.16).

Alberta Environment, 2010. 2010 Reclamation Criteria for Wellsites and Associated Facilities for Cultivated lands. Edmonton, Alberta.

Ben-Dor, E., Banin, A., 1995. Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj1995.03615995005900020014x

Chang, C.-W., Laird, D.A., 2002. NEAR-INFRARED REFLECTANCE SPECTROSCOPIC ANALYSIS OF SOIL C AND N. Soil Sci. https://doi.org/10.1097/00010694-200202000-00003

Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2001.652480x

Ciss, S., 2015. _randomUniformForest: random Uniform Forests for Classification, Regression and Unsupervised Learning_.

Duda, B.M., Weindorf, D.C., Chakraborty, S., Li, B., Man, T., Paulette, L., Deb, S., 2017. Soil characterization across catenas via advanced proximal sensors. Geoderma. https://doi.org/10.1016/j.geoderma.2017.03.017

Government of Alberta, 2010. Environmental Protection and Enhancement Act 1–164.

Hammermeister, A.M., Naeth, M.A., Schoenau, J.J., Biederbeck, V.O., Biederbeck, V.O., 2003. Soil and plant response to wellsite rehabilitation on native prairie in southeastern Alberta, Canada. Can. J. Soil Sci. 83, 507–519. https://doi.org/10.4141/S03-030

Hendershot, W.H., Lalande, H., Duquete, ., 2008. Ion Exchange and Exchangeable Cations, in: Carter, M.R., Gregorich, E.. (Eds.), Soil Sampling and Methods of Analysis. CRC Press, Boca Raton, Florida, pp. 197–206.

Hijmans, R.J., 2016. raster: Geographic Data Analysis and Modeling.

Ji, W., Adamchuk, V.I., Biswas, A., Dhawale, N.M., Sudarsan, B., Zhang, Y., Viscarra Rossel, R.A., Shi, Z., 2016. Assessment of soil properties in situ using a prototype portable MIR spectrometer in two agricultural fields. Biosyst. Eng. 152, 14–27. https://doi.org/10.1016/j.biosystemseng.2016.06.005

Kweon, G., Lund, E., Maxton, C., 2013. Soil organic matter and cation-exchange capacity sensing with on-the-go electrical conductivity and optical sensors. Geoderma. https://doi.org/10.1016/j.geoderma.2012.11.001

McConkey, T., Bulmer, C., Sanborn, P., 2012. Effectiveness of five soil reclamation and reforestation techniques on oil and gas well sites in northeastern British Columbia. Can. J. Soil Sci. https://doi.org/10.4141/cjss2010-019

Mele, P.M., Crowley, D.E., 2008. Application of self-organizing maps for assessing soil biological quality. Agric. Ecosyst. Environ. https://doi.org/10.1016/j.agee.2007.12.008

Nelson, D.W., Sommers, L.E., 1996. Total Carbon, Organic Carbon, and Organic Matter, in: Sparks, D.L., Page, A.., Helmke, P.A., Loeppert, R.H., Soltanpour, P.N., Tabatabai, M.A., Johnston, C.T., Sumner, M.E. (Eds.), Methods of Soil Analysis Part 3—Chemical Methods. Soil Science Society of America, Madison, WI, pp. 961–1011.

Parfitt, R.L., Giltrap, D.J., Whitton, J.S., 1995. Contribution of organic matter and clay minerals to the cation exchange capacity of soils. Commun. Soil Sci. Plant Anal. https://doi.org/10.1080/00103629509369376

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Comput. Geosci. 30, 683–691.

Piikki, K., Sï¿½derstrï¿½m, M., Eriksson, J., John, J.M., Muthee, P.I., Wetterlind, J., Lund, E., 2016. Performance evaluation of proximal sensors for soil assessment in smallholder farms in Embu County,Kenya. Sensors (Switzerland). https://doi.org/10.3390/s16111950

Pinheiro, J., Bates, D., Debroy, S., Sarkar, D., R Core Team, 2016. _nlme: Linear and Nonliner Mixed Effects Models_.

R Core Team, 2018. R: A language and environment for statistical computing.

Rossel, R.A.A.V., Behrens, T., Viscarra Rossel, R.A., Behrens, T., 2010. Using data mining to model and interpret soil diffuse reflectance spectra. Geoderma 158, 46–54. https://doi.org/10.1016/j.geoderma.2009.12.025

Sorenson, P.T., Quideau, S.A., Rivard, B., 2018. High resolution measurement of soil organic carbon and total nitrogen with laboratory imaging spectroscopy. Geoderma. https://doi.org/10.1016/j.geoderma.2017.11.032

Sorenson, P.T., Small, C., Tappert, M.C., Quideau, S.A., Drozdowski, B., Underwood, A., Janz, A., 2017. Monitoring organic carbon, total nitrogen, and pH for reclaimed soils using field reflectance spectroscopy. Can. J. Soil Sci. https://doi.org/10.1139/cjss-2016-0116

Soriano-Disla, J.M., Janik, L.J., Viscarra Rossel, R.A., MacDonald, L.M., McLaughlin, M.J., 2014. The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. Appl. Spectrosc. Rev. https://doi.org/10.1080/05704928.2013.811081

Topp, G.C., Parkin, G.W., Ferré, T.P.., 2008. Soil Water Content, in: Carter, M., Gregorich, E. (Eds.), Soil Sampling and Methods of Analysis. CRC Press, Boca Raton, Florida, pp. 939–961.

Viscarra Rossel, R.A., Lobsey, C.R., Sharman, C., Flick, P., McLachlan, G., 2017. Novel Proximal Sensing for Monitoring Soil Organic C Stocks and Condition. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.7b00889

Wehrens, R., Buydens, L.M.C.L.M.C., 2007. Self- and Super-organising Maps in R: the kohonen package. J. Stat. Softw. 21, 1–19. https://doi.org/10.18637/jss.v021.i05

6. Summary

6.1. Objectives

The objectives of this research were to investigate the following:

1. To determine if shortwave infrared imaging spectroscopy could be used to measure soil organic carbon and total nitrogen on intact and unground samples in the laboratory for a variety of Canadian soil samples. Further, to use the imaging spectroscopy results to characterize the spatial distribution of SOC and TN at the soil aggregate scale, and determine if the distribution of SOC and TN varies based on soil type and horizon at fine spatial scales.
2. To identify at what depth changes could be detected in SOC and TN using imaging spectroscopy. Specifically, following the use of different crop rotations, and if these changes in SOC and TN were associated with a change in the spatial distribution of these parameters at the same fine spatial scales.
3. To determine if reflectance spectroscopy data collected in the field with a drill-rig mounted spectroscopy system could be used to measure SOC, TN and pH as part of reclamation assessments.
4. To investigate if a simple two-band reflectance sensor could be used to generate quantitative SOC results or support reclamation assessments by successfully identifying different zones of soil organic matter content in reclaimed soil.

6.2. High Resolution Soil Organic Carbon and Nitrogen Measurement

Imaging spectroscopy presents novel opportunities to determine the concentration of C and N, and the spatial distribution of these parameters at fine spatial scales. Using shortwave infrared imaging spectroscopy aggregation in soil organic C and N at fine scales is observable. Differences amongst horizons and soil orders are observable in terms of the aggregation of both C and N. This work indicates that imaging spectroscopy has an important role to play in investigating soil parameter relationships at fine scales due to the ability to collect data non-destructively at a sub centimeter scale.

By utilizing imaging spectroscopy, this work enabled detailed investigation of the spatial aggregation of SOC and TN at fine spatial scales. Specifically, TN was significantly more aggregated that SOC and Chernozemic soils had more aggregated SOC compared to Luvisols and Gleysols. The key novel contribution of this work was to illustrate that imaging spectroscopy can be used to measure SOC and TN in Canadian soils, and that the spatial relationships in these parameters can be investigated at spatial scales not possible using conventional techniques.

6.3. Soil carbon Distribution After Introducing Forages to Cultivated Boreal Forest Soils

Gray Luvisolic soils converted to arable crop production present unique challenges compared to the cultivation of Chernozemic soils. The organic matter rich LFH layer begins to decompose following cultivation, leading to a loss in organic matter, and the high clay Bt horizon presents a potential barrier for deep root penetration. The introduction of forages into crop rotation clearly leads to an increase in soil organic carbon in the topsoil. However, the Bt horizon appears to limit the additions of carbon to the subsoil following the introduction of forages. An important consideration from these results is that any loss of topsoil due to erosion or management change will likely lead to losses of any gained carbon and it will not be stored deeper in the horizon where it is less sensitive to management change.

Additionally, the conversion of land to permanent forages leads to an increase in carbon in the top 2 to 3 cm of soil, but not an increase on average throughout the full profile compared to introducing forages into a rotation that also includes grains and pulses. These results indicate that the majority of soil carbon storage benefits can likely be obtained from including forages in rotation, and the additional benefit of complete forage conversion is limited. This is advantageous for producers as a rotation that includes forages will have economic advantages and flexibility in most cases compared to a complete forage conversion.

While there has been previous investigation of these soil parameters at these research plots, imaging spectroscopy represented a novel way to investigate SOC and TN and the influence of rotation on these parameters. The fine spatial resolution data obtainable with imaging spectroscopy enabled an understanding of the precise depth of influence of treatment

effects on SOC and TN and effects on the spatial structure of these parameters throughout the soil profile.

6.4. Reflectance Spectroscopy to Support Reclamation Assessments

Reflectance spectroscopy presents a technological solution that makes quantitative assessment of soil parameters for reclamation assessment possible. However, a few key considerations need to be taken into account. First is the spectral resolution and range of the spectrometer. While key organic matter features are found below 2200 nm, a spectrometer that includes up to 2500 nm has clear advantages as key carbon bonds are present near 2275 nm and throughout the 2300 to 2500 nm range. Second, systems that can collect a high number of spatially referenced data points are valuable. That said, for quantitative measurements a full continuous spectrum that includes the entire shortwave infrared region of the electromagnetic spectra is desirable. This type of system allows for quantitative measurements, and the ability to use tools such as wavelet analysis to remove non-compositional effects from the spectra. Both of these factors are not possible using a simple two-band reflectance sensor.

Overall, this work indicated that sensor data can be used to reliably collect soil parameter data in the field to support quantitative reclamation assessments. Also, even simple two-band reflectance sensors can identify zones with varying concentrations of key soil parameters, when combined with appropriate data analysis methods. However, for truly quantitative results a spectroscopy system is needed, and ideally one that collects to 2500 nm.

6.5. Recommendations for Future Research

Further work is still needed to follow up the work investigating soil property changes following the introduction of forages into crop rotations. Increases in SOC and TN were restricted to the A horizon and shallow B horizon depths in the soils that were investigated. These soils were Luvisolic soils characterized by a Bt horizon with high clay content and bulk density values. The depth of influence of forage crops on SOC contents could be much deeper in Chernozemic soils that have a consistent clay content throughout the profile. Imaging spectroscopy has an important role to play in this investigation to identify the precise depth of influence. Imaging spectroscopy also will enable the investigation of the level of spatial aggregation of the SOC and how much the SOC is associated with clay, both of which provide valuable insight into the stability of SOC additions throughout the profile.

This work has shown that reflectance spectroscopy can be used to successfully measure SOC and TN in reclaimed soils. Further research is needed to investigate the use of reflectance spectroscopy to measure petroleum hydrocarbon concentrations in soil in Western Canada. There is are over 400,000 wellsites in Alberta that will need to be reclaimed. Reflectance spectroscopy tools deployed in the field can reduce costs by preventing over-remediation of sites and reducing equipment standby time. Research is needed to determine the optimal data the viability and optimal data processing techniques for field measurement of petroleum hydrocarbons with reflectance spectroscopy.

Reflectance spectroscopy and imaging spectroscopy are maturing as soil sensing technologies. Recent advances in both the hardware and data processing techniques now make it possible to rapidly and non-destructively collect orders of magnitude more soil data compared to conventional wet chemistry techniques. By reducing the cost to collect large volumes of data, these tools will play an important role in improving soil management and conservation.

References

Alberta Energy Regulator, 2011. Directive for Monitoring the Impact of Sulphur Dust on Soils. Edmonton, Alberta.

Alberta Energy Regulator, 2013. Report 2013-B: Pipeline Performance in Alberta, 1990-2012. Calgary, Alberta.

Alberta Energy Regulator, 2016. Statistical Reports [WWW Document]. URL https://www.aer.ca/data-and-publications/statistical-reports (accessed 5.18.16).

Alberta Energy Regulator, 2016a. ST37: List of Wells in Alberta Monthly Updates [WWW Document]. URL https://www.aer.ca/data-and-publications/statistical-reports/st37 (accessed 4.26.16).

Alberta Energy Regulator, 2016b. Statistical Reports [WWW Document]. URL https://www.aer.ca/data-and-publications/statistical-reports (accessed 5.18.16).

Alberta Environment, 2010. 2010 Reclamation Criteria for Wellsites and Associated Facilities for Cultivated lands. Edmonton, Alberta.

Anderson, D., Cerkowniak, D., 2010. Soil Formation in the Canadian Prairie Region. Prairie Soils Crop. 3, 57–64.

Bartholomeus, H.M., Schaepman, M.E., Kooistra, L., Stevens, A., Hoogmoed, W.B., Spaargaren, O.S.P., 2008. Spectral reflectance based indices for soil organic carbon quantification. Geoderma. https://doi.org/10.1016/j.geoderma.2008.01.010

Ben-Dor, E., Banin, A., 1990. Near-infrared reflectance analysis of carbonate concentration in soils. Appl. Spectrosc. https://doi.org/10.1366/0003702904086821

Ben-Dor, E., Banin, A., 1995. Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj1995.03615995005900020014x

Ben-Dor, E., Taylor, R.G., Hill, J., Demattê, J.A.M., Whiting, M.L., Chabrillat, S., Sommer, S., 2008. Imaging Spectrometry for Soil Applications. Adv. Agron. https://doi.org/10.1016/S0065-2113(07)00008-9

Berg, B., 2000. Litter decomposition and organic matter turnover in northern forest soils. For. Ecol. Manage. 133, 13–22. https://doi.org/http://dx.doi.org/10.1016/S0378-1127(99)00294-7

Bingham, A.H., Cotrufo, M.F., 2016. Organic nitrogen storage in mineral soil: Implications for policy and management. Sci. Total Environ. 551–552, 116–126. https://doi.org/10.1016/j.scitotenv.2016.02.020

Blake, L., Goulding, K.W.T., 2002. Effects of atmospheric deposition, soil pH and acidification on heavy metal contents in soils and vegetation of semi-natural ecosystems at

Rothamsted Experimental Station, UK. Plant Soil. https://doi.org/10.1023/A:1015731530498

Bremner, J.M., 1996. Nitrogen - Total, in: Sparks, D.L., Page, A.., Helmke, P.A., Loeppert, R.H., Soltanpour, P.N., Tabatabai, M.A., Johnston, C.T., Sumner, M.E. (Eds.), Methods of Soil Analysis Part 3—Chemical Methods. Madison, WI, pp. 1085–1123.

Cardinael, R., Eglin, T., Guenet, B., Neill, C., Houot, S., Chenu, C., 2015. Is priming effect a significant process for long-term SOC dynamics? Analysis of a 52-years old experiment. Biogeochemistry 123, 203–219. https://doi.org/10.1007/s10533-014-0063-2

Chabrillat, S., Goetz, A.F.H., Krosley, L., Olsen, H.W., 2002. Use of hyperspectral images in the identification and mapping of expansive clay soils and the role of spatial resolution. Remote Sens. Environ. https://doi.org/10.1016/S0034-4257(02)00060-3

Chakraborty, S., Weindorf, D.C., Morgan, C.L.S., Ge, Y., Galbraith, J.M., Li, B., Kahlon, C.S., 2010. Rapid Identification of Oil-Contaminated Soils Using Visible Near-Infrared Diffuse Reflectance Spectroscopy. J. Environ. Qual. https://doi.org/10.2134/jeq2010.0183

Chang, C.-W., Laird, D.A., 2002. Near-Infrared Reflectance Spectroscopic Analysis of Soil C and N. Soil Sci. 167, 110–116. https://doi.org/10.1097/00010694-200202000-00003

Chang, C.-W., Laird, D.A., 2002. NEAR-INFRARED REFLECTANCE SPECTROSCOPIC ANALYSIS OF SOIL C AND N. Soil Sci. https://doi.org/10.1097/00010694-200202000-00003

Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2001.652480x

Chen, G., Weil, R.R., 2010. Penetration of cover crop roots through compacted soils. Plant Soil 331, 31–43. https://doi.org/10.1007/s11104-009-0223-7

Chen, R., Senbayram, M., Blagodatsky, S., Myachina, O., Dittert, K., Lin, X., Blagodatskaya, E., Kuzyakov, Y., 2014. Soil C and N availability determine the priming effect: Microbial N mining and stoichiometric decomposition theories. Glob. Chang. Biol. 20, 2356–2367. https://doi.org/10.1111/gcb.12475

Ciss, S., 2015. _randomUniformForest: random Uniform Forests for Classification, Regression and Unsupervised Learning_.

Conant, R.T., Cerri, C.E.P., Osborne, B.B., Paustian, K., 2017. Grassland management impacts on soil carbon stocks: a new synthesis. Ecol. Appl. 27, 662–668. https://doi.org/10.1002/eap.1473

Dessureault-Rompré, J., Zebarth, B.J., Burton, D.L., Georgallas, A., 2015. Predicting soil nitrogen supply from soil properties. Can. J. Soil Sci. https://doi.org/10.4141/cjss-2014-057

Doetterl, S., Stevens, A., Van Oost, K., van Wesemael, B., 2013. Soil Organic Carbon Assessment at High Vertical Resolution using Closed-Tube Sampling and Vis-NIR Spectroscopy. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2012.0410n

Duda, B.M., Weindorf, D.C., Chakraborty, S., Li, B., Man, T., Paulette, L., Deb, S., 2017. Soil characterization across catenas via advanced proximal sensors. Geoderma. https://doi.org/10.1016/j.geoderma.2017.03.017

Dyck, M.F., Roberston, J.A., Puurveen, D., 2012. The University of Alberta Breton Plots. Prairie Soils Crop. J. 5, 96–115.

Fontaine, S., Barot, S., Barré, P., Bdioui, N., Mary, B., Rumpel, C., 2007. Stability of organic carbon in deep soil layers controlled by fresh carbon supply. Nature. https://doi.org/10.1038/nature06275

Forrester, S.T., Janik, L.J., McLaughlin, M.J., Soriano-Disla, J.M., Stewart, R., Dearman, B., 2013. Total Petroleum Hydrocarbon Concentration Prediction in Soils Using Diffuse Reflectance Infrared Spectroscopy. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2012.0201

Ge, Y., Morgan, C.L.S., Ackerson, J.P., 2014. VisNIR spectra of dried ground soils predict properties of soils scanned moist and intact. Geoderma. https://doi.org/10.1016/j.geoderma.2014.01.011

Gomez, C., Lagacherie, P., Coulouma, G., 2008a. Continuum removal versus PLSR method for clay and calcium carbonate content estimation from laboratory and airborne hyperspectral measurements. Geoderma. https://doi.org/10.1016/j.geoderma.2008.09.016

Gomez, C., Lagacherie, P., Coulouma, G., 2012. Regional predictions of eight common soil properties and their spatial structures from hyperspectral Vis-NIR data. Geoderma. https://doi.org/10.1016/j.geoderma.2012.05.023

Gomez, C., Viscarra Rossel, R.A., McBratney, A.B., 2008. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. Geoderma. https://doi.org/10.1016/j.geoderma.2008.06.011

Gomez, C., Viscarra Rossel, R.A., McBratney, A.B., 2008b. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. Geoderma. https://doi.org/10.1016/j.geoderma.2008.06.011

Government of Alberta, 2009. Soil Monitoring Directive. Edmonton, Alberta.

Government of Alberta, 2010. Environmental Protection and Enhancement Act 1–164.

Guo, L.B., Gifford, R.M., 2002. Soil carbon stocks and land use change: A meta analysis. Glob. Chang. Biol. https://doi.org/10.1046/j.1354-1013.2002.00486.x

Hammermeister, A.M., Naeth, M.A., Schoenau, J.J., Biederbeck, V.O., Biederbeck, V.O., 2003. Soil and plant response to wellsite rehabilitation on native prairie in southeastern Alberta, Canada. Can. J. Soil Sci. 83, 507–519. https://doi.org/10.4141/S03-030

He, Y., Huang, M., García, A., Hernández, A., Song, H., 2007. Prediction of soil macronutrients content using near-infrared spectroscopy. Comput. Electron. Agric. https://doi.org/10.1016/j.compag.2007.03.011

Hendershot, W.H., Lalande, H., Duquete, ., 2008. Ion Exchange and Exchangeable Cations, in: Carter, M.R., Gregorich, E.. (Eds.), Soil Sampling and Methods of Analysis. CRC Press, Boca Raton, Florida, pp. 197–206.

Hijmans, R.J., 2016. raster: Geographic Data Analysis and Modeling.

Hijmans, R.J., 2017. raster: Geographic Data Analysis and Modeling.

Hobley, E., Steffens, M., Bauke, S.L., Kögel-Knabner, I., 2018. Hotspots of soil organic carbon storage revealed by laboratory hyperspectral imaging. Sci. Rep. 8, 13900. https://doi.org/10.1038/s41598-018-31776-w

Hoyos-Santillan, J., Lomax, B.H., Large, D., Turner, B.L., Boom, A., Lopez, O.R., Sjögersten, S., 2015. Getting to the root of the problem: litter decomposition and peat formation in lowland Neotropical peatlands. Biogeochemistry. https://doi.org/10.1007/s10533-015-0147-7

Indorante, S.J., Jansen, I.J., Boast, C.W., 1981. Surface mining and reclamation : Initial changes in soil character. J. Soil Water Conserv. 36, 347–351.

IUSS Working Group WRB, 2014. World reference base for soil resources 2014. International soil classification system for naming soils and creating legends for soil maps, World Soil Resources Reports No. 106. https://doi.org/10.1017/S0014479706394902

Jastrow, J.D., 1996. PIk soo38-0717(!spo159-x SOIL AGGREGATE FORMATION AND THE ACCRUAL OF PARTICULATE AND MINERAL-ASSOCIATED ORGANIC MATTER. Soil Biol. Biochem 28, 665–676.

Ji, W., Adamchuk, V.I., Biswas, A., Dhawale, N.M., Sudarsan, B., Zhang, Y., Viscarra Rossel, R.A., Shi, Z., 2016. Assessment of soil properties in situ using a prototype portable MIR spectrometer in two agricultural fields. Biosyst. Eng. 152, 14–27. https://doi.org/10.1016/j.biosystemseng.2016.06.005

Jobbágy, E.G., Jackson, R.B., 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecol. Appl. https://doi.org/10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2

Karatzoglou, A., Smola, A., Hornik, K., Seileis, A., 2004. kernlab - An S4 Package for Kernal Methods in R. J. Stat. Softw. 11, 1–20.

Kinoshita, R., Moebius-Clune, B.N., van Es, H.M., Hively, W.D., Bilgilis, A.V., 2012. Strategies for Soil Quality Assessment Using Visible and Near-Infrared Reflectance Spectroscopy in a Western Kenya Chronosequence. Soil Sci. Soc. Am. J. https://doi.org/10.2136/sssaj2011.0307

Klavarioti, M., Kostarelos, K., Pourjabbar, A., Ghandehari, M., 2014. In situ sensing of subsurface contamination-part I: Near-infrared spectral characterization of alkanes, aromatics, and chlorinated hydrocarbons. Environ. Sci. Pollut. Res. https://doi.org/10.1007/s11356-013-2478-z

Kuhn, M., Johnson, K., n.d. Applied Predictive Modeling.

Kuhn, M., Weston, S., Keefer, C., Coulter, N., 2014. Cubist: Rule- and Instance-Based Regression Modeling.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., R Core Team, Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., 2016. caret: Classification and Regression Training.

Kweon, G., Lund, E., Maxton, C., 2013. Soil organic matter and cation-exchange capacity sensing with on-the-go electrical conductivity and optical sensors. Geoderma. https://doi.org/10.1016/j.geoderma.2012.11.001

Laakso, K., Rivard, B., Rogge, D., 2016. Enhanced detection of gossans using hyperspectral data: Example from the Cape Smith Belt of northern Quebec, Canada. ISPRS J. Photogramm. Remote Sens. https://doi.org/10.1016/j.isprsjprs.2016.02.004

Lagacherie, P., Baret, F., Feret, J.B., Madeira Netto, J., Robbez-Masson, J.M., 2008. Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2007.06.014

Lal, R., 2004. Soil Carbon Sequestration Impacts on Global Climate Change and Food Security. Am. Assoc. Adv. Sci. 304, 1623–7. https://doi.org/10.1126/science.1097396

LeBauer, D.S., Treseder, K.K., 2008. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. Ecology 89, 371–379. https://doi.org/10.1890/06-2057.1

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 2, 18–22.

Luo, Z., Wang, E., Sun, O.J., 2010. Can no-tillage stimulate carbon sequestration in agricultural soils? A meta-analysis of paired experiments. Agric. Ecosyst. Environ. 139, 224–231. https://doi.org/10.1016/j.agee.2010.08.006

Malhi, S.S., Brandt, S., Gill, K.S., 2003. Cultivation and grassland type effects on light fraction and total organic C and N in a Dark Brown Chernozemic soil. Can. J. Soil Sci. 83, 145–153. https://doi.org/10.4141/S02-028

Martin, P.D., Malley, D.F., Manning, G., Fuller, L., 2002. Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy. Can. J. Soil Sci. https://doi.org/10.4141/S01-054

McBratney, A.B., Minasny, B., Viscarra Rossel, R., 2006. Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. Geoderma. https://doi.org/10.1016/j.geoderma.2006.03.051

McCarty, G.W., Reeves, J.B., Reeves, V.B., Follett, R.F., Kimble, J.M., 2002. Mid-Infrared and Near-Infrared Diffuse Reflectance Spectroscopy for Soil Carbon Measurement. Soil Sci. Soc. Am. J. 66, 640. https://doi.org/10.2136/sssaj2002.0640

McConkey, T., Bulmer, C., Sanborn, P., 2012. Effectiveness of five soil reclamation and reforestation techniques on oil and gas well sites in northeastern British Columbia. Can. J. Soil Sci. https://doi.org/10.4141/cjss2010-019

McIntire, E.J.B., Fajardo, A., 2009. Beyond description: the active and effective way to infer processes from spatial patterns 90, 46–56.

McLean, E.O., 1982. Soil pH and Lime Requirement, in: Page, A.., Miller, R.H., Keeney, D.R. (Eds.), Methods of Soil Analysis Part 2, Chemical and Microbiological Properties. Madison, WI, pp. 199–225.

Mele, P.M., Crowley, D.E., 2008. Application of self-organizing maps for assessing soil biological quality. Agric. Ecosyst. Environ. https://doi.org/10.1016/j.agee.2007.12.008

Melendez-Pastor, I., Navarro-Pedreño, J., Koch, M., Gómez, I., 2010. Applying imaging spectroscopy techniques to map saline soils with ASTER images. Geoderma. https://doi.org/10.1016/j.geoderma.2010.02.015

Mevik, B.H., Wehrens, R., Liland, K.H., 2015. pls: Partial Least Squares and Principal Component Regression.

Milborrow, S., 2016. earth: Multivariate Adaptive Regression Splines.

Milborrow, S., 2018. earth: Multivariate Adaptive Regression Splines.

Minasny, B., McBratney, a B., Pichon, L., Sun, W., Short, M.G., 2009. Evaluating near infrared spectroscopy for field prediction of soil properties. Aust. J. Soil Res. 47, 664–673. https://doi.org/Doi 10.1071/Sr09005

Minasny, B., McBratney, A.B., 2008. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. Chemom. Intell. Lab. Syst. https://doi.org/10.1016/j.chemolab.2008.06.003

Morellos, A., Pantazi, X.-E., Moshou, D., Alexandridis, T., Whetton, R., Tziotzios, G., Wiebensohn, J., Bill, R., Mouazen, A.M., 2016. Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. Biosyst. Eng. https://doi.org/10.1016/j.biosystemseng.2016.04.018

Nawar, S., Buddenbaum, H., Hill, J., Kozak, J., Mouazen, A.M., 2016. Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy. Soil Tillage Res. https://doi.org/10.1016/j.still.2015.07.021

Nelson, D.W., Sommers, L.E., 1996. Total Carbon, Organic Carbon, and Organic Matter, in: Sparks, D.L., Page, A.., Helmke, P.A., Loeppert, R.H., Soltanpour, P.N., Tabatabai, M.A., Johnston, C.T., Sumner, M.E. (Eds.), Methods of Soil Analysis Part 3—Chemical Methods. Soil Science Society of America, Madison, WI, pp. 961–1011.

Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., Dor, E. Ben, Brown, D.J., Clairotte, M., Csorba, A., Dardenne, P., Demattê, J.A.M., Genot, V., Guerrero, C., Knadel, M., Montanarella, L., Noon, C., Ramirez-Lopez, L., Robertson, J., Sakai, H., Soriano-Disla, J.M., Shepherd, K.D., Stenberg, B., Towett, E.K., Vargas, R., Wetterlind, J., 2015. Soil Spectroscopy: An Alternative to Wet Chemistry for Soil Monitoring, in: Advances in Agronomy. https://doi.org/10.1016/bs.agron.2015.02.002

Okparanma, R.N., Coulon, F., Mouazen, A.M., 2014. Analysis of petroleum-contaminated soils by diffuse reflectance spectroscopy and sequential ultrasonic solvent extraction-gas chromatography. Env. Pollut 184, 298–305. https://doi.org/10.1016/j.envpol.2013.08.039

Ouerghemmi, W., Gomez, C., Naceur, S., Lagacherie, P., 2011. Applying blind source separation on hyperspectral data for clay content estimation over partially vegetated surfaces. Geoderma. https://doi.org/10.1016/j.geoderma.2011.04.019

Parfitt, R.L., Giltrap, D.J., Whitton, J.S., 1995. Contribution of organic matter and clay minerals to the cation exchange capacity of soils. Commun. Soil Sci. Plant Anal. https://doi.org/10.1080/00103629509369376

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Comput. Geosci. 30, 683–691.

Percival, D.B., Walden, A.T., William, A., Percival, D., Constantine, M.W., 2016. Wavelet Methods for Time Series Analysis.

Piikki, K., Sï¿½derstrï¿½m, M., Eriksson, J., John, J.M., Muthee, P.I., Wetterlind, J., Lund, E., 2016. Performance evaluation of proximal sensors for soil assessment in smallholder farms in Embu County,Kenya. Sensors (Switzerland). https://doi.org/10.3390/s16111950

Pinheiro, J., Bates, D., Debroy, S., Sarkar, D., R Core Team, 2016. _nlme: Linear and Nonliner Mixed Effects Models_.

Poirier, V., Angers, D.A., Rochette, P., Chantigny, M.H., Ziadi, N., Tremblay, G., Fortin, J., 2009. Interactive Effects of Tillage and Mineral Fertilization on Soil Carbon Profiles. Soil Sci. Soc. Am. J. 73, 255. https://doi.org/10.2136/sssaj2008.0006

Post, M., Kwon, K.C., 2000. Soil Carbon Sequestration and Land-Use Change: Processes and Potential. Glob. Chang. Biol. 6, 317–328. https://doi.org/10.1046/j.1365-2486.2000.00308.x

R Core Team, 2018. R: A language and environment for statistical computing.

Reeves, D.W., 1997. The role of soil organic matter in maintaining soil quality in continuous cropping systems. soil& Tillage Res. Soil Tillage Res. Reeues/Soil Tillage Res. 43, 131–167.

Ribeiro, P.J., Diggle, P.J., 2016. geoR: Analysis of Geostatistical Data.

Rivard, B., Feng, J., Gallie, A., Sanchez-Azofeifa, A., 2008. Continuous wavelets for the improved use of spectral libraries and hyperspectral data. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2008.01.016

Rodriguez, P.P., Gianola, D., 2016. brnn: Bayesian Regularization for Feed-Forward Neural Networks.

Roger, J.M., Chauchard, F., Bellon-Maurel, V., 2003. EPO-PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. Chemom. Intell. Lab. Syst. 66, 191–204. https://doi.org/10.1016/S0169-7439(03)00051-0

Ross, S.M., Izaurralde, R.C., Janzen, H.H., Robertson, J.A., McGill, W.B., 2008. The nitrogen balance of three long-term agroecosystems on a boreal soil in western Canada. Agric. Ecosyst. Environ. https://doi.org/10.1016/j.agee.2008.04.007

Rossel, R.A.A.V., Behrens, T., Viscarra Rossel, R.A., Behrens, T., 2010. Using data mining to model and interpret soil diffuse reflectance spectra. Geoderma 158, 46–54. https://doi.org/10.1016/j.geoderma.2009.12.025

Scafutto, R.D.P.M., de Souza Filho, C.R., Rivard, B., 2016. Characterization of mineral substrates impregnated with crude oils using proximal infrared hyperspectral imaging. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2016.03.033

Schaumann, G.E., 2006. Soil organic matter beyond molecular structure Part II: Amorphous nature and physical aging. J. Plant Nutr. Soil Sci. https://doi.org/10.1002/jpln.200521791

Shrestha, R.K., Lal, R., 2007. Soil Carbon and Nitrogen in 28-Year-Old Land Uses in Reclaimed Coal Mine Soils of Ohio. J. Environ. Qual. https://doi.org/10.2134/jeq2007.0071

Soil Classification Working Group, 1998. The Canadian System of Soil Classification. Can. Syst. Soil Classif. 3rd ed. Agric. Agri-Food Canada Publ. 1646 187.

Sorenson, P.T., MacKenzie, M.D., Quideau, S.A., Landhausser, S.M., 2017. Can spatial patterns be used to investigate aboveground-belowground links in reclaimed forests? Ecol. Eng. 104, 57–66. https://doi.org/10.1016/j.ecoleng.2017.04.002

Sorenson, P.T., Quideau, S.A., MacKenzie, M.D., Landhäusser, S.M., Oh, S.W., 2011. Forest floor development and biochemical properties in reconstructed boreal forest soils. Appl. Soil Ecol. https://doi.org/10.1016/j.apsoil.2011.06.006

Sorenson, P.T., Quideau, S.A., Rivard, B., 2018. High resolution measurement of soil organic carbon and total nitrogen with laboratory imaging spectroscopy. Geoderma. https://doi.org/10.1016/j.geoderma.2017.11.032

Sorenson, P.T., Small, C., Tappert, M.C., Quideau, S.A., Drozdowski, B., Underwood, A., Janz, A., 2017. Monitoring organic carbon, total nitrogen, and pH for reclaimed soils using field reflectance spectroscopy. Can. J. Soil Sci. https://doi.org/10.1139/cjss-2016-0116

Soriano-Disla, J.M., Janik, L.J., Viscarra Rossel, R.A., MacDonald, L.M., McLaughlin, M.J., 2014. The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. Appl. Spectrosc. Rev. https://doi.org/10.1080/05704928.2013.811081

St. Luce, M., Ziadi, N., Zebarth, B.J., Grant, C.A., Tremblay, G.F., Gregorich, E.G., 2014. Rapid determination of soil organic matter quality indicators using visible near infrared reflectance spectroscopy. Geoderma. https://doi.org/10.1016/j.geoderma.2014.05.023

Steffens, M., Buddenbaum, H., 2013. Laboratory imaging spectroscopy of a stagnic Luvisol profile - High resolution soil characterisation, classification and mapping of elemental concentrations. Geoderma. https://doi.org/10.1016/j.geoderma.2012.11.011

Stevens, A., Ramirez-Lopez, L., 2013. An introduction to the prospectr package.

Strong, W.L., La Roi, G.H., 1983. Root-system morphology of common boreal forest trees in Alberta, Canada. Can. J. For. Res. 13, 1164–1173.

Su, Y.Z., 2007. Soil carbon and nitrogen sequestration following the conversion of cropland to alfalfa forage land in northwest China. Soil Tillage Res. 92, 181–189. https://doi.org/10.1016/j.still.2006.03.001

Tappert, M.C., Rivard, B., Fulop, A., Rogge, D., Feng, J., Tappert, R., Stalder, R., 2015. Characterizing kimberlite dilution by crustal rocks at the Snap Lake diamond mine (Northwest Territories, Canada) using SWIR (1.90-2.36 μm) and LWIR (8.1-11.1 μm) hyperspectral imagery collected from drill core. Econ. Geol. https://doi.org/10.2113/econgeo.110.6.1375

Topp, G.C., Parkin, G.W., Ferré, T.P.., 2008. Soil Water Content, in: Carter, M., Gregorich, E. (Eds.), Soil Sampling and Methods of Analysis. CRC Press, Boca Raton, Florida, pp. 939–961.

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics, Fourth. ed. Springer, New York.

Viscarra Rossel, R.A., Lark, R.M., 2009. Improved analysis and modelling of soil diffuse reflectance spectra using wavelets. Eur. J. Soil Sci. https://doi.org/10.1111/j.1365-2389.2009.01121.x

Viscarra Rossel, R.A., Lobsey, C.R., Sharman, C., Flick, P., McLachlan, G., 2017. Novel Proximal Sensing for Monitoring Soil Organic C Stocks and Condition. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.7b00889

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma. https://doi.org/10.1016/j.geoderma.2005.03.007

Vitousek, P.M., Howarth, R.W., 1991. Nitrogen Limitation on Land and in the Sea : How Can It Occur ? Nitrogen limitation on land and in the sea : How can it occur ? 13, 87–115. https://doi.org/10.1007/BF00002772

Vitousek, P.M., Howarth, R.W., 1991. Nitrogen Limitation on Land and in the Sea : How Can It Occur? Biogeochemistry 13, 87–115. https://doi.org/10.1007/BF00002772

Wang, S., Zhuang, Q., Wang, Q., Jin, X., Han, C., 2017. Mapping stocks of soil organic carbon and soil total nitrogen in Liaoning Province of China. Geoderma. https://doi.org/10.1016/j.geoderma.2017.05.048

Wanit, S.P., Mcgill, W.B., Robertson, J.A., Thurston, J.J., 1994. Improved soil quality and barley yields and fababeans, manure, forages and crop rotation on a Gray Luvisol. Can. J. Soil Sci. 74, 75–84.

Wehrens, R., Buydens, L.M.C.L.M.C., 2007. Self- and Super-organising Maps in R: the kohonen package. J. Stat. Softw. 21, 1–19. https://doi.org/10.18637/jss.v021.i05

Xie, H.T., Yang, X.M., Drury, C.F., Yang, J.Y., Zhang, X.D., 2011. Predicting soil organic carbon and total nitrogen using mid- and near-infrared spectra for Brookston clay loam soil in Southwestern Ontario, Canada. Can. J. Soil Sci. https://doi.org/10.4141/cjss10029