DIGITAL SOIL MAPPING TO ENHANCE CLIMATE CHANGE MITIGATION AND ADAPTATION IN THE LOWER FRASER VALLEY USING REMOTE SENSING

by

Siddhartho Shekhar Paul

B.Sc. (Honors), University of Dhaka, Bangladesh, 2008
M.Sc., University of Dhaka, Bangladesh, 2009
M.Sc., University of Northern British Columbia, Canada, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Soil Science)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2020

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Digital soil mapping to enhance climate change mitigation and adaptation in the Lower Fraser Valley using remote sensing

submitted by Siddhartho Shekhar Paul in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Soil Science

Examining Committee:

Sean Smukler
Supervisor

Nicholas Coops
Supervisory Committee Member

Maja Krzic
Supervisory Committee Member

Cindy Prescott
University Examiner

Les Lavkulich
University Examiner

Additional Supervisory Committee Members:

Mark Johnson
Supervisory Committee Member

Supervisory Committee Member

Abstract

Globally, the agriculture sector is constantly being challenged by multiple climate change-induced stresses while agricultural activities are responsible for a large portion of global greenhouse gas emissions. At the same time, agroecosystems have a sizable potential to mitigate climate change through the sequestration of atmospheric carbon-dioxide as soil organic carbon (SOC); a key soil quality parameter that can also enhance climate change adaptation. Although the dual benefits of SOC are well established, intensive agricultural production and associated land use/land cover (LULC) changes continue to drive large declines in SOC. Alternatively, sustainable LULC practices can potentially reverse this trend and improve SOC stocks. Digital soil mapping (DSM) using remote sensing can help elucidate SOC dynamics associated with LULC change and agricultural management practices by producing spatially explicit information on SOC at the field- and landscape-scales. In this research, I developed and applied innovative DSM techniques to study the spatiotemporal changes in SOC and related soil properties in the Lower Fraser Valley (LFV), one of the most intensive agriculture regions of British Columbia, Canada. At the field-scale, I evaluated various sampling strategies for DSM using unmanned aerial vehicle imagery, mid-infrared spectroscopy and geostatistical models to identify the most cost-effective approach. At the landscape-scale, using Landsat satellite imagery and machine learning tools, I produced maps of soil workability thresholds (WT) for the agricultural lands in Delta and then, assessed the SOC dynamics across the entire LFV since 1984. My analysis identified that 40% of Delta's agricultural lands had a WT of <30%, making them extremely vulnerable to the shifting precipitation patterns expected for the region. In addition, 61% of LFV lost SOC, 12% of the region gained SOC, while 27% remained unchanged between 1984 and

2018. Areas that lost the most SOC were those that had experienced changes in LULC; however, I concluded the majority of SOC loss occurred due to agricultural practices. The dissertation contributes to devising cost-effective approaches to quantify and monitor changes in SOC at the field- and landscape-scales that can help in the development of effective agricultural climate change mitigation and adaptation strategies.

Lay Summary

Agricultural producers have the potential to enhance their climate change mitigation and adaptation capacity through the sequestration of atmospheric carbon dioxide as soil organic carbon. In this dissertation, I sought to improve the understanding of soil organic carbon dynamics at the field- and landscape-scales in the Lower Fraser Valley of British Columbia using remote sensing-based digital soil mapping techniques. At the field-scale, I identified the geospatial statistical approach that provides the most cost-effective mapping of soil properties. At the landscape-scale, I observed an overall decline in soil organic carbon in most parts of the region. This research provides baseline information on the status of soil organic carbon and the rate of changes since 1984. The dissertation contributes to devising cost-effective strategies to monitor changes in soil organic carbon for enhancing climate change mitigation and adaptation in agricultural landscapes.

Preface

I designed this dissertation to include three manuscripts, prepared for publication in peer-reviewed journals, in chapters 2-4. I present a general introduction (Chapter 1) and a general conclusion (Chapter 5) to tie them together. Chapters 2-4 were slightly modified from the original manuscript version for improved readability and minimize repetitions. I developed the overarching goal and specific research objectives of this work in consultation with my supervisory committee. I designed the methodology, conducted the field data collection and data analysis, and wrote the manuscripts while all the co-authors of each manuscript, as listed below, were consulted for critical feedback. In addition, Chandna, A. in Chapter 3 and Lyndsey, D. in Chapter 4 contributed to the field data collection and laboratory analysis. Besides, this work was supported by the research assistants at the Sustainable Agricultural Landscapes Lab, led by Dr. Sean Smukler.

- **Chapter 2:** Paul, S. S., Coops, N. C., Johnson, M., Krzic, M., & Smukler, S. M. (2019). Evaluating sampling efforts of standard laboratory analysis and mid-infrared spectroscopy for cost effective digital soil mapping at field-scale using unmanned aerial vehicle imagery. Geoderma, 356. doi: 10.1016/j.geoderma.2019.113925. (Published)

- **Chapter 3:** Paul, S. S., Coops, N. C., Johnson, M., Krzic, M., Chandna, A., & Smukler, S. M. (2020). Mapping soil organic carbon and clay using remote sensing to predict soil workability for enhanced climate change adaptation. Geoderma. (Published)

- **Chapter 4:** Paul, S. S., Lyndsey, D., Coops, N. C., Johnson, M., Krzic, M., & Smukler, S. M. (2020). Tracking changes in soil organic carbon across the heterogeneous agricultural landscape of the Lower Fraser Valley of British Columbia. (In review)

Table of Contents

List of Tables

List of Figures

List of Equations

List of Abbreviations

AHM – Annual heat-moisture index

AHS – Analytical hill-shading

ALUI – Agricultural land use inventory

AC – Annual crop

BC – British Columbia

BUBL – built-up or bare land

C – Carbon

CA – Catchment area

CCC – Lin's concordance correlation coefficient

CD – Closed depressions

CEC – Cation exchange capacity

CI – Convergence index

CL – Clay

CC – Clear cut

cLHS – Conditioned Latin Hypercube Sampling

CO_2 – Carbon dioxide

CNBL – Channel network base level

CSC – Cross sectional curvature

CV – Coefficient of variation

DEM – Digital elevation model

DGPS – Differential Global Positioning System

DI – Disturbance index

DSM – Digital soil mapping

EC – Electrical conductivity

EV – Experimental variogram

FFP – Forest or forest patches

GBM – Generalized Boosted Regression Model

GHG – Greenhouse gas

GL – Grassland

GLCM – Grey level co-occurrence matrix

IDW – Inverse Distance Weighting

IncMSE – Increase in mean square error

KED – Kriging with external drift

LC – Longitudinal curvature

LFV – Lower Fraser Valley

LPL – Lower plastic limit

LS – Slope-length factor

LULC – Land use/land cover

MAP – Mean annual precipitation

MAT – Mean annual temperature

ML – Machine learning

MIRS – Mid-infrared spectroscopy

MRRTF – Multiresolution index of the ridge top flatness

MRVBF – Multiresolution index of valley bottom flatness

NIR – Near infrared

NOpen – Negative topographic openness

nRMSE – Normalized root-mean-square error

PC – Perennial crop

POpen – Positive topographic openness

PTFs – Pedotransfer functions

RF – Random forest

RGB – Red-green-blue

RMSE – Root mean square error

RSP – Relative slope position

SOC – Soil organic carbon

SOM – Soil organic matter

SLA – Standard laboratory analysis

SWIR – Short wave infrared

TC – Total curvature

TCT - Tasseled cap transformation

TN – Total nitrogen

TRI – Terrain ruggedness index

TRIM – Terrain Resource Information Management

TS – Time step

TWI – Total wetness index

UAV – Unmanned aerial vehicle

USGS – United States Geological Survey

UPL – Upper plastic limit

VD – Valley depth

VI – Variable importance

WL – Wetland

WT – Soil workability threshold

Acknowledgements

The research in this dissertation was made possible with generous funding support from multiple organizations, including Mitacs Accelerate Program, British Columbia Ministry of Agriculture, British Columbia Climate Action Initiative, Delta Farmland and Wildlife Trust, Delta Farmers Institute, and the Natural Sciences and Engineering Research Council of Canada.

I am profoundly indebted to my supervisor Dr. Sean Smukler for his endless guidance, advice, and encouragement during this challenging PhD journey. He always shared his own experiences to inspire me, to alert me, and to make me successful in this endeavor. I am also extremely grateful to my other supervisory committee members – Dr. Mark Johnson, Dr. Maja Krzic, and Dr. Nicholas Coops for their consistent support and intellectual input throughout my PhD.

I would like to acknowledge the continuous and untiring support of my coworkers and undergraduate research assistants at the Sustainable Agricultural Landscapes Lab. I could not finish the research without this immensely helpful team. I want to thank all the farmers who allowed access to their fields for sampling and supported this research by sharing their knowledge. Special thanks to my peers, colleagues, and professors at the Soil Science Group in the Faculty of Land and Food Systems for their support, friendship, and giving me an enjoyable graduate student experience.

To all my teachers since childhood, I am grateful to you for educating me and bridging the path which led me to the last stage of my formal education. And most importantly, to my family, thank you for being with me during this difficult journey and always giving me the opportunity to prosper.

Dedication

To the brave people of Bangladesh

who taught me how to win against all odds and how to survive in the most difficult condition

And

To the amazing and dearest women in my life

who endured and sacrificed to give me love, comfort, and prosperity

Chapter 1: Introduction

1.1 Background

1.1.1 Soil organic carbon

The agricultural sector, globally, has severely altered the Earth's natural landscape through land use change and intensive agricultural practices. At the same time, agriculture is being challenged by multifaceted climate change-induced stresses. Of major concern is the massive reductions in soil organic carbon (SOC) (Lal, 2004; Sanderman et al., 2017), which impairs agricultural productivity and contributes to climate-forcing greenhouse gas (GHG) emissions (Paustian et al., 2016). As such, reversing this trend by sequestering carbon in agricultural soil has been identified as a key solution to reduce climate change-induced stresses and improve food production in the coming decades (Minasny et al., 2017; Rumpel et al., 2018).

Globally, the carbon associated with the SOC is larger than carbon stored in terrestrial vegetation and the atmosphere combined (Stockmann et al., 2013). SOC is a critical component of soil quality and provides resilience in agricultural systems against climatic perturbations, such as droughts or excessive rainfall (Kononova, 2013). SOC enhances nutrient cycling, improves soil water retention, and influences drainage – all of which are essential for agricultural productivity. However, since the widespread adoption of agriculture to produce food, a sizable amount of SOC has been lost to the atmosphere through land cover change. Land use/land cover (LULC) changes from natural ecosystems (e.g. prairies, wetlands, and forests) to intensive agriculture can result in a 20-60% loss of SOC (Lal, 2004). Improved land management practices, nevertheless, can prevent or reverse these losses and mitigate GHG emissions. Management practices like conservation or no-tillage, winter cover cropping, crop rotations, maintaining perennial grass margins and hedgerows, agroforestry or shifting from annual crops to perennial crops can substantially reduce GHG emission from agricultural lands and remove atmospheric CO_2 by sequestering carbon in the soil (Lal et al., 2015; MacHmuller et al., 2015). Although there has been substantial work to quantify the impacts of LULC change and agricultural management on SOC, there are still major gaps in our understanding of the impacts of many types of change in LULC or management practices and how these vary across environmental conditions.

Researchers have observed contrasting effects of global climate change on SOC and it is unclear how LULC change and agricultural management will interact with rising temperatures and atmospheric CO_2 concentrations. For example, Crowther et al. (2016) reported negative impacts of global warming on global SOC stock, while other research has suggested that higher soil microbial activity and plant productivity in warming conditions may result in increased SOC stocks (Koven et al., 2015; Todd-Brown et al., 2014). Thus, understanding SOC dynamics in response to changes in LULC and agricultural management, in the context of climate change is critically important for identifying effective SOC management.

1.1.2 Climate change, agricultural LULC, and SOC in the Lower Fraser Valley

The GHG emissions in the province of British Columbia (BC) increased by 17% between 1990 and 2015 mainly as a result of increased methane from livestock and nitrous oxide from fertilizer and manure use, making BC one of the top five GHG emitting provinces of Canada in 2015 (Environment and Climate Change Canada, 2017). Although the agriculture sector in BC has been estimated to be responsible for 3% of the provincial emissions (BC MOEnvironment, 2014), these figures, however, do not account for changes in SOC either from LULC change or management practices. Additionally, climate models have predicted considerable changes in temperature and precipitation patterns across the province. For example, in the Lower Fraser Valley (LFV), which is one of the most agriculturally intensive regions of BC, a 1.8ºC increase in annual average temperature and 7% rise in annual precipitation was predicted by 2050s (BC Agriculture & Food Climate Action Initiative, 2015). Much of this precipitation is expected to occur in the form of rain during the winter, spring, and autumn causing issues for agricultural drainage management in the wet season and most importantly having an impact on the shoulder seasons when planting and harvest operations are reliant on dry soil. Furthermore, summers are expected to become hotter and drier. As a result, effective soil management in this region has become increasingly challenging for farmers. Soil properties can be highly spatially heterogeneous, and some are also changing rapidly over time, which in turn makes characterizing soil properties challenging. Knowing precise and spatially explicit information of various soil properties across a farm field could enable, for example, more effective site-specific management of fertilizer applications, or the installation of drainage infrastructure, which in turn could lead to increased SOC. Moreover, the precision management of SOC at individual farm fields may provide greater benefits in terms of region-wide climate change mitigation.

Currently, the LFV of BC is one of the most intensive agricultural regions in the province and is responsible for more than 50% of the province's annual gross farm receipts (Crawford and MacNair, 2012). The LFV; however, was once dominated by coastal rain forest and wetlands and beginning in the late 1800s has seen large scale LULC change to agricultural production (Boyle et al., 1997). Particularly, over the past few decades, the LFV has experienced substantial agricultural expansion and consequent LULC changes. The conversion from ecosystems such as forests, wetlands, and grasslands to high input, heavily mechanized agricultural production has likely caused a severe decline in SOC. If these losses in SOC are taken into consideration, the estimation of provincial agricultural GHG emissions could be substantially higher than what is currently estimated (i.e. 3% of the total provincial emissions), which has serious implications for managing provincial GHG emissions as well as the long-term productivity of the soil. Loss of SOC is a prime concern for soil productivity because of its contribution to the formation of soil aggregates and associated influences on the soil structure and related properties (such as bulk density, water retention, and water movement) (Merante et al., 2017). As a result, the decline in SOC will adversely affect key soil functions including reducing the soils' capacity of ameliorating excessive rain derived drainage issues in wet periods and enhancing water holding capacity during dry periods. With the predicted changes in climate conditions, these effects on soil functioning may have significant consequences on agricultural productivity in the LFV.

While understanding the relationship between LULC change and SOC dynamics is critical for developing effective management strategies to maintain SOC, the rate, direction, and magnitude of the SOC change are still not clearly understood in the LFV. Previous research conducted by Boyle & Lavkulich (1997) investigated the carbon pool dynamics in the LFV from 1827 to 1990 using land use maps, air photographs, and previously published data. They estimated a total loss of SOC was 125 Mt from 1827 to 1990. This was a well-reasoned estimate of SOC losses in the LFV, yet the analysis had some important limitations. This research did not specifically focus on the SOC dynamics rather it focused on the changes in biomass carbon and made assumptions about the relative changes in SOC associated with the above-ground carbon. Moreover, since

1990, there have been considerable changes in both LULC and agricultural management in the LFV and the associated impacts on SOC are unknown. An alternative approach to using historical land use maps, based on analysis of archived time-series remote sensing data (e.g. Landsat satellite imagery) and digital soil mapping (DSM), could better fill important gaps in our understanding of LULC change, agricultural management, and SOC dynamics in the LFV. While this approach would not be able to capture the changes occurred before 1972 (when the satellite data came on-line), this time-series remote sensing-based change analysis could be far higher resolution and more accurate.

As explained above, changes in SOC directly impact several fundamental soil functions, especially the capability of the soil to provide efficient drainage during wet seasons. This key soil function has important implications for some of LFV’s highly productive low-lying agricultural lands characterized by fine-textured soil and poor drainage. Heavy rainfall in the shoulder seasons is resulting in overly saturated soils and delaying the field preparation and harvesting with heavy machinery in these agricultural lands (Neufeld et al., 2017). The use of heavy farm machinery during such soil conditions may cause soil compaction and degrade soil structure. Rainfall thus can restrict the number of days these soils are workable with heavy machinery without negatively impacting soil quality. Prior site-specific knowledge of the optimum soil water content for working the soil could be of value for planning the timing of farm management activities to minimize damage to soil structure. Such information, however, is not currently available for the LFV. Soil maps of the region developed in the 1980s describe a drainage class, but given that workability is a function of drainage, soil texture, and soil organic matter, these maps have limited utility to describe soil organic matter changes with time.

Furthermore, these maps are categorical with a spatial resolution that generally precludes field management decisions. Provided that the optimum water content is highly reliant on soil organic matter, which is largely determined by SOC (Obour et al., 2017), the DSM technique can be applied to develop high-resolution maps of areas with the greatest potential for soil moisture related compaction. Hence, understanding SOC dynamics will also help develop climate change adaptation strategies in areas with poor drainage.

1.1.3 Modeling spatial and temporal dynamics of SOC and related soil properties

Due to their highly heterogeneous nature, soil properties are challenging to quantify and evaluate in spatial and temporal contexts. For dynamic soil properties (e.g. SOC, total nitrogen), the primary modeling approach for spatial and temporal assessment has been reliant on mechanistic models (e.g. Century, DNDC). These types of models have been the mainstay for modeling past and future SOC dynamics in agricultural landscapes across the globe. However, these models are data-intensive and for accurate prediction, requiring a large amount of historical input data, which are expensive to gather and often not available. As a result, such modeling efforts use a large number of assumptions or use datasets which may not be at a fine enough scale to make accurate or relevant predictions for heterogeneous soil properties (Grinand et al., 2017). A pragmatic alternative could be 'scorpan' based DSM models, proposed by Minasny et al. (2013), which produce static-empirical prediction based on locally calibrated models using remote sensing data. According to 'scorpan', soil class or attribute that will be modeled is a function of soil (i.e. other or previously measured soil properties), climate (i.e. climate variables), organisms (i.e. LULC), relief (i.e. topographic attributes), parent material (i.e. geological materials on which soil formed), age (i.e. time factor), and spatial or geographic position. Incorporation of remote sensing derived environmental variables reduces the cost of analysis and allows the DSM model to account for the spatiotemporal variation in the soil landscape. Although the DSM approach has been widely used for mapping soil properties at a single time-step (McBratney et al., 2003), this practical approach has not been utilized with full potential for assessing spatiotemporal dynamics of SOC in heterogeneous agricultural landscapes (Yigini and Panagos, 2016). Moreover, integration of the outputs from locally calibrated high spatial resolution land use change model and downscaled climate prediction may enhance the capability of the DSM models for the spatiotemporal projection of SOC dynamics, but such integration has not been recognized or undertaken in the literature. Modeling other related soil properties, like optimum soil water content using remote sensing and DSM has also not been explored, although such modeling can be of critical value for agricultural soil management in areas with poor drainage.

As reported in various DSM studies, integration of remote sensing imagery reduces the overall cost of analysis for mapping various soil properties, however, the cost largely depends on the type and spatial resolution of the remote sensing data in use (McBratney et al., 2003; Mulder et al., 2011). Understanding the variation in soil properties at both field- and landscape-scales is important for efficient management and decision-making since the field-scale variation is mainly attributed to agricultural management practices while the landscape-scale variation may involve both LULC changes and management practices. Accurate soil mapping at field-scale requires high spatial resolution remote sensing imagery (i.e. ≤5 m spatial resolution) to capture the large variability in soil properties (Malone et al., 2013a; Minasny and McBratney, 2016). Although remotely sensed satellite imagery with the high spatial resolution is relatively expensive, images acquired using unmanned aerial vehicles (UAV) can be an effective alternative for field-scale

DSM. Recently, UAV imagery has gained popularity for agroecological monitoring due to its high spatial resolution, cost-effectiveness, and flexibility of image acquisition (Mogili and Deepak, 2018; Tsouros et al., 2019). Although having these practical benefits, the cost-effectiveness and accuracy of field-scale DSM models using UAV imagery has remained unclear and warrants further research to enhance the utility of UAV imagery in DSM. At a landscape-scale, however, acquiring UAV data is not yet cost-effective and requires the use of satellite imagery. Landsat satellite imagery, for example, has a spatial resolution of 30 m, 16 days revisit time, and a data archive at those resolutions from 1984 (poorer spatial resolution and slightly longer temporal resolution since 1972) has been shown to be an effective solution for accurate mapping of different soil properties at landscape-scale, especially for assessing the spatiotemporal dynamics. Thus, incorporation of time-series Landsat imagery in static-empirical 'scorpan' based DSM model can provide an innovative approach for modeling the changes in dynamic soil properties, like SOC across space and time.

Regardless of the type of imagery or the spatial scale, a key challenge that must be first addressed by any digital soil methodology is to identify the optimum number of point samples and modeling strategy required for predicting the continuous soil surfaces with the desired accuracy and cost. In heterogeneous agricultural lands, capturing the spatial variability is critical for producing accurate soil maps at both field- and landscape-scales. Increasing the number of samples improves the accuracy of the models but it also increases the cost and effort required for field data collection and analysis. Developing a sampling strategy that deals with this trade-off between accuracy and cost is largely dependent on the type of laboratory analysis used for soil samples and the type of DSM model. Standard laboratory analysis (SLA) is the conventional and the most accurate approach for soil sample analysis, however, SLA can be a substantial portion of the cost of DSM. In contrast, mid-infrared spectroscopy (MIRS), which predicts soil data based on the spectral responses of SLA samples, provides a cost-effective alternative (Nocita et al., 2015) and has recently been widely used in DSM research. Although MIRS is not as accurate as SLA, it can substantially increase the number of soil samples utilized for DSM at an equivalent cost, thus potentially improving model accuracy. Although MIRS is widely adopted there have been a few studies comparing DSM outcomes derived using the SLA and MIRS soil dataset for investigating the trade-offs in accuracy and cost of DSM. In addition, there has been little validation of DSM models based on MIRS using an independent set of SLA samples. Most DSM studies using the MIRS dataset validate the model performance against another set of MIRS data, which incorporates more uncertainties in the predicted digital soil maps. Furthermore, DSM researchers use a large number of modeling techniques, including generalized linear models, classification and regression trees, neural networks, fuzzy systems, and geostatistics to analyze point soil data and produce digital soil maps (McBratney et al., 2003). The most widely used DSM techniques are decision tree models and geostatistical models, both of which have been subdivided into a few other categories. Thus, selecting a single model from this large number of available options is challenging. Comparative studies using the most commonly used models and investigating for the accuracy at different scales from field to landscape may provide valuable information for future DSM research.

1.2 Research objectives

There are important knowledge gaps related to SOC dynamics and associated soil properties under changes in LULC and climate conditions in the LFV and there is a need to establish new and cost-effective tools to address these gaps. Therefore, in my dissertation, I sought to develop innovative DSM using remote sensing approaches to model SOC and other related soil properties for enhancing climate change mitigation and adaptation in the LFV. The specific research objectives of my dissertation were to:

1) Evaluate the cost and accuracy of geostatistical modeling technique at varying sampling efforts of SLA and MIRS for DSM at the field-scale

2) Compare two machine learning approaches for DSM of SOC and clay (CL) to predict a soil workability threshold (WT) at the landscape-scale

3) Assess the spatiotemporal dynamics of SOC in response to historical changes in LULC, agricultural management, and climate conditions

Each of these objectives will be presented as an individual chapter of my dissertation. Through addressing these objectives, I intended to develop innovative and cost-effective tools for mapping different soil properties at the field- and landscape-scales for enhanced agricultural climate change mitigation and adaptation.

1.3 Overview of the dissertation

I have organized this dissertation into five chapters. I have presented an overall introduction here in Chapter 1 followed by three separate research papers presented as individual chapters that addressed each of the specific research objectives mentioned above. In Chapter 5, I presented my general conclusions. Figure 1 shows the outline of the dissertation chapters.

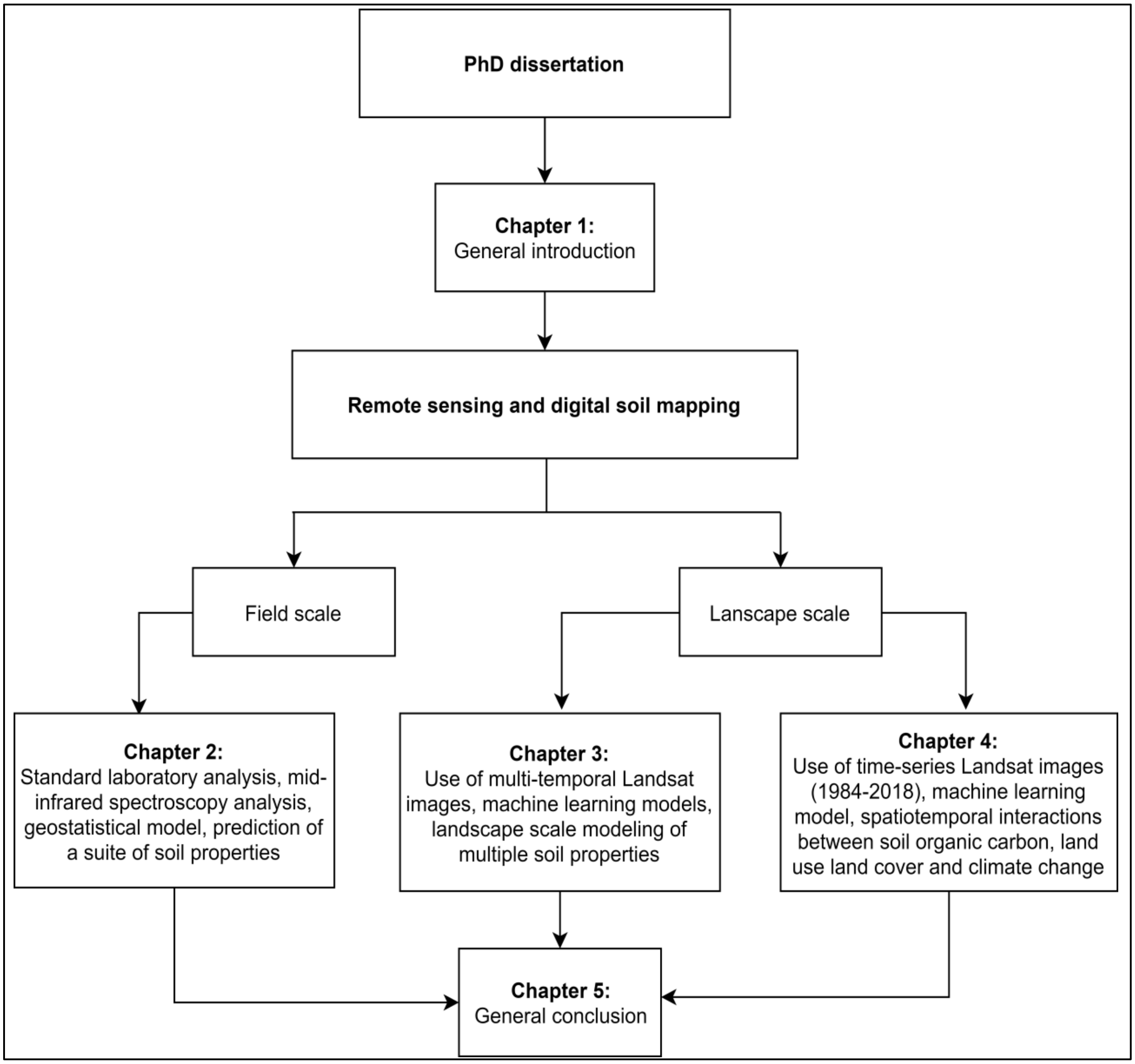

Figure 1: Outline of the dissertation chapters

In Chapter 2 of my dissertation, I evaluated different sampling strategies and analysis techniques for high spatial resolution mapping of a suite of soil properties at the field-scale. I compared between the sampling efforts of SLA and MIRS to identify the most effective sampling efforts at which model accuracy and cost of analysis were optimized. This chapter provides cost-effective strategies for precision agricultural management which is highly important for efficient fertilizer application, organic matter management, drainage and salinity management, all of which contribute to both climate change mitigation and adaptation at a farm level.

In Chapter 3, I compared two machine learning models, random forest and generalized boosted regression model for predicting soil properties at landscape-scale. To do this I utilized multi-temporal Landsat satellite imagery to capture seasonal variation across the agricultural landscape of Delta, BC and predicted maps of SOC and clay using the machine learning models. I applied multiple pedotransfer functions to these predicted maps to produce maps of soil plasticity limits and identify spatially explicit values for optimum soil water content as a threshold for workability. The outcomes of this chapter will help inform effective soil management strategies under expected shifting precipitation patterns and enhance climate change adaptation capability across the agricultural landscape.

In Chapter 4, I applied time-series Landsat satellite images and machine learning models to assess spatiotemporal interactions between SOC, LULC, and climate change (1984 – 2018) in the LFV. I employed a static-empirical approach and calibrated a SOC model using a field dataset I collected in 2018 and then, predicted SOC for the previous years (i.e. 1984, 1990, 1999, and 2009) by updating the dynamic environmental variables. I also used the same Landsat imagery for producing LULC maps for each time step through a combined pixel- and object-based approach. I then evaluated the spatiotemporal relationship between SOC, LULC change, agricultural management, and climatic variables. The results of this analysis identify potential areas and management strategies for SOC sequestration which provides benefits for both climate change mitigation and adaptation across the LFV.

Figure 2 shows schematically the field data collection, analysis, and overall outcomes of the research. Field data collection included soil sampling and collecting ground truth information on LULC types.

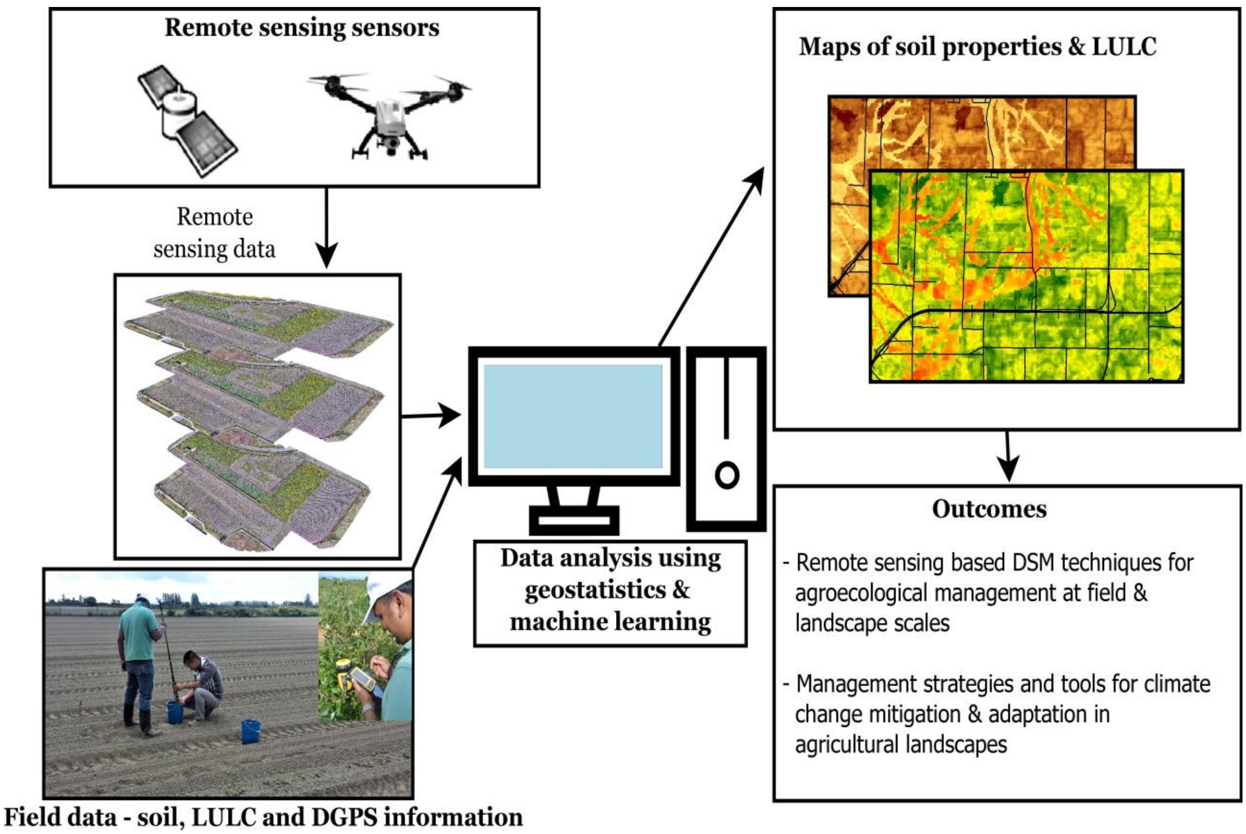

Figure 2: Schematic diagram showing data collection, analysis, and outcomes of research

DGPS, LULC, and SOC refer to Differential Global Positioning System, land use/land cover, soil organic carbon, respectively

Chapter 2: Evaluating sampling efforts of standard laboratory analysis and mid-infrared spectroscopy for cost-effective digital soil mapping at field-scale

2.1 Chapter introduction

Digital soil mapping (DSM) is increasingly being used for managing and monitoring a wide range of soil-derived ecosystem services, including the provisioning of food, fiber and fuel, carbon sequestration and nutrient cycling. DSM combines information from sparsely populated point soil data with geospatial data, such as remotely sensed imagery, to provide continuous predictions of soil properties (Lagacherie, 2008; Li and Heap, 2011). DSM produces seamless spatial interpolation of point soil information at scales ranging from broad global maps to fine scale maps of individual farm fields (Grunwald et al., 2011; Malone et al., 2017). Detailed knowledge of soil properties at different scales can help land managers to make spatially explicit management decisions (Cruz-Cárdenas et al., 2014). Given that soil properties exhibit high spatial heterogeneity, mapping at finer spatial scales may be critical to meet specific farm management objectives, especially for precision agriculture (Malone et al., 2013; Suk Lee and Ehsani, 2015).

Fine scale DSM requires closely spaced point information (Hengl et al., 2004); however, there is no consensus regarding the sampling effort required for the optimum performance of the spatial prediction of soil properties (Brungard and Boettinger, 2010; Ließ, 2015). The number of samples, sample spacings, and the actual locations of the samples are all factors related to the sampling effort that influence the prediction process (Zhu and Lin, 2010). Enhancing the sampling effort by adding more samples will improve the accuracy of the predicted output, but this also increases the time, cost, and data processing required. Most sampling designs for DSM either aim to achieve a well-distributed spatial coverage of the area or to capture the spatial variations of the feature space (Minasny and McBratney, 2006). A number of studies on precision agriculture have explored this at the field-scale and (Kerry et al., 2010) found that a sampling interval of 100-120 m can provide adequate spatial coverage and precise soil management of a farm field. In order to achieve a well-distributed spatial coverage, the sampling effort may be increased but this does not necessarily result in accurate predictions. Alternatively, an optimum sampling effort may be obtained where the spatial variations of the study site can be effectively captured (Brungard and Boettinger, 2010). After reaching the optimum number of samples, increasing the number of samples will not improve the prediction capability of the model; rather, additional samples will result in diminishing returns in terms of improved accuracy of the model. Thus, a spatially optimized sampling effort will provide the most benefits in terms of both prediction accuracy and sampling investments.

The efficacy of the sampling design is likely dependent on the statistical model used for predicting the soil properties. Many studies have suggested that geospatial environmental variables are important for capturing spatial variations and improving the prediction accuracy of the models (Li et al., 2015). Kriging with external drift (KED) is one of the commonly used hybrid geostatistical models which assumes that the value at any given point is spatially dependent on the values of the neighboring points but the variation trend or drift is determined externally as a linear function of a group of ancillary environmental variables (Keskin and Grunwald, 2018; Wackernagel, 2003). KED, a straightforward approach where the trend and residuals are estimated as part of a single system, has been successfully used for a number of DSM studies where it obtained similar or better accuracies than simpler kriging models, like ordinary kriging (Li, 2010) or more complex and newer hybrid models, like regression kriging (Santra et al., 2017). Thus, KED can be used as an effective technique for predicting a suite of soil properties. The outcomes of KED prediction derived using various sampling designs can then be compared to identify the most effective sampling effort.

Standard laboratory analysis (SLA) of the soil samples can be a substantial portion of the overall DSM expense. Recent advances in soil analysis using mid-infrared spectroscopy (MIRS) have shown promise to reduce costs compared to SLA (Nocita et al., 2015). MIRS can produce fast and relatively inexpensive predictions of soil properties, that, although they are improving, are not as accurate as SLA as they are derived from SLA predictions (Viscarra Rossel et al., 2006). Larger sampling efforts, in general, better explain the spatial variability of the soil properties across a study site (Brus and Heuvelink, 2007). Thus, MIRS techniques, which allow the addition of more sample points for a given budget, may make up for their reduced accuracy when producing predictive DSMs. Although MIRS techniques have recently been successfully used for landscape-scale DSMs (Cobo et al., 2010; Vågen et al., 2016; Winowiecki et al., 2016), its performance in conjunction with field-scale DSM for a suite of soil properties is less clear. At the field-scale, soil properties exhibit fine resolution spatial variations requiring a large set of soil data to produce accurate predictive maps. In precision agriculture, the use of visible- and near-infrared spectroscopy for on-the-go proximal soil sensing is widely used, however, use of MIRS is not common because of the high cost and lack of availability of portable MIRS instruments (Ge et al., 2011; Viscarra Rossel et al., 2006). Producing high resolution field-scale DSM using laboratory-based MIRS data could be an effective alternative to current proximal sensing approaches and/or demonstrate the utility of developing portable MIRS technology. However, there is a need to understand the trade-offs in accuracy and costs between using relatively more accurate, but more expensive SLA datasets and using comparatively less accurate, but less expensive MIRS datasets for DSM.

To address this research need, we conducted a study to produce digital maps of a suite of soil properties on a farm field in British Columbia at 5 m resolution using various approaches. The specific objectives of this study were to: (1) compare a range of equivalent sampling efforts (based on their cost for fieldwork and lab analysis) of SLA and MIRS in terms of their relative accuracy for predicting a suite of soil properties, including sand, silt, clay, pH, salinity, soil organic matter (SOM), and total nitrogen (TN) and (2) assess the trade-offs between cost and accuracy to determine the most effective sampling effort for producing predictive maps of these soil properties.

2.2 Materials and methods

To develop DSMs of selected soil properties for a farm field in western Fraser Valley of British Columbia, we used a combination of methods (Figure 3). We sampled soils using a grid design, then analyzed them using two different lab analysis approaches. We also applied multiple statistical and geostatistical tools to evaluate a range of sampling efforts (by pseudo-sampling from the full dataset) and the resulting DSM predictions for the soil properties.

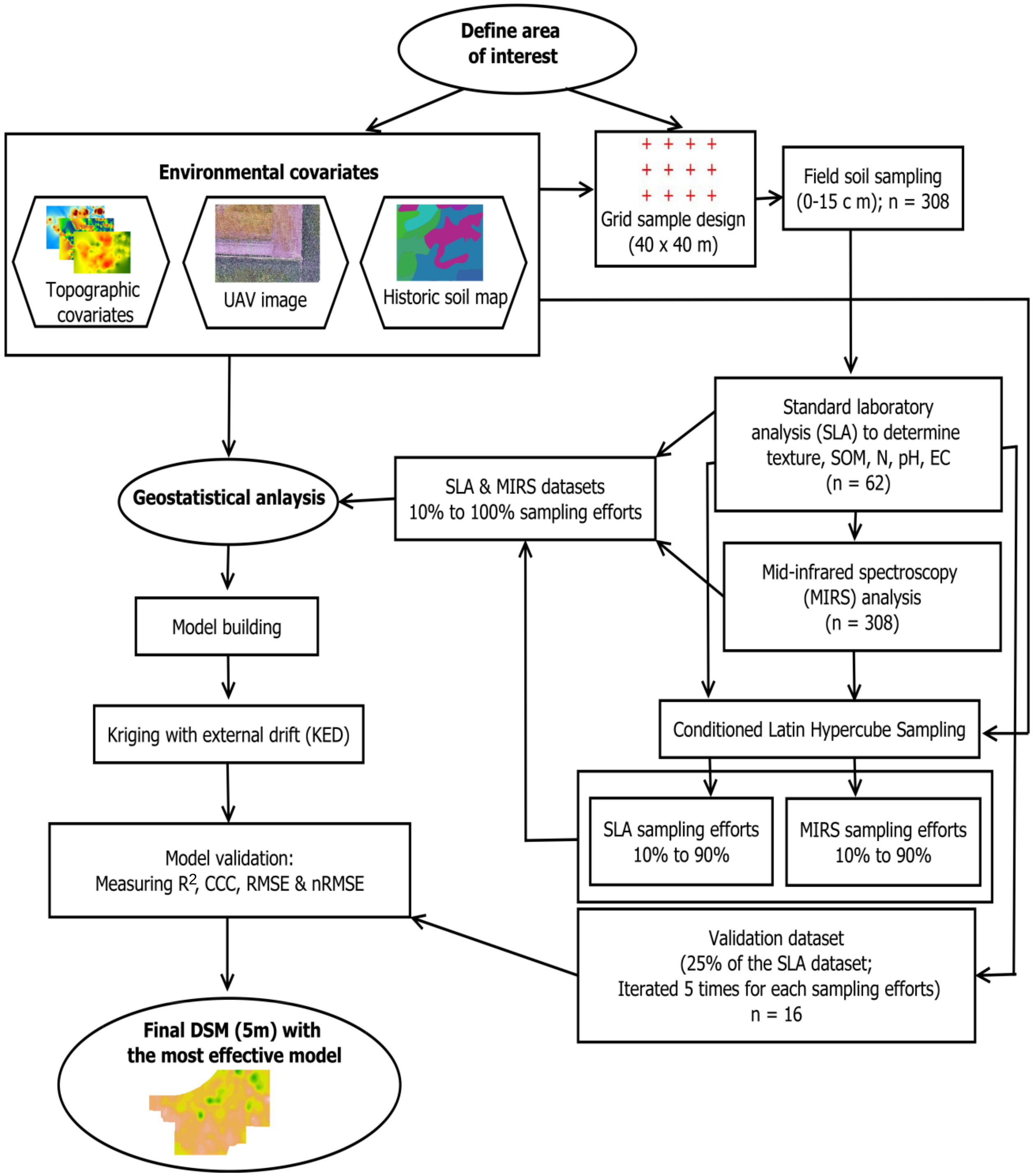

Figure 3: Flowchart showing the methods utilized for producing digital soil maps (DSM) in this study

UAV refers to the unmanned aerial vehicle; R^2, CCC, RMSE, and nRMSE refer to the coefficient of determination, Lin's concordance correlation coefficient, root mean square error, and normalized root-mean-square error respectively.

2.2.1 Study Site

The study site was a 54-hectare agricultural field near the City of Delta, British Columbia, Canada (49.08 N, 123.06 W), about 25 km south of the City of Vancouver (Figure 4). The field had known salinity and drainage problems at the time of sampling and was used for organic vegetable production. The study site was located on Rego Gleysol and Orthic Humic Gleysol (Umbric Gleysol) formed predominantly from fluvial parent materials. The study site is in the Fraser River delta and close to the ocean with elevation ranges from 1.25 to 1.70 meters above mean sea level. This area is characterized by a humid maritime climate with a mean annual temperature of 11.1º C and a mean annual precipitation of 928 mm based on 30-year climate records (Environment Canada, 2019).

2.2.2 Soil Sampling and Analysis

After reviewing the existing soil map (Luttmerding, 1981) and conducting preliminary field observations, a 40 x 40 m grid was developed for soil sampling (Figure 4). In 2015, a total of 308 points were sampled at the 0-15 cm depth across the field. All 308 sample locations were recorded with a GNSS Pro 6H Differential Global Positioning System (DGPS) (Trimble Inc., Sunnyvale, California, USA) with post-processing accuracy ranging from 10 - 50 cm. We derived a sub-sample set of the original 308 locations to use for SLA, and these were generally 120 m apart (compared to the original 40 x 40 m grid) but differed to a limited extent for some locations due to field edges. We included a few additional randomly selected grid locations for SLA analysis so that we achieve an equivalent cost for both SLA and MIRS. In total, 62 of the 308 original samples were retained for the SLA dataset.

All 308 samples were air-dried and sieved to <2 mm. The 62 SLA samples were sent to the Technical Service Laboratory of British Columbia Ministry of Environment for particle size analysis using the hydrometer method and for soil organic carbon (SOC) and TN using the combustion elemental analysis with a Vario EL Cube Elemental Analyzer (Elementar, Langenselbold, Germany). Separate aliquots of the subset of SLA samples were also analyzed at the University of British Columbia lab facility to measure pH in a 1:1 soil-water ratio, and electrical conductivity (EC) in a 1:2 soil-water ratio using an Oakton PC 700 pH/conductivity meter. A conversion factor of 5 was applied to compare our EC values with the values measured using a more typical and expensive 'saturated paste' method (Hanlon et al., 1993). The full 308 sample set was then analyzed with MIRS using a TENSOR 37 spectrometer (Bruker Instruments, Ettlingen, Germany). For MIRS analysis, the samples were prepared by oven drying at 105°C before grinding with a ball-mill. We then analyzed three 1g subsamples of each soil sample and recorded the MIRS spectral response. Later, we calibrated and validated the recorded spectra using the OPUS v7.2 Spectroscopy Software and Partial Least Squares Regression (PLSR) model where the SLA dataset served as the calibration (70%) and validation (30%) data. We performed a log ratio transformation of the texture data to achieve a combined composition of 100% for sand%, silt%, and clay% after PLSR prediction. We used isometric log ratio transformation for this purpose (Niang et al., 2014). We used the 'compositions' package (van den Boogaart and Tolosana-Delgado, 2008) within the R software (version 3.3.2, R Core Team, 2018) for log transformation. Finally, we multiplied the SOC data by 1.72 to compute SOM.

Figure 4: Location map of the study site (49.08 N, 123.06 W)

Map shows the standard laboratory analysis (SLA, n = 62) and mid-infrared spectroscopy (MIRS, n = 308) sample points following a 40 x 40 m grid. The SLA samples were also analyzed with MIRS.

2.2.3 Environmental Covariates

A total of 14 environmental covariates were utilized for this study (Table 1). A 5 m spatial resolution digital elevation model (DEM) was created using the point elevation data (n = 308) collected with the DGPS unit. We used the hydrologically correct DEM interpolation tool in ArcGIS 10.5 software for producing the DEM (Childs, 2004; Hutchinson, 1993). A group of topographic covariates was generated from this DEM using SAGA 2.1.2 software based on the work of Behrens et al. (2010), Lacoste et al. (2014), and Malone et al. (2009). The first and second derivatives, namely aspect, slope, multiresolution index of valley bottom flatness (MRVBF), multiresolution index of the ridge top flatness (MRRTF), positive and negative topographic openness, valley depth, terrain ruggedness index (TRI), total curvature, and total wetness index (TWI) were derived from the DEM.

In July 2016, an unmanned aerial vehicle (UAV) was flown over the study site to capture images in the visible bands (red-green-blue: RGB) of the electromagnetic spectrum. We used a DJI Matrice UAV which had a Zenmuse X3 CMOS sensor of 12.4 megapixels and a 20 mm focal length. The flight altitude was 30 m above ground level capturing images of 2 cm spatial resolution. We processed the images and resampled to derive two covariates at 5 m resolution from this RGB imagery – the green band reflectance was used directly and another covariate (averaged RGB band reflectance) was generated by averaging the reflectance of the three bands for each pixel (Amini et al., 2005; Levin et al., 2005). These two covariates were selected to utilize the variation in soil and vegetation color. The historic polygon-based soil map, developed in 1981 and comprised of only 3 soil classes (Luttmerding, 1981), was used to derive a raster layer of clay content at 5 m resolution to use as an additional covariate. The polygon map consisted of single-component map units for our study site. We extracted values from the polygon soil map using the 40 x 40 m sampling grid and then, utilized Inverse Distance Weighting (IDW) interpolation to produce a raster surface.

Table 1: Environmental covariates used for digital soil mapping in this study

Environmental Covariate Type	Input Representative Data	Source Data
Terrain	Digital elevation model (DEM) Aspect Slope Multiresolution index of valley bottom flatness (MRVBF) Multiresolution index of the ridge top flatness (MRRTF) Positive topographic openness Negative topographic openness Valley depth Terrain ruggedness index (TRI) Total curvature Total wetness index (TWI)	Digital Elevation Model
Vegetation & Management	Green band reflectance Averaged RGB band reflectance	Unmanned aerial vehicle image
Soil type	Clay raster surface	Historic soil map

2.2.4 Sampling Efforts and Conditioned Latin Hypercube Sampling

We used the Conditioned Latin Hypercube Sampling (cLHS) technique to develop several sampling efforts for both MIRS (n=308) and SLA (n=62) datasets using 10 - 90% of the total data points in 10% increments (Table 2). cLHS is a stratified random sampling technique that selects locations representing the spatial variability of the multiple input environmental covariates (Minasny and McBratney, 2006). The ‘clhs’ package (Roudier, 2014) within the R software (version 3.3.2, R Core Team, 2018) was used to design all sampling efforts. We utilized the environmental covariates listed in Table 1 for cLHS analysis. Sampling and analysis costs of each soil sample for both SLA and MIRS analyses were determined based on the cost of the external laboratory analyses, labor, and materials. The costs for SLA analyses, including the field sampling, were ~40 C$/sample whereas the cost of MIRS analysis including the field sampling was ~8 C$/sample. All laboratory fees are expressed on a cost-recovery basis.

Table 2: Number of samples and total cost of various equivalent sampling efforts of standard laboratory analysis (SLA) and mid-infrared spectroscopy (MIRS)

Sampling effort	Number of SLA samples	Number of SLA samples/hectare	Number of MIRS samples	Number of MIRS samples/hectare	Total cost (C$)
100%	62	1.15	308	5.70	2464
90%	56	1.04	277	5.13	2216
80%	50	0.93	246	4.56	1968
70%	43	0.80	216	4.00	1728
60%	37	0.69	185	3.43	1480
50%	31	0.57	154	2.85	1232
40%	25	0.46	123	2.28	984
30%	19	0.35	92	1.70	736
20%	12	0.22	62	1.15	496
10%	6	0.11	31	0.57	248

2.2.5 Prediction and Mapping

We used kriging with external drift (KED) to scale from point data to a continuous map of the entire field. The prediction using KED is based on the spatial correlation between the data points as well as the spatial information derived by the auxiliary environmental variables. As the name suggests, the drift is defined externally by the environmental variables rather than as a function of the coordinates of the data points (Wackernagel, 2003). In KED, prediction at an unknown location is derived by the (Equation 1) and (Equation 2) (Hengl, 2007):

$$\hat{z}(s_0) = \sum_{1=1}^{n} w_i(s_0).\ z(s_i) \qquad \textbf{(Equation 1)}$$

For

$$\sum_{1=1}^{n} w_i(s_0).\ q_k(s_i) = q_k(s_0);\ \ k = 1, 2, \ldots\ldots, p \qquad \textbf{(Equation 2)}$$

Where $\hat{z}(s_0)$ is the target soil property predicted at location s_0, q_k's are the environmental covariates, p is the number of environmental covariates. For kriging, it is critical to determine the spatial autocorrelation of the input data, i.e. semivariance which increases with distance. The distance where it stabilizes within the study area extent determines the range of the spatial autocorrelation (Malone et al., 2013b). We used the 'gstat' package for R software (Pebesma and Heuvelink, 2016) to perform KED interpolations. The KED model used the environmental covariates listed in Table 1. However, we performed a Pearson correlation analysis to evaluate the relationship between the target soil property and environmental covariates. Then, in KED prediction, we only included the variables which are highly correlated ($r \geq 0.20$ or $r \leq -0.20$) with the target soil property. For example, sand content of the soil appeared to have meaningful correlation with averaged RGB band reflectance (r = 0.21), clay raster surface (r = -0.46), DEM (r = -0.22), MRRTF (r = 0.54), MRVBF (r = -0.43), and valley depth (r = -0.26). Thus, we only included these six variables in the KED model for predicting the sand content.

Training and testing of the prediction models: We built and assessed the KED models separately for all the sampling efforts derived from the SLA and MIRS datasets. We randomly separated 25% (n=16) samples from the '100% sampling effort' of the SLA dataset and utilized them for independent validation of all 20 models. Given the small sample size of the validation data set, this was repeated five times using a new set of randomly selected validation data for each iteration. We then reported the mean and standard deviation of the accuracy metrics for the five iterations.

We used four error indices for measuring the model performance: (i) the coefficient of determination (R^2); (ii) the Lin's concordance correlation coefficient (CCC); (iii) the root mean square error (RMSE); and (iv) the normalized root-mean-square error (nRMSE), where RMSE is normalized by dividing by the range of the observed data (Shen et al., 2016). (Equation 3), (Equation 4), (Equation 5), and (Equation 6) below describe these accuracy measures.

$$\mathrm{R}^2 = \frac{\left(\sum_{i=1}^{n}(\mathrm{Y_i} - \overline{\mathrm{X}})(\mathrm{X_i} - \overline{\mathrm{X}})\right)^2}{\sum_{i=1}^{n}(\mathrm{Y_i} - \overline{\mathrm{Y}})^2 \sum_{i=1}^{n}(\mathrm{X_i} - \overline{\mathrm{X}})^2} \qquad \textbf{(Equation 3)}$$

$$\mathrm{CCC} = \frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2 + \left(\overline{\mathrm{X}} - \overline{\mathrm{Y}}\right)^2} \qquad \textbf{(Equation 4)}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\mathrm{Y_i} - \mathrm{X_i})}{n}} \qquad \textbf{(Equation 5)}$$

$$\mathrm{nRMSE} = \frac{1}{Z}\sqrt{\frac{\sum_{i=1}^{n}(\mathrm{Y_i} - \mathrm{X_i})}{n}} \qquad \textbf{(Equation 6)}$$

In (Equation 3), (Equation 4), (Equation 5), and (Equation 6), X and Y represent measured and predicted values, respectively; n is the number of samples; X_i and Y_i are the paired i^{th} values from the measured and predicted data, respectively; $\overline{X}$ and $\overline{Y}$ are the mean of the predicted and observed data, respectively; ρ is the Pearson correlation coefficient between the measured and predicted values; σ_x and σ_y are the corresponding variances of the measured and predicted values; Z is the difference between the maximum and minimum values of the observed data.

2.3 Results and Discussion

2.3.1 Summary of Soil Properties and Prediction Using MIRS

There was considerable variability in different soil properties across the study site (Table 3). The soil was dominated by silt and relatively high clay content. The soil pH was within the optimum tolerance range of the cultivated crops, but the soil salinity (determined by EC) was relatively high. While the mean EC values below the 4 dS/m threshold were identified for crop production in this region (Bertrand, 1991), EC values determined in our study were as high as 16.9 dS/m. The SOM and TN were with the optimum ranges for vegetable production. The skewness of the data indicated close to a normal distribution for most of the soil properties, except for EC. Kriging interpolation, in general, does not perform well for highly skewed data (Ouyang et al., 2003) and thus, normality transformations were performed if data became highly skewed with different cLHS selections of sampling efforts. Appendix A included the full dataset for all the soil properties.

Table 3: Summary statistics of different soil properties

Statistics of sand, silt, clay, pH, electrical conductivity (EC), soil organic matter (SOM), and total nitrogen (TN) from the 0-15 cm depth for standard laboratory analysis (SLA) and mid-infrared spectroscopy (MIRS). CV refers to the coefficient of variation. The MIRS prediction accuracy derived from the SLA using partial least square regression is illustrated by the coefficient of determination (R^2) and root mean square error (RMSE).

Soil property	**SLA analyzed samples (n=62)**			**MIRS analyzed samples (n=308)**			**MIRS prediction accuracy**	
	Mean	CV	Skewness	Mean	CV	Skewness	MIRS - R^2	MIRS - RMSE
Sand (%)	33.2	0.37	0.02	37.1	0.31	0.25	0.79	8.43
Silt (%)	50.4	0.19	0.08	47.6	0.20	-0.01	0.78	6.57
Clay (%)	16.4	0.27	0.73	15.1	0.26	0.85	0.78	2.72
pH	5.5	0.05	0.52	5.6	0.05	0.39	0.71	0.42
EC (dS/m)	2.23	1.07	2.02	5.05	0.72	0.69	0.69	3.13
SOM (%)	5.4	0.20	-1.07	5.1	0.18	-1.02	0.87	0.46
TN (%)	0.28	0.21	-0.91	0.26	0.19	-0.61	0.88	0.04

Models derived from SLA and MIRS spectra using the PLSR were fairly accurate for some soil properties, but not all (Table 3) highlighting the difference between SLA and MIRS data. The best predictions were attained for SOM and TN with R^2 values nearing 90%, while the pH and EC were predicted with lower accuracies (R^2 around 70%). Prediction accuracies for sand, silt, and clay were intermediate with R^2 values of nearly 80% and fell within or close to the range of R^2 values reported by other studies using MIRS and PLSR. For example, Masserschmidt et al. (1999) achieved R^2 of >90% for SOM and Janik and Skjemstad (1995) reported an R^2 of 88% for N. Our findings were consistent with others who reported, consistently lower prediction accuracies for pH and EC. Janik et al. (1998) predicted EC with an R^2 of 23%, while Janik and Skjemstad (1995) found R^2 of 72% for predicting pH using the same techniques. In addition, Janik et al. (1998) predicted sand, silt, and clay with an R^2 of 88%, which is close to the accuracy we achieved in our study. While the reduction in accuracy due to modeling MIRS spectra is clear, particularly for some soil properties (i.e. up to 30% reduction), the cost savings for the differences in accuracy are large. In the present study, MIRS enabled about 5 times the amount of sampling for an equivalent cost.

2.3.2 Semivariogram – The Analysis of Spatial Autocorrelation

The semivariogram analysis was used to compute the experimental variograms (EV), which identified the spatial structure of the data to be used for prediction. It was observed that the spatial structure weakens with decreasing sampling efforts (e.g. from 60% to 20% sampling effort). At the minimum sampling efforts (i.e. 10%, 20%, and 30%) there were limited to no spatial patterns, even after short separation distances, especially for the variograms of the SLA dataset.

As the spatial structure of the data changed with varying sampling efforts, the shape and structure of the EVs also differed. It was evident that the variograms were considerably different from each other for the varying intensities of sampling effort, with clear differences between the 60-100% and 20-40% sampling efforts. An example of how the EVs differed for various sampling efforts of the MIRS for TN is shown in Figure 5. The maximum range of spatial autocorrelation was observed at the 100% sampling effort as 300 m. The range remained close to 300 m with the decreasing sampling intensities until it reached 60% sampling effort but subsequently, the range declined significantly. Moreover, the nugget effect representing the measurement errors or microscale variation causing a discontinuity in the EV near the origin

(0,0) becomes larger with decreasing sample efforts. The higher nugget effect may result from the sparse sample distribution at the decreased sampling effort and consequent decline in spatial autocorrelation (Robinson and Metternicht, 2006). The decline of the spatial structure can be further realized if the nugget-to-sill ratios (N/S) are compared where N/S <0.25 indicates strong spatial dependence and N/S ranging between 0.25 to 0.75 refers to moderate spatial dependence, and finally, N/S >0.75 represents weak spatial dependence indicating poor or meaningless kriging prediction (Cambardella et al., 1994; Duffera et al., 2007). For the example demonstrated in Figure 5, the N/S ratio ranged between 0.25 to 0.49 for the sampling efforts of 60% to 100%, indicating strong to moderate spatial dependence. However, at 40% sampling effort, the ratio declined to 0.81, which refers to poor spatial dependence, and at 20% sampling effort there was no spatial autocorrelation as observed by the level straight line of EV (Liu et al., 2006). Similar results were observed for the other soil properties for which 50% to 60% sampling efforts were found as the optimum mark after which the spatial autocorrelation declined substantially.

We also observed that the variance decreased or flattened out with the increasing lag distance and thus, at some reduced sampling efforts, EV could not be constructed (Kerry and Oliver, 2007). In the case of SLA sampling efforts, which comprise 80% fewer samples than its MIRS counterpart, poor spatial autocorrelation and variograms were detected even at higher sampling efforts.

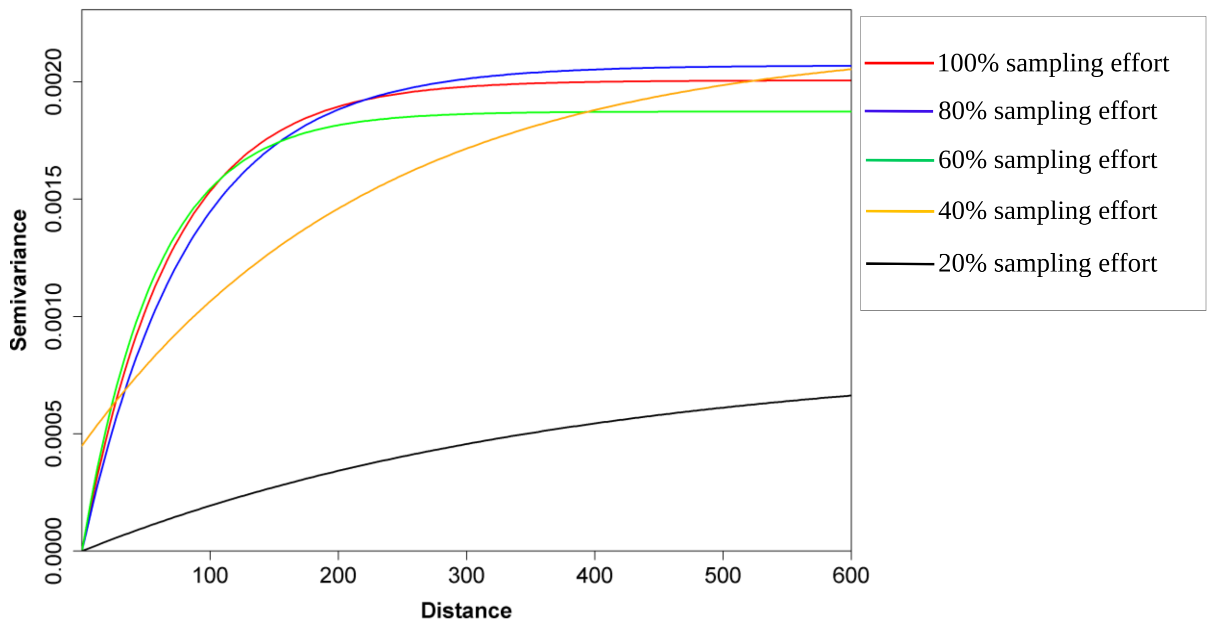

Figure 5: Experimental semivariograms of total soil nitrogen at different sampling efforts
Predicted using kriging with an external drift model. Data obtained by mid-infrared spectroscopy

Figure 6 displays examples of DSM for the study field at 5 m resolution of sand% and organic carbon% produced with SLA and MIRS datasets at 60% sampling effort. DSMs produced with the MIRS dataset provided a more detailed representation of the surface compared to the DSMs produced with the SLA dataset. MIRS DSMs showed clearly different patterns of soil property distribution than the SLA DSMs. In the MIRS DSMs, a greater range of values was predicted and there were fewer fine scale isolated patches of distinct values compared to the SLA DSMs. We also observed higher kriging prediction variance and edge effects for the SLA DSMs. Some linear features that were exhibited in the maps represent wide drainage ditches separating different field plots or the farm access roads.

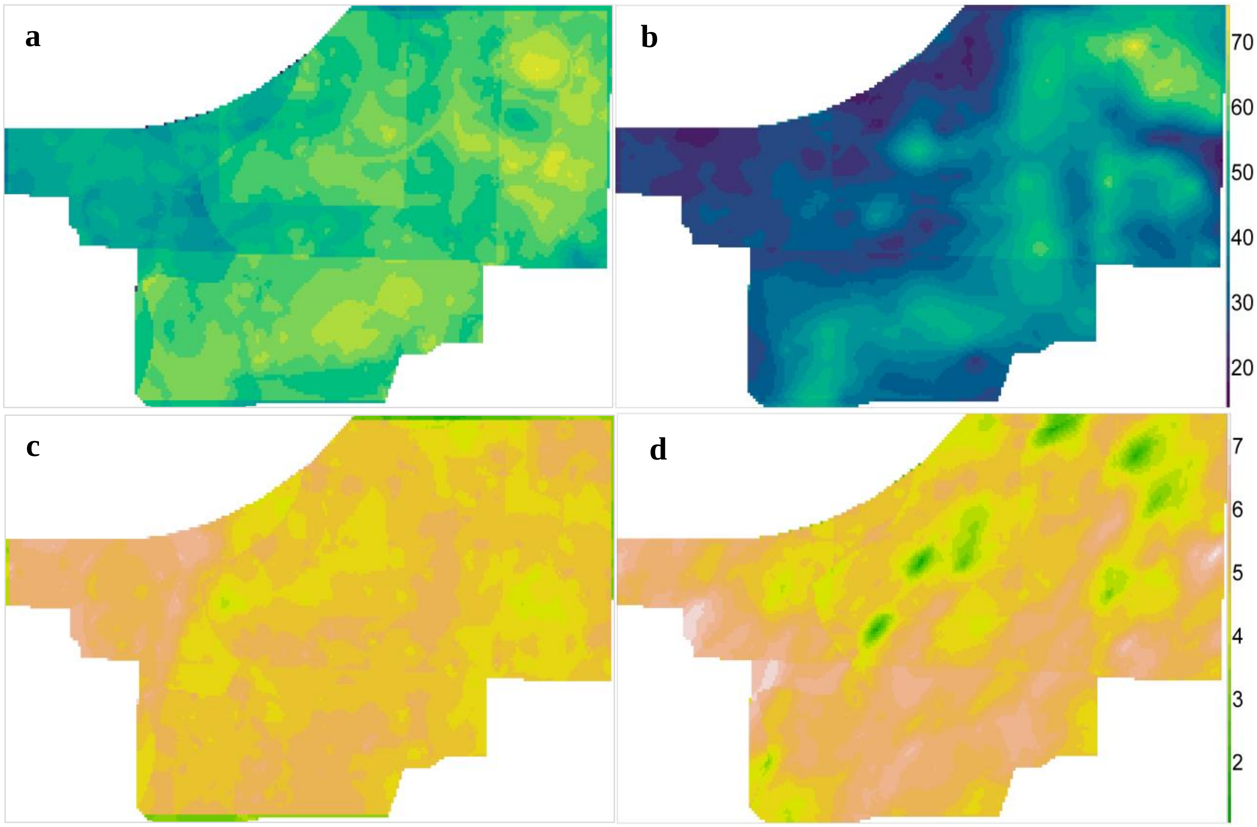

Figure 6: Examples of digital soil maps for the study field

Map shows sand% produced with (a) standard laboratory analysis (SLA) or (b) mid-infrared spectroscopy (MIRS) and organic matter% (c) with SLA or (d) with MIRS at 5 m resolution

2.3.3 Assessment of Model Performance

Comparing Equivalent Sampling Efforts of SLA and MIRS

When comparing the SLA and MIRS sampling efforts for predicting the digital soil maps, our analysis showed that the pattern of prediction accuracy differed widely across the soil properties and there was a clear difference in performance between SLA and MIRS (Figure 7 and Figure 8). It was evident that MIRS always performed better than SLA at an equivalent sampling effort but the differences in accuracy decreased as the sampling efforts got smaller. For example, the difference in R^2 of sand prediction ranged from 0.45 to 0.52 for the 40% to 100% sampling efforts of SLA and MIRS while the difference was as low as 0.14 for <40% sampling efforts. We also observed that accuracy measures, i.e. R^2, CCC, and nRMSE were not always in agreement when SLA and MIRS were compared for a specific model. In terms of R^2, the pH model for the MIRS dataset at 100% sampling effort, for instance, was 69% better than the output using the SLA dataset. However, the same prediction output was only 35% and 18% better when nRMSE and CCC values, respectively, were compared.

Despite the lower lab accuracy as explained in section 2.3.1, the overall performance of the MIRS dataset was substantially better than that of the SLA dataset for each equivalent sampling effort. In our analysis, the MIRS dataset had about 5 times more samples than its SLA counterpart for the same cost, thus capturing more of the spatial variability across the study site and producing stronger prediction performances. If equivalent sampling densities (i.e. number of samples), however, were compared instead of the sampling efforts, we might obtain different results. For example, both SLA-100% and MIRS-20% sampling efforts had a total of 62 samples and from the results, it was clear that SLA-100% performed better than MIRS-20% sampling efforts for all the soil properties. This might be due to the prediction inaccuracies that occurred during producing the MIRS dataset as explained in section 3.1. Yet, our analysis illustrated that the spatial variability was not effectively captured with such a small number of samples (i.e. n=62), resulting in weak prediction performances.

Our findings are similar to several previous studies that also confirmed that the use of the MIRS technique allowed them to utilize more sample points for DSM and the additional points helped them acquire better prediction accuracies without increasing the cost of analysis. A recent regional-scale study conducted in south-west Germany, using MIRS and kriging interpolation, generated SOM map with high accuracy as indicated by the overall similarity of 48 – 69% with the existing digital map (Mirzaeitalarposhti et al. 2017). Mirzaeitalarposhti et al. (2017) also concluded that the use of MIRS significantly reduced the cost of their research as 90% of the samples were analyzed with the MIRS technique although no cost comparison with SLA was provided. Another study by World Agroforestry Center reported that MIRS analysis of SOC data reduced the cost by 70% as compared to that of traditional chemical analysis in their African soil information project (Nocita et al., 2015). O'Rourke and Holden (2011) found spectroscopic analysis was 10 times more cost-effective than SLA. In our analysis, the MIRS dataset was 5 times more cost-effective than the SLA dataset. The low cost for MIRS enabled the use of far more data points than SLA for DSM for an equivalent sampling effort, clearly resulting in better prediction accuracies for all the soil properties.

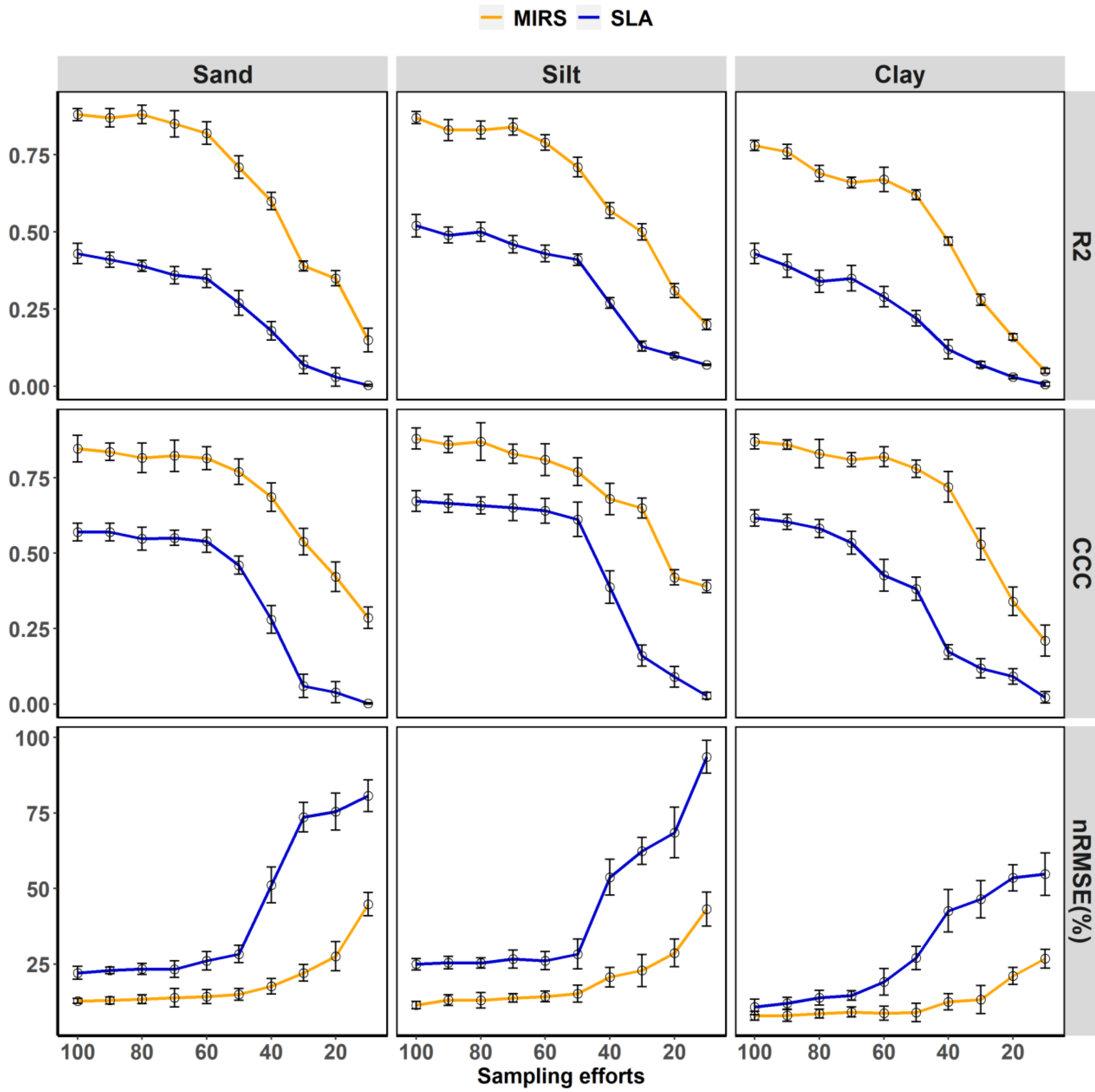

Figure 7: Model assessment of the prediction of sand, silt, and clay

At various equivalent sampling efforts of standard laboratory analysis (SLA) and mid-infrared spectroscopy (MIRS) datasets using kriging with external drift. Means and standard deviation (error bars) of the iterative analysis of the coefficient of determination (R^2), Lin's concordance correlation coefficient (CCC), and normalized root mean square error (nRMSE) are shown.

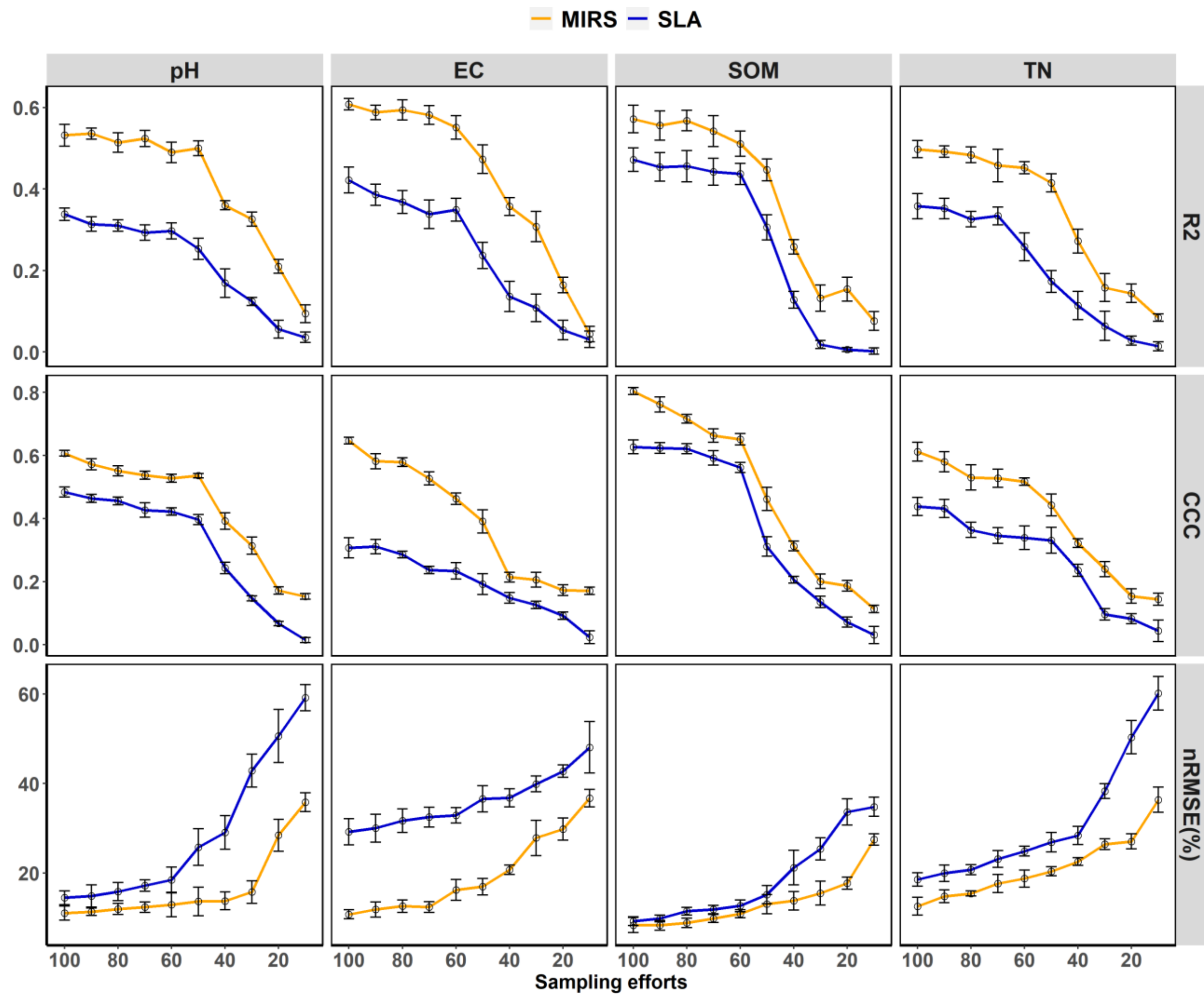

Figure 8: Model assessment of the prediction of soil pH, electrical conductivity (EC), soil organic matter (SOM) and total nitrogen (TN)

At various sampling efforts of standard laboratory analysis (SLA) and mid-infrared spectroscopy (MIRS) datasets using kriging with external drift. Means and standard deviation (error bars) of the iterative analysis of the coefficient of determination (R^2), Lin's concordance correlation coefficient (CCC), and normalized root mean square error (nRMSE) are shown.

Trade-offs between Cost and Accuracy of the Predictive Models

The results of the model assessment described above clearly show that the relationship between sampling efforts and prediction accuracy was generally non-linear for most of the soil properties (Figure 7 and Figure 8). With decreasing sampling efforts, the prediction accuracy declined exponentially for most soil properties regardless of methodology or accuracy metric. In a few cases, the decline was sharper and more linear, e.g. the decline in CCC of silt prediction between 50% and 20% of SLA sampling efforts (Figure 7), and the decline in R^2 of pH prediction between 50% and 10% of MIRS sampling efforts (Figure 8).

We found that the prediction accuracies significantly improved up to the 50% to 60% sampling efforts of both SLA and MIRS and after this point, the accuracy did not equally improve. Similar results were obtained by Simbahan and Dobermann (2006) when they tested the prediction accuracy of a regression kriging model for SOC using sample sizes of 50, 100, 150, and 200 and found that RMSE of prediction did not improve or minimally improved after reaching the sample size of 100. However, as we already mentioned, predictions using the MIRS datasets were more accurate than the predictions using the SLA datasets at equivalent sampling efforts. While data points were added with the increasing sampling efforts, the spatial autocorrelation was captured at the 50% to 60% sampling effort, which resulted in the most effective prediction performance in terms of both the prediction accuracy and cost. For example, the CCC values of sand prediction using the MIRS dataset were improved by 184% between 10% and 60% sampling efforts, whereas between 60% and 100% sampling efforts accuracy only improved by 4%. This 4% improvement in accuracy for CCC between the 60% and 100% sampling efforts cost 984 C$; i.e. 246 C$/1% accuracy improvement. In contrast, an investment of only 7 C$/1% accuracy improvement of CCC was required between 10% and 60% sapling efforts. For a dynamic soil property such as pH, the prediction using MIRS dataset required an investment of 4 C$/1% improvement of CCC between 10% and 50% sampling efforts, whereas the investment was 93 C$/1% improvement between 50% and 100% sampling efforts. Mapping using the full dataset of

308 points with SLA would have cost 12,320 C$, whereas the total cost of the full MIRS dataset was only 2,464 C$, highlighting the cost efficiency of using the MIRS technique. Although the investment for 1% accuracy improvement varies for different soil properties, it was clear that the model performance was most cost-effective at the 50% to 60% sampling efforts considering the high incremental cost and the unequal gain in the prediction accuracy above 60%.

Thus, applying the cLHS sample selection technique and sampling at an intensity of 2-3 samples per hectare (i.e. 50-60% data points of the initial sampling effort derived from a 40 x 40 m grid) provided the most effective sample design for DSM for our study field in terms of the accuracy of the KED model and costs. This sampling density is substantially different than that determined by Kerry et al. (2010) for a study field in Wallingford, England. Using a 30 x 30 m sampling grid for their analysis, they reported that spatial variability was captured most effectively at the 100-120 m sampling interval (i.e. 0.7-1 sample per hectare). Although our sampling intensity was higher than in Wallingford, this may not be the most effective intensity in all cases as soil property, environmental conditions, and management strategy vary from site to site. While it is unlikely that farmers would sample at such high density by hand frequently, our analysis demonstrates the intensity of sampling that could be employed to calibrate hand-held or tractor-mounted MIRS techniques.

2.4 Chapter conclusions

Identifying an effective sampling effort is critical for maximizing the accuracy and minimizing the cost of DSM models. We compared 20 different sampling efforts derived from the datasets of standard laboratory analysis (SLA) and mid-infrared spectroscopy (MIRS) for producing digital maps of a suite of soil properties including sand, silt, clay, pH, EC, SOM and TN. We determined that at an equivalent sampling effort the MIRS dataset produced more accurate maps of selected soil properties as compared to the maps predicted by SLA datasets, although the prediction accuracy varied across the soil properties and by accuracy metric (e.g. R^2, CCC, and nRMSE). Our analysis showed that the cost per improvement in accuracy with increasing sampling efforts was optimized at the 50-60% sampling effort. Thus, a sampling density of 2-3 samples per hectare, selected using a spatial sample selection technique (e.g. cLHS), and analyzed using MIRS in the lab was the most cost-effective approach for the production of accurate DSMs for our study field. However, these findings may vary for mapping in other crop fields with different soils, topography or management history. Hence, further analysis should explore how these findings may differ based on soil type, environmental covariates, or field management.

Chapter 3: Mapping soil organic carbon and clay using remote sensing to predict soil workability for enhanced climate change adaptation

3.1 Chapter introduction

Agricultural production worldwide has become highly mechanized to reduce labor, increase efficiency and to meet the demand of the growing global population. Mechanized farming operations often must be done under a narrow range of weather conditions to avoid adverse impacts because soil conditions at the time of mechanized operation determine the level of degradation of soil structure. This is becoming more challenging as weather patterns are rapidly shifting and becoming more unpredictable in many parts of the world (Chipanshi et al., 2018). Excessive precipitation, for example, can cause poor soil trafficability, restricting the use of farm equipment during critical times of the growing season (Kolberg et al., 2019; Servadio et al., 2016). Soil trafficability is a condition at which a soil provides sufficient tire-traction for the farm vehicles without causing the soil structural deformation (Earl, 1997). The use of farm machinery when trafficable conditions are not optimal destroys soil structure and leads to soil compaction (Müller et al., 2011). This is of particular concern in humid regions where precipitation can cause soil saturation during certain parts of the year, resulting in considerable delays in field preparation or harvesting which may shorten growing season and potentially reduce crop yields. Therefore, the shifting precipitation patterns, projected by various climate models (Fischer and Knutti, 2016), will likely reduce the number of days when agricultural soils are workable with heavy equipment without causing soil degradation (Tomasek et al., 2017). Site-specific optimum water content information for soil workability could be of value for scheduling farm management operations as well as evaluating the impacts of changing climatic conditions and developing associated adaptation strategies.

In the literature, a soil workability threshold (WT) is defined as the optimum soil water content at which mechanical tillage operations lead to a maximum number of soil aggregates (Dexter and Bird, 2001; Müller et al., 2011). WT is a combination of soil trafficability and the capacity of soil to be operated without causing substantial damage to its structure (Earl, 1997). WT depends on a number of soil properties, including texture, soil organic carbon (SOC), bulk density (Obour et al., 2017). The soil bulk density can increase as a result of overburden pressure imparted to the soil by machinery, while the WT is mainly controlled by the balance of precipitation, drainage, and evapotranspiration. However, both bulk density and WT are strongly related to SOC and soil texture (Gupta and Larson, 1979). Given these relationships, many previous studies proposed methodologies to estimate WT using these two key soil properties. For example, Dexter and Bird (2001) applied SOC and texture data and the water retention characteristics of the soil to determine WT, while some authors, like Kretschmer (1996) and Mueller et al. (1990) proposed the use of consistency limits of cohesive soils for this purpose. Additionally, Bueno et al. (2006), Mapfumo and Chanasyk (1998), Rounsevell (1993), and Rutledge and Russell (1971) found that WT is highly correlated to 95-99% of the soil water content at field capacity for different soil types.

Determination of WT using soil consistency is mainly based on the Atterberg's plasticity limits – lower plastic limit (LPL) and upper plastic limit (UPL) (Campbell, 1991). Soil exhibits liquid behavior and can freely flow at the water content above UPL, but it shows friability and breaks apart under pressure at the water content below LPL (Keller and Dexter, 2012). Soil needs to be at this friable stage during mechanical operation for optimum tillage; hence, determining this threshold is critical (Keller et al., 2007; Mueller et al., 2003). Since WT is certainly associated with the soil water content at LPL, the estimation of WT using the plasticity limits has been widely used for different soil types (Mueller et al., 2003; Smedema, 1993). The applications of this technique may not be effective for non-cohesive sandy soils; however, it is of value for cohesive, clay-rich soils that often have poor drainage (Dexter and Bird, 2001). Furthermore, there are well-established pedotransfer functions (PTFs) that can determine LPL and UPL from the soil texture and SOC data. For example, Kværnø et al. (2007) validated PTFs derived LPL and UPL values against field measurements achieving R^2 from 0.94 to 0.97. While the PTFs are extensively utilized, the most direct way of measuring LPL and UPL is from the remoulded soil at the laboratory (Obour et al., 2017). However, these PTFs can provide a considerable advantage in terms of time and cost-effectiveness when landscape-scale spatial variability of WT is of interest.

Given that soil properties are highly heterogeneous, spatially explicit and landscape-scale information on WT would be helpful for designing effective climate-adaptive management strategies. Yet, such information is not widely available to agricultural producers except in some selected regions. Mueller et al. (2003) demonstrated that available soil survey data and PTFs can be combined to assess the landscape-scale spatial variability of WT. Kværnø et al. (2007) presented a Norwegian case study where they examined the nature and extent of variability in soil texture, SOC, and WT within various soil map units using the approach described by Mueller et al. (2003). But the mapping of WT at the landscape-scale, using advanced digital soil mapping (DSM) of SOC and clay (CL) that integrates remote sensing (RS) and machine learning (ML) techniques (e.g. random forest and generalized boosted regression model), is still limited and needs further development. Advanced DSM is currently being used widely for predictive mapping of different soil properties with high accuracies (Heung et al., 2016; McBratney et al., 2003). Hence, landscape-scale mapping of WT using these state-of-the-art tools coupled with the PTFs of plasticity limits can provide particular benefits for devising regional soil management strategies to enhance adaption to changing climatic conditions. It is, however, unclear how effectively PTFs may be used for developing DSMs of WT and what modeling approach is best suited for these types of data.

To address these research gaps, the specific objectives of this study were to (1) produce maps of SOC and CL using advanced RS technique and two ML models including random forest (RF) and generalized boosted regression model (GBM), (2) compare the outcomes of the ML models for mapping SOC and CL, and (3) generate landscape-scale map of WT based on the PTFs described in Mueller et al. (2003). We conducted the study in the highly intensive agricultural landscape of Delta, British Columbia, Canada, where a combination of the soils with moderate to fine texture, poor natural drainage, and increasing spring and fall precipitation has amplified concerns for soil workability.

3.2 Methods

Our approach in this study included a combination of field sampling and geospatial analysis using DSM and RS techniques, and Figure 9 shows the steps involved in our analysis.

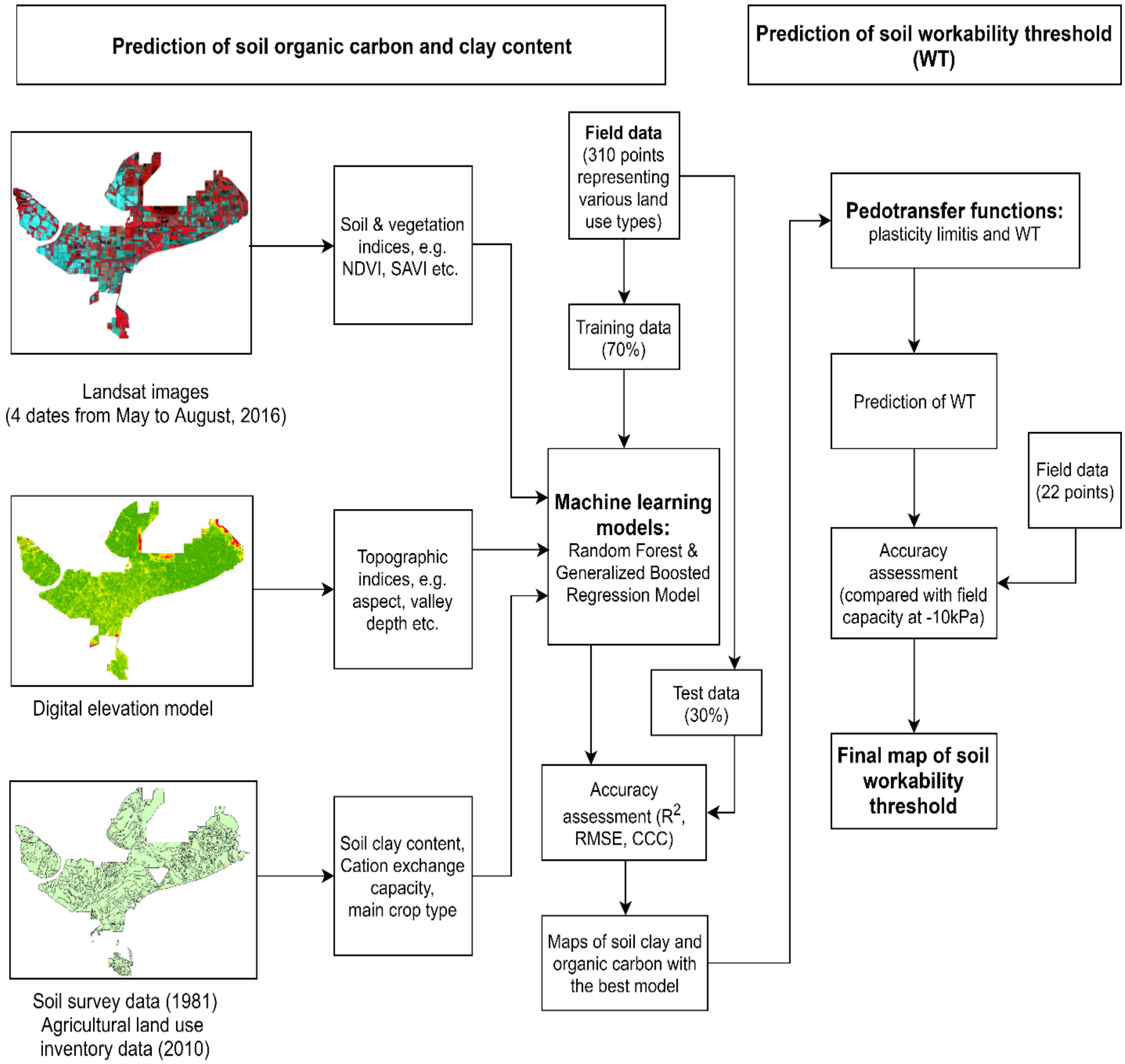

Figure 9: Schematic diagram showing the steps of producing the maps of soil organic carbon (SOC), clay (CL), and the workability threshold

R^2, RMSE, and CCC refer to the coefficient of determination, root mean square error, and concordance correlation coefficient, respectively

3.2.1 Study area

The study area represents the agricultural landscape within the City of Delta (49.08 N, 123.06 W), British Columbia, Canada and contains an area of 120 km^2 (Figure 10). Our analysis was restricted to land within the British Columbia's Agricultural Land Reserve. The study area is in the Fraser River delta and close to the ocean with an average elevation of 10 m above mean sea level. This area comprises a mild, humid climate with a mean annual temperature of 11.1ºC and mean annual precipitation of 1189 mm based on 30-year climate records (Environment Canada, 2019). The area is characterized by highly fertile, silty clay loam to silt loam soil with known issues of poor drainage. Delta is one of the most productive agricultural regions of British Columbia and produces a major share of the province's vegetable and blueberry crops. For the vegetable crops grown in the region, trafficability, is a major concern as a number of mechanized operations are required for preparation of the crop in the spring including disking, tillage, forming beds, nutrient applications, and planting. In that fall, heavy equipment is again used for operations like harvest, disking, tillage and sowing of cover crops. Trafficability is not as much of a problem for blueberry production, as mechanized operations of mowing grasses alleyways and harvesting are done during typically dry summers. During the winter, however, farmers do use heavy equipment to assist with pruning operations.

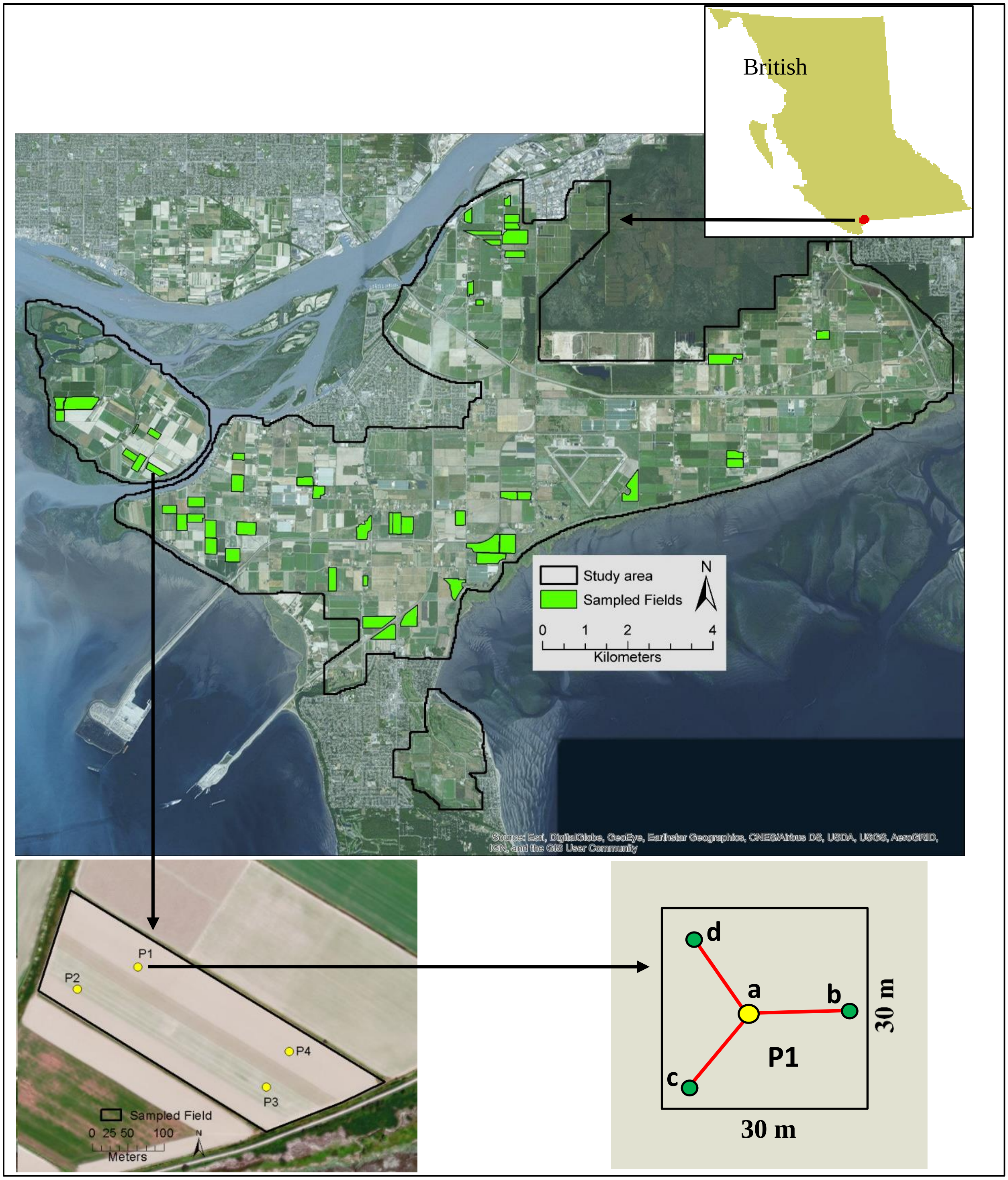

Figure 10: Map showing the fields sampled across the study area

At each field, 4 plots (P1, P2, P3, and P4) were sampled and 4 sub-samples (a, b, c, and d) were composited at each plot to get a representative sample from an area of 900 m2 (an area covered by a Landsat satellite image pixel)

3.2.2 Soil sampling and laboratory analysis

Across the study area, we collected a total of 310 soil samples at the 0–15 cm depth (Table 4) that were representative of different land use types (i.e., various annual and perennial crops, grassland, and hedgerow). At each sampling plot, we collected 4 samples (Figure 10) from an area roughly covering the area of a Landsat image pixel (i.e., 900 m^2). The soils from these 4 samples were composited to get a representative sample for that plot, while the center of the plot was recorded with a GNSS Pro 6H Differential Global Positioning System (DGPS) (Trimble Inc., Sunnyvale, California, USA) with post-processing accuracy varying from 10 - 50 cm.

We sent 25% of the samples (n = 75) to the Technical Service Laboratory of British Columbia Ministry of Environment for determination of SOC using the combustion elemental analysis with a Vario EL Cube Elemental Analyzer (Elementar, Langenselbold, Germany) and for soil clay (CL) using the hydrometer method (Kroetsch and Wang, 2007). We also analyzed all the samples (n = 310) in the TENSOR 37 spectrometer (Bruker Instruments, Ettlingen, Germany) using mid-infrared spectroscopy where the data from the standard laboratory analysis were utilized for calibration and validation in a Partial least square regression (PLSR) model.

Table 4: Total number of soil samples collected at 0-15 cm depth for various land use types

Land use type	Number of samples
Annual crop	193
Perennial crop (blueberry)	60
Grassland	37
Hedgerow	20
Total	**310**

3.2.3 Environmental variables for predicting SOC and CL

Topographic indices: We used the provincial 25 m Terrain Resource Information Management (TRIM) digital elevation model (BC TRIM, 2012) to derive a suite of topographic indices. We resampled the digital elevation model to 30 m. SAGA 2.1.2 software was then utilized to produce analytical hill shading (AHS), aspect, catchment area (CA), channel network base level (CNBL), closed depressions (CD), convergence index (CI), cross sectional curvature (CSC), longitudinal curvature (LC), slope-length factor (LS), multiresolution index of the ridge top flatness (MRRTF), multiresolution index of valley bottom flatness (MRVBF), negative topographic openness (NOpen), positive topographic openness (POpen), relative slope position (RSP), slope, terrain ruggedness index (TRI), total wetness index (TWI), valley depth (VD), and total curvature (TC) from the DEM. The details of the computation of these indices are described here: http://www.saga-gis.org/saga_tool_doc/2.1.3/a2z.html. Behrens et al. (2010), Lacoste et al.

(2014), and Schillaci et al. (2017) utilized some or all these topographic indices for producing DSM at different scales.

Existing soil survey and agricultural land use inventory: We used the existing detailed Canadian soil survey (1981) information and the data from British Columbia Ministry of Agriculture's agricultural land use inventory (2010) to derive environmental predictors. We extracted sand%, silt%, clay%, and cation exchange capacity (CEC) from the polygon-based soil survey map using a 30 m grid and produced raster layers for each of the soil properties using inverse distance weighting (IDW) interpolation. In the case of multi-component map units, we used the dominant (i.e., covering >50% of the unit area) category and assigned values accordingly; however, there were only a few cases with multi-component map units. The data from agricultural land use inventory (ALUI) represents detail land use information of every crop field within the agricultural land reserve in the study area (BC Ministry of Agriculture, 2016). This polygon dataset was directly rasterized to produce the ALUI covariate at 30 m spatial resolution.

Landsat image-derived indices: We downloaded the Landsat 8 Level-2 surface reflectance images (Path 47, Row 26) for 2016 from the United States Geological Survey Earth Explorer data warehouse. Four Landsat scenes captured on the dates May 31, July 02, July 18 and August 19 of 2016 were used in this study to capture the seasonal variability of the agricultural landscape. Acquiring images with minimum or no cloud cover was a significant challenge and limited the choice during the image selection process. We derived a suite of soil and vegetation indices (Table 5) and several image textural variables, including homogeneity (Homo), contrast

(Cont), and dissimilarity (Diss) of the images. These indices were developed for each of the four images of 2016. To produce the textural variables, we used the grey level co-occurrence matrix (GLCM) (Clausi, 2002) and derived the textural variables for the Near-Infrared (B5), Short-wave Infrared-1 (B6), and Short-wave Infrared-2 (B7) bands of the Landsat images. We applied the 'glcm' package in R to generate these textural variables (Zvoleff, 2016).

Table 5: List of soil and vegetation indices

In the formula, R, B, NIR, SWIR1, SWIR2 refer to the red, blue, near-infrared, short wave infrared-1, short wave infrared-2 bands of Landsat 8 satellite imagery, respectively, while L refers to canopy background adjustment factor

Soil and vegetation indices	Formula	References
Normalized Difference Vegetation Index (NDVI)	(NIR – R) / (NIR + R)	Rouse et al. (1974)
Soil Adjusted Vegetation Index (SAVI)	(1+L) (NIR – R) / (NIR+R+L), L=0.5	Huete (1988)
Normalized Difference Moisture Index (NDMI)	(NIR – SWIR1) / (NIR + SWIR1)	Hunt Jr & Rock (1989)
Soil Brightness Index (SBI)	$\sqrt{(R)^2 + (NIR)^2}$	Elvidge & Lyon (1985)
Normalized Difference Tillage Index (NDTI)	(SWIR1–SWIR2) / (SWIR1 + SWIR2)	Van Deventer et al. (1997)
Clay Minerals Ratio (CMR)	SWIR1 / SWIR2	Carranza & Hale (2002)
Bare Soil Index (BSI)	((SWIR1 + R) – (NIR + B)) / ((SWIR1+R) + (NIR + B)) * 100 + 100	Rikimaru et al. (2002)

3.2.4 RF and GBM for predicting SOC and CL

We then used two ML models – RF (Breiman, 2001) and GBM (Friedman, 2001) to predict SOC and CL. These decision tree-based ensemble models are comprised of a number of nodes and leaves where the nodes perform an 'if-then' statement based on the inferred relationships between the dependent variables and a set of predictor variables. The leaves represent the 'end-nodes' where a decision is made for the prediction (Bui and Moran, 2001; Heung et al., 2014). In

RF and GBM, outputs from an ensemble of decision trees are combined to improve the prediction accuracy and thus, they are being utilized for predicting complex soil-landscape where the prediction is usually dependent on a large number of variables (Heung et al., 2016; Schillaci et al., 2017). RF and GBM models have their own procedures for measuring the variable importance (VI). RF calculates the percent increase in mean square error (%IncMSE) of prediction by removing the variables one by one from the model and accordingly, determines the importance of each of the variables (Breiman, 2001). On the other hand, GBM measures the 'relative importance' score for each variable based on the empirical improvement of the model attained by splitting on a variable at the nodes and averaging over all boosted trees (Friedman, 2001). In this study, we used the 'randomForest' (Brieman et al., 2015) and 'gbm' (Ridgeway, 2015) packages in R to implement these models. We used 70% of the field data (n=310) for training the models, while the rest of the data were applied for accuracy assessment. We utilized coefficient of determination (R^2), concordance correlation coefficient (CCC), root mean square error (RMSE), and normalized root mean square error (nRMSE) to assess the accuracy of the predicted outputs. nRMSE is the RMSE value normalized by dividing by the difference between the maximum and minimum values of the observed data (Shen et al., 2016). (Equation 3), (Equation 4), (Equation 5), and (Equation 6) described these accuracy measures in section 2.2.5. Between the two models, the most accurate prediction of SOC and CL were used in the subsequent step for estimating WT.

3.2.5 Prediction and validation of WT

In this study, we used a series of PTFs for predicting WT. At first, we determined the UPL and LPL using (Equation 7) and (Equation 8) where the SOC and CL maps produced were used as the inputs. (Equation 7) and (Equation 8) were modified after Olson (1975) considering that SOC comprises 50% of the total organic matter in the soil (Pribyl, 2010). Although the conventional method considers that SOC constitutes 58% of the total organic matter or a conversion factor of 1.72 (Van Bemmelen, 1890), recent studies have reported stronger theoretical and empirical evidence in support of using a conversion factor of 50% or 2 (Kätterer et al., 2011; Pasley et al., 2020; Pribyl, 2010; Tammeorg et al., 2014). Using an independent field validation, we also observed a stronger correlation with our modeled data when a conversion factor of 2 was used. The maps of UPL and LPL were then applied to (Equation 9) to predict WT which is the maximum soil water content that provides optimum workability (Kværnø et al., 2007; Mueller et al., 2003). In all equations, the soil water content is represented as gravimetric %.

UPL = 11.9 + 0.92 x CL% + 0.08 x SOC% (after Olson, 1975) **(Equation 7)**

LPL = 7.15 + 0.199 x CL% + 1.957 x SOC% (after Olson, 1975) **(Equation 8)**

WT = LPL - 0.15 x (UPL - LPL) (Kretschmer, 1996; Mueller et al., 2003) **(Equation 9)**

To validate the prediction of WT, we collected an additional set of soil samples from 22 locations across the study area using a Conditioned Latin Hypercube Sampling technique (Minasny and McBratney, 2006) which selected stratified random sampling locations based on the spatial variability of the topography and soil types of the study area. The samples were collected as undisturbed cores at 0-7.5 cm depth and analyzed in the lab to determine the soil water content at field capacity at -10 kPa. We herein used -10 kPa matric potential for field capacity since our PTF of WT is assumed to be equal to the water content at this matric potential (Kværnø et al., 2007). We used Richard’s Pressure Plate Apparatus (Richards and Fireman,

1943) for this purpose. The samples were completely hydrated, weighed, placed inside the pressure plate chamber, and allowed to equilibrate at -10 kPa pressure. We then oven-dried the samples at 105ºC for 48 hours and weighed again. We determined the soil bulk density from the mass of oven-dried soils and the volume of the sampling core. Particle density was also calculated using the mass of oven-dried soils and the total volume of soil particles. We then calculated the porosity of soil using (Equation 10). Finally, the soil water content at -10 kPa (or at field capacity) was determined by applying (Equation 11).

$$\text{Porosity of soil} = \left(1 - \frac{\text{Bulk density}}{\text{Particle density}}\right) \text{x } 100 \qquad \textbf{(Equation 10)}$$

$$\text{Soil water content at -10 kPa} = \text{Water lost at -10 kPa x } \frac{\text{Porosity of Soil}}{\text{Total water lost}} \qquad \textbf{(Equation 11)}$$

Thereafter, we extracted the WT values of these 22 locations from the predicted map and tested for the accuracy measures described in section 2.2.5 (i.e., R^2, CCC, RMSE, and nRMSE).

3.3 Results and Discussion

3.3.1 Summary of soil properties from the field plot data

The SOC and CL in the study area varied substantially as would be expected in a region where land use is very heterogeneous and texture is influenced by the dynamics of the large adjacent river (Table 6). The range of SOC values in the sampled field plots may be attributed to the variable nutrient and soil management practices in these fields, the types of crop (i.e., annual vs. perennial crops) and non-production perennial vegetation (i.e., hedgerow, grass margin) scattered in and around the field. On the other hand, the range of the CL content was much larger than that of SOC, with plots closer to the river exhibiting higher CL content.

Table 6: Summary statistics for soil organic carbon (SOC) and clay (CL)

Soil Property	Minimum	Maximum	Mean	Standard Deviation	Range
SOC (%)	0.58	4.76	2.82	0.78	4.18
CL (%)	6	29	18.81	7.82	27

3.3.2 Selection of environmental variables for RF and GBM models

Of the 80 environmental variables derived from multiple sources, including topography (i.e., digital elevation model), soil survey and ALUI, and Landsat imagery, our analysis identified between 29 and 41 variables of high importance depending on the modeling approach. We ran the models for both SOC and CL with including all the environmental variables and then, identified the top predictors based on a threshold value of 5 for %IncMSE in RF and a threshold value of 1 of relative importance score in GBM. This process reduced the number of variables by 35-50% depending on the soil property and the type of model (Figure 11 and Figure 12). Overall, the topographic variables and soil survey and ALUI variables were found to be stronger predictors than the Landsat variables based on the VI scores for both models. It is also important to note that the Landsat indices, derived from the images of pre-growing and post-harvest seasons, when soils were left without cover, explained more variance predictors than the indices developed from the growing season images. The CEC was found to be the most dominant predictor for SOC in both models, while the multiresolution index of the ridge top flatness (MRRTF) and land use types from ALUI were the most important variables for predicting CL using RF and GBM models, respectively. We also performed a Pearson correlation analysis on the selected variables presented in Figure 11 and Figure 12 and removed any variable from the final model if it had a correlation coefficient value of ≥80% with another variable.

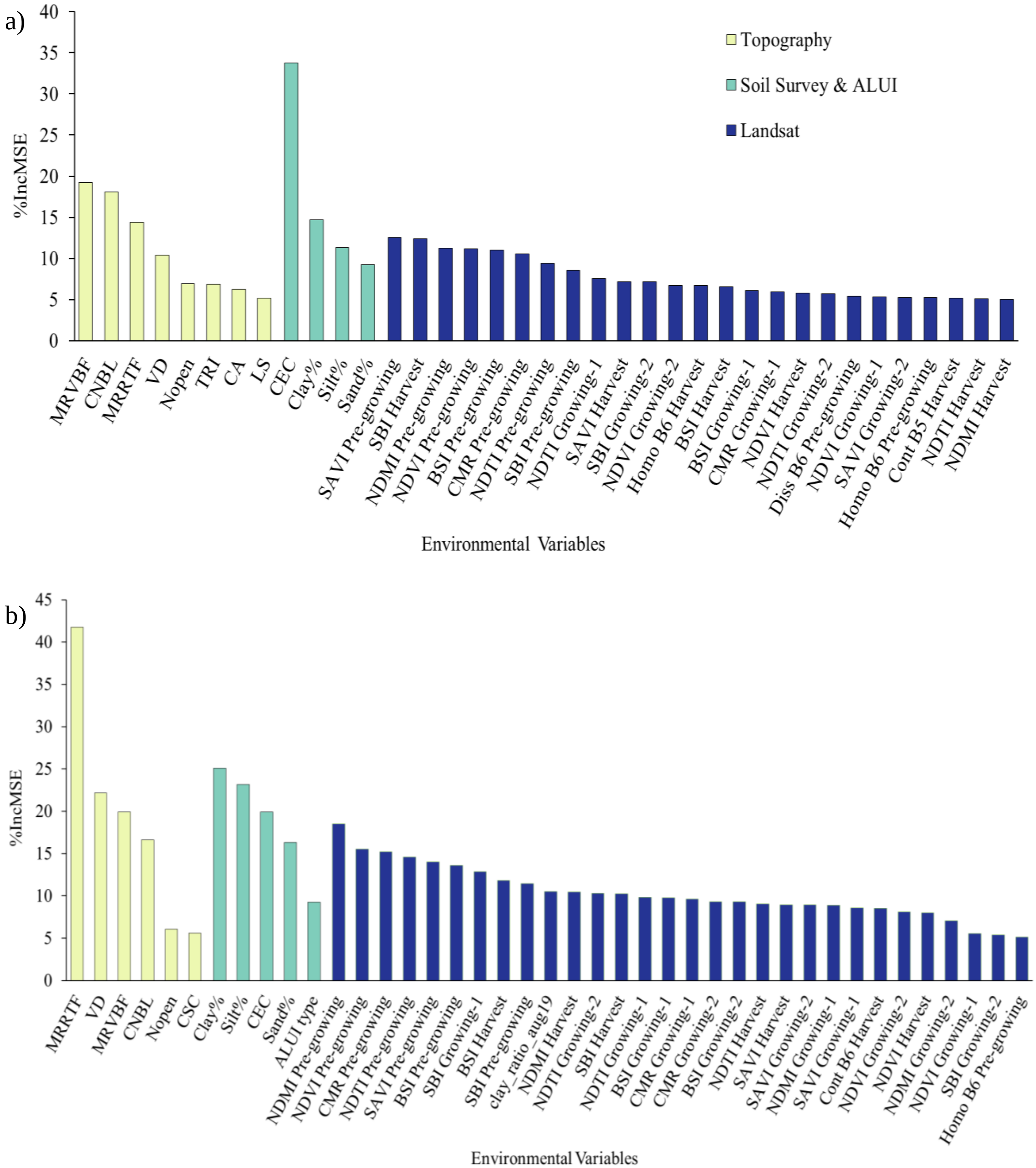

Figure 11: Most important environmental variables in random forest model

a) for predicting soil organic carbon, b) for predicting clay. See section 3.2.3 for the full name of the variables.

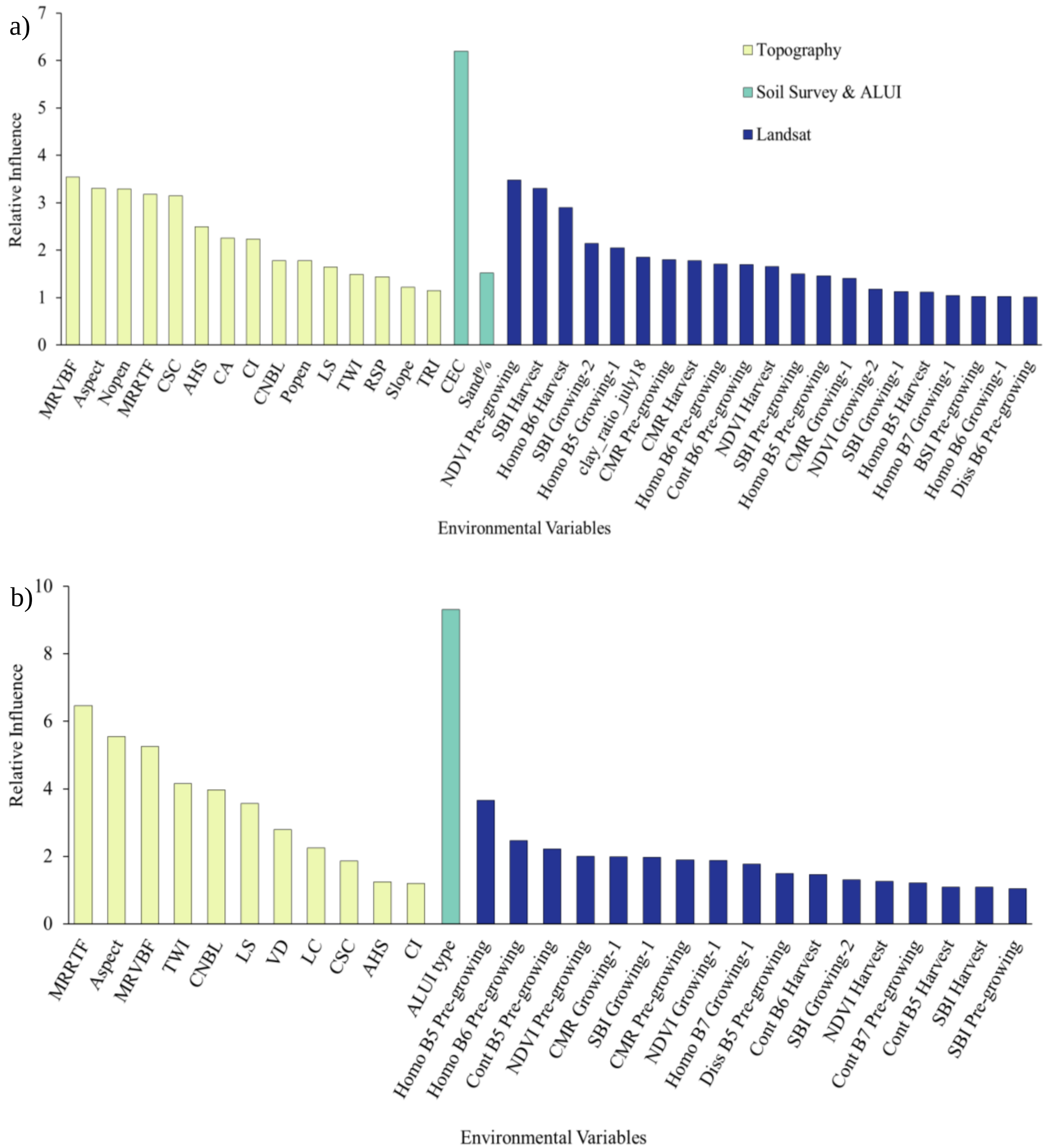

Figure 12: Most important environmental variables in generalized boosted regression model

a) for predicting soil organic carbon, b) for predicting clay. See section 2.3 for the full name of the variables

3.3.3 Assessment of the performance of RF and GBM models

In our study, both ML models performed reasonably well, but there were important differences in their accuracy metrics depending on the soil property. Both RF and GBM models predicted CL more accurately than SOC, although the differences were minor (Table 7). We found that RF outperformed GBM for both SOC and CL for all accuracy metrics (i.e., R^2, CCC, and nRMSE). However, the differences between the models were more obvious when R^2 values were compared for both SOC and CL. The nRMSE and CCC of predictions were somewhat close to each other, especially for SOC. The R^2 and CCC of SOC prediction using RF were 38% and 12% higher, respectively than those using GBM, while the nRMSE was 14% less with the RF model. We found similar results for the prediction of CL where the differences in R^2, CCC, and nRMSE were 51%, 26%, and 25%, respectively between the outcomes of RF and GBM models. We also tested the model accuracy when including all the 80 variables for all cases to test if discarded variables had any effect on the model performance. We found that the model accuracies were not impacted, except the R^2 and CCC of RF-CL prediction were decreased by 3% and 2%.

That the RF model performed better than GF in our study is consistent with Wang et al. (2018) who mapped SOC in a semi-arid rangeland of Australia, predicting SOC with R^2 values ranging between 0.42 and 0.48 and CCC values ranging from 0.56 to 0.62 for various sets of environmental variables, compared with values for R^2 (0.55) and CCC (0.7) for RF in the present study. In contrast, the study by Yang et al. (2016) reported that GBM performed better than RF for modeling SOC in an alpine ecosystem, especially in the areas with greater vegetation cover.

Table 7: Accuracy of prediction of soil organic carbon (SOC) and clay (CL)

Using random forest (RF) and generalized boosted regression model (GBM). R^2, CCC, and nRMSE represent the coefficient of determination, concordance correlation coefficient, and normalized root mean square error (nRMSE) respectively

Model	Accuracy metrics	SOC	CL
RF	R^2	0.55	0.62
	CCC	0.70	0.72
	nRMSE	0.12	0.15
GBM	R^2	0.39	0.41
	CCC	0.55	0.63
	nRMSE	0.14	0.20

Landscape-scale prediction of CL using these techniques is relatively sparse in literature, but our findings support the results of Chagas et al. (2016) who predicted CL with RF and obtained an R^2 of 0.56, which is slightly lower than what we were able to achieve in this study. Our results did differ from those of Sindayihebura et al. (2017) who demonstrated that GBM better predicted CL across Burundi's central plateau in Africa. Such differences in the performance of these ML models in various studies have been mainly attributed to the dissimilarities in landscapes and environmental conditions, the scale of prediction, and the type and quality of environmental variables used (Were et al., 2015). Therefore, selecting a single ML model as the best method for predicting landscape-scale soil properties is difficult, and largely site and case dependent (Ließ et al., 2016). In our study, the RF model likely outperformed GBM because of the RF model's simple parameterization and reduced susceptibility to overfitting as well as greater descriptive power that enables the model to decipher the complex and hierarchical relationships between the environmental variables and the target soil properties (Wang et al., 2018).

We also examined the influence of each variable category on the prediction accuracy of the models where the results reported in Table 7 were used as the benchmark for comparison (Figure 13). The topographic variables were the most important category of predictors in our study, regardless of model or soil property. However, they showed significantly stronger predictive capability for some cases of GBM as compared to RF. For example, when topographic variables were included in the GBM model, the R^2 of SOC prediction was improved by 56% whereas the improvement was 32% for RF. Similarly, R^2 of CL prediction was improved by 105% for GBM and by 27% for RF when the topographical variables were included. The second most important predictor category was the soil survey and ALUI variables followed by the Landsat variables. We also observed that Landsat variables were more influential for RF predictions than those of GBM. For instance, the R^2, CCC, and nRMSE of the RF prediction of SOC was improved by 15%, 7%, and 15%, respectively when Landsat variables were added to the model, but the improvements were 3%, 2%, 6%, respectively when modeled with GBM. We obtained similar results for the prediction of CL.

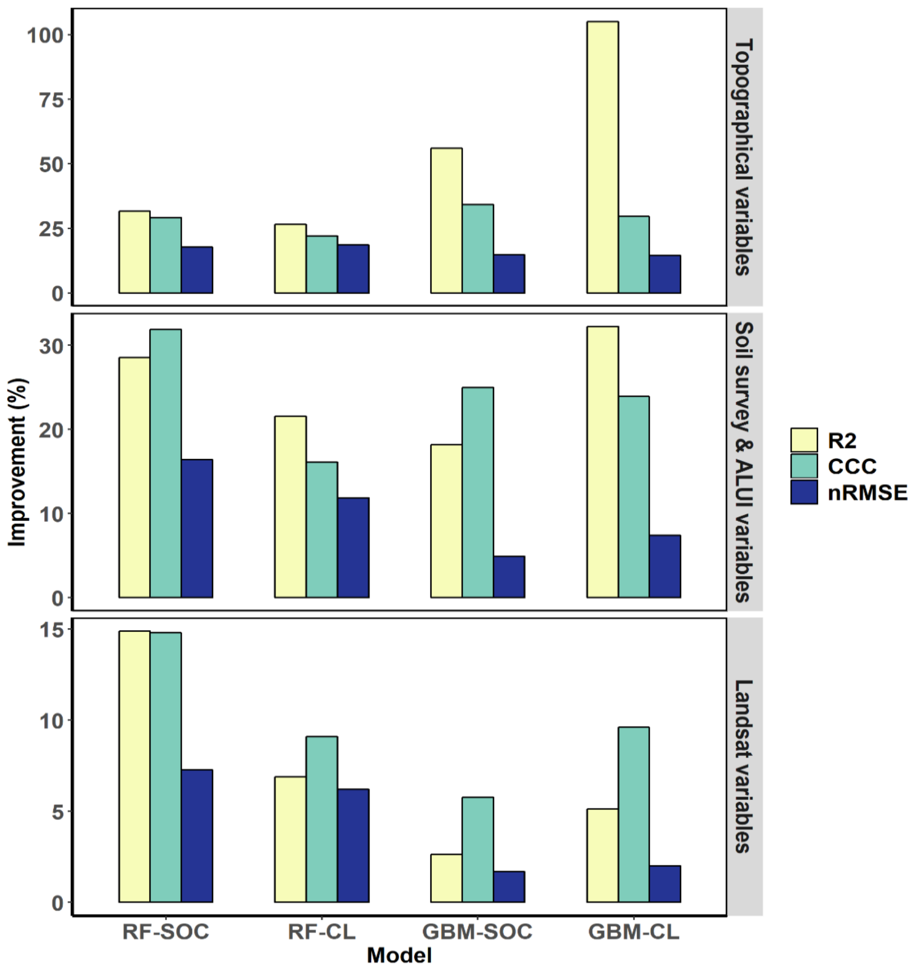

Figure 13: Improvement (%) in model accuracy

In terms of R^2 (coefficient of determination), CCC (concordance correlation coefficient), and nRMSE (root mean square error) for the topographic, soil survey & ALUI, and Landsat variables using both random forest (RF) and generalized boosted regression (GBM) models for predicting soil organic carbon (SOC) and clay (CL)

As mentioned above, the performance of the ML models is highly dependent on-site characteristics, and the predictive capability of the environmental variables differs from case to case. For instance, the strong performance of soil and vegetation information derived from Landsat variables was reported by Chagas et al. (2016) and Grinand et al. (2017) for predicting CL and SOC, respectively using the RF model whereas Grimm et al. (2008) and Schillaci et al.

(2017) found topographic indices along with the soil survey and land use data as the most important predictors of SOC in their studies using both RF and GBM models. The findings of the latter two examples agree with our results. The topography of a site is related to the erosion and deposition of soil materials (Cavazzi et al., 2013); hence, it was expected that topographic variables would have a strong correlation with the target soil properties. In addition, the soil survey data used in our study captured the inherent characteristics of the soil-landscape and a much greater distribution of field-based data than in our study. Although these data were collected almost 40 years ago, soil texture is unlikely to change and changes in SOC are likely to have been correlated with land use. The agricultural land use data informed the model about the agricultural practices of the study area. This would likely be even more useful if we had data on land use changes over time. A study by Schillaci et al. (2017) found that incorporating multiple Landsat images spanning the whole season was more effective for predicting soil properties at landscape-scale, but we did not observe a strong predictive influence for the Landsat variables, even after including a large number of indices derived from four Landsat images representing the whole growing season. Kheir et al. (2010) reported that the removal of the Landsat variables increased the overall accuracy of their predictions. However, in our case, Landsat variables improved the accuracy of the models by 6 to 15% (Figure 13); although the improvement was not large, Landsat variable including contributed to model performance.

Even though the variable selection process substantially reduced the number of the variables and satisfactory predictions were obtained using them, we attempted to assess the model performance with even fewer variables so that we can identify the key predictors and minimize the analysis effort. To accomplish this, we ran the models including only two top predictors from each category. For instance, the RF-SOC model only included MRVBF, CNBL, CEC, Clay%, SAVI pre-growing, and SBI harvest as the predictor variables. We then compared the outcomes of these reduced models with the results of the full model as shown in Table 7. Interestingly, we achieved relatively similar accuracies for both RF and GBM models, with some variation depending on the soil property and accuracy metric. The accuracy remained the same for all the metrics of the RF-SOC model; however, the accuracy of the RF-CL model decreased by 20%, 7%, and 34% in terms of R^2, CCC, and nRMSE, respectively. On the other hand, for the GBM model, the results did not change substantially for either soil property, where R^2 of the predictions decreased by 5% and 9% for SOC and CL, respectively. These findings somewhat agree with the results obtained by Wang et al. (2018) where prediction accuracies for SOC remained unchanged or slightly improved when a more parsimonious model was used, although their variable selection approach was more complex compared with ours. Based on these results, we conclude that identifying only a few key environmental covariates based on the variable importance scores for a specific geographical area can improve the accuracy of digital soil maps. This may substantially reduce the analysis effort for producing the additional covariates.

3.3.4 Spatial distribution of SOC and CL

Given that the RF model performed better than GBM, it was used to predict the SOC and CL for the entire study area at a 30 m spatial resolution. The SOC content in our study area varied from 1.5 to 4.5% (Figure 14). Although there were several patches of crop fields with a higher concentration of SOC distributed across the area; generally, the fields in the north-eastern region had the highest concentration of SOC. This region of the study area borders Burns Bog, which is a unique ombrotrophic raised peat bog in the Fraser River Delta (Hebda et al., 2000). Moreover, this region was dominated by perennial highbush blueberry production which is known for sequestering carbon in the soil (Nemeth et al., 2017). Together, the historic bog ecosystem followed by a perennial cropping system resulted in a high SOC of this small region of the study area. Fields dominated by intensive annual crop production exhibited lower SOC concentration. In addition, the fields with lower SOC, especially on Westham Island have known issues of soil salinity (Lussier et al., 2019). High salinity reduces crop production resulting in inputs of organic matter and in turn, lower SOC concentrations (Rietz and Haynes, 2003).

The study area is characterized by CL content that ranged between 8 and 29% (Figure 15), where the western half of Delta was dominated by fields with higher CL. However, the highest CL values were observed in the fields adjacent to the river which deposited a large amount of fine sediments on those fields over the course of soil formation. Although Clay% from the soil survey data was one of the strongest predictors of SOC in the RF model, the CL and SOC were inversely correlated in our maps with a Pearson correlation coefficient of -0.19. Numerous studies have reported the opposite trend, showing the chemical adsorption of carbon onto the surface of clay minerals led to greater SOC in clay-rich soils than in coarse-textured ones (Johannes et al., 2017; Singh et al., 2018). In our study area, the fields with higher CL are also characterized by poor drainage, which in combination with intensive tillage, often destroys soil structure, compaction, and increased soil erosion (Müller et al., 2011), causing loss of SOC (Lilly et al., 2018). Thus, the observed reverse relationship between CL and SOC suggests that intensive tillage is most likely resulting in SOC losses across the study region. Our maps will provide a base-line for long-term monitoring across the region and enable tracking of future changes in SOC.

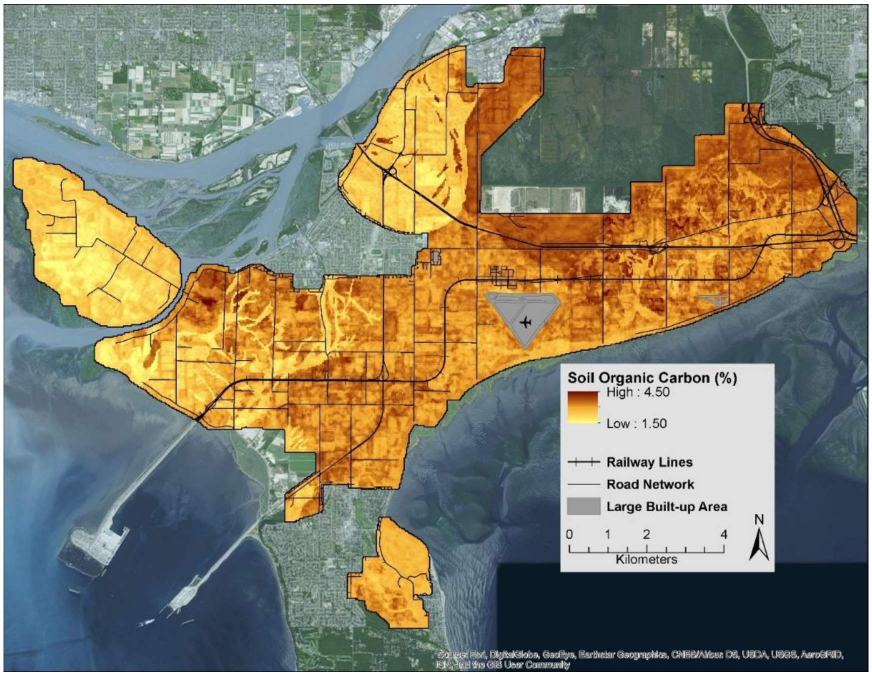

Figure 14: Soil organic carbon map of the study area predicted using random forest model

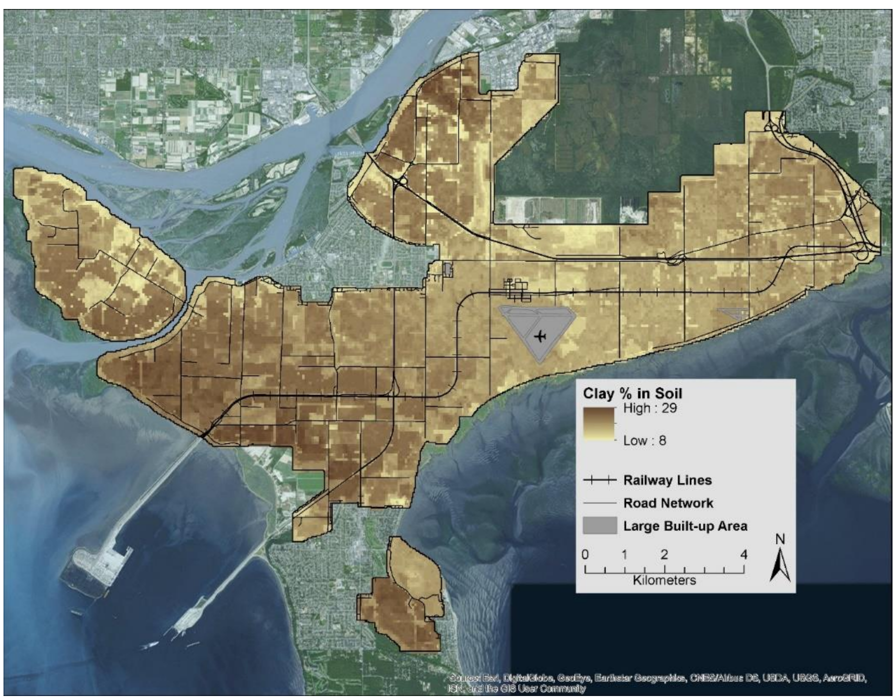

Figure 15: Clay map of the study area predicted using random forest model

3.3.5 Distribution of WT and accuracy of prediction

The prediction of WT for agricultural land in Delta resulted in higher than standard accuracy when validated with independent soil samples analyzed for field capacity at -10 kPa. The validation of the predicted WT map (Figure 16) resulted in R^2 of 0.59, CCC of 0.70, and nRMSE of 0.15. The WT ranged from 20 to 42% across the study area, where fields with high SOC and/or low CL exhibited high WT, and vice versa. Although we have found no other studies that have used advanced DSM to predict WT, Kværnø et al. (2007) produced a map of WT using simple kriging of the estimated values from soil map units; however, they concluded that their WT map did not capture the differences between soil types, especially at the boundaries between two or more soil units. Our map of WT is a substantial improvement in this regard as it effectively captured the variation in soil properties in a more continuous manner across the study area.

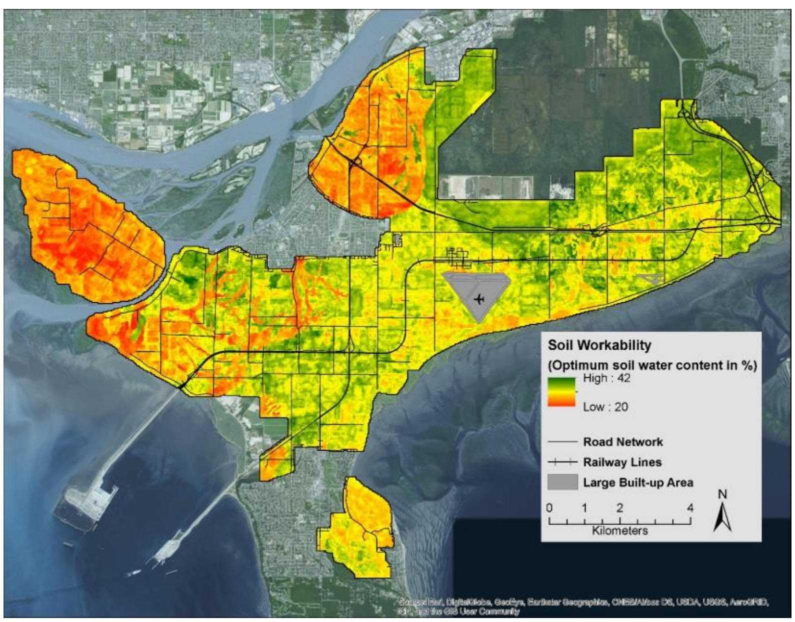

Figure 16: Soil workability (WT) map of the study area

WT values represent the optimum gravimetric soil water content (%) above which soil may be degraded with any mechanical operation

The WT map highlights that a substantial portion of the crop fields in Delta will likely face serious challenges due to poor workable conditions especially during the wet part of the year (i.e., spring and fall). Our map showed that 40% of the crop fields in Delta had a WT <30% (category-1), while 59% had a WT ranging from 30-40% (category-2), and 1% had WT >40% (category-3). Based on a study by Neufeld et al. (2017), which tracked soil water content (at 0 – 15 cm depth) across 26 fields in Delta with both heavy tillage and no-tillage practices, category-2 and -3 fields would have been workable by April 11th in the typical spring conditions of 2016, while fields in category-1 would be workable only a week later. Alternatively, in the unusually wet spring of 2017, fields in category-2 and -3 would not have been workable until May 15th and category-1 fields at least two weeks later. This pattern could also be observed in the fall of 2016 where category 1 fields were workable only until September 15th while category 2-3 fields could have been workable for another month. Wet conditions in the spring and the fall could result in large differences in the number of workable days between category-1 fields and those in category-2 and -3. This situation is expected to become more challenging in the coming years as climate models predicted a 7% increase in precipitation for the region by 2050 occurring mainly in the spring and fall (BC Agriculture & Food Climate Action Initiative, 2015). Adapting to this shifting precipitation pattern will likely require substantial investment at the farm level to enhance SOC or install drainage infrastructure field with low WT and at the regional level to improve water conveyance.

3.4 Chapter conclusions

Producing spatially explicit information on soil workability is critical for effective management and climate change adaptation in agricultural lands. We combined advanced remote sensing and machine learning tools with existing PTFs to produce a digital map of WT for an intensive agricultural landscape in Delta, British Columbia, Canada. We predicted SOC and CL across the landscape using RF and GBM models and found that RF was the best approach for both soil properties. Combining the digital maps of SOC and CL using a number of PTFs to produce a map of WT did not result in much reduction in accuracy. The WT map identified that 40% of the crop fields in the study area had a WT <30%, a threshold that will likely result in substantially fewer workable days than the other fields in the region. Our analysis demonstrates an effective approach to spatially predict WT across a heterogeneous agricultural landscape. These results can be used to formulate efficient farming strategies in the study area for more effective climate change adaptation. However, results may vary depending on the soil type, climate, and agricultural management practices; thus, future analysis should focus on validating the methodology in other geographical contexts.

Chapter 4: Tracking changes in soil organic carbon across the heterogeneous agricultural landscape of the Lower Fraser Valley of British Columbia

4.1 Chapter introduction

Given soil organic carbon (SOC) is the largest pool of carbon in the terrestrial biosphere, its management has important implications for the concentration of carbon dioxide (CO_2) in the atmosphere and the rate of climate change (Stockmann et al., 2013). The fate of SOC is largely the responsibility of agricultural producers who currently manage 38.4% of the global land area (FAO, 2011). To date, it is estimated the nearly 136 Gt of CO_2 have been lost to the atmosphere because of the conversion of natural ecosystems to agricultural production and subsequently intensive high-input, mechanized, management, particularly soil tillage (Lal, 2004). The agricultural sector also contributes a substantial portion of non-CO_2 climate-forcing global greenhouse gas (GHG) emissions (11.2± 0.4% of total emissions in 2010) (Edenhofer, 2014; Tubiello et al., 2015). Agriculture thus could play a critical role in mitigating climate change by reducing GHG emissions and returning previously emitted atmospheric CO_2 to the soil through sustainable or regenerative land use practices (Minasny et al., 2017).

Conservation management practices such as reduced tillage, increased crop diversity, grassland rotations, cover cropping, increased organic inputs or changing land use from annual to perennial cropping, restoration of marginal or degraded lands to forest patches or boundary hedgerow have been shown to sequester atmospheric CO_2 in the soil. Changing from conventional, high input agricultural management practices have been shown to sequester 0.37 to 3.67 Mg of CO_2 per hectare per year in temperate agroecosystems (Conant, et al., 2001; Paustian et al., 2016) while land use change has been shown to sequester 1.1 to 2.49 Mg of CO_2 per hectare per year globally (Deng et al., 2016). In some situations, the rate of SOC change has been observed to be much lower than the reported means and others, much higher (Zhang et al., 2015; Zomer et al., 2017) and this variation has largely been attributed to the differences in LULC or the implementation of land management and/or due to the soil type and climate (Minasny et al., 2017; Stockmann et al., 2015).

Given that SOC dynamics is a microbially mediated process, soil moisture and temperature also play an important role in determining sequestration or loss. While we have a relatively clear understanding of current SOC dynamics in relation to changes in moisture and temperature (Ise and Moorcroft, 2006; Sierra et al., 2015), the nuances presented by feedbacks from C inputs from agricultural management and plants coupled with elevated CO_2 make predicting responses to a changing climate complicated and challenging. A recent study by Crowther et al. (2016) showed that the global stock of SOC was negatively impacted by global warming. Some other studies, however, suggested the opposite, that SOC may increase under the warming condition because of higher soil microbial activity and plant productivity (Koven et al., 2015; Todd-Brown et al., 2014). Thus, tracking changes in SOC in the context of changes in management, LULC, and climatic conditions is critical for improving our understanding of SOC dynamics and the capacity to support practices most likely to increase SOC sequestration and reduce GHG emission from agricultural soil.

Until now, the primary approach for estimating SOC dynamics has been to use process-based models. However, the outcomes of such models may involve substantial uncertainties due to the complex biogeochemical processes and the spatiotemporal variation of environmental factors (e.g. topography, soil type, LULC, climate, etc.) responsible for the production and decomposition of SOC (Huang et al., 2019; Todd-Brown et al., 2014). All of these parameters need to be accurately represented in process-based models for meaningful estimation of SOC and thus, such modeling approaches become significantly data-intensive and often limited by the availability of accurate input data (Grinand et al., 2017). Static-empirical models, however, which use a space-for-time substitution method (Pickett, 1989) may provide a practical alternative to process-based models by establishing an empirical and spatially explicit relationship between SOC and the environmental driving factors for the current time and then, predict for past or future time step based on this relationship (Adhikari et al., 2019).

Minasny et al. (2013) proposed a static-empirical approach, called ‘scorpan’ for digital mapping of soil properties. The ‘scorpan’ approach produces a spatially explicit prediction of soil properties using the factors of climate, organisms (including management and LULC), topography, parent materials, time, space and other soil attributes (McBratney et al., 2003). Since the factors, like topography and parent materials, are static over time, any temporal changes in the dynamic factors, like climate and LULC should be reflected in the locally calibrated SOC empirical model. Although promising, this pragmatic modeling approach has only recently been employed for a few regions for predicting SOC for the past and future time steps, e.g. Adhikari et al. (2019), Huang et al. (2019), Yigini and Panagos (2016). These studies relied on dynamic environmental variables with a spatial resolution ranging from 100 to 1000 m and applied existing LULC data derived from national or continental scale models. As a result, SOC maps produced in these studies might not be fine enough to derive effective SOC enhancing strategies for highly heterogeneous agricultural landscapes where the effects of management and LULC on SOC are likely to be prominent.

Capturing seasonal variation (caused by annual crops) and diversity of vegetation types (i.e. annual and perennial crops, and other non-productive vegetation patches) are crucial for effective SOC estimation in heterogeneous agricultural landscapes. Deriving historical LULC maps and different soil and vegetation indices from the analysis of time series Landsat satellite imagery (30 m spatial resolution) can effectively detect the heterogeneity of such landscapes (Schmidt et al., 2016; Xu et al., 2018). Additionally, downscaling of the modeled climate data at the same resolution can provide a fine scale representation of climatic variability across the landscape. Yet, such addition of time series analysis of Landsat satellite data and high-resolution downscaling of climate data has not been reported in the literature for static-empirical modeling of SOC for heterogeneous agricultural landscapes.

In this study, we employed the static-empirical 'scorpan' modeling technique in a heterogeneous agricultural landscape in the Lower Fraser Valley (LFV), British Columbia, Canada to assess the spatiotemporal changes in SOC under LULC change and climate variability. The LFV provides an important case study for such analysis as it is representative of highly heterogeneous and fragmented agricultural landscapes and that have been farmed with increasing intensity in recent decades. In addition to the conversion of forests, grasslands, and wetlands to agriculture and urban development, the LFV has seen shifts from primarily annual production to perennial crop production along with the adoption of some of the conservation agricultural practices, like perennial grassland rotations, and the establishment of woody perennials on farm edges and along waterways. Thus, there is a need to understand how these shifts in LULC or management practices impacted the SOC dynamics in the LFV. In this connection, the specific objectives of this study were to – (1) produce LULC maps (30 m) of multiple time steps from 1984 to 2018 using Landsat satellite imagery, (2) generate SOC maps (30 m) of the same time steps using the static-empirical 'scorpan' based digital soil mapping (DSM) technique, and (3) using these products, evaluate the impact of changing LULC and climate on SOC across the landscape.

4.2 Methods

4.2.1 Study area

The study area, LFV represents the lower part of Fraser Valley (122° 40' 32.23" W, 49° 11' 19.33" N) in southwestern British Columbia, Canada and contains an area of 3031 km^2 (Figure 17). The LFV comprises a large elevation gradient from 0 m close to the Pacific Ocean to ~1400 m near the coastal mountain range. The region is characterized by a mild, humid maritime climate with a mean annual temperature of 11.1ºC and mean annual precipitation of 1189 mm based on the 30-year climate records (Environment Canada, 2019). Soils in the area were developed on glacial and post-glacial floodplains. The LFV is one of the most productive agricultural regions of British Columbia, responsible for >50% of the gross annual farm receipts of the province (Crawford and MacNair, 2012). To define the study area, we delineated a watershed from the provincial Terrain Resource Information Management (TRIM) digital elevation model (BC TRIM, 2012). Within this watershed, we constrained our analysis to the low-lying areas dominated by the agricultural lands. We then manually edited the eastern boundary to ensure that our study area extent completely captured the provincial agricultural reserve lands.

4.2.2 LULC change analysis

Based on the existing literature and initial reconnaissance survey, we identified seven LULC classes that likely have a strong connection with the SOC dynamics in LFV. Table 8 describes each of these seven classes, including annual crop (AC), perennial crop (PC), grassland (GL), forest/forest patches (FFP), built-up/bare land (BUBL), water, and wetland (WL). According to Yang (2010), the LFV experienced a shift from annual to perennial cropping after 1990 with a substantial change around 2000. Thus, we selected images that ensured that there was a gap of at least 5 years but no more than10 years to capture the LULC changes and images where cloud cover was <10% for the entire scene. Based on these constraints we produced LULC maps of 5 different time steps – 1984, 1990, 1999, 2009, and 2018.

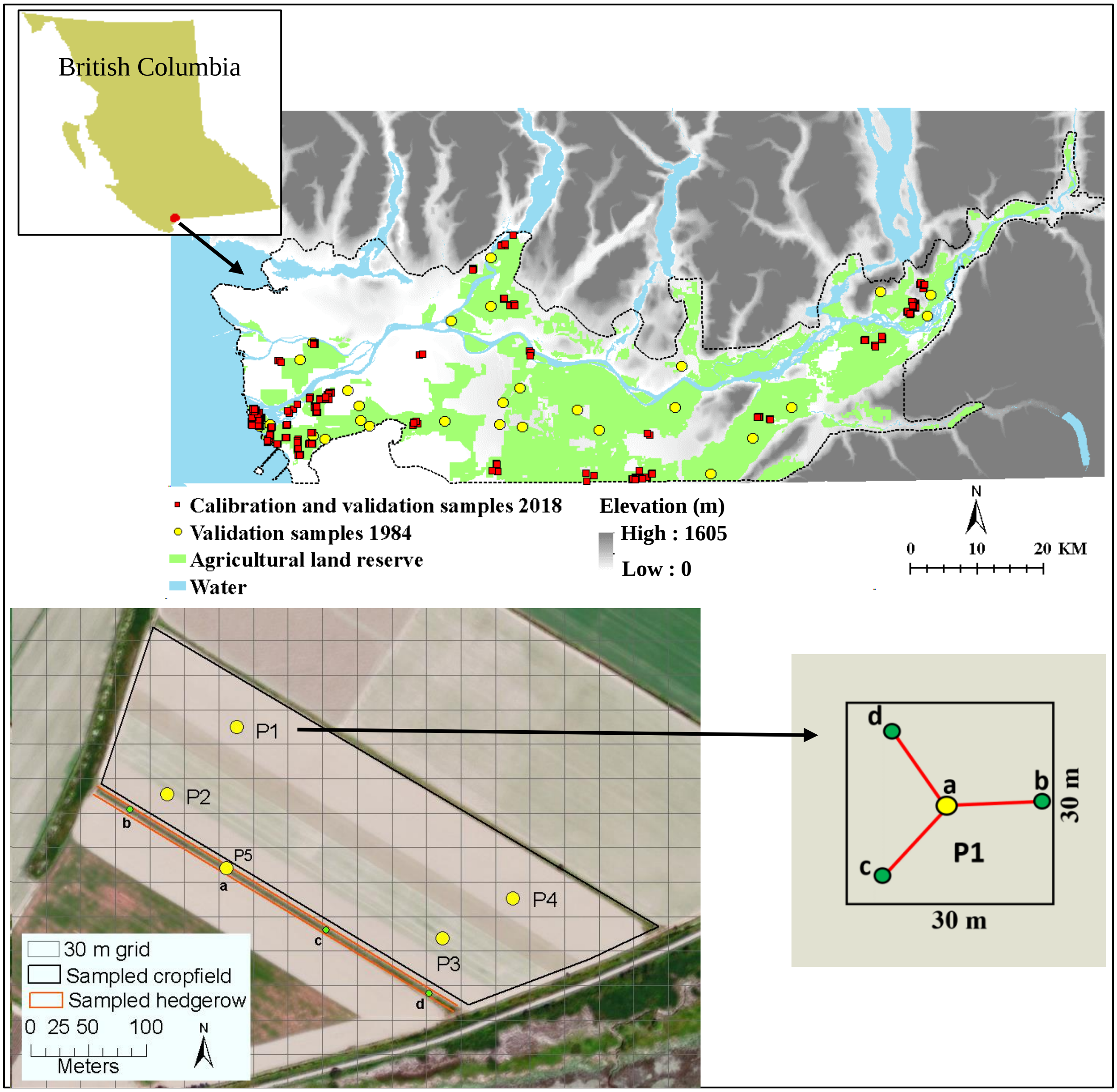

Figure 17: Map showing the study area with the location of soil samples

Upper panel shows the study area with the calibration and validation sample locations of 1984 and 2018. The background map of the upper panel shows the Terrain Resource Information Management (TRIM) digital elevation model (BC TRIM, 2012). Lower panel shows the sampling scheme where at each plot (i.e. P1, P2, P3, and P4), 4 sub-samples (a, b, c, and d) were composited to get a representative sample from an area of 900 m^2 - an area covered by a Landsat satellite image pixel (shown in the grid overlay). P5 plots were only sampled when there were field margins with perennial vegetation, e.g. hedgerow, grass margin, or riparian buffer. At P5 plots, samples were composited from a rectangular strip covering ~900 m^2 area depending on the length and width of the field margin.

We used Landsat 8 Operational Land Imager (for 2018) and Landsat 4-5 Thematic Mapper (for the previous years) surface reflectance images for LULC mapping. The images (Path 47, Row 26) were downloaded from the United States Geological Survey (USGS) Earth Explorer data warehouse. The images were acquired in the summer of each year within a period of 15 days in June or July. We used a hybrid change detection technique for our analysis which combined both pixel and image-based approaches. First, we derived the tasseled cap transformation (TCT) indices – brightness, greenness, and wetness for all images. We then performed an image segmentation on the 2018 TCT images to conduct an object-based image analysis (Blaschke et al., 2014; Paul et al., 2018). The image segments derived from the 2018 imagery were trained with the ground truth information collected from the field and Google Earth high spatial resolution imagery. While collecting the ground truth data (n = 650), we ensured that a minimum of 70 segments was sampled for each LULC type. We used 70% of the ground truth data for each LULC type for training the model and 30% for independent validation. Thereafter, we used a Random Forest (RF) model (Breiman, 2001), a decision tree-based machine learning model to perform the classification and produce a LULC map of 2018. For accuracy analysis, we used the validation data to derive an error matrix and calculate the user's, producer's, and overall accuracies and Kappa coefficient. Paul et al. (2018) provided a detail explanation of these accuracy measures.

Classifying each image back in time and then comparing classifications is not recommended as errors in each classification compound when comparing time periods (Wulder et al., 2018), rather mapped disturbances can be integrated into the 2018 land cover product, to provide land cover change information to the previous time steps. Therefore, we applied the technique described in Arnett et al. (2014) to generate the disturbance index (DI) for each of the previous time steps. The DI of time step 2 (TS2) was subtracted from time step 1 (TS1) to produce a univariate difference image where near-zero values would denote no change, positive value would indicate vegetation loss, and negative value would refer to vegetation increase (Arnett et al., 2014; Healey et al., 2005). We then trained the difference images from each time step to determine the threshold of 'change' and 'no-change' by on-screen sampling from the Landsat imagery and Google Earth imagery. We applied the training samples (n = 600; 200 for 'change' and 400 for 'no change') to derive the change maps showing the changes that occurred between TS1 and TS2 using an RF model. For verification, we then collected 120 random samples from each of the produced maps – 40 for 'change' and 80 for 'no change' and validated the maps using the accuracy measures - user's, producer's, and overall accuracies (Arnett et al., 2014). The change maps were then used as a mask to extract the tasseled cap indices for the pixels that were changed between TS1 and TS2. Using the tasseled cap information, the changed areas of 1984, 1990, 1999, and 2009 were then classified with the validated model of 2018. However, the unchanged areas of a specific year were assigned with LULC information of the subsequent time step, for example, the unchanged areas of 2009 were assigned the LULC types of 2018. Both outcomes were combined to produce a LULC map of that specific year and then, all five LULC maps for 1984, 1990, 1999, 2009, and 2018 were used for pixel by pixel change analysis.

Table 8: Description of the land use/land cover (LULC) types included in the study

LULC type	Description
Annual crop (AC)	Fields with annual crops, including vegetable, grains, and forages that are planted, harvested and typically tilled each production season. These fields may or may not have wintertime cover crops.
Perennial crop (PC)	Fields with perennial crops, that remain in the ground for more than one year, including berries and tree fruit.
Grassland (GL)	Managed and unmanaged grass and shrublands including both grazing land and natural meadows or parkland.
Forest/forest patch (FFP)	Lands with dense forest or tree patches within the built-up or agricultural areas.
Built-up/bare land (BUBL)	Built-up class mainly refers to urban settlement and road. Farmhouses and barns within the cropland areas are also included in the built-up class. Bare lands within the cropland areas primarily include fallow land and within forested areas, they refer to forest clear cut.
Water	Natural and human-made water areas.
Wetland (WL)	All types of wetlands.

4.2.3 Soil sampling and laboratory analysis for SOC measurement

We sampled 309 plots representing different LULC types (Table 9) across the study area and collected soil at the 0–15 cm depth. The AC samples included plots of annual vegetable and cereal crops, grassland fallows, and winter cover crops. Samples from perennial field margins, e.g. hedgerow, riparian buffers, and grass margins were collected in both AC and PC fields. Additionally, GL samples had plots from both managed and unmanaged grasslands. At each sampling plot, we collected a representative sample by compositing soils from four sub-sample locations that would fit within a Landsat image pixel (i.e., 900 m^2) (Figure 17). Each sampling plot was established in a homogenous LULC. To maintain a homogenous LULC, we followed a different protocol while sampling from the perennial field margins due to their limited width. For field margins, we again collected a composite sample from four sub-plots, but the configuration was linear to fit within a rectangular strip covering an area of ~900 m^2 (in some cases the area varied slightly depending on the length and width of the field margin). Thus, Landsat images of 30 m spatial resolution comprised of pixels that could encompass both field margin and in-field areas of either AC or PC (Figure 17). Given that SOC in the field margins is typically higher than that of in-field areas, the 30 m resolution mixed-LULC pixels averaged out the SOC content.

We recorded the center of each plot with a GNSS Pro 6H Differential Global Positioning System (DGPS) (Trimble Inc., Sunnyvale, California, USA) with post-processing accuracy varying from 10 - 50 cm. Out of the total, 25% of the samples (n = 75) were selected randomly within depths and LULC and sent to the Technical Service Laboratory of British Columbia Ministry of Environment. Samples were analyzed for total carbon using combustion elemental analysis (at 950º C) with a Flash 2000 Elemental Analyzer (Thermo Fisher Scientific, 2010) and inorganic carbon using Primacs SNC-100 TN/TC Analyzer (Skalar Analytical, 2019). For both analyses, soil samples were first dried, ground, sieved to a <2mm particle size and then again finely ground prior to analysis. SOC was then determined by subtracting inorganic carbon from total carbon. We also analyzed all the samples (n = 309) using mid-infrared spectroscopy with a TENSOR 37 spectrometer (Bruker Instruments, Ettlingen, Germany). We then utilized the spectra for predicting SOC with a partial least square regression (PLSR) model where the data from elemental measurement were used for calibration and validation.

Table 9: Total number of soil samples collected at 0-15 cm depth for various land use/land cover (LULC) types

LULC type	**Number of samples**
Annual Crop (AC)	132
Perennial Crop (PC)	100
Grassland (GL)	52
Forest/forest patch (FFP)	15
Wetland (WL)	10
Total	**309**

4.2.4 Environmental covariates for DSM

Topographic indices: In our analysis, a suite of topographic indices was generated from the provincial 25 m Terrain Resource Information Management (TRIM) digital elevation model (DEM) (BC TRIM, 2012) using SAGA 2.1.2 software. We resampled the DEM to 30 m and derived the indices – analytical hill-shading (AHS), aspect, catchment area (CA), channel network base level (CNBL), closed depressions (CD), convergence index (CI), cross-sectional curvature (CSC), longitudinal curvature (LC), slope-length factor (LS), multiresolution index of the ridge top flatness (MRRTF), multiresolution index of valley bottom flatness (MRVBF), negative topographic openness (NOpen), positive topographic openness (POpen), relative slope position (RSP), slope, terrain ruggedness index (TRI), total wetness index (TWI), valley depth (VD), and total curvature (TC). The description of the computation of these indices can be found here: http://www.saga-gis.org/saga_tool_doc/2.1.3/a2z.html. Paul et al. (2019a) and Schillaci et al. (2017) applied some or all these topographic indices for DSM at different scales.

Existing soil survey information: We extracted data from the existing detailed Canadian soil survey (AAFC, 2015) to derive environmental predictors for sand, silt, clay, organic carbon, and cation exchange capacity (CEC). We utilized a 30 m grid to extract data for each of these soil properties and produced raster layers using inverse distance weighting (IDW) interpolation. In the case of multi-component map units, the dominant category covering >50% of the unit area was used and assigned values accordingly.

Soil, vegetation, and image textural indices from Landsat imagery: We used three Landsat scenes captured in pre-growing (April/May), growing (June/July), and post-growing (September) seasons for each of the years in the time-series to detect the seasonal variability of the agricultural landscape. We utilized the images described in section 2.2 for the growing season. For the other two seasons, we downloaded images from the USGS data warehouse as mentioned in section 2.2. A group of soil and vegetation indices (Table 10) and several image textural indices, including homogeneity (Homo), contrast (Cont), and dissimilarity (Diss) were derived from each of the Landsat images. We applied the grey level co-occurrence matrix (GLCM) to produce the textural indices (Clausi, 2002) for the Near Infrared, Short-wave Infrared-1, and Short-wave Infrared-2 bands of the Landsat images.

Climate data: We included three climate variables – mean annual temperature (MAT), mean annual precipitation (MAP), and annual heat moisture index (= (MAT+10)/(MAP/1000)) for all study years. The climate data were collected and downscaled using the ClimateBC software which uses data from multiple sources, including PRISM, ANUSPLIN, and other global circulation models from Intergovernmental Panel on Climate Change (T. Wang et al., 2016). ClimateBC performs local downscaling of the climate variables based on the latitude, longitude, and elevation and generates a raster dataset at the resolution of the DEM which is 30 m for our case.

Table 10: Soil and vegetation indices used for predicting soil organic carbon

R, B, NIR, SWIR1, SWIR2 indicate red, blue, near infra-red, short wave infrared-1, short wave infrared-2 bands of Landsat satellite imagery, respectively, while L refers to canopy background adjustment factor

Soil and vegetation indices	Formula	References
Normalized Difference Vegetation Index (NDVI)	(NIR – R) / (NIR + R)	Rouse et al. (1974)
Soil Adjusted Vegetation Index (SAVI)	(1+L) (NIR – R) / (NIR+R+L), L=0.5	Huete (1988)
Normalized Difference Moisture Index (NDMI)	(NIR – SWIR1) / (NIR + SWIR1)	Hunt Jr & Rock (1989)
Soil Brightness Index (SBI)	$\sqrt{(R)^2 + (NIR)^2}$	Elvidge & Lyon (1985)
Normalized Difference Tillage Index (NDTI)	(SWIR1–SWIR2) / (SWIR1 + SWIR2)	Van Deventer et al. (1997)
Clay Minerals Ratio (CMR)	SWIR1 / SWIR2	Carranza & Hale (2002)
Bare Soil Index (BSI)	((SWIR1 + R) – (NIR + B)) / ((SWIR1+R) + (NIR + B)) * 100 + 100	Rikimaru et al. (2002)

4.2.5 SOC prediction and change analysis

At first, we predicted a SOC map for 2018 with RF using the SOC field data described in section 4.2.3 and environmental covariates explained in section 4.2.4. We applied 75% of the field data that were only analyzed with mid-infrared spectroscopy to calibrate the RF model. For model validation, we used the field data derived from the elemental combustion method (the current best practice). We utilized RF's variable importance (VI) measure for selecting the most influencing predictor variables. RF calculates the percent increase in mean square error (%IncMSE) of prediction by removing the variables one by one from the model and accordingly and this way, determines the importance of each variable (Breiman, 2001). We removed any variable from the model if it's VI score was below the threshold of 5% IncMSE. Once the top variables were selected, we performed an independent validation of the prediction using the 30% of the field data that were set aside. We used the coefficient of determination (R^2), concordance correlation coefficient (CCC), root mean square error (RMSE), and normalized root mean square error (nRMSE) to assess the accuracy of the predicted outputs as described in section 2.2.5. nRMSE is the normalized RMSE by the range of the observed data (Shen et al., 2016). We utilized this validated model to predict SOC for the rest of the years. The dynamic variables, i.e. Landsat and climate indices were updated for each prediction of 1984, 1990, 1999, and 2009 to capture the impacts of LULC and climate changes on SOC. We then validated the predicted map of 1984 with a set of archived soil samples collected in 1984. Because these soils were sampled by horizons and thus, were collected from inconsistent depth ranges and up to a maximum depth of 1 m, we needed to model SOC for 0-15 cm. Using mid-infrared spectroscopy, we re-analyzed a group of samples corresponding to different depth ranges and representing a total of 31 sites across the study area. We thereafter fitted a mass-preserving, continuous spline (Brendan P Malone et al., 2009) to predict SOC at 0 – 15 cm depth for all the sites. The spline predicted values were then applied for validating the map of 1984 using the same accuracy measures as before – i.e. R^2, CCC, RMSE, and nRMSE.

Finally, the predicted maps of all years were utilized for identifying the changes in SOC from 1984 to 2018 by pixel to pixel subtraction. We calculated the relative change in SOC (ΔSOC) in % using (Equation 12). We also calculated the relative changes in climatic variables, namely mean annual precipitation (ΔMAP), mean annual temperature (ΔMAT), and annual heat moisture index (ΔAHM). For example, (Equation 13) shows the calculation of ΔAHM. We then tested the pixel by pixel correlation between ΔSOC and the climatic variables to assess the impacts of climate changes on SOC. We applied the Pearson correlation coefficient for this purpose.

$$\Delta\text{SOC} = \left[\frac{(\textbf{SOC of year 2}) - (\textbf{SOC of year 1})}{(\textbf{SOC of year 1})} * \mathbf{100}\right] \qquad \textbf{(Equation 12)}$$

$$\Delta\text{AHM} = \left[\frac{(\textbf{AHM of year 2}) - (\textbf{AHM of year 1})}{(\textbf{AHM of year 1})} * \mathbf{100}\right] \qquad \textbf{(Equation 13)}$$

We performed all of our analysis in R 3.6.0 (R Core Team, 2018) while we utilized a number R packages including ‘raster’ (Hijmans and Van Etten, 2016), ‘rgdal’ (Bivand and Keitt, n.d.), ‘randomForest’ (Brieman et al., 2015), ‘glcm’ (Zvoleff, 2016) and ‘ggplot2’ (Wickham, 2016). We used ArcGIS 10.6 for map generation (ESRI ArcGIS, 2011).

4.3 Results and Discussion

4.3.1 Descriptive statistics of SOC for different LULC types

The plot level data shows that SOC values varied substantially for different LULC types (Table 11). We observed higher SOC values for WL, FFP, and PC although most of the high values of PC were attributed to the cranberry crop which is typically cultivated in organic peat soil. Higher

SOC values in AC class were mostly found in fields with woody vegetation in the non-productive field margins, however, some of the AC fields we observed with high SOC values were likely due to large quantities of organic inputs and potentially the integration of cover crops. The SOC content in the GL soils also exhibited a wide variation of values, but our GL class included both managed and unmanaged GL and higher values were mainly observed in the unmanaged GL within the parks. For WL, the lower SOC values were detected in the degraded WL areas which had been under pressure from the adjoining urban development and agricultural production. The WL within the protected parks and bogs exhibited substantially higher SOC values. The highest SOC values for non-organic soils were collected from FFP, with the lower end of these SOC values found in the forest patches.

Table 11: Plot level descriptive statistics

Statistics showing soil organic carbon (g/kg) for different land use/land cover (LULC) types

LULC type	Minimum	Maximum	Mean	Standard deviation
Annual crop (AC)	10.10	68.20	22.30	10.20
Perennial crop (PC) – without cranberry	9.30	73.90	26.60	16.10
Perennial crop (PC) – only cranberry	120.10	526.50	302.90	113.80
Grassland (GL)	9.10	115.50	24.60	21.30
Forest/forest patch (FFP)	2.00	171.90	80.80	79.40
Wetland (WL)	18.40	496.50	230.20	184.00

4.3.2 LULC changes from 1984 to 2018

Although we observed variable accuracies for different LULC classes in 2018 (Table 12), the changes between different years were consistently predicted with high accuracies (Table 13). The overall accuracy and kappa coefficient of the 2018 LULC classification were 0.81 and 0.77 respectively while the overall accuracies of the change analysis ranged from 0.89 to 0.94. The accuracy of our change analysis was close to the outcomes of Arnett et al. (2014) who used a similar technique based on disturbance index for change analysis of a forested landscape achieving overall accuracies ranging from 0.83 to 0.97. For individual classification, our prediction accuracy was relatively low for AC, GL, and WL due to their spectral similarities, especially for the cultivated grasslands (Eggen et al., 2016). In contrast, the prediction of FFP and water yielded the highest accuracies because of their distinct spectral signature and unique pattern on the landscape which is clearly distinguishable from the other LULC types (Paneque-Gálvez et al., 2013).

Table 12: Accuracy assessment of 2018 land use/land cover (LULC) classification

LULC Type	User's accuracy	Producer's accuracy	Overall accuracy
Annual Crop (AC)	0.78	0.76	
Perennial Crop (PC)	0.80	0.86	0.81
Grassland (GL)	0.74	0.68	
Forest/Forest Patch (FFP)	0.86	0.89	**Kappa coefficient**
Built-up/bare land (BUBL)	0.78	0.75	
Water	0.95	0.95	0.77
Wetland (WL)	0.73	0.83	

Table 13: Accuracy assessment of land use/land cover change detection

Year	Change class	User's accuracy	Producer's accuracy	Overall accuracy
1984-1990	Change	0.93	0.90	0.94
	No change	0.95	0.96	
1990-1999	Change	0.88	0.85	0.91
	No change	0.93	0.94	
1999-2009	Change	0.95	0.86	0.93
	No change	0.93	0.96	
2009-2018	Change	0.85	0.83	0.89
	No change	0.91	0.92	

Our study did not detect large areas of LULC changes during the study period – 1984 to 2018 showing consistently dominant LULC categories were FFP and built-up/bare land (BUBL) followed by a relatively even distribution of agricultural production types (Figure 18). We observed a gradual increase in BUBL from 1984 to 2018, 117.62 km^2 (14.70%) in total. During this period, FFP declined by 128.19 km^2 (13.75%) mainly due to conversion to BUBL and agricultural production. These findings support the results of Shupe (2013) who studied the LULC dynamics in the LFV (1966-2011) using the historical dataset from Canada Land Use Monitoring Program (CLUMP) for the years 1966, 1976 and 1986 and unsupervised classification of Landsat imagery for the years 1993, 2000, and 2011. Shupe (2013) documented that the most significant conversion of forests to BUBL occurred in the rural areas east of the City of Vancouver. We also observed a similar pattern, however, the conversion of forest to

BUBL was much lower in our analysis than that reported by Shupe (2013). It is likely that much of the forest loss reported in Shupe's study occurred before 1984. In addition, our BUBL class includes the bare lands within forested and cultivated areas, thus a portion of the increase in BUBL may be attributed to forest clear-cutting and agricultural production. Our pixel by pixel change analysis suggested that bare lands in the cultivated areas primarily refer to the annual crops that were bare at the time the image was captured but was likely not a change in land use.

Changes in the agricultural production area from 1984 to 2018 were relatively small. We detected a decrease in AC by 28.86 km^2 (8.45%) and GL by 17.50 km^2 (5.75%) but an increase in PC by 65.49 km^2 (30.43%). By 2018, we observed ~32% of the agricultural lands were occupied by PC. This finding is consistent with that of the agricultural land use inventory of Metro Vancouver (ALUI, 2016), which covers a slightly smaller area than our study area. This inventory reported that ~35% of the total agriculture land was used for perennial crop production (i.e. berries, nursery, and tree fruit) in 2016 (ALUI, 2016). Although the area of water body remained the same throughout the study period, we observed a decline of WL by 23.56 km^2 (24.89%). WL was found to be primarily converted to PC and BUBL. These results showing a reduction in the WL area agree with the findings of Wilson (2010) who reported a total decline of 13.6 km^2 of WL area between 1989 and 2009; our analysis detected a decline of 16.86 km^2 between 1990 and 2009. Similar to our observations, Wilson (2010) also described that the major portion of the WL loss was due to conversion to berry production or other agricultural uses.

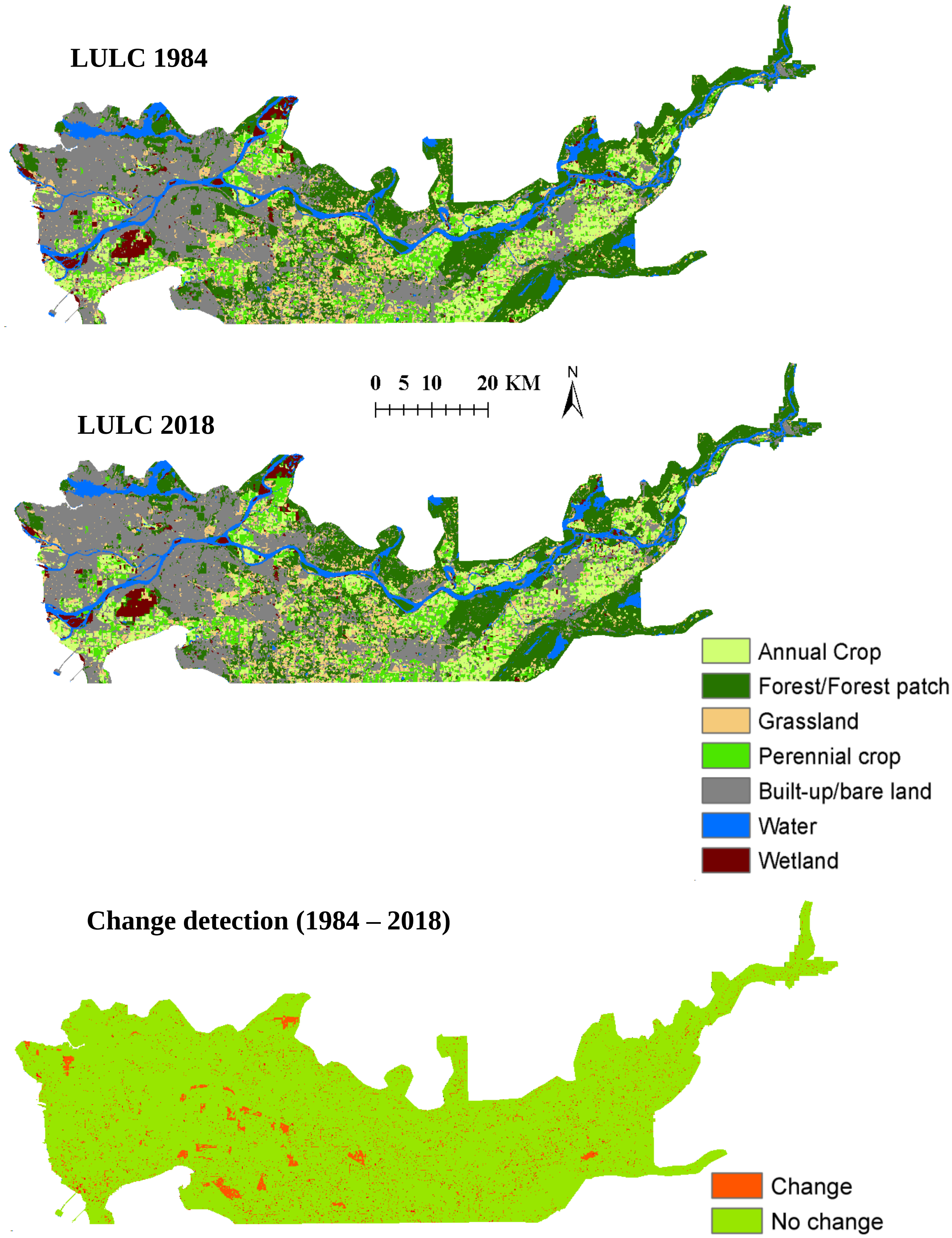

Figure 18: Land use/land cover (LULC) maps of 1984 and 2018 and the change over that period for the Lower Fraser Valley

4.3.3 Accuracy and variable importance for SOC prediction

We achieved variable accuracies for the independent validation of the predicted maps of SOC in 1984 and 2018 (Table 14). When we tested the accuracy of 2018 prediction for two cases, i.e. with and without Landsat indices, we found that accuracy was reduced by 8.35% for R^2, 11.09% for CCC, and 9.40% for nRMSE when all Landsat indices were removed from the prediction. In addition, the predicted map became extremely grainy and patchy in some parts of the region without the Landsat indices. Thus, for smooth, continuous prediction of SOC with higher accuracy, we kept all categories of environmental predictors, including topographic, soil survey, Landsat, and climate variables in the final map generation. Initially, we had a total of 86 variables but by using the threshold of variable important (VI) score (i.e. 5% IncMSE) in RF, we removed the less influential variables and the final model only included 22 variables in total (Figure 19). The historic soil survey variables became the strongest predictors with VI score of 12-19%. The VI scores of topographic predictors ranged from 5-10% while climate variables had a VI score of 5-7%. The VI score of Landsat indices ranged from 5-11% including several variables derived from the images of pre- and post-seasons when bare soils were exposed. Our analysis indicating small differences in the contribution of the various types of environmental predictors was consistent with some studies and not others. For instance, Grinand et al. (2017) reported that Landsat indices were the most important predictors for SOC but Schillaci et al. (2017) found topographic variables were the strongest predictors. Grimm et al. (2008), based on their study in Barro Colorado Island, determined that information from existing soil maps or surveys had little influence on the SOC dynamics rather terrain attributes and current biomass information were stronger predictors of SOC; the opposite of our findings. Based on the evidence to date we can conclude that predictor importance is highly reliant on the particular environmental conditions of the study site, particularly the topography and no consensus can be reached on the best predictors of SOC (Were et al., 2015).

The accuracy of the final 2018 map (Figure 20) was fairly high relative to other studies of SOC. Alternatively, the accuracy of the 1984 prediction was substantially lower than the prediction of 2018. There might be several reasons for the relatively low accuracy of the 1984 prediction. The first, soil sampling in 1984 followed different protocols where samples were taken by soil horizon from soil pits thus the sampling depths were not consistent with our sampling in 2018. To account for this, we used spline prediction to determine the SOC value at 0-15 cm depth which likely introduced an element of error in the 1984 validation dataset. The second reason, in 2018, we collected a composite sample from an area covering that of a Landsat image pixel (i.e. 900 m^2) but in 1984, samples were collected from a single point location and although the X, Y coordinates were recorded, the type of GPS might also contribute to the associated error. As a result, part of the discrepancy between the predicted values and the validation data is likely due to spatial location and scale mismatch. Despite these issues, the accuracy of the SOC map of 1984 was similar to the SOC prediction of others using archived soil samples. Huang et al. (2019), for example, found R^2 and CCC of 0.48 and 0.67 using a similar validation approach for SOC values extracted from archived soil samples collected between 1980 and 2002 for a study in Wisconsin, USA. Alternatively, S. Wang et al. (2016) using the historical soil survey dataset from a study area in China for SOC prediction in1990, yielded an R^2 of only 0.43, by applying 10-fold cross-validation instead of a more rigorous independent validation as we performed.

Table 14: Accuracy of soil organic carbon (SOC) prediction for 2018 and 1984.

R^2, CCC, nRMSE refer to the coefficient of determination, concordance correlation coefficient, and normalized root mean square error respectively

Accuracy Metrics	2018		1984
	With Landsat indices	Without Landsat indices	
R^2	0.670	0.614	0.459
CCC	0.757	0.673	0.584
nRMSE	0.117	0.128	0.183

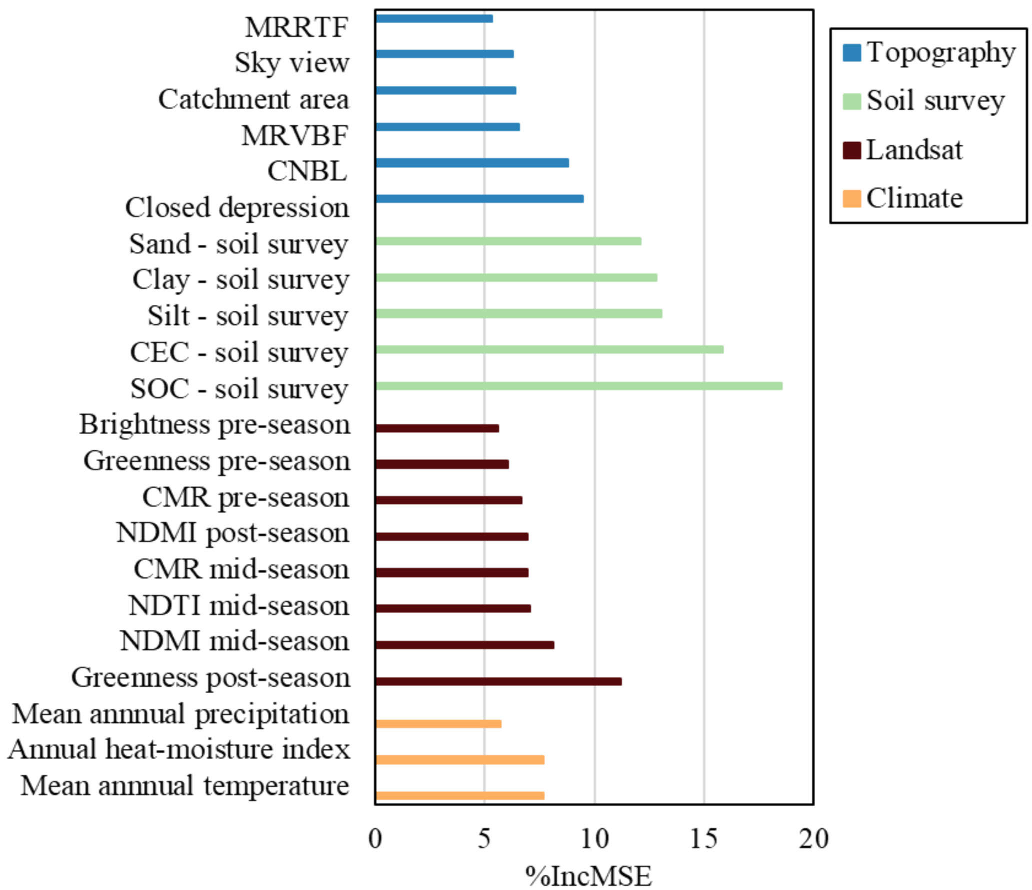

Figure 19: Variable importance of random forest model for predicting SOC in 2018

Different colors represent four categories of environmental variables. See section 4.2.4 for the full names of the variables.

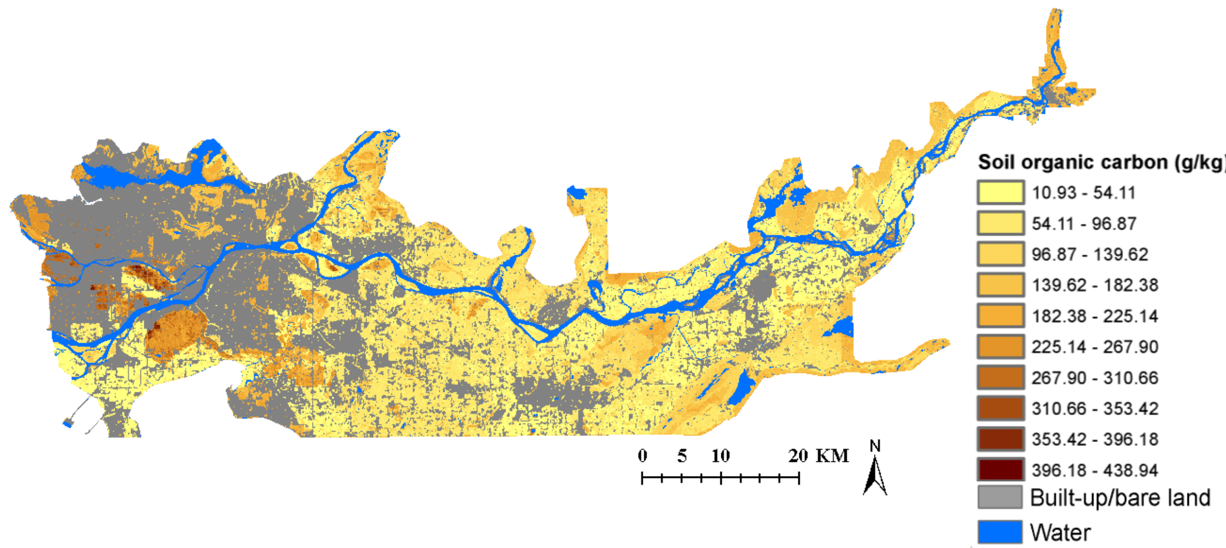

Figure 20: Distribution of soil organic carbon (SOC)

Across the Lower Fraser Valley in 2018. SOC was not predicted in the built-up/bare land and water areas.

4.3.4 Distribution of SOC across the LFV and changes in SOC (1984 – 2018)

The predicted maps showed that the distribution of SOC across the LFV was fairly similar between 1984 and 2018 with higher SOC values concentrated in the WL and cranberry fields in the western and north-central parts of the area. High SOC values distributed along the course of Fraser River were mainly in WL and forested soils. The forests in the southeastern corner of LFV also had higher concentrations of SOC. Low SOC values were observed in croplands; with the lowest values in the croplands in the southwestern and northeastern parts of the valley.

As expected, WL that had not had any change from 1984 to 2018 (i.e. WL-no change; Figure 21a) had the highest predicted mean SOC with 167.88±69.70 g/kg (i.e. mean ± standard deviation). This was followed by FFP-no change, with 114.34±42.31 g/kg. AC-no change fields had far lower SOC with 58.27±44.32 g/kg in 2018 while GL-no change was 73.82±48.65 g/kg and PC-no change was 83.37±59.82 g/kg. The large differences between wildland land covers and agricultural land uses suggest that LULC types are a key determinant of the SOC variability across a landscape. Such large differences in SOC between cultivated and uncultivated areas were also reported in other studies (Priyanka, 2018; Wu et al., 2009). For example, in a landscape-level study in southwestern Yunnan province of China, Liu et al. (2015) observed nearly twice the SOC, at 0-20 cm depth, in forested lands with SOC ranging from 30 to 60 g/kg while cropland and grassland SOC ranged from 10 to 20 g/kg.

Differences in SOC for some LULC classes were even more dramatic when their changes in land use were considered (Figure 21a). For example, AC-no change and AC-from FFP had similar predicted mean SOC in 2018 but it was about 75% higher for AC-to/from GL which indicates AC and GL rotations. Conversely, the predicted mean SOC of GL-from FFP was 20% higher than that of GL-no change in 2018. Similarly, we observed notable variation in the predicted mean SOC of different PC fields with PC-no change having the lowest and PC-from WL having the highest mean values. FFP-no change and FF-to/from clear cut (CC), however, did not exhibit much variation and their predicted mean SOC differed only by 1%.

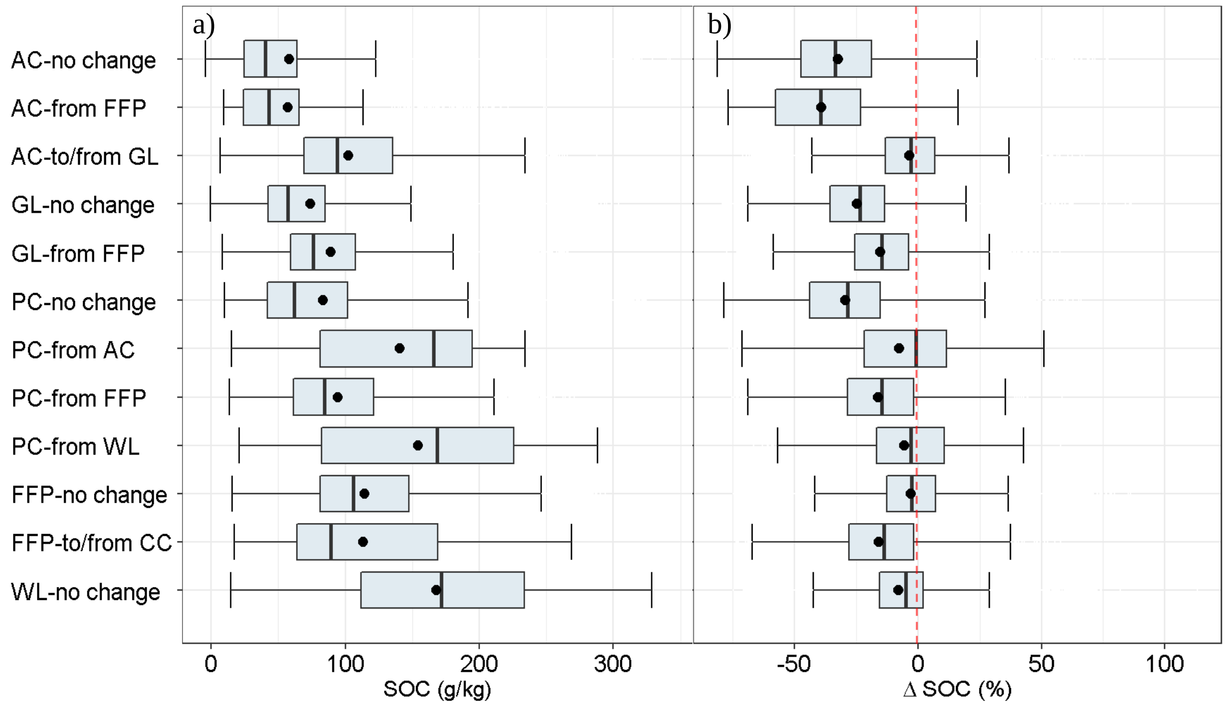

Figure 21: a) SOC in 2018; b) ΔSOC from 1984 to 2018

Boxplots showing the first quartile, median (bar), third quartile, and mean (circles) of soil organic carbon (SOC) content in 2018 and relative changes in SOC (ΔSOC) from 1984 to 2018 for different land use/land cover changes

In the LFV, SOC dynamics from 1984 to 2018 were highly variable with some areas gaining SOC, some with substantial loss, and others with no change. While there was a wide variation in mean SOC change associated with specific LULC conversion, there was also a great deal of variability within each LULC type or change class (Figure 21b). Overall, we observed a decline in the predicted mean SOC for all LULC conditions regardless of whether there were changed or not. From 1984 to 2018, the largest declines were in AC-no change and AC-from FFP with mean ΔSOC of -32.27% and -39.07% respectively. Conversely, FFP-no change had the lowest decline in SOC with mean ΔSOC of -2.92% followed by AC-to/from GL with -3.49%. The conversion of FFP to different agricultural practices had large decreases in SOC with the highest decline for changes to AC (mean ΔSOC of -39.07%), followed by PC (-16.18%) and GL (-15.29%). The SOC decrease for the conversion of other LULC to PC was relatively low compared to the decrease associated with the changes of other LULC to AC. The changes from WL to PC were mainly due to the conversion to cranberry production and despite the intensive agricultural production, we observed only minimal SOC loss for this change class (i.e. PC-from WL) with mean ΔSOC of -4.51%.

Interestingly, we detected large SOC losses for all agricultural practices that had not changed since 1984, with mean ΔSOC ranging from -24.84% to -32.27%. The decline in AC and GL are likely due to intensive management practices, such as heavy tillage (Guo and Gifford, 2002; Haddaway et al., 2017). We did, however, also observe similar losses of SOC in PC-no change where there is minimal annual soil disturbance through tillage. Losses in these perennial systems may not be comparable to AC because of the relatively small biomass or C amendment inputs compared to what is typical for AC (Nemeth et al., 2017). In PC, there may be substantial applications of nitrogen, which may be leading to microbial respiration and loss of SOC (Paal et al., 2011). Another possible explanation for SOC losses across the agricultural land uses could be due to changes in soil drainage conditions. Our study area receives a large amount of rainfall during the winter and many farmers have installed tile drainage to improve soil trafficability at the beginning and end of the production season for AC fields and to keep the water table from saturating roots in PC fields (Neufeld et al., 2017). Improved drainage can result in greater microbial access to oxygen during these periods which could be leading to an increase in the oxidation of carbon and the large decline in SOC (Baker et al., 2007; Schaufler et al., 2010).

Many of these drainage systems may have been installed prior to 1984 but continue to contribute to SOC loss (Lalonde and Hughes-Games, 1997).

Altogether, 42% of the total land area in 2018 (excluding BUBL and water) was agriculture that had not changed since 1984 (Figure 22). Thus, the SOC decline in these LULC classes had a significant influence on the overall SOC dynamics across the region. Alternatively, the FFP-no change, which had minimal SOC loss, was similar in area, accounting for 44% of the study area. Therefore, conserving these FFP areas and/or managing them carefully are highly important for maintaining SOC stock across the LFV. Although mean values indicated SOC losses for each of the LULC types and change classes, it is also important to recognize that for each of these we observed a great deal of variability. For all LULC types and change classes, we also observed large increases in SOC in some parts of the landscape over the study period. In the agricultural lands, the positive ΔSOC may be attributed to various conservation management practices (e.g. the application of organic amendments, crop rotations, winter cover cropping, and establishment of perennial vegetation in field margins). Additionally, the transition of LULC from AC to PC or AC to GL resulted in some of the highest positive ΔSOC values we observed (51% and 38% respectively). It is possible that some of these extreme positive values may represent the outcome of longer duration of LULC change impacts as it usually takes time to exhibit a net positive increase than losses (Poeplau et al., 2011; Xiong et al., 2014). Hence, some of these LULC types or change classes may ultimately show a net gain in SOC in the future based on the type of conversion or specific management practices employed.

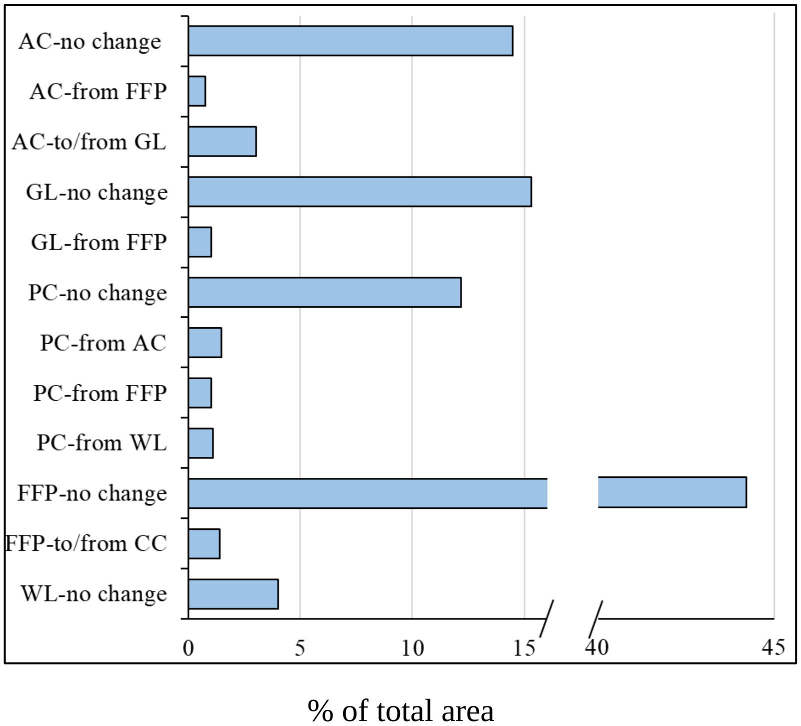

Figure 22: Proportion of different land use/land cover (LULC) change classes in 2018
LULC change classes affecting soil organic carbon dynamics in 2018. Total area excludes built-up/bare land and water classes

Our findings on the impacts of different LULC changes on SOC are in agreement with various local and global scale meta-analysis which reported decreases in SOC for conversion of FFP to crop production and increase in SOC for AC to PC conversions and rotations of AC and fallow or GL (Poeplau et al., 2011; Stockmann et al., 2013; Vaccari et al., 2012). For example, Bruun et al. (2015) reported a SOC decline of 20-40% after 20 years following the conversion of FFP to AC while Chen et al. (2007) documented 19% and 34% increase in SOC for the conversion of AC to fallow and perennial tree fruit production respectively over a study period of 27 years. A Canadian study found that changes from AC to PC increased SOC by 44% at a long-term study site in Breton, Alberta (VandenBygaart et al., 2010)– this value matches with some of the positive ΔSOC we observed in our analysis for AC to PC conversion (Figure 21b). Many of these studies also investigated the impacts of other LULC changes and management on SOC that were not possible to detect given our approach. For example, increases in SOC were detected by Del Galdo et al. (2003) of 23% after 20 years following the conversion of AC to FFP, Chen et al. (2007) found the conversion of AC to permanent GL resulted in 102% after 27 years, and Dimassi et al. (2013) found the transition from heavy tillage to no-tillage practices resulted in a 53% increase after 20 years.

We detected variable and increasing rates of SOC change between different time periods in the LFV (Figure 23a). The mean annual rate of relative change (ΔSOC%/year) for the entire period was -0.41%/year, ranging from -2.18%/year in some parts of the study area to 1.94%/year in others. The highest rate of decline happened during 2009-2018 and 1999-2009 with mean rates of -1.03%/year and -0.54%/year respectively. The SOC decline during 1984-1990 and 1990-1999 was relatively minimal with mean rates of only -0.06% and -0.10% respectively. The variation in SOC decline between different time periods was likely associated with shifting crops or agricultural management practices given that much of the landscape had no LULC change. Areas with the most change, however, would have been driven by LULC change (Figure 24). Our study area falls within two census divisions of Statistics Canada – partly in Metro Vancouver and partly in Fraser Valley Regional District. From the agricultural census data, we identified some shifts in crop production with increasing and decreasing trends in the acreage of various crops between different years from 1996 and 2016 (Statistics Canada, 2016) which likely had connections with the variability we observed in SOC decline. For instance, there was a decline of 13-54% in seeded and natural pasture areas at different years during this period. At the same time, there was an 80-91% increase in more tillage intensive field crop (i.e. grains, potatoes, field peas, etc.) acreage in both census divisions during the same period. These shifts in crop production may have a substantial influence on the SOC decline in our study area. In addition, the census data reported increases in the acreage of perennial fruit, berry and nut production by 135% and 53% in Metro Vancouver and Fraser Valley Regional District respectively – which might be associated with the SOC gain that we captured in some parts of the study area. The absolute SOC loss in the LFV during 1984 – 2018 was detected as high as 126 g/kg, however, in most parts, SOC loss ranged from 10 to 60 g/kg (Figure 24). Overall, the SOC remained unchanged in 27% of the area, declined in 61% and increased in 12%. Assuming bulk density did not change during the study period and taking the mean of bulk density samples (n = 309) collected in 2018, we estimated a total loss of 5.8 Mt of SOC or 0.56 ton of SOC per hectare per year in the LFV from 1984 to 2018. This loss of SOC could be attributed to either as oxidation to the atmosphere as CO2 or movement to the rivers or ocean due to soil erosion. However, our low-lying study area is characterized by more or less uniform topography with 80% of the area having a slope of <5%. Thus, it is likely that SOC loss due to soil erosion was minimal and SOC was mostly lost to the atmosphere.

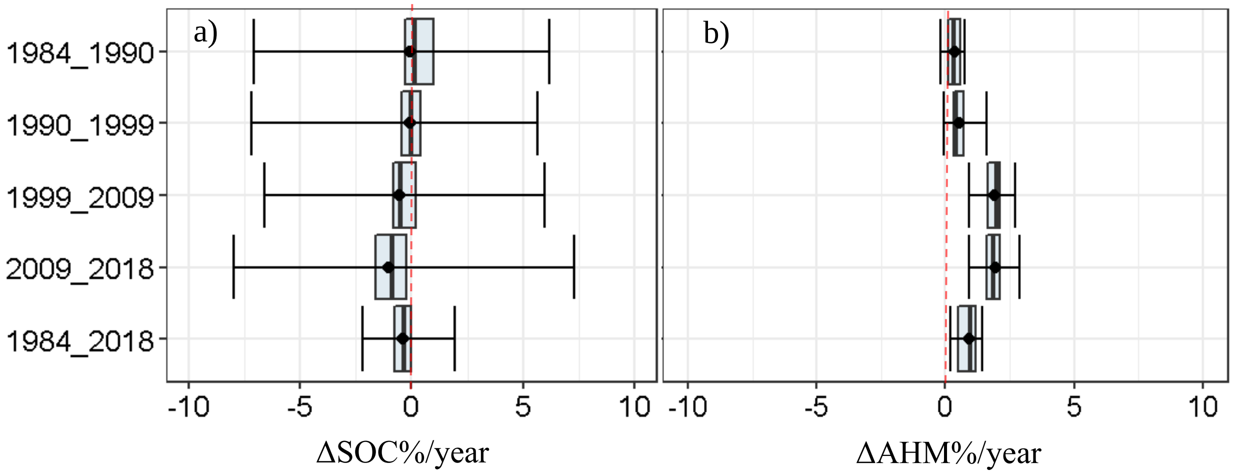

Figure 23: Rate of change in a) ΔSOC and b) ΔAHM between different time periods

Boxplots showing the first quartile, median (bar), third quartile, and mean (circles) of the annual rate of relative change in soil organic carbon (ΔSOC%/year) and an annual rate of relative change in annual heat-moisture index (ΔAHM%/year) between different time periods

It is unlikely that the variability in ΔSOC we observed was related to climate variability (Figure 23b). As explained above, climate variables did not perform as strongly as the other environmental variables for predicting SOC in the LFV using the RF model. Given the small rate of relative change in ΔAHM from 1984 to 2018, which varied between 0.19%/year and 1.44%/year, we observed a poor correlation between ΔSOC and ΔAHM when pixel by pixel changes were analyzed in a linear regression model (data not shown). These results were consistent with Adhikari et al. (2019), who used a similar DSM approach to study future SOC changes across Wisconsin, USA and found that SOC was also highly influenced by LULC and parent material characteristics and not by climate change in 2050. Another researcher in the same study area by Huang et al. (2019) assessed the historical SOC changes and attributed the improved SOC between 1980 and 2002 to sustainable LULC management. Bui et al. (2009) suggested that climate variability may play a critical role in SOC changes at larger regional to continental scales. The impacts of climate change may not have been evident given our study period only spanned 34 years and the extent of the study area was small enough, and level enough, to maintain a relatively consistent temperature and precipitation patterns.

Although our study provided a baseline for landscape-scale SOC estimation in the LFV and changes since 1984, integration of more detail information on agricultural management practices could help identify more specifically which management practices are driving SOC change. Landscape-wide information on tillage practices, fertilizer application, winter cover cropping, etc. could enhance the prediction process of SOC, but this information was not available at a spatial resolution relevant for our study area. Longer-term historical data or larger spatial extents could capture greater variability and thus more effectively simulate the impacts of LULC and climate change on SOC. Finally, although our static-empirical modeling approach performed for our study as well or better than in different parts of the world (Bonfatti et al., 2016; Reyes Rojas et al., 2018; Yigini and Panagos, 2016), it is unclear how this approach compares to process-based models or how this data-limited approach would perform for predicting SOC changes when climatic change is more dramatic.

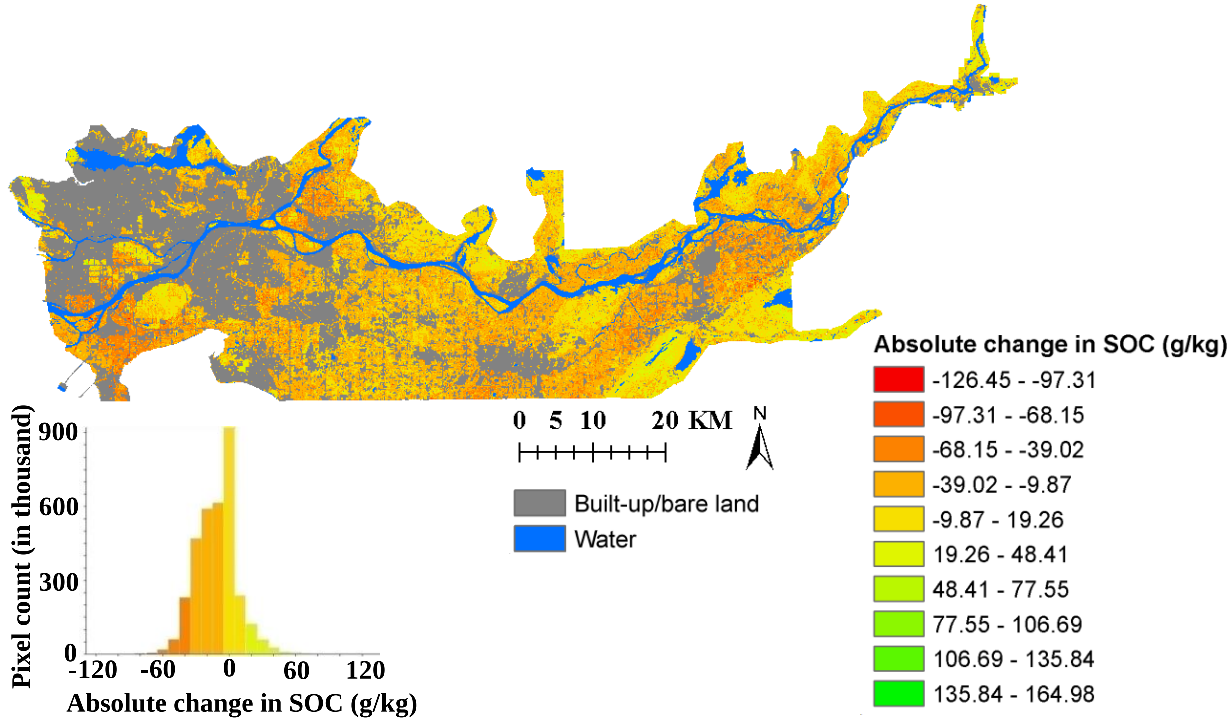

Figure 24: Absolute changes in soil organic carbon (SOC) from 1984 to 2018

The built-up/bare land and water area extent displayed in the map is for 2018

4.4 Chapter conclusions

Spatiotemporal analysis of SOC dynamics in response to changes in LULC, temperature, and precipitation could be an important tool for developing effective climate change mitigation and adaptation strategies for agricultural landscapes. Using remote sensing, digital soil mapping, and machine learning to establish a 'scorpan' based static-empirical SOC model, we predicted and assessed the SOC dynamics in the LFV from 1984 to 2018. During the study period, we detected SOC loss in 61% of the study area while SOC gain and no change were observed in 12% and 27% of the study area respectively. Over the 34-year time period, SOC across the region declined at a rate of 0.41%/year. We identified sizable losses of SOC due to LULC changes, but the majority of the losses were attributed to consistent agricultural land uses. We found LULC type and change classes to be important predictors for SOC changes whereas climate variability had only a minor influence during this study period. This study demonstrates the efficacy of a simple and cost-effective methodology to monitor SOC changes at landscape-scale that can easily be updated by incorporating new data. Cost-effective monitoring of SOC at a landscape scale enables the identification of LULC and management strategies that maximize SOC as well as the possibility of developing regional incentive-based programs for C sequestration.

Chapter 5: General conclusion

In this dissertation, I developed innovative approaches for DSM using remote sensing to monitor different soil properties at the field- and landscape-scales for enhancing climate change mitigation and adaptation in the LFV of BC. The research was intended to address three specific objectives:

1) Evaluate the cost and accuracy of geostatistical modeling technique at varying sampling efforts of SLA and MIRS for DSM at the field-scale,
2) Compare two machine learning approaches for DSM of SOC and clay (CL) to predict a soil workability threshold (WT) at the landscape-scale,
3) Assess the spatiotemporal dynamics of SOC in response to historical changes in LULC and climate conditions.

I addressed each objective as individual research chapters (chapters 2-4). In the following sections, I provide a summary and key research outcomes of each of these chapters, describe the limitation of the study, and identify future research needs.

5.1 Synopsis of Chapter 2

The research presented in Chapter 2 was conducted at a 54-ha field in Delta, BC where I compared different sampling efforts of DSM for predicting a suite of soil properties, including sand, silt, clay, pH, salinity, organic matter, and total nitrogen. This study utilized samples analyzed with standard laboratory analysis (SLA) and mid-infrared spectroscopy (MIRS) and compared their efficacy for the prediction of the soil properties at equivalent sampling efforts. I employed Conditioned Latin Hypercube Sampling (cLHS) to generate a range of sampling efforts from the full SLA (n = 62) and MIRS (n = 308) datasets. I applied a group of environmental variables derived from a high-resolution digital elevation model, unmanned aerial vehicle imagery (UAV), and historical soil survey data. Kriging with external drift model was used for predicting the soil properties at 5 m resolution across the study field.

In this study, DSM outputs were found most effective for model accuracy and cost at 50-60% of the full sampling effort. MIRS predictions of soil properties using a partial least square regression model introduced a sizable amount of error when compared with the SLA dataset. However, DSM outcomes using the MIRS dataset were more accurate than those using the SLA dataset at equivalent sampling efforts since MIRS enabled five times more samples and better captured the spatial variability across the study site. The prediction accuracy for digital soil maps varied across the soil properties. At the optimum sampling effort, the R^2 of prediction ranged from 0.82 (for sand) to 0.45 (for total nitrogen). From this analysis, I concluded that spatially optimized sampling efforts and the use of the MIRS technique substantially improve the model accuracy and cost efficiency of field-scale DSM for multiple soil properties. The cost-effective DSM strategies identified in this chapter can be used for precision management of various soil properties at a farm level and will enhance mitigation and adaptation abilities of the producers.

5.2 Synopsis of Chapter 3

In Chapter 3, I presented landscape-scale mapping of SOC and CL to predict WT for the agricultural lands of Delta, BC, comprising an area of 120 km^2, which has been facing significant challenges due to the changing precipitation pattern and poor drainage. This study utilized multi-temporal Landsat imagery for the year 2016 and two machine learning models, namely random forest (RF) and generalized boosted regression model (GBM) for mapping SOC and CL. I derived a suite of environmental covariates, including topographic indices from the digital elevation model, soil and vegetation indices from multiple Landsat images representing pre-, growing, and post-seasons, variables from existing soil survey and agricultural land use inventory for the mapping of SOC and CL. These maps were then applied to existing pedotransfer functions (PTFs) to predict WT across the landscape.

The independent validation showed that the RF model outperformed GBM for predicting SOC and CL although both models performed reasonably well. The R^2, CCC, and nRMSE of prediction using the RF model were 0.55, 0.70, and 0.12 for SOC, while they were 0.62, 0.72, and 0.15 for CL. Topographic indices were the strongest predictors for both models and for both SOC and CL, while Landsat and existing survey variables also had a notable contribution. The SOC and CL maps were spatially applied to several PTFs to estimate the plasticity limits of the soil and produce the WT map which was then tested against independent field samples of the soil water content at -10 kPa. The validation yielded an R^2 of 0.59, CCC of 0.70, and nRMSE of 0.15. My analysis showed that 40% of the fields in the study area had WT <30% and were particularly vulnerable to increased precipitation in the shoulder seasons, and subsequently, likely to have reduced workable days. My analysis demonstrated an effective approach for producing high-resolution WT maps that could be utilized for enhancing spatial prioritizations for field management or investments for climate change adaptation at the farm to regional scales in regions facing similar drainage challenges as Delta, BC.

5.3 Synopsis of Chapter 4

In Chapter 4, I presented a spatiotemporal analysis of SOC in response to LULC change and climate change from 1984 to 2018 across the LFV, comprising an area of 3031 km^2. The study utilized a static-empirical modeling approach and used time-series Landsat images and climate variables for the years 1984, 1990, 1999, 2009, and 2018 for tracking changes in SOC and LULC. For LULC change detection, I conducted a hybrid analysis, combining both pixel- and object-based approaches, of the Landsat imagery for each of the years using a RF model for 7 LULC classes, namely annual crop (AC), perennial crop (PC), grassland (GL), forest/forest patches (FFP), built-up/bare land (BUBL), water, and wetland (WL). To predict SOC, an RF model was calibrated and validated for 2018 while Landsat indices, climate variables, topographic indices, and soil survey variables were used as environmental covariates. For predicting SOC for the rest of the years, the Landsat and climate variables were updated to represent the variation in those specific years, however, the static variables, i.e. topographic indices and soil survey variables (e.g. soil texture) remained unchanged. The prediction of SOC in 1984 was also validated using a set of archived soil samples.

The LULC classification yielded an overall accuracy of 81% and a kappa coefficient of 0.77 for independent validation using the ground truth data of 2018. However, the accuracy varied for the individual LULC classes with relatively low accuracy observed for AC, GL, and WL while FFP and water classes were predicted with the highest accuracy. The accuracy of LULC changes between different years ranged from 0.89 to 0.94. My predictions of SOC in 2018 resulted in an R^2 of 0.67, CCC of 0.76, and nRMSE of 0.12, while SOC prediction of 1984 yielded R^2, CCC, and nRMSE of 0.46, 0.58, and 0.18, respectively.

My LULC analysis did not identify substantial changes during the study period – 1984 to 2018 in the LFV. The dominant LULC categories were FFP and BUBL while different types of agricultural production were evenly distributed. The area of AC declined by 8.45%, while PC observed an increase of 30.43%. In 2018, the predicted SOC content across the study area ranged from 10.62 g/kg to 371.56 g/kg. I measured a mean annual loss in SOC of 0.41%/year (median - 0.34%/year) across the landscape from 1984 to 2018. I detected SOC loss in 61% of the study area and gain in only 12%, while 27% remained unchanged. The largest losses in SOC were due to LULC changes, yet the majority of the SOC losses across the landscape were attributed to areas that were consistently in the same type of agricultural production. Overall, I observed a decline in the predicted mean SOC for all LULC conditions regardless of whether there were changed or not. Although SOC changes were strongly associated with LULC changes, climate variability did not have a strong effect on SOC changes. The outcomes of this study have provided a solid baseline for the SOC status of the LFV and showed a concerning rate of loss since 1984. The study also identified areas in the LFV where there is a substantial potential to sequester SOC by implementing supportive management strategies for enhancing climate change mitigation and adaptation.

5.4 Key research outcomes

The research presented in this dissertation provided cost-effective methodologies for accurate mapping of a suite of soil properties at the field- and landscape-scales in the LFV. The key research findings include:

- Chapter 2
  1. The integration of spatially optimized sample selection and MIRS techniques substantially reduced the sampling efforts needed for effective field-scale DSM of multiple soil properties.
  2. Two to three samples per hectare (i.e. 50-60% data points of the initial sampling effort derived from a 40 x 40 m grid) were found to be optimal for accuracy of DSM models and the cost of analysis.

- Chapter 3
  1. Random Forest model outperformed Generalized Boosted Regression model for predicting SOC and clay
  2. 40% of the fields in the study area had WT <30% and were particularly vulnerable to increased precipitation in the shoulder seasons
  3. WT map can be used for on-farm and regional agricultural management for effective climate change adaptation

- Chapter 4
    i. ‘Scorpan’ based static-empirical modeling predicted SOC changes (1984 – 2018) across the LFV with R^2 as high as 0.67
    ii. 61% of the LFV area experienced SOC loss, 12% gained SOC, and 27% remained unchanged
    iii. The predicted mean annual loss in SOC was 0.41%/year (median -0.34%/year) from 1984 to 2018

5.5 Limitations and future research

While this research established and applied innovative DSM tools for evaluating and monitoring soil properties at the field- and landscape-scales, it had some methodological limitations and constraints for application. The field-scale analysis, presented in Chapter 2, was only conducted in one crop field. Thus, the findings of the field-scale analysis may differ in other fields with different soil and topographic characteristics. The optimum sampling effort that I identified can be different for soils with lesser or greater variabilities. Therefore, future analysis should replicate the study in multiple fields with variable site conditions and management history to assess how these outcomes may vary based on soil type, environmental predictors, or management practices. In addition, the UAV imagery I used in this analysis, only comprised of visible bands of the electromagnetic spectrum. Including imagery at different infrared bands could better capture the vegetation information (Knoth et al., 2013) and contribute to better prediction accuracy. Future research, thus, should incorporate UAV imagery beyond the visible bands.

The landscape-scale analysis in Chapter 3 produced a continuous map of WT for the agricultural lands in Delta, however, the PTFs used in this analysis were based on Atterberg's plasticity limits, which are mainly suitable for high clay soils. Thus, the technique may not be tenable in agricultural lands dominated by coarse-textured sandy soil. Future research should validate the methodology in other locations with different soil texture or develop PTFs for variable soil conditions. Additionally, my analysis only incorporated maps of SOC and CL for predicting WT across the entire landscape where there was a lot of variation in agricultural management practices. As such, the future analysis may consider incorporating landscape-wide management information for better prediction of WT.

In Chapter 4, my analysis produced a baseline for landscape-scale SOC estimation in the LFV and changes since 1984 in response to changes in LULC and climatic conditions. Although this analysis utilized a comprehensive geospatial dataset for predicting SOC, integration of more detail information on agricultural management practices (e.g. landscape-scale data on tillage practices, fertilizer application, winter cover cropping, etc.) could strengthen the outcomes of the research and specify the management practices responsible for SOC change. Nonetheless, such data was not available for the LFV at a spatial resolution appropriate for my analysis. Future research, therefore, should gather and incorporate long-term datasets on management practices to capture greater variability in SOC and identify the relationships between SOC changes with specific management practices. Furthermore, my analysis only utilized satellite imagery of 30 m spatial resolution, however, some agricultural LULC features may not be wide enough (< 10 m) to be clearly distinguishable at this resolution (e.g. perennial grass margin, hedgerow). These LULC types could have notably different SOC content than the neighboring crop fields. Thus, future analysis should explore finer resolution satellite imagery to capture this variation which may result in a more accurate prediction of SOC. Finally, the static-empirical modeling approach performed for my study was an effective and practical alternative to the data-intensive process-based modeling, but it is not evident yet how the data-limited static-empirical technique compares with the proven process-based modeling approach. Moreover, it would also be important to examine how static-empirical modeling performs in a condition with larger climate variability and greater topographic variation as well as for investigating SOC change over time in larger study areas, e.g. national or continental scales.

References

AAFC, 2015. Detailed Soil Survey. Ottawa, Canada.

Adhikari, K., Owens, P.R., Libohova, Z., Miller, D.M., Wills, S.A., Nemecek, J., 2019. Assessing soil organic carbon stock of Wisconsin, USA and its fate under future land use and climate change. Sci. Total Environ. 667, 833–845.

ALUI, 2016. Agricultural Land Use Inventory - Metro Vancouver 2016. Abbosford, BC.

Amini, M., Afyuni, M., Fathianpour, N., Khademi, H., Flühler, H., 2005. Continuous soil pollution mapping using fuzzy logic and spatial interpolation. Geoderma 124, 223–233.

Arnett, J.T.T.R., Coops, N.C., Gergel, S.E., Falls, R.W., Baker, R.H., Arnett, J.T.T.R., Coops, N.C., Gergel, S.E., Falls, R.W., Arnett, J.T.T.R., Coops, N.C., Gergel, S.E., Falls, R.W., Baker, R.H., 2014. Detecting Stand-Replacing Disturbance using RapidEye Imagery : a Tasseled Cap Transformation and Modified Disturbance Index Detecting Stand-Replacing Disturbance using RapidEye Imagery : a Tasseled Cap Transformation and Modified Disturbance Index. Can. J. Remote Sens. 8992, 1–14. doi:10.1080/07038992.2014.899878

Baker, J.M., Ochsner, T.E., Venterea, R.T., Griffis, T.J., 2007. Tillage and soil carbon sequestration—What do we really know? Agric. Ecosyst. Environ. 118, 1–5.

BC Agriculture & Food Climate Action Initiative, 2015. BC Agriculture & Climate Change Regional Adaptation Strategies series - Fraser Valley. Victoria, BC.

BC Ministry of Agriculture, 2016. Fraser Valley Regional District agricultural land use inventory, Summer 2011 – 2013. Abbosford, BC.

BC TRIM, 2012. British Columbia Terrain Resource Information Mapping (TRIM) Digitial Map Products. Victoria, Canada.

BCMOEnvironment, 2014. British Columbia greenhouse gas inventory report 2012. British Columbia Ministry of Environment, Victoria, BC.

Behrens, T., Zhu, A.X., Schmidt, K., Scholten, T., 2010. Multi-scale digital terrain analysis and feature selection for digital soil mapping. Geoderma 155, 175–185. doi:10.1016/j.geoderma.2009.07.010

Bertrand, R., 1991. Soil Management Handbook for The Lower Fraser Valley, 2nd ed. B.C. Ministry of Agriculture, Fisheries and Food , Abbosford.

Bivand, R., Keitt, T., n.d. Rowlingson B. rgdal: Bindings for the Geospatial Data Abstraction Library. R package version 0.9–1. 2014.

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Queiroz Feitosa, R., van der Meer, F., van der Werff, H., van Coillie, F., Tiede, D., 2014. Geographic Object-Based Image Analysis - Towards a new paradigm. ISPRS J. Photogramm. Remote Sens. 87, 180–191. doi:10.1016/j.isprsjprs.2013.09.014

Bonfatti, B.R., Hartemink, A.E., Giasson, E., Tornquist, C.G., Adhikari, K., 2016. Digital mapping of soil carbon in a viticultural region of Southern Brazil. Geoderma 261, 204–221.

Boyle, A., Lavkulich, L., 1997. Carbon pool dynamics in the Lower Fraser Basin from 1827 to 1990. Environ. Manage. 21, 443–455. doi:10.1007/s002679900041

Boyle, A., Lavkulich, L., Schreier, H., Kiss, E., 1997. Changes in land cover and subsequent effects on Lower Fraser Basin ecosystems from 1827 to 1990. Environ. Manage. 21, 185–196. doi:10.1007/s002679900017

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Brieman, L., Cutler, A., Liaw, A., Wiener, M., 2015. randomForest package. Repos. CRAN.

Brungard, C.., Boettinger, J.., 2010. Conditioned latin hypercube sampling: optimal sample size for digital soil mapping of arid rangelands in Utah, USA, in: Digital Soil Mapping. Springer Netherlands, pp. 67–75. doi:10.1007/978-90-481-8863-5_6

Brus, D.J., Heuvelink, G.B.M., 2007. Optimization of sample patterns for universal kriging of environmental variables. Geoderma 138, 86–95.

Bruun, T.B., Elberling, B., de Neergaard, A., Magid, J., 2015. Organic carbon dynamics in different soil types after conversion of forest to agriculture. L. Degrad. Dev. 26, 272–283.

Bueno, J., Amiama, C., Hernanz, J.L., Pereira, J.M., 2006. Penetration resistance, soil water content, and workability of grasslands soils under two tillage systems. Trans. ASABE 49, 875–882.

Bui, E., Henderson, B., Viergever, K., 2009. Using knowledge discovery with data mining from the Australian Soil Resource Information System database to inform soil carbon mapping in Australia. Global Biogeochem. Cycles 23.

Bui, E., Moran, C., 2001. Disaggregation of polygons of surficial geology and soil maps using spatial modelling and legacy data. Geoderma 103, 79–94.

Cambardella, C.A., Moorman, T.B., Parkin, T.B., Karlen, D.L., Novak, J.M., Turco, R.F., Konopka, A.E., 1994. Field-scale variability of soil properties in central Iowa soils. Soil Sci. Soc. Am. J. 58, 1501–1511.

Campbell, D.J., 1991. Liquid and plastic limits, in: Smith, K., Mullins, C. (Eds.), Soil Analysis–Physical Methods. Dekker Inc., New York, pp. 367–398.

Carranza, E.J.M., Hale, M., 2002. Mineral imaging with Landsat Thematic Mapper data for hydrothermal alteration mapping in heavily vegetated terrane. Int. J. Remote Sens. 23, 4827–4852.

Cavazzi, S., Corstanje, R., Mayr, T., Hannam, J., Fealy, R., 2013. Are fine resolution digital elevation models always the best choice in digital soil mapping? Geoderma 195, 111–121.

Chagas, C., de Carvalho Junior, W., Bhering, S.B., Calderano Filho, B., 2016. Spatial prediction of soil surface texture in a semiarid region using random forest and multiple linear regressions. Catena 139, 232–240.

Chen, L., Gong, J., Fu, B., Huang, Z., Huang, Y., Gui, L., 2007. Effect of land use conversion on soil organic carbon sequestration in the loess hilly area, loess plateau of China. Ecol. Res. 22, 641–648. doi:10.1007/s11284-006-0065-1

Childs, C., 2004. Interpolating surfaces in ArcGIS spatial analyst. ArcUser, July-September 3235, 569.

Chipanshi, A., Fitzmaurice, J., De Jong, R., Bogdan, D., Lewis, M., Kroetsch, D., Lee, D., 2018. Assessment of soil trafficability across the agricultural region of the Canadian Prairies with the gridded climate data set. Soil Tillage Res. 184, 128–141.

Clausi, D.A., 2002. An analysis of co-occurrence texture statistics as a function of grey level quantization. Can. J. Remote Sens. 28, 45–62.

Cobo, J., Dercon, G., Yekeye, T., Chapungu, L., Kadzere, C., Murwira, A., Delve, R., Cadisch, G., 2010. Integration of mid-infrared spectroscopy and geostatistics in the assessment of soil spatial variability at landscape level. Geoderma 158, 398–411. doi:10.1016/j.geoderma.2010.06.013

Conant, R.T., Paustian, K., Elliott, E.T., 2001. Grassland management and conversion into grassland: effects on soil carbon. Ecol. Appl. 11, 343–355. doi:10.1890/1051-0761(2001)011[0343:GMACIG]2.0.CO;2

Crawford, E., MacNair, E., 2012. BC Agriculture Climate Change Adaptation Risk & Opportunity Assessment Series - Fraser Valley & Metro Vancouver. British Columbia

Agriculture & Food Climate Action Initiative, Victoria, Canada.

Crowther, T.W., Todd-Brown, K.E.O., Rowe, C.W., Wieder, W.R., Carey, J.C., Machmuller, M.B., Snoek, B.L., Fang, S., Zhou, G., Allison, S.D., 2016. Quantifying global soil carbon losses in response to warming. Nature 540, 104.

Cruz-Cárdenas, G., López-Mata, L., Ortiz-Solorio, C.A., Villaseñor, J.L., Ortiz, E., Silva, J.T., Estrada-Godoy, F., 2014. Interpolation of mexican soil properties at a scale of 1:1,000,000. Geoderma 213, 29–35. doi:10.1016/j.geoderma.2013.07.014

Del Galdo, I., Six, J., Peressotti, A., Francesca Cotrufo, M., 2003. Assessing the impact of land-use change on soil C sequestration in agricultural soils by means of organic matter fractionation and stable C isotopes. Glob. Chang. Biol. 9, 1204–1213.

Deng, L., Zhu, G. yu, Tang, Z. sheng, Shangguan, Z. ping, 2016. Global patterns of the effects of land-use changes on soil carbon stocks. Glob. Ecol. Conserv. 5, 127–138. doi:10.1016/j.gecco.2015.12.004

Dexter, A.R., Bird, N.R.A., 2001. Methods for predicting the optimum and the range of soil water contents for tillage based on the water retention curve. Soil Tillage Res. 57, 203–212.

Dimassi, B., Cohan, J.-P., Labreuche, J., Mary, B., 2013. Changes in soil carbon and nitrogen following tillage conversion in a long-term experiment in Northern France. Agric. Ecosyst. Environ. 169, 12–20.

Duffera, M., White, J.G., Weisz, R., 2007. Spatial variability of Southeastern US Coastal Plain soil physical properties: Implications for site-specific management. Geoderma 137, 327–339.

Earl, R., 1997. Prediction of trafficability and workability from soil moisture deficit. Soil Tillage Res. 40, 155–168. doi:10.1016/S0167-1987(96)01072-0

Edenhofer, O., 2014. Technical Summary In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Technical Report. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Eggen, M., Ozdogan, M., Zaitchik, B., Simane, B., 2016. Land cover classification in complex and fragmented agricultural landscapes of the Ethiopian highlands. Remote Sens. 8, 1020.

Elvidge, C.D., Lyon, R.J.P., 1985. Influence of rock-soil spectral variation on the assessment of green biomass. Remote Sens. Environ. 17, 265–279.

Environment and Climate Change Canada, 2017. Canadian Environmental Sustainability Indicators Greenhouse Gas Emissions Greenhouse Gas Emissions.

Environment Canada, 2019. Canadian Climate Normals 1981-2010 [WWW Document]. URL http://climate.weather.gc.ca/climate_normals/results_1981_2010_e.html?searchType=stnName&txtStationName=delta&searchMethod=contains&txtCentralLatMin=0&txtCentralLatSec=0&txtCentralLongMin=0&txtCentralLongSec=0&stnID=766&dispBack=0 (accessed 3.23.19).

ESRI ArcGIS, 2011. Release 10. Redlands, CA Environ. Syst. Res. Inst.

FAO, 2011. FAOSTAT Online Database.

Fischer, E.M., Knutti, R., 2016. Observed heavy precipitation increase confirms theory and early models. Nat. Clim. Chang. 6, 986.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232.

Ge, Y., Thomasson, J.A., Sui, R., 2011. Remote sensing of soil properties in precision agriculture : A review. Front. Earth Sci. 5, 229–238. doi:10.1007/s11707-011-0175-0

Grimm, R., Behrens, T., Märker, M., Elsenbeer, H., 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island—Digital soil mapping using Random Forests analysis. Geoderma 146, 102–113.

Grinand, C., Maire, G. Le, Vieilledent, G., Razakamanarivo, H., Razafimbelo, T., Bernoux, M., 2017. Estimating temporal changes in soil carbon stocks at ecoregional scale in Madagascar using remote-sensing. Int. J. Appl. Earth Obs. Geoinf. 54, 1–14. doi:10.1016/j.jag.2016.09.002

Grunwald, S., Thompson, J.A., Boettinger, J.L., 2011. Digital soil mapping and modeling at continental scales: Finding solutions for global issues. Soil Sci. Soc. Am. J. 75, 1201–1213.

Guo, L.B., Gifford, R.M., 2002. Soil carbon stocks and land use change: a meta analysis. Glob. Chang. Biol. 8, 345–360.

Gupta, S.C., Larson, W.E., 1979. Estimating soil water retention characteristics from particle size distribution, organic matter percent, and bulk density. Water Resour. Res. 15, 1633–1635. doi:10.1029/WR015i006p01633

Haddaway, N.R., Hedlund, K., Jackson, L.E., Kätterer, T., Lugato, E., Thomsen, I.K., Jørgensen, H.B., Isberg, P.-E., 2017. How does tillage intensity affect soil organic carbon? A systematic review. Environ. Evid. 6, 30.

Hanlon, E.A., McNeal, B.L., Kidder, G., 1993. Soil and container media electrical conductivity interpretations. Florida Cooperative Extension Service, Institute of Food and Agricultural ….

Healey, S.P., Cohen, W.B., Zhiqiang, Y., Krankina, O.N., 2005. Comparison of Tasseled Cap-based Landsat data structures for use in forest disturbance detection. Remote Sens. Environ. 97, 301–310. doi:10.1016/j.rse.2005.05.009

Hebda, R.J., Gustavson, K., Golinski, K., Calder, A.M., 2000. Burns Bog Ecosystem Review: Synthesis Report for Burns Bog, Fraser River Delta, South-western British Columbia, Canada. Environmental Assessment Office Victoria, BC.

Hengl, T., 2007. A Practical guide to Geostatistical Mapping, Scientific and Technical Research series. doi:10.1016/0277-9390(86)90082-8

Hengl, T., Heuvelink, G.B.M., Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma 120, 75–93. doi:10.1016/j.geoderma.2003.08.018

Heung, B., Bulmer, C.E., Schmidt, M.G., 2014. Geoderma Predictive soil parent material mapping at a regional-scale : A Random Forest approach. Geoderma 214, 141–154. doi:10.1016/j.geoderma.2013.09.016

Heung, B., Ho, H.C., Zhang, J., Knudby, A., Bulmer, C.E., Schmidt, M.G., 2016. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. Geoderma 265, 62–77.

Hijmans, R.J., Van Etten, J., 2016. raster: Geographic data analysis and modeling. R package version 2.5-8. Vienna, Austria R Found. Retrieved from https//CRAN. R-project. org/package= raster.

Huang, J., Hartemink, A.E., Zhang, Y., 2019. Climate and Land-Use Change Effects on Soil Carbon Stocks over 150 Years in Wisconsin, USA. Remote Sens. 11, 1504.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309.

Hunt Jr, E.R., Rock, B.N., 1989. Detection of changes in leaf water content using near-and middle-infrared reflectances. Remote Sens. Environ. 30, 43–54.

Hutchinson, M.F., 1993. Development of a continent-wide DEM with applications to terrain and climate analysis, in: Goodchild, M.F. (Ed.), Environmental Modeling with GIS. Oxford University Press, New York, pp. 392–399.

Ise, T., Moorcroft, P.R., 2006. The global-scale temperature and moisture dependencies of soil organic carbon decomposition: an analysis using a mechanistic decomposition model. Biogeochemistry 80, 217–231.

Janik, L.J., Merry, R.H., Skjemstad, J.O., 1998. Can mid infrared diffuse reflectance analysis replace soil extractions? Aust. J. Exp. Agric. 38, 681–696.

Janik, L.J., Skjemstad, J.O., 1995. Characterization and analysis of soils using mid-infrared partial least-squares. 2. Correlations with some laboratory data. Soil Res. 33, 637–650.

Johannes, A., Matter, A., Schulin, R., Weisskopf, P., Baveye, P.C., Boivin, P., 2017. Optimal organic carbon values for soil structure quality of arable soils. Does clay content matter? Geoderma 302, 14–21.

Kätterer, T., Bolinder, M.A., Andrén, O., Kirchmann, H., Menichetti, L., 2011. Roots contribute more to refractory soil organic matter than above-ground crop residues, as revealed by a long-term field experiment. Agric. Ecosyst. Environ. 141, 184–192.

Keller, T., Arvidsson, J., Dexter, A.R., 2007. Soil structures produced by tillage as affected by soil water content and the physical quality of soil. Soil Tillage Res. 92, 45–52.

Keller, T., Dexter, A.R., 2012. Plastic limits of agricultural soils as functions of soil texture and organic matter content. Soil Res. 50, 7–17.

Kerry, R., Oliver, M.A., 2007. Comparing sampling needs for variograms of soil properties computed by the method of moments and residual maximum likelihood. Geoderma 140, 383–396.

Kerry, R., Oliver, M.A., Frogbrook, Z.L., 2010. Sampling in precision agriculture, in: Geostatistical Applications for Precision Agriculture. Springer, Dordrecht, pp. 35–63.

Keskin, H., Grunwald, S., 2018. Regression kriging as a workhorse in the digital soil mapper's toolbox. Geoderma 326, 22–41.

Kheir, R.B., Greve, M.H., Bøcher, P.K., Greve, M.B., Larsen, R., McCloy, K., 2010. Predictive mapping of soil organic carbon in wet cultivated lands using classification-tree based models: The case study of Denmark. J. Environ. Manage. 91, 1150–1160.

Knoth, C., Klein, B., Prinz, T., Kleinebecker, T., 2013. Unmanned aerial vehicles as innovative remote sensing platforms for high-resolution infrared imagery to support restoration monitoring in cut-over bogs. Appl. Veg. Sci. 16, 509–517.

Kolberg, D., Persson, T., Mangerud, K., Riley, H., 2019. Impact of projected climate change on workability, attainable yield, profitability and farm mechanization in Norwegian spring cereals. Soil Tillage Res. 185, 122–138.

Kononova, M.M., 2013. Soil organic matter: Its nature, its role in soil formation and in soil fertility. Elsevier.

Koven, C.D., Chambers, J.Q., Georgiou, K., Knox, R., Negron-Juarez, R., Riley, W.J., Arora, V.K., Brovkin, V., Friedlingstein, P., Jones, C.D., 2015. Controls on terrestrial carbon feedbacks by productivity vs. turnover in the CMIP5 Earth System Models. Biogeosciences Discuss. 12, 5757–5801.

Kretschmer, H., 1996. Koernung und Konsistenz, in: Blume, H.-P., Felix-Henningsen, P., Fischer, W.R., Frede, H.G., Horn, R., Stahr, K. (Eds.), Handbuch Der Bodenkunde. Ecomed.

Kværnø, S.H., Haugen, L.E., Børresen, T., 2007. Variability in topsoil texture and carbon content within soil map units and its implications in predicting soil water content for optimum workability. Soil Tillage Res. 95, 332–347. doi:10.1016/j.still.2007.02.001

Lacoste, M., Minasny, B., McBratney, A., Michot, D., Viaud, V., Walter, C., 2014. High resolution 3D mapping of soil organic carbon in a heterogeneous agricultural landscape. Geoderma 213, 296–311. doi:10.1016/j.geoderma.2013.07.002

Lagacherie, P., 2008. Digital soil mapping: A state of the art, in: Digital Soil Mapping with Limited Data. Springer Netherlands, Dordrecht, pp. 3–14. doi:10.1007/978-1-4020-8592-5_1

Lal, R., 2004. Soil carbon sequestration impacts on global climate change and food security. Science (80-.). 304, 1623–7. doi:10.1126/science.1097396

Lal, R., Negassa, W., Lorenz, K., 2015. Carbon sequestration in soil. Curr. Opin. Environ. Sustain. 15, 79–86.

Lalonde, V., Hughes-Games, G.A., 1997. BC agricultural drainage manual. BC Ministry of Agriculture, Fisheries and Food, Victoria, Canada.

Levin, N., Ben-Dor, E., Singer, A., 2005. A digital camera as a tool to measure colour indices and related properties of sandy soils in semi-arid environments. Int. J. Remote Sens. 26, 5475–5492.

Li, H., Webster, R., Shi, Z., 2015. Mapping soil salinity in the Yangtze delta: REML and universal kriging (E-BLUP) revisited. Geoderma 237–238, 71–77. doi:10.1016/j.geoderma.2014.08.008

Li, J., Heap, A.D., 2011. A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. Ecol. Inform. 6, 228–241. doi:10.1016/j.ecoinf.2010.12.003

Li, Y., 2010. Can the spatial prediction of soil organic matter contents at various sampling scales be improved by using regression kriging with auxiliary information? Geoderma 159, 63–75.

Ließ, M., 2015. Sampling for regression-based digital soil mapping: Closing the gap between statistical desires and operational applicability. Spat. Stat. 13, 106–122. doi:10.1016/j.spasta.2015.06.002

Ließ, M., Schmidt, J., Glaser, B., 2016. Improving the spatial prediction of soil organic carbon stocks in a complex tropical mountain landscape by methodological specifications in machine learning approaches. PLoS One 11, e0153673.

Lilly, A., Baggaley, N.J., Loades, K.W., McKenzie, B.M., Troldborg, M., 2018. Soil erosion and compaction in Scottish soils: adapting to a changing climate. Dundee.

Liu, D., Wang, Z., Zhang, B., Song, K., Li, X., Li, J., Li, F., Duan, H., 2006. Spatial distribution of soil organic carbon and analysis of related factors in croplands of the black soil region, Northeast China. Agric. Ecosyst. Environ. 113, 73–81.

Liu, S., An, N., Yang, J., Dong, S., Wang, C., Yin, Y., 2015. Prediction of soil organic matter variability associated with different land use types in mountainous landscape in southwestern Yunnan province, China. Catena 133, 137–144. doi:10.1016/j.catena.2015.05.010

Lussier, J.M., Krzic, M., Smukler, S.M., Bomke, A.A., Bondar, D., 2019. Short-term effects of grassland set-asides on soil properties in the Fraser River delta of British Columbia. Can. J. Soil Sci. 99, 136–145.

Luttmerding, H., 1981. Soils of the Langley-Vancouver map area. Kelowna.

MacHmuller, M.B., Kramer, M.G., Cyle, T.K., Hill, N., Hancock, D., Thompson, A., 2015. Emerging land use practices rapidly increase soil organic matter. Nat. Commun. 6. doi:10.1038/ncomms7995

Malone, B., McBratney, A., Minasny, B., 2013a. Spatial Scaling for Digital Soil Mapping. Soil Sci. Soc. Am. J. 77, 890–902. doi:10.2136/sssaj2012.0419

Malone, B., Minasny, B., McBratney, A., 2013b. Using R for Digital Soil Mapping. Springer. doi:10.1007/978-3-319-44327-0

Malone, B. P., McBratney, A.B., Minasny, B., Laslett, G.M., 2009. Mapping continuous depth functions of soil carbon storage and available water capacity. Geoderma 154, 138–152. doi:10.1016/j.geoderma.2009.10.007

Malone, Brendan P, McBratney, A.B., Minasny, B., Laslett, G.M., 2009. Mapping continuous depth functions of soil carbon storage and available water capacity. Geoderma 154, 138–152.

Malone, B.P., Styc, Q., Minasny, B., McBratney, A.B., 2017. Digital soil mapping of soil carbon at the farm scale: A spatial downscaling approach in consideration of measured and uncertain data. Geoderma 290, 91–99. doi:10.1016/J.GEODERMA.2016.12.008

Mapfumo, E., Chanasyk, D.S., 1998. Guidelines for safe trafficking and cultivation, and resistance–density–moisture relations of three disturbed soils from Alberta. Soil Tillage Res. 46, 193–202.

Masserschmidt, I., Cuelbas, C.J., Poppi, R.J., De Andrade, J.C., De Abreu, C.A., Davanzo, C.U., 1999. Determination of organic matter in soils by FTIR/diffuse reflectance and multivariate calibration. J. Chemom. 13, 265–273.

McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping, Geoderma. doi:10.1016/S0016-7061(03)00223-4

Merante, P., Dibari, C., Ferrise, R., Sánchez, B., Iglesias, A., Lesschen, J.P., Kuikman, P., Yeluripati, J., Smith, P., Bindi, M., 2017. Adopting soil organic carbon management practices in soils of varying quality: Implications and perspectives in Europe. Soil Tillage Res. 165, 95–106. doi:10.1016/J.STILL.2016.08.001

Minasny, B., Malone, B.P., Mcbratney, A.B., Angers, D.A., Arrouays, D., Chambers, A., Chaplot, V., Chen, Z., Cheng, K., Das, B.S., Field, D.J., Gimona, A., Hedley, C.B., Young, S., Mandal, B., Marchant, B.P., Martin, M., Mcconkey, B.G., Leatitia, V., Rourke, S.O., Richer-de-forges, A.C., Odeh, I., Padarian, J., Paustian, K., Pan, G., Poggio, L., Savin, I., Stolbovoy, V., Stockmann, U., Sulaeman, Y., Tsui, C., Vågen, T., Wesemael, B. Van, Winowiecki, L., Arrouays, D., Martin, M., Richer-de-forges, A.C., Mulder, V.L., 2017. Soil carbon 4 per mille. Geoderma 292, 59–86. doi:10.1016/j.geoderma.2017.01.002

Minasny, B., McBratney, A.B., 2016. Digital soil mapping: A brief history and some lessons. Geoderma 264, 301–311.

Minasny, B., McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Comput. Geosci. 32, 1378–1388. doi:10.1016/j.cageo.2005.12.009

Minasny, B., Mcbratney, A.B., Malone, B.P., Wheeler, I., 2013. Digital Mapping of Soil Carbon. Adv. Agron. 118. doi:10.1016/B978-0-12-405942-9.00001-3

Mogili, U.M.R., Deepak, B., 2018. Review on application of drone systems in precision agriculture. Procedia Comput. Sci. 133, 502–509.

Mueller, L., Schindler, U., Fausey, N.R., Lal, R., 2003. Comparison of methods for estimating maximum soil water content for optimum workability. Soil Tillage Res. 72, 9–20. doi:10.1016/S0167-1987(03)00046-1

Mueller, L., Tille, P., Kretschmer, H., 1990. Trafficability and workability of alluvial clay soils in response to drainage status. Soil Tillage Res. 16, 273–287.

Mulder, V.L., De Bruin, S., Schaepman, M.E., Mayr, T.R., 2011. The use of remote sensing in soil and terrain mapping—A review. Geoderma 162, 1–19.

Müller, L., Lipiec, J., Kornecki, T.S., Gebhardt, S., 2011. Trafficability and workability of soils, in: Encyclopedia of Agrophysics. Springer, pp. 912–924.

Nemeth, D., Lambrinos, J.G., Strik, B.C., 2017. The effects of long-term management on patterns of carbon storage in a northern highbush blueberry production system. Sci. Total Environ. 579, 1084–1093. doi:10.1016/J.SCITOTENV.2016.11.077

Neufeld, K., Paul, S., Smukler, S., 2017. Improving On-Farm Drainage Management to Reduce the Impacts of Climate Change in Delta, BC. Vancouver, Canada.

Niang, M.A., Nolin, M.C., Jégo, G., Perron, I., 2014. Digital Mapping of soil texture using RADARSAT-2 polarimetric synthetic aperture radar data. Soil Sci. Soc. Am. J. 78, 673–684.

Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann jj, M., Barth es, B., Ben Dor, E., Brown xx, D., Clairotte, M., Csorba, A., Dardenne jjjj, P., Demattê, J., Genot, V., Guerrero, C., Knadel xxx, M., Montanarella, L., Noon, C., Ramirez-Lopez, L., Robertson, J., Sakai jjjjjj, H., Soriano-Disla, J., Shepherd, K., Stenberg xxxx, B., Towett, E., Vargas, R., Wetterlind xxxx, J., Paulo, S., 2015. Soil Spectroscopy: An Alternative to Wet Chemistry for Soil Monitoring. Adv. Agron. 132, 139–159. doi:10.1016/bs.agron.2015.02.002

O'Rourke, S.M., Holden, N.M., 2011. Optical sensing and chemometric analysis of soil organic carbon–a cost effective alternative to conventional laboratory methods? Soil Use Manag. 27, 143–155.

Obour, P.B., Lamandé, M., Edwards, G., Sørensen, C.G., Munkholm, L.J., 2017. Predicting soil workability and fragmentation in tillage: a review. Soil Use Manag. 33, 288–298.

Olson, G.O., 1975. Engineering characteristics of Ohio soil series.

Ouyang, Y., Higman, J., Campbell, D., Davis, J., 2003. Three-dimensional kriging analysis of sediment mercury distribution: A case study. J. Am. Water Resour. Assoc. 39, 689–702. doi:10.1111/j.1752-1688.2003.tb03685.x

Paal, T., Starast, M., Noormets-Šanski, M., Vool, E., Tasa, T., Karp, K., 2011. Influence of liming and fertilization on lowbush blueberry in harvested peat field condition. Sci. Hortic. (Amsterdam). 130, 157–163.

Paneque-Gálvez, J., Mas, J.-F., Moré, G., Cristóbal, J., Orta-Martínez, M., Luz, A.C., Guèze, M., Macía, M.J., Reyes-García, V., 2013. Enhanced land use/cover classification of heterogeneous tropical landscapes using support vector machines and textural homogeneity. Int. J. Appl. Earth Obs. Geoinf. 23, 372–383.

Pasley, H.R., Camberato, J.J., Cairns, J.E., Zaman-Allah, M., Das, B., Vyn, T.J., 2020. Nitrogen rate impacts on tropical maize nitrogen use efficiency and soil nitrogen depletion in eastern and southern Africa. Nutr. Cycl. Agroecosystems 1–12.

Paul, S.S., Coops, N., Johnson, M., Krzic, M., Smukler, S., 2019. Evaluating sampling efforts of standard laboratory analysis and mid-infrared spectroscopy for cost effective digital soil mapping at field scale. Geoderma 356, 113925.

Paul, S.S., Li, J., Wheate, R., Li, Y., 2018. Application of Object Oriented Image Classification and Markov Chain Modeling for Land Use and Land Cover Change Analysis. J. Environ. Informatics 31, 30–40. doi:10.3808/jei.201700368

Paustian, K., Lehmann, J., Ogle, S., Reay, D., Robertson, G.P., Smith, P., 2016. Climate-smart soils. Nature 532, 49–57. doi:10.1038/nature17174

Pebesma, E., Heuvelink, G., 2016. Spatio-temporal interpolation using gstat. RFID J. 8, 204–218.

Pickett, S.T.A., 1989. Space-for-time substitution as an alternative to long-term studies, in: Long-Term Studies in Ecology. Springer, pp. 110–135.

Poeplau, C., Don, A., Vesterdal, L., Leifeld, J., Van Wesemael, B., Schumacher, J., Gensior, A., 2011. Temporal dynamics of soil organic carbon after land-use change in the temperate zone - carbon response functions as a model approach. Glob. Chang. Biol. 17, 2415–2427. doi:10.1111/j.1365-2486.2011.02408.x

Pribyl, D.W., 2010. A critical review of the conventional SOC to SOM conversion factor. Geoderma 156, 75–83. doi:10.1016/J.GEODERMA.2010.02.003

Priyanka, K., 2018. Quantifying total and labile pools of soil organic carbon in cultivated and uncultivated soils in eastern India. Soil Res. 56, 413–420.

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

Reyes Rojas, L.A., Adhikari, K., Ventura, S.J., 2018. Projecting soil organic carbon distribution in central Chile under future climate scenarios. J. Environ. Qual. 47, 735–745.

Richards, L.A., Fireman, M., 1943. Pressure-plate apparatus for measuring moisture sorption and transmission by soils. Soil Sci. 56, 395–404.

Ridgeway, G., 2015. The gbm package. Generalized boosted regression models (Documentation on the R Package ‘gbm’, version 1.6–3.).

Rietz, D.N., Haynes, R.J., 2003. Effects of irrigation-induced salinity and sodicity on soil microbial activity. Soil Biol. Biochem. 35, 845–854.

Rikimaru, A., Roy, P.S., Miyatake, S., 2002. Tropical forest cover density mapping. Trop. Ecol. 43, 39–47.

Robinson, T.P., Metternicht, G., 2006. Testing the performance of spatial interpolation techniques for mapping soil properties. Comput. Electron. Agric. 50, 97–108. doi:10.1016/j.compag.2005.07.003

Roudier, P., 2014. clhs-package Conditioned Latin Hypercube Sampling.

Rounsevell, M.D.A., 1993. A review of soil workability models and their limitations in temperate regions. Soil use Manag. 9, 15–20.

Rouse Jr, J., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS, in: NASA. Goddard Space Flight Center 3d ERTS-1 Symposium. NASA, United States, College Station, TX, pp. 309–317.

Rumpel, C., Amiraslani, F., Koutika, L.-S., Smith, P., Whitehead, D., Wollenberg, E., 2018. Put more carbon in soils to meet Paris climate pledges. Nature 564, 32–34.

Rutledge, P.L., Russell, D.G., 1971. Work day probabilities for tillage operations in Alberta. Edmonton.

Sanderman, J., Hengl, T., Fiske, G.J., 2017. Soil carbon debt of 12,000 years of human land use. Proc. Natl. Acad. Sci. 114, 9575–9580.

Santra, P., Kumar, M., Panwar, N., 2017. Digital soil mapping of sand content in arid western India through geostatistical approaches. Geoderma Reg. 9, 56–72.

Schaufler, G., Kitzler, B., Schindlbacher, A., Skiba, U., Sutton, M.A., Zechmeister-Boltenstern, S., 2010. Greenhouse gas emissions from European soils under different land use: effects of soil moisture and temperature. Eur. J. Soil Sci. 61, 683–696.

Schillaci, C., Acutis, M., Lombardo, L., Lipani, A., Fantappiè, M., Märker, M., Saia, S., 2017. Spatio-temporal topsoil organic carbon mapping of a semi-arid Mediterranean region: The role of land use, soil texture, topographic indices and the influence of remote sensing data to modelling. Sci. Total Environ. 601–602, 821–832. doi:10.1016/j.scitotenv.2017.05.239

Schmidt, M., Pringle, M., Devadas, R., Denham, R., Tindall, D., 2016. A framework for large-area mapping of past and present cropping activity using seasonal Landsat images and time series metrics. Remote Sens. 8, 312.

Servadio, P., Bergonzoli, S., Beni, C., 2016. Soil tillage systems and wheat yield under climate change scenarios. Agronomy 6, 43.

Shen, W., Li, M., Huang, C., Wei, A., 2016. Quantifying Live Aboveground Biomass and Forest Disturbance of Mountainous Natural and Plantation Forests in Northern Guangdong, China, Based on Multi-Temporal Landsat, PALSAR and Field Plot Data Wenjuan. Remote Sens. 8, 1–24. doi:10.3390/rs8070595

Shupe, S., 2013. Land cover trends in Metro Vancouver, Canada over 45 years: mapping, analysis, and visualization, in: Buchroithner, M. (Ed.), Proceedings of the 26th International Cartographic Conference. International Cartographic Association, Dresden, Germany, pp. 901–902.

Sierra, C.A., Trumbore, S.E., Davidson, E.A., Vicca, S., Janssens, I., 2015. Sensitivity of decomposition rates of soil organic matter with respect to simultaneous changes in temperature and moisture. J. Adv. Model. Earth Syst. 7, 335–356.

Simbahan, G.C., Dobermann, A., 2006. Sampling optimization based on secondary information and its utilization in soil carbon mapping. Geoderma 133, 345–362.

Sindayihebura, A., Ottoy, S., Dondeyne, S., Van Meirvenne, M., Van Orshoven, J., 2017. Comparing digital soil mapping techniques for organic carbon and clay content: Case study in Burundi's central plateaus. Catena 156, 161–175.

Singh, M., Sarkar, B., Sarkar, S., Churchman, J., Bolan, N., Mandal, S., Menon, M., Purakayastha, T.J., Beerling, D.J., 2018. Stabilization of soil organic carbon as influenced by clay mineralogy, in: Advances in Agronomy. Elsevier, pp. 33–84.

Skalar Analytical, 2019. PrimacsSNC-100 TN/TC analyzer user manual. Skalar Analytical B.V., Breda, The Netherlands.

Smedema, L.K., 1993. Drainage performance and soil management. Soil Technol. 6, 183–189.

Statistics Canada, 2016. Census of Agriculture. Ottawa, Canada.

Stockmann, U., Adams, M.A., Crawford, J.W., Field, D.J., Henakaarchchi, N., Jenkins, M., Minasny, B., McBratney, A.B., Courcelles, V. de R. de, Singh, K., Wheeler, I., Abbott, L., Angers, D.A., Baldock, J., Bird, M., Brookes, P.C., Chenu, C., Jastrow, J.D., Lal, R., Lehmann, J., O'Donnell, A.G., Parton, W.J., Whitehead, D., Zimmermann, M., 2013. The knowns, known unknowns and unknowns of sequestration of soil organic carbon. Agric. Ecosyst. Environ. 164, 80–99.

Stockmann, U., Padarian, J., McBratney, A., Minasny, B., de Brogniez, D., Montanarella, L., Hong, S.Y., Rawlins, B.G., Field, D.J., 2015. Global soil organic carbon assessment. Glob. Food Sec. 6, 9–16.

Suk Lee, W., Ehsani, R., 2015. Sensing systems for precision agriculture in Florida. Comput. Electron. Agric. 112, 2–9. doi:10.1016/j.compag.2014.11.005

Tammeorg, P., Simojoki, A., Mäkelä, P., Stoddard, F.L., Alakukku, L., Helenius, J., 2014. Short-term effects of biochar on soil properties and wheat yield formation with meat bone meal and inorganic fertiliser on a boreal loamy sand. Agric. Ecosyst. Environ. 191, 108–116.

Thermo Fisher Scientific, 2010. Flash 2000 elemental analyzer operating manual, D. ed. Thermo Fisher Scientific Inc, Cambridge, UK.

Todd-Brown, K.E.O., Randerson, J.T., Hopkins, F., Arora, V., Hajima, T., Jones, C., Shevliakova, E., Tjiputra, J., Volodin, E., Wu, T., 2014. Changes in soil organic carbon storage predicted by Earth system models during the 21st century. Biogeosciences 11, 2341–2356.

Tomasek, B.J., Williams II, M.M., Davis, A.S., 2017. Changes in field workability and drought risk from projected climate change drive spatially variable risks in Illinois cropping systems. PLoS One 12, e0172301.

Tsouros, D.C., Bibi, S., Sarigiannidis, P.G., 2019. A Review on UAV-Based Applications for Precision Agriculture. Information 10, 349.

Tubiello, F.N., Salvatore, M., Ferrara, A.F., House, J., Federici, S., Rossi, S., Biancalani, R., Condor Golec, R.D., Jacobs, H., Flammini, A., Prosperi, P., Cardenas-Galindo, P., Schmidhuber, J., Sanz Sanchez, M.J., Srivastava, N., Smith, P., 2015. The Contribution of Agriculture, Forestry and other Land Use activities to Global Warming, 1990-2012. Glob. Chang. Biol. 21, 2655–2660. doi:10.1111/gcb.12865

Vaccari, F.P., Lugato, E., Gioli, B., D'Acqui, L., Genesio, L., Toscano, P., Matese, A., Miglietta, F., 2012. Land use change and soil organic carbon dynamics in Mediterranean agro-ecosystems: The case study of Pianosa Island. Geoderma 175–176, 29–36. doi:10.1016/j.geoderma.2012.01.021

Vågen, T.-G., Winowiecki, L.A., Tondoh, J.E., Desta, L.T., Gumbricht, T., 2016. Mapping of soil properties and land degradation risk in Africa using MODIS reflectance. Geoderma

263, 216–225. doi:10.1016/j.geoderma.2015.06.023

Van Bemmelen, J.M., 1890. Über die Bestimmung des Wassers, des Humus, des Schwefels, der in den colloïdalen Silikaten gebundenen Kieselsäure, des Mangans usw im Ackerboden. Die Landwirthschaftlichen Versuchs-Stationen 37, e290.

van den Boogaart, K.G., Tolosana-Delgado, R., 2008. "Compositions": a unified R package to analyze compositional data. Comput. Geosci. 34, 320–338.

Van Deventer, A.P., Ward, A.D., Gowda, P.H., Lyon, J.G., 1997. Using Thematic Mapper data to identify contrasting soil plains and tillage practices. Photogramm. Eng. Remote Sensing 63, 87–93.

VandenBygaart, A.J., Bremer, E., McConkey, B.G., Janzen, H.H., Angers, D.A., Carter, M.R., Drury, C.F., Lafond, G.P., McKenzie, R.H., 2010. Soil organic carbon stocks on long-term agroecosystem experiments in Canada. Can. J. soil Sci. 90, 543–550.

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 131, 59–75. doi:10.1016/j.geoderma.2005.03.007

Wackernagel, H., 2003. Multivariate geostatistics: an introduction with applications, 3rd ed. Springer-Verlag, Berlin Heidelberg.

Wang, B., Waters, C., Orgill, S., Cowie, A., Clark, A., Li Liu, D., Simpson, M., McGowen, I., Sides, T., 2018. Estimating soil organic carbon stocks using different modelling techniques in the semi-arid rangelands of eastern Australia. Ecol. Indic. 88, 425–438. doi:10.1016/J.ECOLIND.2018.01.049

Wang, S., Wang, Q., Adhikari, K., Jia, S., Jin, X., 2016. Spatial-Temporal Changes of Soil

Organic Carbon Content in Wafangdian , China 11–14. doi:10.3390/su8111154

Wang, T., Hamann, A., Spittlehouse, D., Carroll, C., 2016. Locally downscaled and spatially customizable climate data for historical and future periods for North America. PLoS One 11, e0156720.

Were, K., Bui, D.T., Dick, Ø.B., Singh, B.R., 2015. A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. Ecol. Indic. 52, 394–403.

Wickham, H., 2016. ggplot2: elegant graphics for data analysis. Springer.

Wilson, S., 2010. Natural capital in BC's Lower Mainland: valuing the benefits from nature. Vancouver, Canada.

Winowiecki, L., Vågen, T.-G., Huising, J., 2016. Effects of land cover on ecosystem services in Tanzania: A spatial assessment of soil organic carbon. Geoderma 263, 274–283. doi:10.1016/j.geoderma.2015.03.010

Wu, H., Guo, Z., Gao, Q., Peng, C., 2009. Distribution of soil inorganic carbon storage and its changes due to agricultural land use activity in China. Agric. Ecosyst. Environ. 129, 413–421. doi:10.1016/j.agee.2008.10.020

Wulder, M.A., Coops, N.C., Roy, D.P., White, J.C., Hermosilla, T., 2018. Land cover 2.0. Int. J. Remote Sens. 39, 4254–4284.

Xiong, X., Grunwald, S., Myers, D.B., Ross, C.W., Harris, W.G., Comerford, N.B., 2014. Interaction effects of climate and land use/land cover change on soil organic carbon sequestration. Sci. Total Environ. 493, 974–982. doi:10.1016/j.scitotenv.2014.06.088

Xu, Y., Yu, L., Peng, D., Cai, X., Cheng, Y., Zhao, J., Zhao, Y., Feng, D., Hackman, K., Huang, X., 2018. Exploring the temporal density of Landsat observations for cropland mapping:

experiments from Egypt, Ethiopia, and South Africa. Int. J. Remote Sens. 39, 7328–7349.

Yang, J., 2010. Economic analysis of blueberry investment in British Columbia. University of British Columbia.

Yang, R., Zhang, G., Liu, F., Lu, Y., Yang, Fan, Yang, Fei, Yang, M., Zhao, Y., Li, D., 2016. Comparison of boosted regression tree and random forest models for mapping topsoil organic carbon concentration in an ... Ecol. Indic. 60, 870–878. doi:10.1016/j.ecolind.2015.08.036

Yigini, Y., Panagos, P., 2016. Assessment of soil organic carbon stocks under future climate and land cover changes in Europe. Sci. Total Environ. 557–558, 838–850. doi:10.1016/j.scitotenv.2016.03.085

Zhang, K., Dang, H., Zhang, Q., Cheng, X., 2015. Soil carbon dynamics following land-use change varied with temperature and precipitation gradients: Evidence from stable isotopes. Glob. Chang. Biol. 21, 2762–2772. doi:10.1111/gcb.12886

Zhu, Q., Lin, H.S., 2010. Comparing ordinary kriging and regression kriging for soil properties in contrasting landscapes. Pedosphere 20, 594–606. doi:10.1016/S1002-0160(10)60049-5

Zomer, R.J., Bossio, D.A., Sommer, R., Verchot, L. V., 2017. Global Sequestration Potential of Increased Organic Carbon in Cropland Soils. Sci. Rep. 7, 1–8. doi:10.1038/s41598-017-15794-8

Zvoleff, A., 2016. Package ‘glcm.’

Appendices

Appendix A

Supplementary table for Chapter 2

Sampling efforts	Soil proper-ty	Met-rics	SLA	SLA SD	MIRS	MIRS SD	Metrics	SLA	SLA SD	MIRS	MIRS SD	Metrics (%)	SLA	SLA SD	MIRS	MIRS. SD
100%	Sand	R^2	0.430	0.033	0.880	0.020	CCC	0.570	0.029	0.847	0.044	nRMSE	22.048	2.119	12.771	0.753
90%	Sand	R^2	0.410	0.024	0.870	0.030	CCC	0.570	0.029	0.836	0.029	nRMSE	22.849	1.130	12.909	1.221
80%	Sand	R^2	0.390	0.018	0.880	0.030	CCC	0.548	0.038	0.816	0.049	nRMSE	23.334	1.849	13.316	1.470
70%	Sand	R^2	0.360	0.028	0.850	0.042	CCC	0.551	0.025	0.823	0.052	nRMSE	23.261	2.735	13.870	3.053
60%	Sand	R^2	0.350	0.030	0.820	0.036	CCC	0.540	0.037	0.815	0.038	nRMSE	26.040	3.094	14.173	2.283
50%	Sand	R^2	0.270	0.040	0.710	0.036	CCC	0.460	0.030	0.770	0.042	nRMSE	28.271	2.873	14.937	1.912
40%	Sand	R^2	0.180	0.030	0.600	0.028	CCC	0.280	0.046	0.686	0.047	nRMSE	51.125	5.923	17.615	2.568
30%	Sand	R^2	0.070	0.029	0.390	0.016	CCC	0.060	0.038	0.538	0.044	nRMSE	73.630	4.892	22.077	2.739
20%	Sand	R^2	0.030	0.030	0.350	0.025	CCC	0.039	0.035	0.422	0.049	nRMSE	75.411	6.138	27.530	4.836
10%	Sand	R^2	0.004	0.003	0.150	0.038	CCC	0.002	0.002	0.286	0.036	nRMSE	80.693	5.302	44.801	3.882
100%	Silt	R^2	0.520	0.037	0.870	0.020	CCC	0.673	0.034	0.880	0.035	nRMSE	24.903	1.857	11.392	1.220
90%	Silt	R^2	0.490	0.026	0.830	0.034	CCC	0.665	0.030	0.860	0.027	nRMSE	25.432	2.119	13.037	1.813
80%	Silt	R^2	0.500	0.031	0.830	0.029	CCC	0.658	0.028	0.870	0.062	nRMSE	25.318	1.733	12.974	2.514
70%	Silt	R^2	0.460	0.028	0.840	0.027	CCC	0.651	0.043	0.830	0.032	nRMSE	26.631	3.025	13.752	1.376
60%	Silt	R^2	0.430	0.027	0.790	0.025	CCC	0.641	0.041	0.810	0.053	nRMSE	26.089	3.017	14.253	1.791
50%	Silt	R^2	0.410	0.018	0.710	0.032	CCC	0.612	0.057	0.770	0.046	nRMSE	28.305	4.931	15.179	2.823
40%	Silt	R^2	0.270	0.017	0.570	0.025	CCC	0.388	0.054	0.680	0.052	nRMSE	53.701	5.923	20.623	3.247
30%	Silt	R^2	0.130	0.016	0.500	0.026	CCC	0.160	0.035	0.650	0.033	nRMSE	62.417	4.451	22.827	5.346
20%	Silt	R^2	0.100	0.008	0.310	0.023	CCC	0.090	0.034	0.420	0.025	nRMSE	68.532	8.378	28.661	4.581
10%	Silt	R^2	0.070	0.002	0.200	0.017	CCC	0.028	0.011	0.390	0.021	nRMSE	93.572	5.404	43.172	5.591
100%	Clay	R^2	0.430	0.033	0.780	0.017	CCC	0.617	0.027	0.870	0.024	nRMSE	10.783	2.547	7.832	1.512
90%	Clay	R^2	0.390	0.037	0.760	0.023	CCC	0.604	0.025	0.860	0.016	nRMSE	11.935	2.059	8.015	2.063
80%	Clay	R^2	0.340	0.036	0.690	0.026	CCC	0.582	0.030	0.830	0.047	nRMSE	13.831	2.392	8.573	1.462
70%	Clay	R^2	0.350	0.041	0.660	0.017	CCC	0.535	0.038	0.810	0.023	nRMSE	14.552	1.586	9.117	1.647
60%	Clay	R^2	0.290	0.033	0.670	0.039	CCC	0.427	0.053	0.820	0.033	nRMSE	19.085	4.373	8.714	2.308
50%	Clay	R^2	0.220	0.025	0.620	0.016	CCC	0.382	0.039	0.780	0.029	nRMSE	26.972	3.813	8.937	3.015
40%	Clay	R^2	0.120	0.031	0.470	0.013	CCC	0.173	0.024	0.720	0.051	nRMSE	42.584	7.004	12.470	2.672
30%	Clay	R^2	0.070	0.010	0.280	0.018	CCC	0.118	0.032	0.530	0.052	nRMSE	46.446	6.191	13.179	4.661
20%	Clay	R^2	0.030	0.006	0.160	0.010	CCC	0.091	0.026	0.340	0.047	nRMSE	53.504	4.338	21.017	2.803
10%	Clay	R^2	0.007	0.007	0.050	0.008	CCC	0.022	0.019	0.210	0.051	nRMSE	54.730	7.074	26.749	3.152
100%	pH	R^2	0.338	0.015	0.532	0.027	CCC	0.484	0.016	0.607	0.009	nRMSE	14.463	1.518	11.043	1.582
90%	pH	R^2	0.314	0.018	0.536	0.014	CCC	0.464	0.013	0.572	0.018	nRMSE	14.861	2.500	11.364	0.834

80%	pH	R²	0.310	0.014	0.514	0.024	CCC	0.456	0.012	0.551	0.016	nRMSE	15.849	2.015	11.973	1.260
70%	pH	R²	0.293	0.019	0.524	0.020	CCC	0.427	0.023	0.537	0.013	nRMSE	17.181	1.284	12.351	1.145
60%	pH	R²	0.297	0.020	0.490	0.025	CCC	0.422	0.012	0.528	0.013	nRMSE	18.447	2.887	12.947	2.731
50%	pH	R²	0.253	0.026	0.500	0.018	CCC	0.397	0.016	0.536	0.007	nRMSE	25.768	4.092	13.628	3.203
40%	pH	R²	0.169	0.035	0.360	0.011	CCC	0.243	0.018	0.392	0.026	nRMSE	29.058	3.739	13.774	2.005
30%	pH	R²	0.124	0.010	0.326	0.017	CCC	0.147	0.008	0.314	0.027	nRMSE	42.869	3.681	15.735	2.525
20%	pH	R²	0.056	0.022	0.210	0.017	CCC	0.067	0.007	0.172	0.012	nRMSE	50.630	5.945	28.436	3.568
10%	pH	R²	0.036	0.013	0.094	0.022	CCC	0.015	0.008	0.153	0.009	nRMSE	59.151	2.919	35.843	2.093
100%	EC	R²	0.422	0.032	0.608	0.014	CCC	0.307	0.032	0.647	0.011	nRMSE	29.194	2.907	10.751	1.020
90%	EC	R²	0.386	0.026	0.588	0.018	CCC	0.311	0.023	0.582	0.024	nRMSE	30.010	3.130	11.851	1.694
80%	EC	R²	0.368	0.028	0.594	0.025	CCC	0.286	0.011	0.579	0.014	nRMSE	31.665	2.629	12.594	1.379
70%	EC	R²	0.338	0.035	0.582	0.023	CCC	0.237	0.012	0.527	0.021	nRMSE	32.477	2.215	12.392	1.173
60%	EC	R²	0.349	0.028	0.551	0.029	CCC	0.234	0.026	0.463	0.018	nRMSE	32.870	1.738	16.220	2.335
50%	EC	R²	0.237	0.032	0.473	0.035	CCC	0.192	0.033	0.391	0.037	nRMSE	36.560	2.931	16.951	1.852
40%	EC	R²	0.136	0.037	0.357	0.022	CCC	0.148	0.017	0.214	0.015	nRMSE	36.780	2.039	20.732	1.074
30%	EC	R²	0.108	0.034	0.308	0.037	CCC	0.126	0.012	0.206	0.023	nRMSE	39.905	1.771	27.822	3.960
20%	EC	R²	0.054	0.024	0.164	0.019	CCC	0.092	0.012	0.173	0.017	nRMSE	42.757	1.415	29.833	2.476
10%	EC	R²	0.031	0.020	0.044	0.019	CCC	0.024	0.020	0.171	0.012	nRMSE	48.091	5.755	36.701	1.965
100%	SOM	R²	0.472	0.029	0.572	0.034	CCC	0.627	0.022	0.804	0.011	nRMSE	9.233	0.992	8.371	1.663
90%	SOM	R²	0.454	0.035	0.556	0.036	CCC	0.624	0.017	0.762	0.024	nRMSE	9.815	0.752	8.299	1.105
80%	SOM	R²	0.456	0.038	0.568	0.025	CCC	0.621	0.016	0.716	0.014	nRMSE	11.458	0.846	8.861	1.048
70%	SOM	R²	0.442	0.033	0.542	0.038	CCC	0.592	0.023	0.663	0.021	nRMSE	11.828	0.959	9.829	0.902
60%	SOM	R²	0.437	0.026	0.511	0.031	CCC	0.562	0.016	0.651	0.018	nRMSE	12.676	1.311	10.884	0.893
50%	SOM	R²	0.306	0.031	0.447	0.027	CCC	0.311	0.031	0.462	0.037	nRMSE	15.168	2.017	13.034	2.117
40%	SOM	R²	0.128	0.021	0.258	0.018	CCC	0.206	0.011	0.313	0.015	nRMSE	21.209	3.894	13.790	2.067
30%	SOM	R²	0.018	0.010	0.132	0.032	CCC	0.136	0.018	0.201	0.023	nRMSE	25.385	2.504	15.480	2.671
20%	SOM	R²	0.006	0.005	0.154	0.029	CCC	0.072	0.016	0.187	0.017	nRMSE	33.633	2.900	17.713	1.332
10%	SOM	R²	0.002	0.008	0.076	0.023	CCC	0.031	0.027	0.113	0.012	nRMSE	34.781	2.143	27.480	1.255
100%	TN	R²	0.358	0.031	0.498	0.021	CCC	0.438	0.029	0.612	0.030	nRMSE	18.547	1.499	12.564	1.980
90%	TN	R²	0.352	0.025	0.492	0.014	CCC	0.432	0.029	0.580	0.032	nRMSE	19.952	1.824	14.805	1.419
80%	TN	R²	0.326	0.019	0.484	0.019	CCC	0.364	0.024	0.530	0.040	nRMSE	20.728	1.155	15.405	0.570
70%	TN	R²	0.334	0.022	0.458	0.040	CCC	0.346	0.025	0.528	0.029	nRMSE	23.132	1.881	17.646	2.005
60%	TN	R²	0.258	0.034	0.452	0.015	CCC	0.339	0.037	0.517	0.012	nRMSE	24.887	1.096	18.738	1.893
50%	TN	R²	0.173	0.027	0.415	0.022	CCC	0.331	0.041	0.443	0.035	nRMSE	26.908	2.168	20.362	1.005
40%	TN	R²	0.114	0.035	0.272	0.029	CCC	0.236	0.019	0.322	0.014	nRMSE	28.378	2.047	22.562	0.893
30%	TN	R²	0.064	0.036	0.158	0.034	CCC	0.096	0.018	0.240	0.024	nRMSE	38.305	1.634	26.452	1.201
20%	TN	R²	0.028	0.011	0.144	0.023	CCC	0.082	0.016	0.154	0.023	nRMSE	50.330	3.717	27.052	1.700
10%	TN	R²	0.014	0.011	0.084	0.009	CCC	0.044	0.034	0.144	0.019	nRMSE	60.134	3.732	36.375	2.796