

Analysis of Prediction Maps and Data Separation Methods for Site-Specific Management of Wild Blueberry

Allegra Johnston

Department of Bioresource Engineering
McGill University
Montreal, Canada

August 2018

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science.

© 2018 Allegra Johnston

Abstract

Wild blueberry (*Vaccinium angustifolium* Ait.) is a key crop in the Lac-Saint-Jean region of Quebec. The industry totals \$45 million annually. Wild blueberry is a lowbush species which flourishes in heterogeneous agronomic conditions where conventional crops cannot. It grows in areas of varying topography on sandy, acidic soils where competition with other plants is limited. Rhizome establishment takes years to develop, thus, bare spots are a common feature of young or poorly managed fields. Given the variation of soil, topography, and crop density, wild blueberry production would benefit from site-specific management, where levels of nutrient input are tailored to local needs based on within-field variation. A classic approach to site-specific management is the delineation of management zones, sub-field areas of relatively homogenous agronomic properties with uniform management rates. A second SSM approach is regression based, where a prescription regression equation based on sampled variables and known crop response to treatment is used for more continuous targeted treatment within the field. This thesis articulates the thematic mapping of agronomic variables and the comparison of two site-specific management strategies for wild blueberry using conventional soil sampling, proximal soil sensors, and multispectral satellite imagery. Two experimental sites were selected, one of varying topography and the other relatively flat. Soil samples were collected in a 33-m grid scheme and tested for chemical and granulometric attributes. Soil apparent electrical conductivity (EC_a) was collected with the non-contact DUALEM-21S sensor (Dualem Inc., Milton, ON) and the contact Veris 3100 sensor (Veris Technologies, Salina, KS). Elevation was mapped with real-time-kinematic (RTK) level global navigation satellite system (GNSS) receiver. Multispectral imagery acquired from the SPOT6 archive was radiometrically and atmospherically corrected, and a number of vegetation indices were derived from the image to map bare spots and compare VIs prediction of vigor to the sampled yield. Thematic maps were predicted from the sampled data using the Ordinary Kriging (OK) method and cross-validated to determine the

strength of various data layers in predicting spatial patterns within-field. Classical statistics and geostatistics were performed on all sampled data. A classic approach to site-specific management through unsupervised classification of management zones was compared with a new regression-based approach which targets four field condition scenarios. Means of all properties in each of four scenarios were tested with ANOVA and Tukey's post-hoc test. In both MZA and the regression-based method, field conditions were most contrasted between scenarios EC_{low} & $Elev_{high}$ and EC_{high} & $Elev_{low}$. The regression-based method separated data similarly or better than the MZA approach, while providing more precise areas to develop a regression-based prescription map.

Résumé

La culture de bleuet nain sauvage (*Vaccinium angustifolium* Ait.) représente une valeur de \$45 millions annuellement dans la région de Lac-Saint-Jean au Québec. Le bleuet nain sauvage fleurit dans des conditions agronomiques hétérogènes où la récolte conventionnelle ne peut pas. Il grandit dans les zones de topographie variée sur les sols sablonneux et acides où la compétition avec d'autres plantes est limitée. L'établissement rhizomique prend quelques années pour se développer, donc les endroits dénudés sont une fonction commune de champs jeunes ou mal gérés. En raison de la variation du sol, de la topographie et de la densité de culture, la production de bleuet nain sauvage profiterait d'une gestion spécifique, où les niveaux d'apport nutritif sont adaptés aux besoins locaux, basés sur les variabilités intra-parcellaires. Une approche classique à la gestion spécifique est la délimitation de zones d'aménagement (ZA), constituées de propriétés agronomiques relativement homogènes avec les niveaux uniformes de gestion. Une seconde approche est à base d'une régression, lorsqu'une équation de régression sur ordonnance basée sur les variables échantillonnées et une réponse de culture connue au traitement est utilisée pour un traitement ciblé plus continu dans le champ. Cette thèse articule la cartographie thématique de variables agronomiques et la comparaison de deux stratégies de gestion spécifique pour le bleuet nain sauvage, utilisant l'échantillonnage de sol conventionnel, les capteurs proximales de sol et les images satellite. Deux sites expérimentaux ont été choisis, une de topographie variée, l'autre relativement plate. Les échantillons de sol ont été rassemblés dans un plan de grille de 33-m et testés pour les attributs granulométriques et le produit chimique. La conductivité électrique apparente du sol (CEA) a été rassemblée avec les capteurs DUALEM-21S et Veris-3100 afin de comparer les deux. L'élévation a été capturée avec la technologie de positionnement global cinématique en temps réel (OTF). L'image satellite a été acquise de l'archive de SPOT6 et a été corrigée pour les effets radiométriques et atmosphériques. Un certain nombre d'indices de végétation (VI) ont été

dérivés de l'image pour comparer avec le rendement mesuré et afin d'identifier les endroits dénudés. Les cartes thématiques ont été cartographiées avec la méthode d'interpolation par krigeage, puis contre-validées afin d'évaluer la variabilité spatiale de diverses propriétés agronomiques. Les statistiques classiques et les géostatistiques ont été exécutés sur toutes les données échantillonnées. La délimitation de ZAs utilisant l'algorithme de classification *k-moyennes* a été comparée avec une nouvelle approche à base de régression qui vise les endroits extrêmes du champ. Les moyens de toutes les propriétés dans chacun de quatre scénarios ont été testés avec ANOVA et le test de post-hoc de Tukey. Avec les deux approches, les deux champs montrent le contraste le plus grand entre les scénarios $EC_{Low}Elev_{High}$ et $EC_{High}Elev_{Low}$. La méthode basée sur la régression a séparé les données de façon similaire ou meilleure que l'approche de ZA, tout en fournissant des zones plus précises pour développer une carte de prescription basée sur la régression.

Acknowledgments

This thesis would not be possible without the intellectual, emotional and instrumental support of several people. This research was funded by Agriculture and Agri-food Canada (AAFC). Field work was jointly executed by the research teams of Mr. Jean Lafond and Dr. Athyna Cambouris of AAFC. Specifically, much gratitude goes to Denis Bourgault, Rosaire Cantin and Daniel Boisclair for yield sampling and chemical analysis, Claude Lévesque, Sarah-Maude Parent, and Mario Deschênes for soil sampling and chemical analysis, Samson Soticonol for DUALEM mapping technical support, and Isabelle Perron for sampling methodology and GIS field boundary mapping. In addition to the valuable data provided by these people, they were all supportive colleagues delightful to work with. A special thanks to the Precision Agriculture and Sensor Systems (PASS) lab group for the exchange of ideas and interdisciplinary talent in the lab. Particularly, thank you to Hsin-Hui Huang for providing ideas and advice for GIS analysis and Md Saifuzzaman for obtaining the SPOT satellite imagery. Thank you to Prof. Pierre Dutilleul for his help with geostatistical analysis. Finally, a great appreciation to my co-supervisors Dr. Athyna Cambouris, Dr. Asim Biswas, and Dr. Viacheslav Adamchuk, for their steady guidance and tremendous support in not only my thesis endeavor but my overall graduate school experience. I have been awed by your generosity of time and expertise. Last, but not least a very warm thank you to my parents, my sister, my friends, and to my voice teacher Tracy Smith and the Schulich School of Music for their support of my endeavor to pursue operatic singing and scientific research simultaneously.

Contribution of Authors

This masters project was co-supervised by Dr. Viacheslav Adamchuk, Dr. Athyna Cambouris, and Dr. Asim Biswas. The project was funded and overseen as part of a larger research project by Dr. Athyna Cambouris of Agriculture and Agrifood Canada (AAFC), with the cooperation and expertise of Mr. Jean Lafond and Ms. Isabelle Perron of AAFC. Results of this study were presented at the 2017 Canadian Soil Science Society Conference and at the 2018 International Conference of Precision Agriculture.

Table of Contents

1	Introduction	15
1.1	Wild Blueberry Production Challenges	15
1.2	Research Objectives	17
2	Literature Review: Site-Specific Management	18
2.1	Selection of Yield-Limiting Factors to Characterize Within-Field Variation	19
2.1.1	Apparent electrical conductivity	19
2.1.2	Elevation and topography	20
2.1.3	Remote sensing	21
2.2	Data Interpolation	22
2.2.1	Ordinary kriging	23
2.2.2	Ordinary kriging vs. kriging with auxiliary variables	24
2.3	Data Separation & Classification	25
2.3.1	Principal components and partial least squares	25
2.3.2	K-means & c-means clustering	26
2.3.3	Hierarchical classification	27
2.3.4	Neighborhood Search Analyst	28
2.4	Validation of Classes or Management Zones	28
2.5	Implementation with Variable Rate Technologies	29
2.6	Persisting Challenges in Site-specific Management	29
3	Materials & Methods	31
3.1	Site Description	31
3.2	Historic Yield and Management Practices	31
3.3	Description of Data Layers	34

3.3.1	Laboratory analysis of chemical and granulometric soil properties	34
3.3.2	Apparent electrical conductivity sampling	36
3.3.3	Satellite imagery	39
3.4	Map Interpolation	42
3.4.1	Pre-processing	42
3.4.2	Spatial prediction	42
3.5	Correlation Analysis & PLS	44
3.6	Classification & Data Separation	44
4	Results & Discussion	46
4.1	Laboratory Analyses	46
4.2	Proximal Sensor Data	49
4.3	SPOT Imagery and Vegetation Indices	52
4.4	Correlation & PLS	54
4.5	Geostatistical Analysis	57
4.6	Data Separation for Regression-Based Approach	65
4.7	MZA vs. Regression-Based Approach	75
5	Conclusions & Recommendations	79
5.1	Research Study Summary	79
5.2	Future Research	80
	Appendices	98
A	Field 21 Tukey Comparison	99
B	Field 140b Tukey Comparison	101
C	Bare Spot Standards Scores	103
D	Field 21 Standard Scores	105
E	Field 140b Standard Scores	107

List of Figures

3.1	Experimental fields 21 and 140b with soil/yield sampling locations and elevation/ EC_a sensor track.	32
3.2	Depiction of data layers and their relationships. TWI is the topographic wetness index derived from elevation and slope data. EC_a is the soil apparent electrical conductivity captured by two sensors, DUALEM and Veris. Multispectral SPOT6 is the multispectral satellite image, comprising four wavelength bands: R (red), G (green), B (blue), and NIR (near infrared). Sample and sensor data encompass yield-determining properties such as cation exchange capacity, water holding capacity, soil porosity, crop density, and element exposure. Soil pH indirectly influences yield by limiting weeds. . .	35
3.3	Dipole arrangement of DUALEM EMI sensor. T is the transmitter coil. H1 and H2 receivers receive the induced current in horizontal co-planar (HCP) arrays at 1-m and 2-m respectively. P1 and P2 receiver receive the current in perpendicular co-planar (PCP) arrays at 1.1-m and 2.1-m respectively. Receivers at a further distance from the transmitter coil capture EC_a at a greater depth.	37
3.4	Veris 3100 configuration of six roller coulter, one pair which passes a current through the soil and two pairs which measure resistance to extract EC_a at two depths (Oguri <i>et al.</i> , 2009).	38
3.5	Veris 3100 set up. Photo credit to Agriculture and AgriFood Canada. . . .	39
4.1	Classification of bare spots in SPOT6 satellite image using the vegetation index derived from Principal Components Analysis.	53
4.2	Field 21 strongly correlated elevation and EC_a	62
4.3	Field 21 moderately correlated EC_a and chemical properties	63

4.4	Field 21 strongly correlated granulometric maps.	64
4.5	Field 21 moderately correlated granulometric maps.	65
4.6	Field 140b strongly correlated elevation and EC _a maps.	66
4.7	Field 140b strongly correlated granulometric maps.	67
4.8	Field 140b moderately correlated granulometric and chemical maps.	68
4.9	Field 21 scatter plot of Elevation vs. EC _a values. The bimodal distribution of elevation is apparent in the scatter plot as two distinct clusters. Bare spots are highlighted in red and correspond with higher elevation.	69
4.10	Field 140b scatter plot of Elevation vs. EC _a values. Bare spots are highlighted in red	70
4.11	Management zones classified with fuzzy c-means in MZA for Field 21 (left) and Field 140b (right). Targeted scenarios from the regression-based approach overlay the zones for comparison.	71
4.12	Tukey plots of Field 21 elevation, EC _a , and slope. Slope is significantly higher in scenario <i>Elev_{high}&EC_{high}</i> than the other three scenarios.	72
4.13	Field 21 attributes that are significantly different based on contrast in elevation. Texture is influenced by elevation.	73
4.14	Field 21 soil attributes that are contrasted by the combination of elevation and EC _a	74
4.15	Field 140b elevation and EC _a separation.	75
4.16	Field 140b variables separated by the contrast of elevation and EC combined.	76
4.17	Field 140b variables separated by the contrast of elevation and EC combined (P) and variables separated by high vs. low EC _a	78
5.1	Schematic of the field trial with one 30 x 30 m plot containing 16 subplots for 4 treatments of 0, 30, 60 and 90 kg N/ha at four repetitions.	81

List of Tables

2.1	Three common variogram models that depict autocorrelation. $\gamma(h)$ is the variance at lag distance h ; a is the distance up to which autocorrelation exists, until constant variance is reached.	23
3.1	Field 21 summary of yield and treatments 2009-2016.	33
3.2	Field 140b summary of yield and treatments 2009-2016.	33
3.3	Derived depths of EC _a instruments where cumulative response is 70% (Mat Su, 2016).	37
3.4	SPOT Multi-spectral and Panchromatic Image wavelength bands.	40
3.5	Summary of ratio-based vegetation indices (VIs) calculated from SPOT6 satellite image. NIR is the near-infrared wavelength range, R is the red wavelength range, G is the green wavelength range, and B is the blue wavelength range. γ is a weighting function for aerosol conditions in the atmosphere. L the canopy background adjustment factor.	41
4.1	Field 21 summary statistics of chemical and granulometric soil attributes; n is the sample size, STD is standard deviation, and CV% is coefficient of variation.	47
4.2	Field 140b summary statistics of chemical and granulometric soil attributes; n is the sample size, STD is standard deviation, and CV% is coefficient of variation.	48

4.3	Field 21 summary statistics for EC_a data and derived topographic attributes; n is the sample size, STD is the standard deviation, and CV% is the coefficient of variation. Offset values are the original sensor values plus the minimum value collected from all depths. The offset was applied to all layers so that depths could be compared.	50
4.4	Field 140b summary statistics for EC_a data and derived topographic attributes; n is the sample size, STD is the standard deviation, and CV% is the coefficient of variation. Offset values are the original sensor values plus the minimum value collected from all depths. The offset was applied to all layers so that depths could be compared.	51
4.5	Pearson's correlation coefficients.	55
4.6	Field 21 total percent variation accounted for by PLS factors.	58
4.7	Field 140b total percent variation accounted for by PLS factors.	59
4.8	Geostatistical summary of agronomic properties in Field 21.	60
4.9	Geostatistical summary of agronomic properties in Field 140b.	61
A1	Field 21 Tukey test comparison between MZA classification and targeted approach.	100
A1	Field 140b Tukey test comparison between MZA classification and targeted approach.	102
A1	Standard score of soil variables in bare patches. The score represents distance from the field mean. A value close to 0 has little difference from mean. $N_{21} = 8$ and $N_{140b} = 2$	104
A1	Field 21 standard scores of soil variables by regression-based separation method.	106
A1	Field 140b standard scores of soil variables by regression-based separation method.	108

List of Acronyms

CEC	Cation Exchange Capacity
CK	Co-Kriging
CV	Coefficient of Variation
DVI	Difference Vegetation Index
EC _a	Apparent electrical conductivity
EMI	Electromagnetic Induction
EVI	Enhanced Vegetation Index
GARI	Atmospherically Resistant Index
GDVI	Green Difference Vegetation Index
GLS	Generalized Least Squares
GPS	Global Positioning System
GNSS	Global Navigation Satellite System
GRVI	Green Ratio Vegetation Index
HCP	Horizontal Co-Planar
KED	Kriging with External Drift
LAI	Leaf Area Index
MNLI	Modified Non-Linear Index
MSAVI2	Modified Soil-Adjusted Vegetation Index
MSE	Mean Square Error
MSR	Modified Simple Ratio
MZ	Management Zones
MZA	Management Zone Analyst
NDVI	Normalized Difference Vegetation Index
NSA	Neighborhood Search Analyst

OSAVI	Optimized Soil Adjusted Vegetation Index
OK	Ordinary Kriging
PA	Precision Agriculture
PASS	Precision Agriculture and Sensor Systems
PCA	Principal Components Analysis
PLS	Partial Least Squares
PRP	Perpendicular Planar
RDVI	Renormalized Difference Vegetation Index
RK	Regression Kriging
RMSE	Root Mean Squared Error
RMSE _{SD}	Standardized Root Mean Squared Error
RTK	Real-Time Kinematic
SSM	Site-Specific Management
TDVI	Transformed Difference Vegetation Index
TWI	Topographic Wetness Index
VI	Vegetation Index
VR	Variable Rate

Chapter 1

Introduction

In the Saguenay-Lac-St-Jean region of Quebec, 27,000 ha of land are used for wild blueberry cultivation, producing a business industry value of over \$45 million annually in Quebec [?]. Wild blueberry (*Vaccinium angustifolium* Aiton) is a low maintenance crop; it grows on low shrubs in naturally acidic, sandy soils and can withstand long, harsh winters [?]. Wild blueberries thrive when competitors are limited, which entails maintaining low pH soils and applying herbicides to control weed growth. Though it is a perennial crop, bushes are cut in the sprout year to produce new stems, and fruit is harvested every second year. Wild blueberries may be cultivated in fields of heterogeneous growing conditions, with local changes in topography, key soil properties, and crop density. With such heterogeneous growing conditions and unique production challenges, wild blueberry cultivation should greatly benefit from precision agriculture techniques and site-specific management.

1.1 Wild Blueberry Production Challenges

Wild blueberry responds positively to Nitrogen (N) and Phosphorus (P) applications. Herbicides are applied as needed to control annual and perennial weeds. Fungicides are applied to control for Septoria leaf spot (*Septoria* spp.) and Valdensinia leaf spot (*Valdensinia heterodoxa*) in affected fields. N is the principle limiting nutrient in wild blueberry cultivation [?](Lafond and Ziadi, 2011). However, excess N application will cause an overgrowth of leaves, delayed and stunted fruiting, and disease susceptibility (Percival and Sanderson, 2004; "Soil fertility and fertilizers for wild blueberry production", 2013). Furthermore, mis-application of nutrients and pesticides threaten the surrounding ecosystem. When N

fertilization exceeds crop needs and soil pH is above average, N will volatilize as ammonia from the soil and crop (Jones and Jacobsen, 2005). Volatilized ammonia deposits in the environment as ammonium, acidifying soils (Jones and Jacobsen, 2005; Istas *et al.*, 1988). High N application in conjunction with heavy rainfall over sandy soils may also cause nitrate (NO_3^-) leaching, threatening water quality (Favaretto *et al.*, 2006; Saleem *et al.*, 2013). Site-specific nutrient management based on variability of within-field conditions ideally reduces both the cost of inputs and environmental waste.

A secondary challenge to wild blueberry production is the occurrence of bare spots in young and/or poorly managed fields. One study by Zaman *et al.* (2008) found the percentage of bare spots in their wild blueberry study sites varied between 30 and 50%. Bare spots may be found in areas of varying terrain where plant growth is sparse and more exposed to winter frost (“Filling Bare Spots in Wild Blueberry Fields”, n.d.). Wild blueberries have a shallow rooting depth, generally less than 40 cm with lateral rhizomes (Hicklenton, 2000). Frost will destroy blueberry rhizomes which take multiple years to recover in agriculture fields. Producers typically manage large, contiguous bare spots separately from the rest of the field because they require different nutrient rates and less frequent cutting. Identifying these areas saves resources and optimizes bare spot treatment.

Due to heterogeneous growing conditions and the cultivation challenges discussed, wild blueberry crops would benefit from within-field site-specific management (SSM), where nutrient inputs are tailored to specific regions of the field depending on their site-specific needs. To date, few studies have applied SSM practices to wild blueberry cultivation. Saleem *et al.* (2013) delineated sub-fields, or management zones (MZs), for SSM in New Brunswick using slope and soil property data in conjunction with bare spot maps. Similarly, Farooque *et al.* (2012) developed MZs based on soil apparent electrical conductivity (EC_a), digital color photography, and sampled soil properties. Present research has established a link between soil EC_a , soil properties, and wild blueberry yield (Farooque *et al.*, 2012; Saleem *et al.*, 2013). However, previous studies have focused on the conventional MZ delineation approach to SSM, whereby sub-field classified zones are treated with uniform rates. This research project draws from previous studies’ utilization of topographical and EC_a data to characterize within-field variation of yield and nutrient requirements. However, this project proposes a new approach which only identifies areas of extreme contrast within-

field in order to develop a regression-based prescription map. Furthermore, this project investigates the use of satellite imagery to supplement yield data and to identify bare spots within wild blueberry fields.

1.2 Research Objectives

The goal of this study as proposed by Agriculture and Agrifood Canada was to characterize soil spatial variability in wild blueberry fields and investigate the potential of SSM of wild blueberry based on proximally-sensed EC_a . Building upon this principle objective, the specific objectives of this masters project were to:

1. Develop thematic maps to characterize and classify within-field variability based on sampled soil data and measurements from proximal contact and non-contact EC_a sensors. It was hypothesized that spatial variability patterns in soil texture and chemistry would correlate with yield patterns.
2. Evaluate the use of satellite imagery to supplement yield information and to identify bare patches in the fields. It was hypothesized that ratio-based green vegetation indices would correlate strongly with yield and that bare spots would be delineated by satellite imagery.
3. Identify extreme agronomic conditions within-field using elevation and EC_a data in order to develop a regression-based prescription for SSM. It was hypothesized that identified extreme areas would be spatially distinct as well as distinct in sampled yield and soil attributes.
4. Compare results of this regression-based approach to the conventional MZ approach for SSM. It was hypothesized that the regression-based approach would be more distinct in yield and soil attributes than the MZ approach.

Chapter 2

Literature Review: Site-Specific Management

Site-specific management (SSM) is the practice of treating crop within a field according to local needs. In the 1990s, advancement in variable rate (VR) seeding and fertilization equipment allowed for more precise measurement and input application, spurring researchers and farmers to map and classify within-field variation of several yield-limiting factors for SSM practices (Sawyer, 1994; Wollenhaupt, 1994; Khosla *et al.*, 2008).

The conventional approach to SSM in Precision Agriculture is to subdivide a field into management zones (MZ) of relatively homogenous characteristics which influence yield and require similar treatment levels. The methodology behind MZ delineation varies in degree of complexity, but the fundamental procedure involves (1) selecting and sampling possible yield-limiting field parameters; (2) separating parameters using various numerical methods such as unsupervised classification; (3) validating separability; and finally, (4) implementing SSM within classified zones with the prescribed treatments. Detailed in this literature review is a description of the classic methodology of MZ delineation, followed by its existing challenges, some proposed solutions, and finally, a newer regression-based approach to site-specific management which producers and researchers should consider.

2.1 Selection of Yield-Limiting Factors to Characterize Within-Field Variation

Due to the complex biotic and abiotic factors affecting variability in yield – from seasonal fluctuations to historic management to soil quality – MZ delineation based on yield alone is often unreliable (Kerry *et al.*, 2016). Yield is challenging to quantify particularly in specialty crops where standard yield sensors are not applicable. For example, Zaman *et al.* (2008) identified debris, uneven topography, and fragility of fruit as limitations to monitoring wild blueberry yield. Instead, MZ delineation is more often based on temporally stable soil variables which have been demonstrated to directly affect yield. Mzuku *et al.* (2005) divided corn fields by productivity potential and compared soil attributes across zones. Mallarino and Wittry (2004) investigated how traditional soil sampling captured within-field yield variability. However, soil sampling is expensive, time consuming, and does not necessarily capture spatial patterns if sampled at too great an interval (Bianchini and Mallarino, 2002; Lauzon, *et al.*, 2005). Researchers have turned to proximal and remote sensing for quick, dense sampling of data that is auxiliary to soil and yield samples. Most commonly soil apparent electrical conductivity (EC_a), field elevation, and aerial or satellite imagery provide high-density data sets to be used either alone, or in conjunction, with sampled soil attributes to create thematic maps and delineate MZs when correlation with soil attributes and/or yield are well established (Khosla *et al.*, 2008; Kerry *et al.*, 2016).

2.1.1 Apparent electrical conductivity

Soil EC_a represents conductance through soil particles and soil water solution that envelops soil particles (Rhoades, 1993). Conductance is higher in finer textured particles because increased porosity in fine textured soils has a greater water to air ratio (Robain *et al.*, 1996). Additionally, soils with higher clay content and organic matter will have higher cation exchange capacity (CEC) due to their negatively charged surfaces. Conductance is higher in soils with greater CEC because of the higher concentration and mobility of ions (Helfferrich, 1962). A number of studies have demonstrated soil EC_a to be correlated with soil organic matter (SOM), CEC, moisture content, and soil texture, characteristics which directly affect crop yield (Johnson *et al.* 2001; Carroll and Oliver, 2005; Kitchen and Sudduth, 2005; Whelan and McBratney, 2003; Friedman, 2005). Soil EC_a is measured either with

non-contact sensors (Hedley *et al.*, 2004; Kachanoski *et al.*, 1990) or with contact sensors (Farahani *et al.*, 2004; Halvorson *et al.*, 1976). Non-contact sensors utilize electromagnetic induction (EMI) to measure EC_a while contact sensors are electrode-based. In non-saline soils, EC_a variability is largely due to variations in soil water content (Kachanoski *et al.*, 1988). When soil is saturated EC_a measurements indicate water storage potential in various parts of the field (Kachanoski *et al.*, 1990). Unsaturated non-saline soil will indicate variations in both moisture availability and soil texture (Kitchen and Sudduth, 2005). The depth at which soil conductivity is measured is related to the distance and configuration of the current transmitters and receivers. In many instruments, EC_a is measured at multiple depths simultaneously, providing insight into soil characteristics through the crop rooting zone to the parent material. Saey *et al.* (2009) used multiple depths of EC_a to map depth to clay in a field and Park and Vlek (2002) used multiple depths of EC_a data to model soil variability in three dimensions.

A number of studies have compared contact and non-contact EC_a sensors (Serrano *et al.*, 2013; Sudduth, 2001; Saey *et al.*, 2009). When paired with a global navigation satellite system (GNSS), both types of sensors will record and provide georeferenced data almost instantaneously. EMI sensors offer the benefit of being non-invasive. Furthermore, the reliability of EMI sensors depends on the thermal stability of the instrument (Abdu and Robinson, 2007). Drift is a common issue in EMI sensors as the instrument temperature increases with time. Sudduth *et al.* (2001) suggest that a calibration transect be taken which crosses other transects so any drift over time may be detected and corrected in data post-processing. Myers *et al.* (2008) combines EMI and penetrative EC_a sensors to further improve accuracy.

2.1.2 Elevation and topography

Elevation and topographic information derived from elevation data (topographic wetness index, slope, aspect) are also commonly used in MZ delineation (Khosla *et al.*, 2002; Fraisse *et al.*, 2001; Vitharana *et al.*, 2008). Topography indirectly affects yield by influencing water holding capacity, nutrient storage, and water movement. Elevation and slope are correlated with texture and organic matter as they affect the distribution of fine and coarse-grained soils, in turn indirectly influencing EC_a measurements.

Kravchenko and Bullock (2000) studied how topographical information correlated with soybean yield and soil properties. They found that elevation had the greatest influence on yield, with higher yield in lower landscapes. They further found curvature, slope, and flow accumulation significantly affected yield in extreme topographical locations with very high or very low precipitation. Elevation can be measured at high precision and accuracy with real-time kinematic (RTK) and GNSS sensors. Similarly, automated slope measurement and mapping systems have been used to proximally measure field slope (Zaman *et al.*, 2008). Studies have had success relating yield to topography when combined with soil attribute data (Kravchenko and Bullock 2000; Nolan *et al.* 2000). Given the movement of N_2 in low drainage areas and NO_3^- leaching in coarse textured areas, Khosla *et al.* (2002) developed MZs for N management based on topography, historical yield, soil color, and aerial photographs. Additionally, a review by Vitharana *et al.* (2008) concluded that combining pH, EC_a , and elevation provided the most important properties for defining MZs in combinable crops.

2.1.3 Remote sensing

Remote sensing in precision agriculture is most often applied to yield mapping. One study by McCann *et al.* (1996) demonstrated that aerial photography effectively evaluated soil color, but the focus of research remains on supplementing crop yield data (Anderson and Yang 1996). Zaman *et al.* (2010) successfully developed a method for mapping wild blueberry yield with digital color photography calibrated to ground truth data. Kerry *et al.* (2016) used multispectral satellite images to delineate cranberry bogs and disaggregate yield estimates. Multispectral imagery that includes near-infrared (NIR) reflectance has also been used to map vegetation indices (VIs) (Aparicio *et al.*, 2000; Shanahan *et al.*, 2001; Curran, 1980). VIs are ratios or linear combinations of band reflectance which serve to quantify image vegetation properties, such as the photosynthetic activity, vigor, and other growth factors. The near-infrared (NIR) range of reflectance is often utilized in VIs because vegetation reflectance increases dramatically from the red to NIR wavelength ranges in a pattern referred to as the “red edge” due to chlorophyll absorption and leaf internal scattering (Dawson and Curran, 1998). The normalized difference vegetation index (NDVI) is often used to characterize canopy growth of green vegetation, but other VIs have been

developed to account for noise in reflectance data which can affect NDVI. Soil reflectance will impact both red and NIR reflectance in a linear relationship referred to as the “soil line” (Baret *et al.*, 1993; Richardson *et al.*, 1977). The modified soil-adjusted vegetation index (MSAVI2) is one of many soil-adjusted VIs which factor in the soil line (Qi *et al.*, 1994b). Other VIs, such as the new atmospheric effect resistant vegetation index (IAVI), will correct for atmospheric effects. Liu and Huete (1995) found an interaction between the influence of soil and the atmosphere, such that a reduction in one increased the other. They proposed the enhanced vegetation index (EVI) as a feedback to balance the two. VIs have been well researched and are especially robust for quantifying yield in green crops like maize, but little research has investigated satellite imagery for yield estimation in fruit (Lee *et al.*, 2010). Limiting factors to remote sensing are spatial resolution when dealing with specialty crops like wild blueberries, cost of data, and quantifying yield. Mapping yield alone is not often recommended for MZ delineation unless long term temporal patterns have been established (Kerry *et al.*, 2016). However, it is useful when combined with other data layers (Lark and Stafford *et al.* 1998).

2.2 Data Interpolation

To predict sampled variables as continuous values on a map, geostatistical principles are most commonly applied. A key principle of geostatistics is that points which are located more closely together in space are more similar to each other than points further apart. This concept is known as the principle of autocorrelation. The variogram model illustrates the principle of autocorrelation. As lag distance between two spatial points increases, variance will increase at a decreasing rate until it reaches a constant variance. The rate at which it increases varies, and several theoretical variogram models exist to characterize it. The main three are the Gaussian, Spherical, and Exponential models (Table 2.1).

If no autocorrelation is apparent, the variance is a constant value at all distances and is called a pure nugget model, where the only variance occurring is due to random nugget effects. The theoretical variogram comprises three components: the range, the partial sill, and the nugget. The range is the distance up to which autocorrelation exists. The partial sill is the variance exhibited from the autocorrelation, and the nugget is the variance exhibited from a random component. The partial sill and nugget together make up the total

sill, which represents the constant variance.

Table 2.1 Three common variogram models that depict autocorrelation. $\gamma(h)$ is the variance at lag distance h ; a is the distance up to which autocorrelation exists, until constant variance is reached.

Model	Equation
Gaussian	$\gamma(h) = 1 - \frac{2}{3}\exp(-\frac{3h^2}{a^2})(2.1)$
Spherical	$\gamma(h) = \frac{1}{3} + \frac{2}{3}(1.5\frac{h}{a} - 0.5\frac{h^3}{a^3})(2.2)$
Exponential	$\gamma(h) = 1 - \frac{2}{3}\exp(-\frac{3h}{a})(2.3)$

The Gaussian model increases to constant variance more slowly while the Exponential model increases to a constant variance more quickly. The spherical model increases at rate faster than the Gaussian model but slower than the Exponential model.

2.2.1 Ordinary kriging

The geostatistical principles outlined are utilized in the kriging method of interpolation. The kriging method was developed in the 1951 by D.G. Krige and later finessed by Matheron (1962). The most standard version, Ordinary Kriging (OK), is predicted as follows:

$$\hat{z}_{OK}(\mathbf{s}_0) = \sum_{i=1}^n w_i(\mathbf{s}_0) \cdot z(\mathbf{s}_i) = \lambda_0^T \cdot \mathbf{z} \quad (2.4)$$

where λ_0 is the vector of kriging weights derived from the variogram, w_i at new locations $\mathbf{s}_0$, and $\mathbf{z}$ is the vector of observations at primary locations $\mathbf{s}_i$ (Hengl 2009). OK depends on some basic assumptions:

1. Mean stationarity: the mean between two samples is independent of location so that mean is constant;
2. The variogram is constant in the entire area;
3. The target variable is roughly normally distributed (Hengl, 2009).

These assumptions are rarely met in practice. External factors cause drift in data so that the mean stationarity is not met. This is often ignored for practical reasons. Further, data that is not normally distributed can be transformed either with a log or box-cox transformation.

2.2.2 Ordinary kriging vs. kriging with auxiliary variables

Proximally and remotely sensed data sets may be used as auxiliary variables in predicting yield and soil attributes. Kriging with auxiliary variables falls into three categories: Co-kriging (CK), Kriging with External Drift (KED), and Regression Kriging (RK). The latter two methods yield the same results using slightly different methodologies. The simplest method is CK which use a cross-variogram model to in addition to the variogram model in order to predict a variable correlated to the densely sampled auxiliary variable. Co-kriging EC_a with soil properties has been successful when correlations are well established (De Caires *et al.*, 2015; Frogbrook and Oliver, 2003). However, Knotters *et al.* (1995) warns CK relies on over-simplified relationships between variables.

The other two kriging methods, KED and RK, address the condition of mean stationarity which is a required assumption for OK. Mean stationarity is seldom met in the natural environment as external factors will cause drift in the data. KED uses auxiliary variables to model the deterministic component in kriging. Similarly, RK adds a model of the drift to the deterministic component in the prediction, and residuals of the model are interpolated with simple kriging. The distinction between the two methods is that RK models drift and residuals separately and then combines the two (Wackernagel, 2003). Regression coefficients for the model may be derived from fitting generalized least squares (GLS) (Hengl, 2007). If no spatial autocorrelation exists in the residuals, the equation simply reflects a multiple linear regression. Similarly, if there is no correlation between the auxiliary variables, the model mathematical resolves to the OK model. Regression kriging has become a favorite method for soil thematic mapping (Minasny and McBratney, 2007). Carré and Girard (2002) utilized multiple linear regressions of terrain attributes from elevation and land cover attributes from multispectral bands to represent soil types and to kriging residuals. Several studies have found the prediction power of RK exceeds OK and CK (Odeh *et al.*, 1995; Simbahan *et al.*, 2006).

The question of whether or not to model drift depends on the objective of the thematic mapping and quality of soil sampling. RK will produce a more coarse but accurate map to OK. A drawback of RK is the processing time to determine how many and which variables to use (Hengl *et al.*, 2004). Additionally, output in RK may be outside the physical range if predicted values are negative (Goovaerts, 1997). Consensus is generally that OK is simpler to use and more accurate when spatial structure is strong in the data, but otherwise RK with auxiliary variables can yield higher accuracy prediction maps (Zhu and Lin, 2010; Goovaerts, 1997, Li and Heap, 2014). For the purposes of this research, OK was selected.

2.3 Data Separation & Classification

Once yield-limiting variables have been selected and measured, the data is separated using one or more numerical techniques. For multivariate datasets, the number of variables are reduced using principal components analysis (PCA) or partial least squares (PLS) regression in order to determine the greatest sources of variation in the dataset.

2.3.1 Principal components and partial least squares

PCA treats multi-variate data to determine which variables contribute most to within-field variation. PCA reduces multivariate datasets to uncorrelated vector components representing the greatest source of variation in descending order. Principal components (PCs) reduce the overlap of multivariate data attributed to pairwise correlation and can be used to determine the smallest possible number of significant predictors of a target variable. Hengl (2004) recommends reduction of data to principal components prior to mapping and classification. The top two or three components are used to classify management zones. Fraisse *et al.* (2001) used PCs with highly correlated variables EC_a , N, OM, CEC, and NDVI to develop MZs that maximized variation. Moral *et al.* (2010) used PCs to map field variability.

PLS is similar to PCA in that it reduces datasets and multicollinearity. However, while PCA maximizes the variation in predictor variables, PLS maximizes the variation in both the predictor and the target variables by maximizing covariance. Bronson *et al.* (2005) utilized PLS regression analysis to determine the percentage of variation observed in EC_a

measurements that were explained by sampled soil attributes. Components may be further separated with unsupervised classification methods.

With the classic method of SSM, MZs are delineated by classifying sampled variables into clusters according to similarity. Numerical methods of classification may be supervised or unsupervised. Briefly, a distinction should be noted between management zones and classes - multiple classes in a field may belong to the same management zone if they share the same prescribed treatment.

Many statistical methods for clustering spatial data exist; the most common clustering methods in the MZ delineation literature are unsupervised, specifically k-means or c-means clustering (Fraisie *et al.*, 2001; Li *et al.*, 2007; Moral *et al.*, 2010; Fridgen *et al.*, 2000; Ortega *et al.*, 2007). K-means and c-means have been used interchangeably, but they consist of slightly different algorithms. Both algorithms optimize similarity of sample points and can be adjusted to include a fuzziness component where sample points may belong to more than one class.

2.3.2 K-means & c-means clustering

The k-means clustering algorithm is an unsupervised classification in which an initial k number of classes is set and k centroids are defined for each class in the data (MacQueen, 1967). With each iteration, each object in space is assigned to the closest centroid, and the positions of the k centroids are recalculated until they no longer move. The algorithm aims to minimize a squared error function that characterizes within-cluster variability:

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (2.5)$$

where $\|x_i^{(j)} - c_j\|$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster center c_j is an indicator of the distance of the n data points from their respective cluster centers (MacQueen, 1967). The algorithm minimizes within-cluster variability and maximizes diagonal distance between the mean cluster values of each data layer. The original hard k-means-clustering algorithm has since been adapted to include a fuzziness component where points may belong partially in multiple classes (Gruijter and McBratney, 1988;

Boydell and McBratney; 2002; Lark and Stafford, 1997). The free software FuzME (v.3.0, Australian Centre for Precision Agriculture, Sydney, Australia) applies the fuzzy k-means algorithm and retains the option for hard k-means clustering. This algorithm is commonly used in MZ delineation (Taylor *et al.*, 2002; Vrindts *et al.*, 2005; Davatgar *et al.*, 2012).

Another popular classification software, Management Zone Analyst (MZA) (v.1, US Department of Agriculture, Washington, D.C., USA) employs unsupervised classification with the fuzzy c-means algorithm (Fridgen *et al.*, 2004). Fuzzy c-means was developed by Dunn (1973) and fine-tuned by Bezdek (1984) and is essentially identical to fuzzy k-means. The algorithm seeks to minimize the sum of squared distances from all data points in a cluster to the cluster center:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2, 1 \leq m < \infty \quad (2.6)$$

where m is any real number greater than 1, u_{ij} is the degree of membership of x_i in the cluster j , x_i is the i^{th} of d-dimensional measured data, c_{ij} is the d-dimension center of the cluster, and $\|*\|$ is any norm expression the similarity between any measured data and the center (Bezdek, 1984).

Like k-means, the algorithm updates cluster centers c_j with every iteration until a termination criterion between 0 and 1 is reached. The addition of the vector u_{ij}^m distinguishes fuzzy c-means from hard k-means because it provides a weighting to the point's association with a cluster. MZA has been used in many precision agriculture studies (Moral *et al.* 2010; De Caires *et al.* 2015; Zhang *et al.*, 2010). The software includes summaries of the fuzziness performance index (Odeh *et al.*, 1992) and normalized classification entropy (Bezdek, 1984) which allows the user to decide how many classes they consider optimal.

2.3.3 Hierarchical classification

Hierarchical classification is less commonly used in MZ delineation. More often an agglomerative hierarchical classification is used, where two clusters are merged where they have the smallest distance between two points in a space. Ward's method of hierarchical classification merges clusters when the smallest increase in the combined error sum of squares

from a one-way univariate ANOVA for each variable is minimized. It has been used to separate yield performance classes (Farooque *et al.* 2012). Fleming *et al.* (2000) used this technique to separate EC_a data into classes prior to interpolation.

2.3.4 Neighborhood Search Analyst

An alternative method of unsupervised classification which accounts for spatial structure is the Neighborhood search analyst (NSA). NSA groups adjacent cells, ranking classes according to their strength of similarity. The algorithm calculates the field mean square error (MSE), then iteratively clusters adjacent cells and re-calculates the field MSE (Dhawale *et al.*, 2014). If the new MSE is less than the MSE of the entire field, more neighboring cells are added until adding new cells no longer results in a lower MSE. Then a new cluster search is initiated until the criteria is met again. When clustering no longer reduces the MSE of the entire field, the algorithm ends. NSA differs from the more commonly used MZA in that cells are not necessarily assigned to a class if they do not meet a certain degree of similarity. Instead classes are ranked in NSA by how distinct they are, producing more site-specific regions of similarity.

2.4 Validation of Classes or Management Zones

MZs are finalized by validating the significant difference between classes. Attributes tested for significant difference include soil chemical attributes, texture, and yield (De Caires, 2015; Urretavizcaya, 2015). Yield and yield limiting variables can be tested with a two-way analysis of variance (ANOVA) to determine if the classes are significantly different (Urretavizcaya *et al.*, 2015; Farooque *et al.*, 2012). Tukey's post-hoc means comparison test uses ANOVA results to compare zones and individual variables. The Tukey test is useful when a large set of variables are being compared.

The Pearson correlation coefficient measures linear correlation between two variables. Pearson's correlation is ubiquitously used in data analysis, but it is limited because it evaluates linear relationships. Prior to correlation analysis, a box-cox transformation may be applied to normalize data so that the linear model is applicable (Box, 1964). The box-cox transformation is preferable to a log transformation because it allows for more subtle changes in the data distribution. Pearson's correlation is used to evaluate the correlation

among soil parameters and yield in many studies as a preliminary step in MZ delineation (e.g. Moral, 2010; Farooque, 2012).

2.5 Implementation with Variable Rate Technologies

Site-specific management is implemented with variable rate (VR) applicators. The automatic section control feature of many VR technologies allows spreaders, planters and sprayers to automatically turn on, or off, based on their location in the field. When properly calibrated, these technologies allow for greater precision of application and can accommodate irregularly shaped areas. With the classic approach to MZ delineation, VR applicators apply uniform nutrient levels in each delineated zone. However, VR applicators are capable of applying a more continuously variable treatment through the field using isolines. A regression-based approach to SSM would ultimately produce a surface prescription map in which treatment applied is proportional to changes in field condition. The variations in this output surface can then be converted to isolines and input into a VR applicator for more empirical and precise treatment.

2.6 Persisting Challenges in Site-specific Management

While SSM and MZ delineation have been developed and thoroughly researched for decades, debate remains over the technical feasibility of SSM (Tisseyre and McBratney, 2008). Apart from the cost of VR equipment, sampling and analysis can be time consuming and expensive. The expectation is that savings attributed to reduced inputs and/or increased yields will make up for the costs of implementing SSM. A number of researchers have attempted to address financial concerns by developing low-cost, simple methodologies for SSM (Taylor *et al.*, 2007) in an effort to make precision agriculture practices more accessible and wide-spread.

Accessibility of precision agriculture practices is an important step to noticeable and lasting environmental mitigation. While individual farms may benefit economically from SSM, the environmental benefits are negligible unless SSM is adopted on a wide scale. To start, McBratney *et al.* (2005) suggest studies of SSM precision agriculture be applied to whole farms rather than single fields. Additionally, more studies should include the assess-

ment of environmental indicators before and after site-specific treatment in order to quantify environmental effects. Furthermore, research should focus on developing SSM strategies that are replicable and scalable for greater widespread use, should the strategies be found to be environmentally beneficial. McBratney *et al.* (2005) recognize Precision Agriculture (PA) strategies are tailored to individual farms but the development of decision support systems (DSS) and adaptable SSM strategies would assist farmers in adopting PA practices.

Finally, recent debate questions the efficacy of subdividing fields into zones with categorical attributes. Fuzzy classification recognizes that attributes can belong in more than one class, but MZ delineation still subdivides fields into smaller uniform application sites. The NSA clustering algorithm by Dhawle *et al.* (2014) suggests that entire fields need not be subdivided into MZs to identify regions of distinction. Data separated to maximize contrast of field conditions is more desirable when enacting regression-based site-specific management. With this in mind, this study proposes a new regression-based approach to separating topographical and EC_a data to be used to develop a continuous prescription map proportional to the changes in field condition. This regression-based prescription based on field conditions is not only more empirical, it is also less computationally intensive and has the potential to be streamlined into a DSS for producers.

Chapter 3

Materials & Methods

3.1 Site Description

For the purposes of this study, two experimental fields were selected (Fig. 3.1). The experimental blueberry fields lie 6 km southwest of Normandin, QC (48.8369° N, 72.5279° W), north of the Chamouchouane River. Soil here is primarily podzolic, mixed with finer eolian deposits. Field 140b (11.3 ha) represents a uniform low-lying topography ranging from 123–125 m and Field 21 (13.2 ha) represents a more heterogeneous topography with elevation ranging from 127–136 m.

3.2 Historic Yield and Management Practices

During the autumn of the crop year, the study sites were mowed in alternating strips so that un-mowed strips would retain snow on the field. In the sprout year, the blueberry plants were pruned in springtime, and fertilizer was applied uniformly. Herbicides were applied uniformly as well as in spot treatments where weed presence was visually assessed. Insecticides were applied to control for *Altica sylvia* mallcoch. No pesticides or fertilizers were applied in the harvest year. Nitrogen (N) applications varied season to season as illustrated in Tables 3.1 & 3.2. Leaf analyses were performed in the growth year. Phosphorus (P) and boron (B) were applied when leaf analyses indicated nutrient deficiency.

Yield in 2016 was a record high for wild blueberry producers in Lac-Saint-Jean. Despite

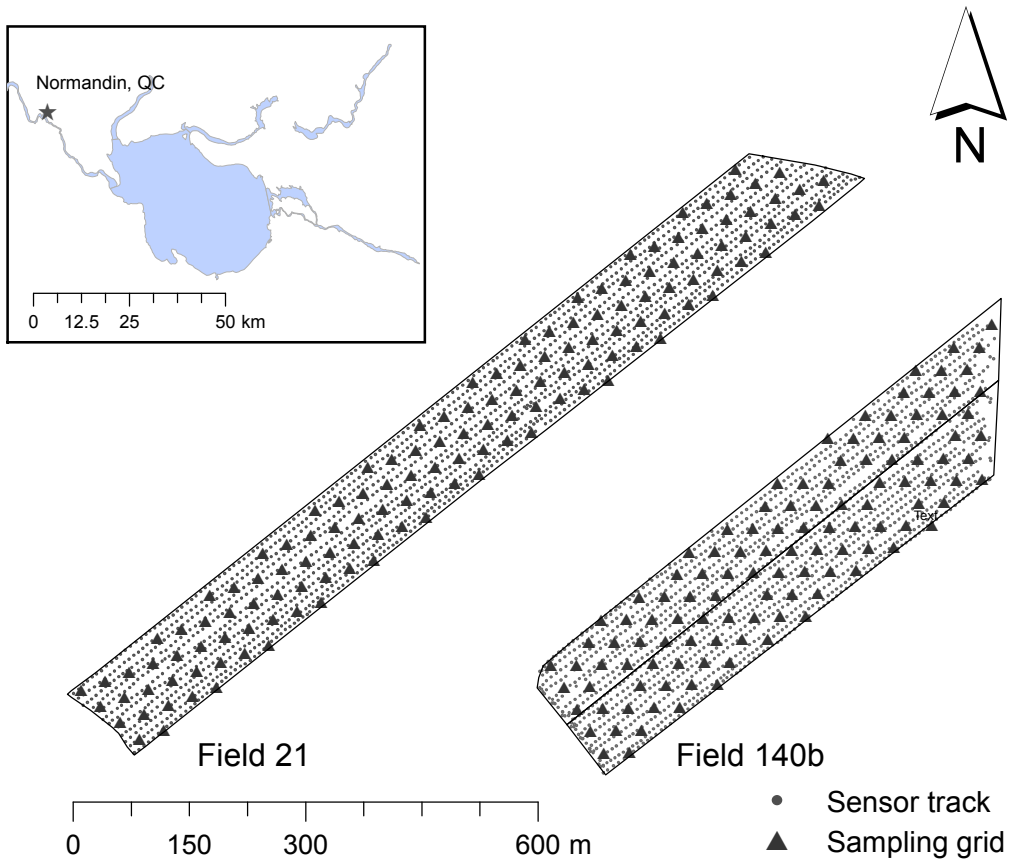


Fig. 3.1 Experimental fields 21 and 140b with soil/yield sampling locations and elevation/EC_a sensor track.

Table 3.1 Field 21 summary of yield and treatments 2009-2016.

Year	Average Yield (kg ha ⁻¹)	N (kg ha ⁻¹)	P (kg ha ⁻¹)	K (kg ha ⁻¹)	B (kg ha ⁻¹)	Herbicide	Herb. Rate (kg ha ⁻¹)
2016	6,398	—	—	—	—	—	—
2015	—	30.2	20.1	20.1	0	Velpar ^x	2
2014	4,530	—	—	—	—	—	—
2013	—	30.5	0	0	0.7	Velpar	2
2012	3,313	—	—	—	—	—	—
2011	—	26.3	21	15.8	0	Velpar	2
2010	1,337	—	—	—	—	—	—
2009	—	NA	NA	NA	NA	Pronone ^y	2

NB: Fertilizer is applied as a mix of N-P-K ratio. Boron is added when leaf analyses show it to be deficient.

^x Active ingredient: Hexazinone (DuPontTM)

^y Active ingredient: Hexazinone (DuPontTM)

Table 3.2 Field 140b summary of yield and treatments 2009-2016.

Year	Average Yield (kg ha ⁻¹)	N (kg ha ⁻¹)	P (kg ha ⁻¹)	K (kg ha ⁻¹)	B (kg ha ⁻¹)	Herbicide	Herb. Rate (kg ha ⁻¹)
2016	3,788	—	—	—	—	—	—
2015 [*]	—	NA	NA	NA	NA	—	—
2014 [*]	4,250	NA	NA	NA	NA	—	—
2013	—	30.5	0	0	0.7	Velpar ^x	2
2012	2,052	—	—	—	—	—	—
2011	2,931	—	—	—	—	—	—
2010	—	14.5	11.7	8.6	0.4	Pronone ^y	16
2009	4,020	—	—	—	—	—	—

NB: Fertilizer is applied as a mix of N-P-K ratio. Boron is added when leaf analyses show it to be deficient.

^{*}Penergetic was an experimental fertilizer/herbicide mix which resulted in 75% loss of crop.

^x Active ingredient: Hexazinone (DuPontTM)

^y Active ingredient: Hexazinone (DuPontTM)

its more heterogeneous topography, Field 21 has historically been more productive than Field 140b. In circumstances of low yield, the producer will extend the cycle to a third year, as the producers did in Field 140b in 2011 and 2012 (Table 3.2). Producers tested Penergetic, a commercial product, as a replacement to fertilizer and herbicide and lost 75% of crop in the 2014-2015 cycle.

3.3 Description of Data Layers

Data layers can be divided into proximally and remotely sensed data, soil samples analyzed in the lab, and sampled yield (Fig. 3.2). The selected data layers are meant to encompass the various properties and processes which affect yield. Figure 3.2 summarizes the interrelationships between crop yield, sensor data, and chemical and granulometric data.

Soil and yield samples were obtained in both fields with a 33 x 33 m (0.1-ha) grid sampling scheme for a total of 136 points in Field 21 and 116 points in Field 140b. Yield samples were collected on August 8-9, 2016 before the fields were harvested. Blueberries were combed from a square meter of blueberry bush at each point, and the weight of the fresh blueberries was measured and recorded on site for every sample. Satellite imagery was acquired on August 11, 2016. EC_a was sampled the last week of September 2016, just after mowing. Soil samples were collected one week later in the beginning of October 2016 at two depths (0-5 cm and 5-15 cm).

3.3.1 Laboratory analysis of chemical and granulometric soil properties

Soil samples were dried and ground to 2 mm for textural and chemical laboratory analysis. All soil samples were weighed. Both depths were analyzed for nutrient content. A Mehlich-III soil extractant was used to extract nutrients (Ziadi and Tran 2007). P content was determined by colorimetry (Lachat Instruments, model 8500, series 2, Loveland, USA) (Murphy and Riley 1962). Potassium (K) content was determined with spectrophotometry flame emission (Isaac and Kerber, 1971). Calcium (Ca) and Magnesium (Mg) were determined with atomic absorption spectrophotometry (Agilent Technologies, model 200, series AA, Santa Clara, USA). Total Carbon (C) and Nitrogen (N) content were evaluated with the Elemental vario MAX CN analyzer (Elemental Analysensysteme, GmbH, Hanau,

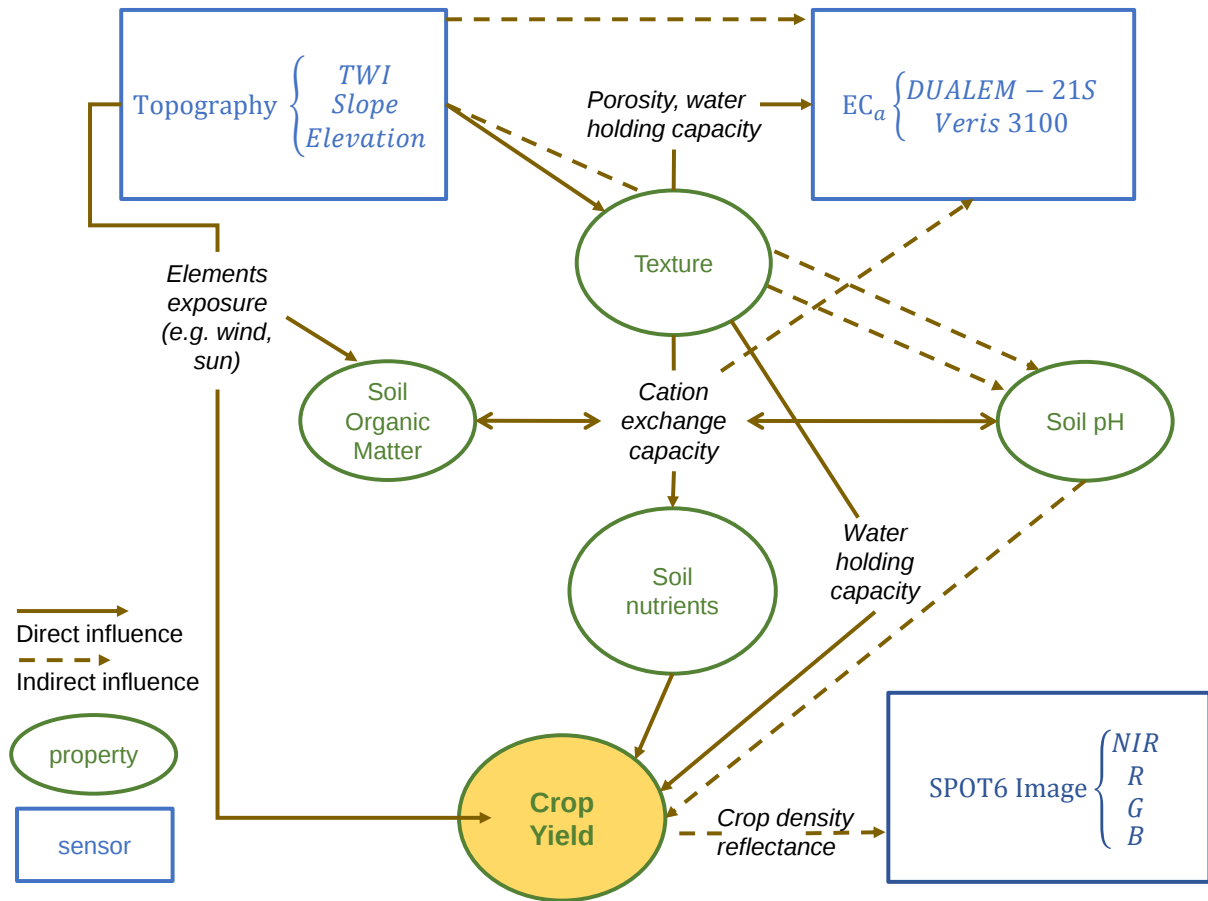


Fig. 3.2 Depiction of data layers and their relationships. TWI is the topographic wetness index derived from elevation and slope data. EC_a is the soil apparent electrical conductivity captured by two sensors, DUALEM and Veris. Multispectral SPOT6 is the multispectral satellite image, comprising four wavelength bands: R (red), G (green), B (blue), and NIR (near infrared). Sample and sensor data encompass yield-determining properties such as cation exchange capacity, water holding capacity, soil porosity, crop density, and element exposure. Soil pH indirectly influences yield by limiting weeds.

Germany), and soil organic matter (S.O.M.) was calculated from the total C percentage.

Soil texture was analyzed for all soil samples at the 5-15 cm depth using the pipette method (Day, 1965). Texture was categorized in terms of g per kg of very coarse sand, coarse sand, medium sand, fine sand, very fine sand, total silt, and total clay according to the Canada Soil Survey Committee standards (Sheldrick, 1984). Descriptive statistics on all attribute data were summarized.

3.3.2 Apparent electrical conductivity sampling

EC_a was collected with two sensors, the DUALEM-21S (Dualem Inc., Milton, ON) and the Veris 3100 (Veris Technologies, Inc., Salina, KS). Measurements were taken at 1 Hz. The depth of investigation of EC_a measurements depends on the configuration of the transmitter and receiver coils. The DUALEM-21S has one transmitter coil and four receiving coils to capture four depths. Coils arranged in the horizontal co-planar (HCP) receive lower depths than the perpendicular co-planar (PRP). Additionally, coil spacing affects the depths received. The DUALEM-21S configuration has two PRP coils at 1.1 m and 2.1 m from the transmitter, and two HCP coils 1 m and 2 m from the transmitter (Fig. 3.3). Receiver coils closer to the transmitter have a shallower depth of investigation. Depth is determined at 70% of the cumulative response in a column of heterogeneous soil (Table 3.3).

Transects were spaced about 10-m apart, guided with a GPS steering guidance system. The DUALEM-21S was run for twenty minutes before being calibrated to reduce the possibility of drift in sensor data. It was pulled on a sled by a John Deere Gator at a relatively constant speed to maximize contact with the ground. At the end of sampling, the sensor was passed over previous transects so that data could be reviewed for evidence of drift.

The Veris 3100 is a galvanic contact resistivity sensor and derives conductivity from its inverse relationship with electrical resistivity. The Veris 3100 is configured with six rolling coulter electrodes (Lund *et al.*, 1999). Electrical current flows through the second and fifth coulters. The voltage drop is measured between the third and fourth coulters and first and sixth coulters (Sudduth *et al.*, 2003). The electrodes are equally spaced in a Wenner array so that resistance is measured at two depths (Fig. 3.4).

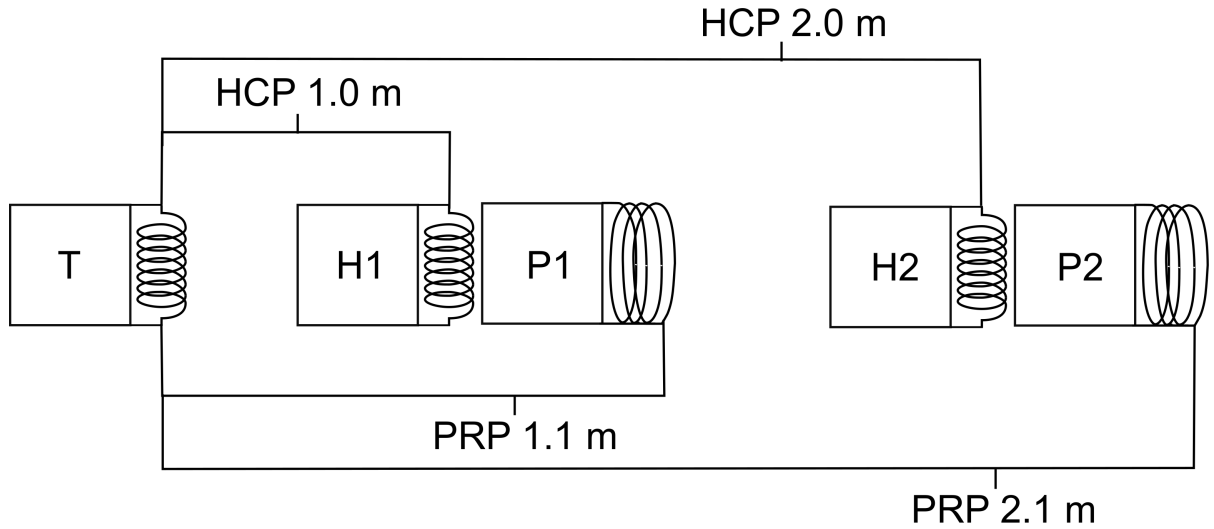


Fig. 3.3 Dipole arrangement of DUALEM EMI sensor. T is the transmitter coil. H1 and H2 receivers receive the induced current in horizontal co-planar (HCP) arrays at 1-m and 2-m respectively. P1 and P2 receiver receive the current in perpendicular co-planar (PCP) arrays at 1.1-m and 2.1-m respectively. Receivers at a further distance from the transmitter coil capture EC_a at a greater depth.

The depth of investigation is related to about one-third the outer electrode spacing (Rhoades, 1993) and summarized in Table 3.3. Sensor transects were 3-m apart, and measurements were taken at a density of 1 sample/s. The sensor was pulled by an SUV at a relatively constant speed (Fig. 3.5).

Table 3.3 Derived depths of EC_a instruments where cumulative response is 70% (Mat Su, 2016).

Measurement	Effective sensing depth (m)
Veris Shallow	0.30
Veris Deep	0.90
DUALEM PRP 1.1	0.54
DUALEM PRP 2.1	1.03
DUALEM HCP 1.0	1.55
DUALEM HCP 2.0	3.18

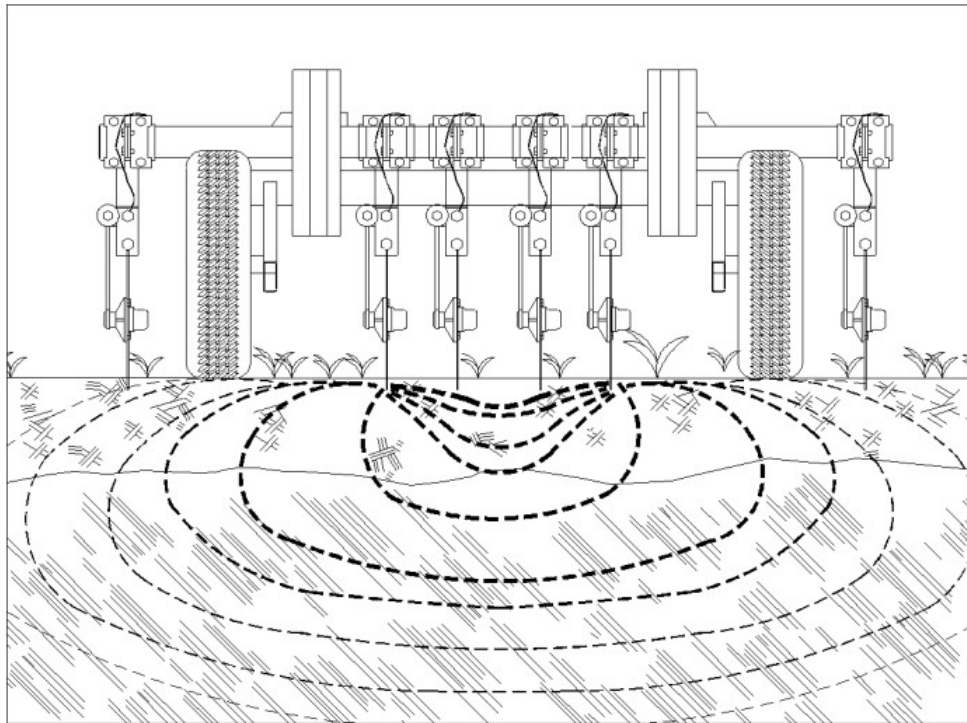


Fig. 3.4 Veris 3100 configuration of six roller coulters, one pair which passes a current through the soil and two pairs which measure resistance to extract EC_a at two depths (Oguri *et al.*, 2009).



Fig. 3.5 Veris 3100 set up. Photo credit to Agriculture and Agrifood Canada.

3.3.3 Satellite imagery

A multi-spectral satellite image of the fields was obtained from the Airbus SPOT-6 satellite archive (Airbus Defense and Space, Ottobrunn, Germany), for the month of August 2016 - shortly after yield sampling occurred and a month before soil sampling. The SPOT-6 satellite measures reflectance in five wavelength ranges referred to as bands (Table 3.4). The SPOT-6 images were delivered georeferenced, corrected for off-nadir acquisition and terrain effects using the standard Reference3-D model for ground corrections (Astrium Services, 2013). The panchromatic and multispectral images were simultaneously acquired, allowing for geospatially accurate pan-sharpening of the satellite images to 1.5-m resolution.

The pansharpened image was radiometrically and atmospherically corrected in ENVI image analysis software (Exelis, Inc., Boulder, CO). Several broadband greenness vegetation indices (VIs), which detect vegetation and vigor were calculated from the multispectral bands using the band math function (Table 3.5). Many of the VIs compare changes in the near-infrared and red bands (e.g., NDVI, DVI, TDVI, MSR, RDVI). NDVI is a widely used VI because it is resistant to topography changes and sensitive to biomass, while TDVI is

Table 3.4 SPOT Multi-spectral and Panchromatic Image wavelength bands.

Band	Wavelength Range (nm)
Pan	450 – 745
Blue	450 – 520
Green	530 – 590
Red	625 – 695
Near-infrared	760 – 890

less saturated than NDVI at close canopy and may detect subtler variations. Other VIs utilize ratios between the near-infrared and green bands (e.g., GARI, GDVI, GRVI), in order to detect a greater range of chlorophyll concentrations. A number of indices which account for the soil line were tested (e.g., MSAVI2, OSAVI, MNLI) in order to reduce noise attributed to soil reflectance. The leaf area index (LAI) was also tested to characterize canopy cover.

In addition to a number of ratio-based indices, Principal Components Analysis (PCA) was performed in ENVI to transform the band space and reduce the multispectral image to two or three principal components which maximize variation and reduce noise. The second principal component is often recommended as a seasonal VI (Eklundh and Singh, 1993; Townshend, 1985). Once all VIs were calculated, sample points were imported into ENVI as regions of interest so that the values of the VIs at the sample points could be extracted and compared statistically to other measured attributes at the same location. Pearson's correlation coefficient was calculated for VIs and yield to assess how well VIs correlated with sampled yield. The best performing VI was classified by the Jenks optimization method to delineate bare patches within the field (Jenks, 1967). The effectiveness of bare patch prediction was assessed by calculating Pearson's correlation with a binary classification of sampled yield where yield values of 0 kg ha⁻¹ were assigned a 0 and all other values were assigned a 1.

Table 3.5 Summary of ratio-based vegetation indices (VIs) calculated from SPOT6 satellite image. NIR is the near-infrared wavelength range, R is the red wavelength range, G is the green wavelength range, and B is the blue wavelength range. γ is a weighting function for aerosol conditions in the atmosphere. L the canopy background adjustment factor.

Name	Formula	Reference
Difference Vegetation Index (DVI)	$= NIR - R$	Tucker (1979)
Enhanced Vegetation Index (EVI)	$= 2.5 * \frac{(NIR-R)}{NIR+6*R-7.5*B+1}$	Huete <i>et al.</i> (2002)
Green Atmospherically Resistant Index (GARI)	$= \frac{NIR-[G-\gamma(B-R)]}{NIR+[G-\gamma(B-R)]}, \gamma = 1.7$	Gitelson, Kaufman, and Merzylak (1996)
Green Difference Vegetation Index (GDVI)	$= NIR - G$	Gitelson and Merzylak (1998)
Green Ratio Vegetation Index (GRVI)	$= \frac{NIR}{G}$	Sripada <i>et al.</i> (2006)
Leaf Area Index (LAI)	$= (3.618 * EVI - 0.118)$	Boegh <i>et al.</i> (2002)
Optimized Soil Adjusted Vegetation Index (OSAVI)	$= \frac{(NIR-R)}{(NIR+R+0.16)}$	Rondeaux, Steven, and Baret (1996)
Modified Non-Linear Index (MNLI)	$= \frac{(NIR^2-R)*(1+L)}{NIR^2+R+L}, L = 0.5$	Yang, Willis, and Mueller (2008)
Modified Simple Ratio (MSR)	$= \frac{(NIR/R)-1}{\sqrt{NIR/R+1}}$	Chen (1996)
Normalized Difference Vegetation Index (NDVI)	$= \frac{NIR-R}{NIR+R}$	Rouse <i>et al.</i> (1973)
Renormalized Difference Vegetation Index (RDVI)	$= \frac{NIR-R}{\sqrt{NIR+R}}$	Roujean and Breon (1995)
Transformed Difference Vegetation Index (TDVI)	$= \sqrt{0.5 + \frac{NIR-R}{NIR+R}}$	Bannari, Asalhi, and Teillet (2002)
Modified Soil Adjusted Vegetation Index (MSAVI2)	$= \frac{2*NIR+1-\sqrt{(2*NIR+1)^2-8*(NIR-R)}}{2}$	Qi <i>et al.</i> (1994b)

3.4 Map Interpolation

3.4.1 Pre-processing

Elevation and EC_a data were averaged to 1 value per 5 s. Data distribution was examined for normality and values outside of two standard deviations were removed. Slope and TWI were calculated from the elevation data using SAGA GIS (v.6.3, System for Automated Geoscientific Analyses, Hamburg, Germany). TWI models field water storage based on slope and catchment area modeling (Beven and Kirkby, 1979).

3.4.2 Spatial prediction

Sampled chemical, granulometric, EC_a , and elevation data were interpolated to three-dimensional, continuous surfaces using the Ordinary kriging (OK) method in R statistical software (R Foundation for Statistical Computing, Vienna, Austria). Predicted surfaces were cross validated with the original samples to assess strength of the prediction. Geo-statistics were calculated on all data layers. An R script utilizing the ‘gstat’ package (Pebesma and Graeler, 2017) standardized the process of interpolation as follows:

1. The data was automatically fit with a box-cox transformation (Box, 1964). The ‘forecast’ package in R (Hyndman *et al.*, 2017) applies Guerrero’s (1993) method to assign a transformation value which minimizes the coefficient of variation for the dataset.
2. A theoretical variogram model was fitted to the experimental variogram of the box-cox-transformed data. Possible models were gaussian, spherical, exponential, or pure nugget. The fitting method used to fit the theoretical variogram to sampled data was an iterative reweighted least squares estimation which prioritized variances of point pairs at shorter lag distances in the experimental variogram (Pebesma and Graeler, 2017). Range, nugget, partial sill, and total sill were calculated for the best fit model.
3. The nugget to sill ratio was calculated from the theoretical nugget and partial sill variances to determine the degree of spatial structure (Camberdella *et al.*, 1994). A smaller nugget to sill ratio indicates stronger autocorrelation in the dataset. A ratio of 100% is indicative of a pure nugget and considered a random spatial structure with no autocorrelation.

4. The fitted variogram was used in the OK method with the box-cox transformed data.
5. Predicted maps were cross validated with the leave-one-out (LOO) or k-fold method. LOO is an iterative process where the value of a point is predicted using all of the dataset with the exception of that point. This was repeated for every point in the set so that the error could be estimated. The k-fold method divided the original data set into k parts to determine error. The k-fold method was used with dense EC_a and elevation data, where $k=10$.
6. The root mean square prediction error was calculated:

$$RMSE = \sqrt{\frac{1}{l} \sum_{j=1}^l [\hat{z}(x, y) - z^*(x, y)]^2} \quad (3.1)$$

where l is the number of validation points, $\hat{z}(x, y)$ is the estimate value at location (x, y) , and $z^*(x, y)$ is the actual observation at location (x, y) . $RMSE$ quantifies the accuracy of a prediction map. To compare $RMSE$ among several variables, it is divided by the total variation s_z :

$$RMSE_{SD} = \frac{RMSE}{s_z} \quad (3.2)$$

A standardized $RMSE_{SD}$ value around 40% is considered satisfactorily accurate. A value greater than 71% signifies that less than 50% of variability of the validation points were represented in the model, so a majority of points are inaccurate (Hengl, 2009, p. 25).

7. The kriging output grid was back-transformed from the box-cox transformation to initial values so that predicted vs. sampled values could be plotted and compared. A cross validation plot was generated with a best fit line using the back-transformed values. Correlation coefficients were calculated for both the transformed and the back-transformed data.

3.5 Correlation Analysis & PLS

Pearson's correlation matrix was calculated with all box-cox transformed soil properties, EC_a layers, and yield. The sample Pearson correlation coefficient (r) measures linear correlation between two sampled variables. It is defined as:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3.3)$$

where n is the sample size, x_i and y_i are individual sample points indexed in i , and $\bar{x}$ and $\bar{y}$ are the sample means. Significance of the correlation was tested using,

$$t = \frac{r}{\sqrt{\frac{1-r^2}{N-2}}} \quad (3.4)$$

where N is sample size, r is the correlation coefficient, and t is the distribution of two perfectly correlated variables given the population.

Data was also analyzed with partial least squares (PLS) regression with the PROC PLS function in SAS Statistical software (SAS Institute Inc., Cary, North Carolina, USA) to determine sources and degree of variation in predictor variables EC_a, topography, and VI and dependent soil attributes. Five PLS models were analyzed in each field, separating dependent variables by chemical attributes (0-5 cm) and (5-15 cm), overall texture, sand texture, and yield. Wold's criterion (1994) was used to determine the extent of contribution a variable makes to the PLS model.

3.6 Classification & Data Separation

Data layers were initially classified using fuzzy c-means in the software Management Zone Analyst (MZA)(v.1, US Department of Agriculture, Washington, D.C., USA). A second data separation methodology for regression-based site-specific management was applied which equally weights elevation and Shallow EC_a data. Veris EC_a shallow values and elevation values were extracted from their raster grids to the sample points so they could be compared with other sample attributes. Elevation vs. Shallow EC_a values were then projected onto a scatter plot, and ten points in the four corners of the scatter plot were

sub-set to represent four extreme growing conditions of $EC_{Low}Elev_{Low}$, $EC_{Low}Elev_{High}$, $EC_{High}Elev_{Low}$, $EC_{High}Elev_{High}$. Shallow EC_a was selected because its depth of response (30 cm) most closely corresponds with the depth of soil samples. A Two-way Analysis of Variance (ANOVA) was calculated to compare the four scenarios for significant differences. Additionally, Tukey's post-hoc test was used to compare the significant difference of the means of individual soil properties in the four scenarios.

Chapter 4

Results & Discussion

4.1 Laboratory Analyses

Summary statistics of the laboratory analyses are presented in Tables 4.1 & 4.2. Results indicate considerable variability in soil properties, except for total sand ($CV_{21}=8.6\%$, $CV_{140b}=3.4\%$) and pH ($CV_{21}=6.0\%$, $CV_{140b}=3.5\%$). While total sand is not highly variable, sand grain size (e.g. very fine vs. very coarse) does vary considerably in both fields.

The high variability among soil properties and yield indicate both sites could benefit from site-specific management. One property that stands out is soil pH (Tables 4.1 and 4.2). Soil pH is important in wild blueberry for controlling weed growth, and the optimal range of pH is between 4.6 and 5.2 (NBDAAF, 1998). The average pH in both fields is within the acceptable range, but the maximum pH exceeds the optimal range. For example, the thematic pH maps of Field 21 (Fig. 4.3) show a concentrated area of the field where pH is 5.5 - 5.6, above optimal range. This same area coincides with finer texture soil. Excess N application here has a greater risk of volatilization. This is one example of how tailored nutrient application could theoretically benefit the field. N may be limited in this area to reduce the risk of volatilization, and/or site-specific sulfur application may be considered to lower pH.

Similar recommendations for nutrient prescription of N, P, or K cannot be drawn from the soil sample data because crop nutrient levels for wild blueberry are presently determined with leaf analysis rather than soil analysis. Soil analyses were used for this project instead

Table 4.1 Field 21 summary statistics of chemical and granulometric soil attributes; n is the sample size, STD is standard deviation, and CV% is coefficient of variation.

	Unit	n	Min	Max	Mean	STD	CV%
Soil Particle size (5-15 cm)							
Clay	g kg ⁻¹	136	12.0	37.3	23.5	5.2	22.1
Silt	g kg ⁻¹	136	35.4	345.8	119.7	75.6	63.1
Sand	g kg ⁻¹	136	636.4	948.1	856.8	74.1	8.6
Very coarse sand	g kg ⁻¹	136	0.0	76.6	12.0	13.7	113.7
Coarse sand	g kg ⁻¹	136	0.9	333.7	99.9	88.8	88.9
Medium sand	g kg ⁻¹	136	3.7	570.1	284.8	163.2	57.3
Fine sand	g kg ⁻¹	136	105.2	679.8	312.2	123.2	39.5
Very fine sand	g kg ⁻¹	136	25.7	509.0	147.8	130.3	88.1
Chemical analysis (0-5 cm)							
S.O.M.	g kg ⁻¹	136	17.9	489.6	169.6	99.5	58.7
Total N	%	136	0.06	1.50	0.46	0.27	59.1
Soil pH_{water}	–	136	3.8	6.9	4.7	0.5	10.6
P	mg kg ⁻¹	136	5.5	265.2	63.4	54.0	85.1
K	mg kg ⁻¹	136	13.6	388.8	107.3	70.1	65.3
Total C	%	136	1.2	32.0	11.1	6.5	58.7
Ca	mg kg ⁻¹	136	104.9	564.2	361.1	76.4	21.2
Mg	mg kg ⁻¹	136	5.3	410.4	107.3	71.9	67.0
Al	mg kg ⁻¹	136	489.5	2,104	889.0	287.0	32.3
Fe	mg kg ⁻¹	136	60.0	5,370	1,502	932.6	62.1
Chemical analysis (5-15 cm)							
S.O.M.	g kg ⁻¹	136	9.7	59.4	19.7	8.1	40.9
Total N	%	136	0.04	0.17	0.07	0.02	35.5
Soil pH_{water}	–	136	4.5	6.5	5.1	0.3	6.0
P	mg kg ⁻¹	136	1.1	249.1	67.0	48.5	72.3
K	mg kg ⁻¹	136	8.1	256.8	38.7	25.0	64.4
Total C	%	136	0.64	3.88	1.29	0.53	40.93
Ca	mg kg ⁻¹	136	106.8	644.3	295.3	103.6	35.1
Mg	mg kg ⁻¹	136	2.7	118.2	7.9	10.0	127.8
Al	mg kg ⁻¹	136	749.9	2238	1653	260.3	15.7
Fe	mg kg ⁻¹	136	19.3	1393	148.2	155.9	105.2

Table 4.2 Field 140b summary statistics of chemical and granulometric soil attributes; n is the sample size, STD is standard deviation, and CV% is coefficient of variation.

	Unit	n	Min	Max	Mean	STD	CV%
Soil Particle size (5-15 cm)							
Clay	g kg ⁻¹	116	9.9	38.1	26.5	6.1	23.1
Silt	g kg ⁻¹	116	19.2	257.6	77.5	30.5	39.3
Sand	g kg ⁻¹	116	718.9	968.4	896.0	30.2	3.4
Very coarse sand	g kg ⁻¹	116	1.4	73.1	25.4	15.3	60.0
Coarse sand	g kg ⁻¹	116	12.0	335.3	170.1	88.9	52.3
Medium sand	g kg ⁻¹	116	81.6	552.5	356.8	103.1	28.9
Fine sand	g kg ⁻¹	116	99.8	633.7	280.3	126.2	45.0
Very fine sand	g kg ⁻¹	116	17.0	271.9	63.3	49.3	77.9
Chemical analysis (0-5 cm)							
S.O.M.	g kg ⁻¹	116	17.3	391.3	135.0	77.6	57.5
Total N	%	116	0.08	1.19	0.44	0.25	56.7
Soil pH_{water}	–	116	3.7	5.6	4.5	0.4	7.8
P	mg kg ⁻¹	116	3.0	411.6	38.7	48.0	124.2
K	mg kg ⁻¹	116	16.1	290.3	92.9	56.1	60.4
Total C	%	116	1.1	25.6	8.8	5.1	57.5
Ca	mg kg ⁻¹	116	107.2	691.5	386.8	99.1	25.6
Mg	mg kg ⁻¹	116	6.7	256.3	77.5	53.2	68.7
Al	mg kg ⁻¹	116	477.3	2,655	939.1	293.9	31.3
Fe	mg kg ⁻¹	116	3.5	1,805	465.4	338.2	72.7
Chemical analysis (5-15 cm)							
S.O.M.	g kg ⁻¹	116	3.0	63.0	17.8	8.7	48.8
Total N	%	116	0.02	0.22	0.08	0.03	35.0
Soil pH_{water}	–	116	4.6	5.8	5.0	0.2	3.5
P	mg kg ⁻¹	116	1.1	134.3	24.3	21.4	88.0
K	mg kg ⁻¹	116	3.4	95.2	40.4	18.4	45.6
Total C	%	116	0.20	4.12	1.17	0.57	48.8
Ca	mg kg ⁻¹	116	41.5	390.5	213.7	79.0	37.0
Mg	mg kg ⁻¹	116	1.7	23.5	6.4	2.9	45.6
Al	mg kg ⁻¹	116	1307	2925	2015	273.4	13.6
Fe	mg kg ⁻¹	116	7.9	1000	205.8	164.5	79.9

of leaf analyses in order to investigate relationships between soil EC_a , soil properties, and yield. Prescribed nutrient level needs will be determined in field trials, detailed further in the concluding chapter.

4.2 Proximal Sensor Data

An initial objective of the project was to compare the results from the DUALEM and Veris sensors. Prior to filtering, the DUALEM data showed more noise, likely due to contact with air and changes in driving speed. However, the datasets were comparable in distribution post-filtering. The coefficient of variation (CV) values reflect roughly normally distributed data among the EC_a measurements except in the Veris Deep layer which was poorly distributed and skewed left in both fields (Tables 4.4 & 4.3). Both sensors satisfactorily captured EC_a variability in the field. Pearson's correlation analysis showed Veris Shallow (0.30 m depth) to be significantly correlated ($p < 0.0001$) with DUALEM PRP1.1 (0.54 m depth) and PRP2.1 in both fields (1.03 m) (Table 4.5).

Both the Veris and DUALEM datasets returned a number of negative EC_a values (Tables 4.3 & 4.4). Negative EC_a values are sometimes considered anomalous and filtered, but the negative values belonged to the normally distributed data. In fact, both PRP 1.1 and Veris Deep means were negative. Overlapping transects collected at the beginning and end of the field sampling were similar, confirming there was no sensor drift in the DUALEM measurements from the beginning to the end of the data collection process. Even after offset, both sensors showed low EC_a values, indicating a very sandy soil. This was validated by the granulometric analysis of sampled soil. Higher EC_a values in HCP1.0 and HCP2.0 may suggest a finer texture parent material under the sandy soil surface. EC_a values are also higher in Veris Deep than Veris Shallow. However, further conclusions of soil stratification cannot be drawn from the data without further three dimensional analyses. For the purposes of this project, the spatial relativity of EC_a values is of interest, so both the Veris depths and all the DUALEM depths were offset by their minimum values to correct for negative readings, and the relative field variation was analyzed.

Table 4.3 Field 21 summary statistics for EC_a data and derived topographic attributes; n is the sample size, STD is the standard deviation, and CV% is the coefficient of variation. Offset values are the original sensor values plus the minimum value collected from all depths. The offset was applied to all layers so that depths could be compared.

	Unit	n	Min	Max	Mean	STD	CV%
HCP1.0	$mS\ m^{-1}$	2005	-1.07	4.49	2.90	0.73	–
HCP1.0 offset	$mS\ m^{-1}$	2005	0.32	5.88	4.29	0.73	17.04
PRP1.1	$mS\ m^{-1}$	2005	-1.39	0.70	-0.05	0.14	–
PRP1.1 offset	$mS\ m^{-1}$	2005	0.00	2.08	1.34	0.14	10.7
HCP2.0	$mS\ m^{-1}$	2005	1.50	3.48	2.46	0.31	–
HCP2.0 offset	$mS\ m^{-1}$	2005	2.88	4.87	3.85	0.31	8.06
PRP2.1	$mS\ m^{-1}$	2005	-0.28	0.80	0.27	0.11	–
PRP2.1 offset	$mS\ m^{-1}$	2005	1.11	2.19	1.66	0.11	6.90
Elevation	m	1749	127.5	136.6	132.2	2.55	1.93
Slope	deg	17493	0.00	25.4	1.91	2.56	134.1
TWI	–	17517	-5.14	13.6	6.45	3.33	51.65
Veris Shallow	$mS\ m^{-1}$	6850	0.14	0.68	0.30	0.08	–
Veris Shallow offset	$mS\ m^{-1}$	6850	3.05	3.59	3.21	0.08	2.45
Veris Deep	$mS\ m^{-1}$	6850	-2.92	3.33	-0.02	0.64	–
Veris Deep offset	$mS\ m^{-1}$	6850	0.00	6.23	2.89	0.64	22.01

Table 4.4 Field 140b summary statistics for EC_a data and derived topographic attributes; n is the sample size, STD is the standard deviation, and CV% is the coefficient of variation. Offset values are the original sensor values plus the minimum value collected from all depths. The offset was applied to all layers so that depths could be compared.

	Unit	n	Min	Max	Mean	STD	CV%
HCP1.0	mS m ⁻¹	1608	1.41	4.62	3.15	0.35	11.09
HCP1.0 offset	mS m ⁻¹	1608	2.52	5.73	4.26	0.35	8.20
PRP1.1	mS m ⁻¹	1608	-1.11	0.84	-0.08	0.11	—
PRP1.1 offset	mS m ⁻¹	1608	0.00	1.95	1.03	0.11	10.54
HCP2.0	mS m ⁻¹	1608	1.10	2.72	1.84	0.22	12.14
HCP2.0 offset	mS m ⁻¹	1608	2.21	3.83	2.95	0.22	7.59
PRP2.1	mS m ⁻¹	1608	-0.34	0.74	0.20	0.11	—
PRP2.1 offset	mS m ⁻¹	1608	0.76	1.85	1.31	0.11	8.61
Elevation	m	1521	123.23	125.47	124.25	0.52	0.42
Slope	deg	26602	0.00	10.22	0.89	1.17	130.43
TWI	—	26638	-9.15	14.08	4.97	2.80	56.45
Veris Shallow	mS m ⁻¹	6181	0.10	0.42	0.26	0.06	23.82
Veris Shallow offset	mS m ⁻¹	6181	2.51	2.83	2.67	0.06	2.34
Veris Deep	mS m ⁻¹	6181	-2.36	2.72	-0.14	0.78	—
Veris Deep offset	mS m ⁻¹	6181	0.05	5.13	2.27	0.78	34.2

4.3 SPOT Imagery and Vegetation Indices

Of the VIs tested, a few were significantly correlated with yield, but no VIs were strongly correlated. In Field 21, the best performing VIs were TDVI ($r=0.29$) and MSR ($r=0.26$). The VI derived from the second principal component (PC2) using Principal Components Analysis outperformed the ratio-based indices. In Field 21, correlation with yield was $r=-0.41$, and in Field 140b, $r=-0.36$, where bare soil represents a higher eigenvalue.

The VI derived from PC2 was used to classify bare patches in the field. In distinguishing bare from vegetation, Field 21 correlation between PC2 and sampled yield was $r=0.68$ and Field 140b was $r=.40$. Based on the classification using PCA, 75.5 m^2 or 8.5% of Field 21 is bare and 29.3 m^2 or 10.7% of Field 140b is bare. Figure 4.1 shows the bare spot classification derived from the PC2 VI. Low correlation coefficients indicate VIs alone cannot capture yield patterns in the field, in part because greener and denser growth do not necessarily indicate higher blueberry yield but may represent weed patches or more leaf growth than fruiting in blueberry bush.

One challenge with the VIs was that especially dry soil has a high reflectance similar to vegetation. Soil adjusted indices like OSAVI and MSAVI2 are meant to correct for this, but a number of bare soil pixels were still indexed at high values with vegetation. Topography will also affect soil reflectance. Bare patches of soil coincide with topographic changes, explaining why some of the bare patches were mis-classified as dense vegetation.

The PC2 correlation values are still not satisfactorily high. Validating the VIs with sampled yield proved challenging because SPOT imagery pixels were pan-sharpened to 1.5 m^2 resolution while ground-truthed yield was sampled at 1 m^2 . This may explain why the PC2 classification did not capture smaller bare spots. Furthermore, sampled yield had high variance due to the sampling process itself which left room for measurement error. Sampling was done by combing blueberry bushes. Blueberries were not always thoroughly removed from the bush or sometimes twigs and leaves would be collected with the blueberries. The clearest issue, however, is that the satellite image was captured 2-3 days after yield sampling was conducted. Areas of the field which were sampled were already disturbed when the image was taken, affecting surrounding biomass.

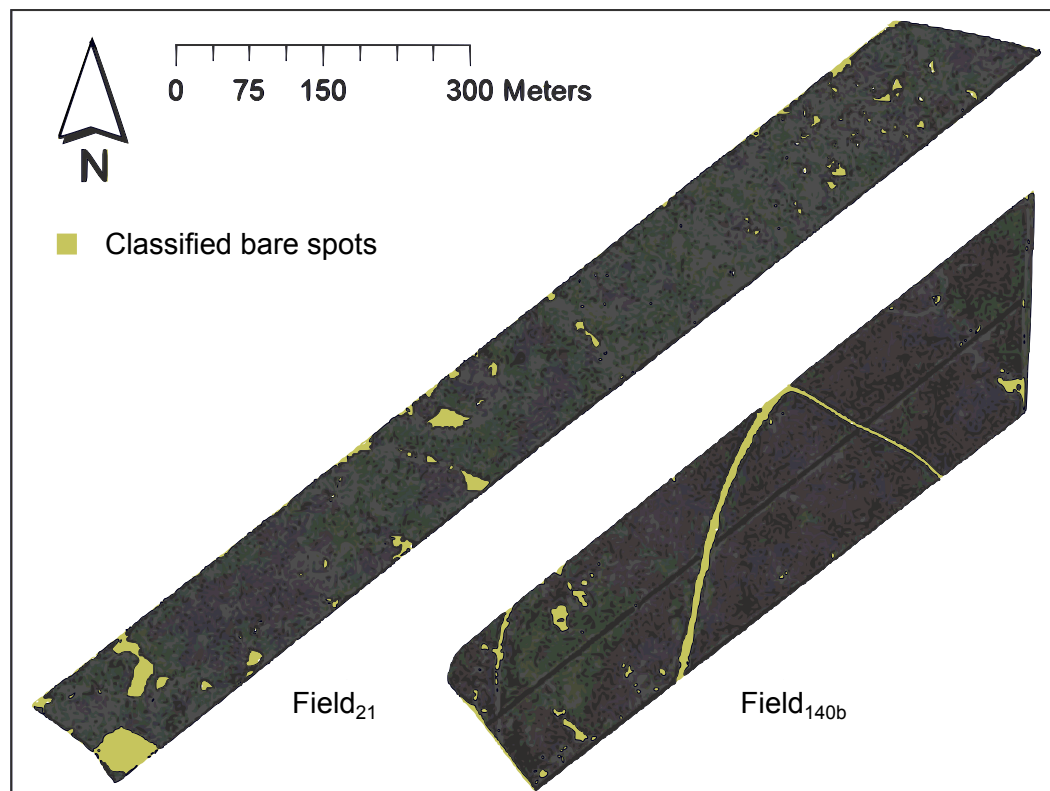


Fig. 4.1 Classification of bare spots in SPOT6 satellite image using the vegetation index derived from Principal Components Analysis.

The VI derived from PCA was nevertheless useful for classifying larger bare patches to be excluded from the regression approach. Average soil conditions in bare patches differ from general field conditions and will not respond linearly to a regression-based SSM treatment. The standard scores of average soil conditions in bare patches are summarized in C. The z-score represents distance from the field mean, so a value close to 0 has little difference from mean. $N_{21} = 8$ and $N_{140b} = 2$ (Appendix C). The average soil conditions in bare patches were found to have a slightly lower Shallow EC_a ($z_{21}=-0.143$, $z_{140b}=-0.090$), yet soil at the 5-15 cm depth showed higher than average pH ($z_{21}=0.33$, $z_{140b}=2.53$) and Al ($z_{21}=0.21$, $z_{140b}=0.13$) and lower than average TC ($z_{21}=-0.27$, $z_{140b}=-1.19$), K ($z_{21}=-0.59$, $z_{140b}=-1.347$), and Ca ($z_{21}=-0.25$, $z_{140b}=-1.725$). Given the non-linear relationship between EC_a and soil nutrient levels, excluding bare patches from the regression-based is recommended.

4.4 Correlation & PLS

Due to the high variability of yield, it was difficult to establish strong correlations between yield and other agronomic factors. Nevertheless, yield was significantly correlated with shallow measurements of total C, total N, Mg, and SOM in both fields (Table 4.5). These correlations were higher in Field 21 than Field 140b. Of all sampled properties, total N was most correlated with yield ($r_{21}=0.40$, $r_{140b}=0.34$) (Table 4.5).

Veris Shallow EC_a was significantly correlated with chemical attributes Mg and Fe at the 0-5 cm depth, and pH, Mg, and Fe at 5-15 cm depth (Table 4.5). Correlation with pH was higher in Field 21 than Field 140b ($r_{21}=0.38$, $r_{140b}=0.25$). In both fields, Veris Shallow EC_a was significantly correlated with soil texture properties, particularly sand ($r_{21}=-0.34$, $r_{140b}=-0.31$), medium sand ($r_{21}=-0.45$, $r_{140b}=-0.57$), and very fine sand ($r_{21}=0.48$, $r_{140b}=0.54$), indicating finer texture soil is positively correlated with EC_a . Thus, in this growing environment, EC_a appears to be an adequate predictor of soil texture and other agronomic attributes like nutrients. EC_a was not significantly correlated with TWI in Field 21 and was weakly correlated in Field 140b ($r_{140b}=0.25$). It is possible that due to the well-drained soils, moisture had less effect on EC_a readings and the wetness index was not significant.

Table 4.5 Pearson's correlation coefficients.

	Field 21			Field 140b		
	Yield	Elevation	Shallow EC _a	Yield	Elevation	Shallow EC _a
Yield	1.00	-0.22*	-0.2	1.00	-0.06	0.18
Elevation	-0.22*	1.00	-0.50***	-0.06	1.00	-0.28*
Veris Shallow	-0.02	-0.50***	1.00	0.18	-.28**	1.00
Veris Deep	0.02	-0.40***	0.52***	0.14	-0.19*	0.17
PRP 1.1	0.05	-0.73***	0.70***	0.09	-0.30*	0.65***
HCP 1.0	0.20*	-0.16	-0.05	0.14	-0.16	-0.04
HCP 2.0	0.21*	-0.69***	0.33**	0.06	-0.36***	-0.07
PRP 2.1	0.12	-0.85***	0.61***	0.28*	-0.15	0.54***
TWI	-0.02	0.16	-0.16	0.06	-0.22*	0.25*
Slope	-0.14	0.29**	0.05	-0.03	0.04	0.13
0-5 cm						
Total C	0.37***	0.00	0.00	0.33**	-0.15	0.39***
Total N	0.40***	-0.03	-0.01	0.34**	-0.19	0.44***
pH	-0.14	-0.34***	0.33***	-0.16	-0.30*	0.07
P	-0.183*	-0.29**	0.35**	-0.09	-0.31**	0.10
K	0.36***	-0.05	0.06	0.28*	-0.02	0.20*
Ca	-0.01	-0.06	0.16	0.15	-0.08	-0.14
Mg	0.21*	-0.11	0.20*	0.28*	-0.24*	0.38***
Al	-0.17*	-0.04	0.14	-0.06	-0.29**	0.15
Fe	0.12	-0.20*	0.28*	0.12	-0.22*	0.42***
S.O.M.	0.37***	-0.01	0.00	0.33**	-0.15	0.39***
5- 15 cm						
Total C	0.10	-0.02	0.16	0.25*	0.00	0.28*
Total N	0.08	-0.04	0.14	0.23	-0.14	0.36***
pH	-0.04	-0.50***	0.38***	-0.18	-0.46***	0.25*
P	-0.02	-0.11	0.04	0.02	-0.31**	0.17
K	0.25*	0.22*	-0.07	0.25*	0.04	0.16
Ca	0.10	0.02	0.14	0.26*	-0.07	0.21*
Mg	-0.01	-0.14	0.33**	0.03	-0.12	0.28*
Al	0.01	0.22*	-0.16	0.15	-0.03	0.16
Fe	-0.04	-0.32**	0.48***	0.05	-0.22**	0.38***
S.O.M.	0.10	-0.02	0.16	0.10	-0.02	0.15
Total Sand	-0.16	0.60***	-0.34***	-0.20	0.12	-0.31**
Total Silt	0.19*	-0.69***	0.38***	0.12	-0.06	0.25*
Total Clay	-0.14	0.55***	-0.28*	-0.02	0.15	-0.12
V. Coarse Sand	-0.13	0.34***	-0.22*	-0.02	-0.12	0.03
Coarse Sand	-0.11	0.63***	-0.44***	-0.15	0.39***	-0.32**
Med. Sand	-0.11	0.73***	-0.45***	-0.09	0.49***	-0.57***
Fine Sand	-0.08	-0.22*	0.26	0.12	-0.52***	0.38***
V. Fine Sand	0.12	-0.74***	0.48***	0.17	-0.35**	0.54***

NB: T-test significance denoted as *p < 0.05, **p < 0.001, ***p < 0.0001.

Elevation showed significant correlation with pH in both fields at both depths with higher elevations correlating to lower pH. Lower pH is usually considered better for wild blueberry growth as it suppresses competitors, but analysis showed this not to be the case at the experimental sites. One possibility is that although pH was significantly lower in high elevation areas, other agronomic factors corresponding with high elevation areas limited yield. Elevation was negatively correlated with P (0-5 cm) ($r_{21}=-0.29$, $r_{140b}=-0.31$) and Fe (5-15 cm) ($r_{21}=-0.32$, $r_{140b}=-0.22$).

Gravity causes redistribution of soil texture, which explains why the correlation between soil texture and elevation was significant. Very fine sand, fine sand, and silt contents were negatively correlated with elevation while medium to very coarse sands and clay were positively correlated with elevation. Finally, given their respective correlations with texture, it is unsurprising the EC_a and elevation are significantly correlated ($r_{21}=-0.50$, $r_{140b}=-0.28$). The relationship was stronger in Field 21, likely because the topography varied more dramatically (Table 4.5).

The intent of modeling with Partial Least Squares (PLS) regression was to observe the relationship of variability of multiple dependent agronomic variables and variability of sensor data. In Field 21 the model accounted for about 65% of EC_a and topographic features and about 23% of yield or 20% of 0-5 cm chemical attributes. 62% of the variance of EC_a and topographic features accounted for about 14% variance of 5-15 cm chemical attributes. All factor percentages are summarized in Table 4.6. By Wold's criterion, the variables of greatest influence in the regression for texture and sand content were Elevation and PRP 2.1; for chemical attributes (5-15 cm) influential variables were PRP 2.1 and slope; and for chemical attributes (0-5 cm) and yield, the VI was most influential. However, the PLS model only accounted for a small percentage of total variability in the target soil properties, indicating that there are external processes affecting the targeted properties which could not be modeled in PLS.

In Field 140b 24% percent variance of EC_a and topography accounted for about 14% of yield variance, notably less than in Field 21 where topography varies much more dramatically. The other factors of independent variables in Field 140b, summarized in Table 4.7,

captured a smaller percentage of the total variance than in Field 21, but more percent variance of the dependent variables were captured. By Wold's criterion, the VI and PRP 2.1 most influenced the yield model; The VI and Shallow EC_a most influenced the model of chemical attributes at both depths; Elevation, Shallow EC_a, and PRP 1.1 most influenced the sand content model; and the VI, PRP1.1 and Shallow EC_a most influenced the model of total texture. Once again, the PLS model only accounted for a small percentage of total variance in the target soil properties. Thus in both fields, external processes affect the target properties.

The exercise of PLS modeling is a reminder that external factors which were not sampled play a role in agronomic conditions. A certain degree of yield variability is to be expected in physical environments due to environmental factors such as weather fluctuations. For this reason yield maps will vary year to year. Historically factors such as topography and EC_a have been used to predict yield variability because they are more temporally stable. However sensor noise and environmental factors will also cause some variability in these measurements. This can explain why correlation is not always strong between yield, EC_a and topography variables. In Field 21, PLS accounted for 65.4% of EC_a, VI, and topography variance, as opposed to the 24% in Field 140b. With this model, 23% of yield variance could be accounted for. Other studies have used the same variables to capture 70% yield variability (Guo et al. 2012). This indicates latent factors exist which caused the high variability in yield. The year that yield was collected was a record year for yields in the region due to favorable weather patterns. The study would benefit from multiple years of yield sampling to determine a more temporally stable yield pattern within the field.

4.5 Geostatistical Analysis

A geostatistical summary produced from the R script is found in Tables 4.8 and 4.9. In general, the box-cox transformation effectively fit the data to a normal distribution. In some variables, such as pH, the automatic box-cox transformation was redundant, so it was not used. Tables 4.8 and 4.9 include the box-cox transformation variable λ to show which variables were transformed. A λ value of 1 signifies no transformation. Automation of the process allowed for a consistent methodology for interpolating different variables. The theoretical variogram model selection initially included the gaussian model. However,

the maps produced with the gaussian model were overly smooth and devoid of meaningful spatial patterns. When a spherical model was applied instead, the practical range of the variogram model better represented field variation patterns.

Strong spatial dependence was not observed in several properties (Tables 4.8 and 4.9), so OK could not accurately predict most chemical soil properties, with $RMSE_{SD}$ values greater than 0.90. Spatial dependence is classified by the nugget to sill ratio according to standards outlined by Camberdella *et al.* (1994). A variable belonging to a stronger spatial class should yield a better prediction from OK because the autocorrelation can be modeled. Unsurprisingly, elevation showed the strongest spatial dependence ($R^2_{21} = 1.00, R^2_{140b} = 1.00$). Maps were interpolated with $RMSE_{SD}=0.02$ and $RMSE_{SD}=0.08$ in Fields 21 and 140b, respectively. Most soil chemical attributes showed poor spatial structure. In Field 21, total C, K, total N (0-5cm), Ca (0-5 cm), and Mg (0-5cm) were completely random, fitted with a pure nugget model, and could not be predicted by kriging. In Field 140b, spatial dependence was moderate so that soil properties could be predicted except for total sand. However high $RMSE_{SD}$ and low R^2 for most chemical attributes indicate the properties predicted with OK were not truly accurate.

Table 4.6 Field 21 total percent variation accounted for by PLS factors.

Model	EC _a and Topo.			Dependent variables		
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
Yield	31.3	54.2	65.4	14.9	21.2	23.0
Chemical (0-5 cm)	37.8	52.1	65.2	8.3	18.2	20.1
Chemical (5-15 cm)	40.2	49.2	62.6	8.6	12.0	13.6
Sand Content	40.9	—	—	35.8	—	—
Total Texture	40.9	—	—	37.8	—	—

According to the correlation coefficients calculated in the cross-validation process, in Field 21 well predicted maps were Silt ($R^2=0.70$), Medium sand ($R^2=0.80$), Very fine sand ($R^2=0.85$), HCP10 ($R^2=0.92$), HCP20 ($R^2=0.95$), and Veris Shallow ($R^2=0.73$). Moderately accurately predicted maps were Sand ($R^2=0.68$), Coarse sand ($R^2=0.65$), pH at both depths ($R^2=0.56$ and 0.64 , respectively), PRP1.1 ($R^2=0.62$), and PRP2.1 ($R^2=0.67$). In Field 140b well predicted maps were coarse sand ($R^2=0.76$), medium sand ($R^2=0.75$), fine

Table 4.7 Field 140b total percent variation accounted for by PLS factors.

Model	EC _a and Topo.			Dependent variables		
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
Yield	24.0	–	–	13.8	–	–
Chemical (0-5 cm)	25.2	39.3	50.0	18.9	26.7	31.8
Chemical (5-15 cm)	24.6	36.8	45.6	9.00	23.3	29.7
Sand Content	25.4	38.0	48.7	29.3	35.5	39.7
Total Texture	26.0	39.2	–	14.0	19.4	–

sand ($R^2=0.77$), very fine sand ($R^2=0.76$), HCP1.0 ($R^2=0.88$), and HCP2.0 ($R^2=0.81$). Moderately accurately predicted maps were Clay ($R^2=0.68$), Al at both depths ($R^2=0.60$ and 0.62 , respectively), Fe (5-15 cm) ($R^2=0.53$), and Veris Shallow EC_a ($R^2=0.70$). They are presented in Figures 4.2 - 4.7.

Coefficients of variation were generally greater in properties in Field 21 than in Field 140b. Geostatistical analysis also showed random spatial structure among a number of chemical properties in Field 21, including S.O.M., total N, total C, K, Ca, and Mg at the 0-5 cm depth (Table 4.8). Yield values were high but sampled agronomic properties were randomly distributed, so establishing a relationship between yield variability and soil property variability was limited. Therefore, developing a management regime on the soil attributes alone could not reliably predict yield.

Conversely, several nutrients showed greater spatial structure in Field 140b (Table 4.9), and certain nutrients could be mapped with moderate cross validation results, namely Al at both depths and Fe (5-15 cm). EC_a was significantly correlated with a number of chemical attributes in Field 140b in addition to soil texture as in Field 21. Chemical attributes included TC, TN, Mg, Fe, and OM at both depth, and K (0-5 cm), Ca (5-15 cm). Tukey results show a number of these attributes significantly separated with the regression based method (Figures 4.15-4.17).

Based on the spatial structure and variability, the most useful prediction maps for the blueberry fields are the soil texture maps, EC_a maps, and topography. Texture maps can be compared to EC_a to verify EC_a as a predictor of soil texture. Given that nutrient avail-

Table 4.8 Geostatistical summary of agronomic properties in Field 21.

	Model ^a	Nugget ratio	Spatial class ^b	Range (m)	R^2	$RMSE_{SD}$	λ
Yield	p.n.	1.0	R	—	—	—	1.0
Elevation	sph	0.0	S	3032	1.00	0.02	1.0
Veris Shallow	sph	0.5	M	405	0.73	0.69	1.0
Veris Deep	sph	0.7	M	131	0.45	0.89	1.0
PRP 1.1	exp	0.5	M	239	0.62	0.78	1.0
HCP 1.0	sph	0.1	S	121	0.92	0.39	1.0
HCP 2.0	exp	0.0	S	87	0.95	0.32	1.0
PRP 2.1	sph	0.1	S	5423	0.67	0.74	1.0
0-5 cm							
S.O.M.	p.n.	1.0	R	—	—	—	0.0
Total N	p.n.	1.0	R	—	—	—	0.2
Total C	p.n.	1.0	R	—	—	—	0.1
Soil pH_{water}	exp	0.35	M	69	0.56	0.83	1.0
P	sph	0.6	M	29	0.36	0.93	-0.2
K	p.n.	1.0	R	—	—	—	-0.3
Ca	p.n.	1.0	R	—	—	—	1.0
Mg	p.n.	1.0	R	—	—	—	-0.1
Al	sph	0.3	M	18	0.23	0.97	-0.9
Fe	p.n.	1.0	R	—	—	—	-0.2
5- 15 cm							
S.O.M.	p.n.	1.0	W	—	—	—	-1.0
Total N	sph	0.9	W	376	0.13	0.99	-1.0
Total C	sph	1.0	R	—	—	—	-1.0
Soil pH_{water}	sph	0.36	M	416	0.64	0.76	1.0
P	sph	0.8	W	168	0.16	0.97	0.1
K	p.n.	1.0	R	—	—	—	-0.4
Ca	exp	0.9	W	26	0.20	0.98	1.0
Mg	sph	0.4	M	5537	0.26	0.96	-0.8
Al	sph	0.5	M	138	0.45	0.89	2.0
Fe	sph	0.3	M	2629	0.50	0.86	-0.3
Clay	sph	0.3	M	3223	0.55	0.83	0.5
Silt	sph	0.0	S	13725	0.72	0.69	0.0
Sand	exp	0.3	M	616	0.67	0.74	0.0
V. coarse sand	sph	0.58	M	608	0.40	0.92	0.1
Coarse sand	sph	0.0	S	8885	0.78	0.62	0.2
Medium sand	sph	0.0	S	4986	0.81	0.58	0.7
Fine sand	sph	0.8	W	313	0.29	0.95	0.5
V. fine sand	sph	0.1	S	749	0.85	0.53	0.1

^a p.n.= pure nugget (p.n.), sph.= spherical, exp.= exponential. ^b S = strong, M = moderate, W = weak, R = random.

Table 4.9 Geostatistical summary of agronomic properties in Field 140b.

	Model ^a	Nugget ratio	Spatial class ^b	Range (m)	R^2	$RMSE_{SD}$	λ
Yield	p.n.	1.0	R	—	—	—	1.0
Elevation	exp	0.0	S	167	1.00	0.08	1.0
Veris Shallow	sph	0.47	M	110	0.70	0.72	1.0
Veris Deep	sph	0.81	W	112	0.23	0.97	1.0
PRP 1.1	exp	0.8	W	157	0.38	0.92	1.0
HCP 1.0	sph	0.1	S	191	0.87	0.48	1.0
HCP 2.0	sph	0.2	S	249	0.81	0.58	1.0
PRP 2.1	exp	0.9	W	52	0.31	0.95	1.0
0-5 cm							
S.O.M.	exp	0.7	M	63	0.28	0.96	-0.1
Total N	exp	0.6	M	87	0.35	0.93	-0.2
Total C	sph	0.7	M	63	0.28	0.96	-0.1
Soil pH_{water}	exp	0.3	M	857	0.48	0.87	1.0
P	sph	0.6	M	564	0.42	0.90	-0.1
K	sph	0.7	M	2298	0.11	0.99	0.0
Ca	sph	0.1	S	5570	0.33	0.94	1.0
Mg	sph	0.4	M	4331	0.23	0.97	0.0
Al	exp	0.1	S	55	0.62	0.78	-0.8
Fe	exp	0.1	S	11820	0.39	0.92	0.2
5- 15 cm							
S.O.M.	sph	0.2	S	3926	0.37	0.93	-0.2
Total N	sph	0.3	M	1175	0.42	0.91	-0.4
Total C	p.n.	1.0	R	3927	0.37	0.93	-0.2
Soil pH_{water}	sph	0.1	S	4416	0.50	0.86	1.0
P	sph	0.5	M	604	0.50	0.86	0.0
K	sph	0.6	M	3026	0.18	0.98	0.5
Ca	exp	0.4	M	25	0.21	0.97	1.0
Mg	sph	0.6	M	147	0.33	0.94	-0.4
Al	sph	0.2	S	82	0.54	0.84	-1.0
Fe	sph	0.7	M	410	0.38	0.92	0.2
Clay	sph	0.3	M	112	0.61	0.79	2.0
Silt	exp	0.8	W	25	0.23	0.97	0.0
Sand	p.n.	1.0	R	—	—	—	2.0
V. coarse sand	sph	0.4	M	125	0.55	0.83	0.2
Coarse sand	sph	0.1	S	240	0.79	0.61	1.4
Medium sand	sph	0.2	S	542	0.76	0.65	2.0
Fine sand	sph	0.1	S	345	0.80	0.59	0.3
V. fine sand	sph	0.1	S	469	0.85	0.51	-0.6

^a p.n.= pure nugget (p.n.), sph.= spherical, exp.= exponential. ^b S = strong, M = moderate, W = weak, R = random.

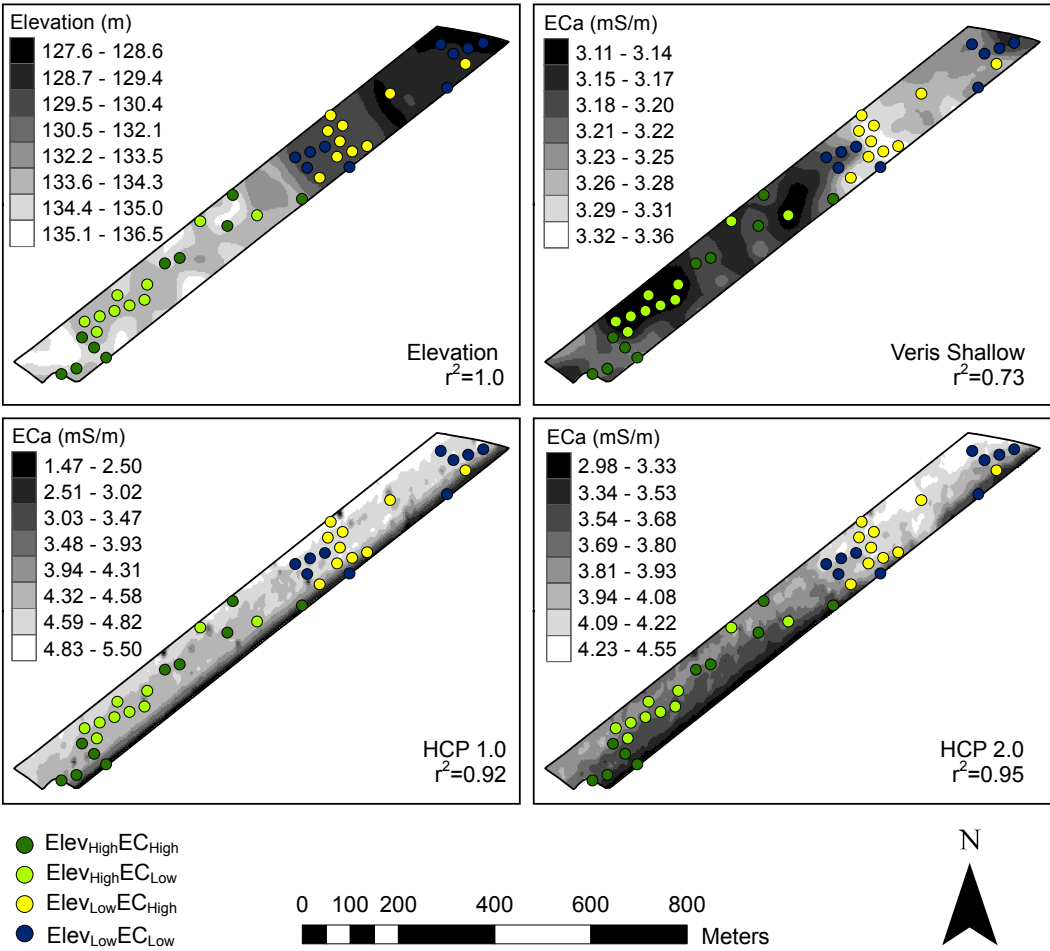


Fig. 4.2 Field 21 strongly correlated elevation and EC_a

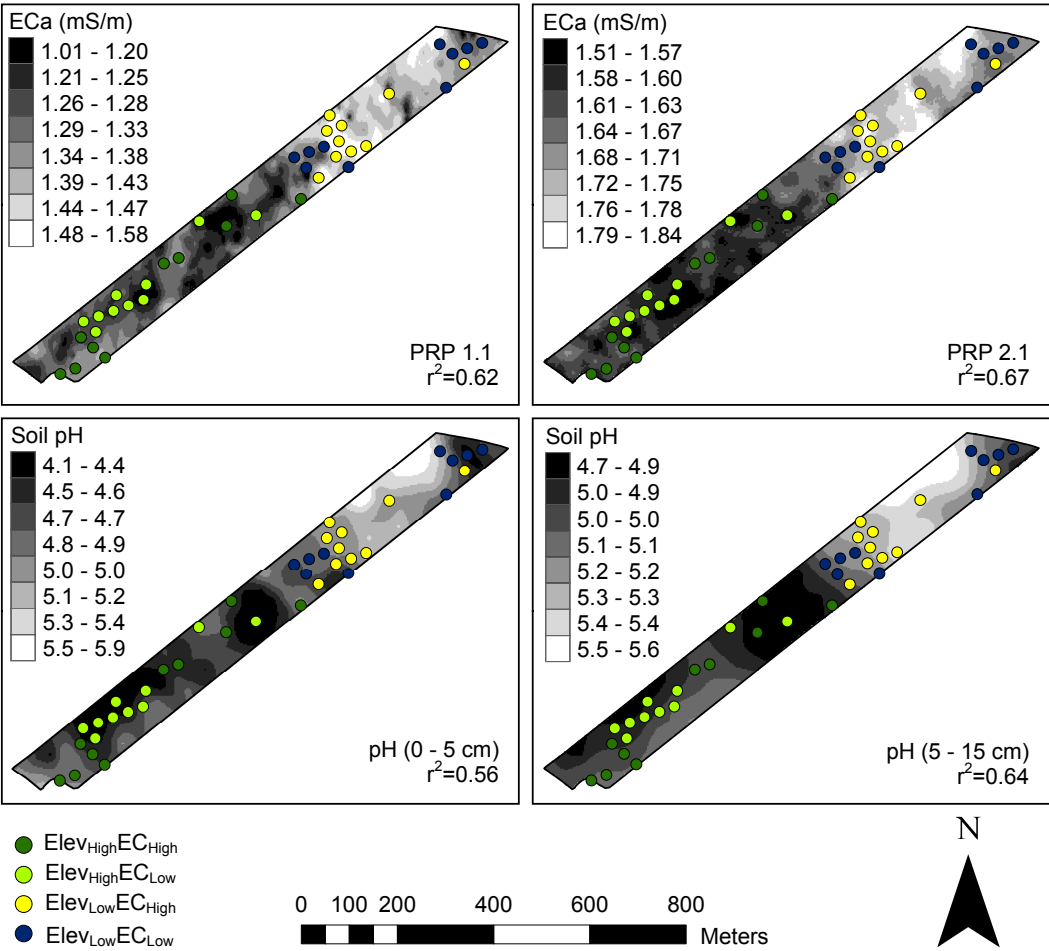


Fig. 4.3 Field 21 moderately correlated EC_a and chemical properties

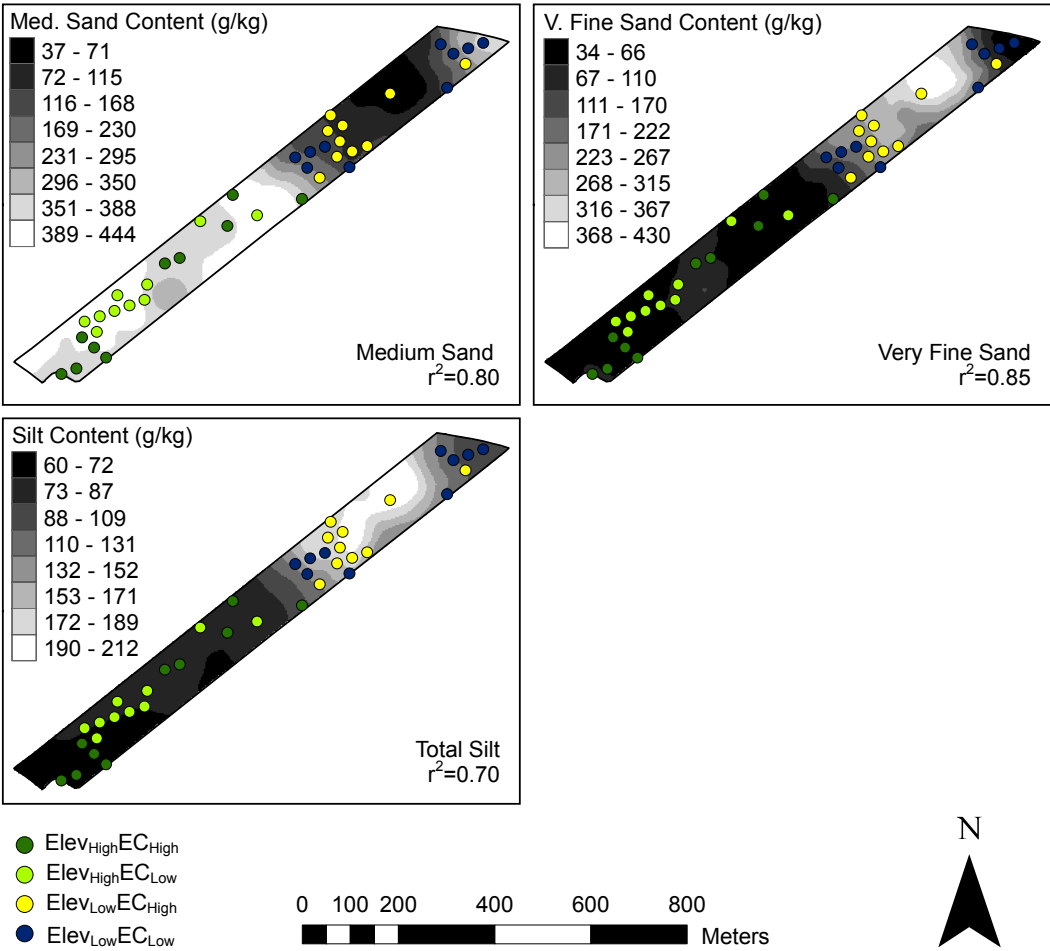


Fig. 4.4 Field 21 strongly correlated granulometric maps.

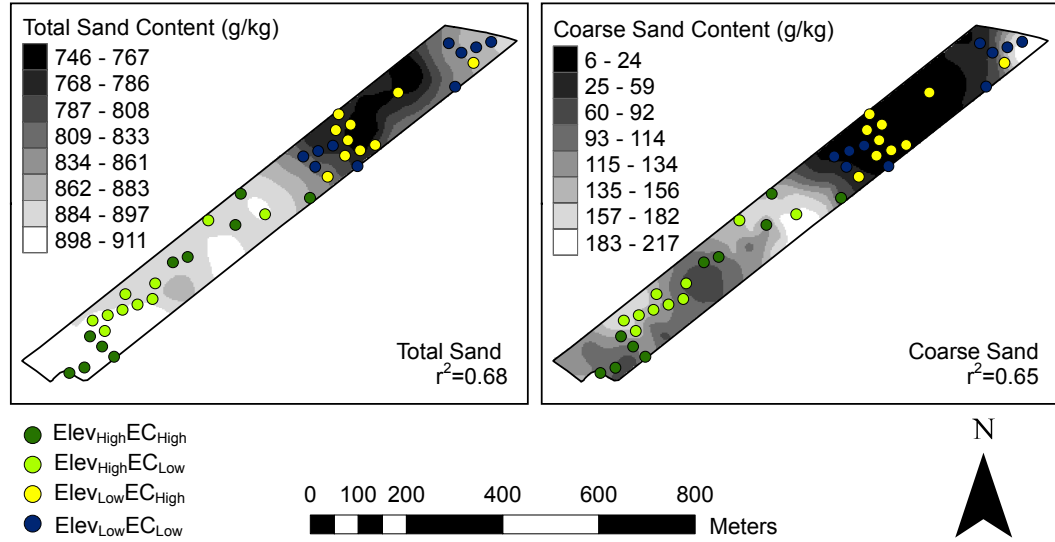


Fig. 4.5 Field 21 moderately correlated granulometric maps.

ability is linked with EC_a , texture, and elevation (Fig. 3.2), the thematic maps produced encompass both physical and chemical yield factors.

Maps with poor prediction may benefit from de-trending. However, co-kriging with elevation did not dramatically improve the accuracy of chemical variables. Correlation between elevation and chemical attributes was not strong enough to improve the prediction in either field. Regression kriging may also be considered but is only useful in variables which show strong spatial structure.

4.6 Data Separation for Regression-Based Approach

The scatter plots in Figures 4.9 and Figure 4.10 illustrate how extreme field conditions were identified for the regression-based approach. The two distinct clusters in Figure 4.9 represent the bimodal distribution and high variability of elevation in Field 21. EC_a is generally lower in the high elevation cluster, so EC_a is significantly different among the four scenarios according to Tukey's test (Fig. 4.12). The red points highlighted in the scatterplot represent occurrences of zero yield. They all occur in the high elevation cluster and are mostly distributed among higher EC_a . Tukey's post-hoc test revealed slope to also

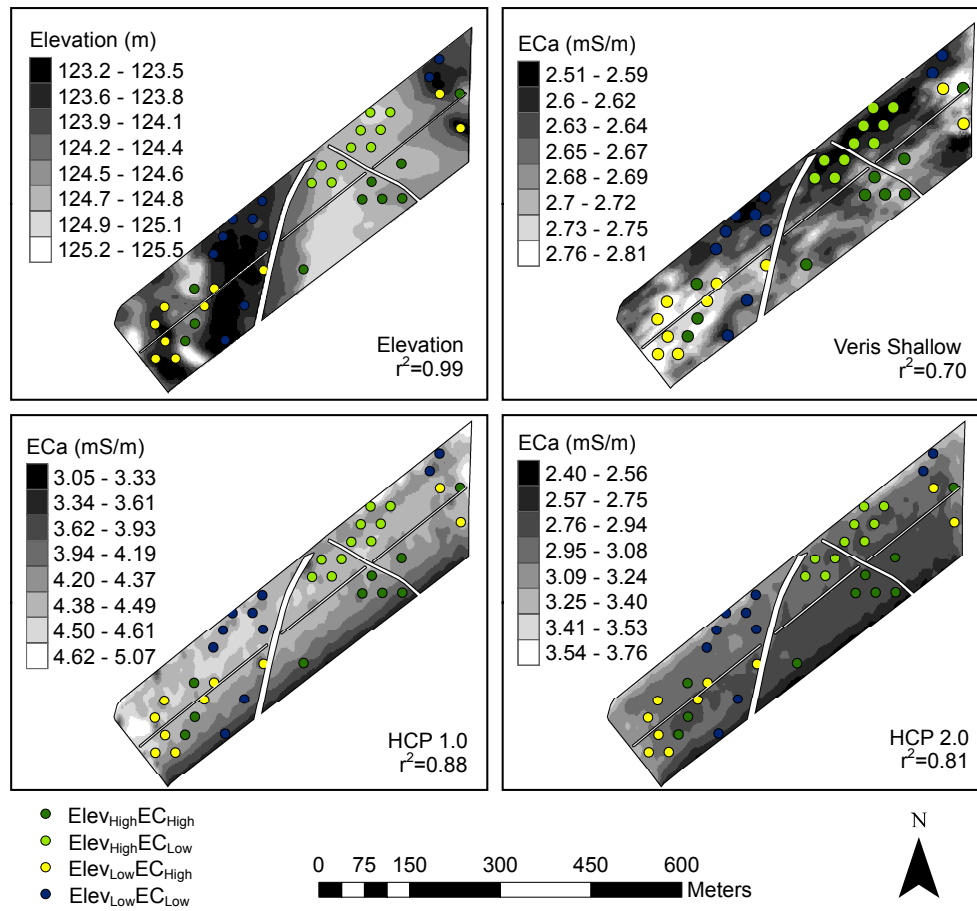


Fig. 4.6 Field 140b strongly correlated elevation and EC_a maps.

be significantly higher in the $Elev_{high} EC_{high}$ scenario. Bare patches may coincide with steep slope among other factors.

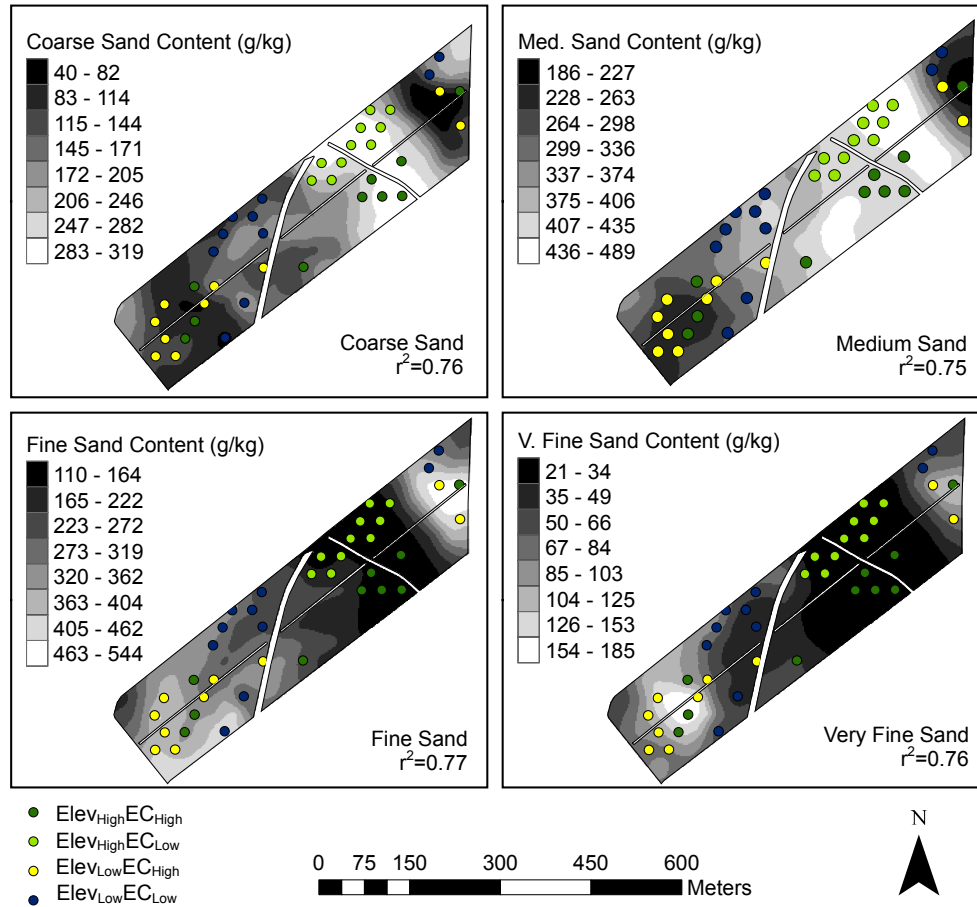


Fig. 4.7 Field 140b strongly correlated granulometric maps.

Analysis of variance (ANOVA) in Field 21 did not show a significant difference between the four scenarios ($p=0.81$). Tukey results showed a significant difference in soil texture among high and low elevation scenarios. Figure 4.16 shows medium sand content, silt, and very fine sand content to be separated by elevation, but not distinguished by EC_a . Figure 4.14 shows the combination of EC_a and elevation distinguishes other granulometric and chemical attributes (sand, coarse sand, Fe (0-5 cm), and pH at both depths). Three important field conditions were separated by this method: 1) slope, 2) pH, and 3) texture. Average pH was significantly higher in $Elev_{Low} & EC_{High}$, slightly above the optimal range.

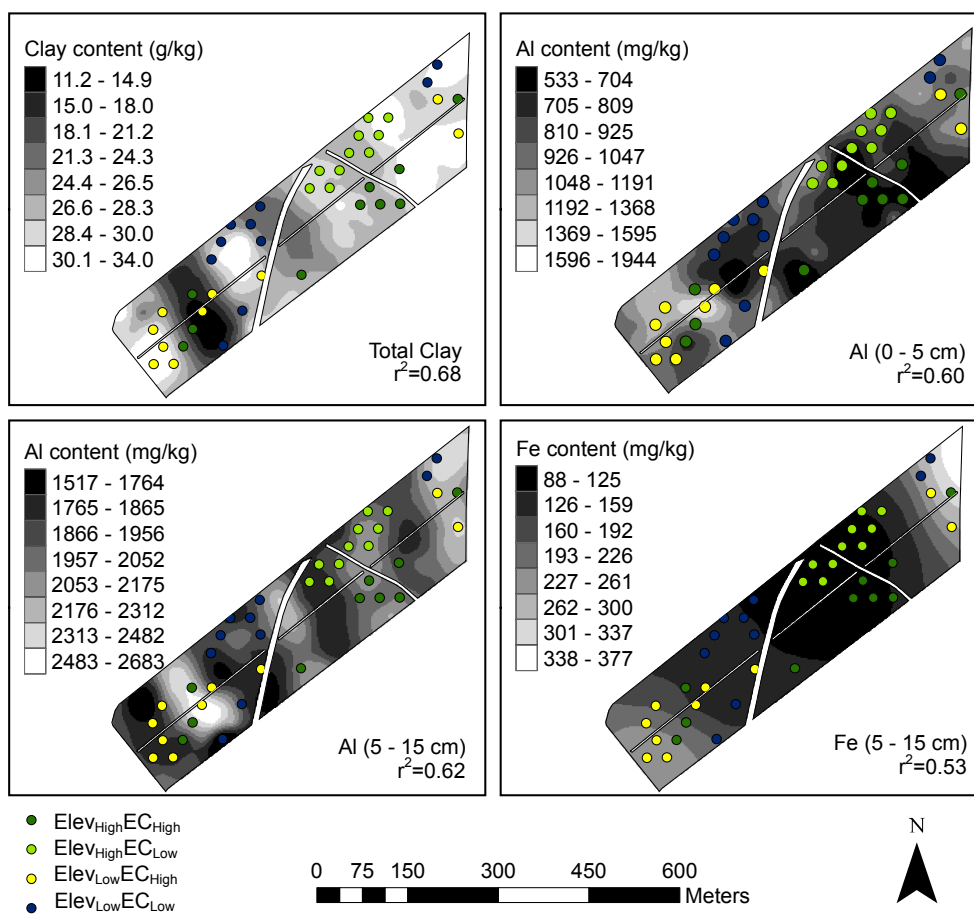


Fig. 4.8 Field 140b moderately correlated granulometric and chemical maps.

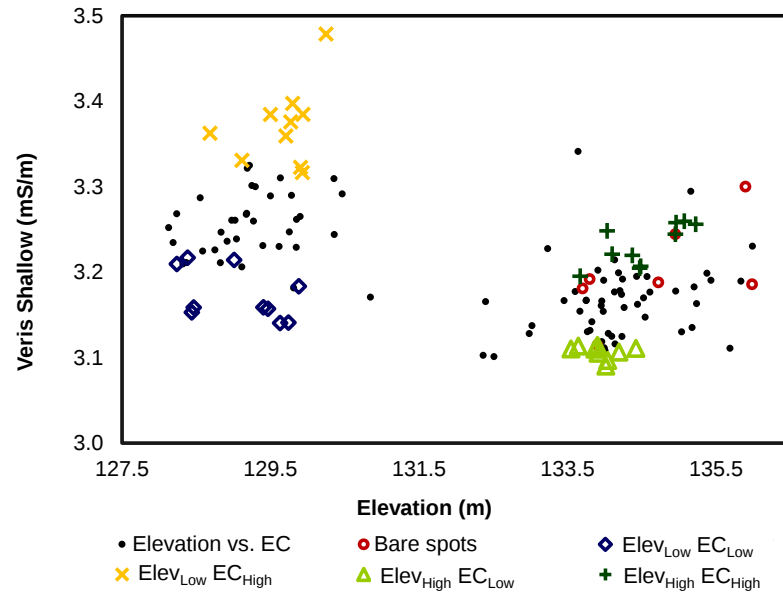


Fig. 4.9 Field 21 scatter plot of Elevation vs. EC_a values. The bimodal distribution of elevation is apparent in the scatter plot as two distinct clusters. Bare spots are highlighted in red and correspond with higher elevation.

Finally, texture was most correlated with yield in Field 21 and represents the field's nutrient and water storage potential.

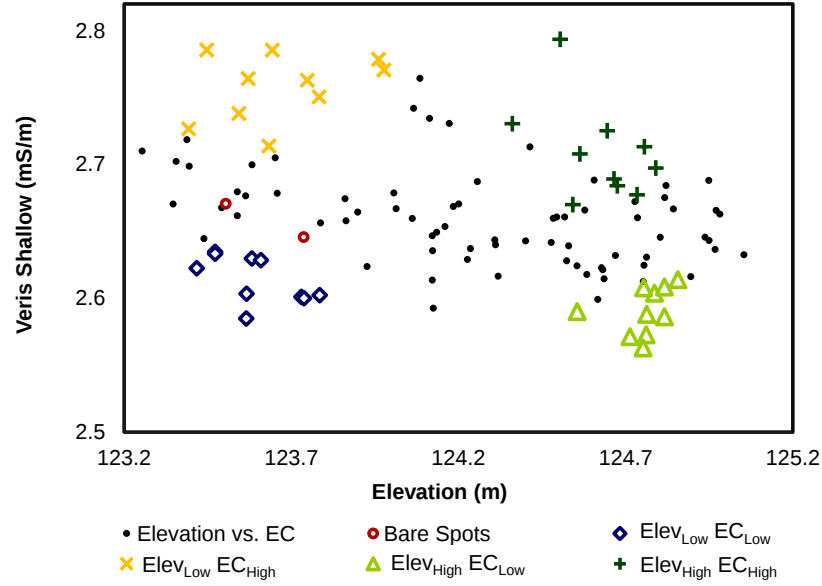


Fig. 4.10 Field 140b scatter plot of Elevation vs. EC_a values. Bare spots are highlighted in red

ANOVA in Field 140b showed an overall significant difference between the four scenarios at 90% confidence ($p=0.06$). Elevation in Field 140b is unimodal, but EC_a is still slightly lower in high elevation areas (Fig. 4.10). Like Field 21, the greatest distinction is between scenarios $Elev_{Low}EC_{High}$ and $Elev_{High}EC_{Low}$. The variables total sand, total silt, very fine sand, total C (5-15 cm), total N (5-15 cm), P (5-15 cm), and S.O.M. (5-15 cm), are all distinguished by these combinations (Fig. 4.15 and 4.16). A number of chemical attributes are also distinguished by high vs. low EC_a (Fig. 4.16 and 4.17). pH was found to be significantly different in $Elev_{High}EC_{Low}$ but still within ideal range. Elevation was less variable in Field 140b, so EC_a was more useful in separating data.

Based on the results of the Tukey post-hoc test, the greatest contrast in both fields

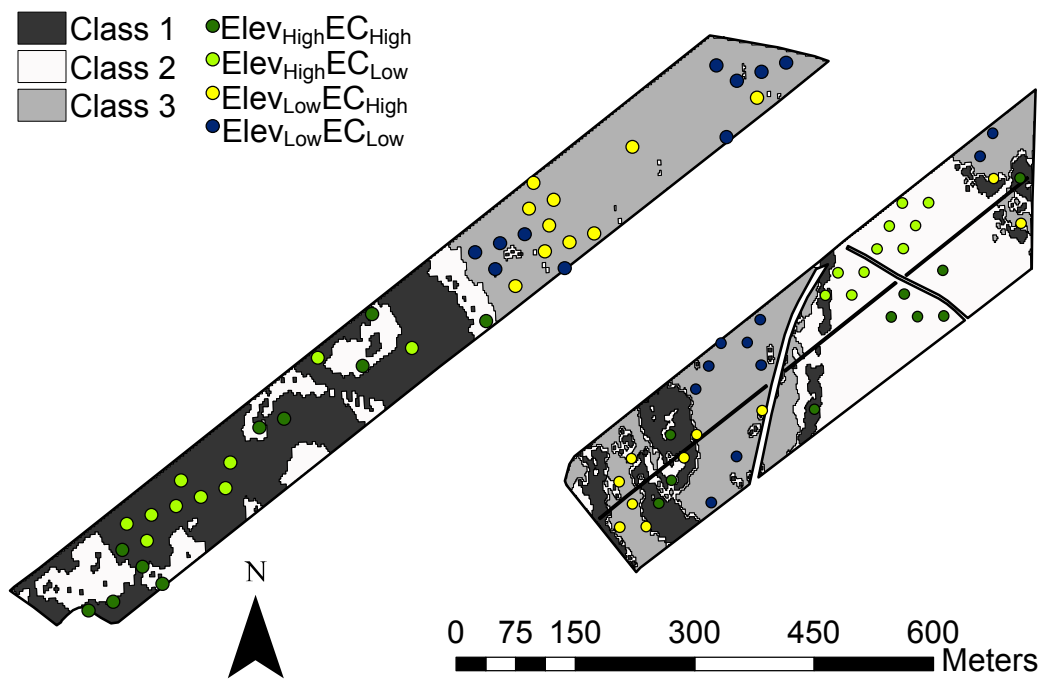


Fig. 4.11 Management zones classified with fuzzy c-means in MZA for Field 21 (left) and Field 140b (right). Targeted scenarios from the regression-based approach overlay the zones for comparison.

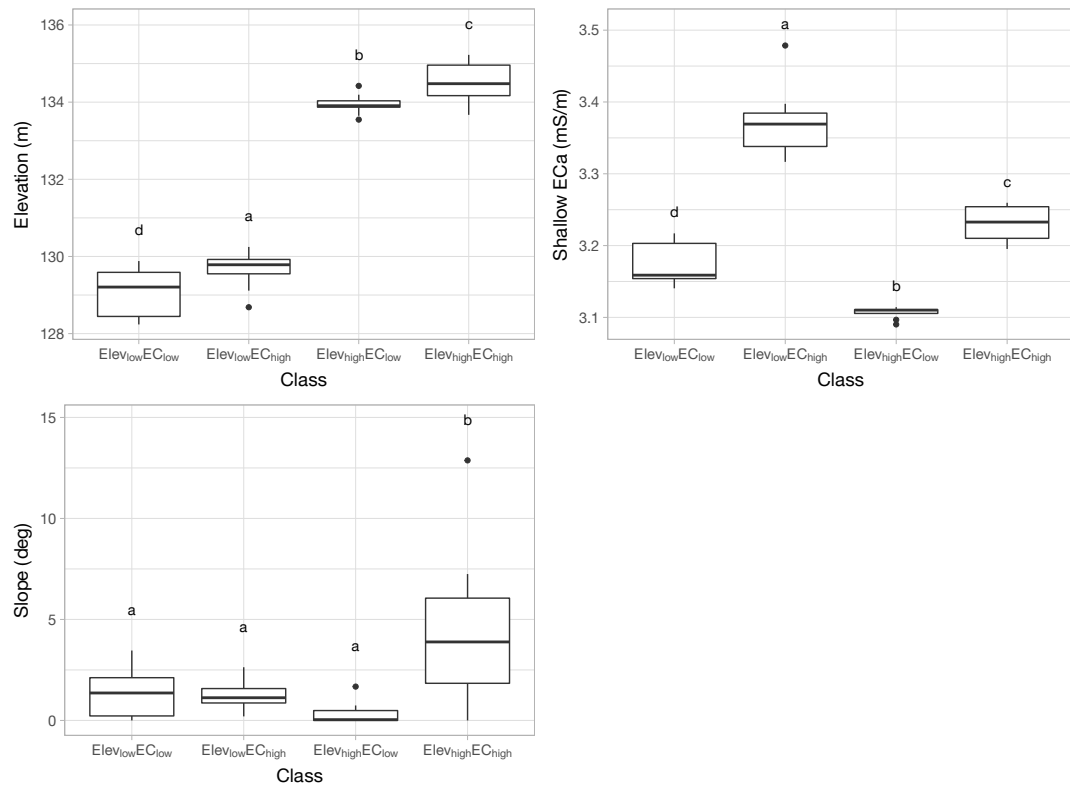


Fig. 4.12 Tukey plots of Field 21 elevation, EC_a , and slope. Slope is significantly higher in scenario $Elev_{high}EC_{high}$ than the other three scenarios.

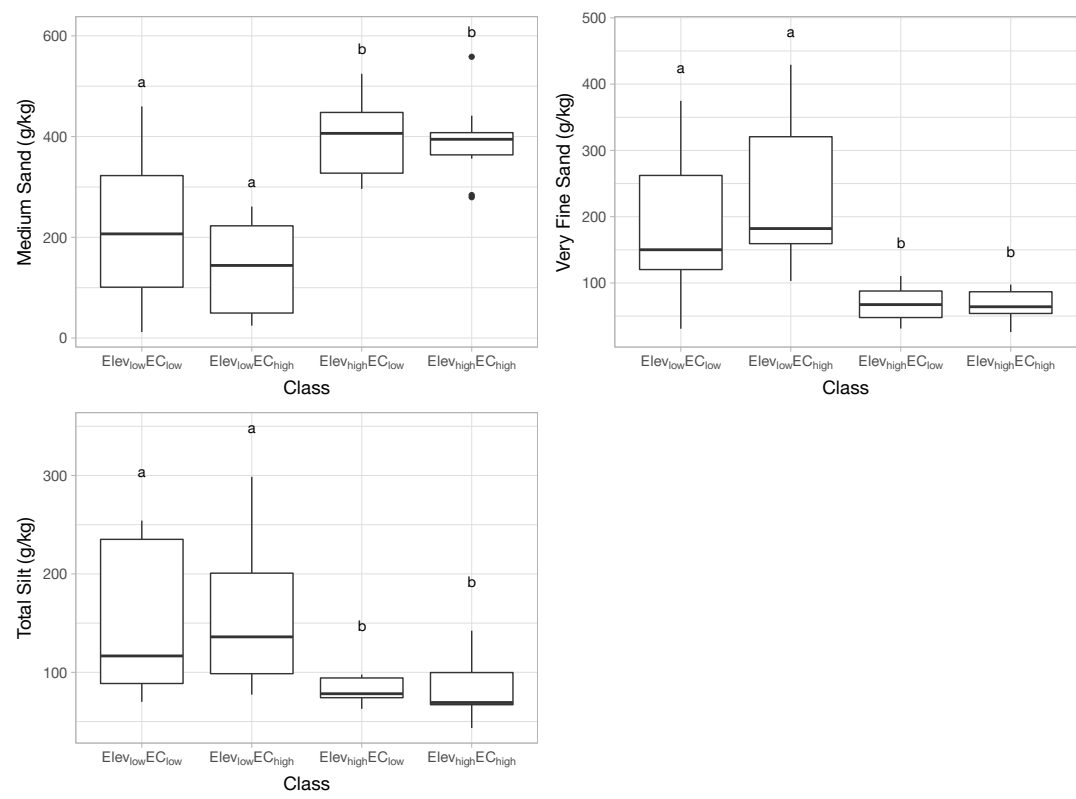


Fig. 4.13 Field 21 attributes that are significantly different based on contrast in elevation. Texture is influenced by elevation.

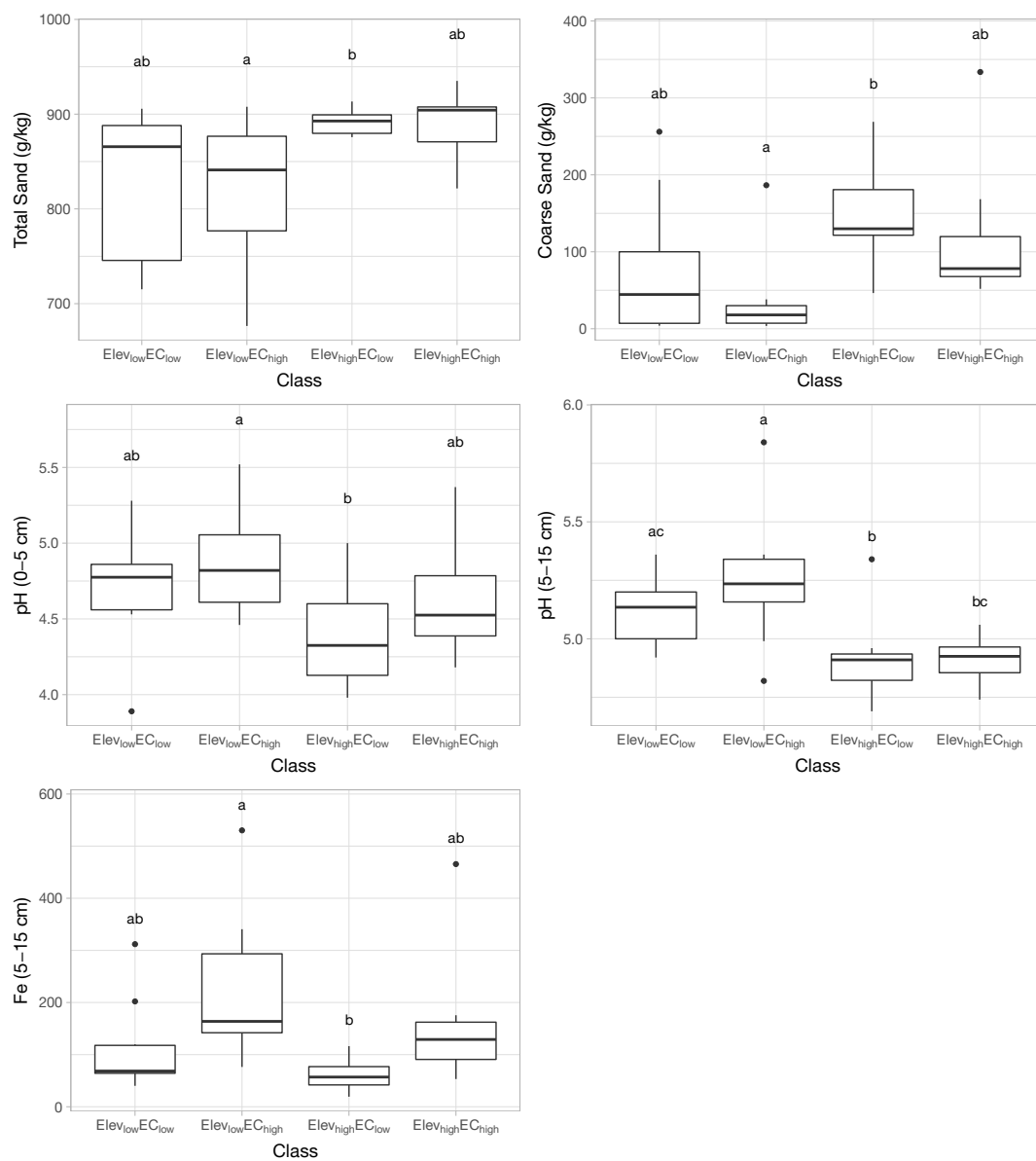


Fig. 4.14 Field 21 soil attributes that are contrasted by the combination of elevation and EC_a.

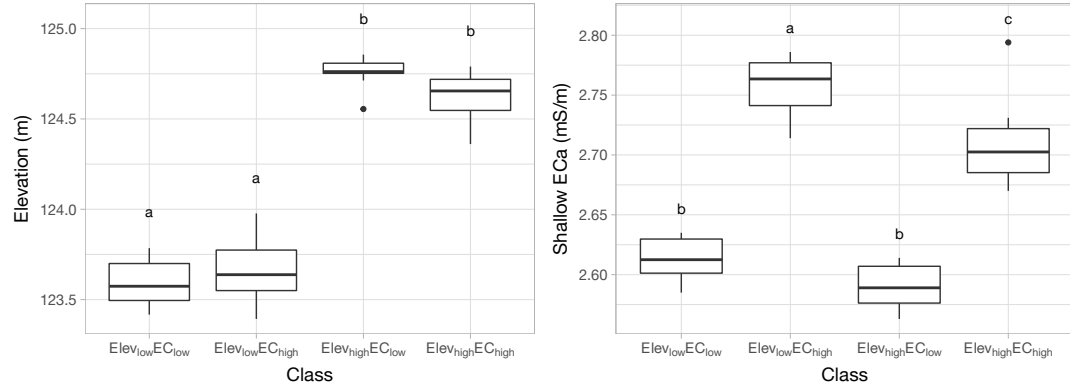


Fig. 4.15 Field 140b elevation and EC_a separation.

occurs between $Elev_{Low}EC_{High}$, and $Elev_{High}EC_{Low}$. These contrasts can be used to determine a regression-based treatment. Tables in Appendices A2 and A3 show the average values of each soil attribute in the four scenarios, standardized to the field average and standard deviation in order to illustrate the degree of difference from the average.

4.7 MZA vs. Regression-Based Approach

Three zones were selected in both fields using the MZA methodology (Fig. 4.11). In Field 21, Class 1 is a high elevation, low EC_a, low slope scenario; Class 2 is a high elevation, low EC_a, steep slope scenario, and Class 3 is low elevation, high EC_a, low slope scenario.

The higher elevation portion of the field somewhat coincides with the scenarios $Elevation_{High}EC_{High}$ and $Elevation_{Low}EC_{Low}$. $Elevation_{High}EC_{High}$ occurs close to but not within Class 2 while $Elevation_{Low}EC_{Low}$ falls within Class 1. Scenarios $Elev_{Low}EC_{High}$ and $Elev_{Low}EC_{Low}$ are encompassed by Class 3. These overlaps are largely due to the changes in topography and the fact that EC_a is correlated to elevation in this field. However, Class 2 does not identify the areas of highest EC_a. The greatest contrast exists between classes 1 and 3, irrespective of slope. This contrast exists in the variables total sand, total silt, total clay, very coarse sand, medium sand, very fine sand, pH deep, K (5-15 cm), P, and fine sand (Appendix A). Yield was significantly different between Class 2 and Class 3, likely because bare patch incidence coincides with steep slope in Field 21. In both the scatter plot separation method and the MZA method, the low elevation, low EC_a scenario is not

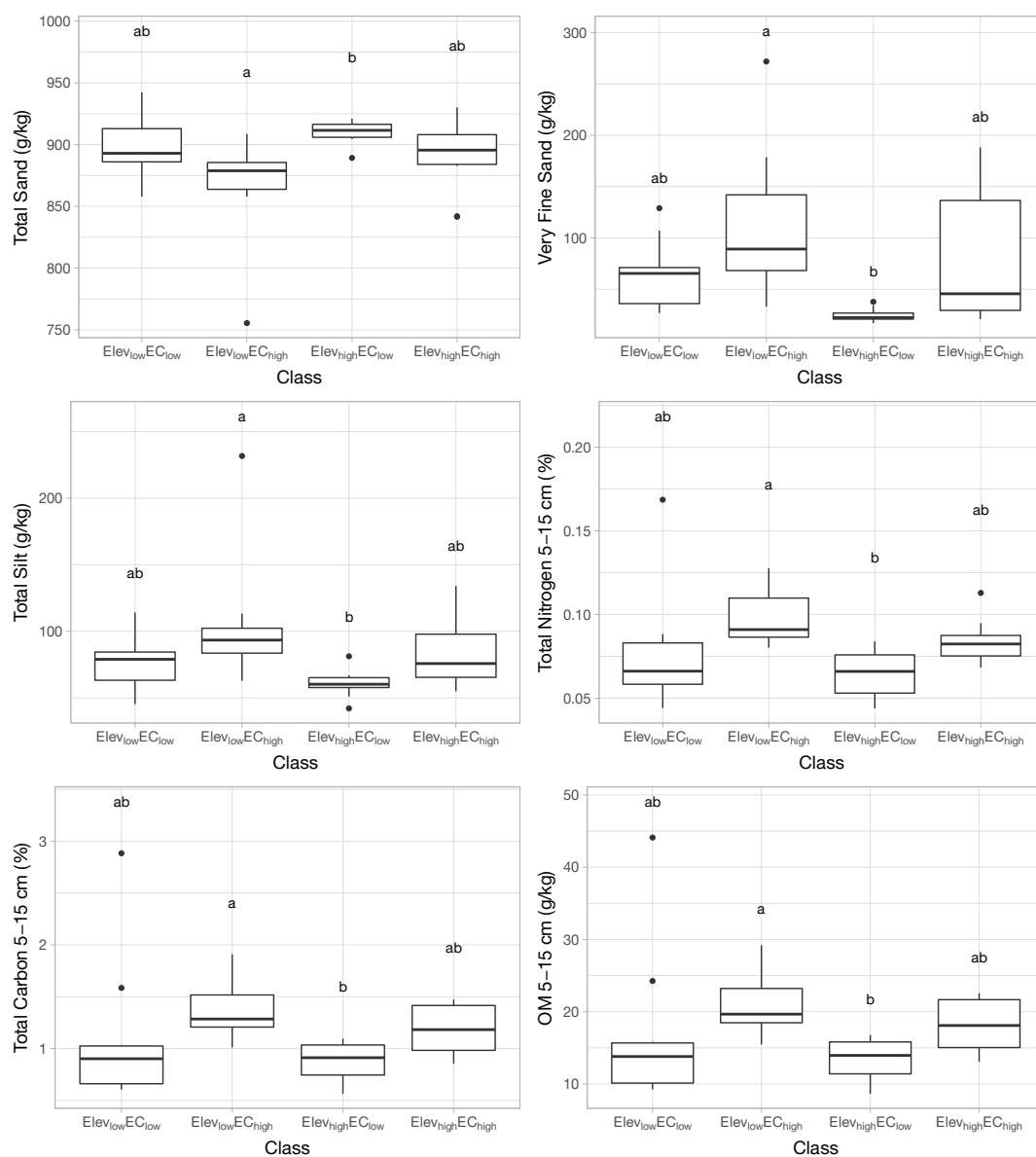


Fig. 4.16 Field 140b variables separated by the contrast of elevation and EC combined.

significant, and steep slope areas are isolated.

Like the scatter plot data separation method, the greatest contrast from the MZA output occurs between $Elev_{High}EC_{Low}$ and $Elev_{Low}EC_{High}$. Because Field 21 is topography driven, MZA performs well in separating the data. However, it fails to distinguish the areas of contrasting EC_a , which would be used in a regression-based approach.

In Field 140b, the classes are separated by elevation, slope & TWI, and to a lesser extent EC_a (Appendix B). Class 1 represents high elevation, high slope, high TWI, and high EC_a ; class 2 represents high elevation, low EC_a ; and class 3 represents low elevation and low EC_a (Appendix B). Scenarios $Elevation_{High}EC_{High}$ and $Elevation_{High}EC_{Low}$ coincide with Class 2 and parts of Class 1. Scenario $Elevation_{Low}EC_{Low}$ coincides with Class 3. Scenario $Elevation_{Low}EC_{High}$, however, is not meaningfully distinguished by MZA. Between classes 1 and 2 silt, clay, TC (0-5 cm), TN (0-5 cm), and OM (0-5 cm) are separated, indicating a distinction of agronomic properties by slope and EC_a . Between classes 2 and 3, Fe (5-15 cm), Mg, P, and pH are separated, indicating a distinction of agronomic properties by elevation (Appendix B). Because topography varies less in Field 140b, fewer agronomic properties were separated by MZA than the targeted approach.

Topographic features were more dominant in the MZA classification. Given the correlation of soil texture and properties with elevation, and given particularly the variability of topography in Field 21, this method of unsupervised classification performs well if the objective is to segment the field into smaller sub-fields. However, analyses show that areas of contrasting soil EC_a were not distinguished in either experimental field. For a prescription regression based on both soil EC_a and topography, the scatter plot selection method is more rapid and precise in identifying areas of extreme contrast.

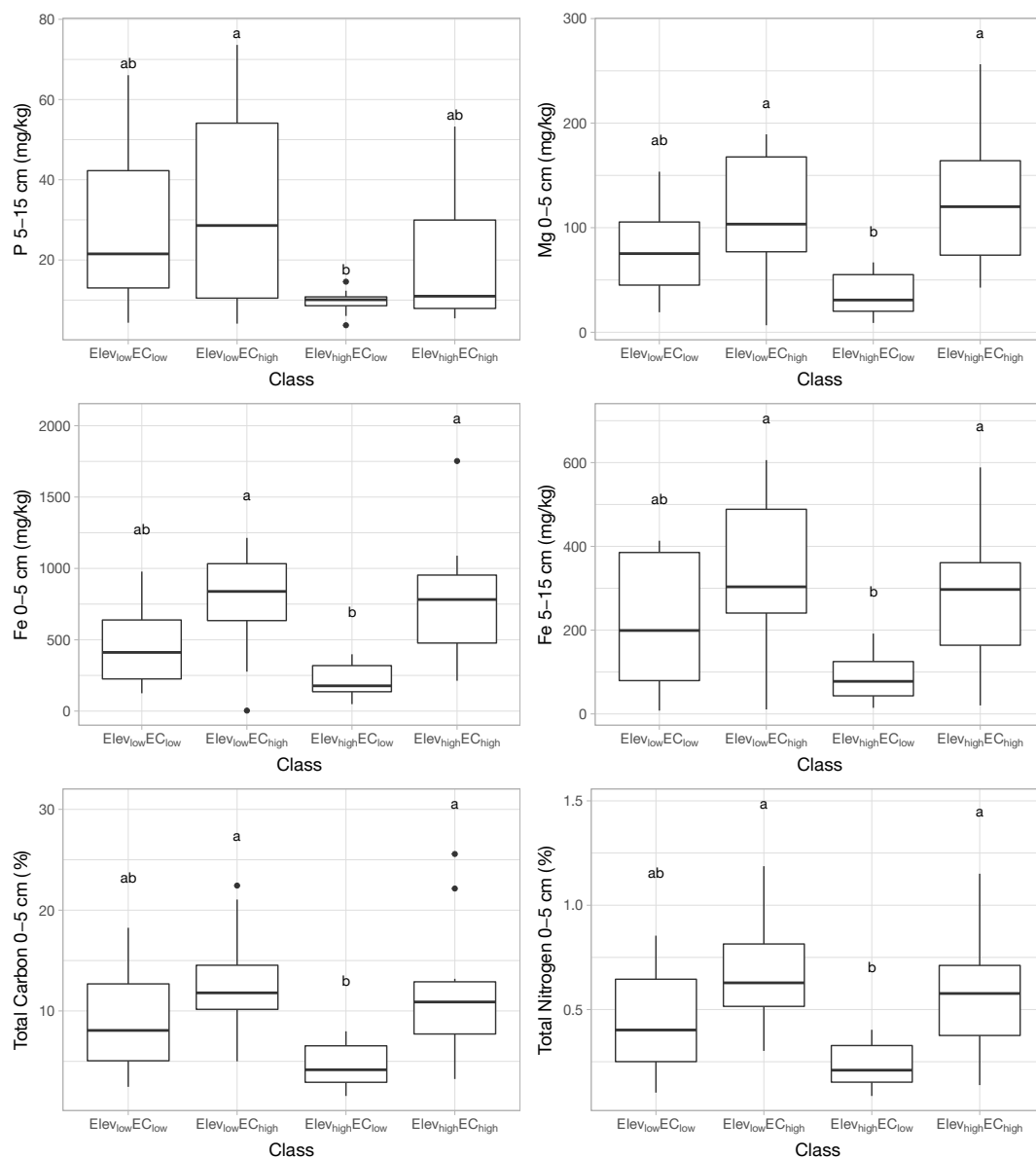


Fig. 4.17 Field 140b variables separated by the contrast of elevation and EC combined (P) and variables separated by high vs. low EC_a.

Chapter 5

Conclusions & Recommendations

5.1 Research Study Summary

The conclusions of the study which relate to the study objectives are:

1. Spatial variability could be characterized through soil EC_a and topographic data. Soil EC_a and topography showed strong spatial dependence and significantly correlated with a number of physical and chemical soil properties as well as yield in both fields. However yield and a number of soil properties showed weak or random spatial structure, so thematic maps of properties using Ordinary kriging produced high error. Thematic mapping of soil nutrients using Ordinary kriging generally yielded low-accuracy maps ($RMSE_{SD} > 0.90$) for chemical attributes, but physical soil attributes and pH showed greater spatial dependence and could be mapped with high accuracy. In the future, to map chemical attributes, strategic soil sampling should be implemented to try to detect spatial dependence at a shorter interval, and/or Regression kriging should be investigated using elevation and EC_a data where there is a correlation between proximally sensed variables and the target soil attribute.
2. Vegetation indices were not strongly correlated with yield partly due to the noise of soil reflectance and topography and partly because validating yield was challenging when yield was sampled at $1\ m^2$ and image resolution was $1.5\ m^2$. The greatest challenge in validating yield was that yield sampling was done two days prior to the date the satellite image was captured, disturbing biomass. However, the PCA-derived

vegetation index did successfully classify large contiguous bare spots which may be separated from the regression-based treatment.

3. While unsupervised classification with MZA software divided the field into three distinct areas by topographical traits, it did not distinguish the areas of most contrasting EC_a in either field. A comparison of means in the four targeted scenarios showed the greatest contrast in soil attributes to be between $Elev_{Low}EC_{High}$ and $Elev_{High}EC_{Low}$. These findings indicate an integration of elevation and EC_a data improves targeting within-field variation in order to develop a regression-based prescription map. Based on the agronomic properties in these contrasting scenarios, it is expected that areas in the field of low elevation and high EC_a will require less nutrient input than areas of the field of high elevation and low EC_a . Field trials are the next step to test this hypothesis.

5.2 Future Research

Fertilization recommendations in wild blueberry are based on leaf nutrient measurements rather than soil nutrient levels. Therefore, nutrient application rates cannot be inferred from the sampled soil data. Instead, field trials will test yield response in the selected areas within each of the four scenarios. In May 2017, AAFC began treatment in the sprout year of the wild blueberry fields. The study design is one 30 x 30 m experimental plot within each of the four scenarios in both fields. Within each plot, there are four 4 x 4 m subplots with a buffer for a total of 16 trials in each scenario. A variable rate sprayer applied levels of 0, 30, 60, and 90 kg N ha⁻¹ at four repetitions (Fig. 5.1). Crop response to the treatments will be used to model prescribed treatments in each of the four scenarios.

In addition to concluding the fertilization trials at this site and the implementation of a management plan based on the conclusions of that study, future research should investigate the holistic effects of SSM in wild blueberry plots, including tracking both environmental and economic impacts in the short and the long term. The costs and savings of SSM implementation should be carefully studied and quantified once fertilization levels are determined.

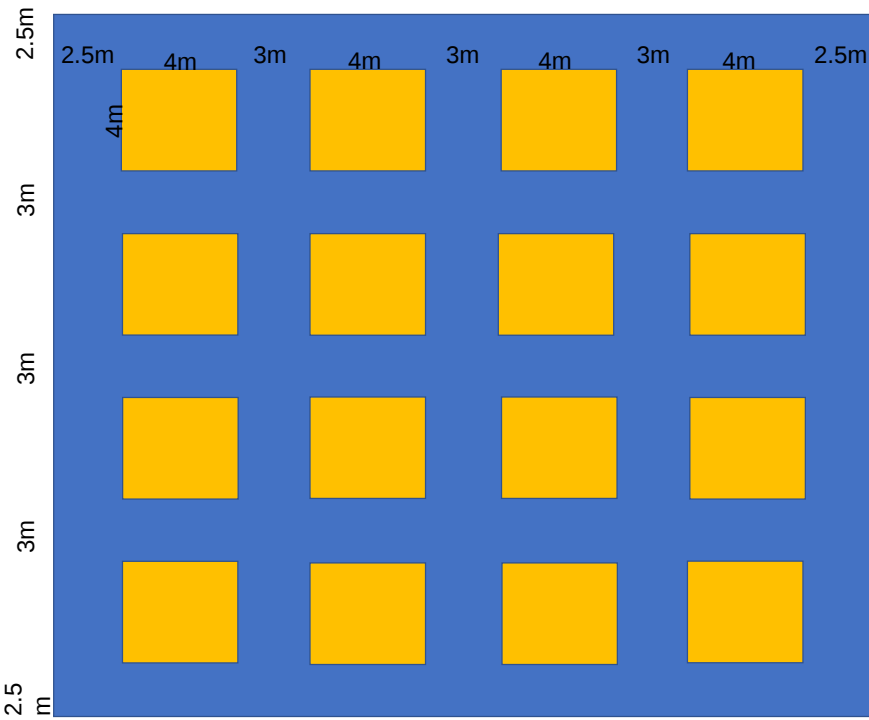


Fig. 5.1 Schematic of the field trial with one 30 x 30 m plot containing 16 subplots for 4 treatments of 0, 30, 60 and 90 kg N/ha at four repetitions.

Future studies should address the question of scalability of site-specific management strategies on this specialty crop in this region. Furthermore, the regression-based fertilization strategy lends itself to the development of a decision support system for determining fertilization levels based on sampled EC_a and elevation data.

Finally, future research should investigate how remote sensing can better predict crop density in wild blueberry fields. Higher spatial resolution imagery may improve VI utility. Archived satellite images would also allow for temporal mapping of yields and may be useful for determining the rate of rhizome development in bare patches of the field over time.

References

1. X. Zhang, L. Shi, X. Jia, G. Seielstad and C. Helgason, "Zone mapping application for precision-farming: a decision support tool for variable rate application," *Precision Agriculture*, vol. 11, pp. 103-114, 2010.
2. Q. U. Zaman, A. W. Schumann and D. C. Percival, "Development of an automated slope measurement and mapping system," in 2008 Providence, Rhode Island, June 29–July 2, 2008, 2008.
3. Q. U. Zaman, K. C. Swain, A. W. Schumann and D. C. Percival, "Automated, low-cost yield mapping of wild blueberry fruit," *Applied engineering in agriculture*, vol. 26, pp. 225-232, 2010.
4. D. E. Yarborough, "Establishment and management of the cultivated lowbush blueberry (*Vaccinium angustifolium*)," *International Journal of Fruit Science*, vol. 12, pp. 14-22, 2012.
5. Z. Yang, P. Willis and R. Mueller, "Impact of band-ratio enhanced AWIFS image to crop classification accuracy," *Pecora 17, The Future of Land Imaging*, Denver, pp. 18-20, 2008.
6. J. Xue and B. Su, "Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications," *Journal of Sensors*, vol. 2017, 2017.
7. N. C. Wollenhaupt, R. P. Wolkowski and M. K. Clayton, "Mapping soil test phosphorus and potassium for variable-rate fertilizer application," *Journal of Production Agriculture*, vol. 7, pp. 441-448, 1994.

8. B. M. Whelan, J. Cupitt, A. B. McBratney, P. C. Robert and others, "Practical definition and interpretation of potential management zones in Australian dryland cropping.," in *Proceedings of the 6th International Conference on Precision Agriculture and Other Precision Resources Management*, Minneapolis, MN, USA, 14-17 July, 2002., 2003.
9. B. M. Whelan and A. B. McBratney, "Definition and interpretation of potential management zones in Australia," in *Proceedings of the 11th Australian Agronomy Conference*, 2003.
10. H. Wackernagel, "External drift," in *Multivariate Geostatistics*, Springer, 2003, pp. 283-299.
11. E. Vrindts, A. M. Mouazen, M. Reyniers, K. Maertens, M. R. Maleki, H. Ramon and J. De Baerdemaeker, "Management zones based on correlation between soil compaction, yield and crop data," *Biosystems Engineering*, vol. 92, pp. 419-428, 2005.
12. U. W. A. Vitharana, M. Van Meirvenne, D. Simpson, L. Cockx and J. De Baerdemaeker, "Key soil and topographic properties to delineate potential management classes for precision agriculture in the European loess area," *Geoderma*, vol. 143, pp. 206-215, 2008.
13. C. J. Tucker, "Red and photographic infrared linear combinations for monitoring vegetation," *Remote Sensing of Environment*, vol. 8, pp. 127-150, 1979.
14. J. R. G. Townshend, T. E. Goff and C. J. Tucker, "Multitemporal dimensionality of images of normalized difference vegetation index at continental scales," *IEEE Transactions on geoscience and remote sensing*, pp. 888-895, 1985.
15. G. Thyssen, D. Percival, D. Burton and K. Sanderson, "Effect of nitrogen fertilizers on ammonia volatilization in wild blueberry production," *Canadian journal of plant science*, vol. 86, pp. 1383-1386, 2006.
16. J. A. Taylor, A. B. Mcbratney and R. Ciavarella, "Predicting and mapping winegrape quality from multiple must properties," in *Proc. of the 6th Int. Conf. on Precision Agriculture*, 2002.

17. J. A. Taylor, A. B. McBratney and B. M. Whelan, "Establishing management classes for broadacre agricultural production," *Agronomy Journal*, vol. 99, pp. 1366-1376, 2007.
18. K. A. Sudduth, N. R. Kitchen, G. A. Bollero, D. G. Bullock and W. J. Wiebold, "Comparison of electromagnetic induction and direct sensing of soil electrical conductivity," *Agronomy Journal*, vol. 95, pp. 472-482, 2003.
19. K. A. Sudduth, S. T. Drummond and N. R. Kitchen, "Accuracy issues in electromagnetic induction sensing of soil electrical conductivity for precision agriculture," *Computers and electronics in agriculture*, vol. 31, pp. 239-264, 2001.
20. A. S. M. Su, "Application of Proximal Soil Sensing For Environmental Characterization of Agricultural Land," 2016.
21. R. P. Sripada, R. W. Heiniger, J. G. White and A. D. Meijer, "Aerial color infrared photography for determining early in-season nitrogen requirements in corn," *Agronomy Journal*, vol. 98, pp. 968-977, 2006.
22. G. C. Simbahan and A. Dobermann, "Sampling optimization based on secondary information and its utilization in soil carbon mapping," *Geoderma*, vol. 133, pp. 345-362, 2006.
23. B. H. Sheldrick, Analytical methods manual 1984, Research Branch, Agriculture Canada Ottawa,, Canada, 1984.
24. J. F. Shanahan, J. S. Schepers, D. D. Francis, G. E. Varvel, W. W. Wilhelm, J. M. Tringe, M. R. Schlemmer and D. J. Major, "Use of remote-sensing imagery to estimate corn grain yield," *Agronomy Journal*, vol. 93, pp. 583-589, 2001.
25. J. M. Serrano, S. Shahidian and J. R. M. Da Silva, "Apparent electrical conductivity in dry versus wet soil conditions in a shallow soil," *Precision Agriculture*, vol. 14, pp. 99-114, 2013.
26. J. E. Sawyer, "Concepts of variable rate technology with considerations for fertilizer application," *Journal of Production Agriculture*, vol. 7, pp. 195-201, 1994.

27. S. R. Saleem, Q. U. Zaman, A. W. Schumann, A. Madani, A. A. Farooque and D. C. Percival, "Impact of variable rate fertilization on subsurface water contamination in wild blueberry cropping system," *Applied engineering in agriculture*, vol. 29, pp. 225-232, 2013.
28. T. Saey, D. Simpson, H. Vermeersch, L. Cockx and M. Van Meirvenne, "Comparing the EM38DD and DUALEM-21S sensors for depth-to-clay mapping," *Soil Science Society of America Journal*, vol. 73, pp. 7-12, 2009.
29. J. Rouse Jr, R. H. Haas, J. A. Schell and D. W. Deering, "Monitoring vegetation systems in the Great Plains with ERTS," 1974.
30. J.-L. Roujean and F.-M. Breon, "Estimating PAR absorbed by vegetation from bidirectional reflectance measurements," *Remote sensing of Environment*, vol. 51, pp. 375-384, 1995.
31. G. Rondeaux, M. Steven and F. Baret, "Optimization of soil-adjusted vegetation indices," *Remote sensing of environment*, vol. 55, pp. 95-107, 1996.
32. H. Robain, M. Descloitres, M. Ritz and Q. Y. Atangana, "A multiscale electrical survey of a lateritic soil system in the rain forest of Cameroon," *Journal of applied Geophysics*, vol. 34, pp. 237-253, 1996.
33. A. J. Richardson and C. L. Wiegand, "Distinguishing vegetation from soil background information," *Photogrammetric engineering and remote sensing*, vol. 43, pp. 1541-1552, 1977.
34. J. D. Rhoades, "Electrical conductivity methods for measuring and mapping soil salinity," *Advances in agronomy*, vol. 49, pp. 201-251, 1993.
35. Z. H. U. Qing and H. S. Lin, "Comparing ordinary kriging and regression kriging for soil properties in contrasting landscapes," *Pedosphere*, vol. 20, pp. 594-606, 2010.
36. D. Percival and K. Sanderson, "Main and interactive effects of vegetative-year applications of nitrogen, phosphorus, and potassium fertilizers on the wild blueberry," *Small Fruits Review*, vol. 3, pp. 105-121, 2004.

37. B. Pelletier, P. Dutilleul, G. Larocque and J.W. Fyles, "Coregionalization analysis with a drift for multi-scale assessment of spatial relationships between ecological variables 1. Estimation of drift and random components," *Environmental and ecological statistics*, vol. 16, p. 439, 2009.
38. E. Pebesma, B. Graeler and M. E. Pebesma, "Package 'gstat'," 2017.
39. S. J. Park and P. L. G. Vlek, "Environmental correlation of three-dimensional soil spatial variability: a comparison of three adaptive techniques," *Geoderma*, vol. 109, pp. 117-140, 2002.
40. R. A. Ortega and O. A. Santibáñez, "Determination of management zones in corn (*Zea mays* L.) based on soil fertility," *Computers and Electronics in agriculture*, vol. 58, pp. 49-59, 2007.
41. G. Oguri, P. Andrade-Sanchez and J. T. Heun, "Potential use of the Veris apparent EC sensor to predict soil texture under the semi-arid conditions of central Arizona," in 2009 Reno, Nevada, June 21-June 24, 2009.
42. I. O. A. Odeh, D. J. Chittleborough and A. B. McBratney, "Fuzzy-c-means and kriging for mapping soil as a continuous system," *Soil Science Society of America Journal*, vol. 56, pp. 1848-1854, 1992.
43. I. O. A. Odeh, A. B. McBratney and D. J. Chittleborough, "Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging," *Geoderma*, vol. 67, pp. 215-226, 1995.
44. S. C. Nolan, T. W. Goddard, G. Lohstraeter, G. M. Coen and others, "Assessing managements units on rolling topography," in *Proceedings of the 5th International Conference on Precision Agriculture*, Bloomington, Minnesota, USA, 16-19 July, 2000, 2000.
45. A. New Brunswick Department of Agriculture and Fisheries, "Soil Fertility and Fertilizers for Wild Blueberry," <http://perlebleue.ca/images/documents/amenagement/guideanglais/e> 1998.

46. M. Mzuku, R. Khosla, R. Reich, D. Inman, F. Smith and L. MacDonald, "Spatial variability of measured soil properties across site-specific management zones," *Soil Science Society of America Journal*, vol. 69, pp. 1572-1579, 2005.
47. S. Morrison, J. M. Smagula and W. Litten, "Morphology, growth, and rhizome development of *Vaccinium angustifolium* Ait. seedlings, rooted softwood cuttings, and micropropagated plantlets," *HortScience*, vol. 35, pp. 738-741, 2000.
48. F. J. Moral, J. M. Terrón and J. R. M. Da Silva, "Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques," *Soil and Tillage Research*, vol. 106, pp. 335-343, 2010.
49. B. Minasny and A. B. McBratney, "Incorporating taxonomic distance into spatial prediction and digital mapping of soil classes," *Geoderma*, vol. 142, pp. 285-293, 2007.
50. J. D. McNeill, Electrical conductivity of soils and rocks, Geonics Limited, 1980.
51. B. L. McCann, D. J. Pennock, C. Van Kessel and F. L. Walley, "The development of management units for site-specific farming," *Precision Agriculture*, pp. 295-302, 1996.
52. A. McBratney, B. Whelan, T. Ancev and J. Bouma, "Future directions of precision agriculture," *Precision agriculture*, vol. 6, pp. 7-23, 2005.
53. A. B. McBratney and I. O. A. Odeh, "Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions," *Geoderma*, vol. 77, pp. 85-113, 1997.
54. G. Matheron, *Traité de géostatistique appliquée*. 1 (1962), vol. 1, Editions Technip, 1962.
55. K. K. Mann, A. W. Schumann, T. A. Obreza, W. G. Harris and S. Shukla, "Spatial variability of soil physical properties affecting Florida citrus production," *Soil science*, vol. 175, pp. 487-499, 2010.
56. A. P. Mallarino and D. J. Wittry, "Efficacy of grid and zone soil sampling approaches for site-specific assessment of phosphorus, potassium, pH, and organic matter," *Precision Agriculture*, vol. 5, pp. 131-144, 2004.

57. J. MacQueen and others, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, 1967.
58. E. D. Lund, C. D. Christy and P. E. Drummond, "Practical Application of Soil Electrical Conductivity Mapping," 1999.
59. H. Q. Liu and A. Huete, "A feedback based modification of the NDVI to minimize canopy background and atmospheric noise," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, pp. 457-465, 1995.
60. J. Li and A. D. Heap, "Spatial interpolation methods applied in the environmental sciences: A review," *Environmental Modeling & Software*, vol. 53, pp. 173-189, 2014.
61. F. Li, Y. Miao, S. D. Hennig, M. L. Gnyp, X. Chen, L. Jia and G. Bareth, "Evaluating hyperspectral vegetation indices for estimating nitrogen concentration of winter wheat at different growth stages," *Precision Agriculture*, vol. 11, pp. 335-357, 2010.
62. W. S. Lee, V. Alchanatis, C. Yang, M. Hirafuji, D. Moshou and C. Li, "Sensing technologies for precision specialty crop production," *Computers and electronics in agriculture*, vol. 74, pp. 2-33, 2010.
63. J. D. Lauzon, I. P. O'Halloran, D. J. Fallow, A. P. Bertoldi and D. Aspinall, "Spatial variability of soil test phosphorus, potassium, and pH of Ontario soils," *Agronomy Journal*, vol. 97, pp. 524-532, 2005.
64. R. M. Lark and J. V. Stafford, "Information on within-field variability from sequences of yield maps: multivariate classification as a first step of interpretation," in *Soil and Water Quality at Different Scales*, Springer, 1998, pp. 277-281.
65. R. M. Lark and J. V. Stafford, "Classification as a first step in the interpretation of temporal and spatial variation of crop yield," *Annals of Applied Biology*, vol. 130, pp. 111-121, 1997.
66. J. Lafond and N. Ziadi, "Nitrogen and phosphorus fertilization in wild lowbush blueberry in Quebec," *Canadian Journal of Plant Science*, vol. 91, pp. 535-544, 2011.

67. A. N. Kravchenko and D. G. Bullock, "Correlation of corn and soybean grain yield with topography and soil properties," *Agronomy Journal*, vol. 92, pp. 75-83, 2000.
68. B. Koch, R. Khosla, W. M. Frasier, D. G. Westfall and D. Inman, "Economic feasibility of variable-rate nitrogen application utilizing site-specific management zones," *Agronomy Journal*, vol. 96, pp. 1572-1580, 2004.
69. M. Kotters, D. J. Brus and J. H. O. Voshaar, "A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations," *Geoderma*, vol. 67, pp. 227-246, 1995.
70. N. R. Kitchen, K. A. Sudduth, D. B. Myers, S. T. Drummond and S. Y. Hong, "Delineating productivity zones on claypan soil fields using apparent soil electrical conductivity," *Computers and Electronics in Agriculture*, vol. 46, pp. 285-308, 2005.
71. R. Khosla, K. Fleming, J. A. Delgado, T. M. Shaver and D. G. Westfall, "Use of site-specific management zones to improve nitrogen management for precision agriculture," *Journal of Soil and Water Conservation*, vol. 57, pp. 513-518, 2002.
72. R. Khosla, D. Inman, D. G. Westfall, R. M. Reich, M. Frasier, M. Mzuku, B. Koch and A. Hornung, "A synthesis of multi-disciplinary research in precision agriculture: site-specific management zones in the semi-arid western Great Plains of the USA," *Precision Agriculture*, vol. 9, pp. 85-100, 2008.
73. R. Kerry, P. Goovaerts, D. Giménez, P. Oudemans and E. Muñiz, "Investigating geostatistical methods to model within-field yield variability of cranberries for potential management zones," *Precision agriculture*, vol. 17, pp. 247-273, 2016.
74. R. G. Kachanoski, I. J. Van Wesenbeeck and J. De, "Field scale patterns of soil water storage from non-contacting measurements of bulk electrical conductivity," *Canadian Journal of Soil Science*, vol. 70, pp. 537-542, 1990.
75. R. G. Kachanoski, I. J. Van Wesenbeeck and E. G. Gregorich, "Estimating spatial variations of soil water content using noncontacting electromagnetic inductive methods," *Canadian Journal of Soil Science*, vol. 68, pp. 715-722, 1988.

76. C. Jones and J. Jacobsen, "Nitrogen cycling, testing and fertilizer recommendations," *Nutrient Management Module*, vol. 3, 2005.
77. C. K. Johnson, J. W. Doran, H. R. Duke, B. J. Wienhold, K. M. Eskridge and J. F. Shanahan, "Field-scale electrical conductivity mapping for delineating soil condition," *Soil Science Society of America Journal*, vol. 65, pp. 1829-1837, 2001.
78. G. F. Jenks, "The data model concept in statistical mapping," *International yearbook of cartography*, vol. 7, pp. 186-190, 1967.
79. J. R. Istas, "Effect of Ammonia on the Acidification of the Environment," 1988.
80. D. Inman, R. Khosla, R. Reich and D. G. Westfall, "Normalized difference vegetation index and soil color-based management zones in irrigated maize," *Agronomy journal*, vol. 100, pp. 60-66, 2008.
81. R. J. Hyndman, M. O'Hara-Wild, C. Bergmeir, S. Razbash, E. Wang and M. R. Hyndman, "Package 'forecast'," Online] <https://cran.r-project.org/web/packages/forecast/forecast.pdf>, 2017.
82. A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao and L. G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote sensing of environment*, vol. 83, pp. 195-213, 2002.
83. A. R. Huete, "A soil-adjusted vegetation index (SAVI)," *Remote sensing of environment*, vol. 25, pp. 295-309, 1988.
84. A. Hornung, R. Khosla, R. Reich, D. Inman and D. G. Westfall, "Comparison of site-specific management zones," *Agronomy Journal*, vol. 98, pp. 407-415, 2006.
85. P. R. Hicklenton, J. Y. Reekie and R. J. Gordon, "Physiological and morphological traits of lowbush blueberry (*Vaccinium angustifolium* Ait.) plants in relation to post-transplant conditions and water availability," *Canadian journal of plant science*, vol. 80, pp. 861-867, 2000.
86. T. Hengl, G. B. M. Heuvelink and D. G. Rossiter, "About regression-kriging: from equations to case studies," *Computers & geosciences*, vol. 33, pp. 1301-1315, 2007.

87. T. Hengl, *A practical guide to geostatistical mapping*, vol. 52, 2009.
88. T. Hengl, G. B. M. Heuvelink and A. Stein, "A generic framework for spatial prediction of soil variables based on regression-kriging," *Geoderma*, vol. 120, pp. 75-93, 2004.
89. T. Hengl, G. Heuvelink and A. Stein, "Comparison of kriging with external drift and regression kriging," 2003.
90. F. G. Helfferich, *Ion exchange*, Courier Corporation, 1962.
91. C. B. Hedley, I. J. Yule, C. R. Eastwood, T. G. Shepherd and G. Arnold, "Rapid identification of soil textural and management zones using electromagnetic induction sensing of soils," *Soil Research*, vol. 42, pp. 389-400, 2004.
92. A. D. Halvorson and J. D. Rhoades, "Field Mapping Soil Conductivity to Delineate Dryland Saline Seeps with Four-electrode Technique 1," *Soil science society of America journal*, vol. 40, pp. 571-575, 1976.
93. V. M. Guerrero, "Time-series analysis supported by power transformations," *Journal of forecasting*, vol. 12, pp. 37-48, 1993.
94. P. Goovaerts, *Geostatistics for natural resources evaluation*, Oxford University Press on Demand, 1997.
95. N. S. Goel and W. Qin, "Influences of canopy architecture on relationships between various vegetation indices and LAI and FPAR: A computer simulation," *Remote Sensing Reviews*, vol. 10, pp. 309-347, 1994.
96. A. A. Gitelson, Y. J. Kaufman and M. N. Merzlyak, "Use of a green channel in remote sensing of global vegetation from EOS-MODIS," *Remote sensing of Environment*, vol. 58, pp. 289-298, 1996.
97. A. A. Gitelson and M. N. Merzlyak, "Remote sensing of chlorophyll concentration in higher plant leaves," *Advances in Space Research*, vol. 22, pp. 689-692, 1998.
98. S. Gagnon, *Wild Blueberry Production Guide...in a Context of Sustainable development*, R. Willmot, Ed., CRAAQ, 2013.

99. Z. L. Frogbrook and M. A. Oliver, "Comparing the spatial predictions of soil organic matter determined by two laboratory methods," *Soil Use and Management*, vol. 17, pp. 235-244, 2001.
100. S. P. Friedman, "Soil properties influencing apparent electrical conductivity: a review," *Computers and electronics in agriculture*, vol. 46, pp. 45-70, 2005.
101. J. J. Fridgen, N. R. Kitchen, K. A. Sudduth and others, "Variability of soil and landscape attributes within sub-field management zones," in *Proceedings of the 5th International Conference on Precision Agriculture*, Bloomington, Minnesota, USA, 16-19 July, 2000, 2000.
102. J. J. Fridgen, N. R. Kitchen, K. A. Sudduth, S. T. Drummond, W. J. Wiebold and C. W. Fraisse, "Management zone analyst (MZA)," *Agronomy Journal*, vol. 96, pp. 100-108, 2004.
103. C. W. Fraisse, K. A. Sudduth and N. R. Kitchen, "Delineation of site-specific management zones by unsupervised classification of topographic attributes and soil electrical conductivity," *Transactions of the ASAE*, vol. 44, p. 155, 2001.
104. K. L. Fleming, D. F. Heermann and D. G. Westfall, "Evaluating soil color with farmer input and apparent soil electrical conductivity for management zone delineation," *Agronomy Journal*, vol. 96, pp. 1581-1587, 2004.
105. N. Favaretto, L. D. Norton, B. C. Joern and S. M. Brouder, "Gypsum amendment and exchangeable calcium and magnesium affecting phosphorus and nitrogen in runoff," *Soil Science Society of America Journal*, vol. 70, pp. 1788-1796, 2006.
106. A. A. Farooque, Q. U. Zaman, A. W. Schumann, A. Madani and D. C. Percival, "Delineating management zones for site specific fertilization in wild blueberry fields," *Applied engineering in agriculture*, vol. 28, pp. 57-70, 2012.
107. H. J. Farahani and G. W. Buchleiter, "Temporal stability of soil electrical conductivity in irrigated sandy fields in Colorado," *Transactions of the ASAE*, vol. 47, p. 79, 2004.
108. L. Eklundh and A. Singh, "A comparative analysis of standardised and unstandardised principal components analysis in remote sensing," *International Journal of Remote Sensing*, vol. 14, pp. 1359-1370, 1993.

109. J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," 1973.
110. J. J. De Gruijter and A. B. McBratney, "A modified fuzzy k-means method for predictive classification," 1988.
111. S. A. De Caires, M. N. Wuddivira and I. Bekele, "Spatial analysis for management zone delineation in a humid tropic cocoa plantation," *Precision Agriculture*, vol. 16, pp. 129-147, 2015.
112. P. R. Day, "Particle fractionation and particle-size analysis," *Methods of soil analysis. Part 1. Physical and mineralogical properties, including statistics of measurement and sampling*, pp. 545-567, 1965.
113. T. P. Dawson and P. J. Curran, "Technical note A new technique for interpolating the reflectance red edge position," 1998.
114. N. Davatgar, M. R. Neishabouri and A. R. Sepaskhah, "Delineation of site specific nutrient management zones for a paddy cultivated area based on soil fertility using fuzzy clustering," *Geoderma*, vol. 173, pp. 111-118, 2012.
115. P. Curran, "Multispectral remote sensing of vegetation amount," *Progress in physical geography*, vol. 4, pp. 315-341, 1980.
116. J. M. Chen, "Evaluation of vegetation indices and a modified simple ratio for boreal applications," *Canadian Journal of Remote Sensing*, vol. 22, pp. 229-242, 1996.
117. Z. L. Carroll and M. A. Oliver, "Exploring the spatial relations between soil physical properties and apparent electrical conductivity," *Geoderma*, vol. 128, pp. 354-374, 2005.
118. F. Carré and M. C. Girard, "Quantitative mapping of soil types based on regression kriging of taxonomic distances with landform and land cover attributes," *Geoderma*, vol. 110, pp. 241-263, 2002.
119. C. A. Cambardella, T. B. Moorman, T. B. Parkin, D. L. Karlen, J. M. Novak, R. F. Turco and A. E. Konopka, "Field-scale variability of soil properties in central Iowa soils," *Soil science society of America journal*, vol. 58, pp. 1501-1511, 1994.

120. K. F. Bronson, J. D. Booker, S. J. Officer, R. J. Lascano, S. J. Maas, S. W. Searcy and J. Booker, "Apparent electrical conductivity, soil properties and spatial covariance in the US Southern High Plains," *Precision Agriculture*, vol. 6, pp. 297-311, 2005.
121. B. Boydell and A. B. McBratney, "Identifying potential within-field management zones from cotton-yield estimates," *Precision agriculture*, vol. 3, pp. 9-23, 2002.
122. G. E. P. Box and D. R. Cox, "An analysis of transformations," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 211-252, 1964.
123. E. Boegh, H. Soegaard, N. Broge, C. B. Hasager, N. O. Jensen, K. Schelde and A. Thomsen, "Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture," *Remote sensing of Environment*, vol. 81, pp. 179-193, 2002.
124. R. S. Bivand, E. Pebesma and V. Gomez-Rubio, *Applied Spatial Data Analysis with R*, vol. 10, Springer Science Business Media, 2013.
125. G. S. Birth and G. R. McVey, "Measuring the color of growing turf with a reflectance spectrophotometer," *Agronomy Journal*, vol. 60, pp. 640-643, 1968.
126. A. A. Bianchini and A. P. Mallarino, "Soil-sampling alternatives and variable-rate liming for a soybean-corn rotation," *Agronomy Journal*, vol. 94, pp. 1355-1366, 2002.
127. J. C. Bezdek, R. Ehrlich and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, pp. 191-203, 1984.
128. K. J. Beven and M. J. Kirkby, "A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant," *Hydrological Sciences Journal*, vol. 24, pp. 43-69, 1979.
129. F. Baret, S. Jacquemoud and J. F. Hanocq, "The soil line concept in remote sensing," *Remote Sensing Reviews*, vol. 7, pp. 65-82, 1993.
130. A. Bannari, H. Asalhi and P. M. Teillet, "Transformed difference vegetation index (TDVI) for vegetation cover mapping," in *Geoscience and Remote Sensing Symposium*, 2002. IGARSS'02. 2002 IEEE International, 2002.

131. N. Aparicio, D. Villegas, J. Casadesus, J. L. Araus and C. Royo, "Spectral vegetation indices as nondestructive tools for determining durum wheat yield," *Agronomy Journal*, vol. 92, pp. 83-91, 2000.
132. G. L. Anderson and C. Yang, "Multispectral videography and geographic information systems for site-specific farm management," *Precision Agriculture*, pp. 681-692, 1996.
133. H. Abdu, D. A. Robinson and S. B. Jones, "Comparing bulk soil electrical conductivity determination using the DUALEM-1S and EM38-DD electromagnetic induction instruments," *Soil Science Society of America Journal*, vol. 71, pp. 189-196, 2007.
134. R Core Team, "R: A Language and Environment for Statistical Computing," Vienna, 2013.
135. 11.1 - *Principal Component Analysis (PCA) Procedure*, Pennsylvania State University, 2017.
136. K. L. Fleming, D.G. Westfall, D.W. Wiens and M.C. Brodahl, "Evaluating farmer defined management zone maps for variable rate fertilizer application," *Precision Agriculture*, vol. 2, no. 2, pp. 201-215., 2000.
137. S. Wold, "PLS for Multivariate Linear Modeling," in *QSAR: Chemometric Methods in Molecular Design. Methods and Principles in Medicinal Chemistry.*, H. Waterbeemd, Ed., Verlag-Chemie, 1994.
138. I. Urretavizcaya, C. Miranda, J. Royo and L. Santesteban, "Within-vineyard zone delineation in an area with diversity of training systems and plant spacing using parameters of vegetative growth and crop load," *Precision agriculture 15*, p. 479-486, 2015.
139. T. S. Tran and N. Ziadi, "Mehlich 3-Extractable Elements," *Soil Sampling and Methods of Analysis, Second Edition*, 3 2007.
140. B. Tisseyre and A. B. Mcbratney, "A technical opportunity index based on mathematical morphology for site-specific management: an application to viticulture," *Precision Agriculture*, vol. 9, p. 101-113, 2008.

141. J. Murphy and J. Riley, "A modified single solution method for the determination of phosphate in natural waters," *Analytica Chimica Acta*, vol. 27, p. 31-36, 1962.
142. J. Lafond, *Filling Bare Spots in Wild Blueberry Fields*, CRAAQ.
143. R. A. Isaac and J. D. Kerber, "Atomic Absorption and Flame Photometry: Techniques and Uses in Soil, Plant, and Water Analysis," *Instrumental Methods for Analysis of Soils and Plant Tissue ACSESS publications*, 1971.
144. W. Guo, S. J. Maas and K. F. Bronson, "Relationship between cotton yield and soil electrical conductivity, topography, and Landsat imagery," *Precision Agriculture*, vol. 13, p. 678-692, 2012.
145. N. M. Dhawale, V. I. Adamchuk, S. O. Prasher, P. R. L. Dutilleul and R. B. Ferguson, "Spatially Constrained Geospatial Data Clustering for Multilayer Sensor-Based Measurements," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vols. XL-2, p. 187-190, 11 2014.
146. New Brunswick Agriculture, Aquaculture and Fisheries, "Soil fertility and fertilizers for wild blueberry production", Factsheet D.2.0, 2013.
147. Astrium Services, *SPOT6 User Manual*, 2013.

Appendices

Appendix A

Field 21 Tukey Comparison

Table A1 Field 21 Tukey test comparison between MZA classification and targeted approach.

MZA										Targeted					
Property	Unit	MZ1		MZ2		MZ3		El _L EC _L		El _L EC _H		El _H EC _L		El _H EC _H	
Yield	kg ha ⁻¹	6259	ab	4827	b	7198	a	7172	a	6316	a	6714	a	5426	a
RTK	m	134.03	b	135.07	a	129.29	c	129.07	d	129.66	c	133.94	b	134.53	a
TWI	–	6.87	a	7.24	a	5.69	a	6.70	a	5.54	a	6.60	a	6.94	a
Slope	deg	0.96	b	6.26	a	1.53	b	1.50	b	1.23	b	0.31	b	4.79	a
Shallow	mS m ⁻¹	3.16	c	3.20	b	3.26	a	3.17	c	3.37	a	3.11	d	3.23	b
Deep	mS m ⁻¹	2.60	c	2.83	b	3.13	a	3.01	a	3.23	a	2.25	b	2.69	ab
PRP11	mS m ⁻¹	1.26	b	1.28	b	1.41	a	1.35	b	1.46	a	1.23	c	1.31	b
HCP10	mS m ⁻¹	4.28	ab	4.06	b	4.43	a	4.29	a	4.42	a	4.51	a	4.12	a
PRP21	mS m ⁻¹	1.59	b	1.60	b	1.73	a	1.69	b	1.74	a	1.58	c	1.61	c
HCP20	mS m ⁻¹	3.69	b	3.62	b	4.07	a	3.99	a	4.04	a	3.75	b	3.64	b
0-5 cm															
Total C	%	11.62	a	10.82	a	10.64	a	9.69	a	13.12	a	9.24	a	13.91	a
Total N	%	0.50	a	0.37	a	0.44	a	0.42	a	0.54	a	0.39	a	0.58	a
pH	–	4.48	b	4.80	a	4.97	a	4.72	ab	4.89	a	4.38	b	4.61	ab
P	mg kg-1	44.47	b	64.71	ab	82.24	a	61.40	a	81.48	a	51.65	a	39.93	a
K	mg kg-1	121.30	a	90.77	a	99.05	a	99.05	a	119.58	a	116.48	a	162.67	a
Ca	mg kg-1	364.18	a	334.42	a	367.82	a	370.84	a	395.26	a	367.35	a	331.73	a
Mg	mg kg-1	102.85	a	107.70	a	111.75	a	89.92	ab	136.46	ab	79.06	b	171.11	a
Al	mg kg-1	883.61	a	875.69	a	899.30	a	880.16	a	906.00	a	831.42	a	762.81	a
Fe	mg kg-1	1153.90	b	1779.60	a	1754.90	a	1339.10	a	1995.40	a	1052.30	a	2003.30	a
S.O.M.	g kg-1	177.79	a	165.50	a	162.74	a	148.25	a	200.72	a	141.31	a	212.77	a
5-15 cm															
Total C	%	1.17	b	1.48	a	1.33	ab	1.13	a	1.16	a	1.05	a	1.53	a
Total N	%	0.06	a	0.07	a	0.07	a	0.07	a	0.06	a	0.06	a	0.07	a
pH	–	4.90	b	5.01	b	5.27	a	5.13	ab	5.25	a	4.91	c	4.92	bc
P	mg kg-1	52.54	b	71.32	ab	80.19	a	78.12	a	61.85	a	38.35	a	71.10	a
K	mg kg-1	46.08	a	37.76	ab	31.58	b	30.17	a	34.91	a	40.07	a	45.54	a
Ca	mg kg-1	285.28	a	324.89	a	294.66	a	276.30	ab	330.62	ab	234.92	b	358.60	a
Mg	mg kg-1	7.68	a	8.73	a	7.71	a	6.19	a	8.57	a	5.41	a	8.83	a
Al	mg kg-1	1737.23	a	1560.11	b	1600.85	ab	1742.32	a	1623.52	a	1749.40	a	1639.08	a
Fe	mg kg-1	93.81	b	184.48	a	190.21	a	109.65	ab	215.98	a	60.25	b	152.09	ab
S.O.M.	g kg-1	17.96	b	22.61	a	20.33	ab	17.35	a	17.81	a	16.10	a	23.46	a
Sand	g kg-1	896.00	a	897.49	a	801.82	b	823.96	ab	819.01	b	892.17	a	887.83	ab
Silt	g kg-1	78.15	b	76.35	b	177.93	a	153.62	a	159.41	a	81.18	b	84.68	b
Clay	g kg-1	25.85	a	26.16	a	20.24	b	22.43	ab	21.58	b	26.65	ab	27.49	a
STG	g kg-1	16.29	a	17.72	a	5.58	b	4.70	a	8.14	a	15.19	a	10.67	a
SG	g kg-1	141.85	a	140.30	a	42.39	b	75.41	ab	34.08	b	151.77	a	113.22	ab
SM	g kg-1	387.31	a	385.64	a	143.40	b	216.33	b	142.61	b	395.14	a	390.76	a
SF	g kg-1	284.06	b	285.32	ab	350.71	a	342.88	a	398.92	a	261.19	a	306.73	a
STF	g kg-1	66.48	b	68.50	b	259.75	a	184.65	a	235.26	a	68.89	b	66.44	b
PC2	–	-39.72	ab	105.05	a	-91.47	b	-83.26	ab	-237.81	b	-25.99	a	-58.43	ab

Appendix B

Field 140b Tukey Comparison

Table A1 Field 140b Tukey test comparison between MZA classification and targeted approach.

MZA										Targeted					
Property	Unit	MZ1		MZ2		MZ3		El _L EC _L		El _L EC _H		El _H EC _L		El _H EC _H	
Yield	kg ha-1	4846	a	3717	a	4058	a	3893	a	5487	a	3359	a	4755	a
RTK	m	124.45	a	124.60	a	123.71	b	123.59	b	123.67	b	124.76	a	124.62	a
TWI	–	7.46	a	4.60	b	5.13	b	4.41	a	6.04	a	4.97	a	5.90	a
Slope	deg	2.33	a	0.65	b	0.65	b	0.78	a	0.59	a	0.44	a	1.10	a
Shallow	mS m-1	2.69	a	2.65	b	2.67	ab	2.61	c	2.76	a	2.59	c	2.71	b
Deep	mS m-1	2.07	a	2.22	a	2.28	a	2.14	a	2.48	a	2.26	a	2.13	a
PRP11	mS m-1	1.04	a	1.01	a	1.03	a	1.01	bc	1.08	a	0.97	c	1.05	ab
HCP10	mS m-1	4.24	a	4.25	a	4.33	a	4.33	a	4.37	a	4.40	a	4.26	a
PRP21	mS m-1	1.33	a	1.29	b	1.30	ab	1.28	b	1.34	a	1.26	b	1.34	a
HCP20	mS m-1	2.94	b	2.89	b	3.05	a	3.07	a	3.01	ab	2.97	ab	2.97	b
0-5 cm															
Total C	%	11.78	a	7.70	b	9.34	ab	9.30	ab	12.94	a	4.72	b	12.10	a
Total N	%	0.60	a	0.37	b	0.47	ab	0.45	ab	0.70	a	0.24	b	0.59	a
pH	–	4.56	ab	4.40	b	4.67	a	4.58	a	4.64	a	4.51	a	4.28	a
P	mg kg-1	28.20	a	30.23	a	52.62	a	40.63	a	70.40	a	26.42	a	24.89	a
K	mg kg-1	95.87	a	91.45	a	93.91	a	96.39	ab	99.83	ab	56.78	b	133.58	a
Ca	mg kg-1	340.22	b	384.08	ab	404.81	a	456.92	a	375.85	a	377.77	a	385.56	a
Mg	mg kg-1	97.28	a	61.60	b	91.46	ab	77.91	ab	113.26	a	36.33	b	124.37	a
Al	mg kg-1	1001.56	a	857.40	a	1023.22	a	990.60	ab	1232.30	a	898.60	ab	792.20	b
Fe	mg kg-1	551.89	a	374.09	a	554.12	a	454.70	ab	766.20	a	211.30	b	775.20	a
S.O.M.	g kg-1	180.19	a	117.77	b	142.87	ab	142.35	ab	197.99	a	72.14	b	185.13	a
5-15 cm															
Total C	%	1.26	a	1.12	a	1.20	a	1.09	ab	1.37	a	0.87	b	1.19	ab
Total N	%	0.09	a	0.08	a	0.09	a	0.08	ab	0.10	a	0.07	b	0.08	ab
pH	–	5.00	ab	4.92	b	5.10	a	5.08	a	5.05	a	4.96	b	4.90	ab
P	mg kg-1	26.31	a	19.46	a	29.81	a	27.19	ab	33.00	a	9.59	b	19.13	ab
K	mg kg-1	40.97	a	41.00	a	39.40	a	43.21	ab	46.55	ab	30.83	b	55.16	a
Ca	mg kg-1	225.17	a	197.32	a	230.97	a	233.32	a	217.34	ab	147.26	b	269.10	a
Mg	mg kg-1	6.70	a	5.66	a	7.25	a	7.89	a	8.11	a	4.97	a	7.19	a
Al	mg kg-1	2068.21	a	1986.07	a	2034.08	a	1920.70	a	2057.10	a	1982.90	a	2100.60	a
Fe	mg kg-1	226.31	ab	156.82	b	261.48	a	219.07	ab	323.80	a	87.27	b	276.50	a
S.O.M.	g kg-1	19.22	a	17.07	a	18.40	a	16.69	ab	20.93	a	13.34	b	18.14	ab
Sand	g kg-1	881.37	b	902.18	a	892.68	ab	898.09	ab	867.55	b	910.24	a	889.95	ab
Silt	g kg-1	95.39	a	81.37	ab	70.15	b	76.93	ab	103.98	a	60.59	b	85.67	ab
Clay	g kg-1	23.24	b	27.67	a	25.95	ab	24.98	a	28.47	a	29.17	a	24.38	a
STG	g kg-1	19.91	a	23.96	a	29.05	a	26.43	a	31.74	a	30.57	a	32.46	a
SG	g kg-1	121.70	b	202.30	a	144.46	b	143.41	b	126.09	b	277.73	a	195.36	ab
SM	g kg-1	309.78	b	408.29	a	306.25	b	334.33	ab	223.96	b	442.70	a	336.00	a
SF	g kg-1	326.03	a	224.39	b	336.81	a	329.99	a	373.04	a	134.65	b	246.49	ab
STF	g kg-1	103.96	a	43.24	b	76.11	a	63.93	ab	112.72	a	24.60	b	79.64	ab
PC2	–	96.59	a	49.54	a	52.22	a	54.16	ab	-150.31	c	88.77	a	-73.97	bc

Appendix C

Bare Spot Standards Scores

Table A1 Standard score of soil variables in bare patches. The score represents distance from the field mean. A value close to 0 has little difference from mean. $N_{21} = 8$ and $N_{140b}=2$.

	z_{21}	z_{140b}
HCP10	-0.641	-1.378
PRP11	-0.465	0.168
HCP20	-0.880	-0.725
PRP21	-0.941	-1.022
Veris Shallow	-0.143	-0.090
Veris Deep	0.077	-1.245
Elevation	1.090	-1.159
Yield	-1.838	-1.768
Slope	0.676	0.471
TWI	-0.371	0.140
0-5 cm		
S.O.M.	-1.144	-1.240
Total N	-1.169	-1.163
Total C	-1.144	-1.240
Soil pH_{water}	0.845	2.287
P	0.530	1.196
K	-1.047	-1.141
Ca	-0.337	-1.461
Mg	-0.821	-1.092
Al	1.352	0.524
Fe	-0.404	-0.805
5-15 cm		
S.O.M.	-0.266	-1.186
Total N	0.087	-1.239
Total C	-0.266	-1.186
Soil pH_{water}	0.329	2.530
P	-0.297	1.478
K	-0.587	-1.347
Ca	-0.254	-1.725
Mg	-0.041	-0.297
Al	0.212	0.125
Fe	0.053	-0.574
Clay	0.329	-1.570
Silt	-0.594	-1.124
Sand	0.583	1.453
Very coarse sand	-0.378	0.220
Coarse sand	0.010	-0.642
Medium sand	0.437	-0.639
Fine sand	0.408	1.321
Very fine sand	-0.569	-0.064

Appendix D

Field 21 Standard Scores

Table A1 Field 21 standard scores of soil variables by regression-based separation method.

	$Elev_{Low}EC_{Low}$	$Elev_{Low}EC_{High}$	$Elev_{High}EC_{Low}$	$Elev_{High}EC_{High}$
Yield	0.212	-0.033	0.081	-0.287
Shallow EC	-0.466	2.185	-1.357	0.311
Deep EC	0.317	0.776	-1.264	-0.345
PRP1.1	0.301	1.426	-1.039	-0.145
HCP1.0	-0.025	0.176	0.317	-0.290
PRP2.1	0.545	1.152	-0.978	-0.542
HCP2.0	0.558	0.743	-0.309	-0.690
Elevation	-1.206	-0.978	0.666	0.894
Slope	-0.275	-0.282	-0.655	0.948
TWI	0.135	-0.280	0.046	0.280
0-5 cm				
Total C	-0.214	0.313	-0.284	0.434
Total N	-0.135	0.320	-0.234	0.449
pH_{water}	-0.039	0.301	-0.713	-0.254
P	-0.037	0.334	-0.218	-0.435
K	-0.117	0.176	0.132	0.791
Ca	0.127	0.447	0.082	-0.384
Mg	-0.242	0.405	-0.393	0.887
Al	-0.031	0.059	-0.200	-0.439
Fe	-0.175	0.529	-0.483	0.537
OM	-0.214	0.313	-0.284	0.434
5-15 cm				
Total C	-0.288	-0.230	-0.444	0.471
Total N	-0.108	-0.269	-0.321	0.227
pH_{water}	0.185	0.566	-0.541	-0.519
P	0.229	-0.107	-0.592	0.084
K	-0.343	-0.153	0.054	0.273
Ca	-0.184	0.341	-0.583	0.611
Mg	-0.167	0.070	-0.244	0.096
Al	0.344	-0.112	0.371	-0.052
Fe	-0.247	0.435	-0.564	0.025
OM	-0.288	-0.230	-0.444	0.471
Total Sand	-0.443	-0.510	0.478	0.419
Total Silt	0.449	0.526	-0.510	-0.463
Total Clay	-0.216	-0.379	0.597	0.760
V. Coarse Sand	-0.536	-0.284	0.231	-0.099
Coarse Sand	-0.276	-0.742	0.584	0.150
Medium Sand	-0.420	-0.871	0.676	0.649
Fine Sand	0.249	0.704	-0.414	-0.044
V. Fine Sand	0.283	0.672	-0.606	-0.625

Appendix E

Field 140b Standard Scores

Table A1 Field 140b standard scores of soil variables by regression-based separation method.

	$Elev_{Low}EC_{Low}$	$Elev_{Low}EC_{High}$	$Elev_{High}EC_{Low}$	$Elev_{High}EC_{High}$
Yield	-0.041	0.666	-0.278	0.342
Shallow EC	-0.929	1.371	-1.306	0.589
Deep EC	-0.167	0.270	-0.008	-0.183
PRP1.1	-0.170	0.485	-0.511	0.259
HCP1.0	0.199	0.333	0.424	0.017
PRP2.1	-0.197	0.337	-0.413	0.287
HCP2.0	0.543	0.246	0.090	-0.314
Elevation	-1.274	-1.127	0.972	0.714
Slope	-0.072	-0.267	-0.433	0.263
TWI	-0.295	0.355	-0.073	0.300
0-5 cm				
Total C	0.094	0.811	-0.810	0.645
Total N	0.023	1.048	-0.805	0.599
pH_{water}	0.170	0.336	-0.039	-0.685
P	0.041	0.660	-0.255	-0.287
K	0.062	0.123	-0.645	0.725
Ca	0.707	-0.111	-0.091	-0.013
Mg	0.008	0.672	-0.774	0.881
Al	0.175	0.997	-0.138	-0.500
Fe	-0.031	0.889	-0.751	0.916
OM	0.094	0.811	-0.810	0.645
5-15 cm				
Total C	-0.133	0.354	-0.518	0.034
Total N	-0.153	0.558	-0.575	0.081
pH_{water}	0.457	0.319	-0.574	-0.214
P	0.135	0.407	-0.688	-0.242
K	0.154	0.335	-0.518	0.803
Ca	0.248	0.046	-0.841	0.700
Mg	0.512	0.585	-0.488	0.272
Al	-0.343	0.155	-0.116	0.314
Fe	0.081	0.717	-0.721	0.430
OM	-0.133	0.354	-0.518	0.034
Total Sand	0.070	-0.942	0.472	-0.200
Total Silt	-0.020	0.867	-0.556	0.266
Total Clay	-0.244	0.327	0.442	-0.341
STG	0.065	0.412	0.335	0.459
SG	-0.300	-0.495	1.210	0.284
SM	-0.218	-1.289	0.833	-0.202
SF	0.394	0.735	-1.154	-0.268
STF	0.012	1.001	-0.785	0.331